INFO 654 – Information Technologies

Week 3: Networking and the Internet

“What is Internet anyway…?”

-- “Today” show, 1995

Week 3 Agenda

I. Updates, report-back
II. An overview of basic telecommunications
III. An overview of networking
IV. “What is Internet?” and development of the Internet

<break>

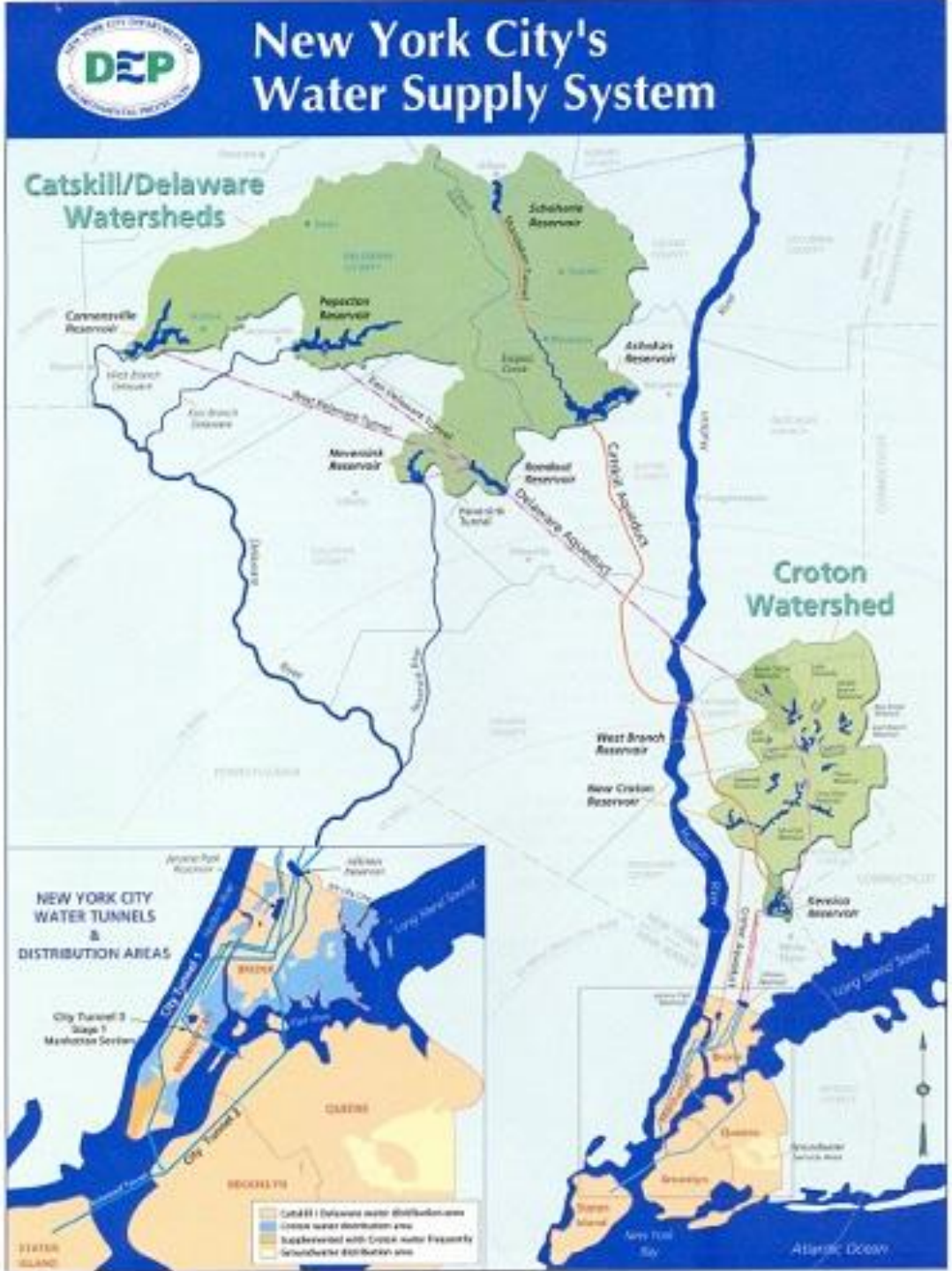
I. More on TCP/IP
II. The World Wide Web
III. Discuss upcoming topics + current events presentations
IV. The Internet Archive & Wayback Machine

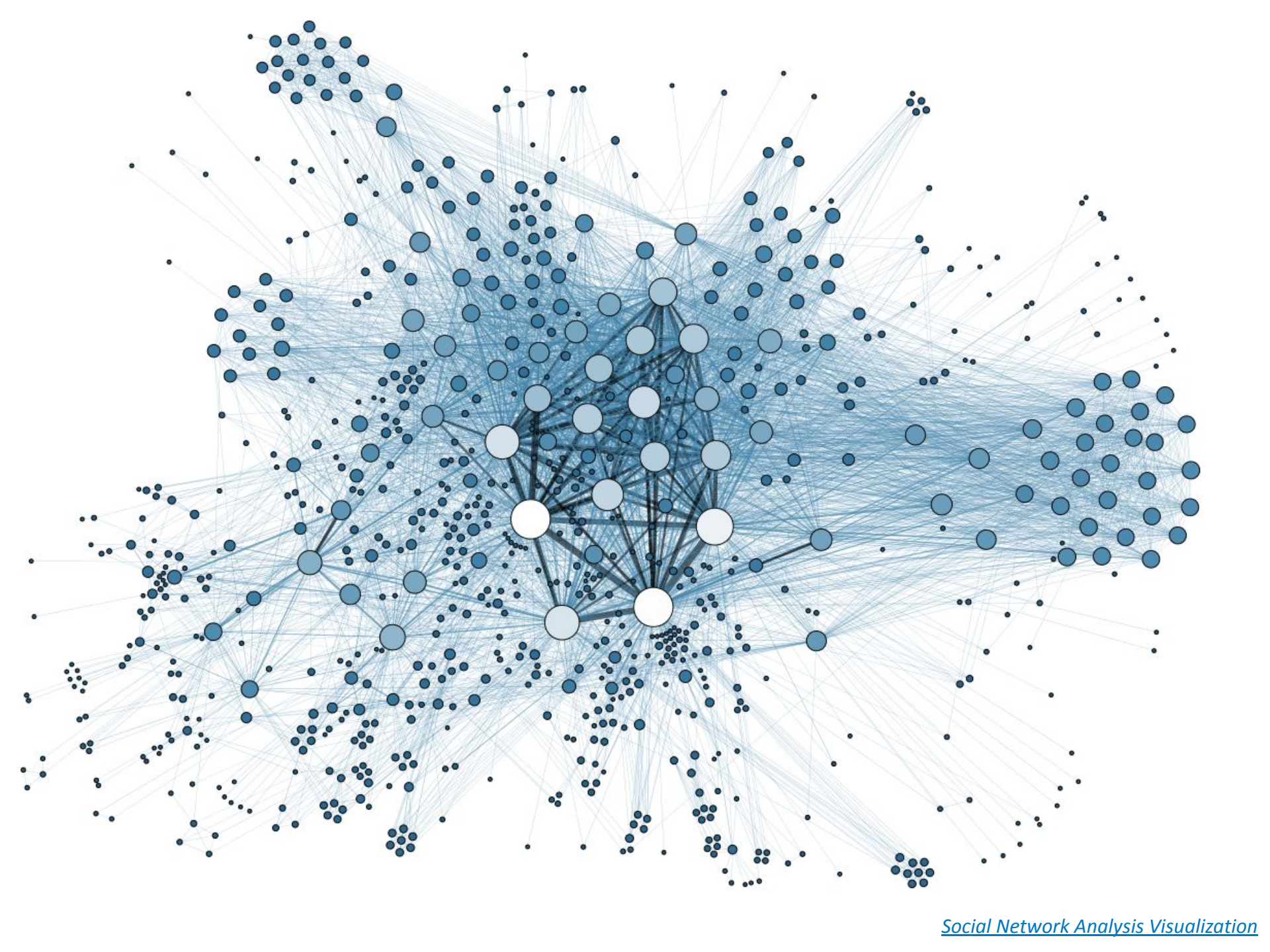
What did you find?

CAREERS AND INFORMATION TECHNOLOGY

What Is a Network…?
(conceptually)

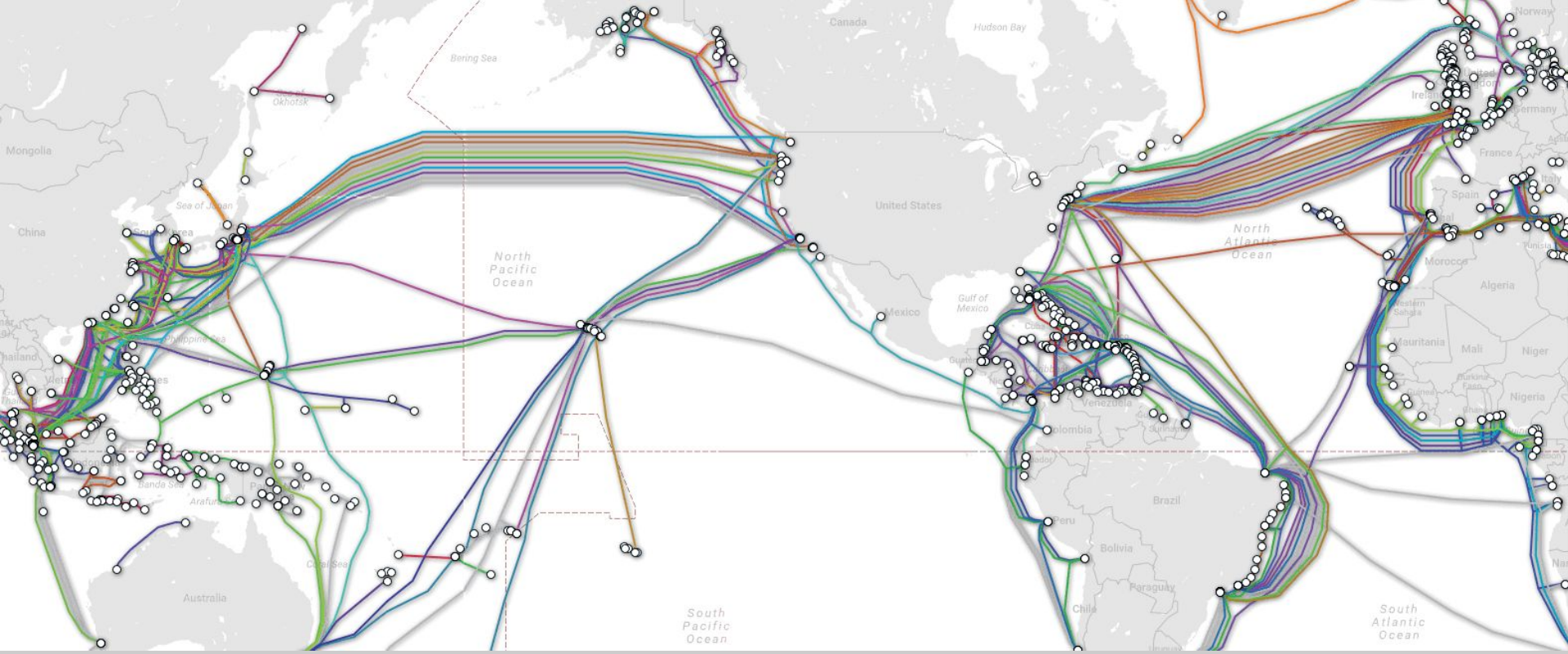
New York City's Water Supply System

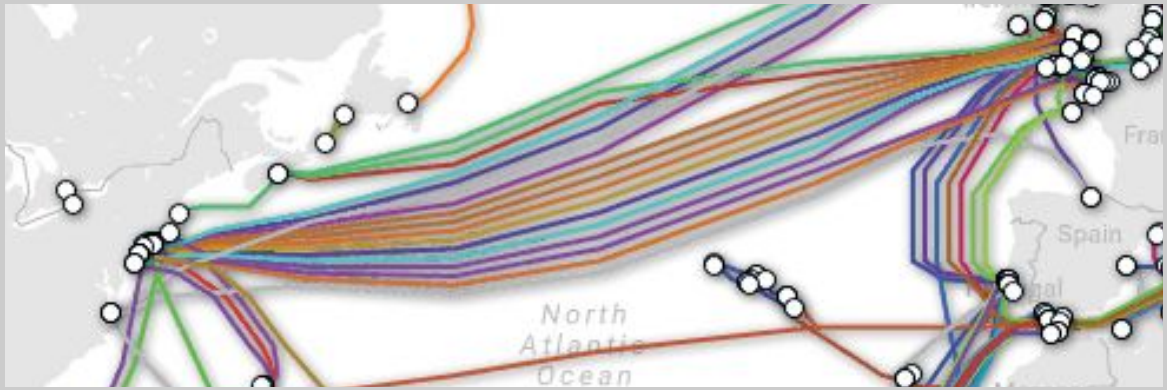


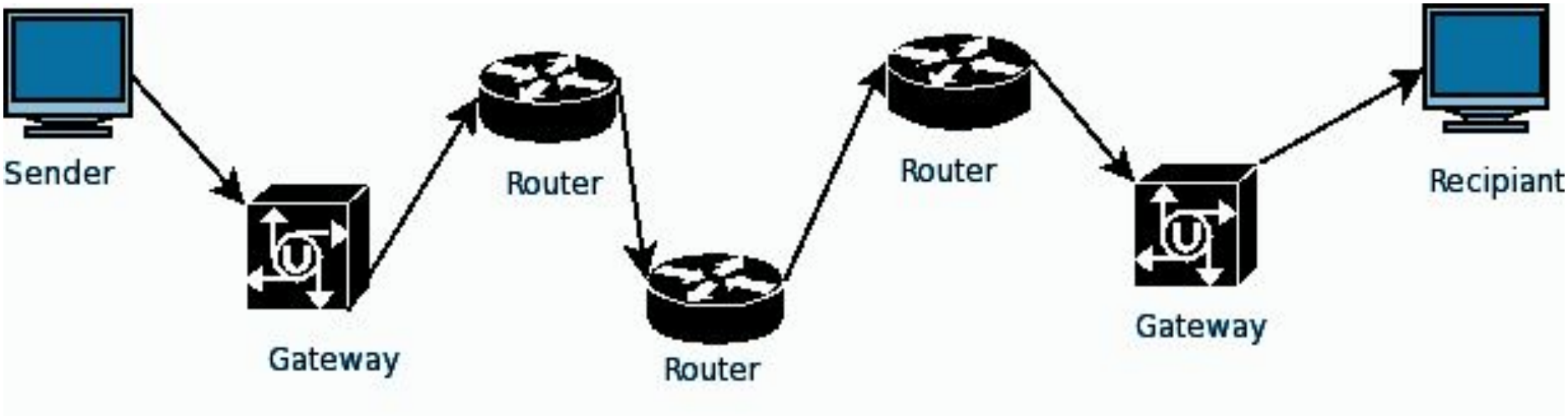


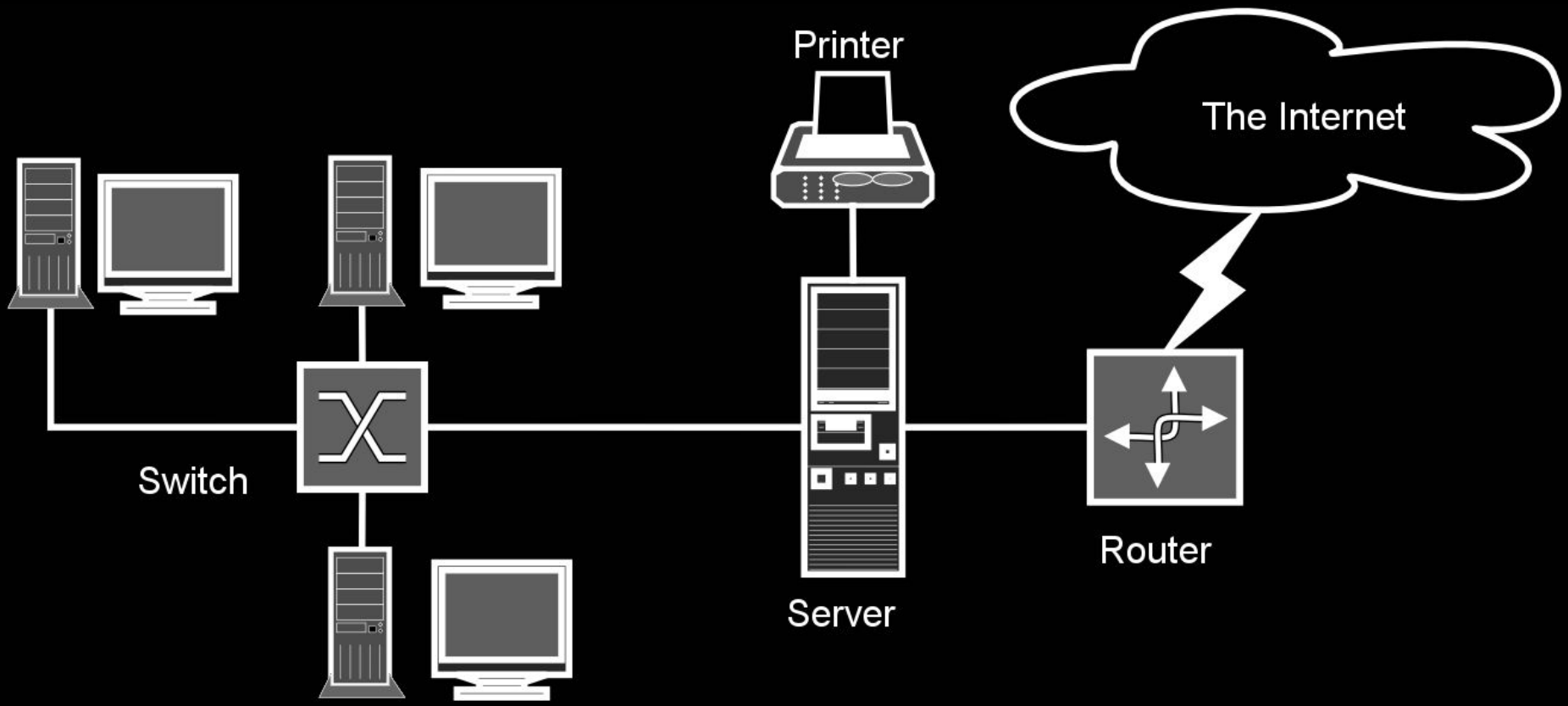
Social Network Analysis Visualization

Photo by the CTBTO Preparatory Commission


Canada
Hudson Bay
Bering Sea
Okhotsk
Mongolia
China
Sea of Japan
United States
North Pacific Ocean
Gulf of Mexico
Mexico
Philippine Sea
Banda Sea
Arafura Sea
Coral Sea
Australia
South Pacific Ocean
North Atlantic Ocean
Norway
Ireland
France
Spain
Morocco
Algeria
Western Sahara
Mauritania
Mali
Niger
Nigeria
Venezuela
Colombia
Brazil
Peru
Bolivia
Paraguay
Chile
South Atlantic Ocean


Spain
North Atlantic Ocean


Sender
Gateway
Router
Router
Router
Gateway
Recipiant


Printer
The Internet
Switch
Server
Router

AN OVERVIEW OF TELECOMMUNICATIONS

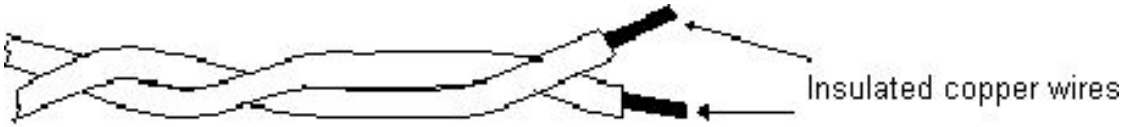
basic telecommunication concepts

Key Terminology

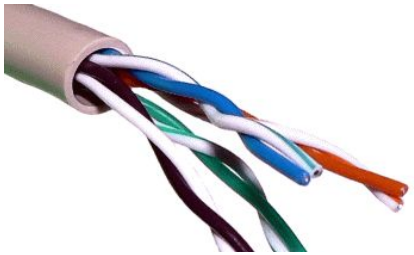
- **Telecommunications** is the electronic transmission of signals for communications
  - Communicating data from a sender(s) to a receiver(s) requires a pathway or **communication channel**

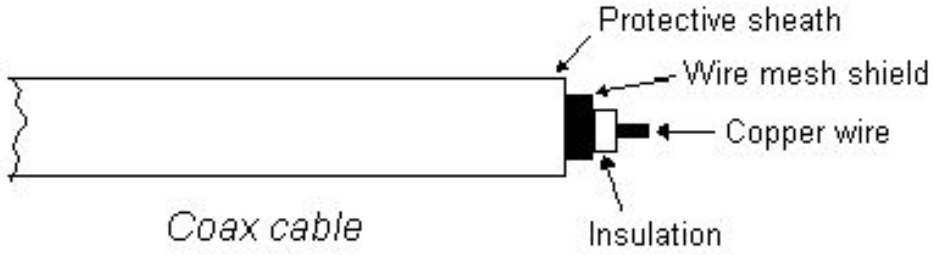
- **Networking protocol** is a set of rules, algorithms, messages, and other mechanisms that enable software and hardware in networked devices to communicate effectively, i.e. a standard

Guided Transmission Media

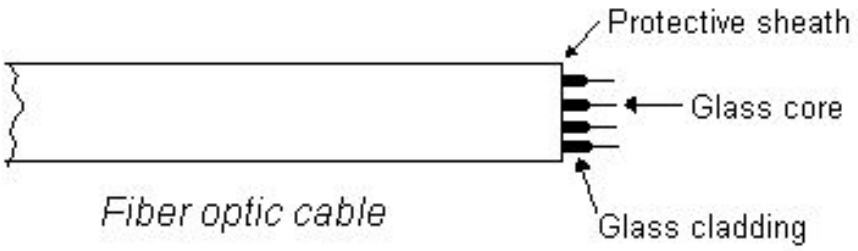


Twisted-pair cable

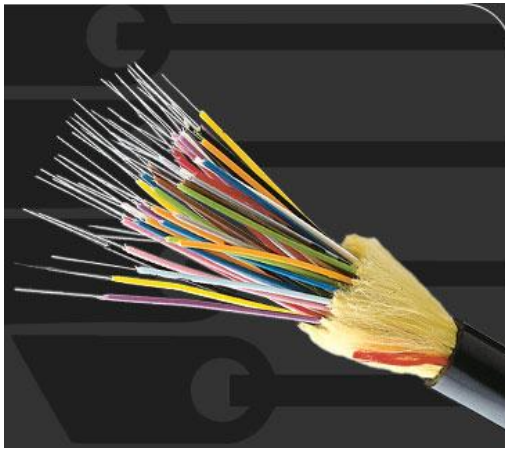




Coax cable



Fiber optic cable



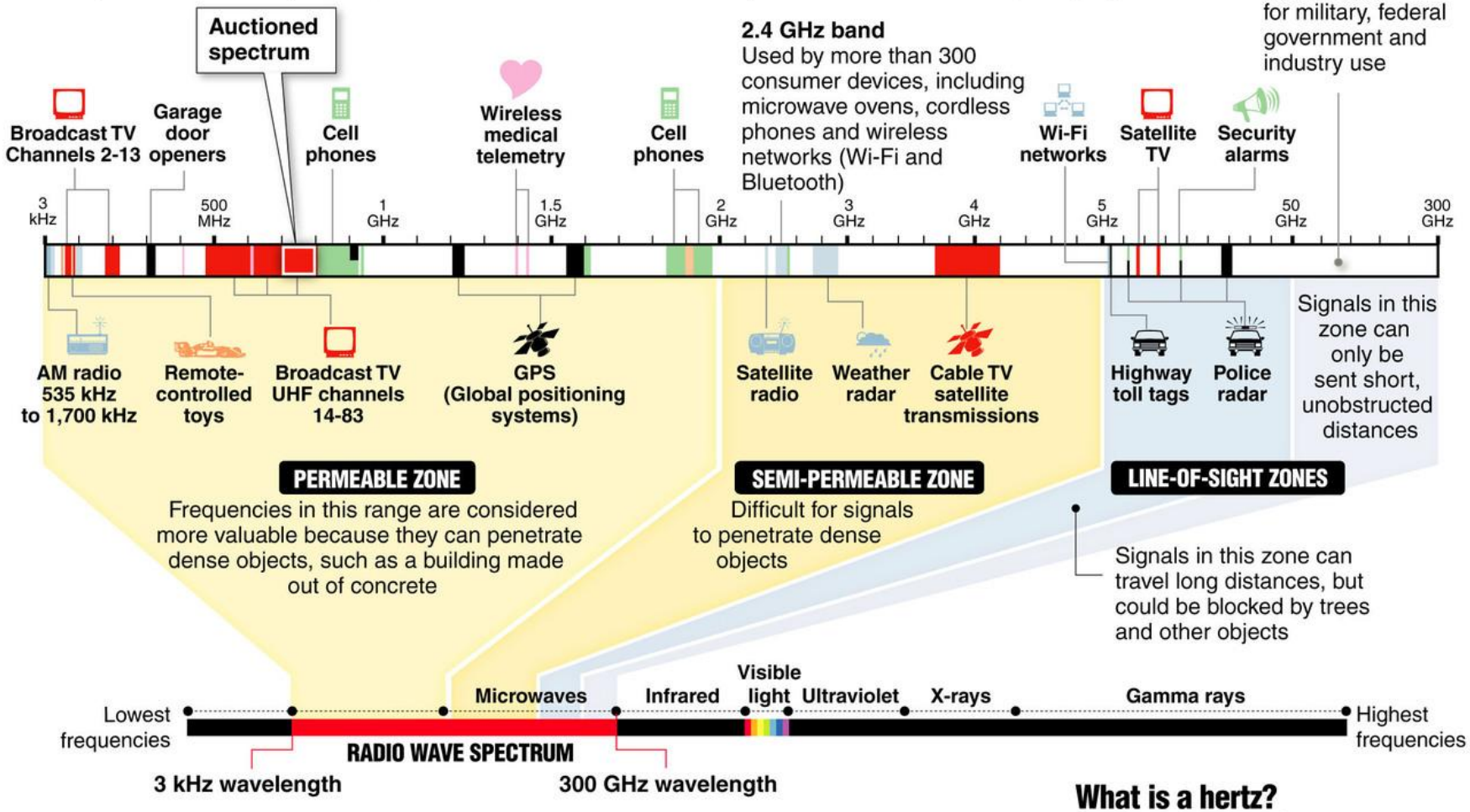
Cable types

Wireless Communication

- **Wireless** transmission involves the *broadcast* of communications in one of three frequency ranges
  - Radio
  - Microwave
  - Infrared frequencies
- In some cases, use of wireless communications is *regulated*
  - Signal must be broadcast within a specific frequency range to avoid interference with other wireless transmissions

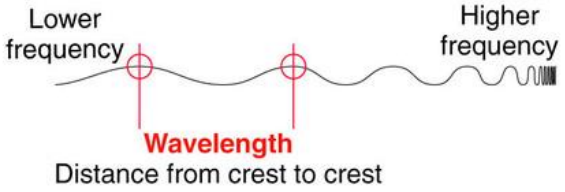
Inside the radio wave spectrum

Almost every wireless technology – from cell phones to garage door openers – uses radio waves to communicate. Some services, such as TV and radio broadcasts, have exclusive use of their frequency within a geographic area. But many devices share frequencies, which can cause interference. Examples of radio waves used by everyday devices:



The electromagnetic spectrum

Radio waves occupy part of the electromagnetic spectrum, a range of electric and magnetic waves of different lengths that travel at the speed of light; other parts of the spectrum include visible light and x-rays; the shortest wavelengths have the highest frequency, measured in hertz



What is a hertz?

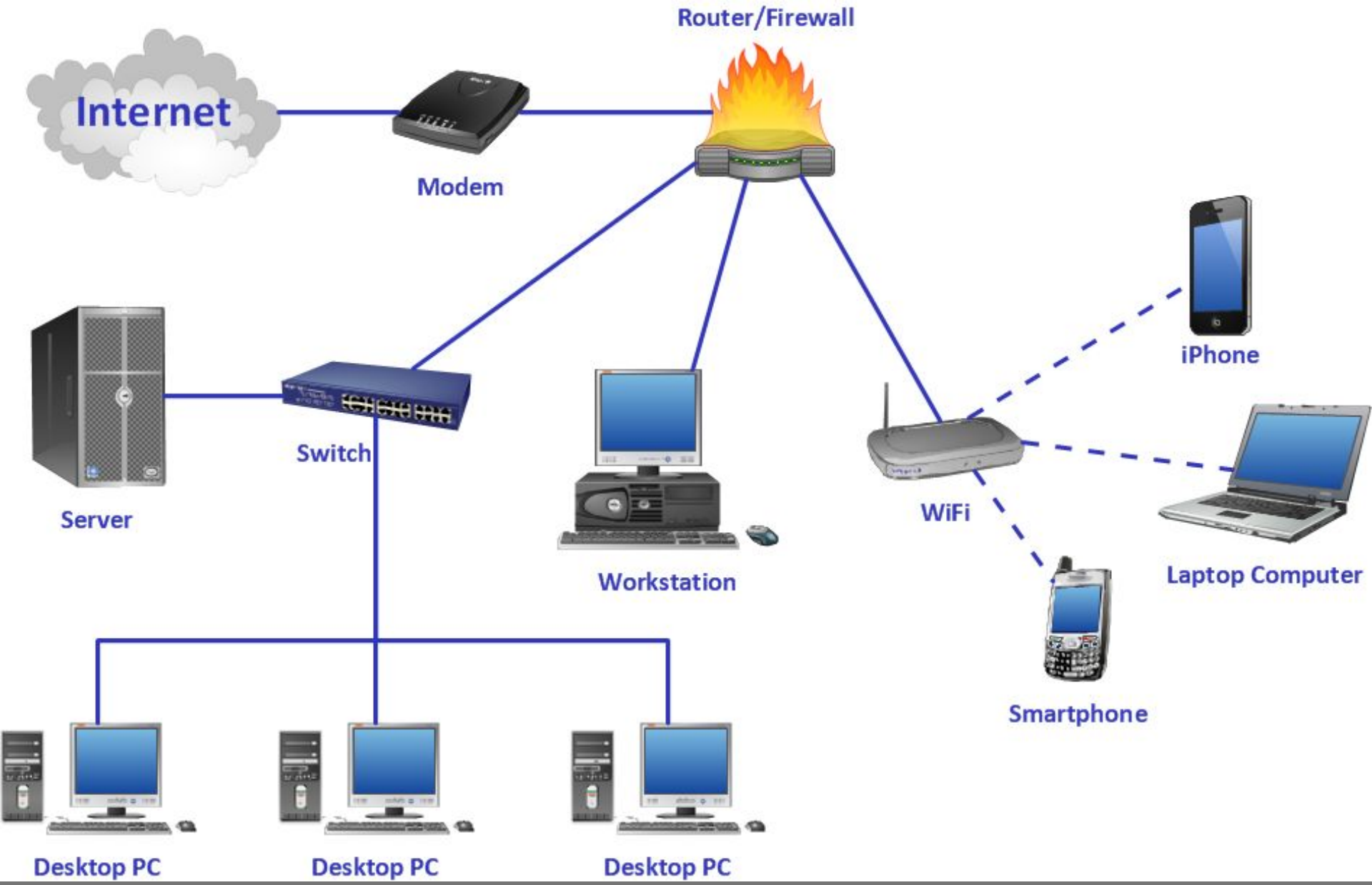
One hertz is one cycle per second. For radio waves, a cycle is the distance from wave crest to crest

1 kilohertz (kHz) = 1,000 hertz

1 megahertz (MHz) = 1 million hertz

1 gigahertz (GHz) = 1 billion hertz

Source: New America Foundation, MCT, Howstuffworks.com

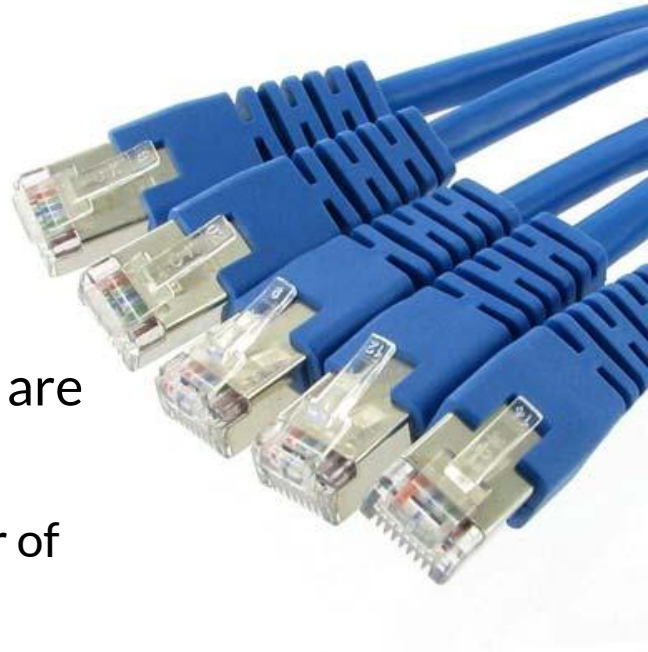

Router/Firewall
Internet
Modem
iPhone
Server
Switch
Workstation
WiFi
Laptop Computer
Smartphone
Desktop PC
Desktop PC
Desktop PC

NETWORKS

Network Types

- **Personal area networks (PAN)**
  - Support interconnection of information technology within a range of about 33 feet, e.g. Bluetooth headset
- **Local area networks (LAN)**
  - Connect computer systems and devices within a small area, e.g., office or home
- **Metropolitan area networks (MAN)**
  - Connect users and their devices in a geographical area that spans a campus or city, e.g. a cable tv network
- **Wide area networks (WAN)**
  - Connect large geographic regions

Local Area Network (LAN)



- A **local area network (LAN)** is when computers are geographically close
  - Usually they can be linked by a single cable or pair of wires

- **Ethernet** is the main technology for local area networks
  - Used for connecting all the computers in a lab or building
  - Signal (data sent as electricity) can only travel limited distances
  - Signal read by all devices on the network, take turns transmitting

Wireless Networks

- Variation of a **LAN** connection
- The router is:
  - Physically connected to an ISP's modem
  - Connected to the Internet
  - Capable of *broadcasting* **and** *receiving* signals, usually radio frequency (rf) signals

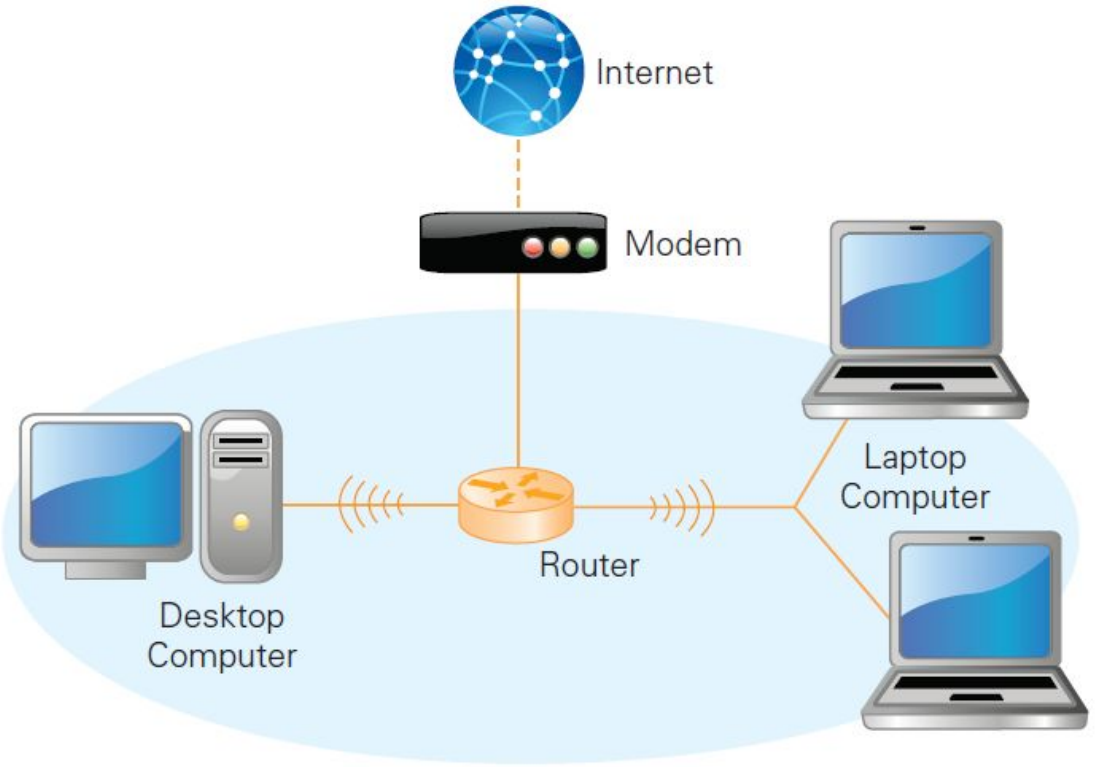
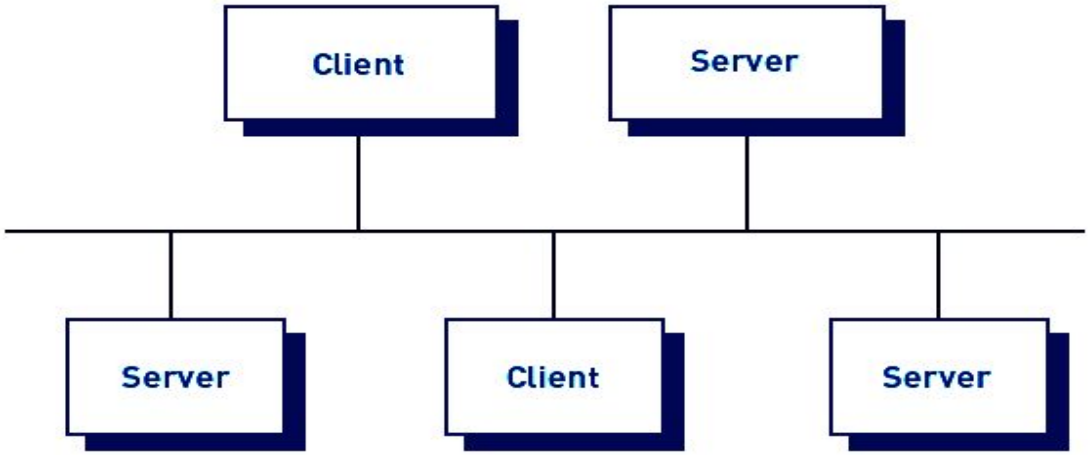
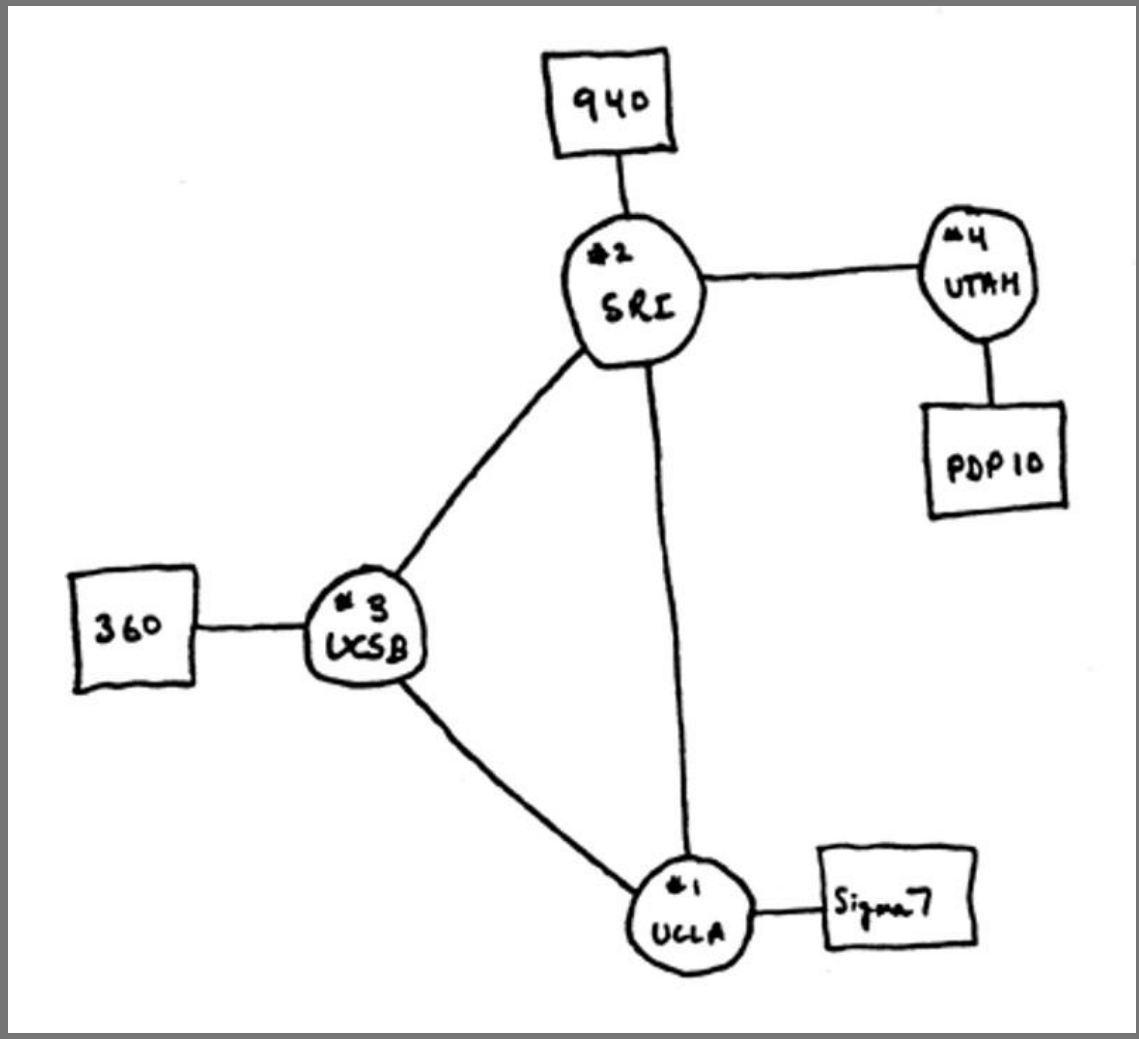


Figure 3.11 Standard Wi-Fi network configuration. A wireless router is connected via the modem to the ISP's Internet modem; laptops and other wireless-enabled devices connect by radio signals to the router.

Client/Server Architecture



- **Client/server architecture:**
  - Multiple computer platforms are dedicated to special functions
- **Client:**
  - Any computer that sends messages requesting services from the servers on the network, e.g. web browser requesting a web page
- **Server:**
  - Provider of a resource or service, e.g. web server, printer, etc.



THE INTERNET

“What is Internet anyway…?”

-- “Today” show, 1995

https://youtu.be/UlJku_CSyNg

https://youtu.be/4aIkMwUeL_Q

What Is The Internet?

- The **Internet** is a *worldwide* system of interconnected computer networks that use the **TCP/IP** set of network protocols to communicate with each other
  - Made up of computers, network hardware such as routers and fiber-optic cables, software, and the **TCP/IP** protocols…
- **Internet Protocol (IP):**
  - Enables computers to route communications traffic from one network to another
- **Transmission Control Protocol (TCP):**
  - Transport-layer protocol that most Internet applications use with IP
- The Internet supports **point-to-point asynchronous communication** – more on this shortly!

Evolution of the Internet

The Internet originated as **ARPANET** in September 1969 and had two main goals:

Allow scientists at different physical locations to share information and work together

Function even if part of the network were disabled or destroyed by a disaster

The “Intergalactic Computer Network” -- the underlying concept

ADVANCED RESEARCH PROJECTS AGENCY

Washington 25, D.C. April 23, 1963

MEMORANDUM FOR: Members and Affiliates of the Intergalactic Computer Network

FROM: J. C. R. Licklider

SUBJECT: Topics for Discussion at the Forthcoming Meeting

“...It seems to me to be interesting and important, nevertheless, to develop a capability for integrated network operation. If such a network as I envisage nebulously could be brought into operation, we would have at least four large computers, perhaps six or eight small computers, and a great assortment of disc files and magnetic tape units–not to mention the remote consoles and teletype stations–all churning away…”

*J.C.R. “Lick” Licklider writes memos to his network of researchers about his “**Intergalactic Computer Network**” concept: everyone connected able to access data, computing power, and tech resources, citing both military and research needs.*

1969 – ARPANET

- Developed by the Department of Defense
  - Connected universities and defense bases
  - Advanced Research Projects Agency (ARPA) of the U.S. DOD, a funder of ‘high-risk, high-gain’ research
- First ***packet switching*** network
  - Developed by computer scientists Leonard Kleinrock *(pictured)* and JCR Licklider *(previous slide)*
- At first consisted of 4 computers
  - One each located at UCLA, UCSB, Stanford Research Institute, and University of Utah

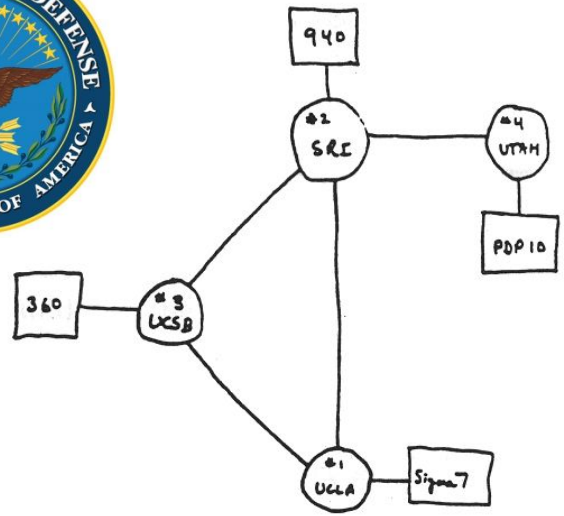


FIGURE 6.2 Drawing of 4 Node Network (Courtesy of Alex McKenzie)


The Initial ARPANET–1969
SRI
University
of Utah
UC Santa Barbara
UCLA

What did the first ARPANET message say?

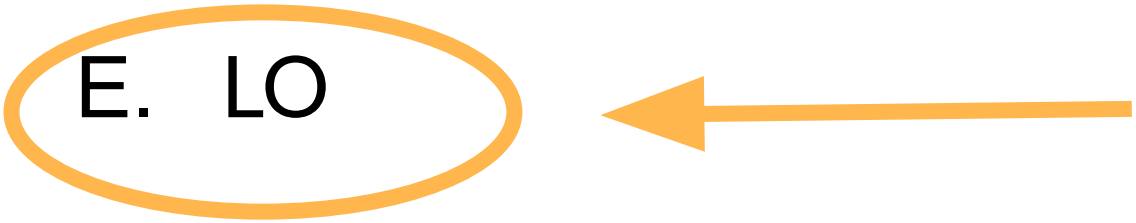
A. LOL

B. SOS

C. LOGIN

D. HELLO WORLD

E. LO

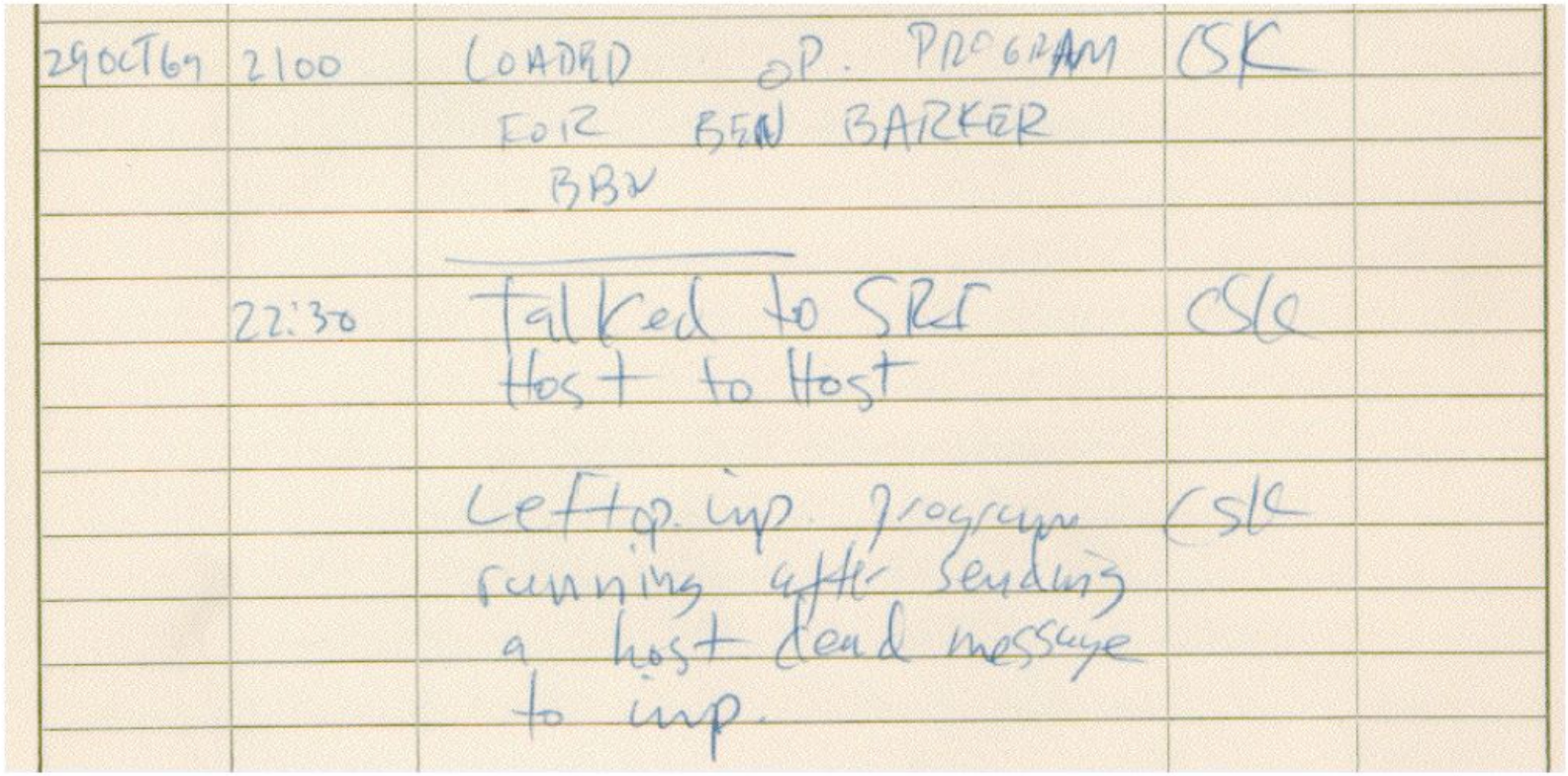


the first host-to-host connection, from UCLA to SRI, is made on October 29, 1969

the first “login” crashes the SRI host, but the next attempt works

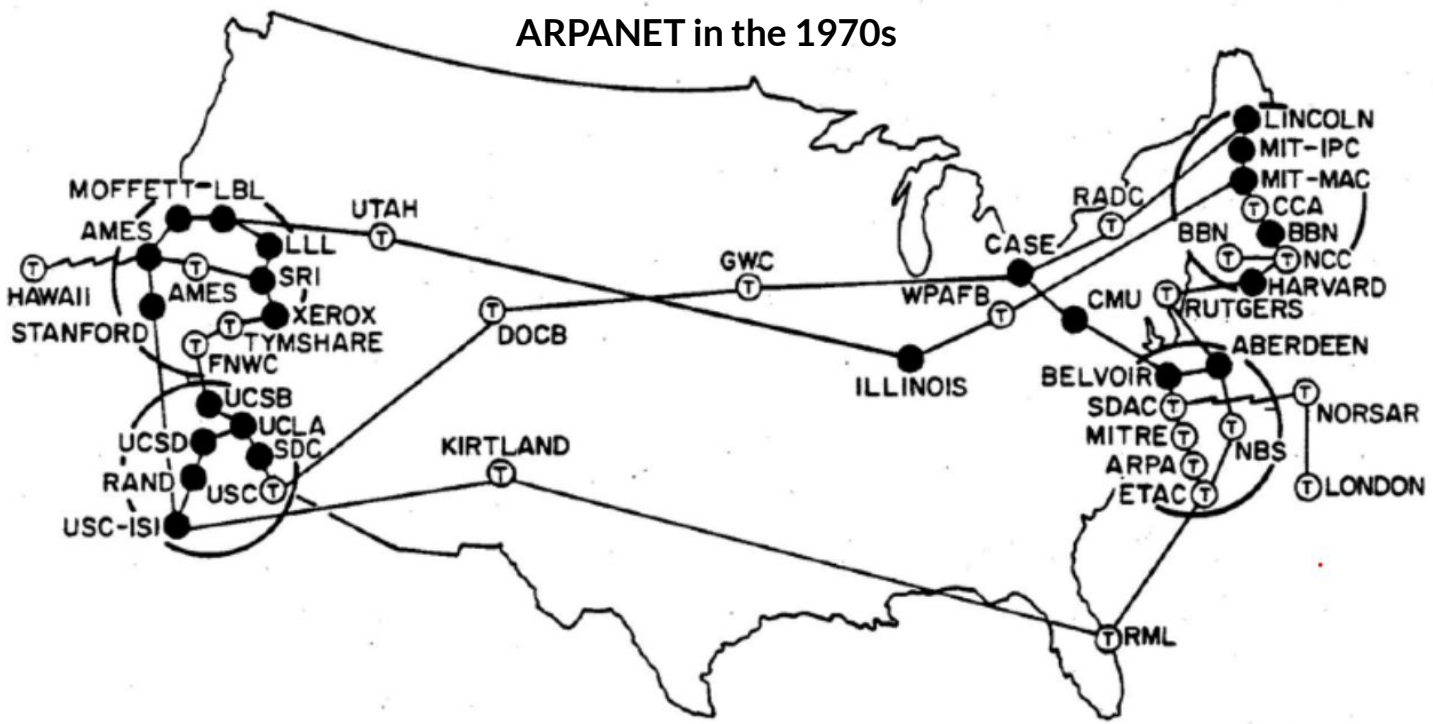
The Day the Infant Internet Uttered its First Words

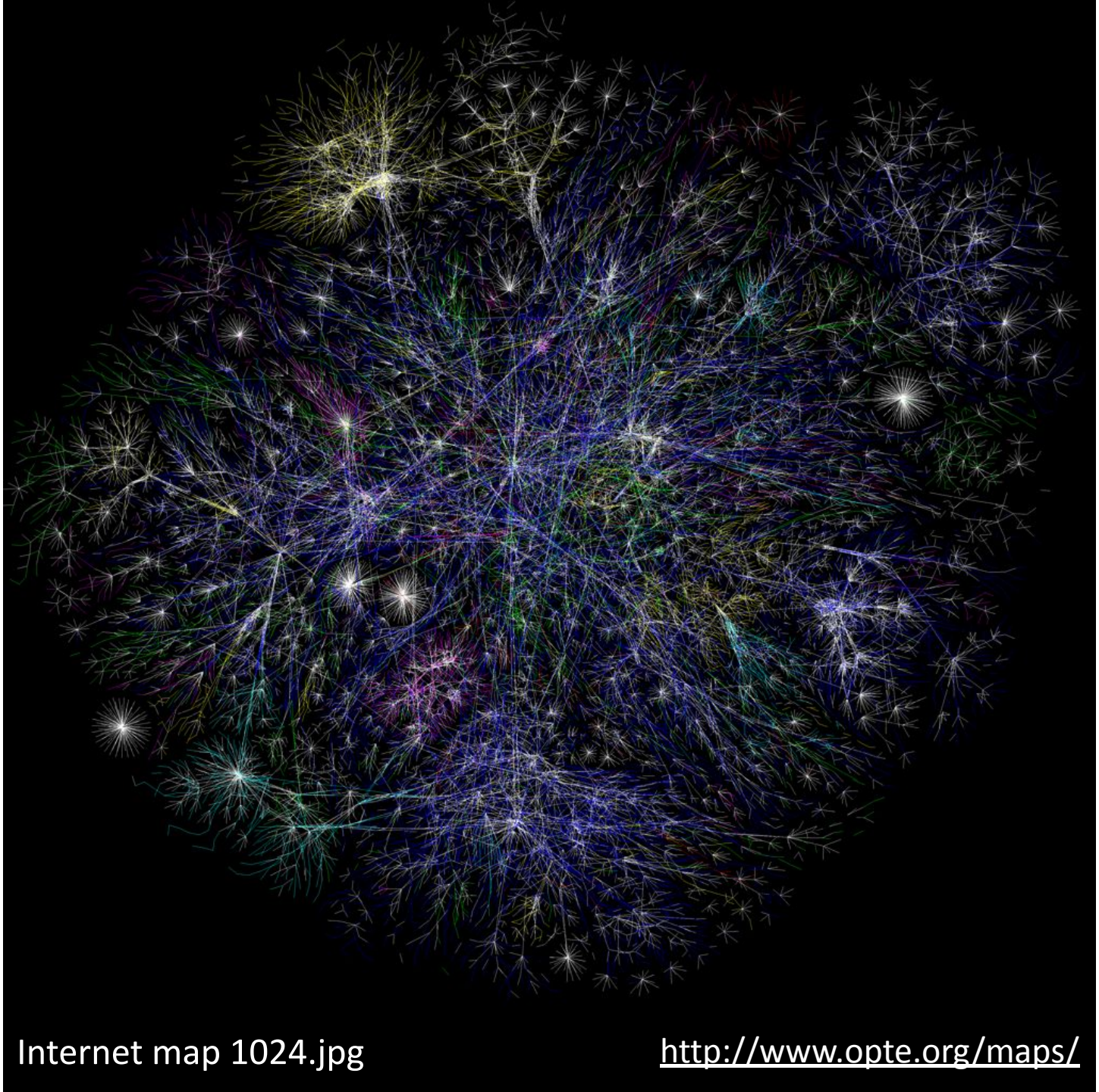
Below is a record of the first message ever sent over the ARPANET. It took place at 22:30 hours on October 29, 1969. This record is an excerpt from the "IMP Log" that was kept at UCLA. Professor Kleinrock was supervising his student/programmer Charley Kline (CSK) and they set up a message transmission to go from the UCLA SDS Sigma 7 Host computer to another programmer, Bill Duvall, at the SRI SDS 940 Host computer. The transmission itself was simply to "login" to SRI from UCLA. They succeeded in transmitting the "l" and the "o" and then the system crashed! Hence, the first message on the Internet was **"lo",** as in "lo and behold! They were able to do the full login about an hour later.



artifacts of Internet history from the UCLA profile page for Leonard Kleinrock

ARPANET in the 1970s

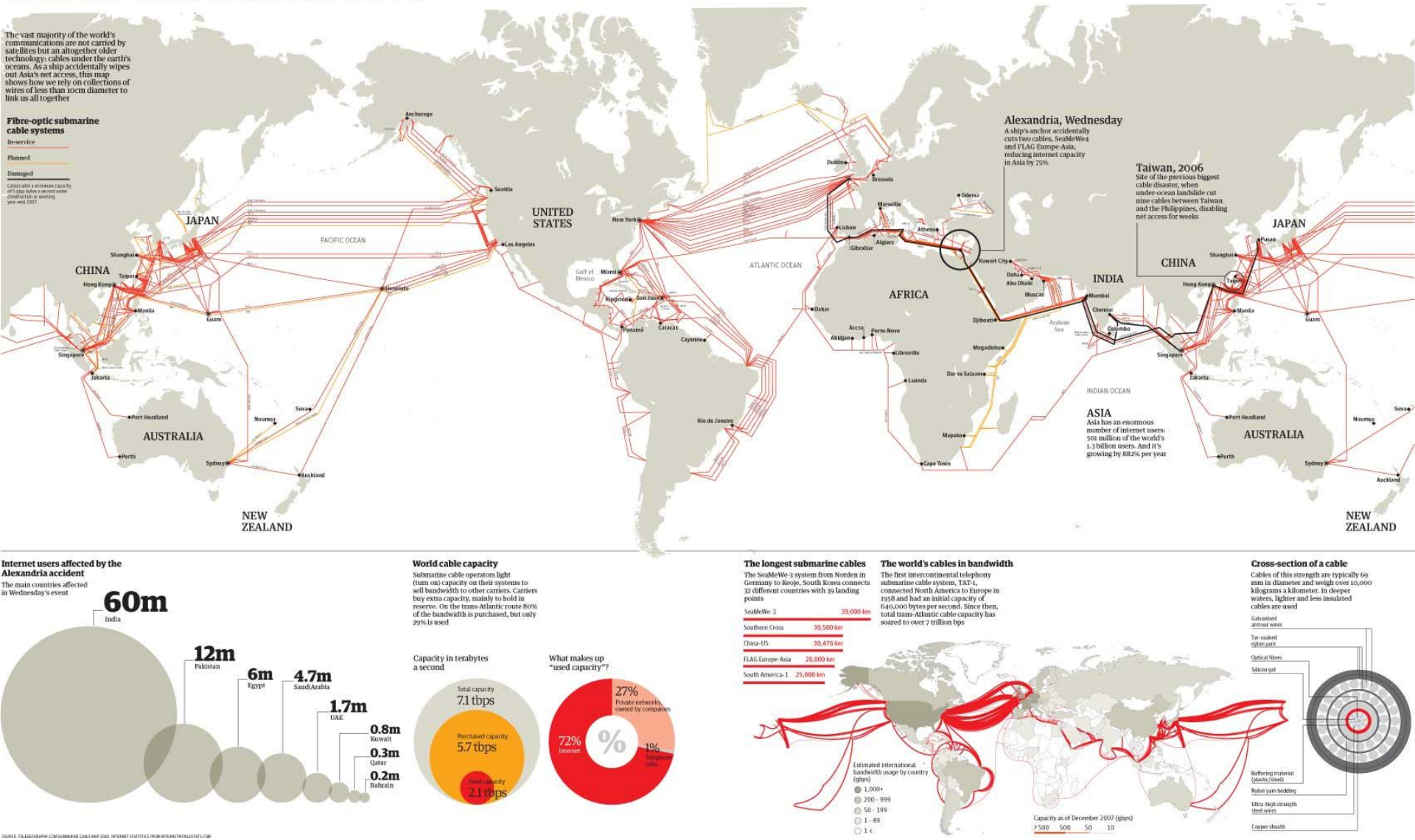




Internet map 1024.jpg

http://www.opte.org/maps/

The internet's undersea world



Internet users affected by the Alexandria accident

The main countries affected in Wednesday's event

World cable capacity

Submarine cable operators light (turn on) capacity on their systems to sell bandwidth to other carriers. Carriers buy extra capacity, mainly to hold in reserve. On the trans-Atlantic route 80% of the bandwidth is purchased, but only 29% is used

Capacity in terabytes a second

What makes up "used capacity"?

The longest submarine cables

The SeaMeWe-3 system from Norden in Germany to Keoje, South Korea connects 32 different countries with 39 landing points

SeaMeWe-3	39,000 km
Southern Cross	30,500 km
China-US	30,476 km
FLAG Europe-Asia	28,000 km
South America-1	25,000 km

The world's cables in bandwidth

The first intercontinental telephony submarine cable system, TAT-1, connected North America to Europe in 1958 and had an initial capacity of 640,000 bytes per second. Since then, total trans-Atlantic cable capacity has soared to over 7 trillion bps

Cross-section of a cable

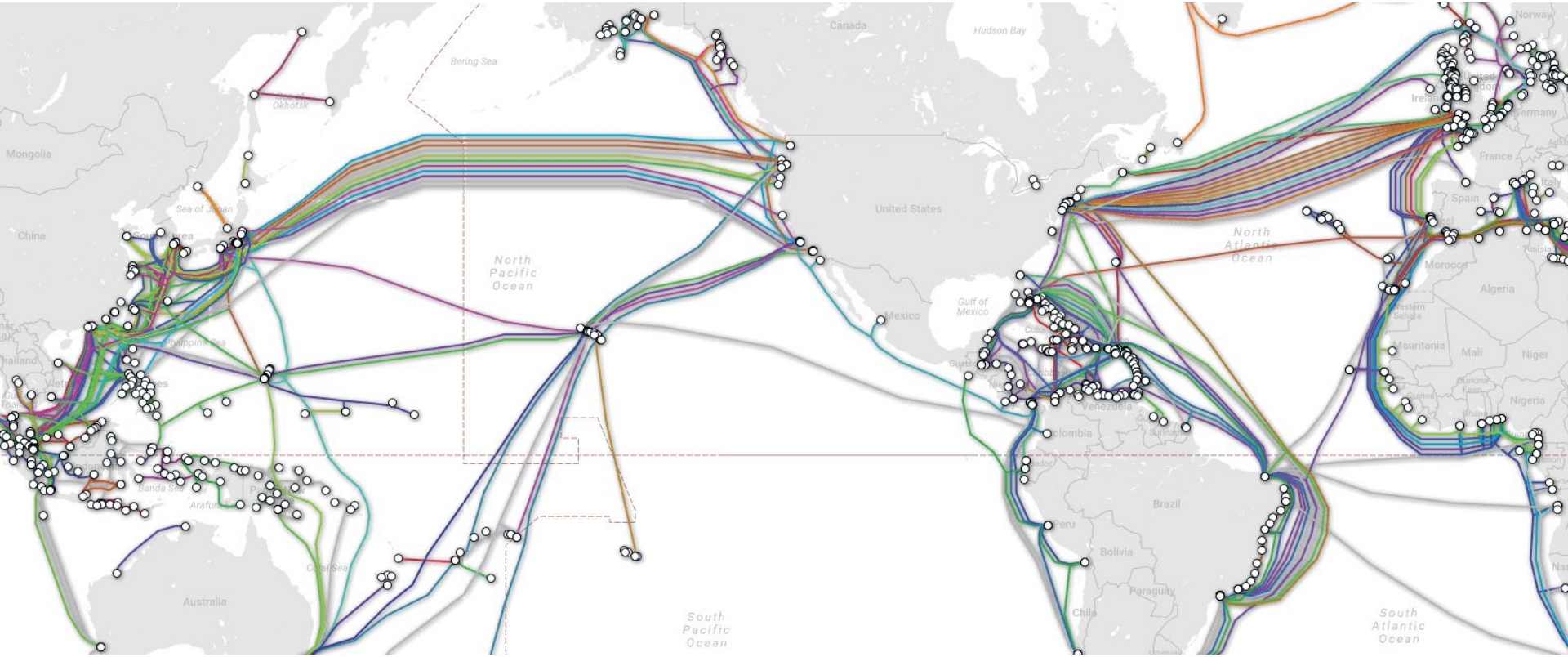
Cables of this strength are typically 69 mm in diameter and weigh over 10,000 kilograms a kilometer. In deeper waters, lighter and less insulated cables are used

SOURCE: TELEGEOGRAPHY.COM SUBMARINE CABLE MAP 2008 INTERNET STATISTICS FROM INTERNETWORLDSTATS.COM

http://image.guardian.co.uk/sys-images/Technology/Pix/pictures/2008/02/01/SeaCableHi.jpg?

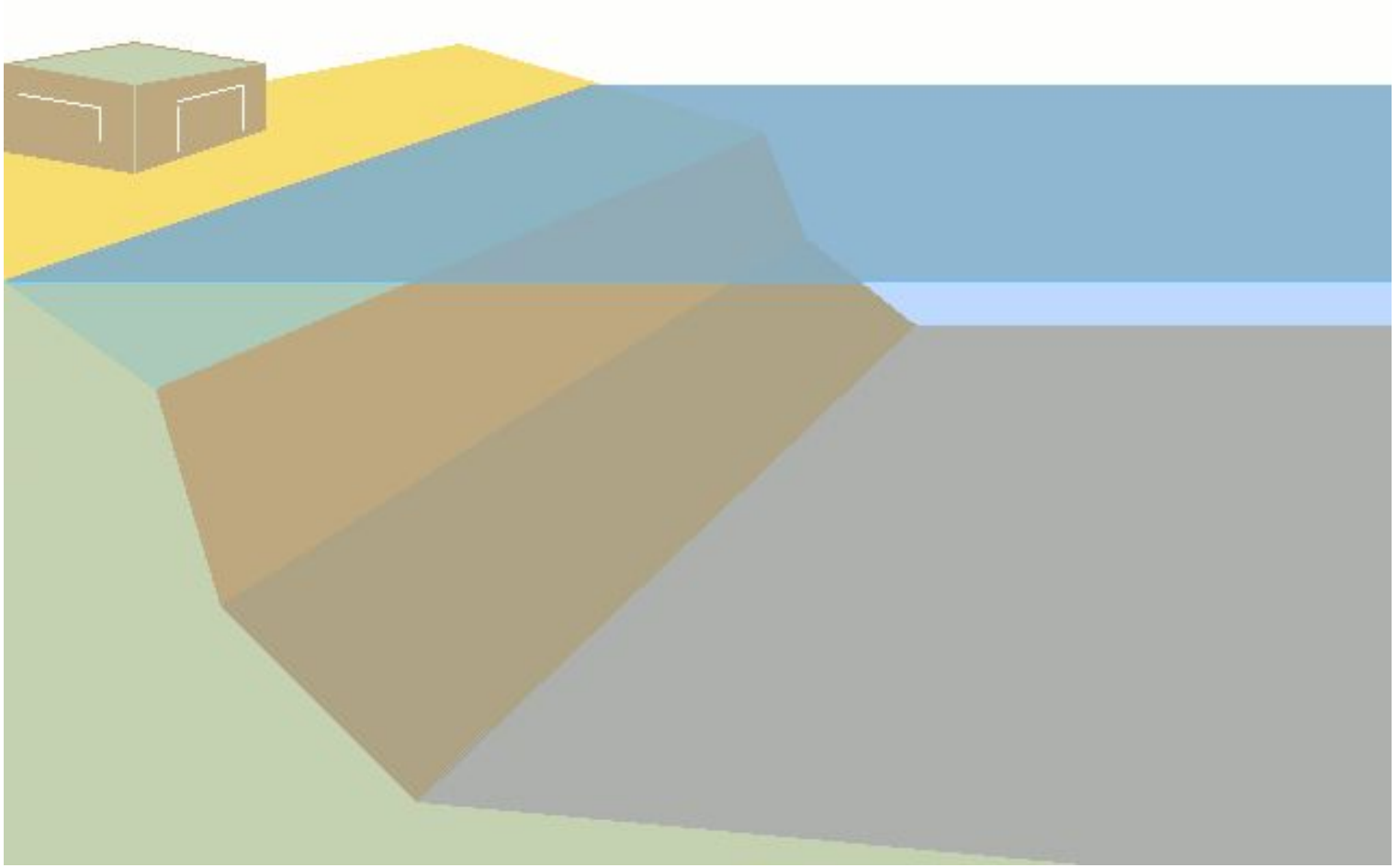
Submarine cables

Large networked systems are connected to each other via undersea cables. Explore the online version of the site -- find one or more countries you've been to and trace their cables.



source: submarinecablemap.com

Undersea Cable (continued)



http://upload.wikimedia.org/wikipedia/commons/8/8f/Undersea_cable_laying.gif

Taryn Simon

Transatlantic Sub-Marine Cables Reaching Land, VSNL International, Avon, New Jersey

These VSNL sub-marine telecommunications cables extend 8,037.4 miles across the Atlantic Ocean

From the article reading “A Brief History of the Internet”

(Leiner et al., 2003), (emphasis added)

“This history revolves around four distinct aspects. There is the **technological evolution** that began with early research on **packet switching** and the ARPANET (and related technologies), and where current research continues to expand the horizons of the **infrastructure** along several dimensions, such as scale, performance, and higher level functionality. There is the **operations and management** aspect of a global and complex **operational infrastructure**. There is the **social aspect**, which resulted in a **broad community of Internauts working together** to create and evolve the technology. And there is the **commercialization** aspect, resulting in an extremely effective transition of research results into a broadly deployed and available information infrastructure.

[The Internet’s] history is complex and involves many aspects - **technological, organizational, and community**. And its influence reaches not only to the technical fields of computer communications but throughout society as we move toward increasing use of online tools to accomplish electronic commerce, information acquisition, and community operations.”

<break!>
(20 mins)

A closer look at how data moves around the Internet

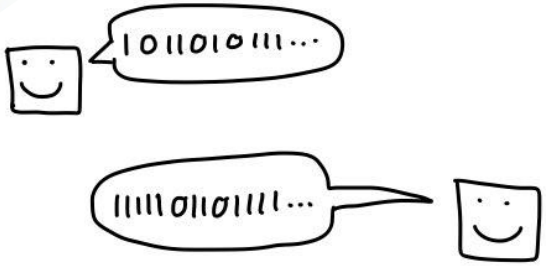
TCP/IP

Julia Evans
@b0rk

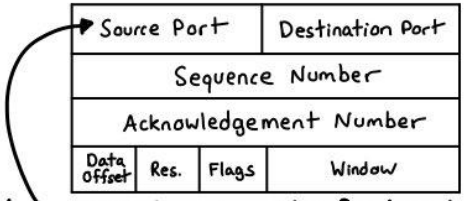
network protocols

computers communicate a lot

internet communication is made of bytes

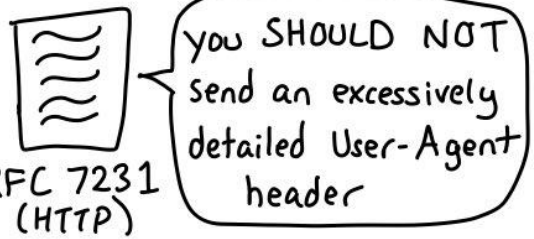


a protocol specifies what the bytes mean



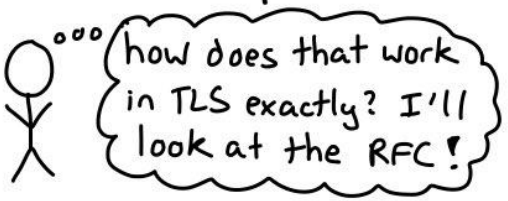
in a TCP header, the first 2 bytes are the source port

protocols also say which requests are allowed



Some network protocols

RFCs define internet protocols



The IETF publishes new RFCs

More Standards!

- The Internet is built on the **TCP/IP** family of standards **(Transmission Control Protocol / Internet Protocol)**

- TCP/IP was standardized via the Internet Engineering Task Force and are a set of rules that govern how data is transferred over the internet. (See also related IP and infrastructure global management roles of the IANA and IAB.)

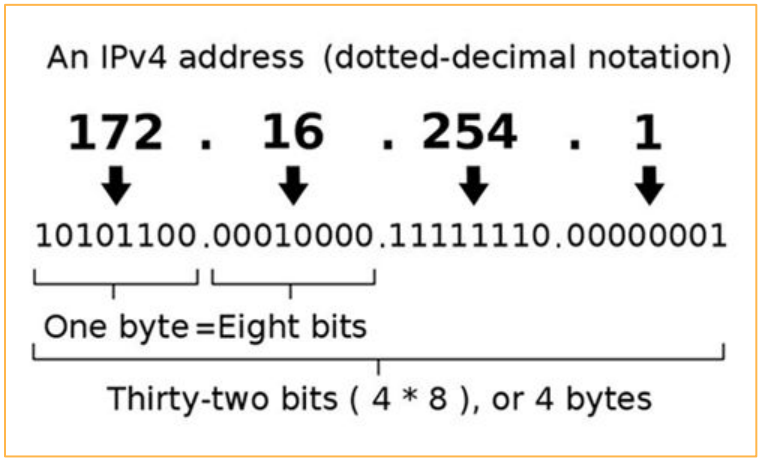
- These are *free and open*, *vendor-neutral standards*
  - They have been incredibly successful!

4 layers

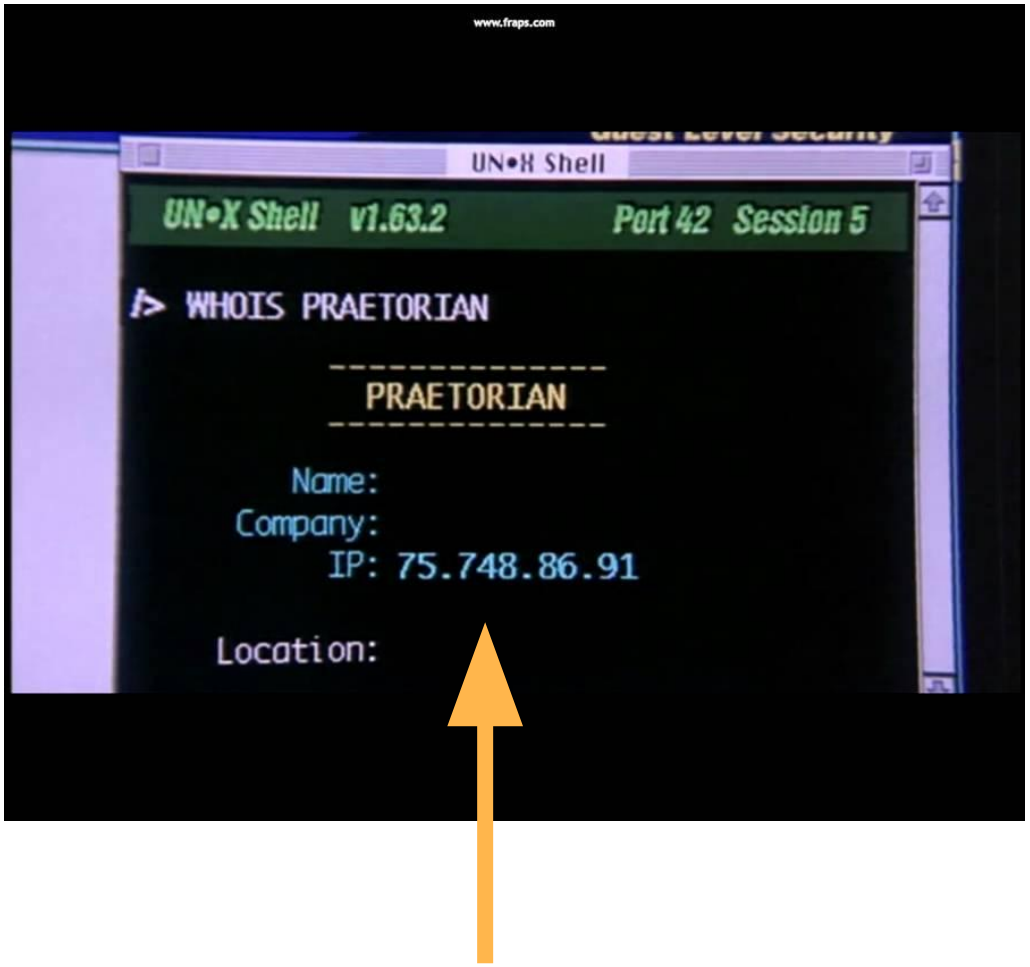
- Application layer (sends information to transport layer)
- Transport (TCP) layer (breaks packets down, ensures reliable byte stream)
- Internet (IP) layer (pair each packet to proper IP destination)
- Link layer (sends packets between IP and hosts)

Internet Protocol (IP)

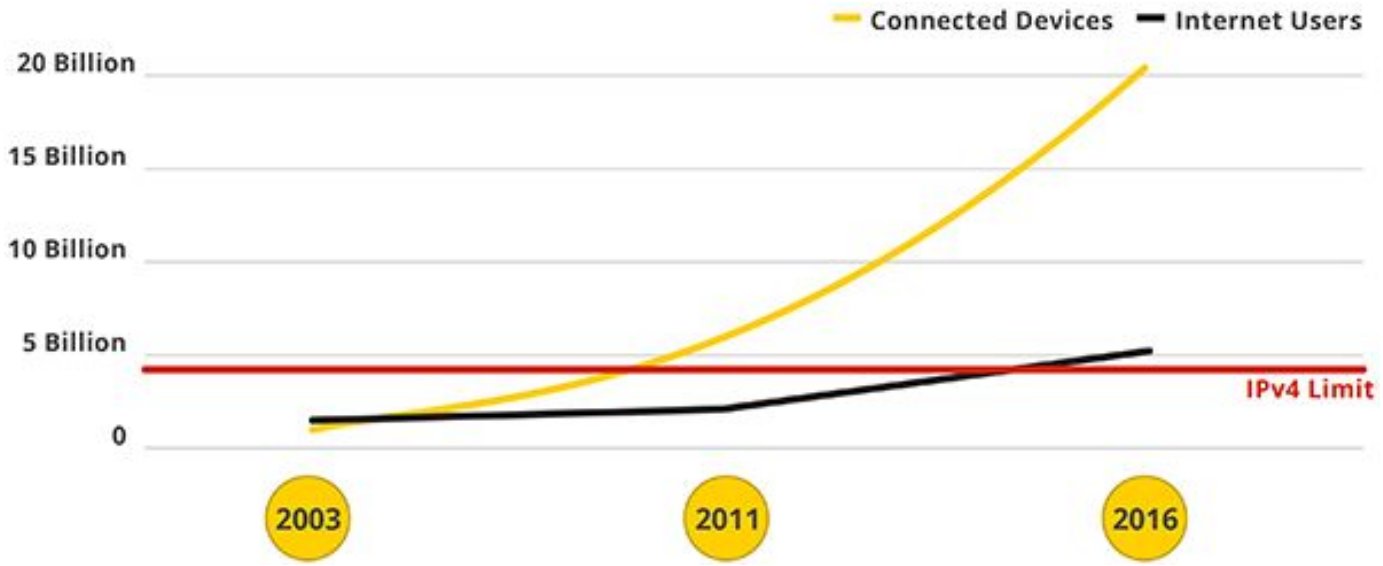
- Enables computers to **route communications traffic** from *one network to another*
- Each computer connected to the Internet is given a unique address called its **IP address**
  - An **IPv4** address is a series of four numbers (one byte each) separated by dots
  - The range of each of these numbers (0–255) allows for billions of IP addresses
- New IP addresses (IPv4) are in *short supply*!
  - Moving towards **IPv6**
  - IPv6 uses a 128-bit addresses



A Misrepresentation in the film “The Net”…

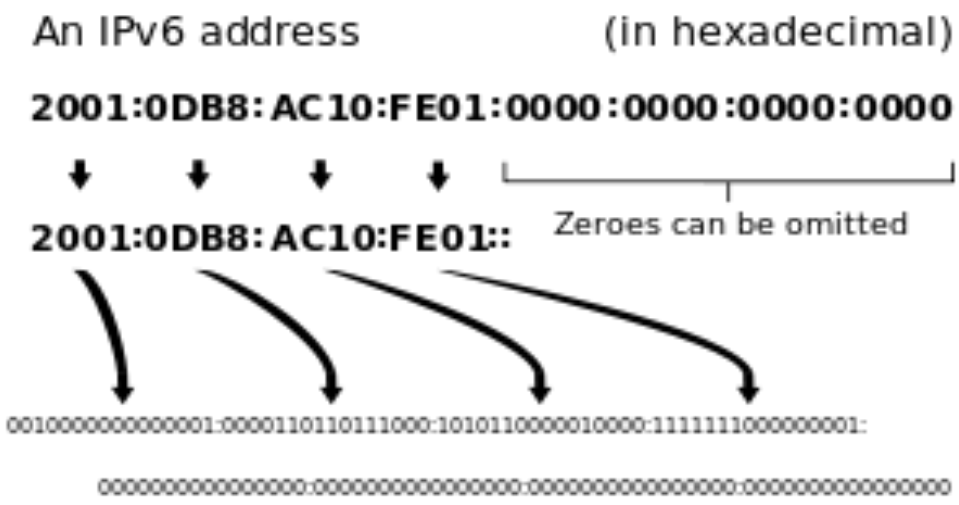


IPv4 is Running Out of Room!

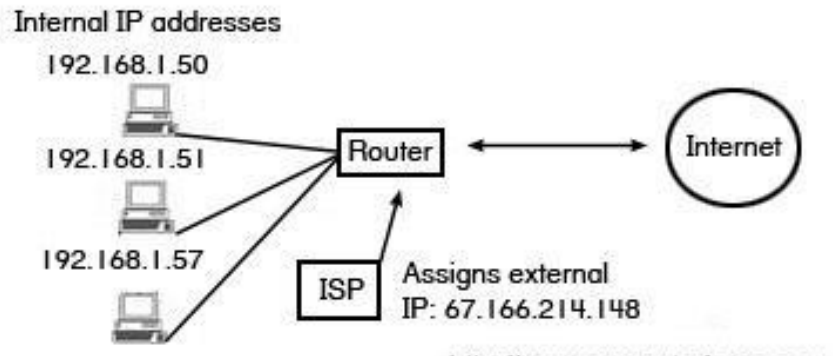


http://www.google.com/intl/en/ipv6/

An **IPv6 address** is represented in eight groups of four hexadecimal digits (eight 16-bit blocks) separated by colons (:)



Activity: check your IP: http://www.whatismyip.com/



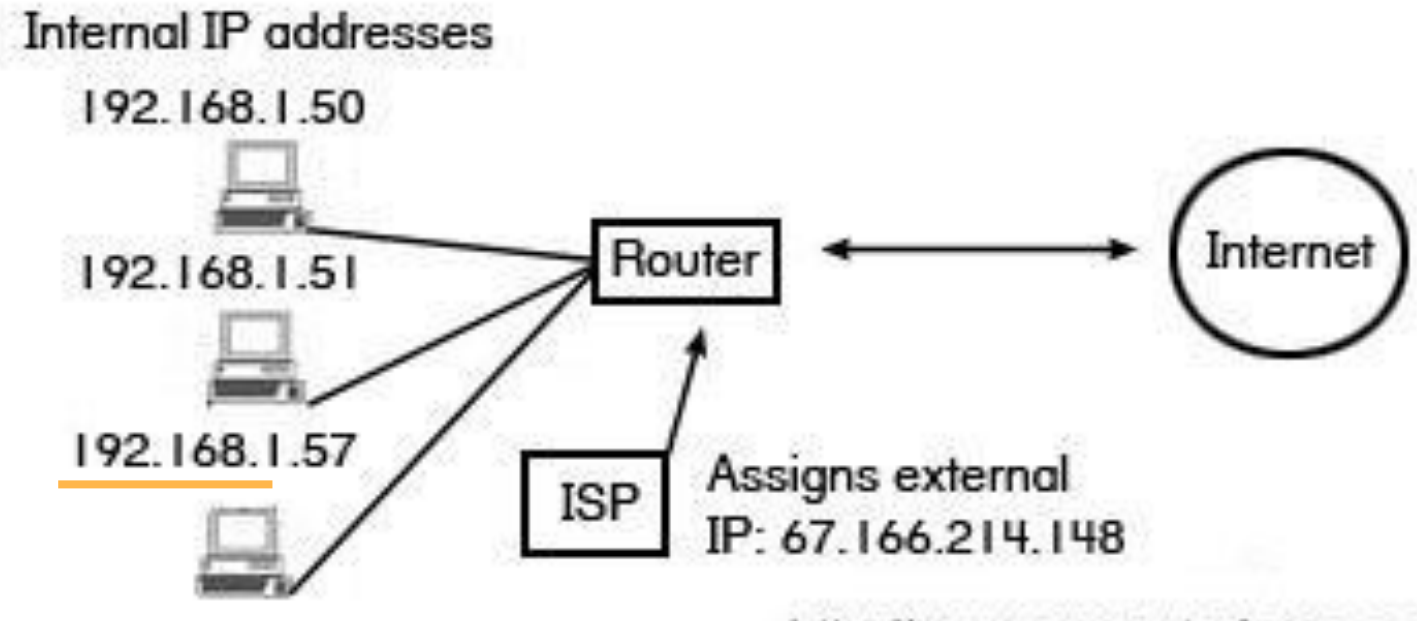
- What does the above web page say? Compare with the people next to you. Try checking a phone or laptop too!
- Do these numbers have anything in common? Why do you think this is?

Dynamic Host Configuration Protocol (DHCP)

Dynamic Host Configuration Protocol (DHCP) is a network protocol that enables a server to automatically assign an IP address to a computer

The router supplies a private IP address for your computer to use

On the Internet everyone sees your external IP address



IP Addresses & Domain Names

It is hard to remember the **numeric IP address** of all the computers we communicate with!

- E.g. *google.com* is much easier to remember than *195.13.189.20*
- domain names are how humans engage with the IP system effectively

The Internet uses *human-readable symbolic names* for computers that are based on a hierarchy of *domains*

- A **domain** is a related group of networked computers

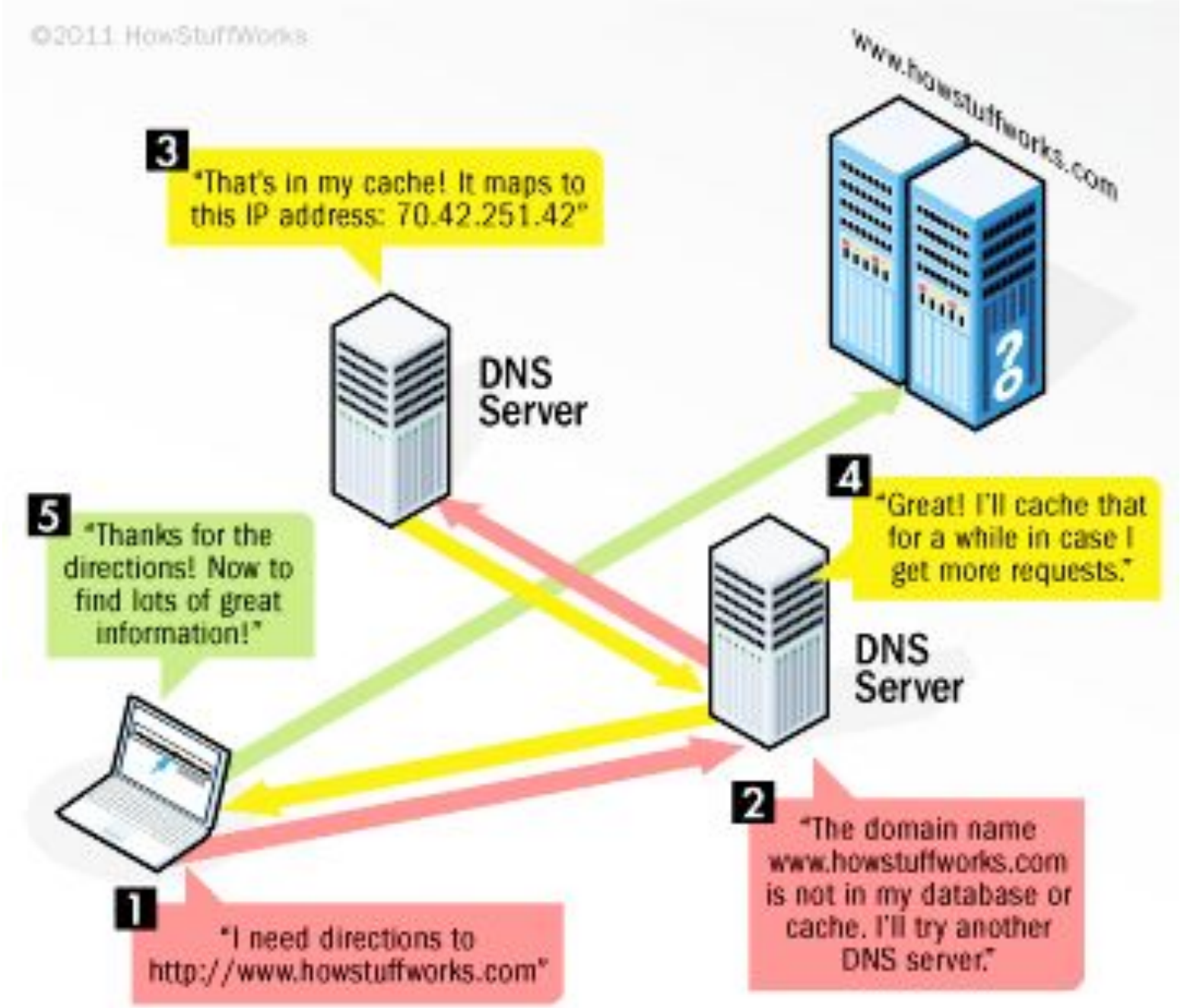
Top-Level Domains

- **Top-level domain** names (**TLDs**):
  - .edu for educational groups
  - .com for commercial enterprises
  - .org for organizations
  - The full list can be found at www.icann.org

- As of October 1, 2016, the US government no longer supervises the DNS, with the **Internet Corporation for Assigned Names and Numbers (ICANN)** taking control without US oversight
  - DNS is moving out of the stewardship of the **National Telecommunications and Information Administration (NTIA)**
  - A global panel of engineers, from the United States, Europe and Asia, will oversee the system and guard security
- How do we get ("register") a new domain name for a project site or organization…? Let's take a look at options at: https://www.hostgator.com/domains

DNS Servers

- The **Domain Name System (DNS)** translates the hierarchical, human-readable names into the four-number IP address
  - Every Internet host knows the IP address of its nearest **DNS name server**
  - Your computer asks the DNS server, which looks up the corresponding IP address
- If the address is new (and not stored on the DNS server), the server asks an **authoritative name server**
- *"DNS: Domain Name System is like an address book for websites…. The browser needs to find out which server the website lives on, so it can send HTTP messages to the right place" -- MDN*



http://computer.howstuffworks.com/dns.htm

Transmitting Data on the Internet

- The **Transmission Control Protocol (TCP)** and **User Datagram Protocol (UDP)** are protocols that **run on top of the *IP* network protocol**
- **TCP** is a *connection-oriented protocol* and is designed to allow two computers to establish a *reliable* connection and *exchange* data
  - *TCP* does the work of ensuring data is delivered and processed in the correct order
  - *TCP* relies on IP to move the data to the correct location

Activity: Mock Register a Domain Name

Imagine your project launch preparations…

Pretend you are launching a new website for a project or organization. Make up a potential domain name. (short, catchy)

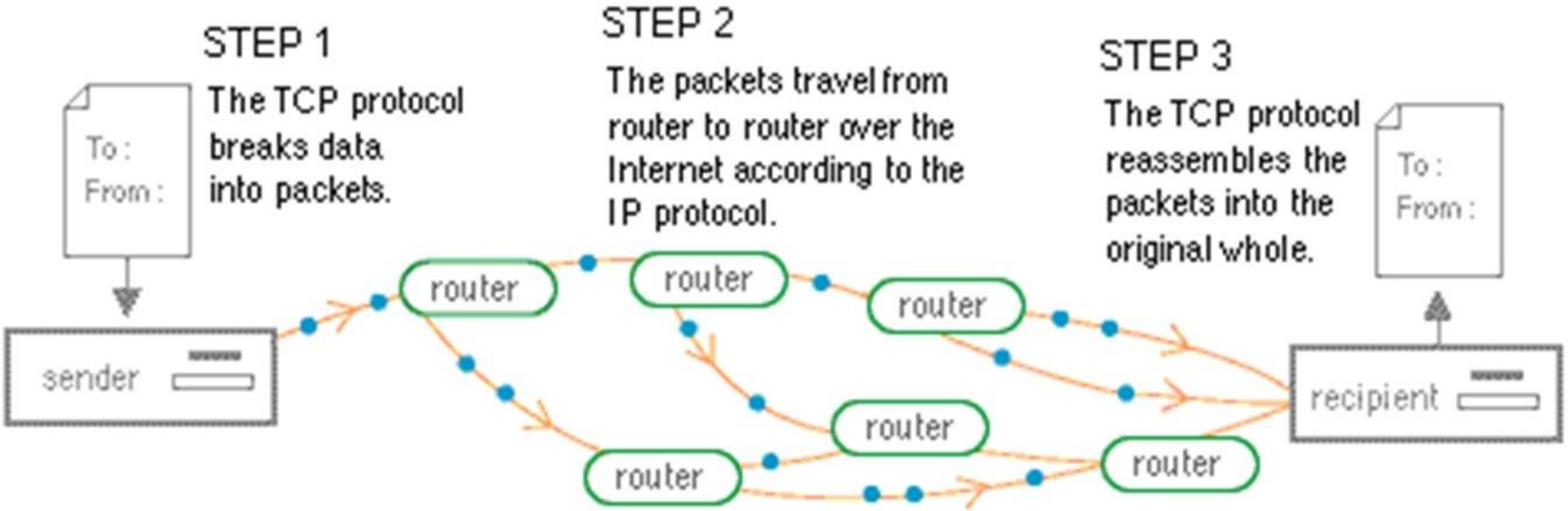
Open your browser and navigate to a reputable web hosting service like:

- https://www.a2hosting.com/
- https://www.bluehost.com/
- https://www.hostgator.com/

Consider:

- Is your domain name available?
- What Top Level Domain options are available? Which would you get?
- What do the costs look like and for how long will the registration last? What are budgeting implications?

Network Packet Example



Moving Packets

- Because each packet can take a different route, congestion and service interruptions ***do not* delay transmissions**
  - Each TCP/IP packet is *independent*, so the TCP/IP protocol works under adverse conditions
- The Internet uses telephone carriers for long-distance connections, fiber optics, and separate dedicated lines for connections
  - The computers do not know or care ***how*** the packet is sent, as long as it can be ***sent and received***
  - Transmissions may rely on ***multiple technologies*** as the packets move across the Internet

https://youtu.be/ewrBalT_eBM

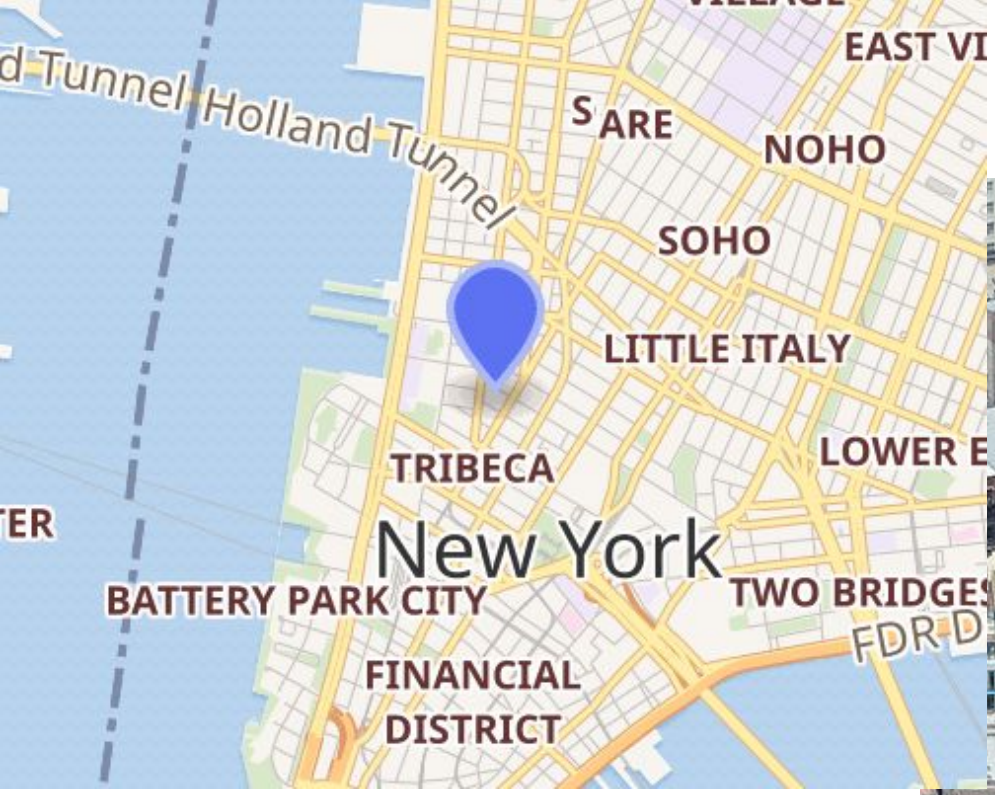
The Internet Lives in a Huge Hotel in Manhattan

Security is drum-tight, with security guards, security cameras, biometric security checkpoints, and even man traps.

THERE IS NO shortage of cool stuff to see on the Internet, but the Internet itself—the networks and servers and cables tying it all together—is pretty mundane. Peter Garritano discovered as much when he went behind the scenes at some of New York's big Internet hubs to see how it all works.

Garritano always knew the Internet is as much a place as it is a thing, one where vast networks in frigid rooms move enormous amounts of data around the world. But he didn't have a real understanding of how all that info moves from one place to another, and he wanted to find out. He's spent the last few months doing just that, shooting at five "carrier hotels" where many networks converge to form a single, larger network. That's why they called it the Internet. It allows an unlimited number of different networks to come together as one.

https://www.wired.com/2015/11/peter-garritano-where-the-internet-lives/

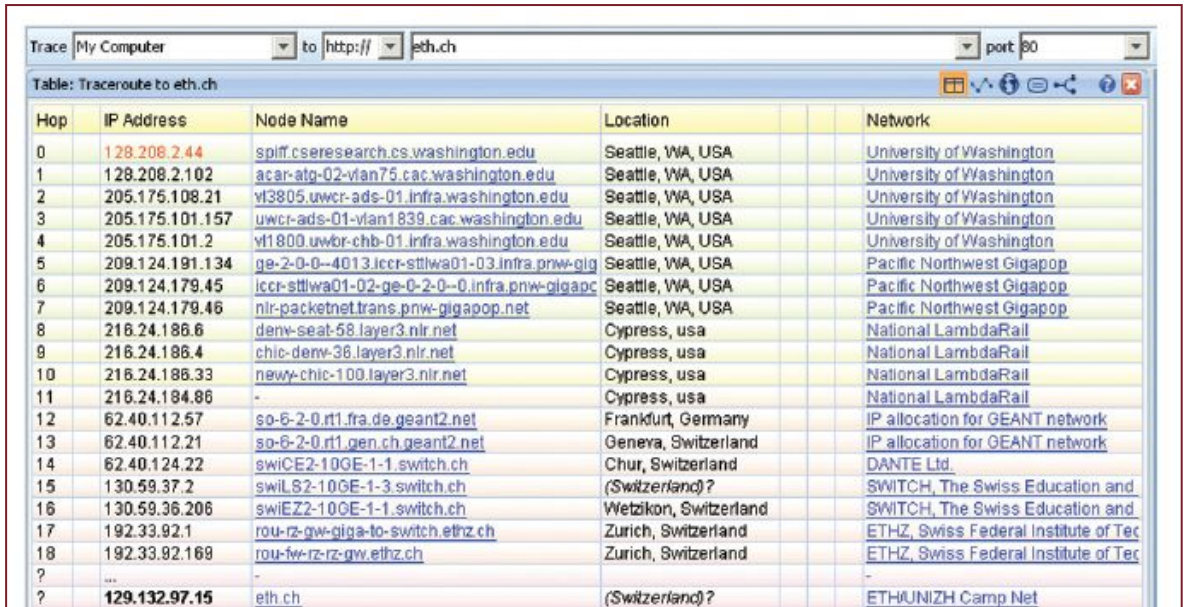


Colocation data center New York
at 60 Hudson Street

https://www.digitalrealty.com/data-centers/new-york/60-hudson-st-new-york-ny

Far and Near: WAN and LAN

- The Internet is a collection of **wide area networks (WAN)**
  - These are networks that are *not* geographically close
- The Internet is a collection of **point-to-point** channels
  - Meaning packets must visit a sequence of computers (or **hops**) before they reach their destination



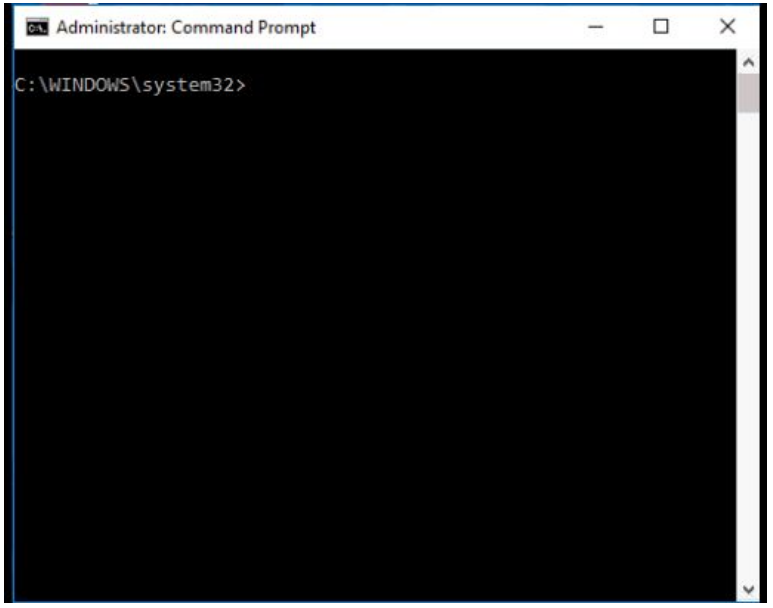
Trace: My Computer to http:// eth.ch port 80

Table: Traceroute to eth.ch

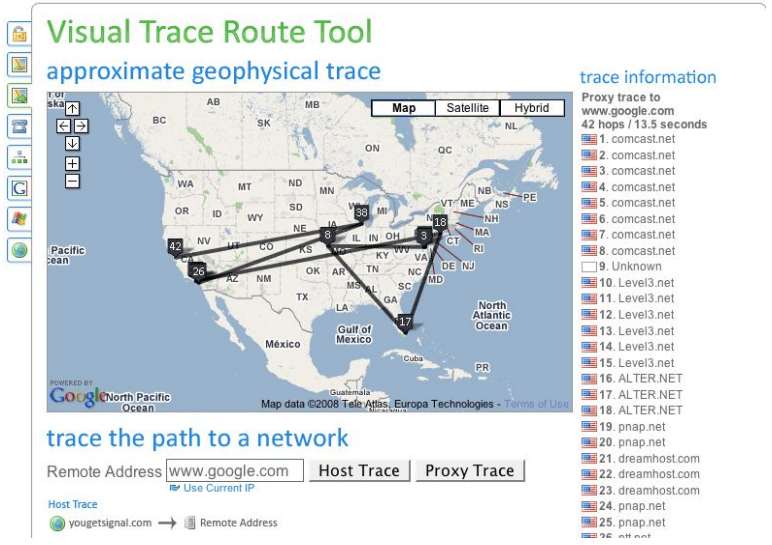
Hop	IP Address	Node Name	Location	Network
0	128.208.2.44	spiff.cseresearch.cs.washington.edu	Seattle, WA, USA	University of Washington
1	128.208.2.102	acar-atg-02-vlan75.cac.washington.edu	Seattle, WA, USA	University of Washington
2	205.175.108.21	vl3805.uwcr-ads-01.infra.washington.edu	Seattle, WA, USA	University of Washington
3	205.175.101.157	uwcr-ads-01-vlan1839.cac.washington.edu	Seattle, WA, USA	University of Washington
4	205.175.101.2	vl1800.uwbr-chb-01.infra.washington.edu	Seattle, WA, USA	University of Washington
5	209.124.191.134	ge-2-0-0--4013.iccr-sttlwa01-03.infra.pnw-gig	Seattle, WA, USA	Pacific Northwest Gigapop
6	209.124.179.45	iccr-sttlwa01-02-ge-0-2-0--0.infra.pnw-gigapc	Seattle, WA, USA	Pacific Northwest Gigapop
7	209.124.179.46	nlr-packetnet.trans.pnw-gigapop.net	Seattle, WA, USA	Pacific Northwest Gigapop
8	216.24.186.6	denv-seat-58.layer3.nlr.net	Cypress, usa	National LambdaRail
9	216.24.186.4	chic-denv-36.layer3.nlr.net	Cypress, usa	National LambdaRail
10	216.24.186.33	newy-chic-100.layer3.nlr.net	Cypress, usa	National LambdaRail
11	216.24.184.86	-	Cypress, usa	National LambdaRail
12	62.40.112.57	so-6-2-0.rt1.fra.de.geant2.net	Frankfurt, Germany	IP allocation for GEANT network
13	62.40.112.21	so-6-2-0.rt1.gen.ch.geant2.net	Geneva, Switzerland	IP allocation for GEANT network
14	62.40.124.22	swiCE2-10GE-1-1.switch.ch	Chur, Switzerland	DANTE Ltd.
15	130.59.37.2	swiLS2-10GE-1-3.switch.ch	*(Switzerland)?*	SWITCH, The Swiss Education and
16	130.59.36.206	swiEZ2-10GE-1-1.switch.ch	Wetzikon, Switzerland	SWITCH, The Swiss Education and
17	192.33.92.1	rou-rz-gw-giga-to-switch.ethz.ch	Zurich, Switzerland	ETHZ, Swiss Federal Institute of Tec
18	192.33.92.169	rou-fw-rz-rz-gw.ethz.ch	Zurich, Switzerland	ETHZ, Swiss Federal Institute of Tec
?	...	-		-
?	**129.132.97.15**	eth.ch	*(Switzerland)?*	ETH/UNIZH Camp Net

Exploring The Internet

Ping – Allows us to see if a computer on the network is available (*can the current computer connect/"talk to" the other computer*)



Traceroute – Displays the route (path) and transit delays of packets across an internet protocol network



Activity

Run a traceroute!

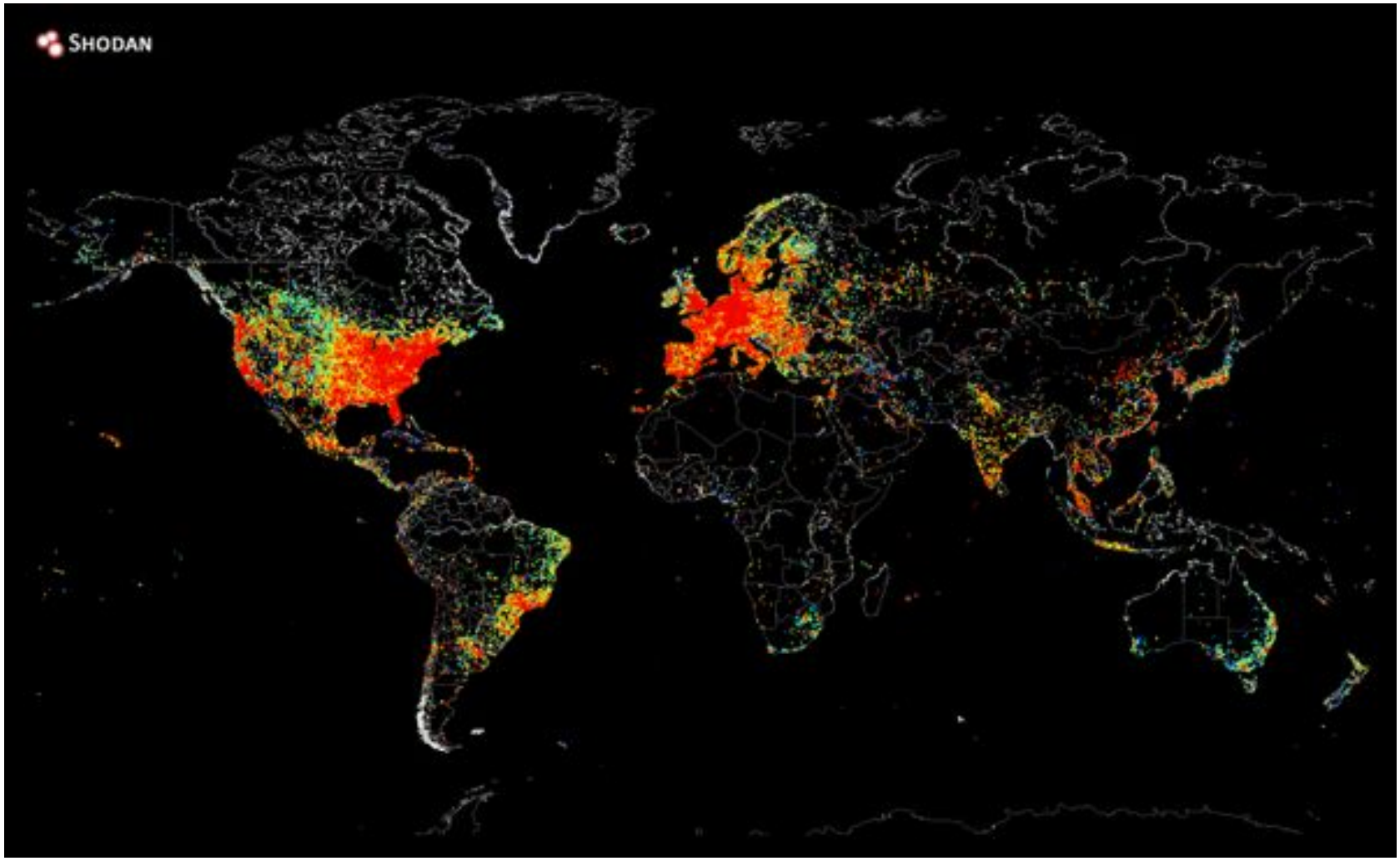
*Note: Traceroute is a command that runs tools used for network diagnostics. These tools trace the paths data packets take from their source to their destinations, allowing administrators to better resolve connectivity issues. On a **Windows** machine, this command is called **tracert**; on **Linux** and **Mac**, it's called **traceroute***

- On a Mac open **Terminal** (go to your Applications folder, then inside the Utilities folder, and you'll see Terminal)
- Note what it says in the Terminal window. Do you see your user name….?
- Type in :

  traceroute

- then a **space**
- then **a URL of your choice** (such as www.nypl.org, www.loc.gov, or any you like)
- for example:
  traceroute www.pratt.edu
- Observe the route for several hops (**see p. 182 in the textbook PDF**)
- (Use "Control-C" when you want it to stop)

Result of Pinging All Devices Connected to the Internet



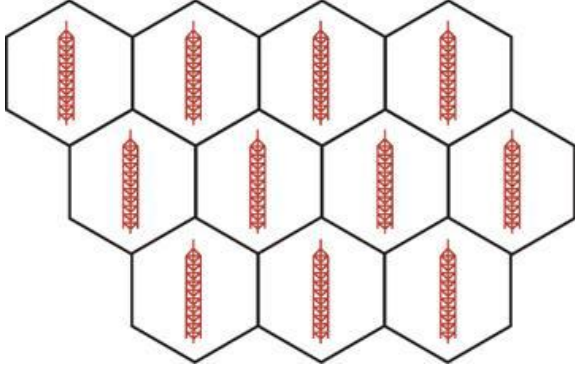
Connecting to the Internet

1. Connection via an Internet service provider (ISP)
   - The company places a modem at your house (frequently in a package purchased with other services)
     - Modems convert the bits a computer outputs into a form that is compatible with the carrier

2. Connection provided by a campus or enterprise network
   - The organization connects to the Internet by a gateway

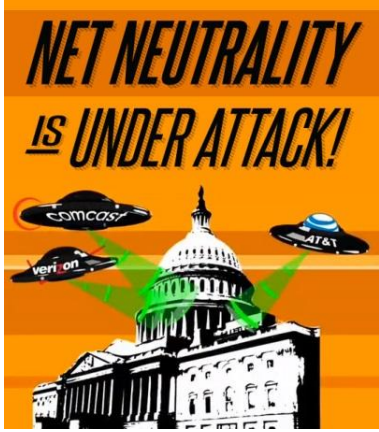
- Most of us use both kinds of connections

Wireless vs. Cellular Data



- Two major categories of wireless Internet access for smartphones, tablets, laptops, etc.
- **Cellular wide area networks (WANs)**
  - Base stations serve cells, land area served by cell tower(s)
  - Data cellular networks allow connection to the public Internet
    - Mobile phone company acts as the ISP
- **Wi-Fi local area networks (LANs)**
  - i.e. simply connecting to a Wi-Fi router in your home, a coffeeshop, at school, etc.

Peering and Transit



- Most network connections are *indirect*
  - The Internet is a “network of networks”; it is nearly impossible to interconnect *directly* with all networks on the globe!
  - This creates the need for ***interconnect agreements***
- **Peering**: Two or more autonomous networks interconnect directly to *exchange* traffic
  - Often done *without charging* for the interconnection/traffic
- **Transit**: Allowing network traffic to cross a computer network
  - Often used to connect a smaller Internet service provider (ISP) to the larger Internet
- Increased traffic due to the rise in video streaming and competition for cable-TV-providers has lead to **peering disputes** and the debate around *net neutrality*

THE WORLD WIDE WEB

How does the Web work?

WWW's early logo, created by Robert Cailliau in 1990 - made of three Ws in Optima Bold

The Internet VS. World Wide Web

- What's the difference?
  - The **web** is a *service* that operates *over the Internet*
    - E.g. as does email
  - The **internet** *predates* the web by several decades
    - All the way back to the early days of computing!

- The **Web**:
  - Consists of server and client software, the Hypertext Transfer Protocol (http), standards, and mark-up languages that combine to deliver information and services over the Internet

The Internet

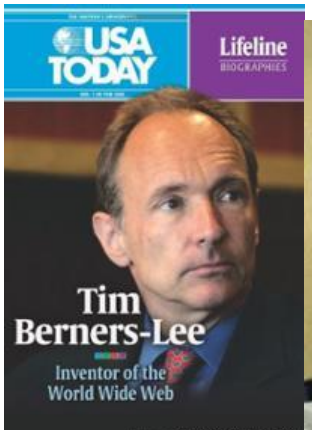
a NETwork for information communication, bound up by material infrastructures and packet configuration and movement

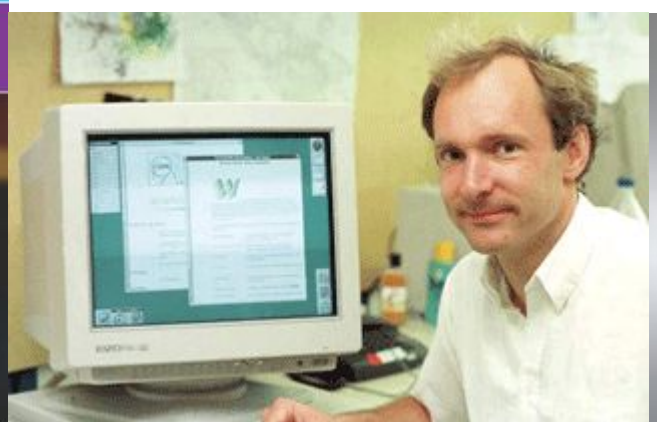
The World Wide Web

a WEB of communications, facilitated through software, standards, and mark-up

A Brief History of the Web

- **Tim Berners-Lee** was a software engineer at CERN, a large particle physics laboratory in Switzerland, late 1980s - 1990s
  - Identified the need for scientists to *exchange data*
- In 1990, he proposed three technologies that would become the foundation of the web
  - Proposed *HTML*, *URI*, *and HTTP* (more on these shortly!)
- **Browser was released to the general public in 1991 and CERN made the Web protocol/code available royalty free in 1993,** enabling widespread use.

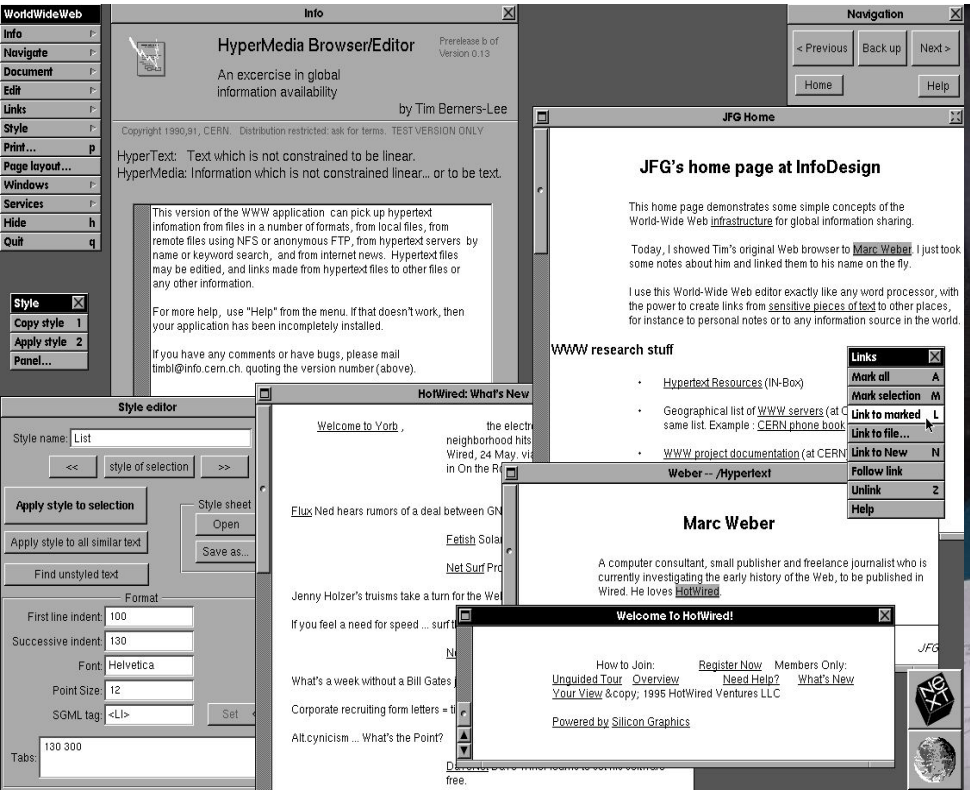




Berners-Lee Introduced:

The web browser

The web server



and… the Hypertext Transfer Protocol (HTTP)

Hypertext Transfer Protocol (HTTP)

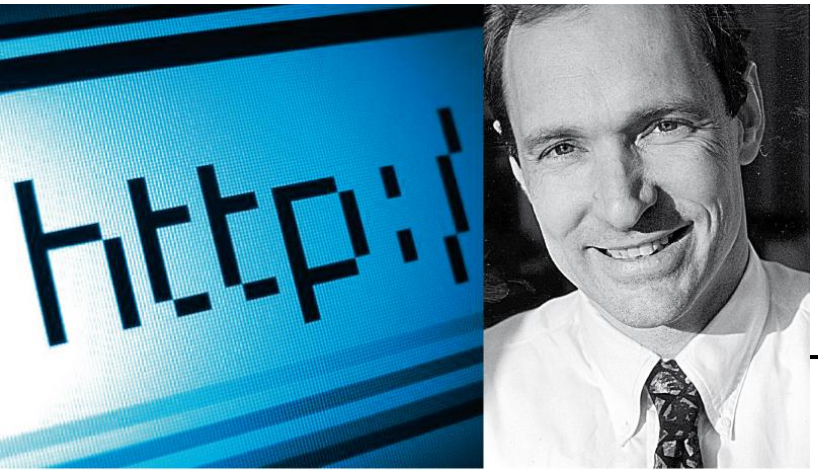
HTTP is a *communication standard* governing the requests and responses that take place between the *web browser* running on the end user's computer and the *web server*

The first website, recreated:

http://info.cern.ch/

A simulation of browsing the web 1990 style:

https://worldwideweb.cern.ch/



World Wide Web

The WorldWideWeb (W3) is a wide-area hypermedia information retrieval initiative aiming to give universal access to a large universe of documents.

Everything there is online about W3 is linked directly or indirectly to this document, including an executive summary of the project, Mailing lists , Policy , November's W3 news , Frequently Asked Questions .

What's out there?
: Pointers to the world's online information, subjects , W3 servers, etc.

Help
: on the browser you are using

Software Products
: A list of W3 project components and their current state. (e.g. Line Mode ,X11 Viola , NeXTStep , Servers , Tools , Mail robot , Library)

Technical
: Details of protocols, formats, program internals etc

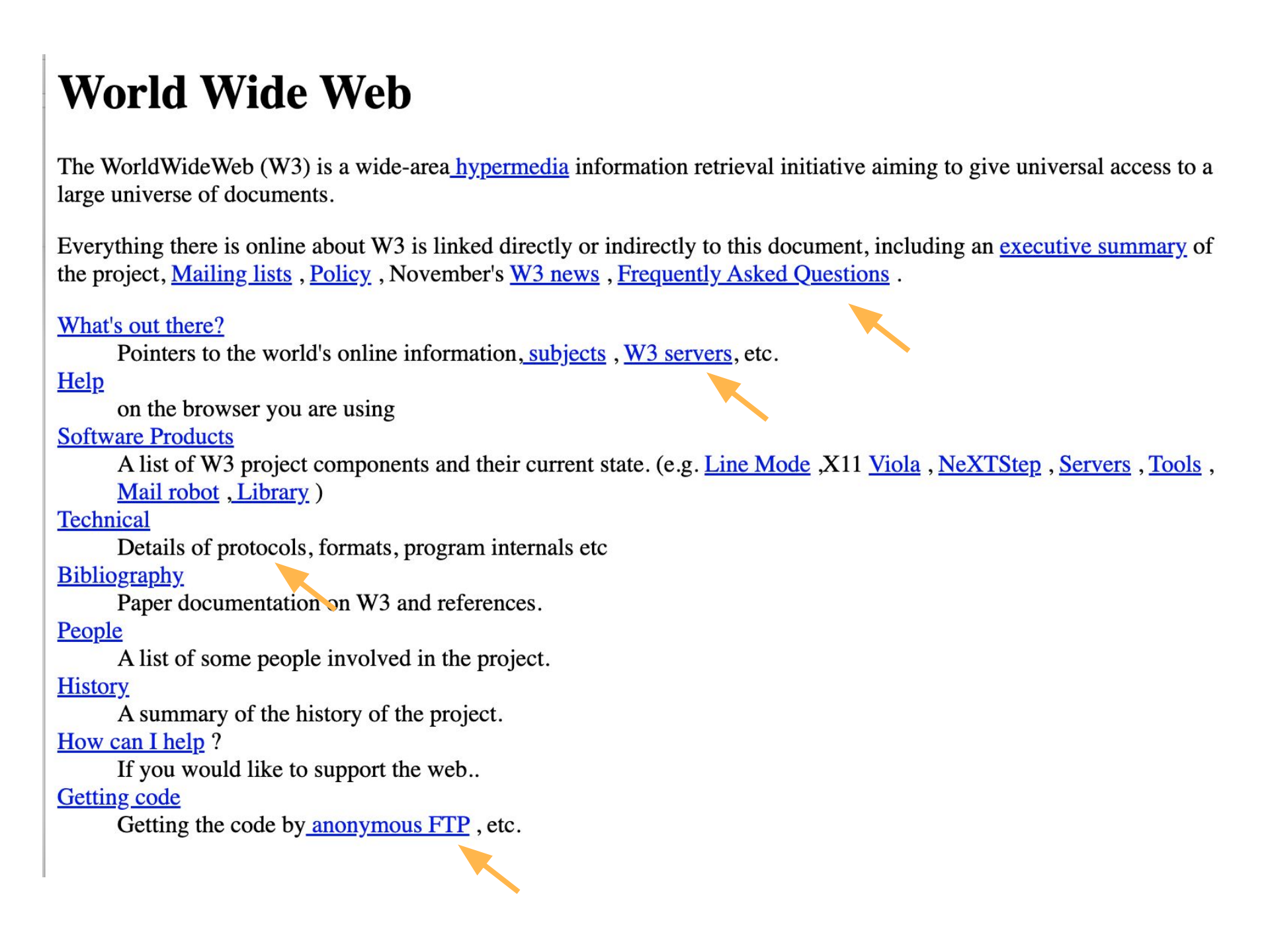
Bibliography
: Paper documentation on W3 and references.

People
: A list of some people involved in the project.

History
: A summary of the history of the project.

How can I help ?
: If you would like to support the web..

Getting code
: Getting the code by anonymous FTP , etc.



Back Then...

```
The MAXIMUS Computer-Based Conversation System, v1.00 (non-commercial)
Copyright 1989, 1990 by Scott J. Dudley of 1:250/814.  All rights reserved.
Supplementary and OS/2 code by Peter Fitzsimmons of 1:250/628.

Compiled on Mar 03 1990 at 21:39:48 under Turbo C

Computer: Generic MSdos-class 0x0

OS: DOS 5.00
FOSSIL: BNU FOSSIL Communications Driver v1.70
Remaining memory in heap: 20384 bytes


MAIN:
Message Areas       File Areas          Change Setup        Goodbye (log off)
Statistics          Yell for SysOp      UserList            Version of BBS
Bulletin Menu       !Remote DOS Shell   @User Editor        ?Help
Select: _
```

What is HyperText

Hypertext is text which is not constrained to be linear.

Hypertext is text which contains links to other texts. The term was coined by Ted Nelson around 1965 (see History).

HyperMedia is a term used for hypertext which is not constrained to be text: it can include graphics, video and sound , for example. Apparently Ted Nelson was the first to use this term too.

Hypertext and HyperMedia are concepts, not products.

from the reproduction of the original WWW page: http://info.cern.ch

Hypertext Transfer Protocol (HTTP)

HTTP is a *communication standard* governing the requests and responses that take place between the *web browser* running on the end user's computer and the *web server*

HTTPS (today)

HTTPS is a protocol developed on top of HTTP that takes extra steps to ensure security, privacy, and integrity of the communication channel between you and the website. *(today's standard/best practice)*

https://

By The Early 1990s…

- 1992
  - Viola browser
  - 50 web servers
- 1993
  - 500 web servers
  - Mosaic web browser
    - First browser to robustly display graphics!

By The Mid-1990s…

- 1994
  - 2,500 web servers
  - WWW Consortium (W3C) began
  - Netscape 1.0
  - Yahoo! Web directory started by Jerry Yang and David Filo


What's New


Check Email

YAHOO!


Personalize


Help

Yahoo! Auctions
Pokemon, Beanies

Yahoo! Mail
free email for life

Search advanced search

Yahoo! Shopping - Apparel, Computers, Videos, DVDs, CDs, Toys, Electronics and more

Shopping - **Auctions** - Yellow Pages - People Search - Maps - Travel - Classifieds - Personals - Games - Chat - **Clubs**
Mail - Calendar - Messenger - **Companion** - My Yahoo! - News - Sports - Weather - TV - Stock Quotes - more...

Arts & Humanities
Literature, Photography...

Business & Economy
Companies, Finance, Jobs...

Computers & Internet
Internet, WWW, Software, Games...

Education
College and University, K-12...

Entertainment
Cool Links, Movies, Humor, Music...

Government
Elections, Military, Law, Taxes...

Health

News & Media
Full Coverage, Newspapers, TV...

Recreation & Sports
Sports, Travel, Autos, Outdoors...

Reference
Libraries, Dictionaries, Quotations...

Regional
Countries, Regions, US States...

Science
Animals, Astronomy, Engineering...

Social Science
Archaeology, Economics, Languages...

Society & Culture

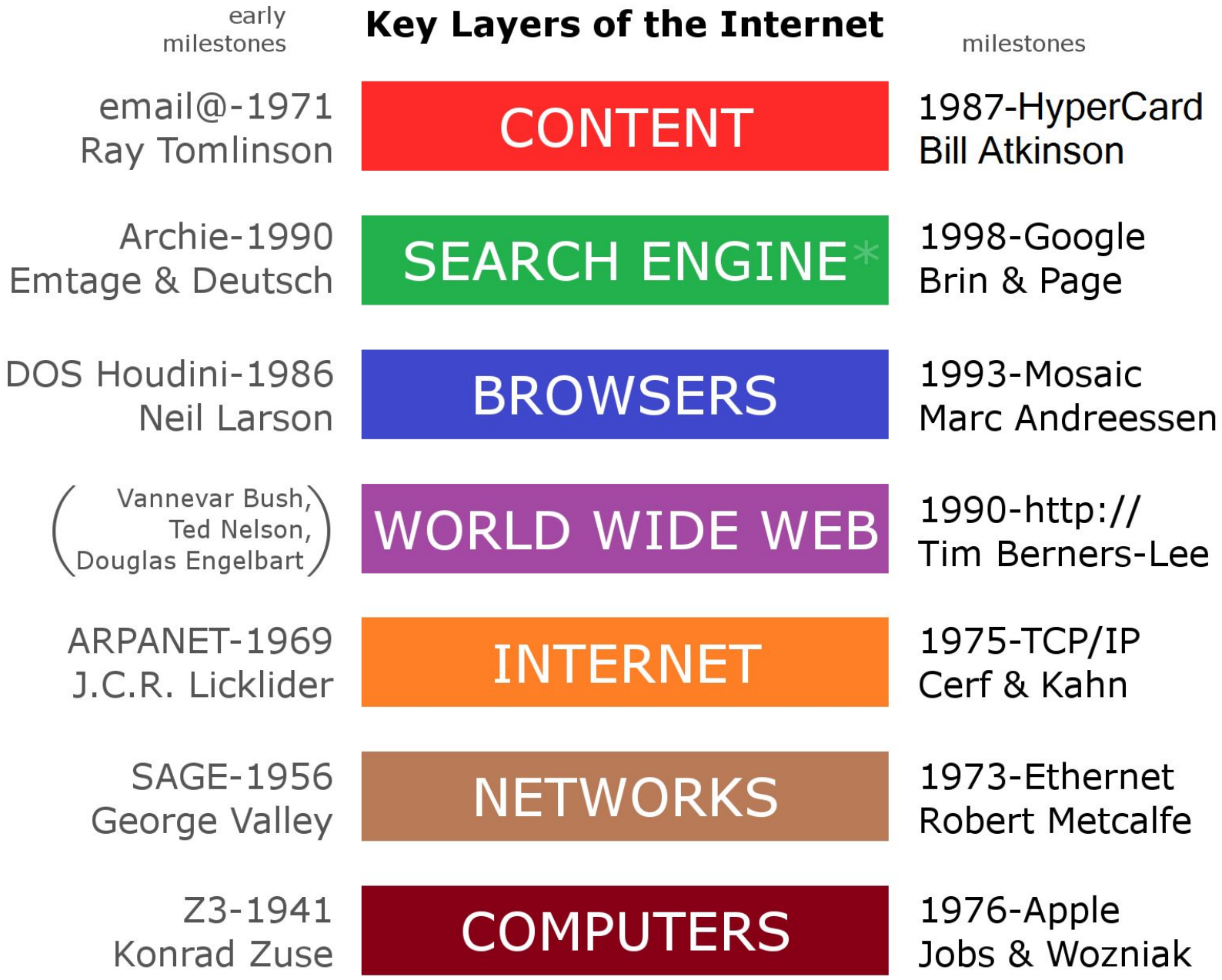
In the News
- Report: EgyptAir crash to become criminal probe
- ATM fee bans blocked by judge
- Comdex

more...

Marketplace
- Y! Travel - plan your holiday travel
- Looking for a car? job? house?

more...

Inside Yahoo!
- Yahoo! GeoCities - build your free home page
- Y! Games - hearts, backgammon, chess



(image source: Wikipedia)

Requesting a Web Page

- Your computer is the client computer and the computer with the Web page is the server (Web server)
  - The client gets services from the server
- When you click a Web link, your computer gets the page for you...beginning the client/server interaction
  - When the page is returned, the operation is completed and the client/server relationship ends

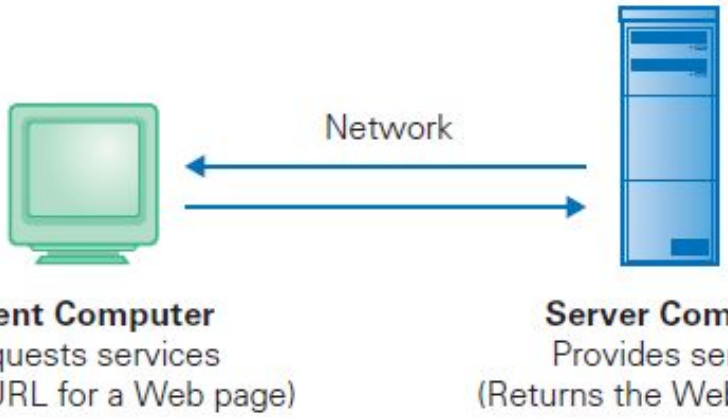


Figure 3.2 The basic client/server interaction, as illustrated by the browser (client) requesting Web pages provided by the Web server.

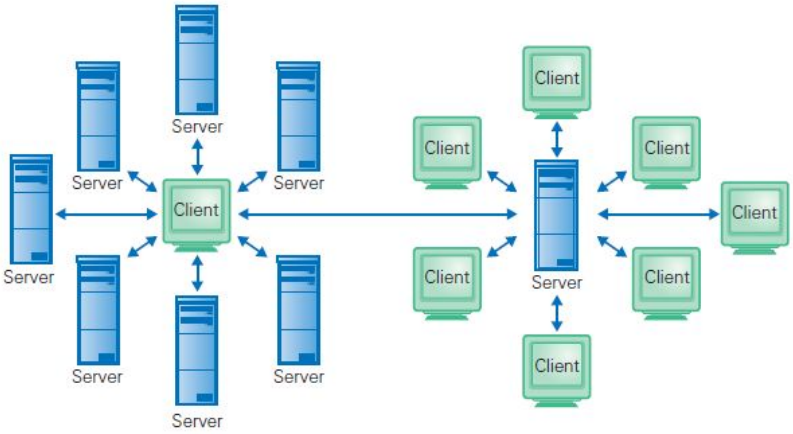


Figure 3.3 Client/server relationships as they might evolve over time.

The Uniform Resource Locator (URL)

https://libguides.pratt.edu/internet-data-privacy

- The **URL** has three main parts:
  - **Protocol**
    tells the computers how to handle the file
  - **Server computer's name**
    or the name given by the domain hierarchy
  - **Page's pathname**
    tells the server which file (page) is requested and where to find it

MDN's web site accessing allegory:
"like you are walking from home to a shop to retrieve the thing you want"

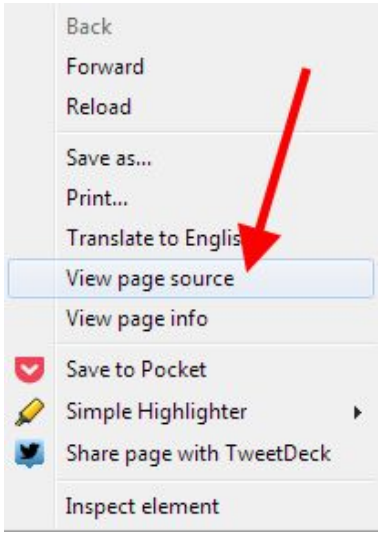
- "When **you type a web address into your browser** (for our analogy that's like walking to the shop):
- The **browser goes to the DNS server,** and finds the real address of the server that the website lives on (you find the address of the shop).
- The **browser sends an HTTP request message to the server,** asking it to send a copy of the website to the client (you go to the shop and order your goods). This message, and all other data sent between the client and the server, is sent across your internet connection using TCP/IP.
- If the server approves the client's request, **the server sends the client a "200 OK" message**, which means "Of course you can look at that website! Here it is", and **then starts sending the website's files to the browser as a series of small chunks called data packets** (the shop gives you your goods, and you bring them back to your house).
- The **browser assembles the small chunks into a complete web page** and displays it to you (the goods arrive at your door — new shiny stuff, awesome!)."

from developer.mozilla.org "getting started with the web"

HTML - Describing a Web Page

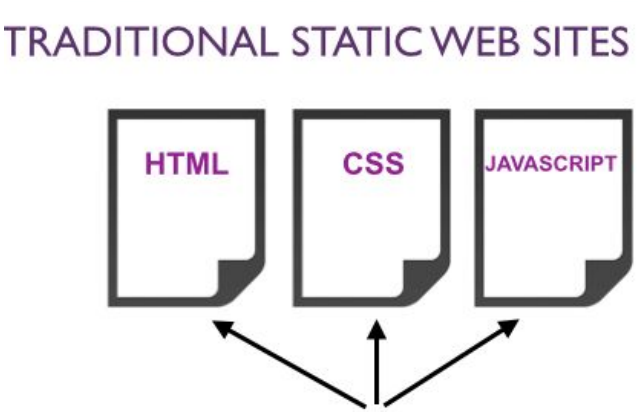
- Servers do not store Web pages in the form seen on our screens
- The **pages are stored as a *description of how they should appear on the* screen**
  - Web pages are written in HyperText Markup Language (HTML), also designed by Berners-Lee!
- The **browser receives the description/source file and creates the Web page** image that is described

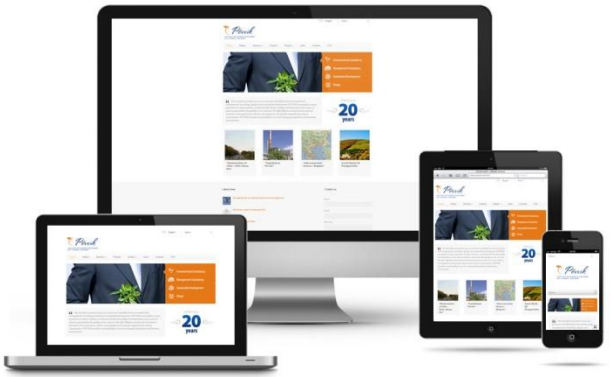
Example – right-click in your browser window and chose "view source". What do you see?



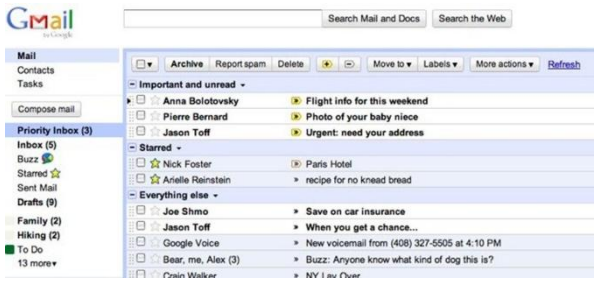
The Modern Web

- Things have *changed*, but many constants remain, e.g. HTML, HTTP, URL, etc.
  - Responsive web, rich web applications, and single page web applications have become common
  - CMS
  - Websites connected to large data stores





AND DYNAMIC SERVER SIDE GENERATED WEB SITES



Activity!

Network Inspection
(Chrome/Firefox tool)

(here if we have time, otherwise at home)

- Open your Chrome or Firefox browser
- Press **option+command+I** (or right-click, and select "inspect" from the context menu)
- Check out the inspection of elements of code (HTML and other code /resources called in via HTML)
- Then **click on the "Network" tab** with the up/down arrow icon (refresh browser)
- View the specifications for loaded resources -- do you see a lot of “200”...?

Visual allegories for HTTP status codes: Find your HTTP status cat/dog

https://http.cat/

https://httpstatusdogs.com/

HTTP status codes come in "series"
200 series (OK)
30x series (moved)
40x series (user errors)
50x series (server errors)

e.g.:
USER ERRORS (400 series)
- Unauthorized 401
- Forbidden 403
- Not found 404

overall, through history and present day configurations, we can see:

- infrastructure, physicality, materiality
- vital importance of protocols/standards
- funding interests & institutional support
- interoperability agreements
- people at the intersections of networks, visions of potential futures

Preview!

What's Coming Up...

- Weeks 4 - 5:
  - HTML/CSS: hands-on coding!
  - Prep for creating your web site
  - hosting options: PrattMySite and GitHub Pages will be discussed

- Lab/Quiz: ***Networks & Internet***
  Will be made live tomorrow before 3:00 PM in Canvas
  DUE: two weeks later 9/27 at class time*
  ***(deadline extended!)**

- Review current events schedule for your specific dates
  First presentations **Week 5**
  (these are in-class only presentations)

Next Week: The Coding Begins!!

- Next week – web fundamentals, part one + HTML; materials will include both tutorials and readings throughout the upcoming hands-on-focused weeks

  We will spend 5 weeks (+ 1 holiday week) luxuriating in web development basics and having fun making stuff!

- Note the Readings & Tutorials in Canvas for upcoming week
  - Hands-on learning HTML via FreeCodeCamp's tutorials
  - Reading: MDN web resource sections
  - Reading: Castro & Hyslop book PDF
  - Reading: "Interneting Is Hard" web resource
- Recommended approach:
  - **Read** one reading that describes the **CONCEPTS** (either Castro & Hyslop PDF or MDN online site), **THEN**
  - Do some of the **hands-on work** in FreeCodeCamp
  - **THEN intersperse** the rest of the readings with the hands-on

- Remember: Tech tutor (Tk) is available for additional help outside class + Open Lab Hours Monday afternoons

Let's preview the web development environment

Let's preview 2 resources
There are also others in the module

Open Canvas and go to next week's module. Click on the FreeCodeCamp link: https://www.freecodecamp.org/learn/2022/responsive-web-design/

Also the MDN (Mozilla Developer Network) links: https://developer.mozilla.org/en-US/docs/Learn/HTML/Introduction_to_HTML

Soon: We will use a coding environment on Pratt desktops/your laptops

As we begin the web development and coding section of the class, you'll want to start thinking about setting up your coding environment on the computer you use at home.

FreeCodeCamp provides all of the coding environment you'll need for the coming week. You will see they provide a way to see the code you write in a text editor-like window alongside the rendered web page in another window next to it. We'll go over setting up your coding environment for work beyond the tutorials/code camp next week.

We will use VS Code in next class (or you may use your preferred text editing tool) + a modern browser of your choice

Network & Databases Lab/Quiz

- Available on Canvas tomorrow at 3:00pm
- **Due by the start of class two weeks from today SEPT 27**
  - Full description doc posted on the LMS

- Open book/open class slides (as you would also look up information from trusted sources in a real-world environment)
- 10% of grade, submit as a PDF or DOC as in instructions

CURRENT EVENTS PRESENTATIONS

- Check that you are signed up on Sched and note that date in your calendar
  (If you need to revise your date, pick an empty slot or reach out to colleagues to arrange a swap)
- A current event that relates to information technology and has a current (or potential) impact on the field of information

- *OPTION*: If you want to you may present on the same topic with another person who has a shared interest but each must present a separate section and be the equivalent content of an individual presentation

Current Events

Each student will participate in one current events presentation spaced throughout weeks 5 and 14 of the semester.

- You may present individually or with one of your classmates (“panel” would need to address different aspects of the topic in as full detail & time as individuals)

For the presentation, find a current event that relates to information technology and has a current (or potential) impact on the field of information.

- Some starting points are in the Technology News Sources doc (link on Canvas)

In your presentation (**10-15 minutes** for each person) you should:

- Provide a **summary** of the current event
- Describe and explain the **technologies** involved in this event, using appropriate technical terminology
- Explain the **impact or relevance to the field of information**, in the present and/or future
- You may include short audio-visual materials as needed (all materials must be presented from the front-of-classroom computer)
- In-class presentation ONLY (nothing else to submit)

Full assignment description and schedule on Canvas

Closing Activity!

The Internet Archive
&
Wayback Machine

- Visit the Internet Archive's Wayback Machine: **https://archive.org/web/**
- Put a URL in their Wayback Machine search tool and explore earlier iterations (university and library sites are good to start with! what did Pratt.edu or nypl.org look like in the 1990s and early 2000s….?)
- Also read more about the Internet Archive under "Projects": https://archive.org/
- (if you are in MLIS, check out their Archive-It tool! https://archive-it.org/)

Closing Activity!

Alternate:
Art Net Anthology

- Rhizome's Art Net Anthology
- https://anthology.rhizome.org/
- Rhizome projects: https://rhizome.org/software/
- and related exhibition description from The New Museum

thank you for being here!
and see you next week

questions? come to office hours (right now in 607)
or email me: rdaniell@pratt.edu