CS433/533: Computer Networks

Y. Richard Yang

http://zoo.cs.yale.edu/classes/cs433/

8/30/2018

Outline

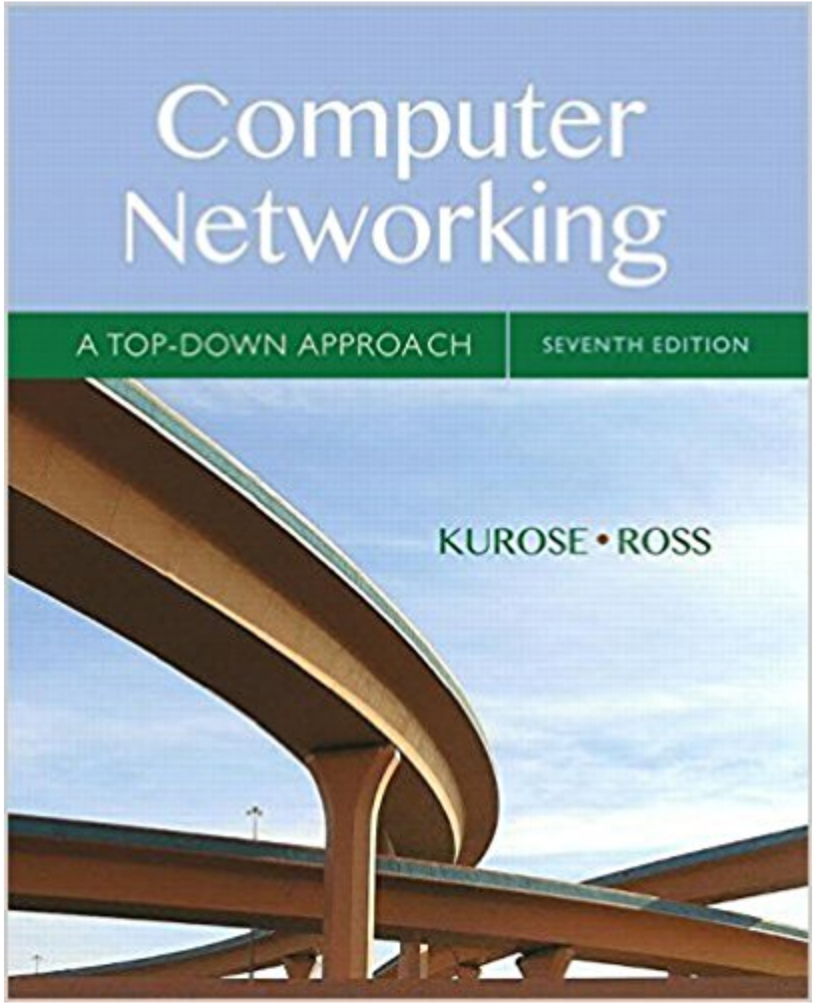
- *Administrative trivia's*
- What is a network protocol?
- A brief introduction to the Internet: past and present
- Challenges of Internet networks and apps
- Summary

Personnel

- Instructor
  - Y. Richard Yang, yry@cs.yale.edu, AKW 208A
    - office hours
      - TTh 2:30-3:30 pm or by appointment
      - please feel free to stop by if you see I am in my office
- Teaching assistant
  - Geng Li, AKW 204
    - office hours MW 2:30-4:00 pm (may change)

Textbook

- Textbook
  - *Computer Networking: A Top-Down Approach, 7/e*
    by Jim Kurose and Keith Ross
- Reference books
  - *Computer Networks*
    by Tanenbaum and Wetherall on
  - *Computer Networks, A Systems Approach*
    by Larry Peterson and Bruce Davie
  - *TCP/IP Illustrated, Volume 1: The Protocols*
    by W. Richard Stevens
  - Java Network Programming,
    by Elliotte Harold
- Resources
  - http://zoo.cs.yale.edu/classes/cs433



What are the Goals of this Course?

- Learn design principles and techniques of:
  - network infrastructures, e.g.,
    - the global Internet infrastructure,
    - the infrastructures of providers (e.g., Google, FB)
    - emerging infrastructures (e.g., wireless)
  - large-scale applications, e.g.,
    - traditional client-server apps (e.g., email, Web, DNS)
    - distributed, p2p apps (e.g., Blockchain, overlays)

- Focus on how the principles and techniques apply and adapt in real world

CPSC433/533 or EENG 452

- CPSC433/533: More from a computer science's view, emphasizing more on systems design and programming design

- EENG 452: More from an engineering's view, emphasizing more on analysis

What Do You Need To Do?

- Please return the class background survey
  - help us determine your background
  - help us determine the depth, topics, and assignments
  - suggest topics that you want to be covered (if you think of a topic later, please send me email)
- Your workload
  - homework assignments
    - written assignments
    - programming assignments
      - one HTTP 1.1 server, one TCP, and one robot network using BeagleBone
  - two exams

Grading

Exams	30%
Assignments	60%
Class Participation	10%

- Subject to change after I know more about your background
- More important is what you realize/learn than the grades !!

Questions?

Outline

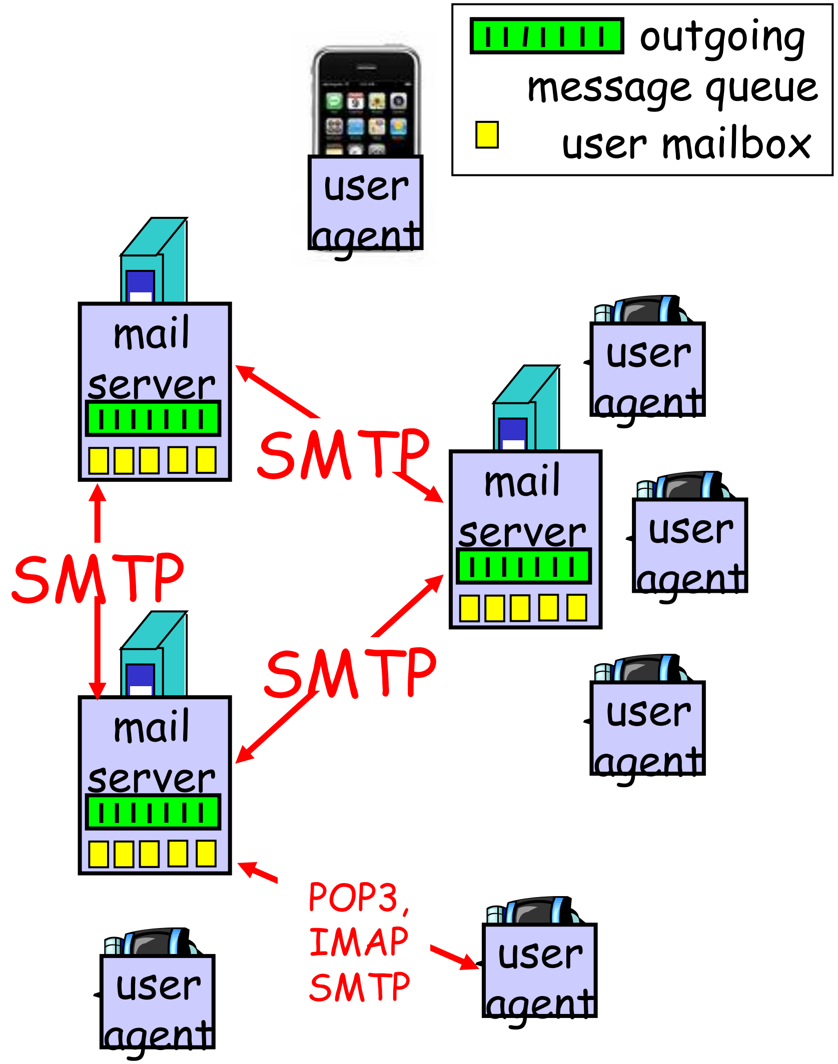
- Administrative trivia's
- *What is a network protocol?*

What is a Network Protocol?

- A **network protocol** defines the format and the order of messages exchanged between two or more communicating entities, as well as the actions taken on the transmission and/or receipt of a message or other events.

Example Protocol: Simple Mail Transfer Protocol (SMTP)

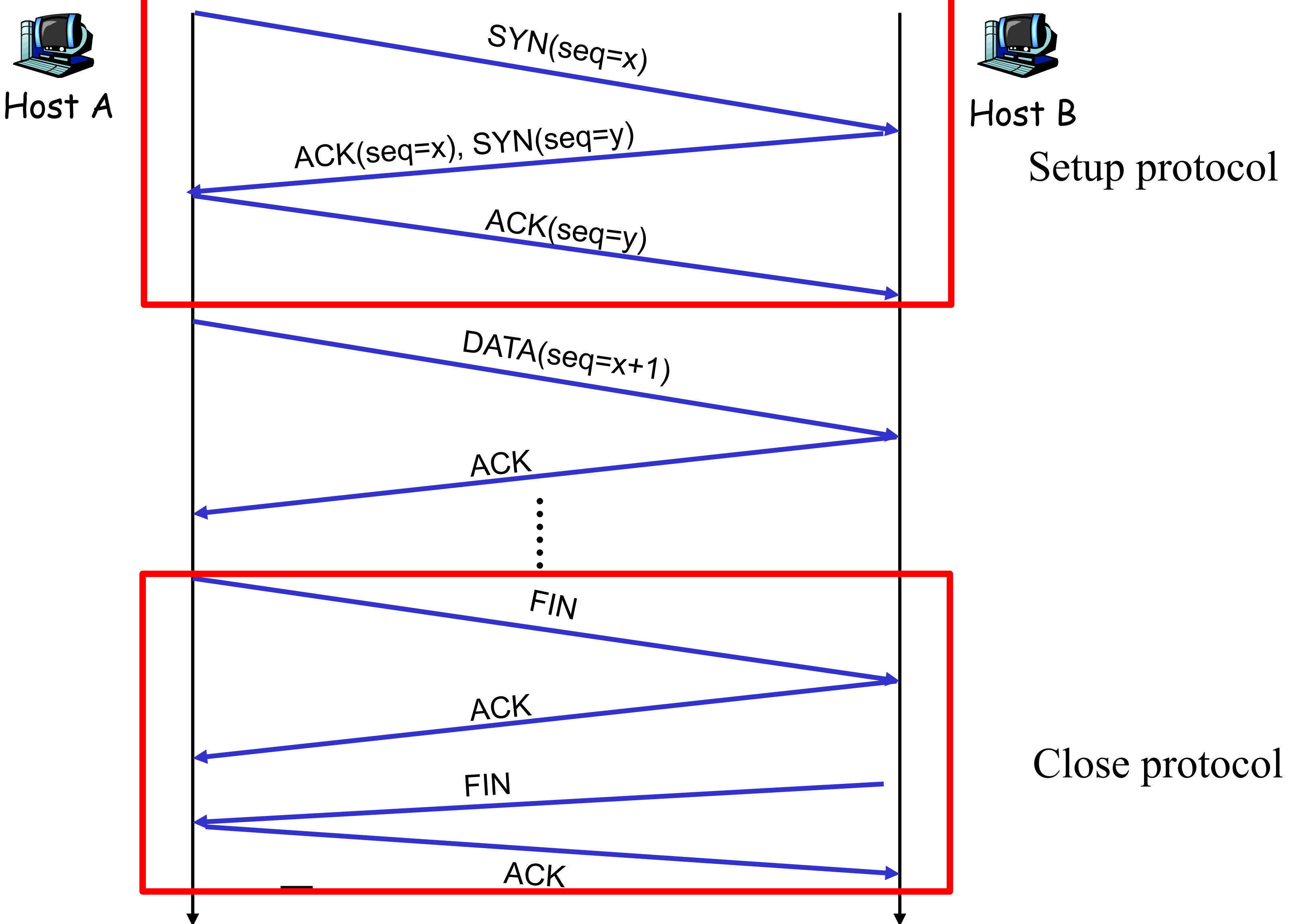
- Messages from a client to a mail server
  - HELO
  - MAIL FROM: <address>
  - RCPT TO: <address>
  - DATA
    <This is the text end with a line with a single .>
  - QUIT
- Messages from a mail server to a client
  - status code
    - The first digit of the response broadly indicates the success, failure, or progress of the previous command.
      - 1xx - Informative message
      - 2xx - Command ok
      - 3xx - Command ok so far, send the rest of it.
      - 4xx - Command was correct, but couldn't be performed for some reason.
      - 5xx - Command unimplemented, or incorrect, or a serious program error occurred.
  - content



Command: %telnet netra.cs.yale.edu smtp

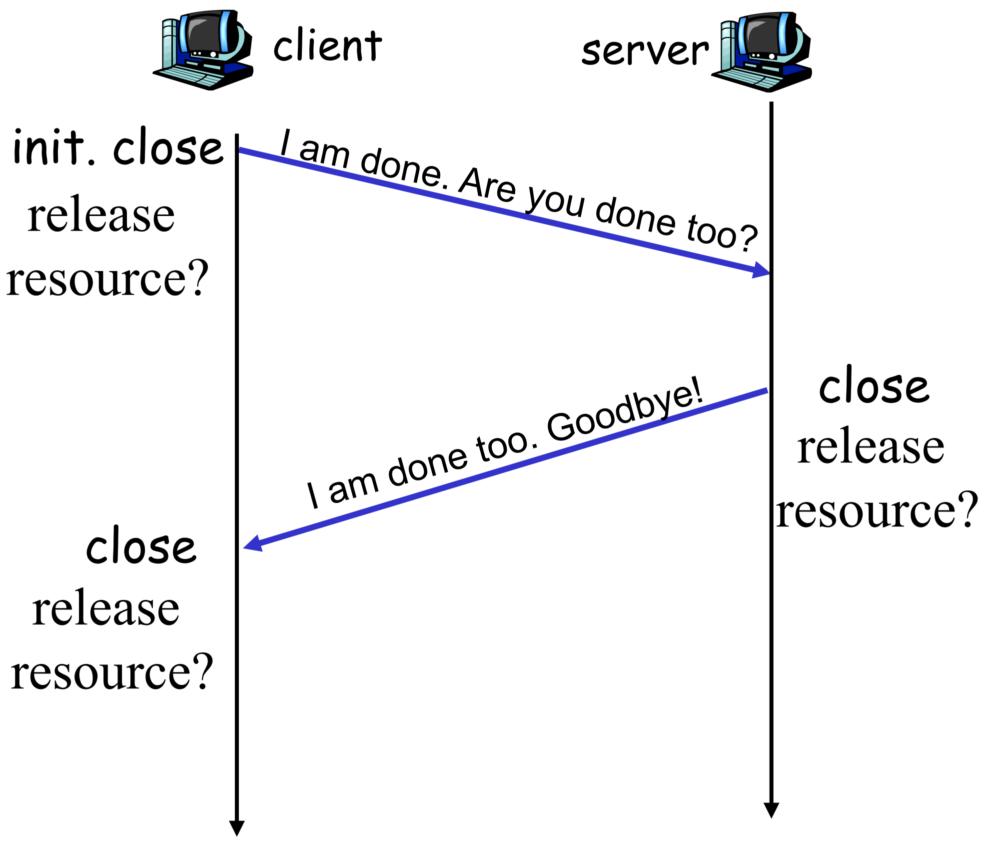
Wireshark capture: port smtp

Example Protocol: Complete SMTP



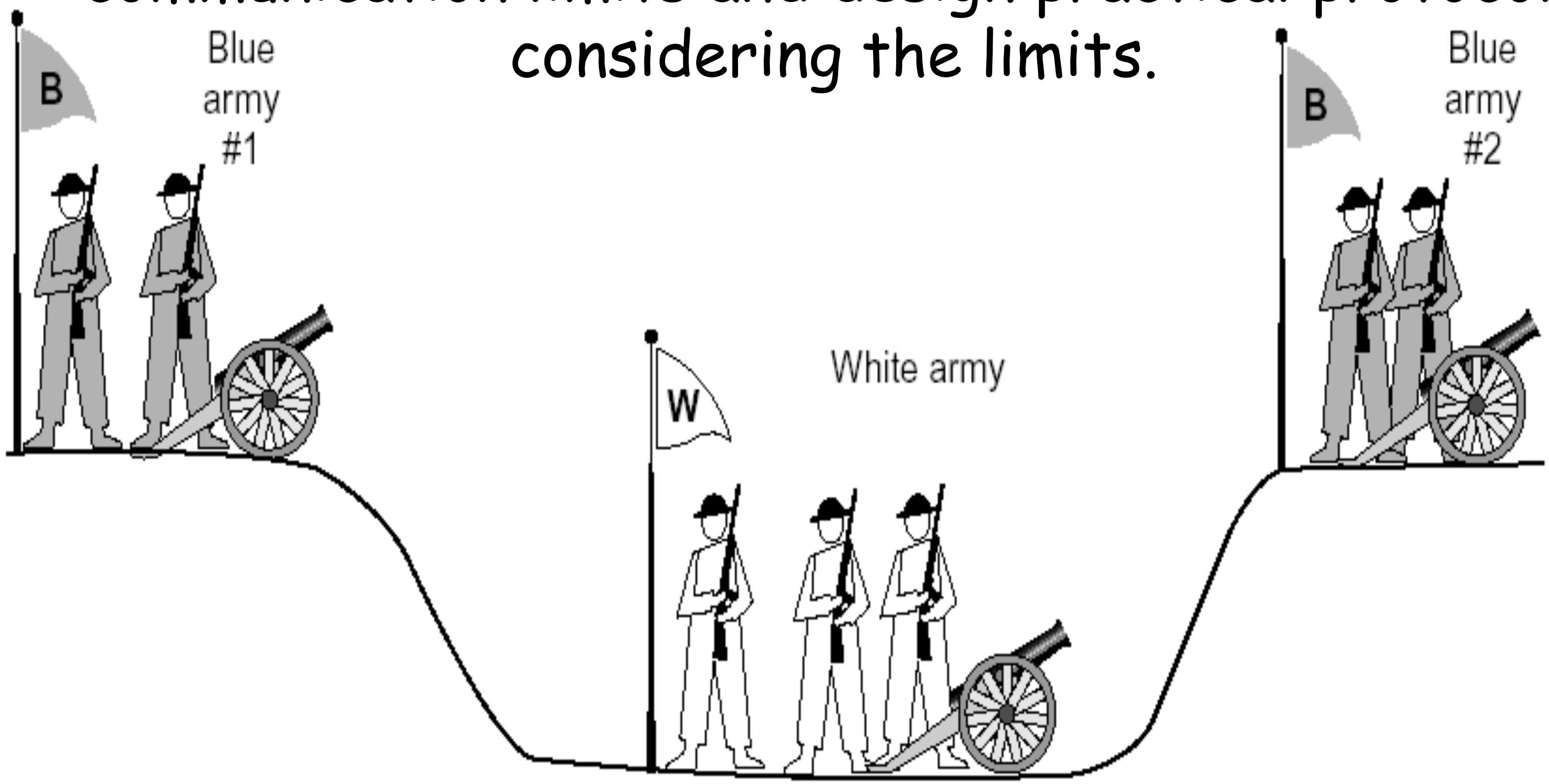
Example Protocol: Connection Close Protocol

- Why connection close protocol?
  - so that each side can release resource and remove state about the communication



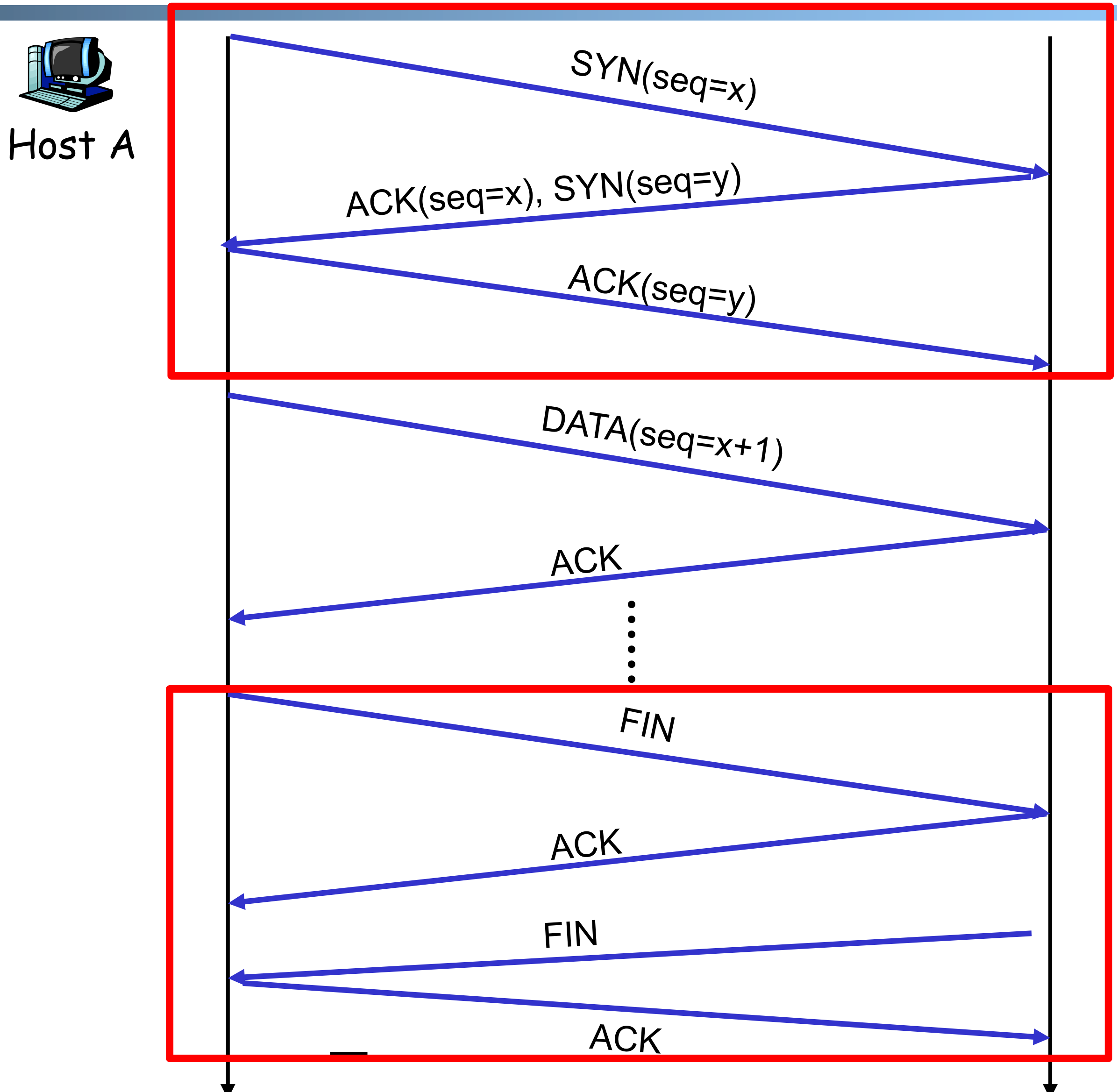
General Case: The Two-Army Problem

Protocol Design needs to understand fundamental communication limits and design practical protocols considering the limits.



The gray (blue) armies need to agree on whether or not they will attack the white army. They achieve agreement by sending messengers to the other side. If they both agree, attack; otherwise, no. Note that a messenger can be captured!

Example Protocol: Complete SMTP



Setup protocol

- Obvious thing you do not like the setup protocol?

Close protocol

Between Google services and Chrome / mobile apps

Example: Google' new QUIC

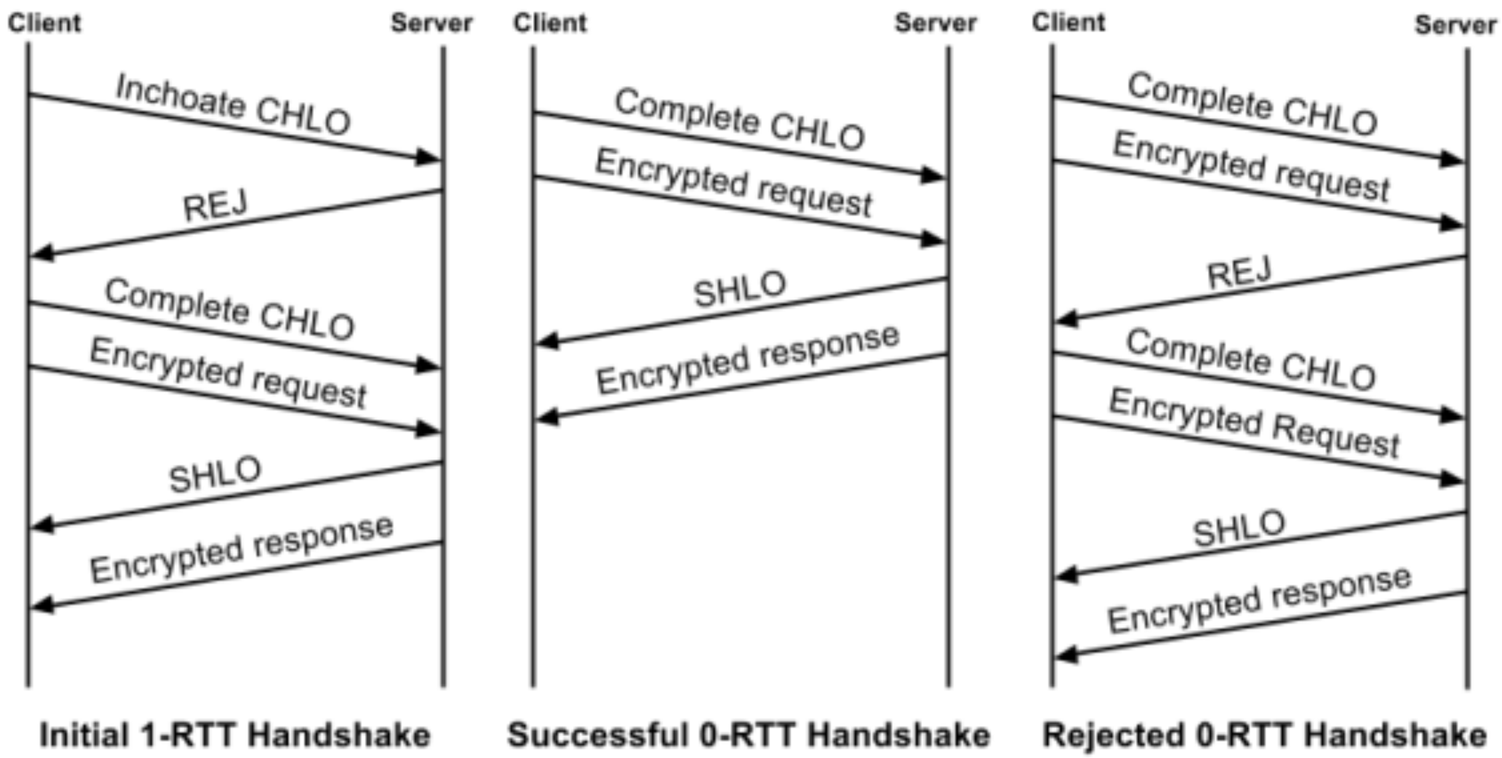


Figure 4: Timeline of QUIC's initial 1-RTT handshake, a subsequent successful 0-RTT handshake, and a failed 0-RTT handshake.

https://static.googleusercontent.com/media/research.google.com/en//pubs/archive/46403.pdf
http://conferences.sigcomm.org/sigcomm/2017/files/program/ts-5-1-QUIC.pdf

Protocol Standardization

- Most widely used protocols are defined in standards
- **Why standard?**

Internet Standardization Process

- All standards of the Internet are published as RFC (Request for Comments)
  - e.g., the SMTP protocol is specified in RFC821
  - but not all RFCs are Internet Standards: http://zoo.cs.yale.edu/classes/cs433/cs433-2018-fall/readings/interestingrfcs.html

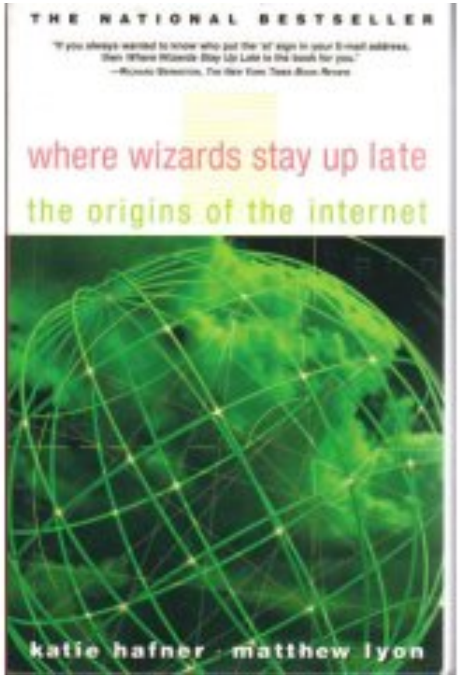
Internet Standardization Process

- All standards of the Internet are published as RFC (Request for Comments)
  - e.g., the SMTP protocol is specified in RFC821
  - but not all RFCs are Internet Standards: http://zoo.cs.yale.edu/classes/cs433/cs433-2018-fall/readings/interestingrfcs.html
- A typical (but not the only) way of standardization:
  - Internet draft
  - RFC
  - proposed standard
  - draft standard (requires 2 working implementations)
  - Internet standard (declared by Internet Architecture Board)
- David Clark, 1992:

  We reject: kings, presidents, and voting. We believe in: rough consensus and running code.

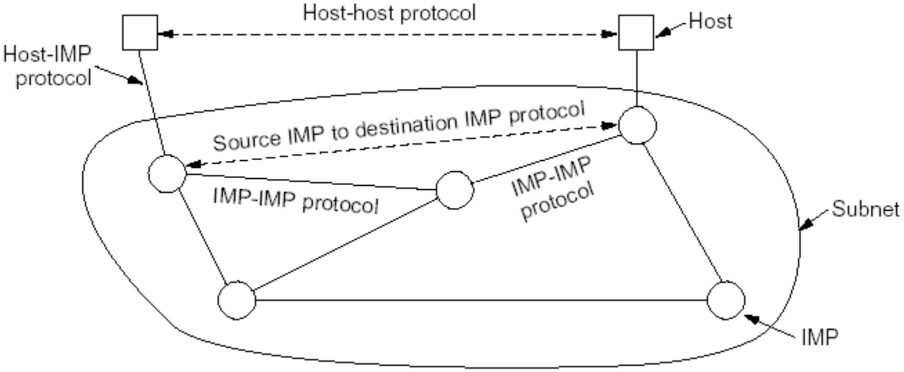
Outline

- Administrative trivia's
- What is a network protocol?
- *A brief introduction to the Internet*
  - *past (a brief history)*



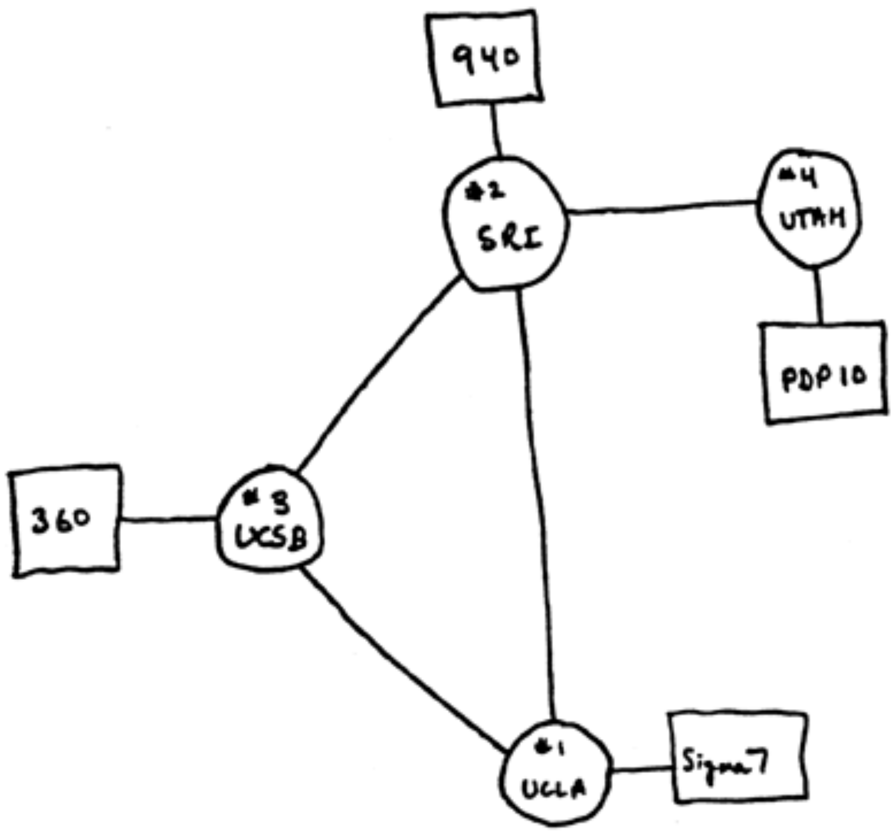
Prelude: Packet Switching and ARPANET

- **1957**
  - USSR launched Sputnik; US DoD formed Advanced Research Projects Agency (ARPA)
- **1961**
  - First paper by Len Kleinrock on packet switching theory
- **1964**
  - Paul Baran from RAND on design of packet switching networks
- **1965-1968**
  - ARPANET plan
  - Bolt Beranek and Newman, Inc. (BBN), a small company, was awarded Packet Switch contract to build Interface Message Processors (IMPs)

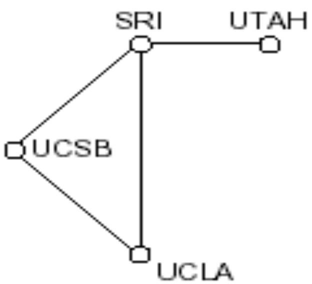


Internet 1.0: Initial ARPANET

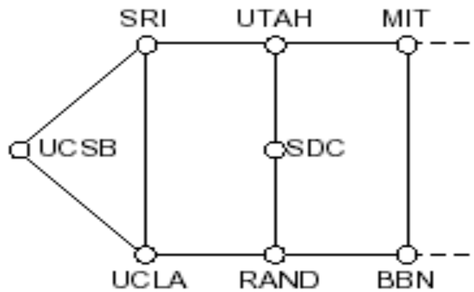
- 1969
  - ARPANET commissioned: 4 nodes, 50kbps



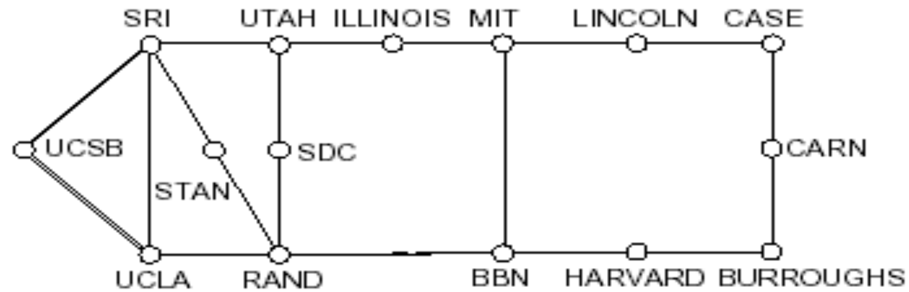
Initial Expansion of the ARPANET



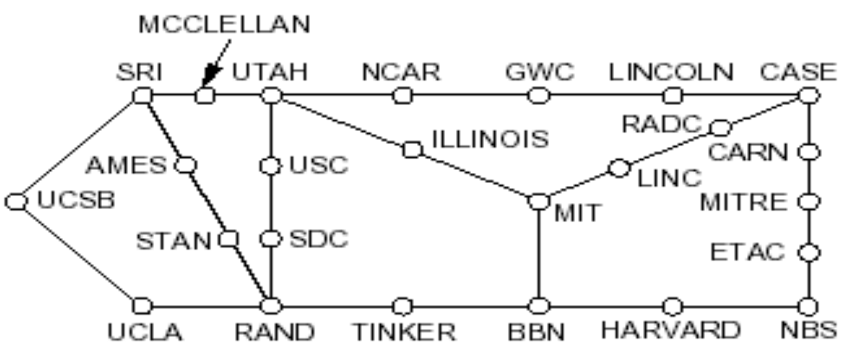
Dec. 1969



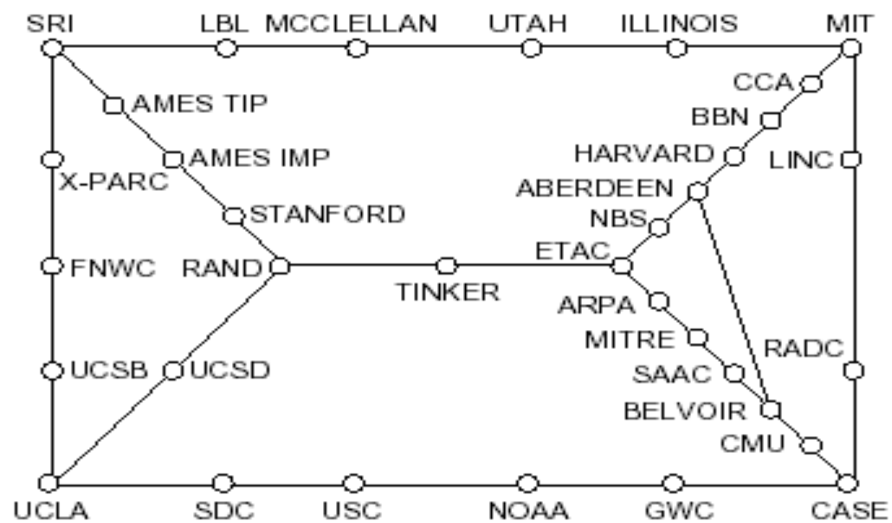
July 1970



Mar. 1971



Apr. 1972



Sept. 1972

RFC 527: ARPAWOCKY; RFC 602: The Stockings Were Hung by the Chimney with Care

The Internet Becomes a Network of Networks

- 1970: ALOHAnet, the first packet radio network, developed by Norman Abramson, Univ of Hawaii, became operational
- 1973: Bob Kahn posed the Internet problem---how to connect ARPANET, packet radio network, and satellite network
- 1974: Vint Cerf, Bob Kahn published initial design of TCP (NCP) to connect multiple networks – the Internet

Growth of the Internet

- 1981: BITNET (Because It's Time NETwork) between CUNY and Yale
- 1986: NSF builds NSFNET as backbone, links 6 supercomputer centers, 56 kbps; this allows an explosion of connections, especially from universities
- 1987: 10,000 hosts
- 1988: Internet congestion collapse; TCP congestion control
- 1989: 100,000 hosts

RFC 1121: Act One – The Poem

WELCOME by Leonard Kleinrock

We've gathered here for two days to examine and debate
And reflect on data networks and as well to celebrate.
To recognize the leaders and recount the path we took.

We'll begin with how it happened; for it's time to take a look.
Yes, the history is legend and the pioneers are here.
Listen to the story - it's our job to make it clear.
We'll tell you where we are now and where we'll likely go.
So welcome to ACT ONE, folks.
Sit back - enjoy the show!!

Internet 2.0: Web, Commercialization, Social Networking of the Internet

- 1990: ARPANET ceases to exist
- 1991: NSF lifts restrictions on the commercial use of the Net; Berners-Lee of European Organization for Nuclear Research (CERN) released World Wide Web
- 1992: 1 million hosts (RFC 1300: Remembrances of Things Past)
- 1998: Google was founded
- 2004: Facebook was founded
- 2006: Amazon AWS cloud computing

For a link of interesting RFCs, please see
http://zoo.cs.yale.edu/classes/cs433/cs433-2018-fall/readings/interestingrfcs.html

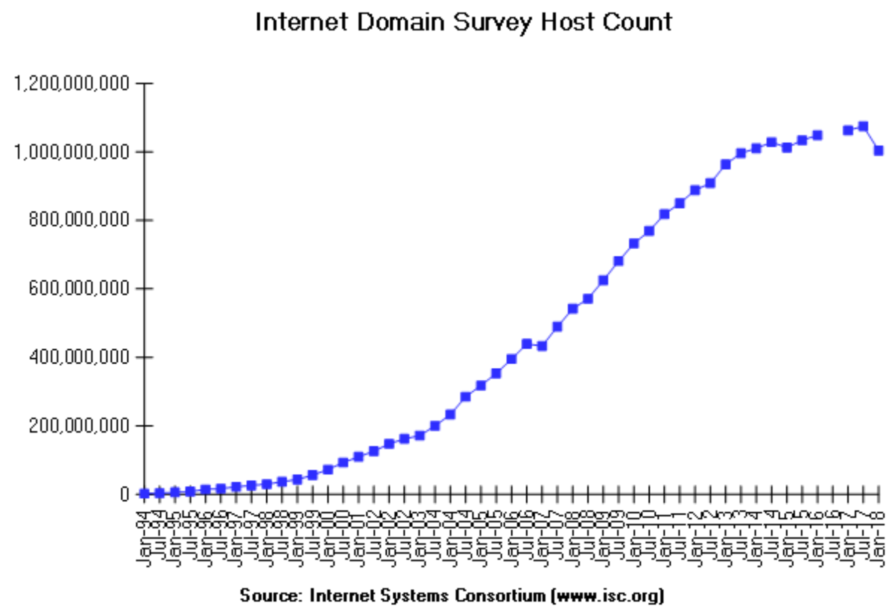
For more on Internet history, please see
http://www.zakon.org/robert/internet/timeline/

Growth of the Internet in Terms of Number of Hosts

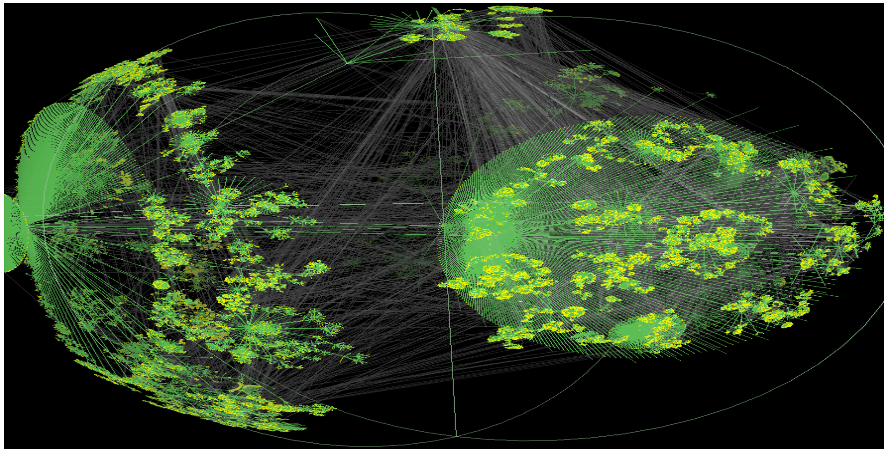
Number of Hosts on the Internet:

Aug. 1981	213
Oct. 1984	1,024
Dec. 1987	28,174
Oct. 1990	313,000
Jul. 1993	1,776,000
Jul. 1996	12,881,000
Jul. 1999	56,218,000
Jul. 2002	162,128,493
Jul. 2005	353,284,187
Jul. 2007	489,774,269
Jul. 2010	768,913,036
Jul. 2013	996,230,757
Jul. 2015	1,033,836,245
Jul. 2018	1,015,787,389

http://ftp.isc.org/www/survey/reports/current/



CAIDA router level view



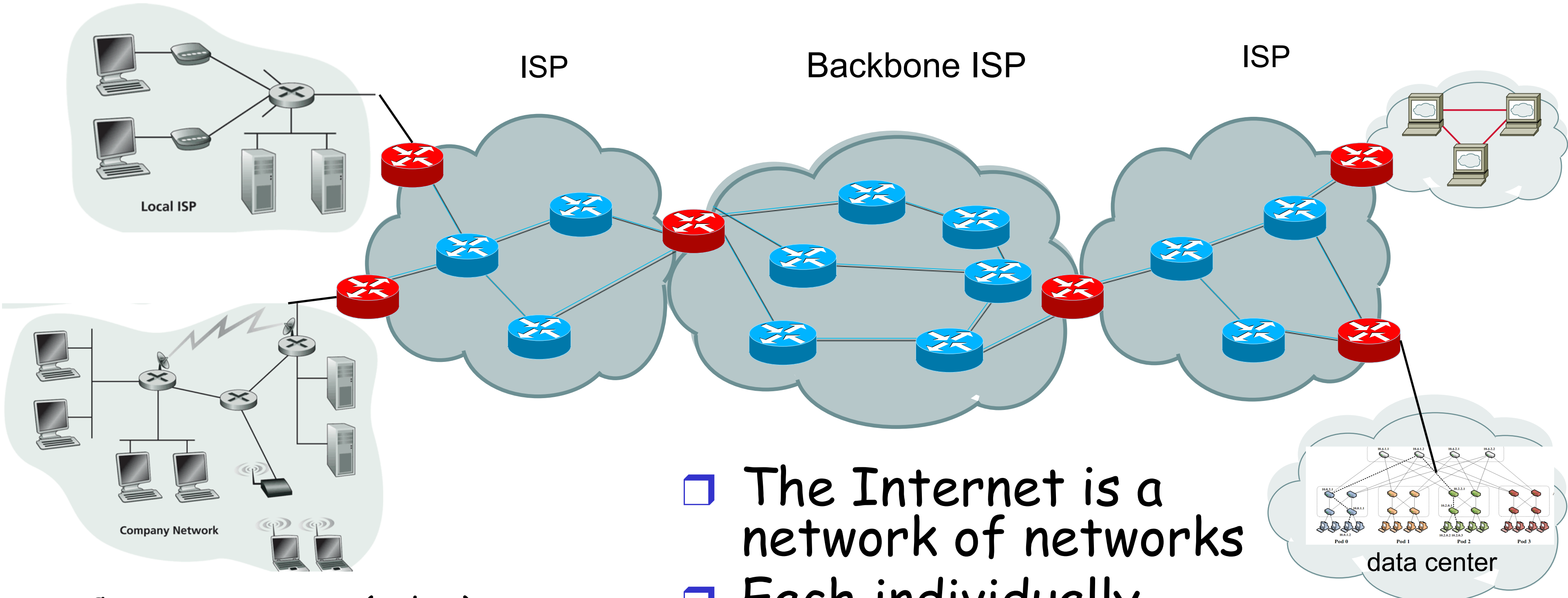
Outline

- Administrative trivia's
- What is a network protocol?
- *A brief introduction to the Internet*
  - past
  - *present*

Internet Physical Infrastructure

Residential access (edge)

- Cable, Fiber, DSL, Wireless

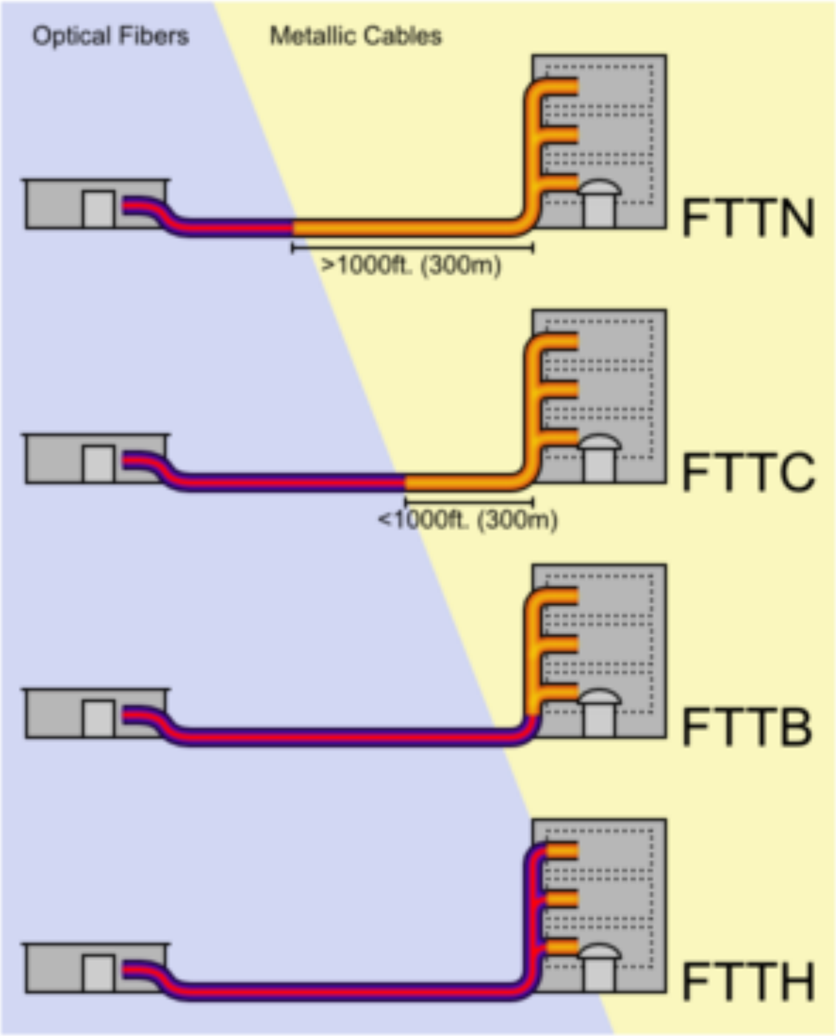


Campus access (edge), e.g.,

- Ethernet
- Wireless

- The Internet is a network of networks
- Each individually administrated network is called an Autonomous System (AS)

Access: Fiber to the x

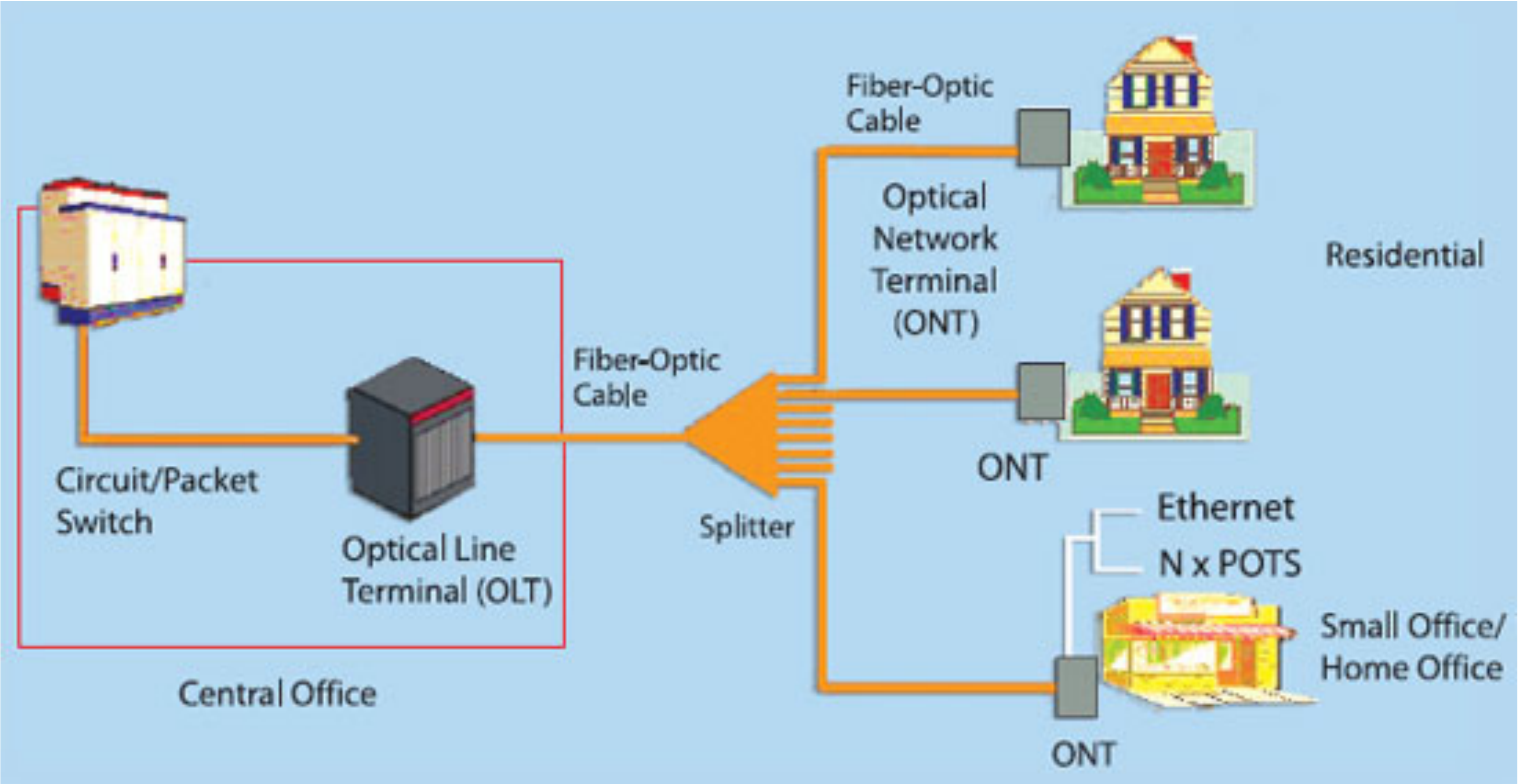


http://en.wikipedia.org/wiki/Fiber_to_the_x

Access: Fiber to the Premises (FTTP)

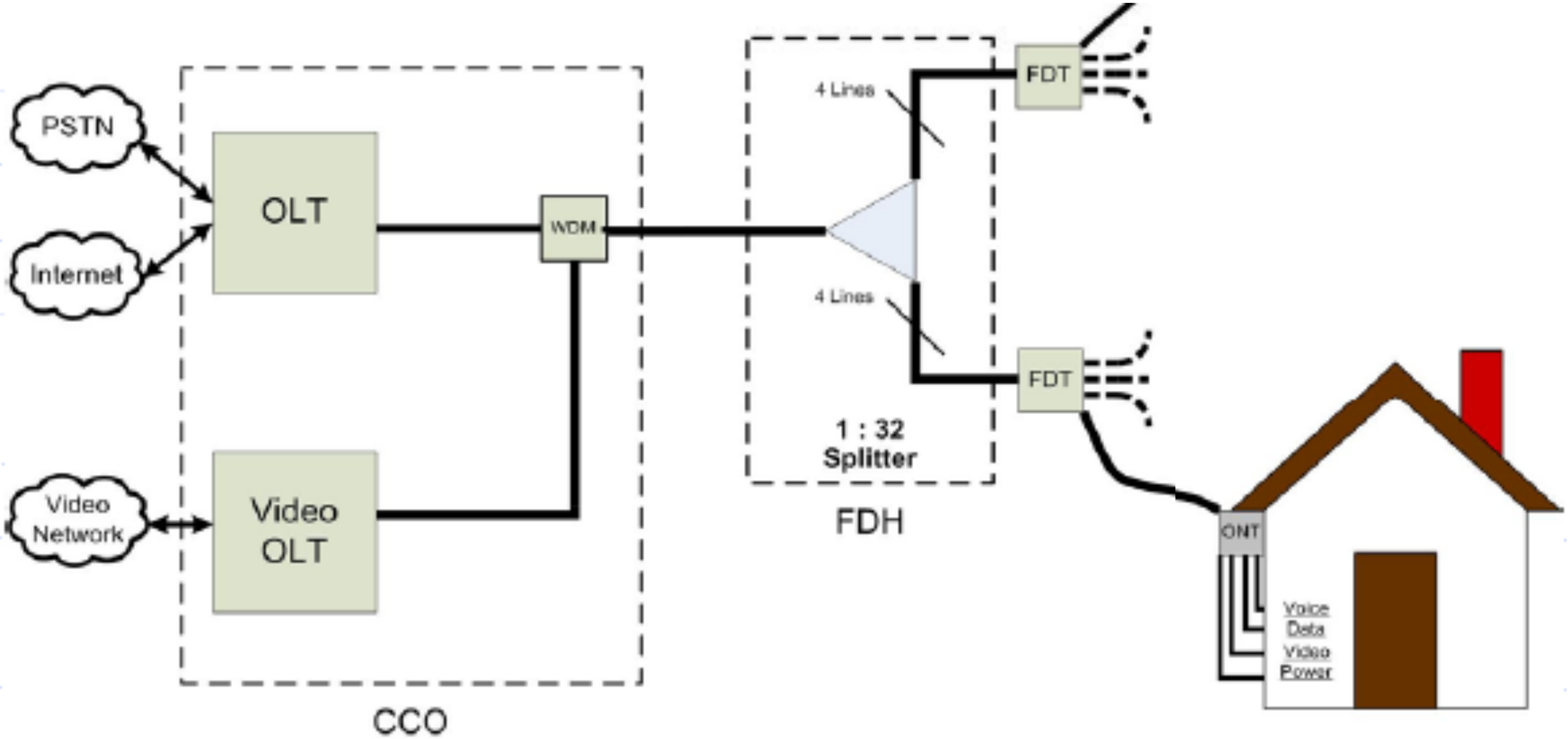
- Deployed by Verizon, AT&T, Google
- One of the largest comm. construction projects

FTTP Architecture

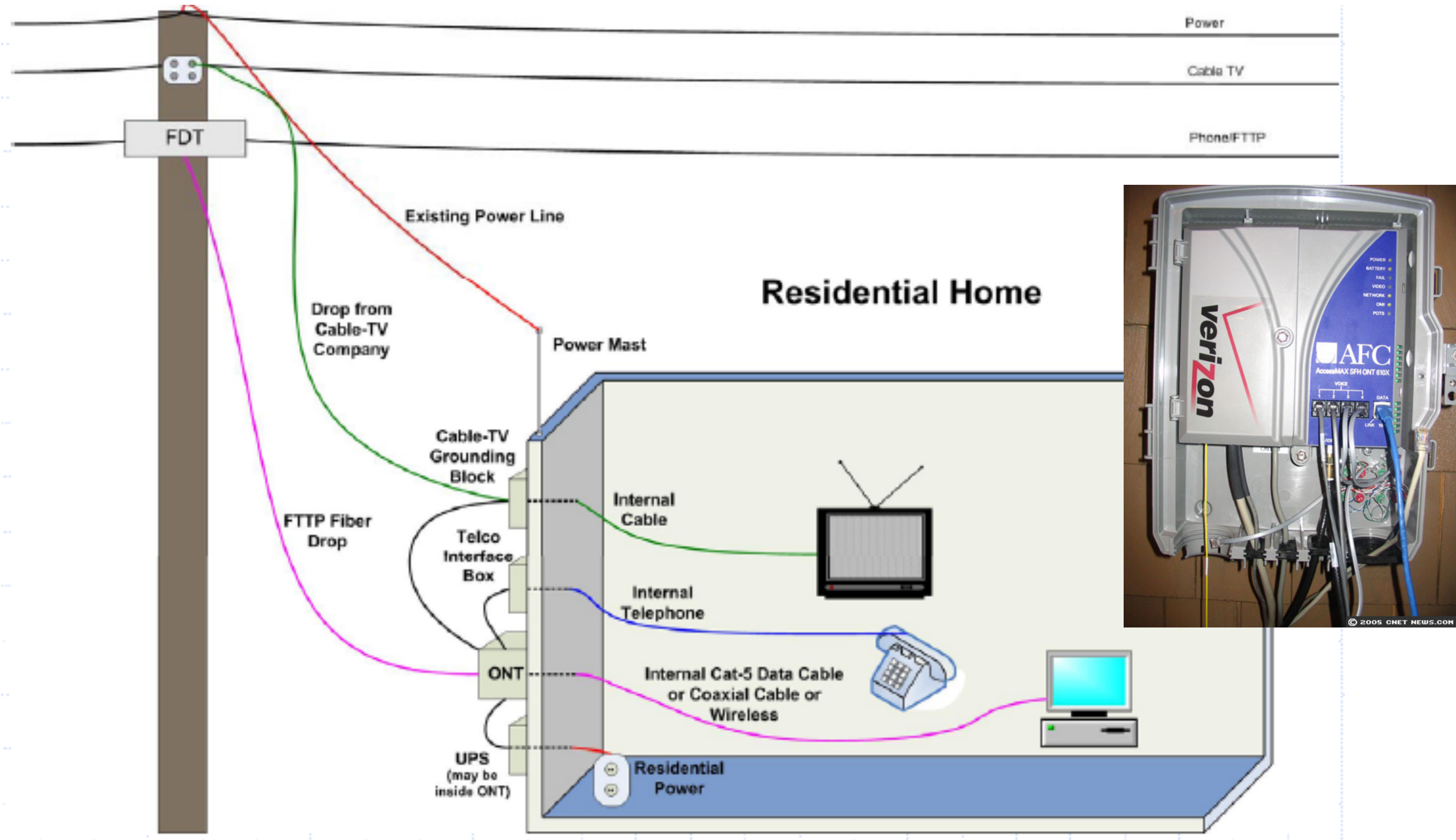


FTTP Architecture

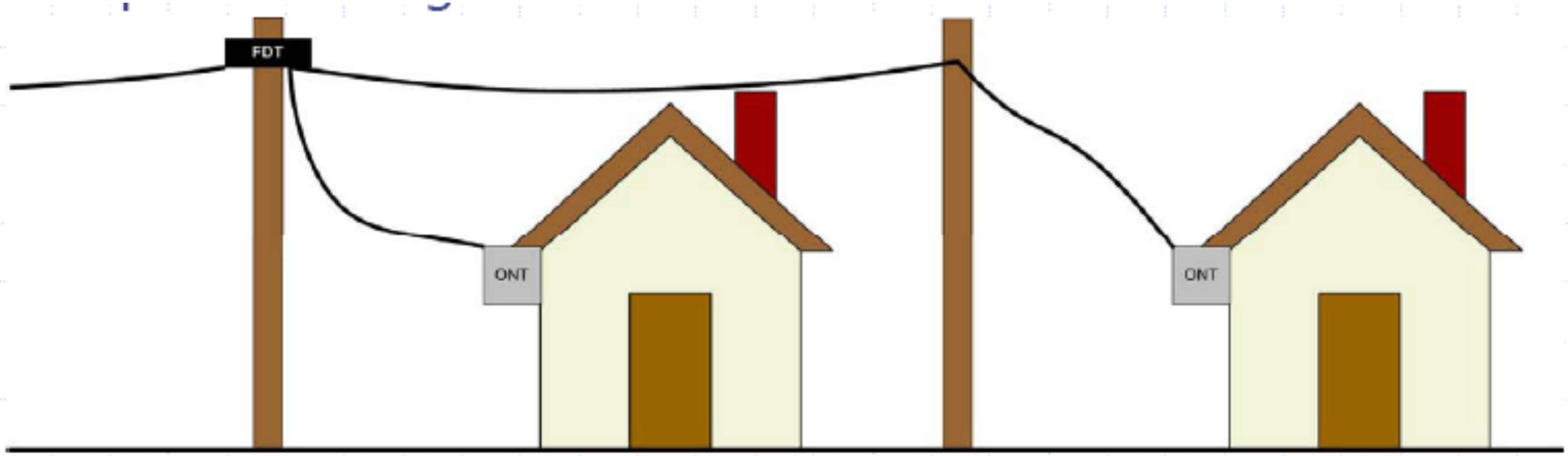
- Optical Network Terminal (ONT) box outside dwelling or business
- Fiber Distribution Terminal (FDT) in poles or pedestals
- Fiber Distribution Hub (FDH) at street cabinet
- Optical Line Terminal (OLT) at central office



FTTP Architecture: To Ho



FTTP Architecture: Fiber Distribution Terminal (FDT)



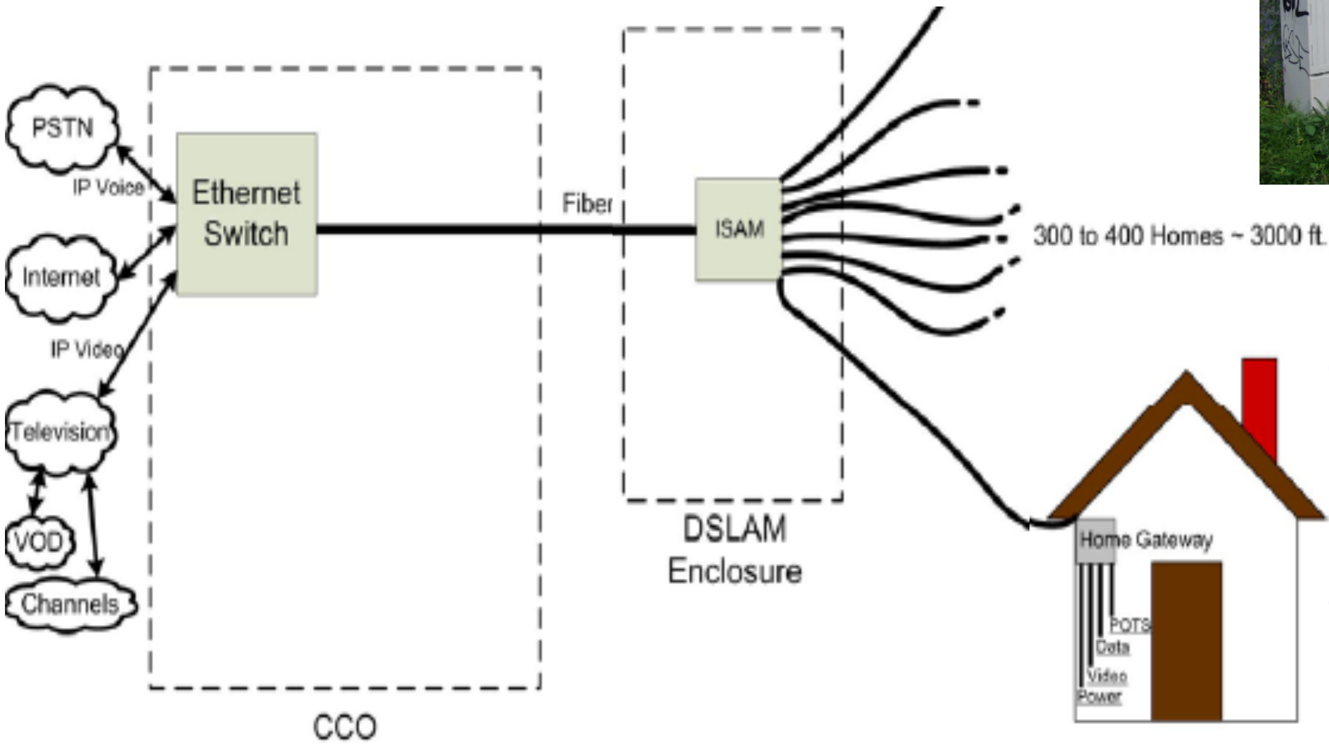
FTTP Architecture: Central to Fiber Distribution Hub (FDH)

- Backbone fiber ring on primary arterial streets (brown)
- Local distribution fiber plant (red) meets backbone at cabinet

FDH

Access: DSL

- Compared with FTTP, copper from cabinet (DSLAM) to home



DSLAM

Access: Wireless

https://x.company/loon/

theguardian

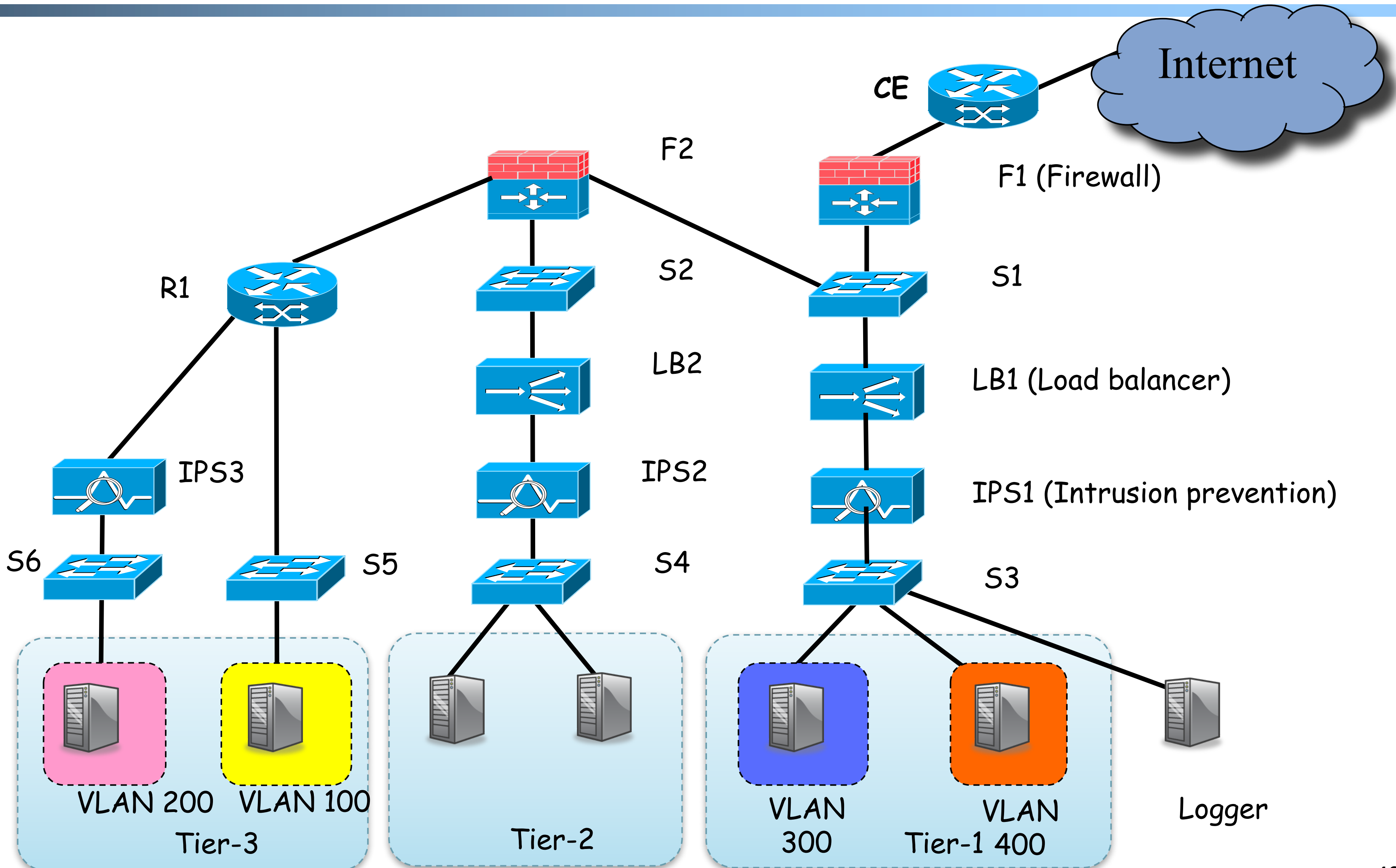
home › tech US politics world opinion sports soccer arts lifestyle fa all

Elon Musk

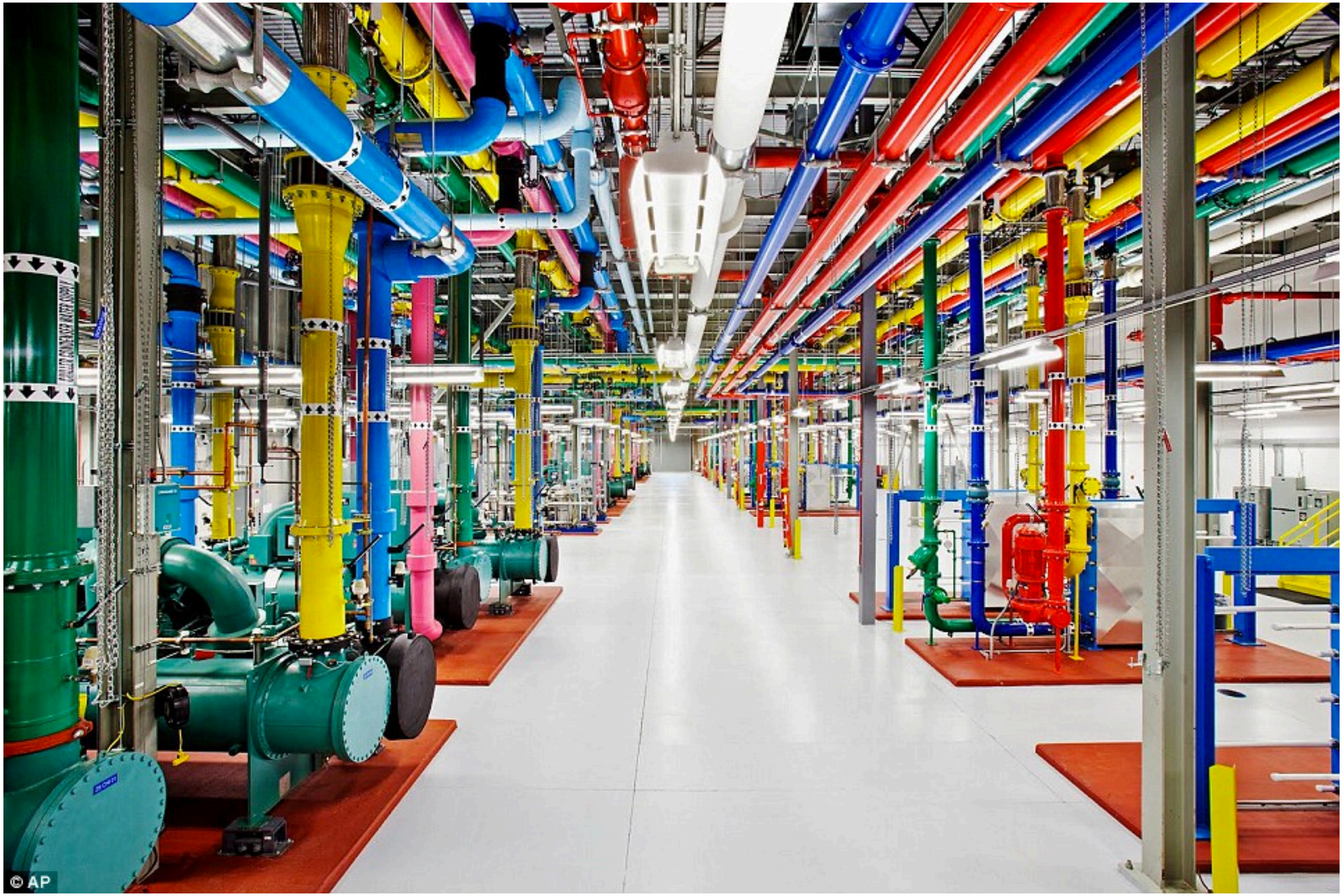
Elon Musk wants to cover the world with internet from space

SpaceX requests permission from US government to operate network of 4,425 satellites to provide high-speed, global internet coverage

Campus Network

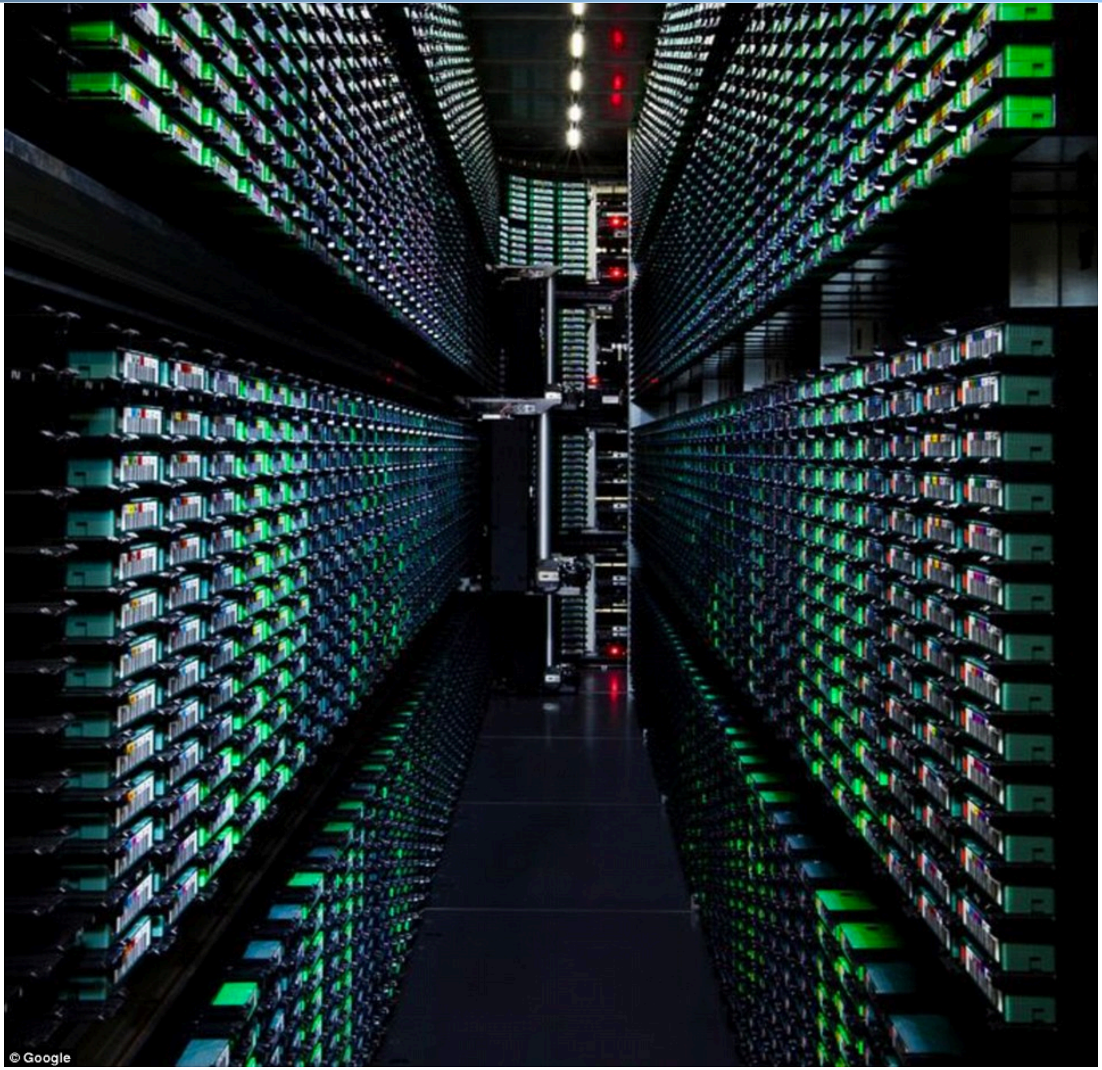


Data Center Networks



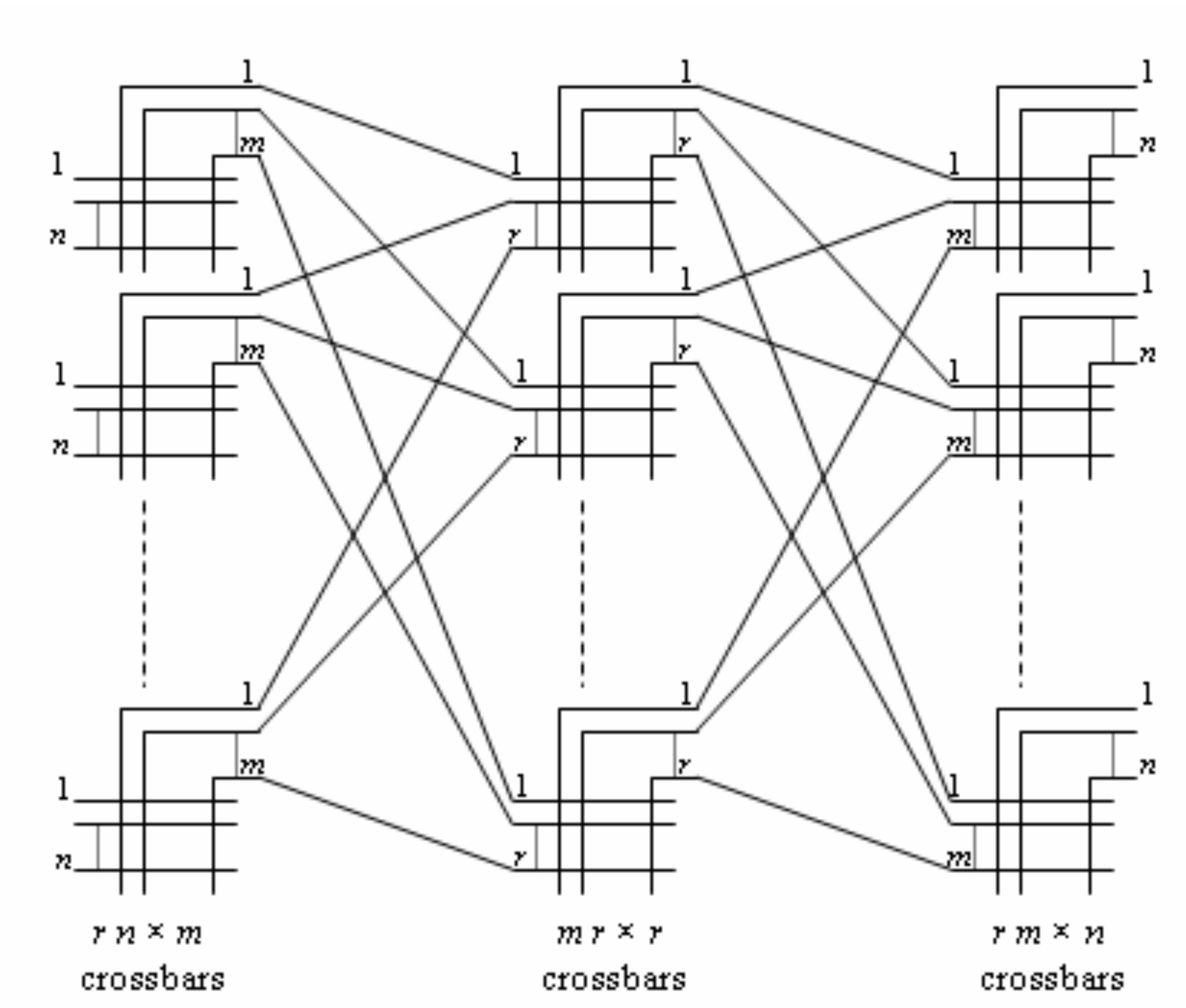
http://www.dailymail.co.uk/sciencetech/article-3369491/Google-s-plan-world-Search-engine-build-half-billion-dollar-data-center-US.html

Data Center Networks



http://www.dailymail.co.uk/sciencetech/article-3369491/Google-s-plan-world-Search-engine-build-half-billion-dollar-data-center-US.html

Foundation of Data Center Networks: Clos Networks



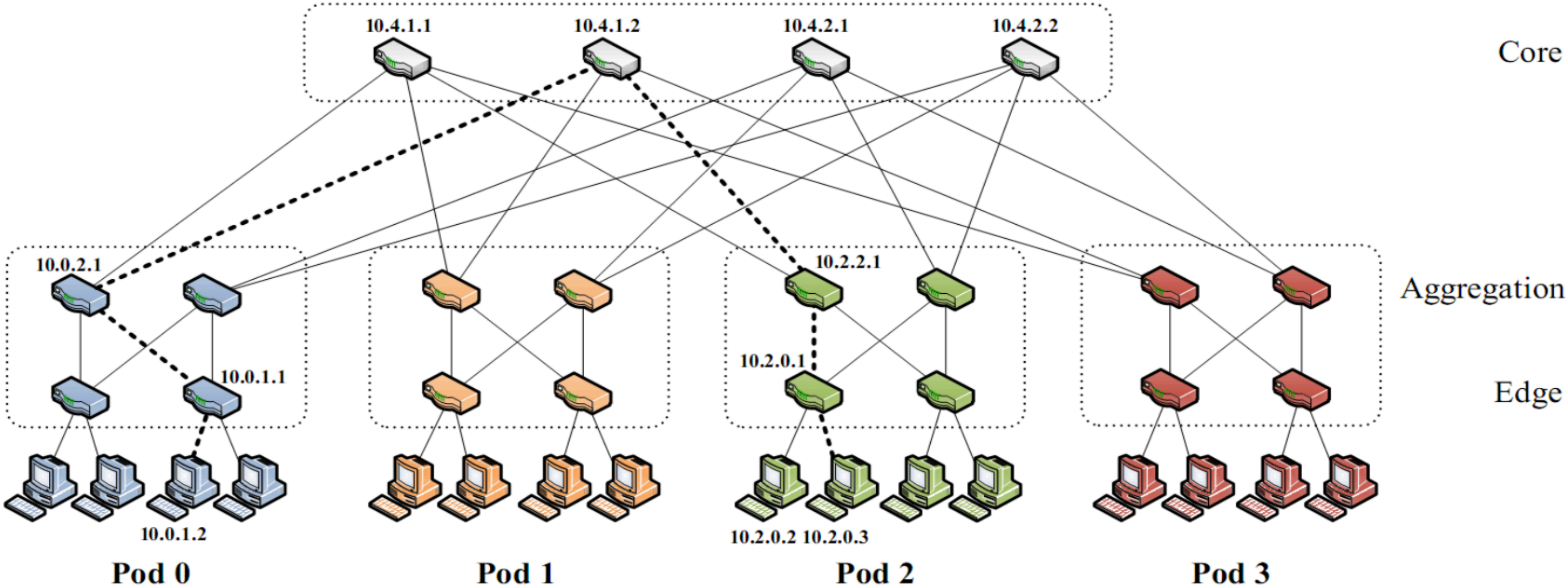
Q: How big is m so that each new call can be established w/o moving current calls?

433/533 Labor day problem: If you can move existing calls, it is only m >= n.

https://en.wikipedia.org/wiki/Clos_network

Data Center Networks: Fat-tree Networks

- K-ary fat tree: three-layer topology (edge, aggregation and core)
  - k pods w/ each pod consists of $(k/2)^2$ servers & 2 layers of k/2 k-port switches
    - each edge switch connects to k/2 servers & k/2 aggr. switches
    - each aggr. switch connects to k/2 edge & k/2 core switches
  - $(k/2)^2$ core switches: each connects to k pods

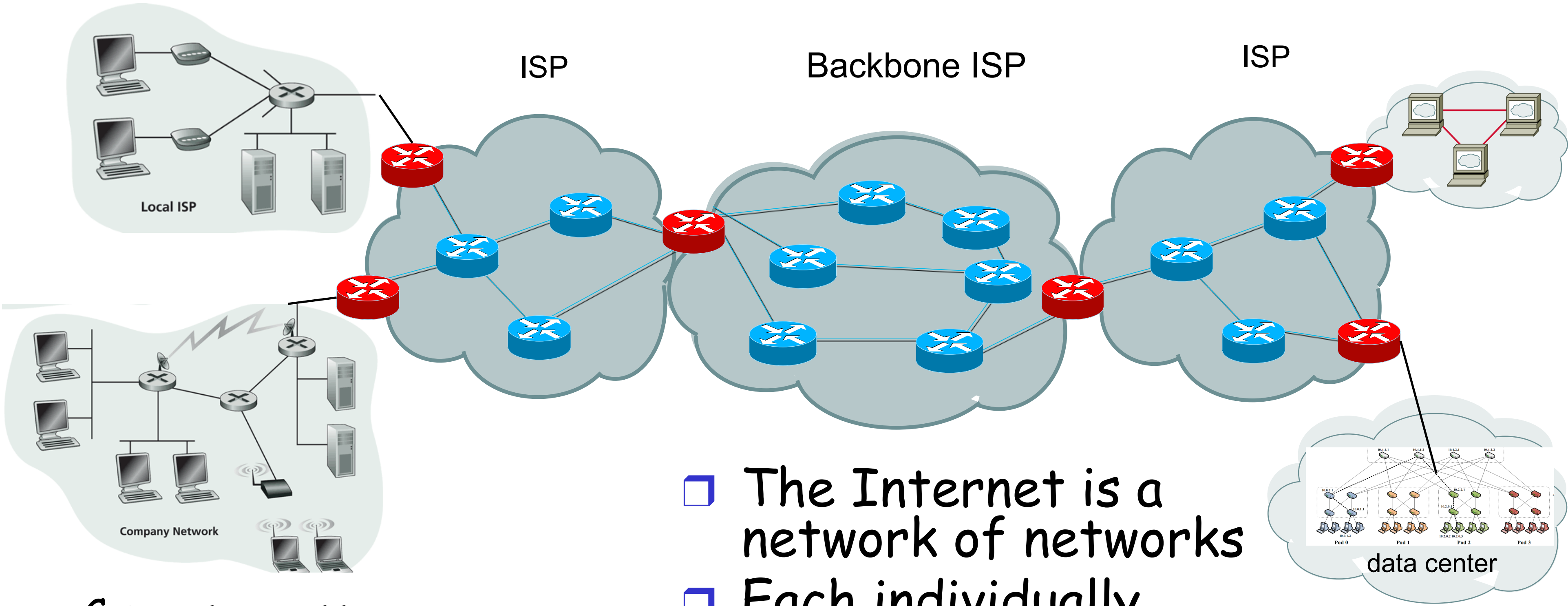


http://www.cs.cornell.edu/courses/cs5413/2014fa/lectures/08-fattree.pdf

Q to think offline: How large a network can k-ary support using k-port switches?

Recall: Internet Physical Infrastructure

Residential access, e.g.,
- Cable, Fiber, DSL, Wireless



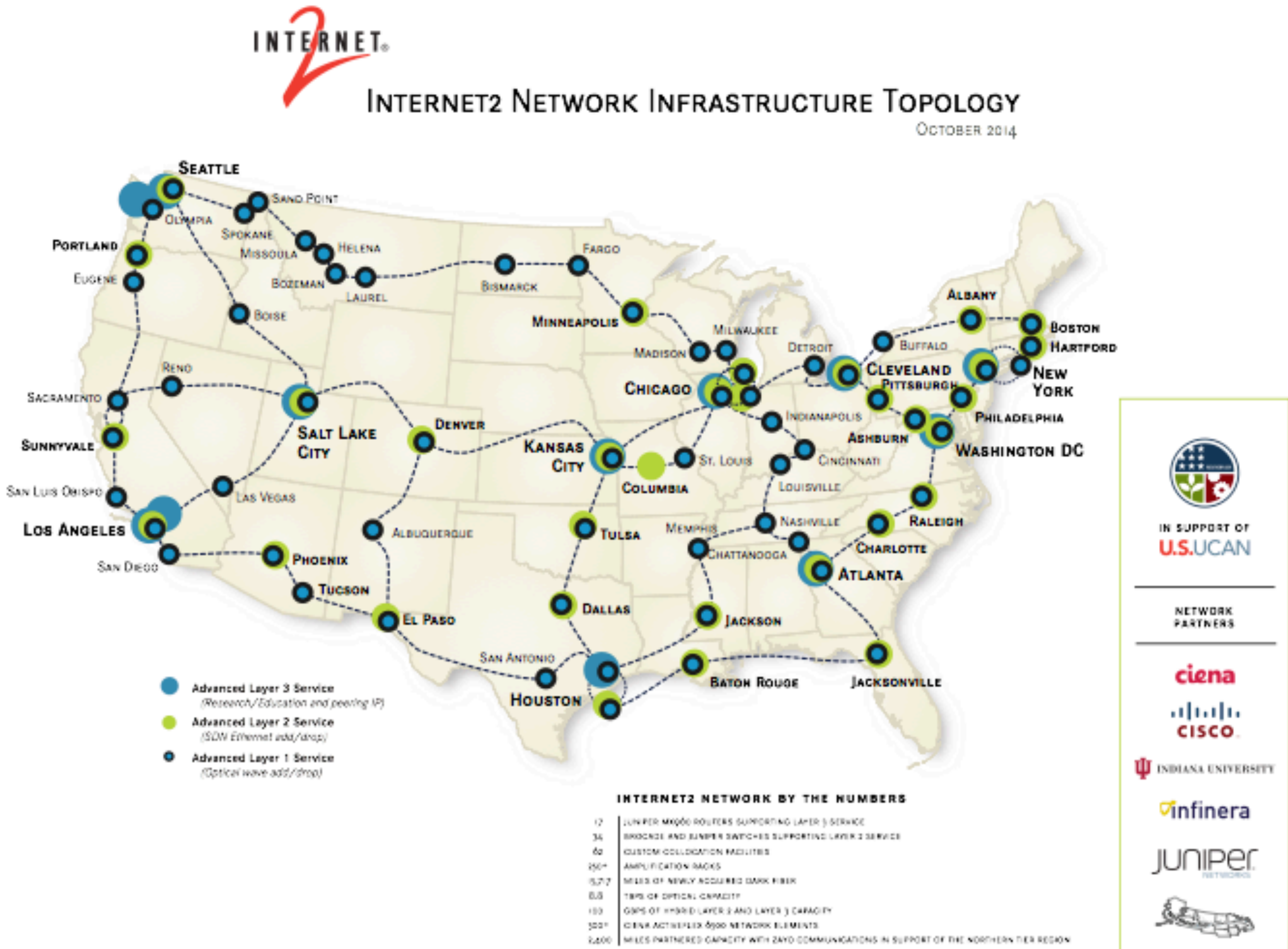
Campus access, e.g.,
- Ethernet, Wireless

- The Internet is a network of networks
- Each individually administrated network is called an Autonomous System (AS)

Yale Internet Connection

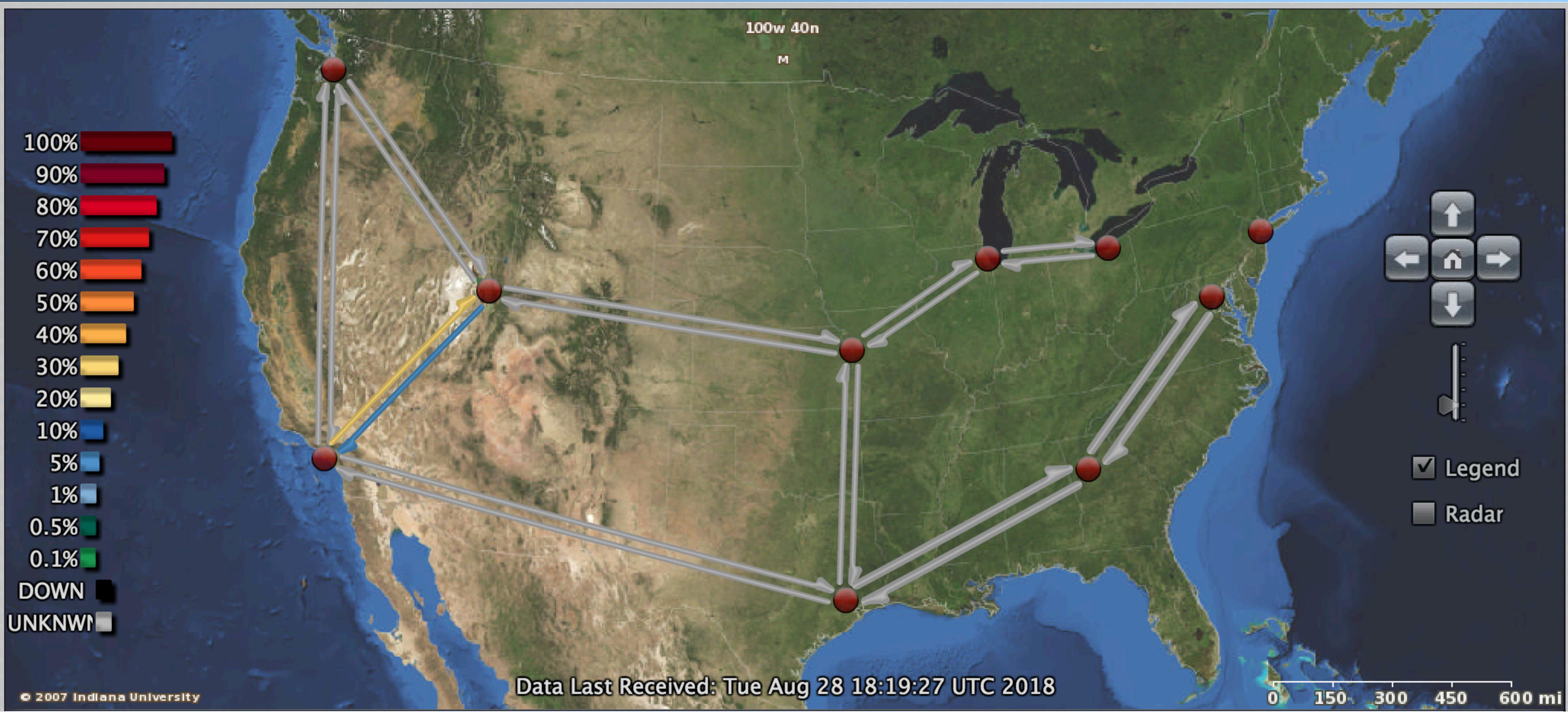
```
cicada:~% traceroute www.tsinghua.edu.cn
 1  college.net.yale.internal (172.28.201.65)  1.440 ms  1.227 ms  1.453 ms
 2  10.1.1.13 (10.1.1.13)  1.359 ms  1.153 ms  1.173 ms
 3  level3-10g-asr.net.yale.internal (10.1.4.40)  2.786 ms  6.110 ms  2.547 ms
 4  cen-10g-yale.net.yale.internal (10.1.3.102)  2.646 ms  3.242 ms  2.576 ms
 5  * * *
 6  enrt064hhh-9k-te0-3-0-5.net.cen.ct.gov (67.218.83.254)  5.169 ms  3.797 ms  6.891 ms
 7  198.71.46.215 (198.71.46.215)  3.615 ms  3.742 ms  3.931 ms
 8  et-10-0-0.1180.rtsw.newy32aoa.net.internet2.edu (198.71.46.214)  6.661 ms  6.532 ms  6.310 ms
 9  et-4-0-0.4079.sdn-sw.phil.net.internet2.edu (162.252.70.103)  8.658 ms  8.714 ms  8.666 ms
10  et-1-1-0.4079.rtsw.wash.net.internet2.edu (162.252.70.119)  11.787 ms  30.111 ms  11.900 ms
11  et-8-1-0.4079.sdn-sw.ashb.net.internet2.edu (162.252.70.62)  12.428 ms  16.654 ms  15.862 ms
12  et-7-1-0.4079.rtsw.chic.net.internet2.edu (162.252.70.61)  28.898 ms  28.999 ms  28.908 ms
13  et-3-1-0.4070.rtsw.kans.net.internet2.edu (198.71.47.207)  40.084 ms  39.958 ms  39.695 ms
14  et-8-0-0.4079.sdn-sw.denv.net.internet2.edu (162.252.70.10)  50.195 ms  50.562 ms  50.258 ms
15  et-4-1-0.4079.rtsw.salt.net.internet2.edu (162.252.70.9)  59.707 ms  60.261 ms  59.762 ms
16  et-7-0-0.4079.sdn-sw.lasv.net.internet2.edu (162.252.70.30)  67.555 ms  67.539 ms  67.312 ms
17  et-4-1-0.4079.rtsw.losa.net.internet2.edu (162.252.70.29)  72.419 ms  72.428 ms  72.376 ms
...
```

Internet2



http://noc.net.internet2.edu/i2network/maps-documentation/maps.html#Internet2 Combined Infrastructure Network Map

Internet2



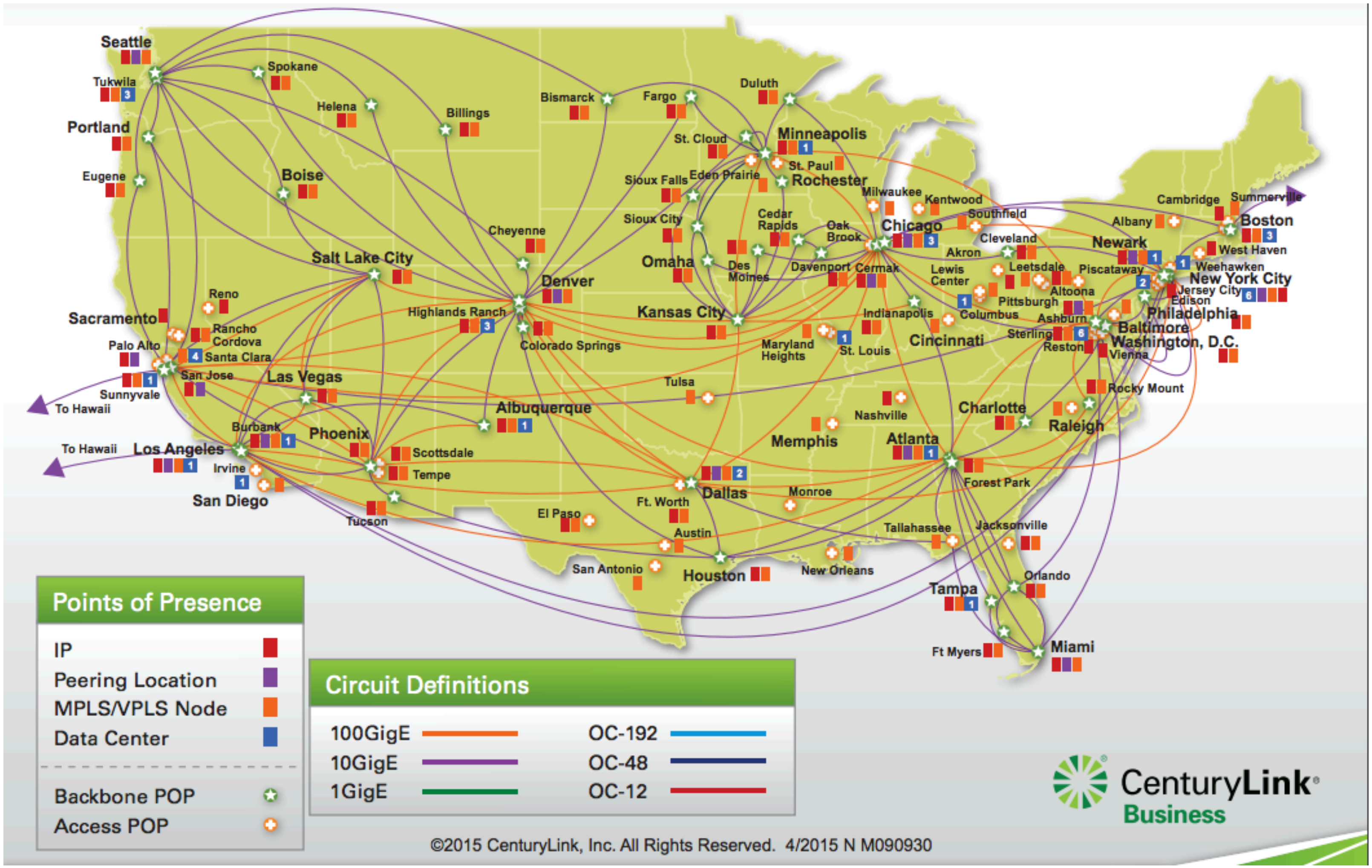
http://atlas.grnoc.iu.edu/atlas.cgi?map_name=Internet2%20IP%20Layer

Yale Internet Connection

Try traceroute from Yale to

- www.microsoft.com
- www.google.com
- www.amazon.com

Qwest (CentryLink) Network Maps



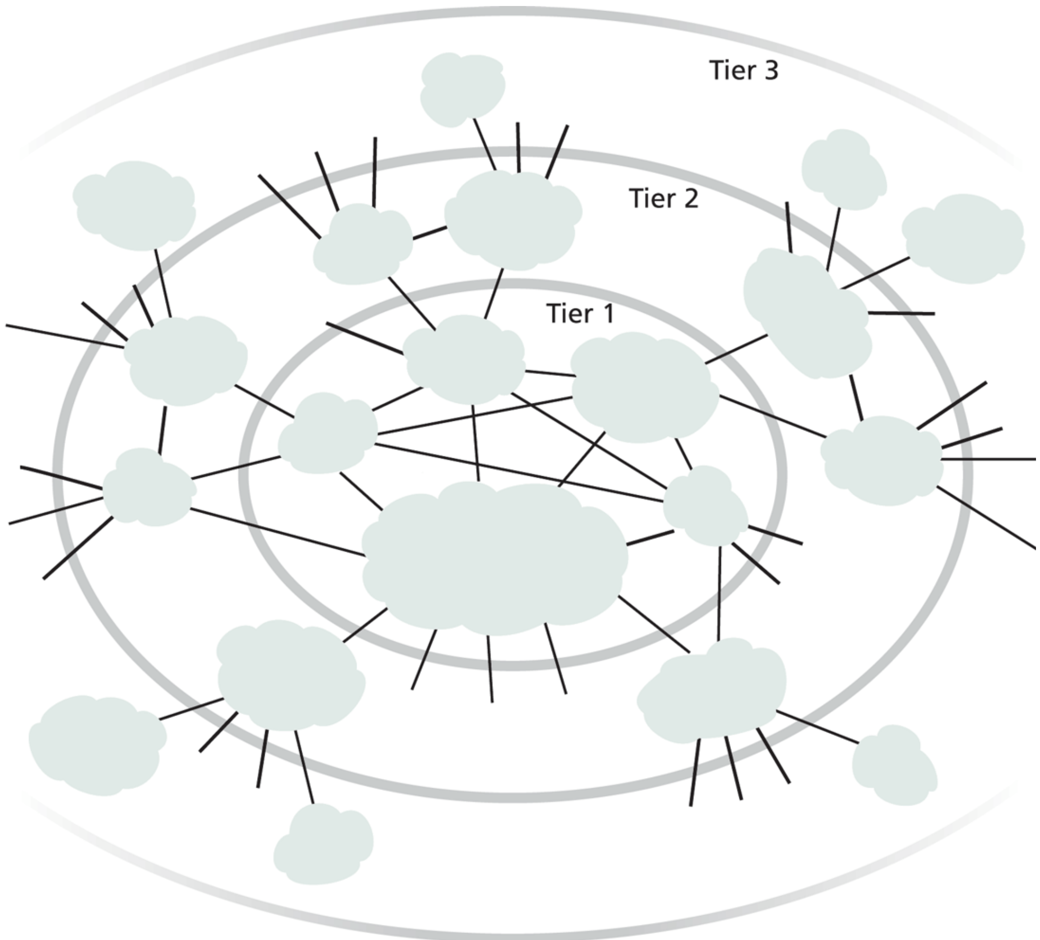
http://www.centurylink.com/business/asset/network-map/ip-mpls-network-nm090930.pdf

Qwest Backbone Map

http://www.centurylink.com/business/resource-center/network-maps/

Internet Infrastructure Connectivity

- Roughly hierarchical
  - Divided into tiers
  - Tier-1 ISPs are also called backbone providers, e.g., AT&T, Verizon, Sprint, Level 3, Qwest
- An ISP runs (private) Points of Presence (PoP) where its customers and other ISPs connect to it
- ISPs also connect at (public) Internet Exchange Point (IXP)
  - public peering



http://en.wikipedia.org/wiki/List_of_Internet_exchange_points_by_size

Outline

- Administrative trivia's
- What is a network protocol?
- *A brief introduction to the Internet*
  - past
  - *Present*
    - *topology*
    - *traffic*

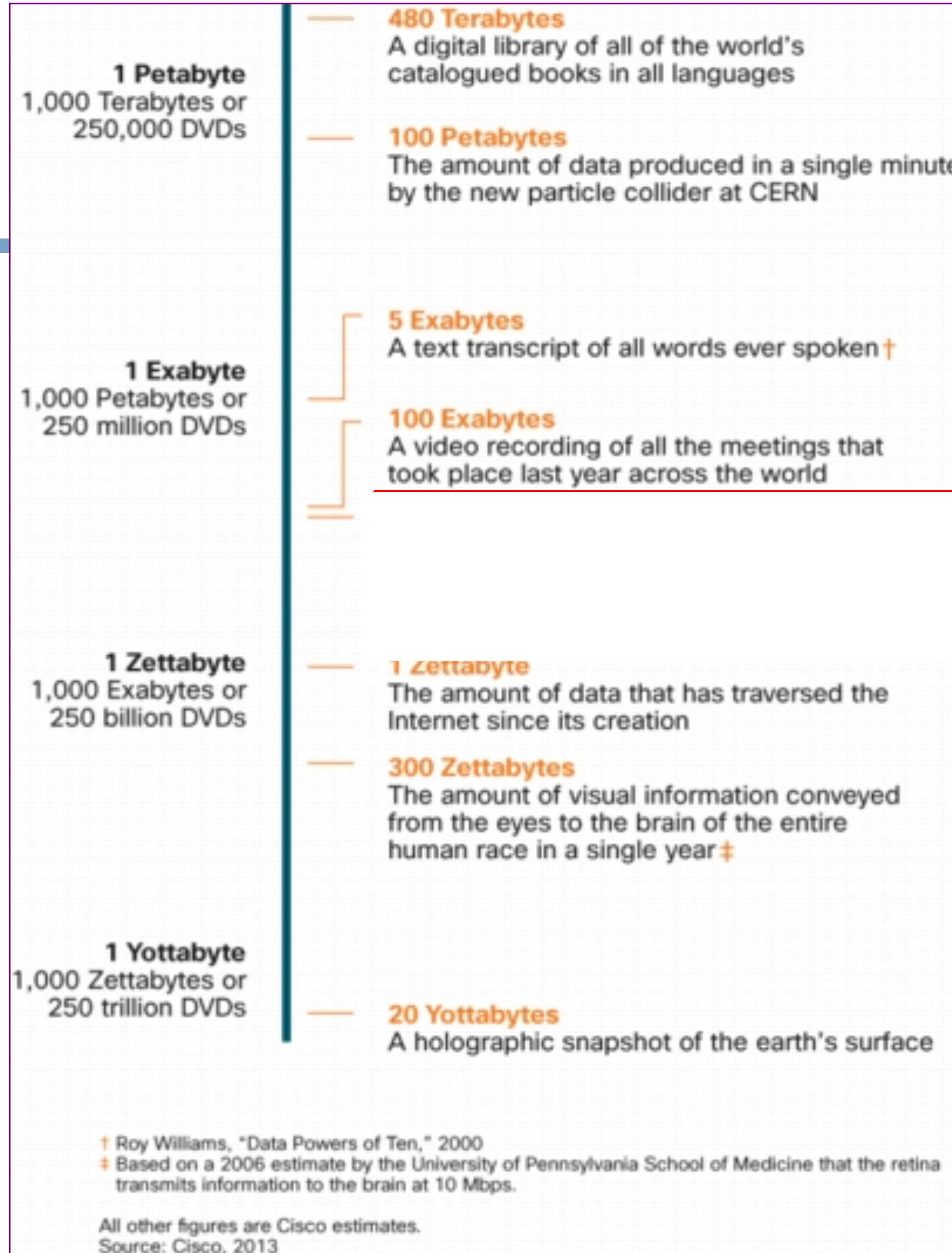
Internet (Consumer) Traffic

Consumer Web, Email, and Data Traffic, 2016–2021	2016	2017	2018	2019	2020	2021	CAGR 2016–2021
By Network (PB per Month)							
Fixed web and data	6,795	7,467	8,569	9,610	10,706	11,337	11%
Mobile web and data	2,263	3,214	4,295	5,509	6,796	8,201	29%
By Geography (PB per Month)							
Asia Pacific	3,393	4,102	5,072	6,160	7,398	8,453	20%
North America	2,578	2,863	3,149	3,410	3,631	3,792	8%
Central and Eastern Europe	1,302	1,404	1,598	1,790	1,994	2,095	10%
Western Europe	693	901	1,177	1,450	1,692	1,882	22%
Middle East and Africa	469	732	1,038	1,358	1,728	2,189	36%
Latin America	624	680	831	953	1,059	1,128	13%
Total (PB per Month)							
Consumer web, email, and data	9,059	10,681	12,864	15,120	17,502	19,538	17%

Source: Cisco VNI,

Internet Traffic in Perspective

640K ought to be enough for anybody.



Outline

- Administrative trivia's
- What is a network protocol?
- A brief introduction to the Internet: past and present
- *Challenges of network infrastructure and application design*

Scale

"Developers who have worked at the small scale might be asking themselves why we need to bother when we could just use some kind of out-of the-box solution. For small-scale applications, this can be a great idea. We save time and money up front and get a working and serviceable application. The problem comes at larger scales—there are no off-the-shelf kits that will allow you to build something like Amazon... There's a good reason why the largest applications on the Internet are all bespoke creations: no other approach can create massively scalable applications within a reasonable budget."

http://www.evontech.com/symbian/55.html

Largest Internet Sites in U.S.

RANK	SITE	MONTHLY PEOPLE
1	google.com	296,034,944
2	youtube.com	260,789,632
3	facebook.com	258,765,232
4	amazon.com	255,453,120
5	wikipedia.org	202,303,936
6	twitter.com	178,084,096
7	yahoo.com	153,547,408
8	nytimes.com	127,581,920
9	ebay.com	120,911,472
10	reddit.com	108,105,512
11	yelp.com	88,001,608
12	walmart.com	61,330,156
13	paypal.com	56,964,048
14	apple.com	53,835,600
15	bing.com	52,061,444

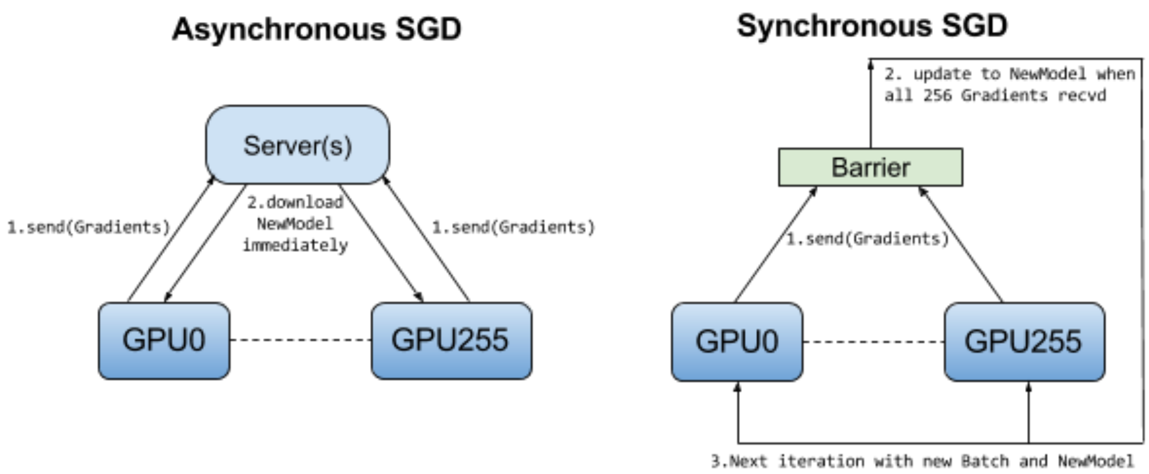
https://www.quantcast.com/top-sites

Complexity from Robustness

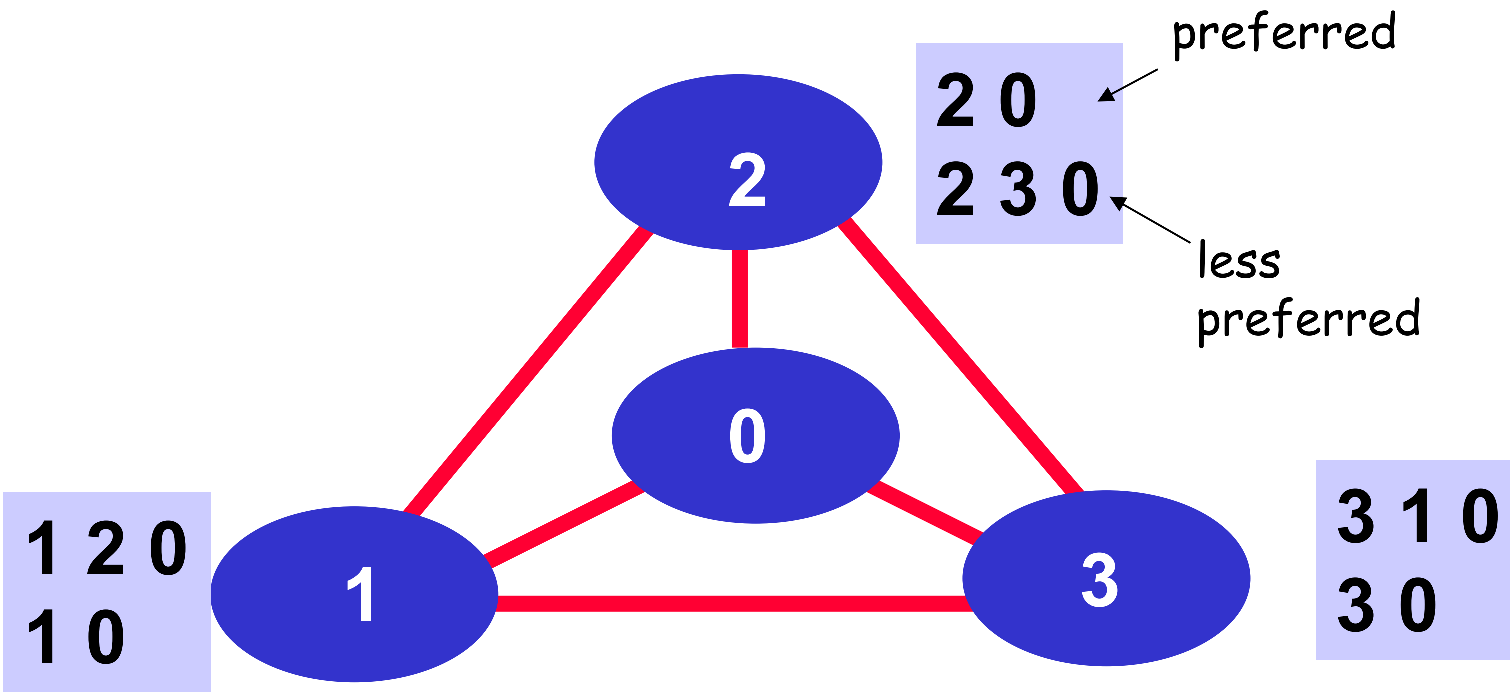
- Complexity in highly organized systems arises primarily from design strategies intended to create *robustness* to uncertainty in their environments and component parts.
  - Scalability
    - robustness to changes to the size and complexity of a system as a whole.
  - Evolvability
    - robustness of lineages to large changes on various (usually long) time scales.
  - Reliability
    - robustness to component failures.
  - Modularity
    - robustness to component rearrangements.
  - Asynchrony
    - robustness to uncertainty of performance.
  - Decentralization
    - robustness to single-point of failure.

Synchronous vs Asynchronous Design

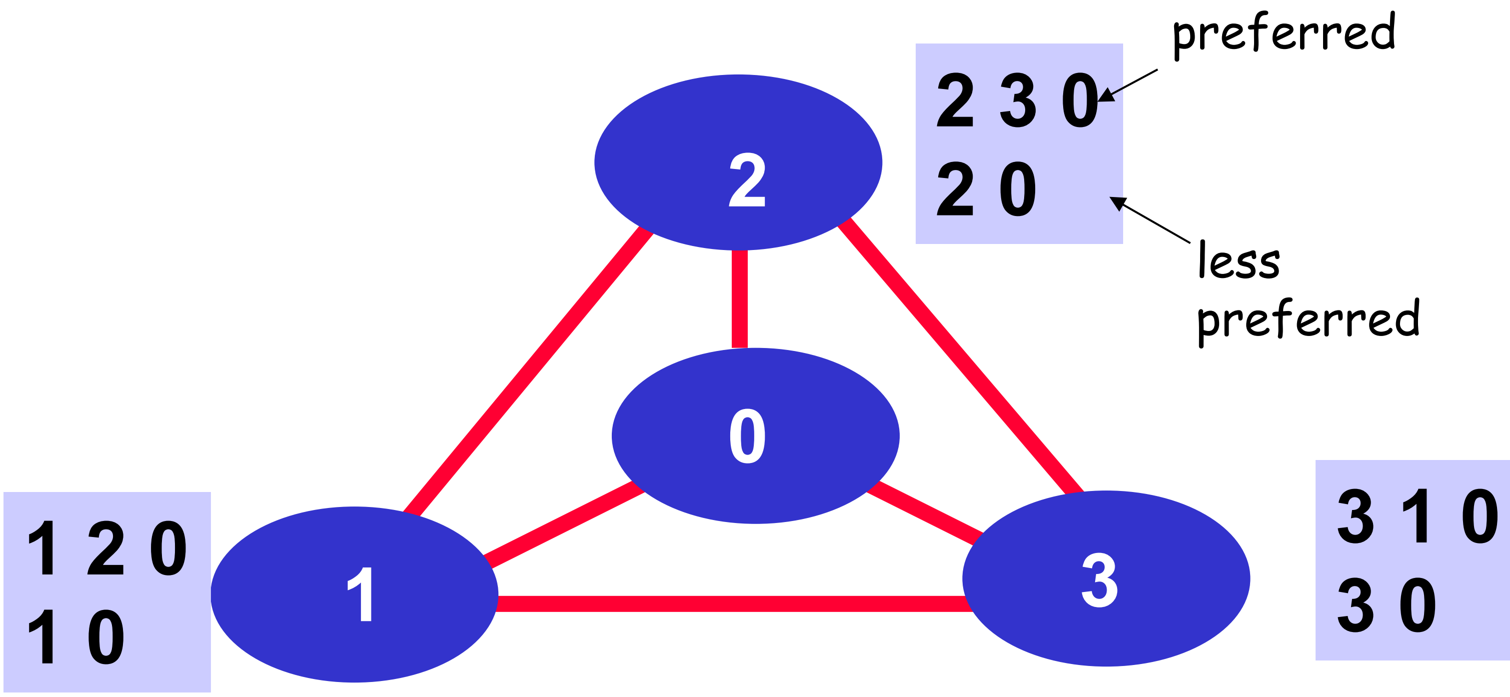
- Synchronous vs asynchronous network
  - Synchronous Bellman-Ford vs Asynchronous Bellman-Ford
- Synchronous server vs asynchronous server
  - Synchrous HTTP server vs Asynchronous HTTP server
- Synchronous app vs asynchronous app



Centralized vs Decentralized Design



Centralized vs Decentralized Design



Not a toy, but a major issue in interdomain routing.

Centralized vs Decentralized /Distributed Design

- Fundamentally, a networked system is a distributed, resource sharing system
  - A set of flows F
  - If x_f is the rate of flow f, assume the utility to flow f is $U_f(x_f)$
  - Maximize aggregate utility, subject to resource constraints

$$
\begin{array}{ll}
\text{max} & \sum_{f \in F} U_f\left(x_f\right) \\
\text{subject to} & \sum_{f : f \text{ uses link } l} x_f \le c_l \text{ for any link } l \\
\text{over} & x \ge 0
\end{array}
$$

Centralized vs Decentralized/Distributed Systems

- Key challenge across the course
  - Distributed design is preferred if possible
    - but distributed computing is hard, e.g., FLP Impossibility Theorem
  - Decentralized design is preferred if possible
    - but decentralized design is hard e.g., Arrow's Impossibility Theorem

What Will We Cover?

- A tentative schedule is posted at class schedule page
- Network architecture and design principles
  - Layered network arch; e2e principle

- Application architecture and design principles
  - application paradigms; high performance network app; queueing analysis
  - HTTP/Web, Email, DNS, P2P, Blockchain, Content distribution; Distributed data analytics; Distributed learning

- Transport
  - transport services
  - reliability; distributed resource allocation; primal-dual
  - transport protocols: TCP/UDP

What Will We Cover?

- Network and link layers
  - distributed, asynchronous, autonomous routing algorithms; scalable router design; IP/IPv6; mobile IP; cellular networks
  - multiple access;
  - cloud and data center design

- Network security
  - security primitives; BAN logic, SSL

Summary

- Course administration
- A protocol defines the format and the order of messages exchanged between two or more communicating entities, as well as the actions taken on the transmission and/or receipt of a message or other events.
- The past:
  - facts:
    - The Internet started as ARPANET in late 1960s
    - The initial link bandwidth was 50 kbps
    - The number of hosts at the end of 1969 was 4
  - some implications of the past:
    - ARPANET is sponsored by ARPA → design should survive failures
    - Initial IMPs made by simple company → keep the network simple
    - Many networks → need a network to connect networks
- Current:
  - The number of hosts connected to the Internet is around 1 billions
  - The backbone speed of the current Internet is about 40/100 Gbps
  - The Internet is roughly hierarchical where ISPs interconnect at PoP and IXP
  - Needs to handle scale, complexity, decentralization, security

Preview

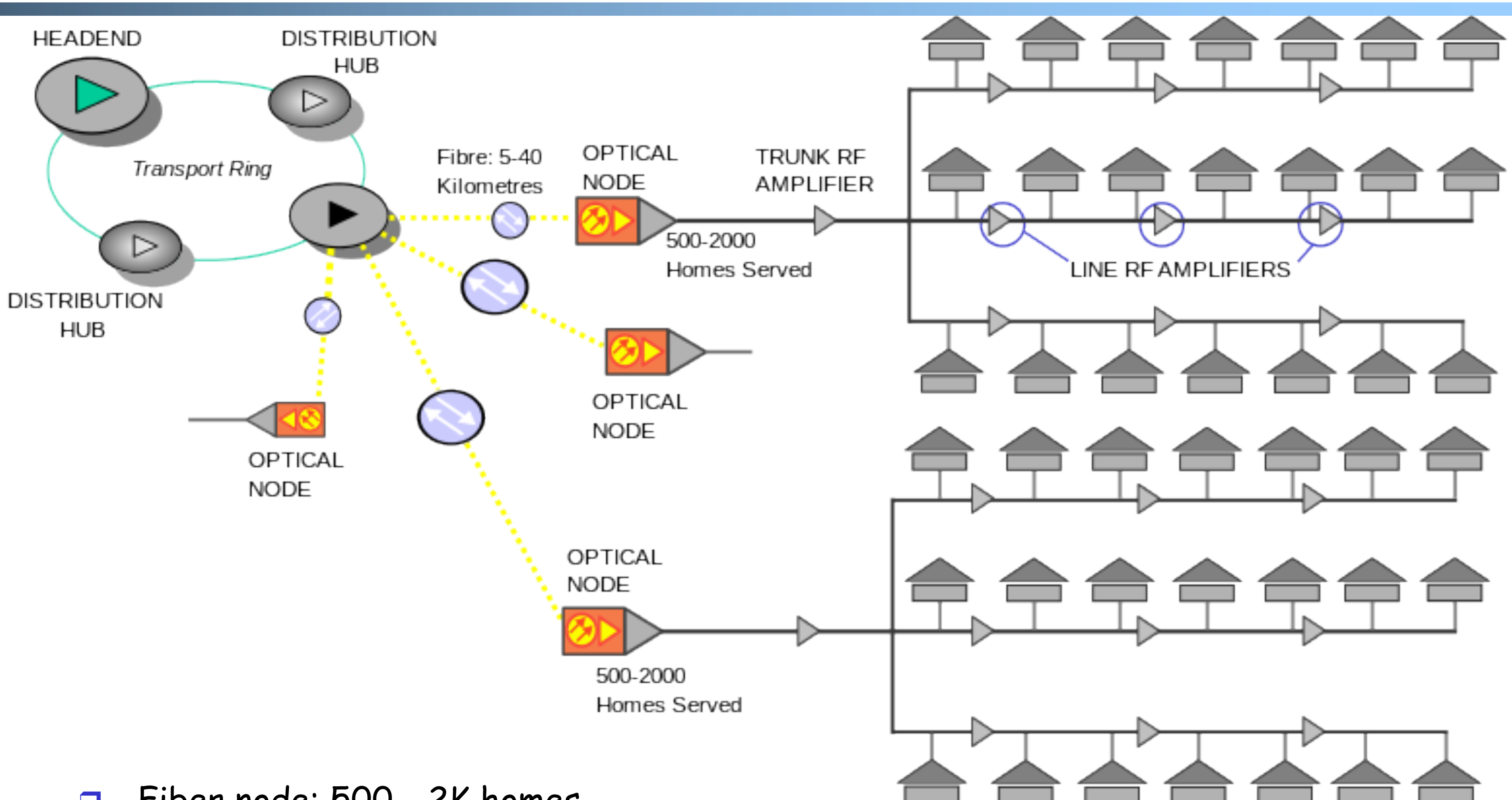
- We have only looked at the topology/connectivity of the Internet
  - a communication network is a mesh of interconnected devices

- *A fundamental question:* how is data transferred through a network?

Backup Slides

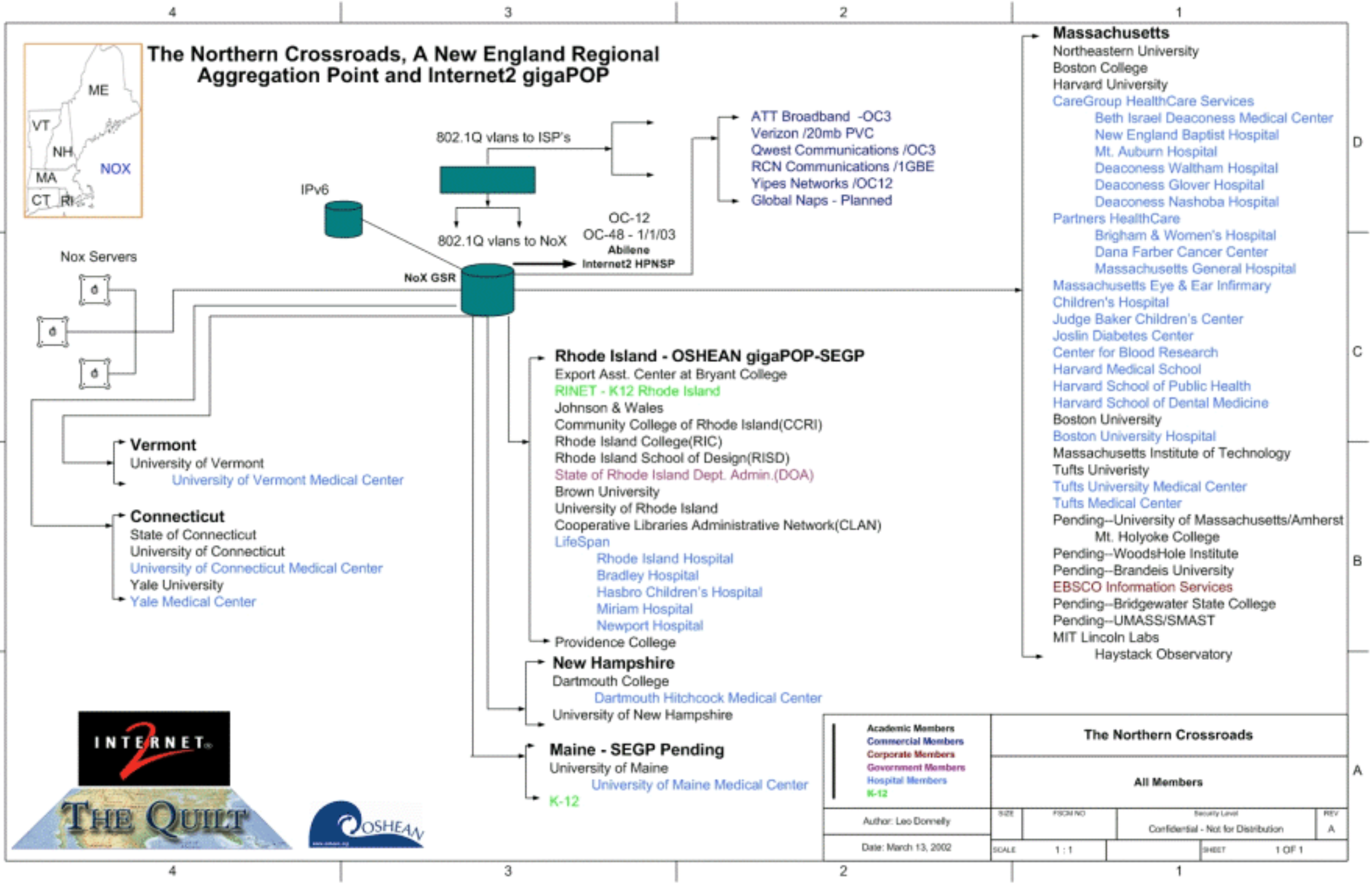
Access: Cable

Also called Hybrid Fiber-coaxial Cable (HFC)



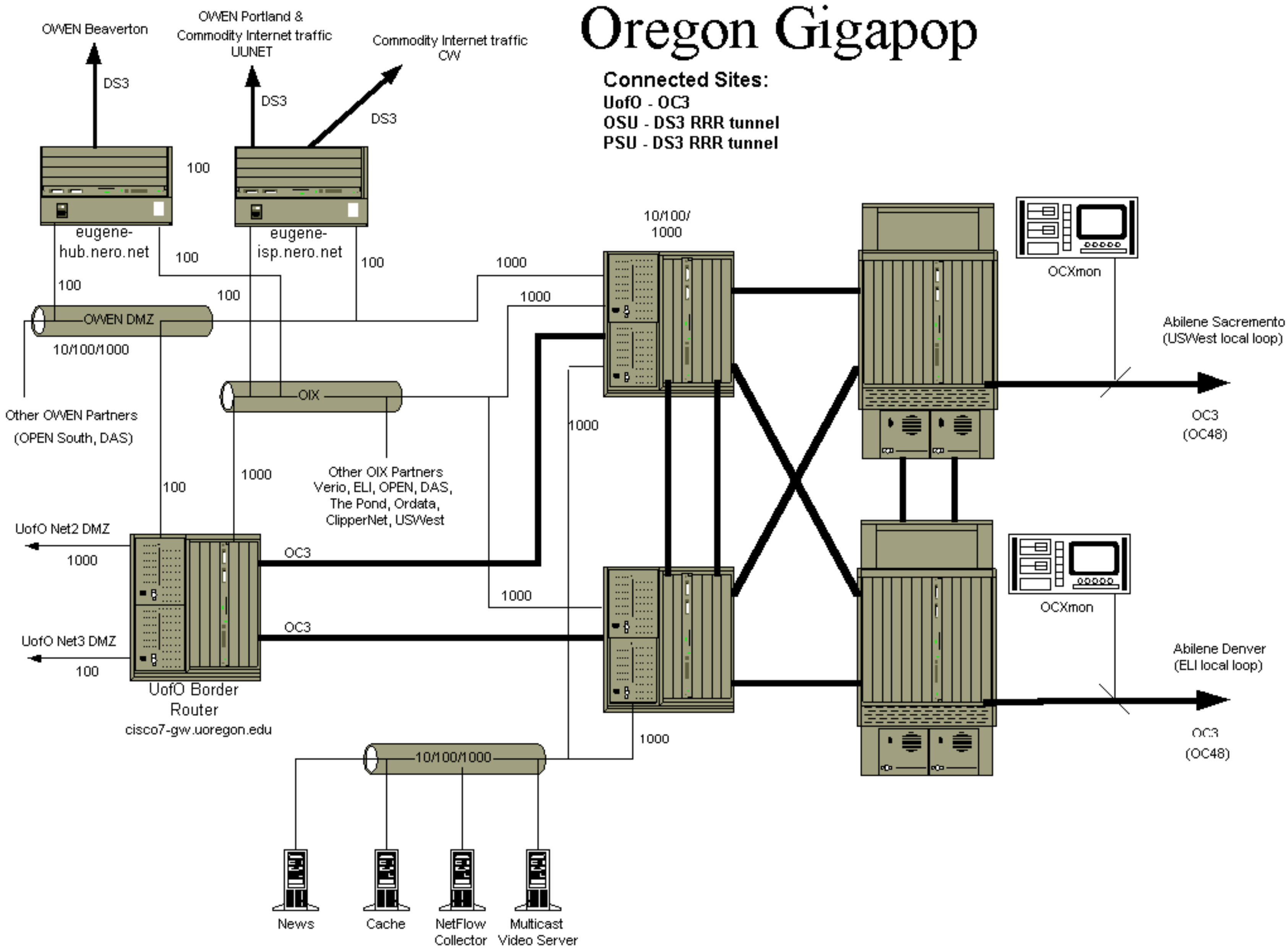
- Fiber node: 500 - 2K homes
- Distribution hub: 20K - 40 K homes
- Regional headend: 200 K - 400 K homes

Northern CrossRoads (NoX) Aggregation Point (AP)



http://www.uis.harvard.edu/emerging_technologies/Northern_Crossroads_Map.gif

Oregon Gigapop



http://www.oregon-gigapop.net/images/OregonGigapop2.gif