The Internet

CSE 120, Winter 2020

breadcrumbs.samwolfson.com

Instructor: Sam Wolfson

Teaching Assistants: Yae Kubota, Eunia Lee, Erika Wolfe

CenturyLink, Frontier took FCC cash, failed to deploy all required broadband

"CenturyLink and Frontier Communications have apparently failed to meet broadband-deployment requirements in numerous states where they are receiving government funding to expand their networks in rural areas."

"CenturyLink and Frontier were among 10 ISPs that accepted funding in the FCC's 2015 Connect America Fund auction in exchange for promises to deploy Internet service with speeds of at least 10Mbps downstream and 1Mbps upstream. CenturyLink is receiving $505.7 million in annual support for six years to deploy service to 1,174,142 homes and businesses in 33 states."

- https://arstechnica.com/tech-policy/2020/01/centurylink-frontier-took-fcc-cash-failed-to-deploy-all-required-broadband/

Administrivia

- Assignments
  - Lego Family [submit] due **tonight**!
  - Animal Functions [submit] due Tuesday (1/28)
  - Continue adding projects to your portfolio

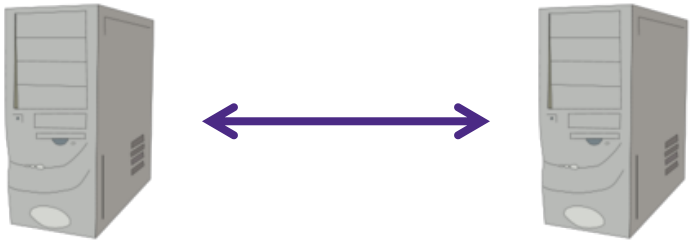
Outline

- **Networks**
- Growth of the Internet
- Sending Information
- Encryption

Communication Channels

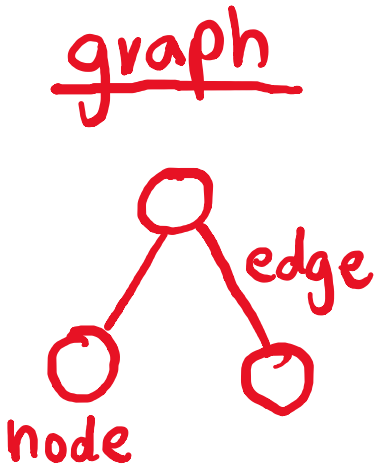
- We often transmit sequences of bits between computers – why? communicate, exchange info, etc.
  - Only capability we need because of binary encoding!
  - Via *wire*: Ethernet
  - Via *wireless*: Wi-Fi, 3G/4G/5G, Bluetooth

- A network is a group of computing devices connected together, either by wire or wirelessly

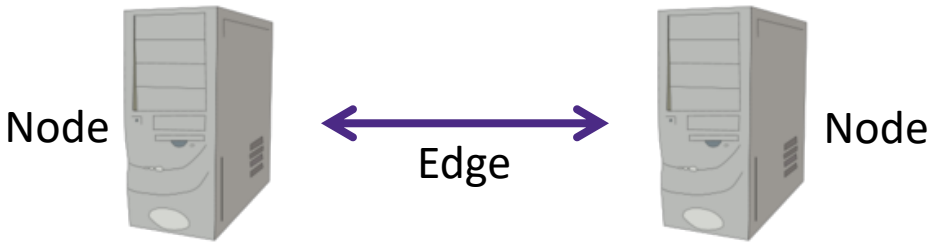


A Simple Model for Networks

- One way to represent computer networks is a **graph**
  - Each **node** represents one machine on the network
  - Each **edge** represents a connection between two machines

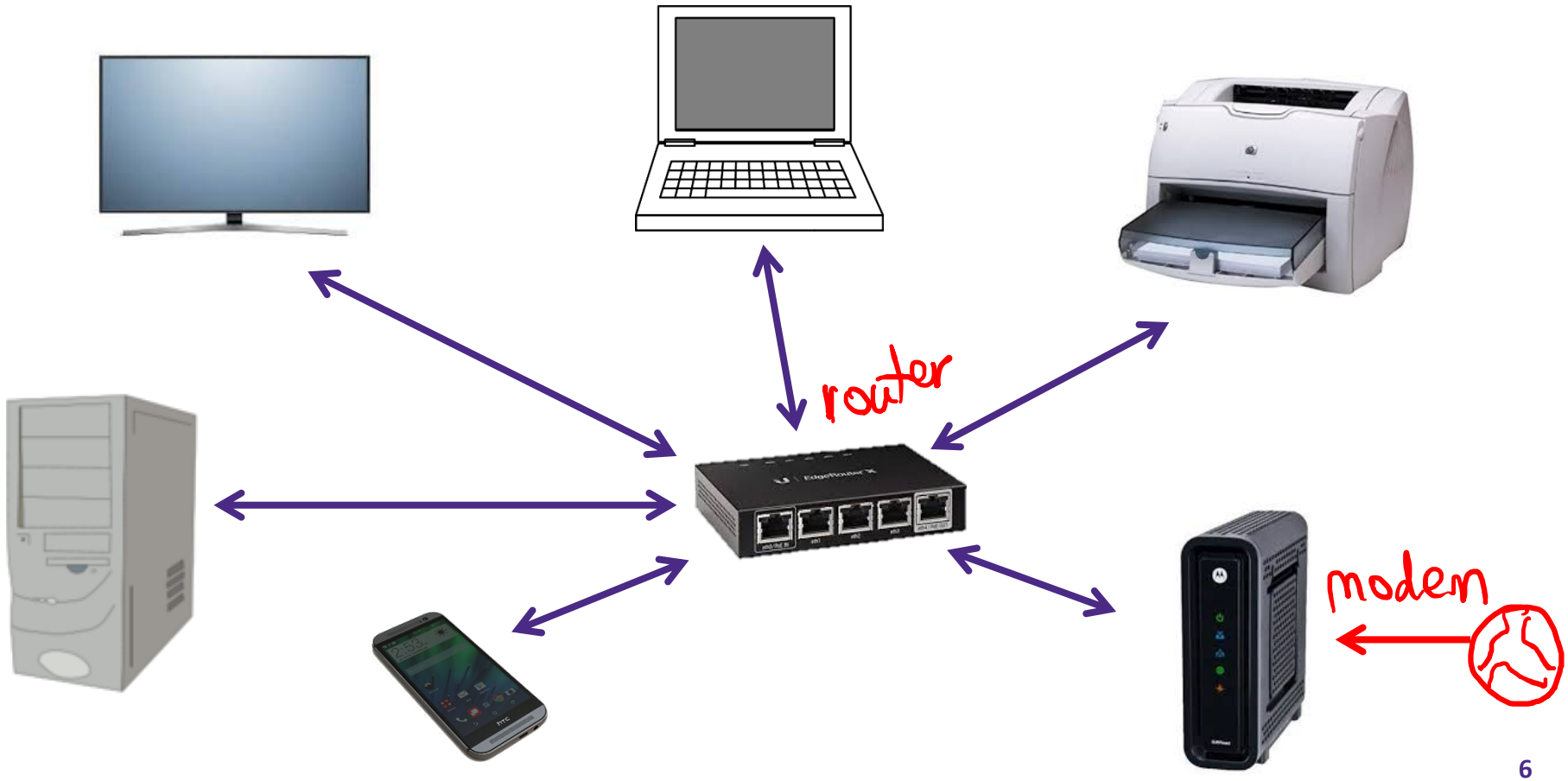


- Below is a network with just two computers
  - 2 nodes, and 1 edge

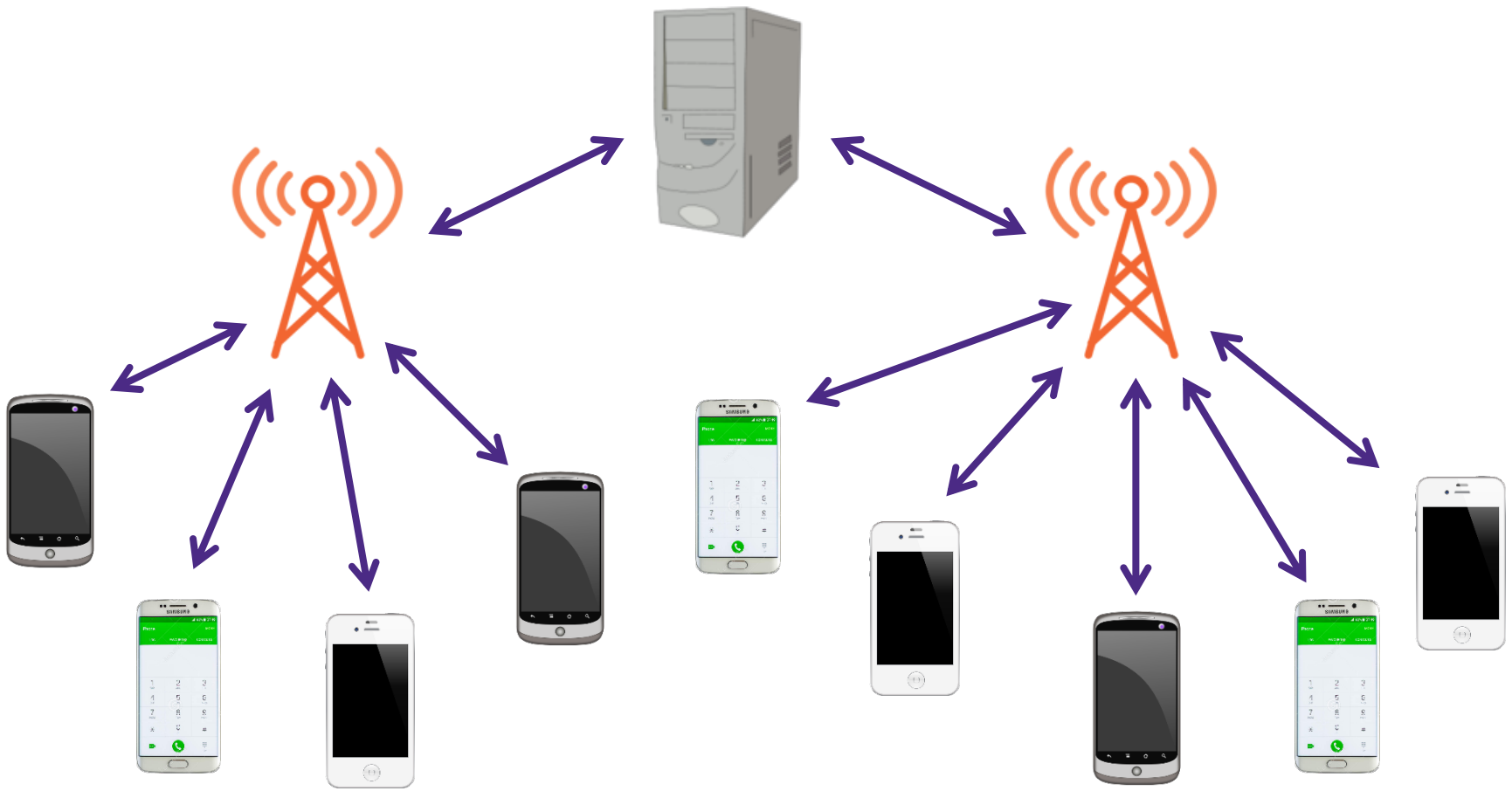


Example: Home Network

- The network at my house: 7 nodes, 6 edges
  - Not counting the outside world

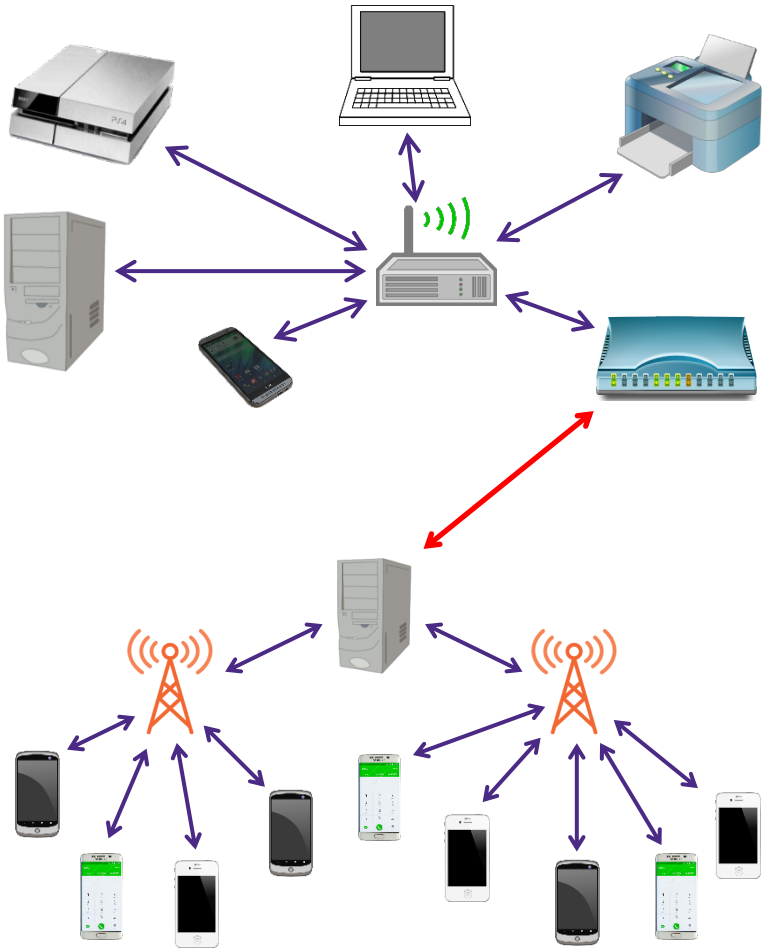


Example: Cellular Network



Internetworking

- If you connect two networks, you still have a network
  - Sometimes called an "internetwork"
- The largest network of networks on the planet is usually called "The Internet"



The Interwebs?

- The Internet: All of the *hardware* and *data* associated with the network of all networks (wires, fibers, switches, routers, servers, files, etc.)

- The World Wide Web: The system used to *access* the Internet (data transmission via browsers, web servers, web services, etc.)

Internet Accessibility

- Can now get Internet almost anywhere:
  - On a bus
  - On a plane
  - On a mountain
  - In outer space

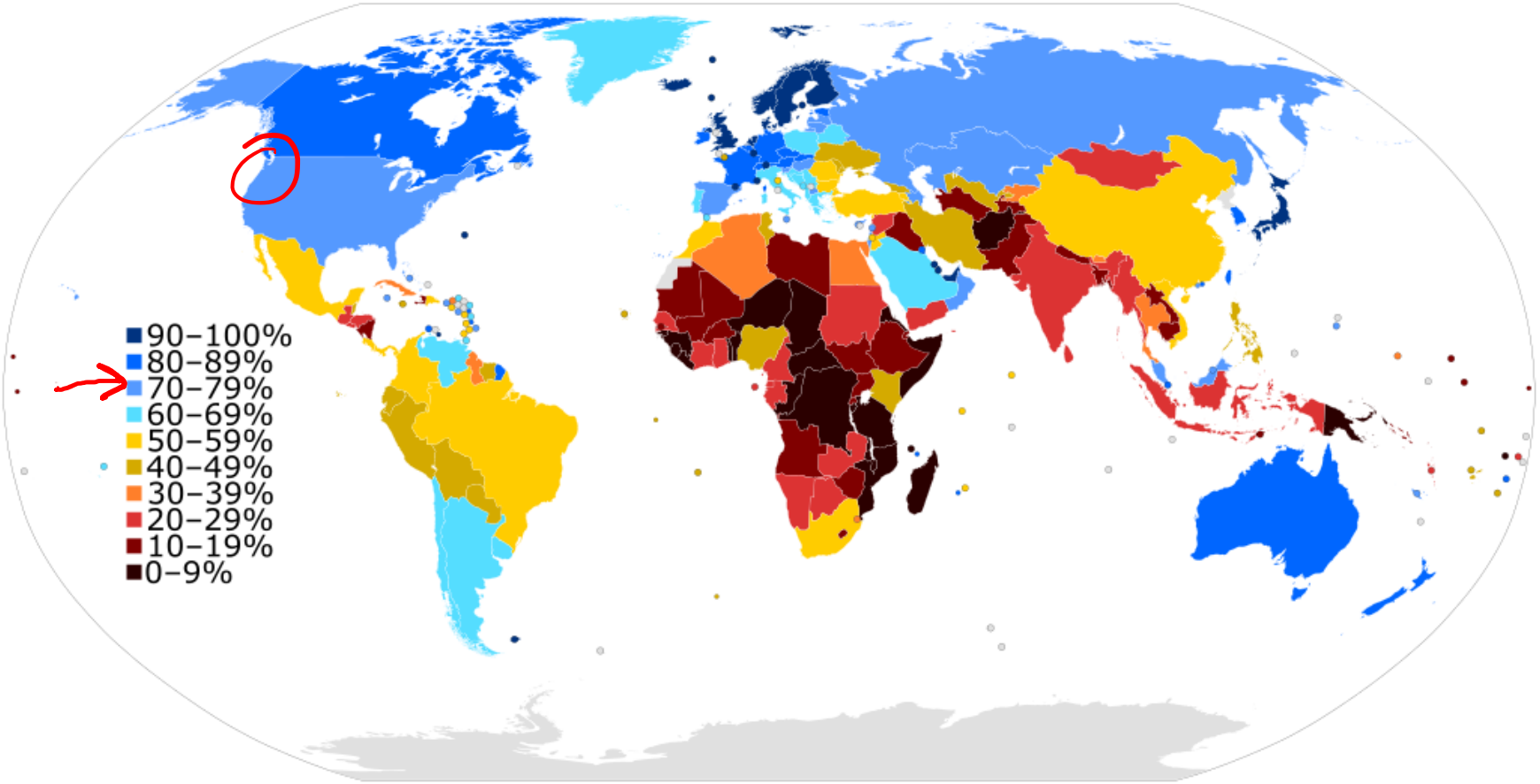
StarTribune

united.com

Slash Gear

NASA

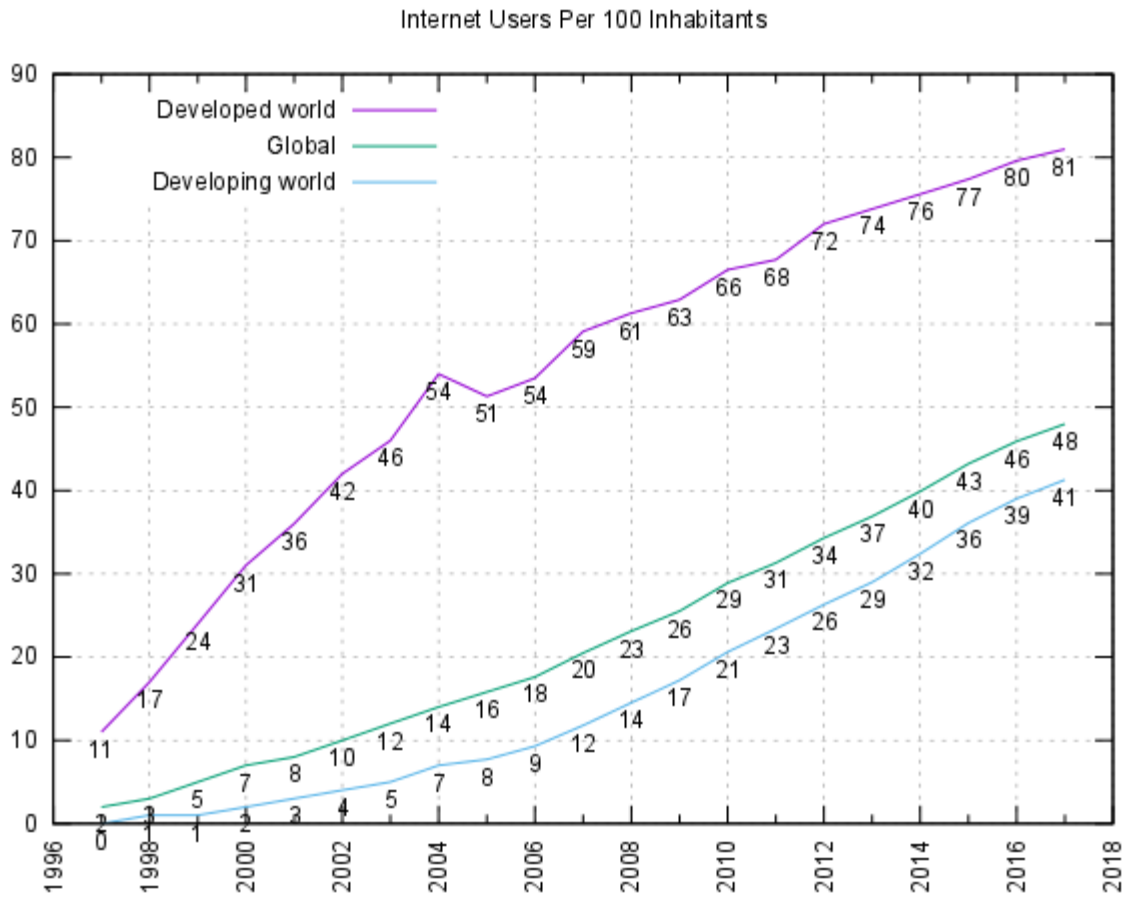
The Internet Today (well, sort of)



Internet Usage as a Percentage of Population (2015)

By Jeff Ogden (W163) - Own work, based on figures from the Wikipedia: List of countries by number of Internet users article in the English Wikipedia, which is in turn based on figures from the International Telecommunications Union (ITU) for 2010 (updated to use figures for 2012 on 28 June 2013). The source code of this SVG is valid. This vector image was created with a text editor. This vector image includes elements that have been taken or adapted from this: BlankMap-World6.svg., CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=19202338

The Internet Today



Internet users per 100 inhabitants

By Jeff Ogden (W163) and Jim Scarborough (Ke4roh) - Own work, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=18972898

Outline

- Networks
- **Growth of the Internet**
- Sending Information
- Encryption
- Data Storage

The DoD and Computer Networks

- The Department of Defense (DoD) observed that central offices made communication network vulnerable to attack
  - 1950s – The Cold War
  - Can we build a more robust, decentralized system?

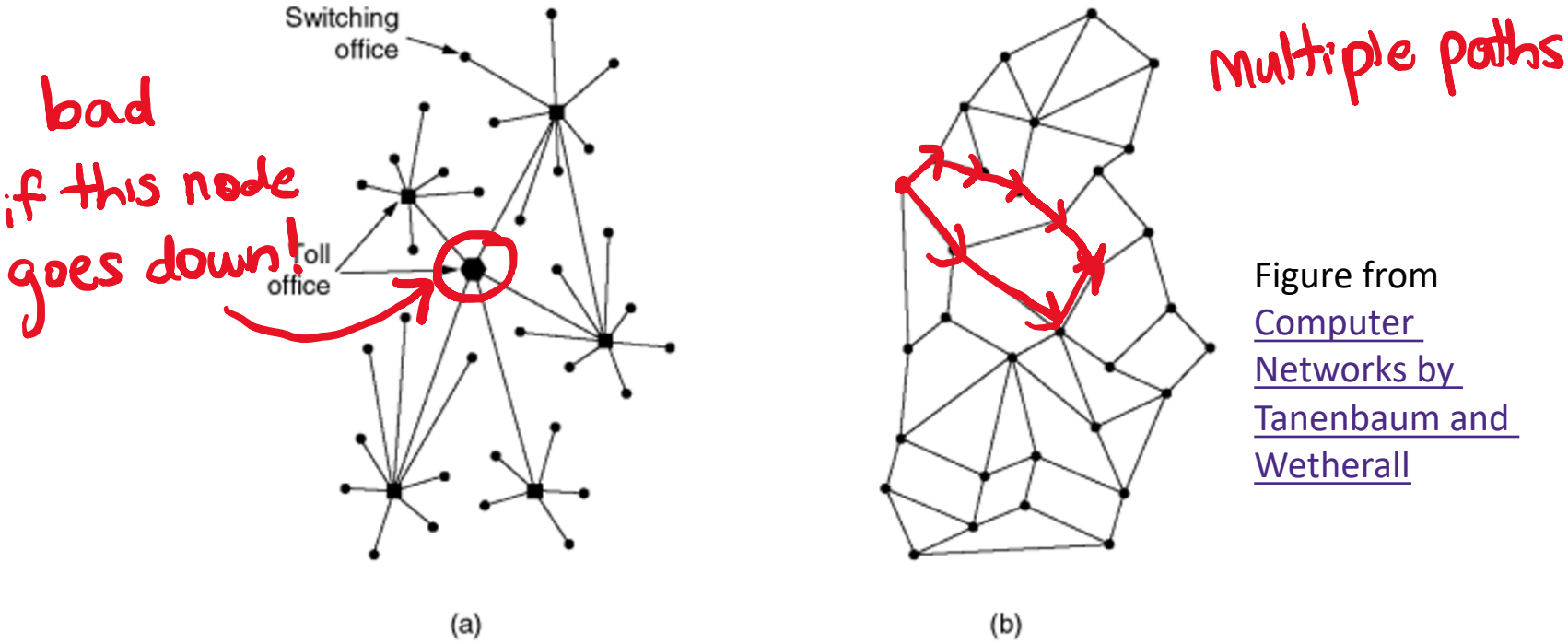
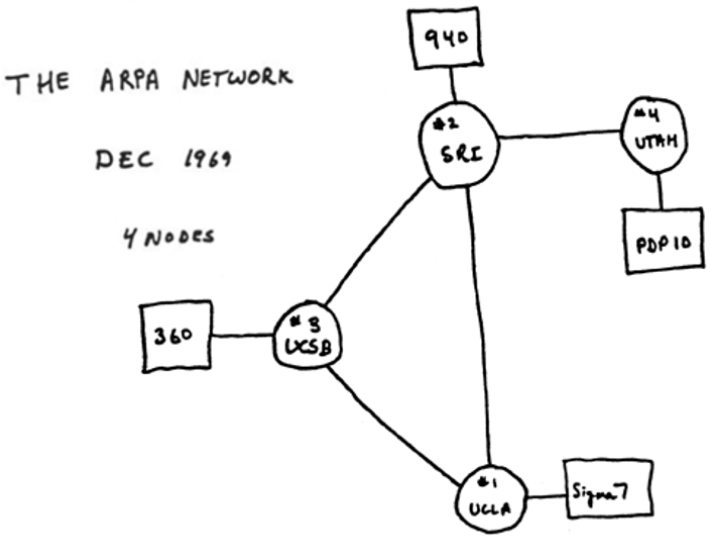


Figure from Computer Networks by Tanenbaum and Wetherall

Figure 1-25. (a) Structure of the telephone system. (b) Baran's proposed distributed switching system.

ARPANET

- First 4 nodes of ARPANET connected in 1969
  - Stanford, UC Los Angeles, UC Santa Barbara, Utah
- By Sept. 1971, there were 18 nodes across the US
  - Grew exponentially from there for a long, long time
  - ARPANET superseded by NSFNET in '86, Internet in '91



Growth of the Internet

- The major point in building networks is *agreement* 
  - The only way to get seamless integration

- Open standards/protocols enabled rapid growth
  - Internet Engineering Task Force (IETF)
    - Request for Comments (RFC)
  - World Wide Web Consortium (W3C)
    - HTML
  - International Standards Organization (ISO)
    - JPEG, MPEG
  - Institute of Electrical and Electronics Engineers (IEEE)
    - Wi-Fi

Outline

- Networks
- Growth of the Internet
- **Sending Information**
- Encryption
- Data Storage

Analogy: Mailing a Letter

- I want to send a letter to my friend in France
  1) Write her unique address on the envelope
  2) Stamp it
  3) Drop it in a mailbox

- I rely on the *abstraction* that the postal service will magically deliver the letter to the specified address

1895

1913

1964

present

IP Addresses

how many IPv4? $2^{32} \simeq 4$ billion
new standard: IPv6 is 128 bits! $2^{128} = 3 \times 10^{38}$

- In 1974, Vint Cerf and Bob Kahn completed the specifications for the Internet Protocol (IP)
  - Every device given a unique 32-bit address (IP address) IPv4
    - Large entities (e.g. companies, universities) can keep an IP address forever and allocate to physical machines as desired
    - For home networks, IP address is typically not permanent
  - Address is used to get information to the right computer on a network

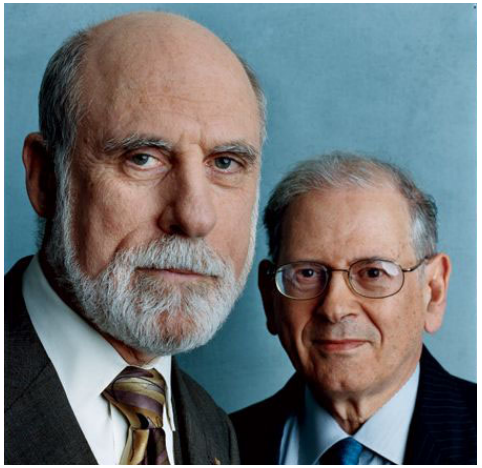
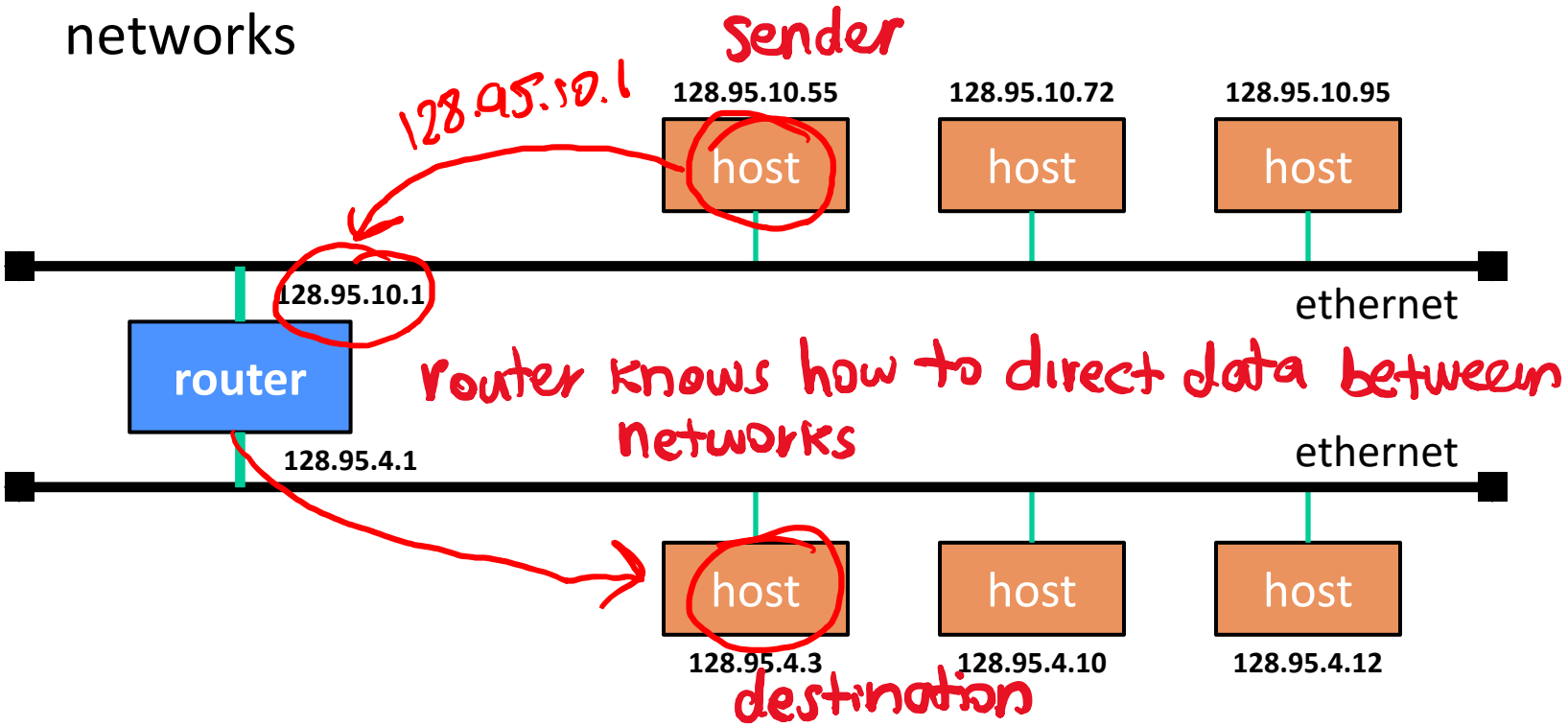


Image: Archeologia Informatica

- Check your IP address: https://www.whatismyip.com

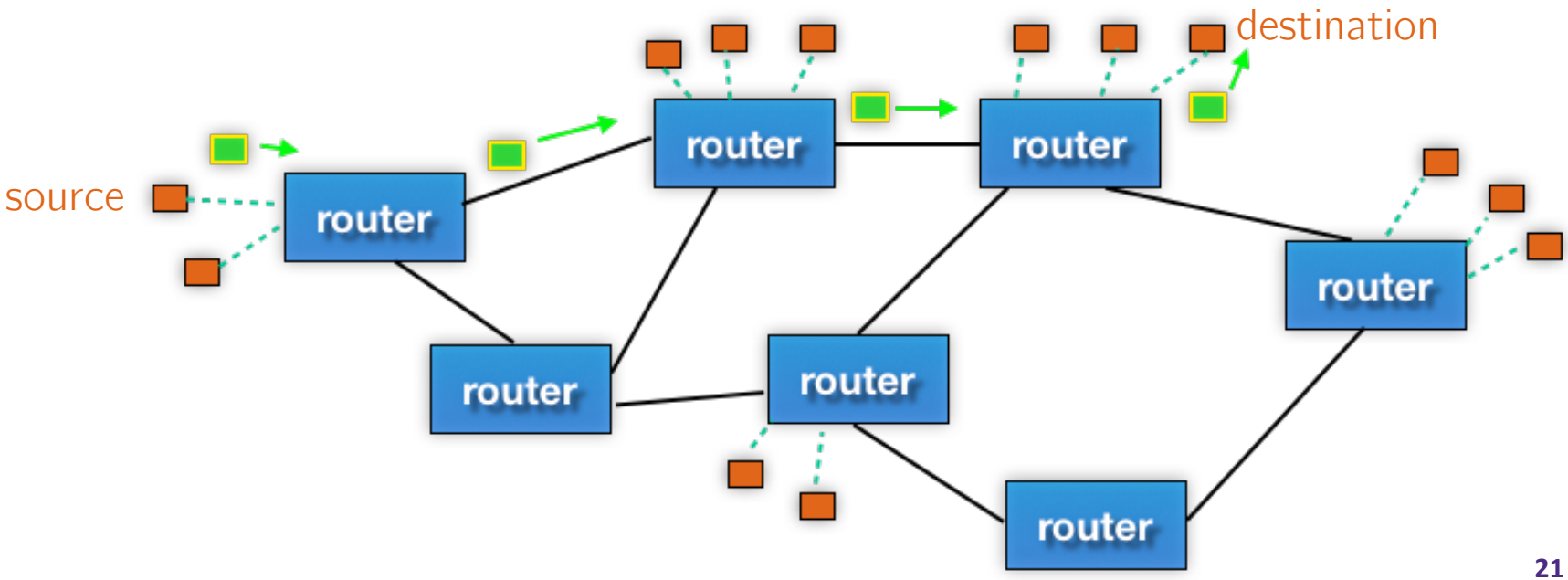
The Internet Protocol

- Internet Protocol (IP) routes data across multiple networks
  - Every computer has a unique IP address
  - Individual networks are connected by routers that span networks



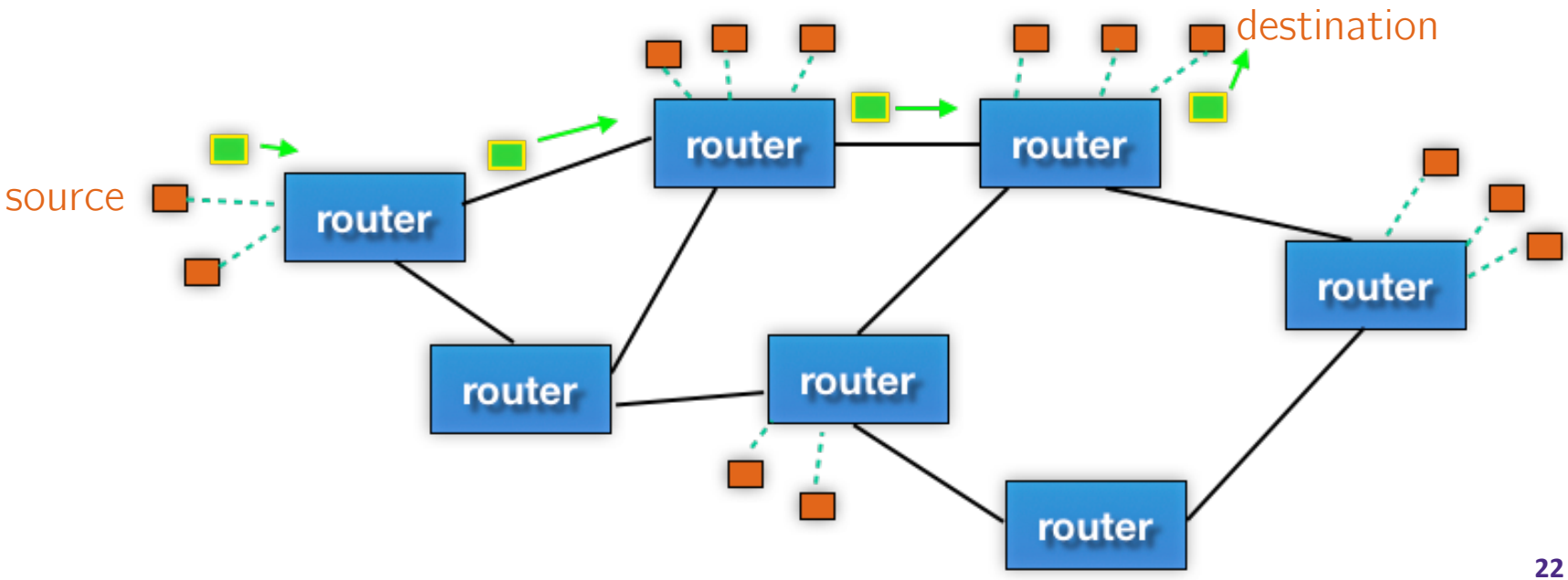
Internet Communication

1) Break the information into lots of tiny pieces called packets, about 1500 bytes long each
2) Packets are sent through the network (passing through many different machines) to their destination
3) The packets are reassembled on the other side



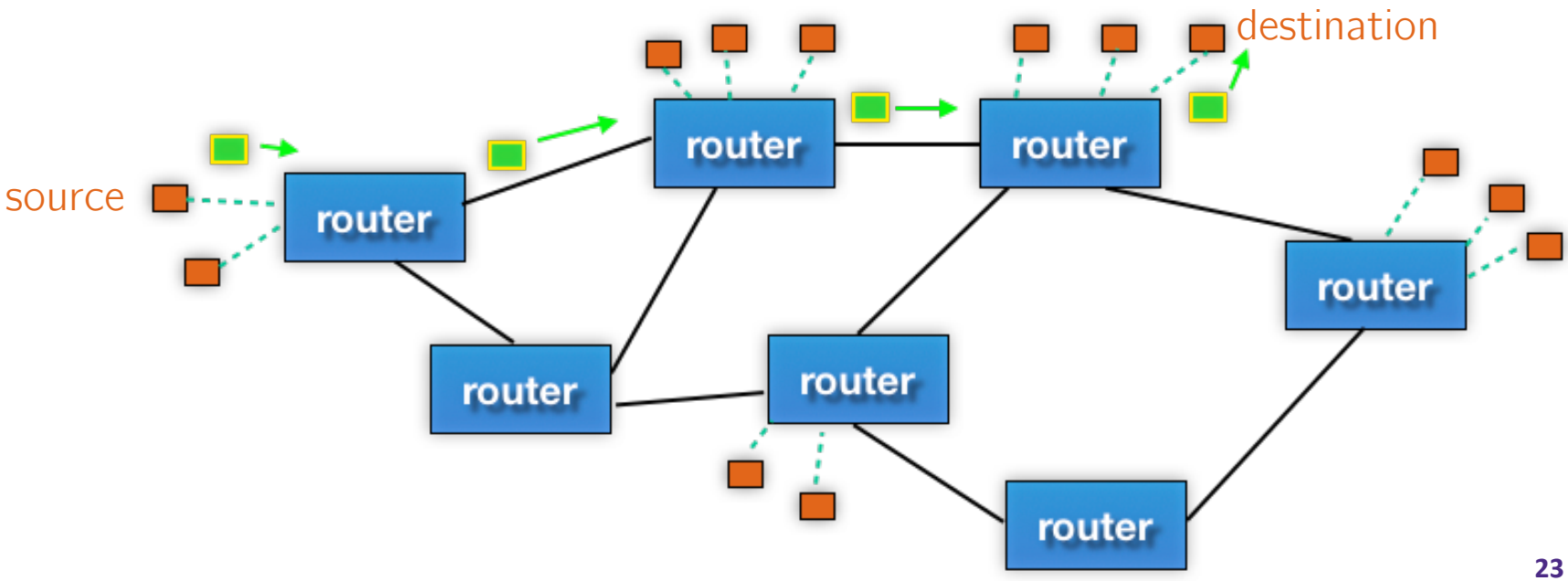
Internet Communication

- Packets must contain:
  - Destination address
  - Sequence/piece number
  - Content/data



Internet Communication

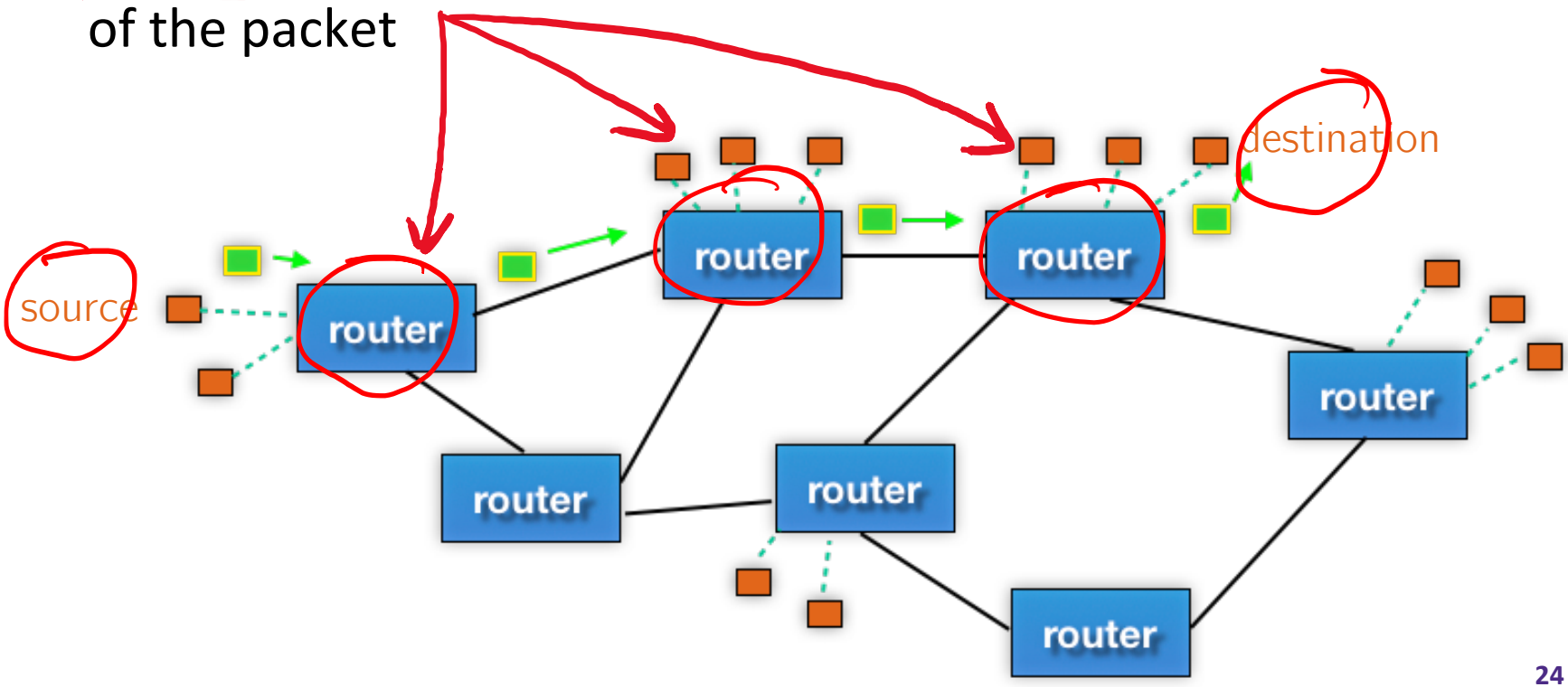
- Advantages:
  - Packets can take separate routes
    - Can even originate from different locations
  - If packet is lost, only must resend small amount of info



Internet Communication

- Disadvantages:
  - Extra transmission data
    - *e.g.* same destination address for many packets
  - Every computer (!!) along a packet's path sees the content of the packet

addr	#	data



Domain Name System

- Remembering IP addresses would be brutal for humans
  - Instead we use domain names, which are human-readable and more flexible
    - *e.g.* `cs.washington.edu` instead of 128.208.3.88

- Computers find IP address for a domain name from the domain name system (DNS)
  - Another computer that acts as an IP address book
    - Your computer *does* need to know the IP address of the DNS server
  - DNS is an automatic directory search – it's huge!

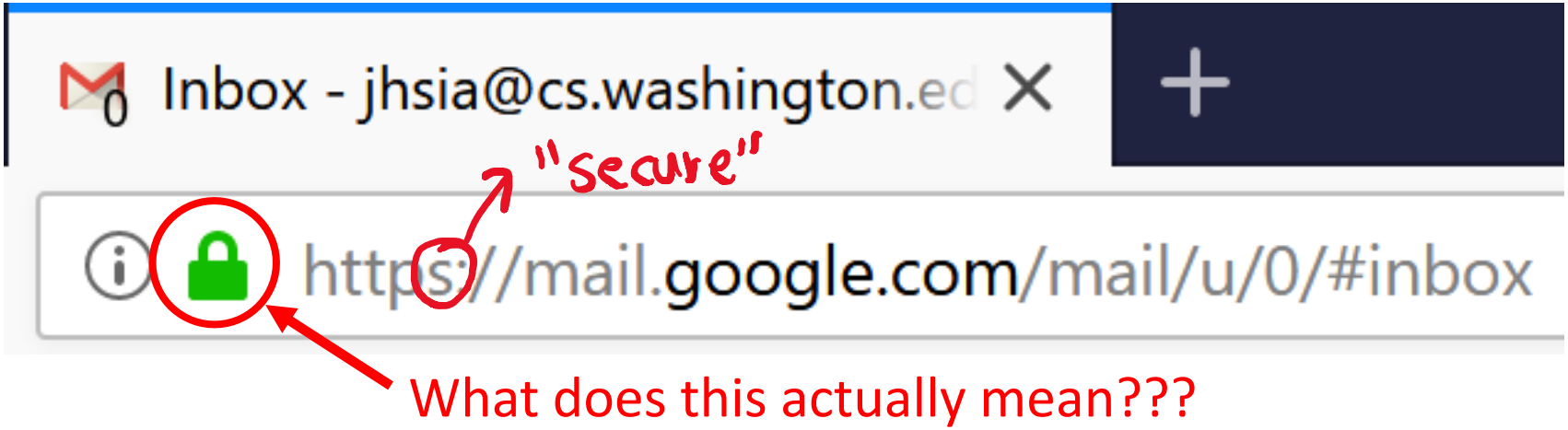
like a phone book: "google.com" ⇒ 8.8.8.4

Outline

- Networks
- Growth of the Internet
- Sending Information
- **Encryption**
- Data Storage

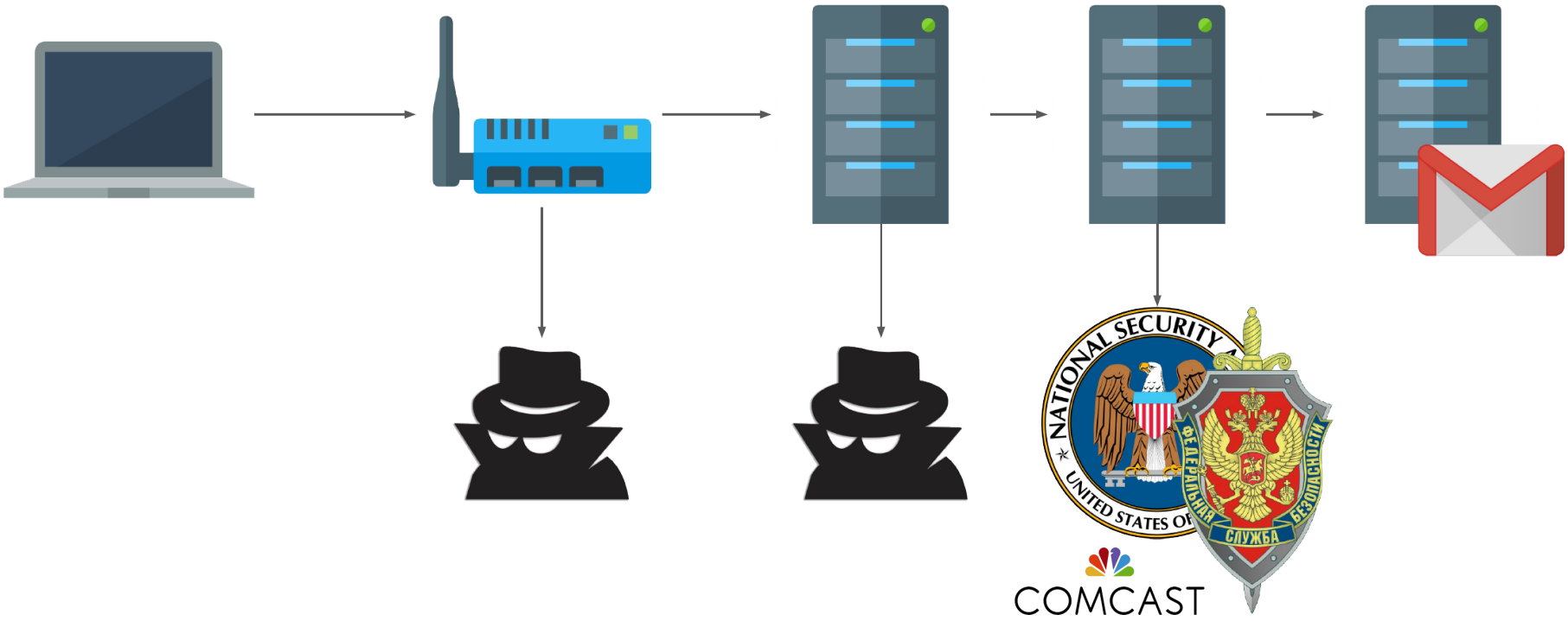
The following slides are courtesy of Prof. Franzi Roesner and Eric Zeng from the Security and Privacy Research Lab at UW CSE.

Something You May Not Have Noticed



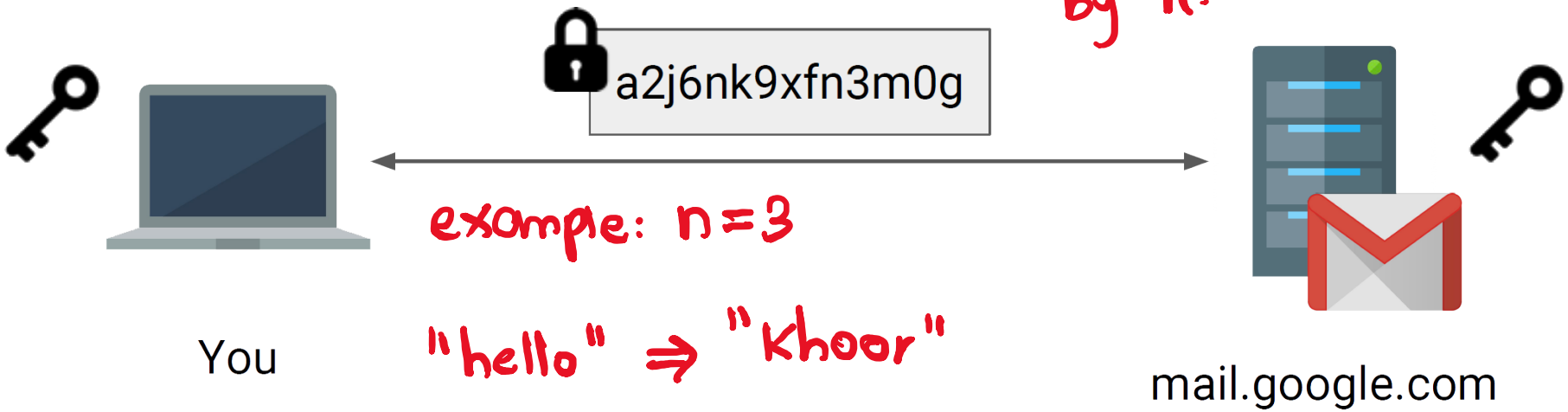
- Your communication with Google is encrypted
- You know that you're talking to Google, as opposed to someone *pretending* to be Google (probably)

Why Encryption?



Symmetric-Key Encryption

Simple example: caesar cipher. shift all characters by n.

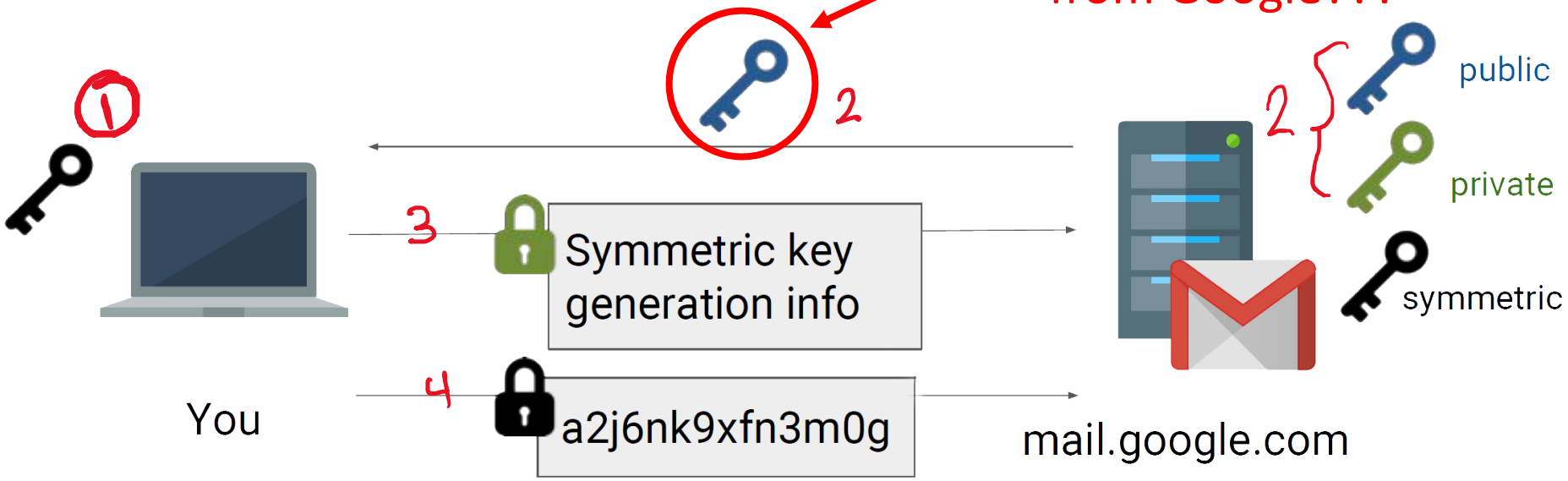


* only recipient knows to shift back by 3

- Use a secret key to both *encrypt* and *decrypt* the message/data
  - Both parties must have access to the secret key
- How do we exchange keys???

Asymmetric Encryption

How do we know that this key is really from Google???

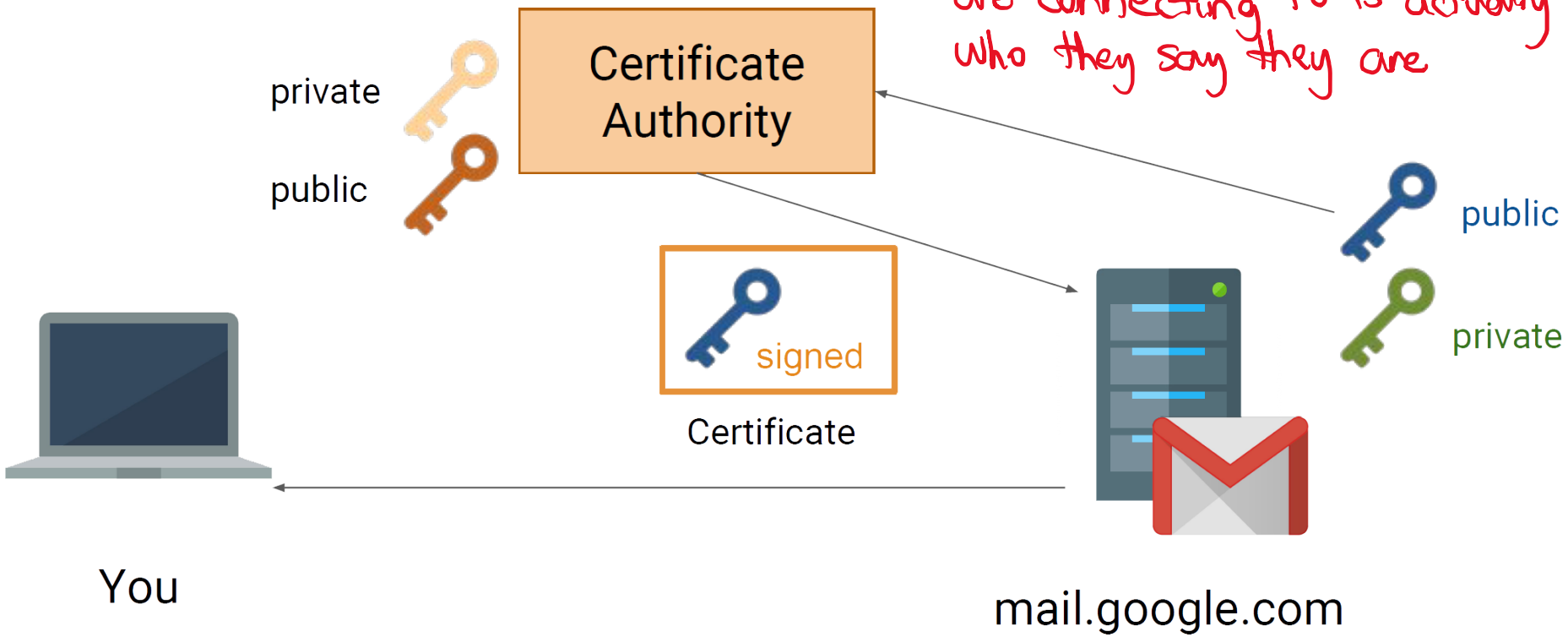


- Use public key encryption to bootstrap symmetric key
  - Much slower to do public key encryption than symmetric

1) I generate symmetric key to use with Google
2) Server generates public-private keypair & sends me public key → (can decrypt but not encrypt)
3) I encrypt symmetric key w/ public key & send it back. Only google can decrypt.
4) Now we can communicate w/ private key

Certificate Authorities

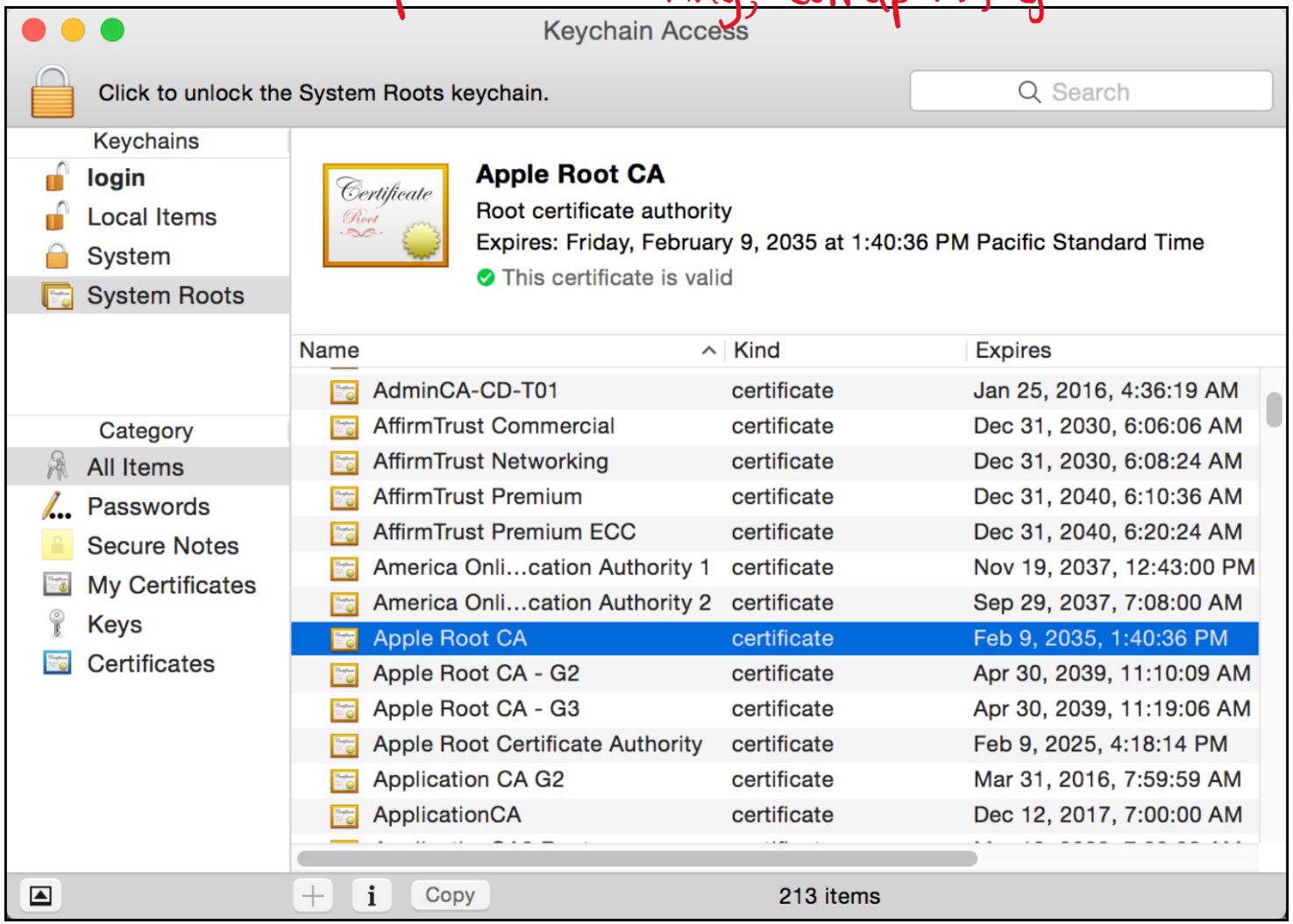
third-parties that can confirm that the server you are connecting to is actually who they say they are



- Your browser knows some trusted authorities!

Trusted(?) Certificate Authorities

not perfect—hacking, corruption, gov. intervention



It's turtles all the way down

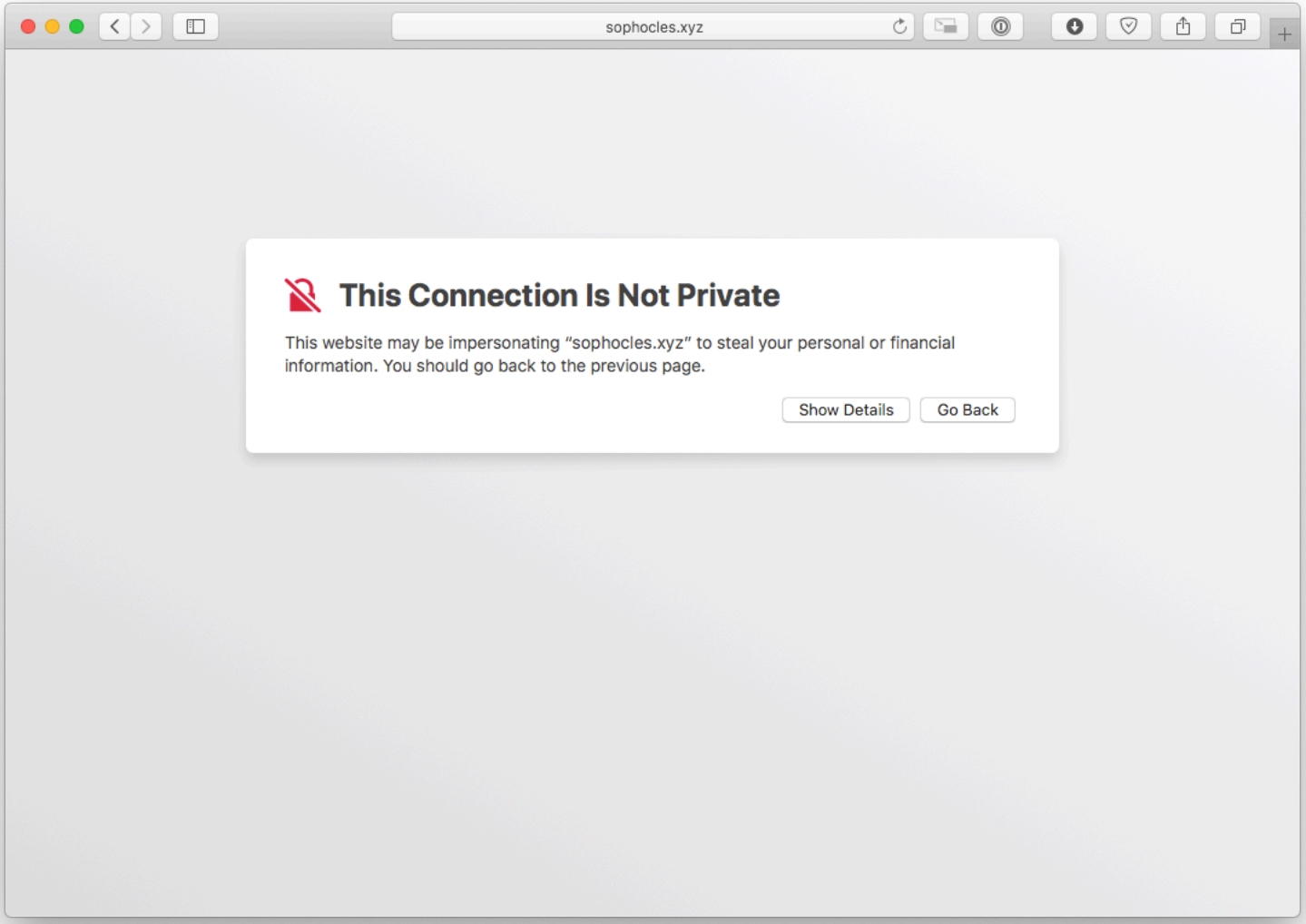
- Used to describe a problem that seems to have infinite dependencies

- How do we know we can trust certificate authorities?
  - How to verify the verifiers?
  - How to verify the verifier-verifiers?
  - How to verify the verifier-verifier-verifiers?

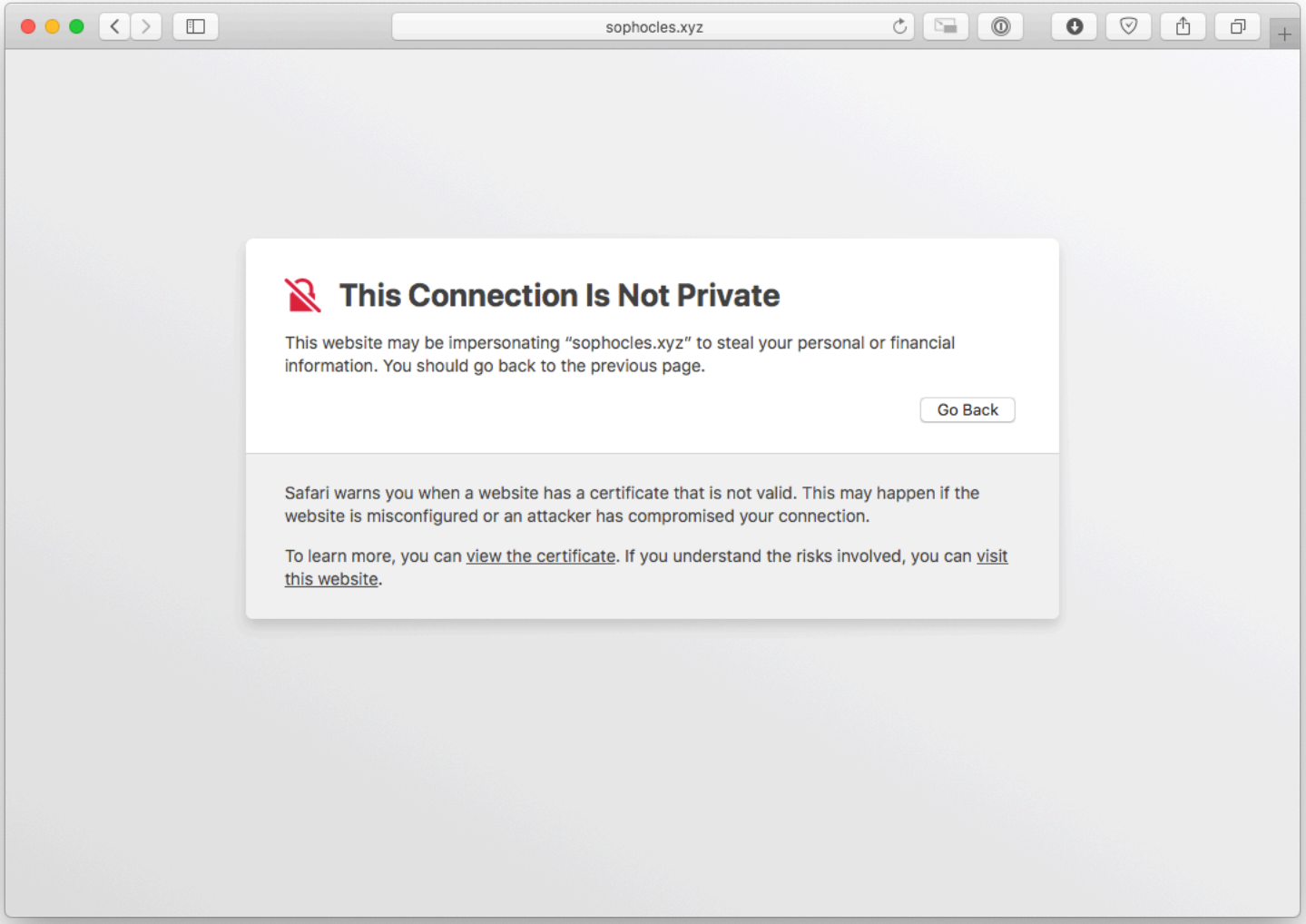
- At some point, we just have to trust without verification :/

What If the Certificate is Bad?

gmail.com



What If the Certificate is Bad?



Summary

- A network is a group of computing devices connected together, either by wire or wirelessly
  - The Internet is the largest network of networks
- The Internet grew rapidly
  - Highly fault-tolerant due to decentralization
  - Growth aided by open standards (agreement)
- Data is passed between computing devices in small pieces called packets
  - The domain name system translates from domain names to IP addresses in order to reach a specific device
- Encryption helps us secure data transmissions between devices