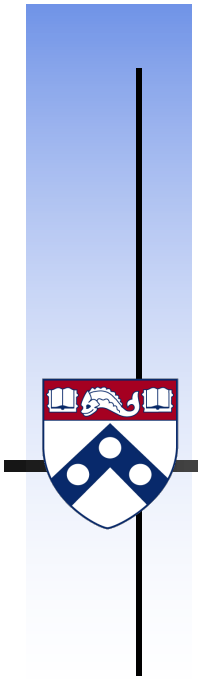


CIS 5530: Networked Systems

Introduction

August 30, 2023

Your team this semester

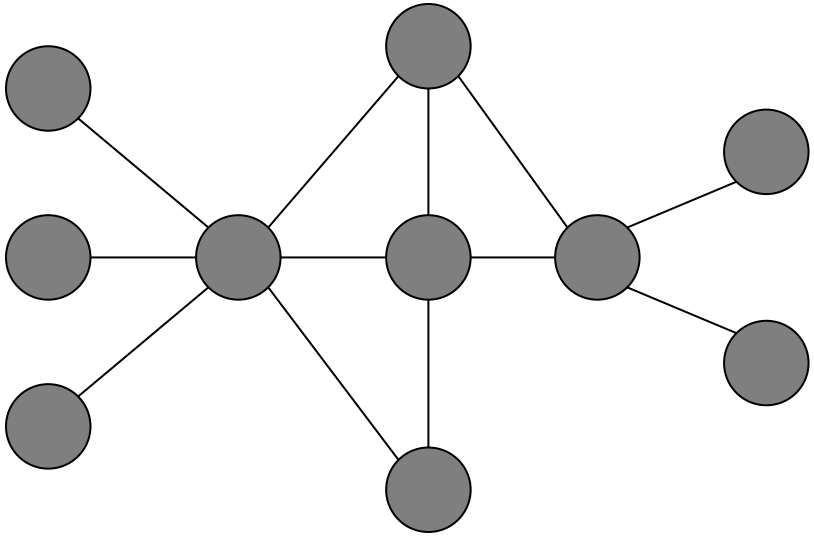
Instructor: Vincent Liu (Me)
liuv@seas
Office Hours: W 2-3 pm ET @ Levine 574

TAs:

- Yinda Zhang (yindaz@seas) C
- Naveen Albert (naveen23@seas) C
- Hugo Genevriere (hgenevri@seas) C P
- Harry Weinstock (harrykw@seas) C P
- Angineh Keshishian (angkesh@seas) P
- Vijay Koundinya (vkoundin@seas) P
- Wenjing Wei (wewe@seas) P

What this course is about

- Networks: A system of links that interconnect computers in order to move data



What this course is about

Mostly about the Internet!

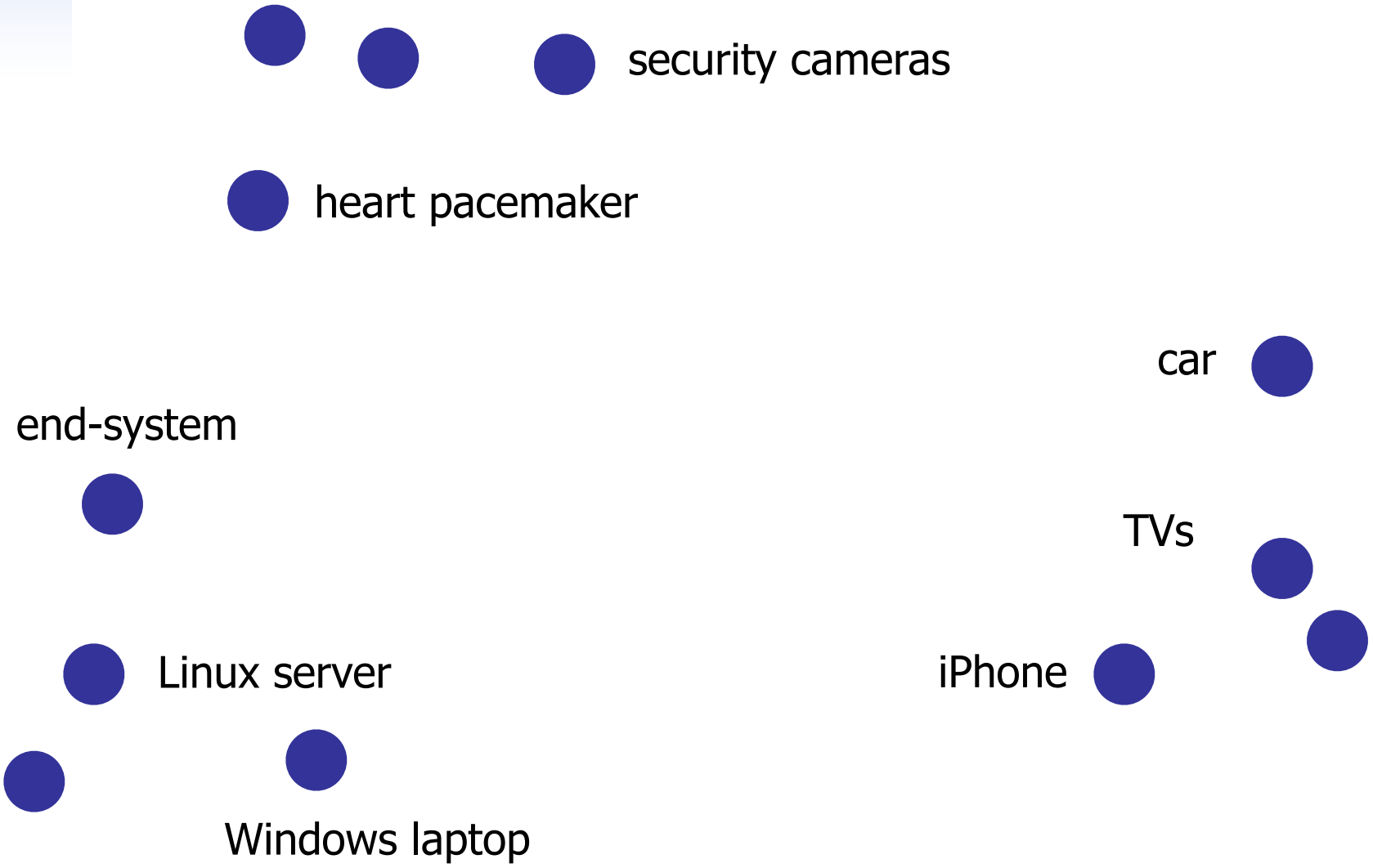
Two different definitions of "Internet":

1. The networking infrastructure that links all connected devices
2. The entire ecosystem of networked applications that uses the basic connectivity provided by the definition above

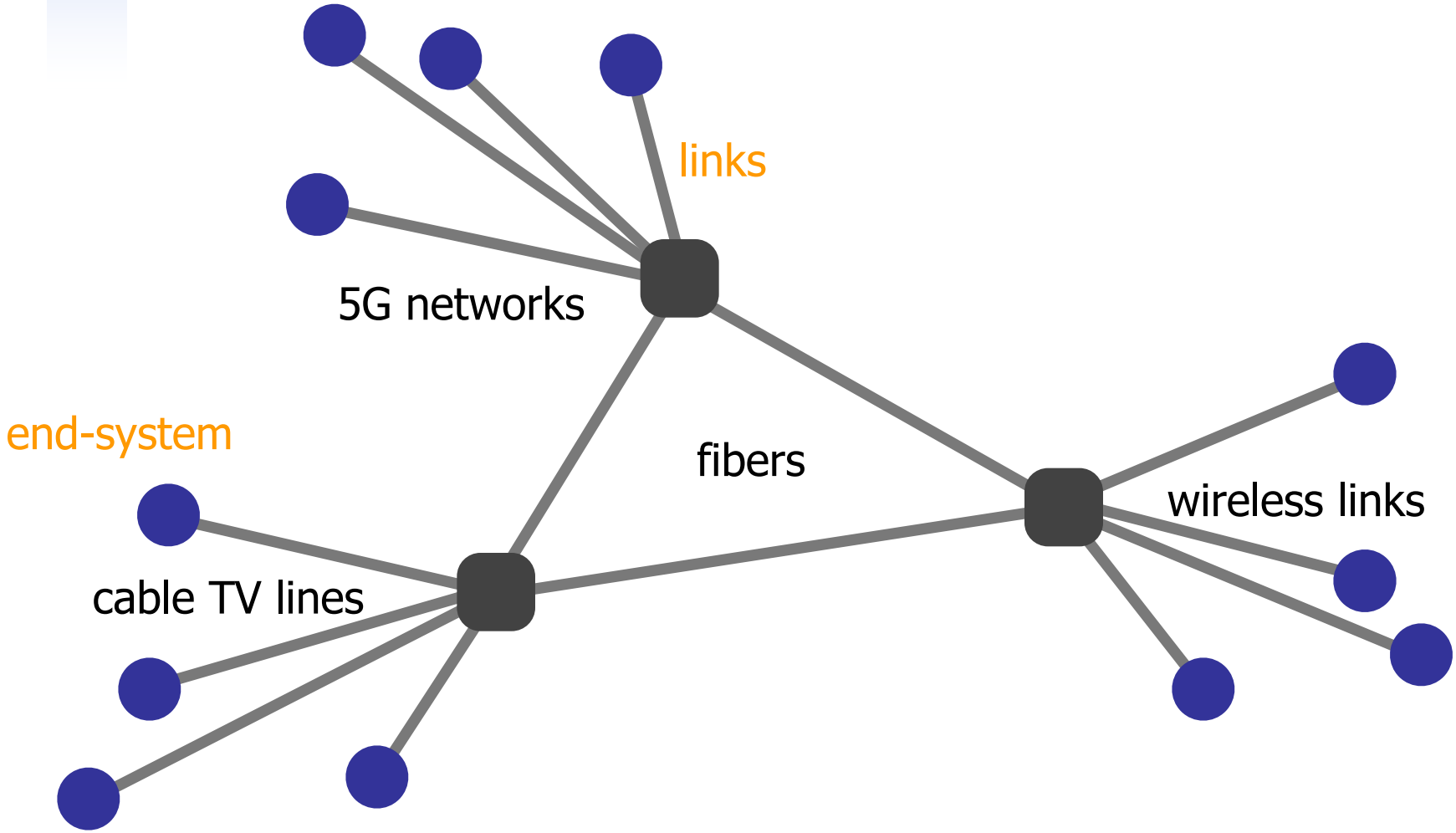
In this class, we'll be using the first

What's so hard about building a network?

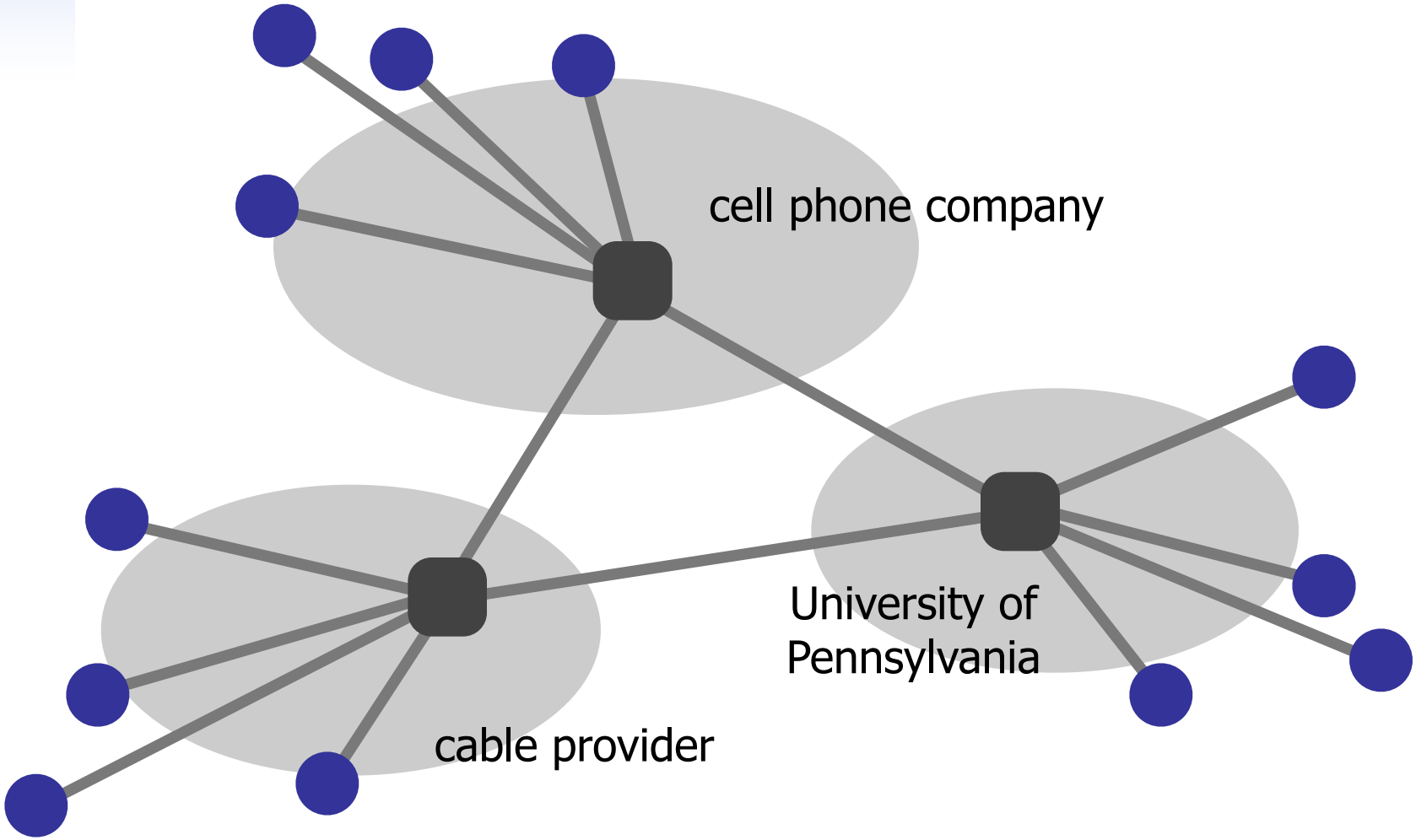
The Internet needs to take end-hosts

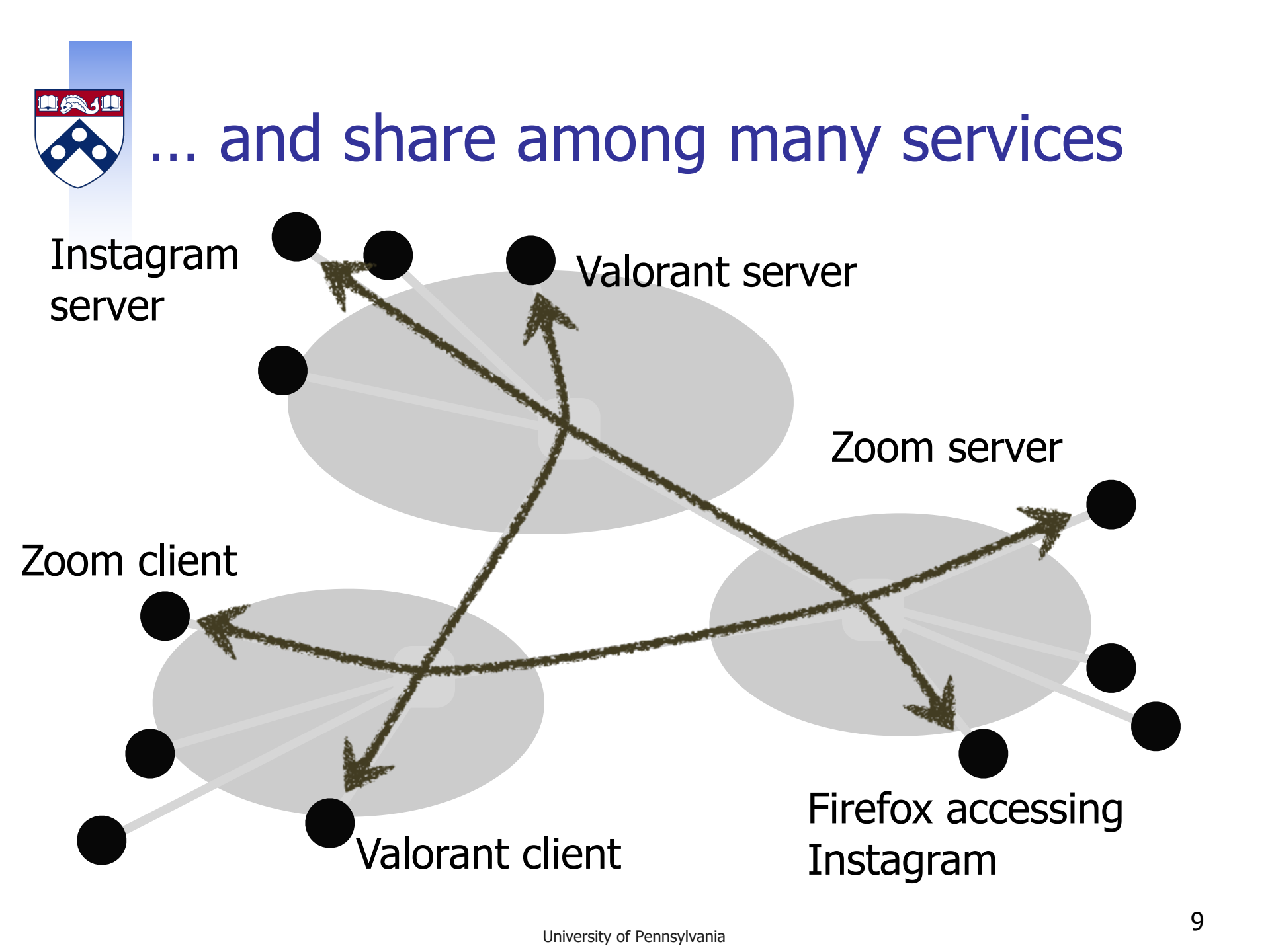


...connect them



…deal with diverse organizations





… and share among many services

The complexity of streaming a video

- How do we know where to send requests for netflix.com?
- Why do their servers not get overwhelmed?
- How does traffic get from here to there? Reliably?
- Why can I continue to browse the web while I watch?
- How do I know that I'm actually talking to Netflix?
- Why is my streaming sometimes choppy and what can I do about it?
- How do they keep costs down even though they make up the majority of residential Internet traffic.

Why should you study it?

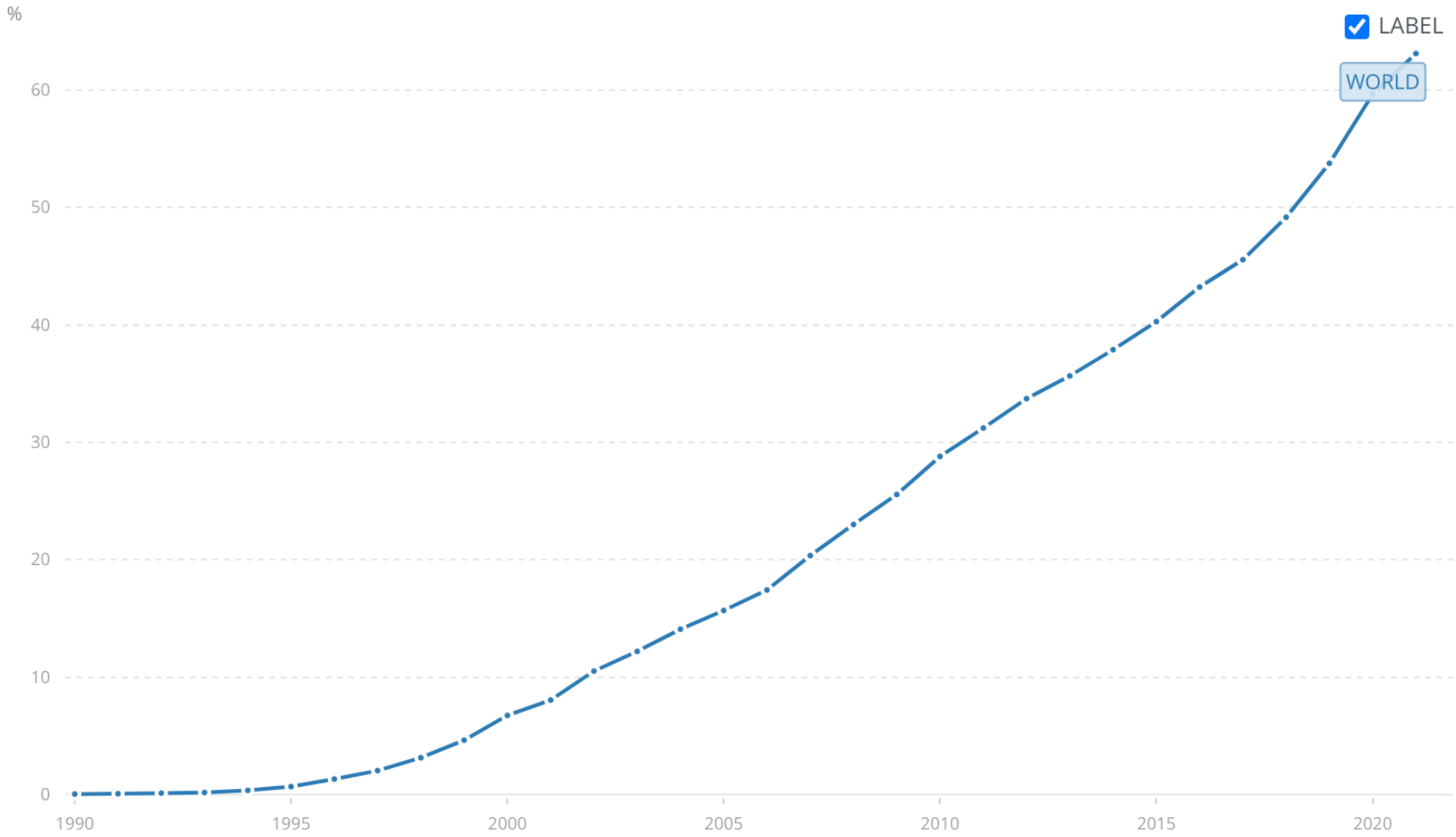
The Internet has transformed everything

- The way we do business
  - E-commerce, advertising, cloud-computing
- The way we have relationships
  - Instagram followers, 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

It's also the largest computer system in the world

- 5.4 Billion users (~68% of world population)
- 1.88 Billion websites
- 347.3 Billion emails sent per day
- 8.5 Billion Google searches per day
- 2.96 Billion Facebook users
- 1 Billion hours of YouTube watched per day
- Routers that switch 25 Terabits/second
- Links that carry 400 Gigabits/second

With growth still ongoing



Source: data.worldbank.org

The Internet Is…

A case study of a *very* complex (and successful) system

- Constantly evolving
  - Technology: Optical, wireless, satellite, copper
  - Endpoint devices: wearable devices to datacenters and supercomputers
  - Applications: Video streaming, social networking, file transfer, Zoom, live TV, gaming, remote medicine, IM
- Decentralized
  - Many parties with (often conflicting) interests
- Failure-prone
  - Physical errors, logic errors, human errors, etc.

What you'll learn in this course

- For network builders: How the Internet functions
  - Protocols (Ethernet, IP, TCP/UDP)
  - Components (hubs, switches, routers)
  - Concepts (naming, security, etc.)
- For network users: How to use/reason about the Internet
  - Socket programming
  - Analysis of networks
  - Measurement
- For computer scientists: How we got here
  - The end-to-end principle
  - Layering
  - Soft state

Relationship to other courses at Penn

- NETS 2120, CIS 5050, 5450, 5550
  - The theory and operations of distributed systems
  - Basics of MapReduce, Spark, and data-parallel programming
  - These will touch on the higher layers of the network stack
- CIS 5510
  - The security of networks and distributed systems
  - Approaches the problem from a different angle
- CIS 3800, 5480, 5710
  - Concepts behind operating system design and computer architecture
  - We will mostly operate above these layers.

Agenda

- Introduction 
- Course logistics 

- A brief history of the Internet...

Course format

- Class time: MW 3:30pm – 5:00pm
- Optional Textbooks:
  - *Computer Networks: A Systems Approach (5th Edition).* Larry L. Peterson and Bruce S. Davie.
  - *Computer Networking: A Top-Down Approach (6th Edition).* James F. Kurose and Keith W. Ross
  - *Computer Networks (5th Edition).* Andrew S. Tanenbaum and David J Wetherall.
- Grading
  - Two midterms: 25% each
  - Two programming projects: 22.5% each
  - Participation: 5%

Participation

- Please come to class, ask questions, engage!
- Don't worry about the participation grade
  - Almost everyone will get a standard score, e.g., 2.5/5
  - Higher/lower scores are reserved for a handful of extreme cases
- However, strong correlation between participation and success

The structure of this course

- First half:
  - The basic protocols and key concepts that run in the middle of the Internet
- Second half:
  - The protocols and concepts that run on the endpoints of the Internet
  - How the Internet has evolved over time

Tools

- Website (https://www.cis.upenn.edu/~cis5530/)
  - Schedule, homework assignments, readings, lecture slides
- Canvas (https://canvas.upenn.edu/courses/1724057)
  - Grades, central hub for all class resources
- Gradescope (https://www.gradescope.com/courses/559790)
  - Homework turn-in, autograding support
- Ed Discussion (https://edstem.org/us/courses/41651/discussion/)
  - Discussion, announcements
- OHQ (TBD)
  - Project Office Hours

Ed Discussion

- Ed Discussion will also be used for
  - Announcements, e.g., cancelled classes (if necessary)
  - Supplemental materials, e.g., links to relevant papers or articles
  - Corrections and clarifications, e.g., bugs in the homework handouts
  - Please check the group frequently!
- You may ask questions or answer them
  - Incentive: top 3 students by endorsed answer count will get 2 points on their final grade!
- Please sign up at
  - https://edstem.org/us/courses/41651/discussion/

Programming projects

- In C++ using NS-3 and a provided github repo
- Done in groups of three

- Project 1: Implementing classic routing protocols
- Project 2: Implementing a custom overlay network

- For now: Form teams ASAP (deadline: Sept 13)
  - Project 1 will be released when teams are finalized

Policies: Collaboration

- All projects are only with your group and no one else
  - All the code you submit has to be your own
    - Only exception: Code we have provided or explicitly authorized
    - **NO** code you have found on the web. **NO** sharing with others.
  - Penn's Code of Academic Integrity applies
    - No cheating, plagiarism, fabrication, multiple submissions, gaining an unfair advantage, or facilitating (!) academic dishonesty
  - **It's not worth it!!** Penalties can be severe: http://www.upenn.edu/academicintegrity/ai_violations.html
- All exams are to be done individually
- Zero tolerance policy to ensure fairness
- Yes, we're serious!

Policies: Extensions

- Everyone gets 4 'slip days'
  - Deducted in increments of 1 day
  - No questions asked
- Absolutely no extensions without slip days
- Some recommendations:
  - Please start the homeworks early!!!
  - Save your slip days for project 2 and unforeseen events
    - Interview calls, deadlines in other courses, tricky bugs in your code, ...

Waitlist

- Lots of demand given the fate of NETS 2120
- I *think* we can handle everyone that needs the class to graduate this year
  - Come talk to me if you are graduating and are still on the waitlist

Agenda

- Introduction 
- Course logistics 
- A brief history of the Internet... 

Internet Prehistory (late 50's)

- Telephone network is largest communication system
  - Uses circuit switching
- Need to build networks for other tasks
  - Defense
  - Computers
  - …
- But people knew that circuit switching was:
  - Inefficient (for bursty loads)
  - Not resilient
    - Which is why AT&T worked so hard to make components reliable

Phase 1: Humble Beginnings

1961 Baran, Kleinrock advocate packet switching

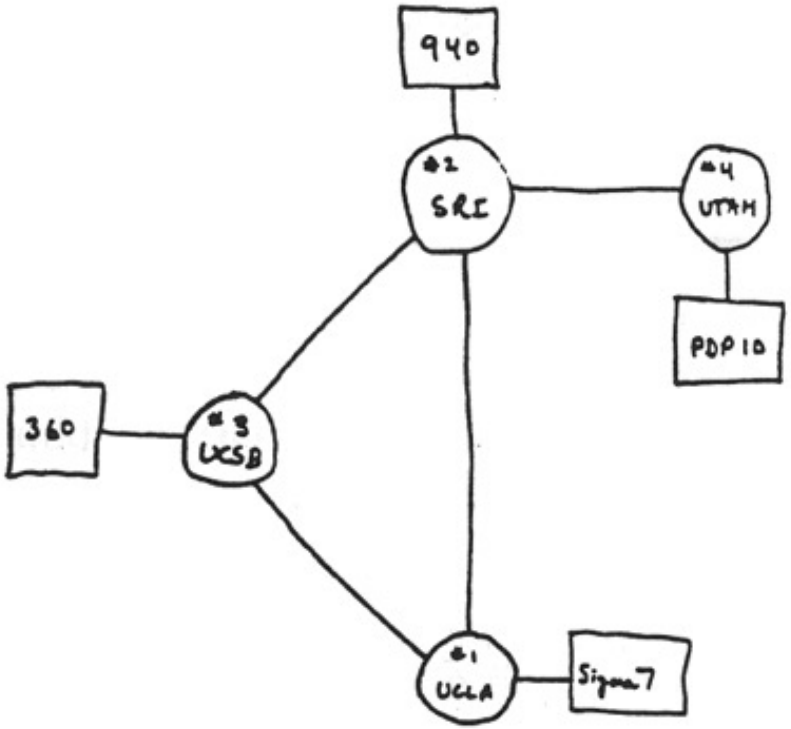
1962 Licklider's vision of Galactic Network

1965 Roberts connects two computers via phone

1967 Roberts publishes vision of ARPANET

1969 BBN installs first IMP at UCLA

IMP: Interface Message Processor



THE ARPA NETWORK

DEC 1969

4 NODES

The opening of the Internet revolution

- Kleinrock's group at UCLA tried to log on to Stanford computer: His recollection of the event…
- We typed the L…
  - "Do you see the L?"
  - "Yes, we see the L."
- We typed the O…
  - "Do you see the O?"
  - "Yes, we see the O."
- Then we typed the G…

…and the system crashed!

Phase 2: Internetworking

1971 Network Control Program (protocol)
1972 Public demonstration of ARPANET
1972 Email invented
1972 Telnet introduced
1973 FTP introduced
1973 Ethernet invented (Xerox PARC)
1974 Cerf and Kahn paper on TCP/IP
1980 TCP/IP adopted as defense standard
1983 Global NCP to TCP/IP flag day

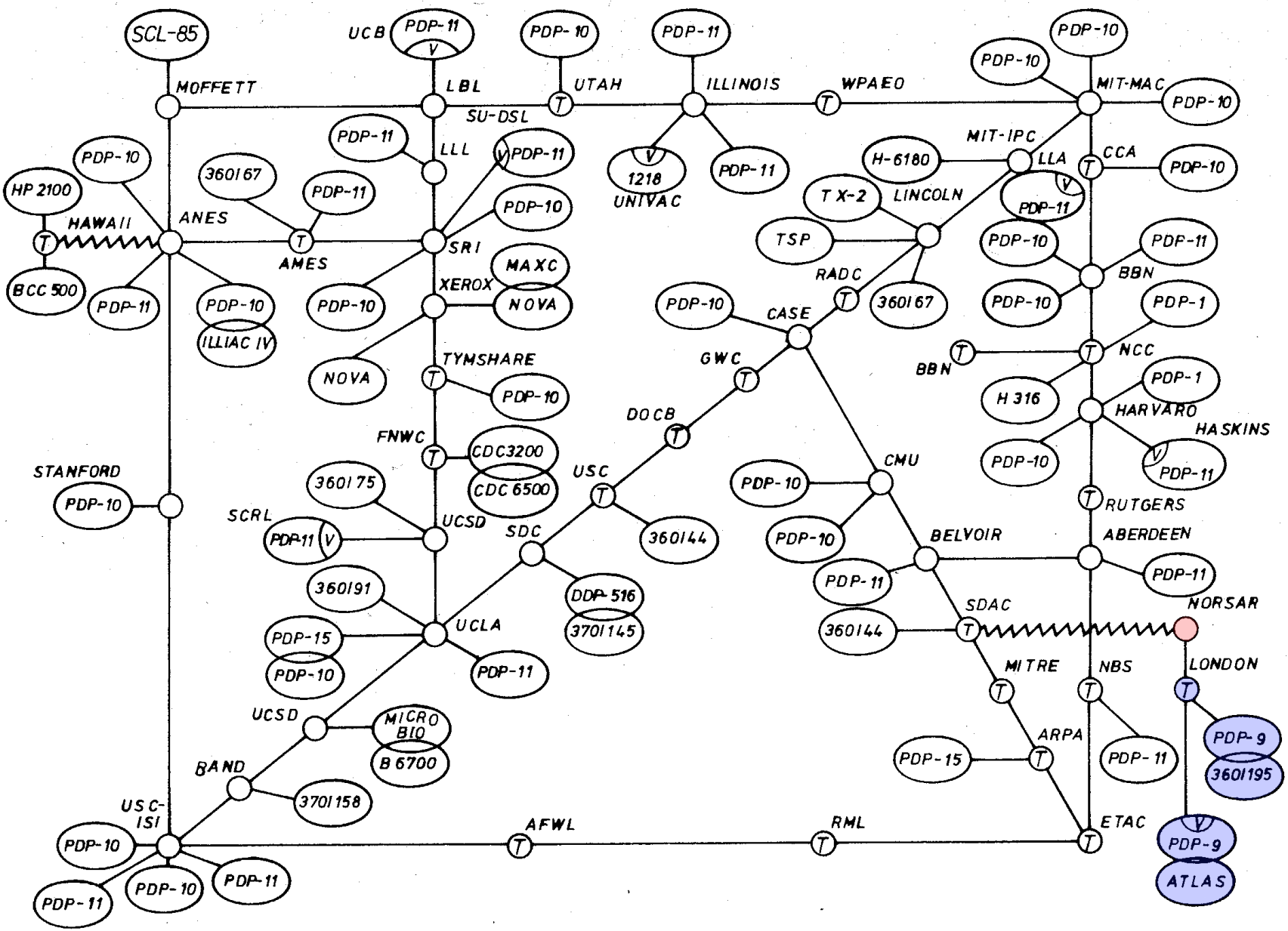


Abb. 4 ARPA NETwork, topologische Karte. Stand Juni 1974.

Phase 3: Rapid growth

198x XNS, DECbit, and other protocols

1984 Janet (British research network)

198x Internet meltdowns due to congestion

1986 Van Jacobson saves the Internet (BSD TCP)

1988 Dave Clark steps down from IAB

1989 Birth of the web….Tim Berners-Lee (CERN)

He invented HTTP

Why did it take a physicist to invent web?

- Computer scientists were trying!
  - Invented Xanadu (Ted Nelson)

- "The World Wide Web … trivializes our original hypertext model with one-way ever-breaking links and no management of version or contents." – Ted Nelson

- Users didn't need what we wanted to invent

Phase 4: The Web

1993 Search engines invented (Excite)

1994 Internet goes commercial

199x ATM rises and falls (as internetworking layer)

Telcos try to kill the Internet

199x QoS rises and falls

1998 IPv6 specification

1998 Google reinvents search

2000 Dot-com bubble burst

Phase 5: The Cloud

1997 The term "Cloud Computing" is coined

1998 Rackspace founded

1999 Salesforce founded

2006 Amazon EC2 and AWS are launched

Went from a bookseller to a cloud computing company

2006 YouTube purchased by Google

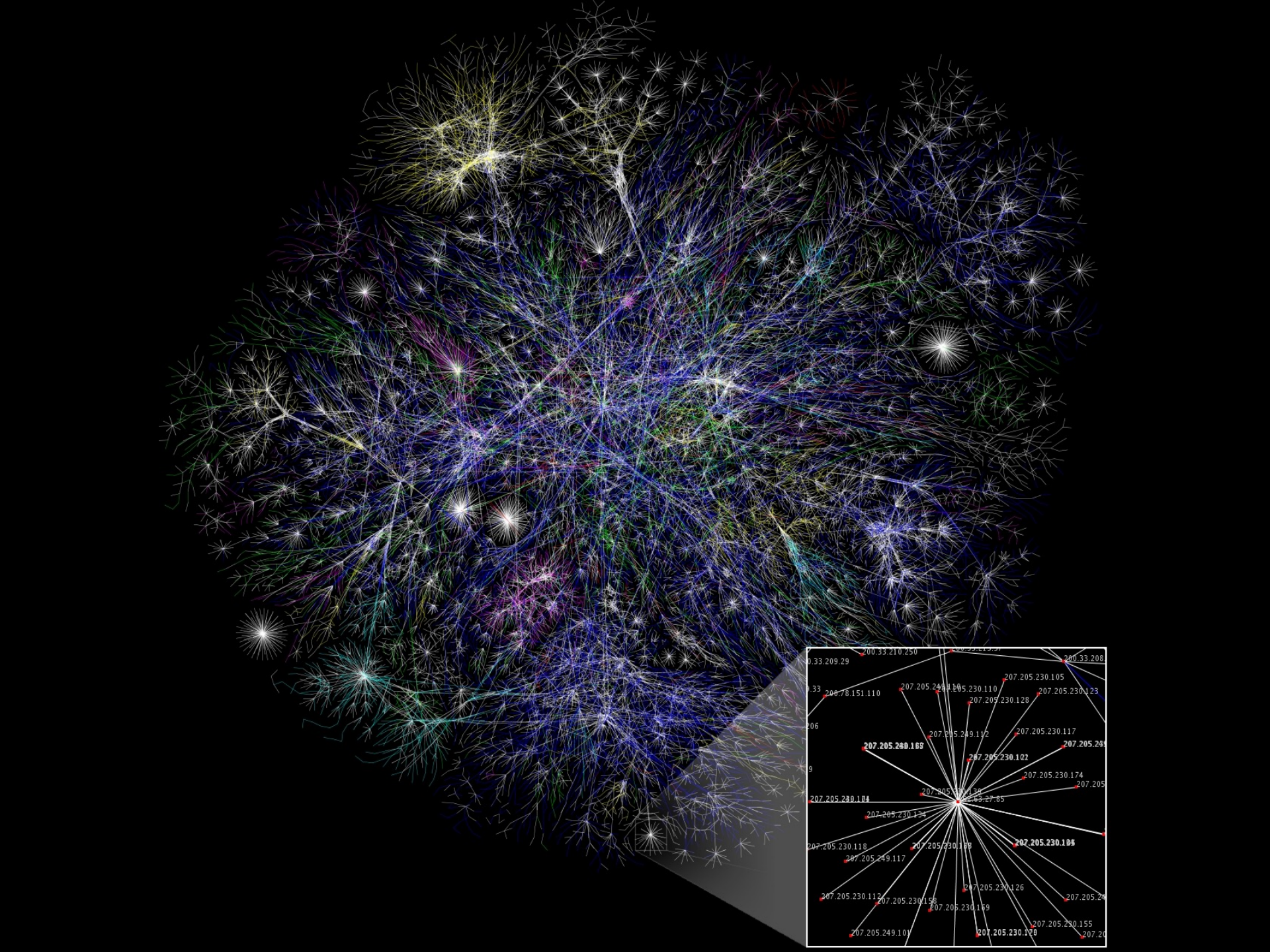
2007 Rise of the iPhone, netflix starts streaming

2010 Instagram founded

Phase 6: Decentralization?

2016 Netflix rolls out its own CDN

2020 Zoom's stock price goes up…


200.33.210.250
207.205.230.105
200.78.151.110
207.205.230.110
207.205.230.123
207.205.230.128
207.205.249.112
207.205.230.117
207.205.230.174
207.205.230.134
207.205.230.118
207.205.249.117
207.205.230.126
207.205.230.112
207.205.230.158
207.205.230.159
207.205.230.155
207.205.249.101

Have we found the right solution?

- We don't really know
- 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
- Still, it is just one design with many questions