COMPUTER NETWORKS

Distributed Computing Group

Roger Wattenhofer

Summer 2007

Chapter 1 INTRODUCTION

Distributed Computing Group

Computer Networks

“Summer” 2007

Overview

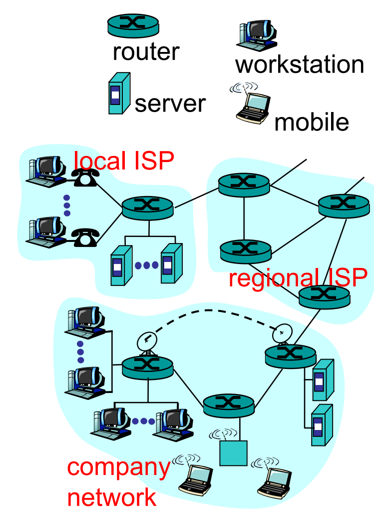
- What’s the Internet?
- What’s a protocol?
- Network edge vs. core
- Access net, physical media
- Performance: loss, delay
- Protocol layers, service models
- Backbones, NAPs, ISPs
- History & Future

- Goal: get context, overview, “feeling” of networking, postpone details.

What’s the Internet: “nuts and bolts” view

- Millions of connected computing devices: Hosts, End-Systems
  - PC’s, workstations, servers
  - PDA’s, phones, toasters

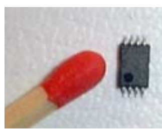
  running network applications
- Communication links
  - fiber, copper, radio
- Routers
  - forward packets (chunks) of data through network



“Cool” Internet appliances

IP picture frame
[www.ceiva.com]

Web-enabled toaster and weather forecaster



World’s smallest web server

Streaming
(Video, Audio, VoIP)

Plus Web “2.0” Stuff

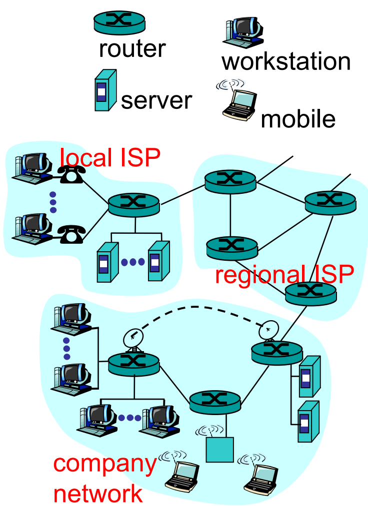
- e.g. “I kind of look like…”

=

- Plus lots of social networking sites
  - myspace, flickr, last.fm, secondlife, etc.

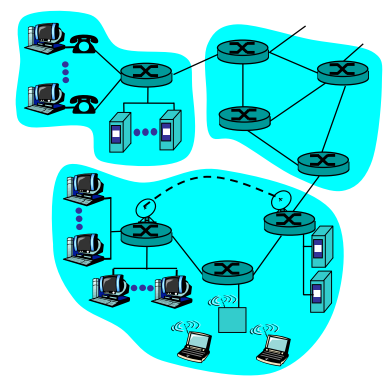
What’s the Internet: “nuts and bolts” view

- protocols: control sending, receiving of messages
  - TCP, IP, HTTP, FTP, PPP
- Internet: “network of networks”
  - loosely hierarchical
  - public Internet versus private Intranet
- Internet standards
  - RFC: Request for comments
  - IETF: Internet Engineering Task Force



What’s the Internet: a service view

- communication infrastructure enables distributed applications
  - WWW, email, games, e-commerce, databases, voting, file sharing
- communication services provided
  - connectionless
  - connection-oriented



- cyberspace [Gibson]:
  “a consensual hallucination experienced daily by billions of operators, in every nation, ..."

What's a protocol?

Human protocols
- "what's the time?"
- "I have a question"
- introductions

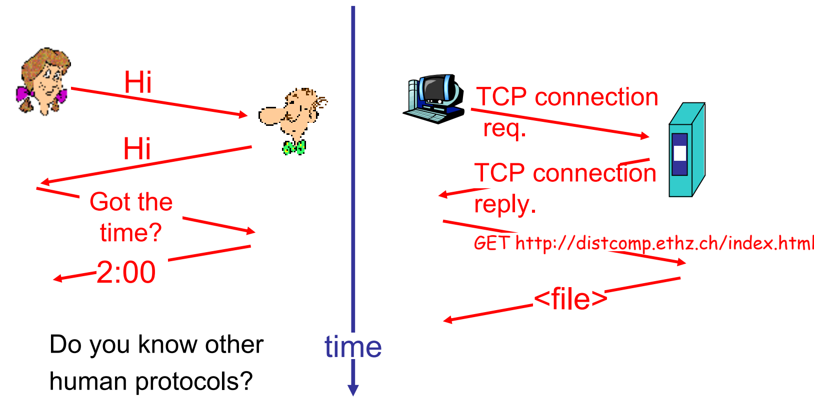
... specific msgs sent
... specific actions taken when msgs received, or other events

Network protocols
- machines rather than humans
- all communication activity in Internet governed by protocols

> *protocols define format, order of msgs sent and received among network entities, and actions taken on msg transmission, receipt*

What's a protocol?

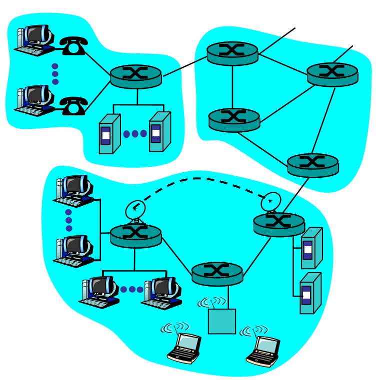
a human protocol and a computer network protocol



Do you know other human protocols?

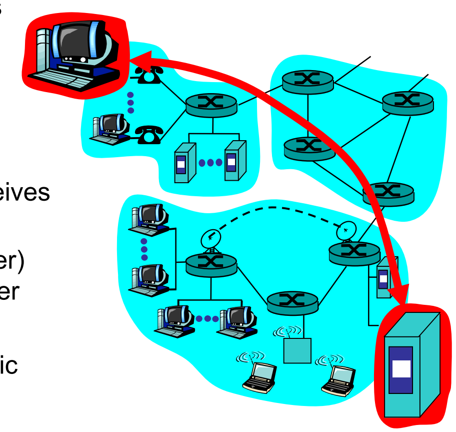
A closer look at network structure

- network edge
  - hosts and applications
- network core
  - routers
  - network of networks
- access networks, physical media
  - communication links



The network edge

- end systems (hosts)
  - run application programs
  - e.g. WWW, email
  - at "edge of network"
- client/server model
  - client host requests, receives service from server
  - e.g. WWW client (browser) /server; email client/server
- peer-to-peer model
  - host interaction symmetric
  - e.g. BitTorrent, Skype



Network edge: connection-oriented service

Goal: data transfer between end systems
- handshaking: setup (prepare for) data transfer ahead of time
  - "Hello, hello back" human protocol
  - set up "state" in two communicating hosts
- TCP
  - Transmission Control Protocol
  - connection-oriented service of the Internet

TCP [RFC 793]
- reliable, in-order byte-stream data transfer
  - loss: acknowledgements and retransmissions
- flow control
  - sender won't overwhelm receiver
- congestion control
  - senders "slow down sending rate" when network congested

Network edge: connectionless service

Goal: data transfer between end systems
- same as before!

- UDP - User Datagram Protocol [RFC 768]
  - Internet's connectionless service
  - unreliable data transfer
  - no flow control
  - no congestion control

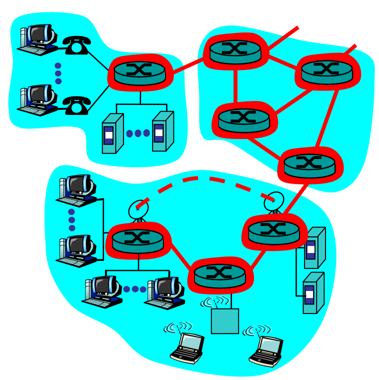
App's using TCP
- HTTP (WWW)
- FTP (file transfer)
- Telnet (remote login)
- SMTP (email)

App's using UDP
- streaming media
- teleconferencing
- Internet telephony

The network core

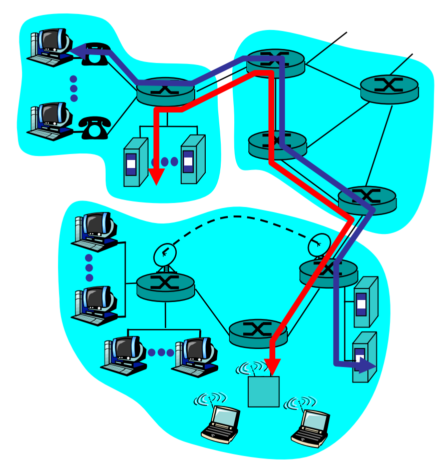
- "graph" of interconnected routers
- the fundamental question: how is data transferred through net?

- Circuit switching
  - dedicated circuit per call
  - telephone network
- Packet switching
  - data sent through network in discrete "chunks"



Circuit Switching

- End-end resources reserved for "call"
- Divide link bandwidth into "pieces"
  - Frequency division
  - Time division
- dedicated resources no sharing; "piece" is idle if not used by user
- circuit-like (guaranteed) performance
- call setup required

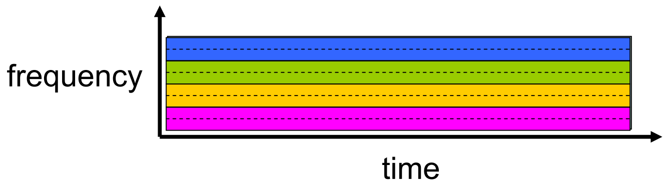


Frequency Division and Time Division Multiple Access

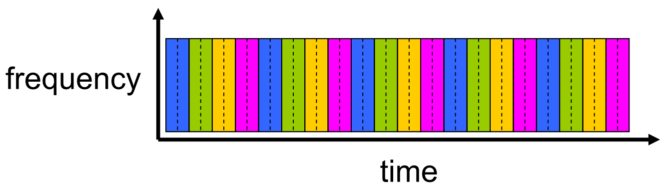
Example:

4 users

FDMA

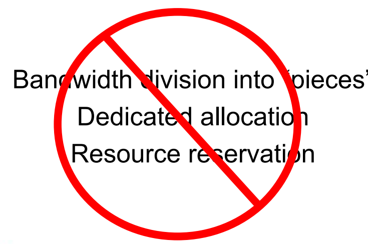


TDMA



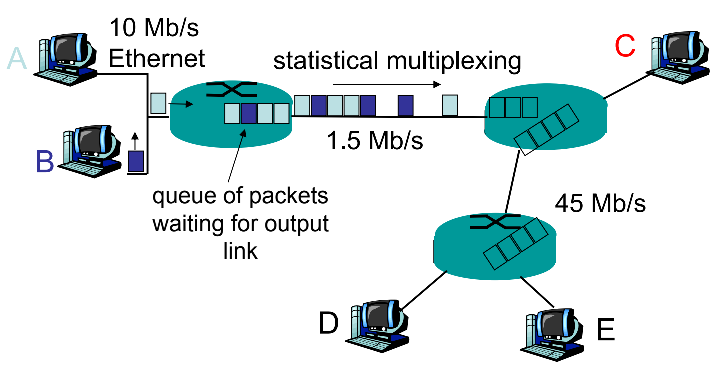
Packet Switching

- each end-end data stream divided into packets
- packets share network resources
- each packet uses full link bandwidth
- resources used as needed



- resource contention
  - aggregate resource demand can exceed amount available
- congestion
  - packets queue
  - wait for link use
- store-and-forward
  - packets move one hop at a time
  - router receives whole packet before sending the first bit over the next link

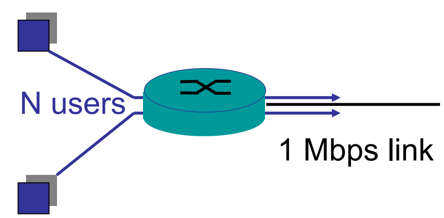
Packet Switching



- Real-world example for packet switching: Cafeteria (ETH Mensa)

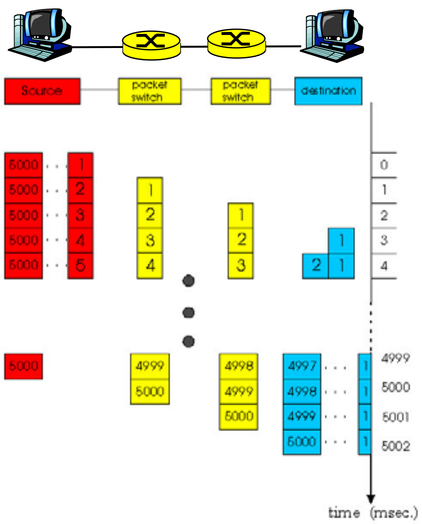
Circuit switching vs. Packet switching

- 1 Mbit link
- each user
  - 100Kbps when "active"
  - active 10% of time



- circuit-switching
  - 10 users
- packet switching:
  - with 50 users, Pr[more than 10 users active] < 1%
  - with 100 users, Pr[more than 10 users active] $\approx$ 42%
- Packet switching allows more users… Really?

Packet Switching



- Source breaks message into smaller chunks: “packets”
- Store-and-forward: switch waits until one chunk has completely arrived, then forwards/routes
- What if message was sent as single unit?

Circuit switching vs. Packet switching

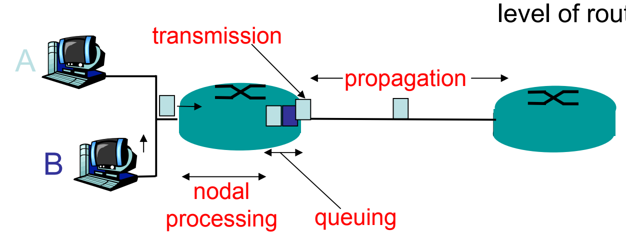
- Is packet switching a “slam dunk winner“?
- Great for bursty data
  - resource sharing
  - no call setup
- But: Excessive congestion: packet delay and loss
  - protocols needed for reliable data transfer
  - header overhead
  - congestion control
- How to provide circuit-like behavior?
  - bandwidth guarantees needed for audio/video apps
  - still an unsolved problem

Packet-switched networks: Routing

- Goal: move packets among routers from source to destination
- We later study several path selection algorithms
- datagram network
  - destination address determines next hop
  - routes may change during session
  - analogy: driving, asking directions
- virtual circuit network
  - each packet carries tag (virtual circuit ID)
  - tag determines next hop
  - fixed path determined at call setup time, remains fixed
  - routers maintain per-call state

Delay in packet-switched networks

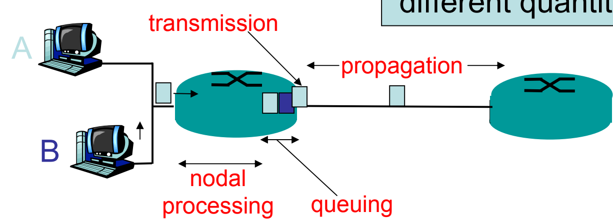
- packets experience delay on end-to-end path
- four sources of delay at each hop
- Nodal processing
  - check bit errors
  - determine output link
- Queuing
  - time waiting at output link for transmission
  - depends on congestion level of router



Delay in packet-switched networks

- Transmission delay:
  - R=link bandwidth (bps)
  - L=packet length (bits)
  - time to send bits into link = L/R
- Propagation delay:
  - d = length of physical link
  - s = propagation speed in medium (~2x10^8 m/sec)
  - propagation delay = d/s

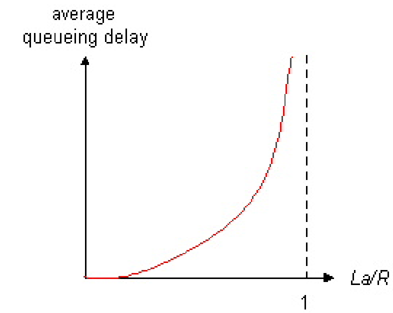
Note: s and R are different quantities!



Queuing delay

- R=link bandwidth (bps)
- L=packet length (bits)
- a=average packet arrival rate (packets per second)

- Arrival rate λ = La (bps)
- Service rate μ = R (bps)
- Traffic intensity $\rho = \lambda / \mu$



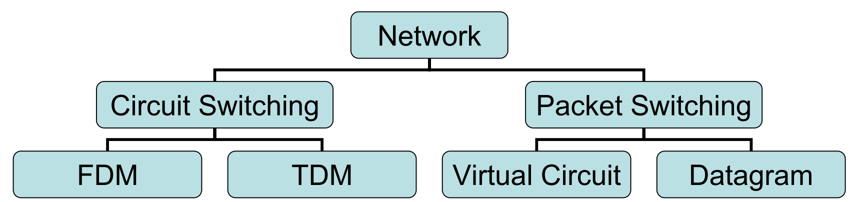
- ρ small: average queuing delay small
- $\rho \to 1$: delays become large
- $\rho \geq 1$: more "work" arriving than can be serviced, average delay grows infinitely!

"Real" Internet delays and routes: traceroute

Tracing route from photek.ethz.ch [129.132.13.122] to google.com [216.239.35.100]:

1	<10 ms	<10 ms	<10 ms	rou-ifw-1-inf-vs.ethz.ch [129.132.13.65]
2	<10 ms	<10 ms	<10 ms	rou-gw-switch-1-mega-transit-2.ethz.ch [129.132.99.213]
3	<10 ms	<10 ms	<10 ms	swiez2.ethz.ch [192.33.92.11]
4	<10 ms	<10 ms	<10 ms	swiIX1-G2-3.switch.ch [130.59.36.250]
5	<10 ms	<10 ms	<10 ms	zch-b1-geth4-1.telia.net [213.248.79.189]
6	<10 ms	10 ms	<10 ms	ffm-b1-pos5-3.telia.net [213.248.77.133]
7	10 ms	20 ms	20 ms	213.248.68.90
8	10 ms	20 ms	20 ms	de-cix.fra.above.net [80.81.192.226]
9	<10 ms	10 ms	<10 ms	so-0-1-0.cr1.fra1.de.mfnx.net [216.200.116.213]
10	10 ms	20 ms	10 ms	pos9-0.cr1.cdg2.fr.mfnx.net [64.125.31.161]
11	40 ms	41 ms	50 ms	so-5-0-0.cr1.lhr3.uk.mfnx.net [64.125.31.154]
12	100 ms	100 ms	100 ms	so-7-0-0.cr1.dca2.us.mfnx.net [64.125.31.186]
13	170 ms	180 ms	170 ms	so-3-0-0.mpr3.sjc2.us.mfnx.net [208.184.233.133]
14	170 ms	180 ms	180 ms	so-0-0-0.mpr4.sjc2.us.mfnx.net [64.125.30.2]
15	170 ms	180 ms	180 ms	so-1-0-0.cr2.sjc3.us.mfnx.net [208.184.233.50]
16	170 ms	180 ms	170 ms	pos1-0.er2a.sjc3.us.mfnx.net [208.185.175.198]
17	160 ms	150 ms	160 ms	sjni1-2-3.net.google.com [216.239.48.238]
18	170 ms	170 ms	160 ms	sjbi1-1-1.net.google.com [216.239.47.162]
19	151 ms	150 ms	160 ms	www.google.com [216.239.35.100]

Networking Taxonomy



- We concentrate on right-hand path (predominant in Internet)

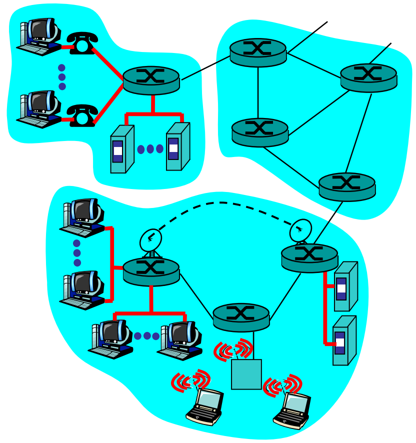
Access networks and physical media

Q: How to connect end systems to edge router?

- residential access nets
- institutional access networks (school, company)
- mobile access networks

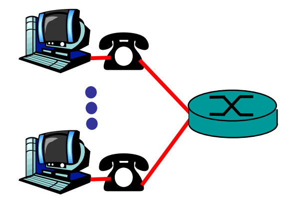
Keep in mind

- bandwidth (bits per second) of access network?
- shared or dedicated?



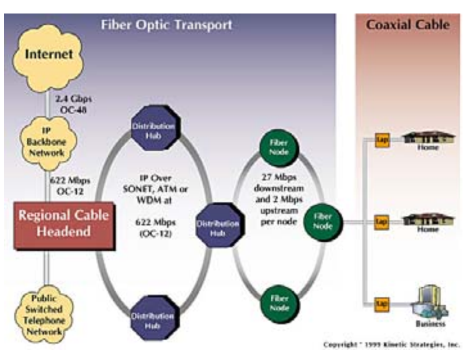
Residential access: point to point access

- Dialup via modem
  - up to 56Kbps direct access to router (conceptually)
- ISDN
  - integrated services digital network
  - 128Kbps all-digital connect to router
- ADSL
  - asymmetric digital subscriber line
  - up to 1 Mbps home-to-router
  - up to 8 Mbps router-to-home
  - ADSL deployment: happening



Residential access: cable modems

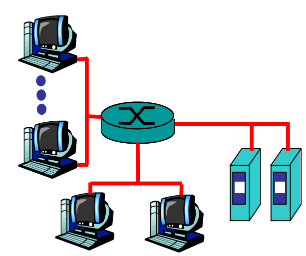
- Other forms of cable modems
  - Power line: e.g. Ascom Powerline
  - TV cable modem: e.g. CableCom, Glattnet
  - Satellite with feedback on phone line
  - Wireless local loop



[http://www.cabledatacomnews.com/cmic/diagram.html]

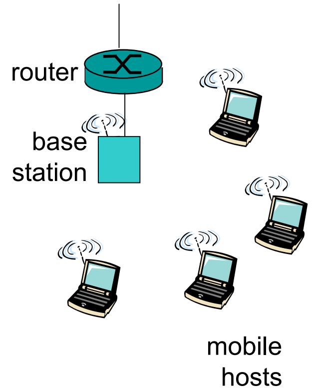
Institutional access: local area networks

- company/university local area network (LAN) connects end system to edge router
- Example: Ethernet
  - shared or dedicated cable connects end systems and router
  - 10 Mbps, 100Mbps, Gigabit Ethernet, etc.
- deployment: institutions, home LANs happening now



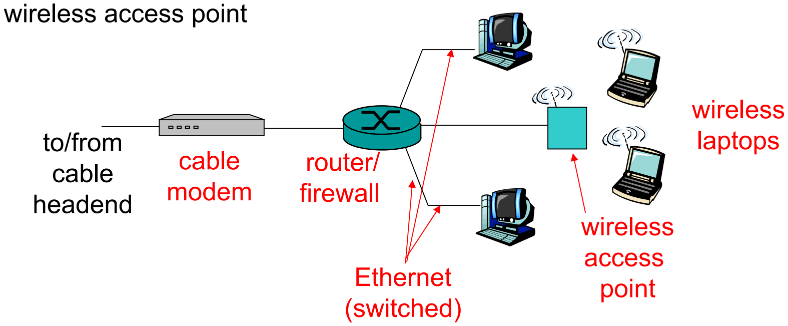
Wireless access networks

- shared *wireless* access network connects end system to router
- wireless LANs
  - radio spectrum replaces wire
  - 802.11b with up to 11 Mbps
  - 802.11a/g with up to 54 Mbps
- wider-area wireless access
  - GSM: wireless access to ISP router via cellular network



Home networks

Typical home network components
- ADSL or cable modem
- router/firewall
- Ethernet
- wireless access point



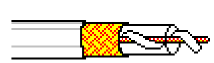
Physical Media

- physical link
  - transmitted data bit propagates across link
- guided media
  - signals propagate in solid media: copper, fiber
- unguided media
  - signals propagate freely, e.g. radio

- Twisted Pair TP (UTP, STP)
  - two insulated copper wires
  - Category 3
    - traditional phone wires
    - 10 Mbps Ethernet
  - Category 5
    - 100Mbps Ethernet
  - Category 6
    - 1Gbps Ethernet

Physical Media: coax, fiber

- Coaxial cable:
  - wire (signal carrier) within a wire (shield)
  - variant baseband ("50Ω")
    - single channel on cable
  - variant broadband ("75Ω")
    - multiple channels on cable
  - bidirectional
  - 10Mbps Ethernet



- Fiber optic cable:
  - glass fiber carrying light pulses
  - high-speed operation: 100Mbps Ethernet
  - high-speed point-to-point transmission (>10 Gbps)
  - low error rate

Physical media: Radio

- signal carried in electromagnetic spectrum
- no physical “wire”
- bidirectional
- propagation environment effects:
  - reflection
  - obstruction by objects
  - interference

- Radio link types:
- microwave
  - e.g. up to 45 Mbps
- wireless LAN (802.11)
  - 2Mbps, 11Mbps, 54Mbps
- wide-area (e.g. cellular)
  - GSM, 10’s Kbps
  - UMTS, Mbps
- satellite
  - up to 50Mbps channel (or multiple smaller channels)
  - GEO: 270 msec end-end delay
  - geosynchronous vs. LEO‘s

Networks are complex!

- many “pieces”
  - hosts
  - routers
  - links of various media
  - applications
  - protocols
  - hardware
  - software

- Questions:
- Is there any hope of organizing the structure of a network?
- Or at least our discussion of networks?

Organization of air travel

ticket (purchase)	ticket (complain)
baggage (check)	baggage (claim)
gates (load)	gates (unload)
runway takeoff	runway landing
airplane routing	airplane routing

airplane routing

Organization of air travel: a different view

ticket (purchase)	ticket (complain)
baggage (check)	baggage (claim)
gates (load)	gates (unload)
runway takeoff	runway landing
airplane routing	airplane routing

airplane routing

- Layers: each layer implements a service
  - via its own internal-layer actions
  - relying on services provided by layer below

Layered air travel: services

- Counter-to-counter delivery of person+bags
- baggage-claim-to-baggage-claim delivery
- people transfer: loading gate to arrival gate
- runway-to-runway delivery of plane
- airplane routing from source to destination

Distributed implementation of layer functionality

Departing airport

- ticket (purchase)
- baggage (check)
- gates (load)
- runway takeoff
- airplane routing

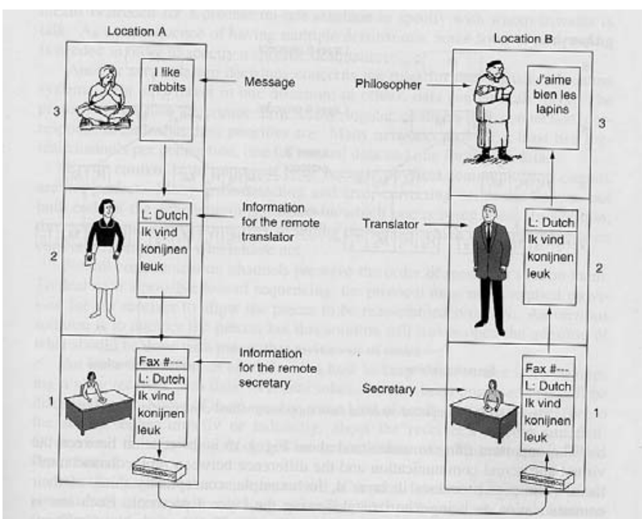
Arriving airport

- ticket (complain)
- baggage (claim)
- gates (unload)
- runway landing
- airplane routing

intermediate air traffic sites

- airplane routing
- airplane routing
- airplane routing

Another example of layering



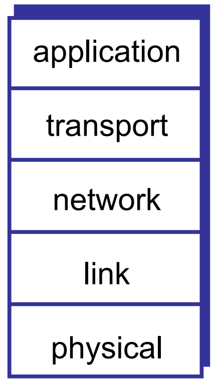
[Tanenbaum]

Why layering?

- Dealing with complex systems
- Explicit structure allows identification, relationship of complex system's pieces
  - layered reference model for discussion
- Modularization eases maintenance, updating of system
  - change of implementation of layer's service transparent to rest of system
  - e.g. change in gate procedure doesn't affect rest of system

Internet protocol stack (TCP/IP reference model)

- application:
  - ftp, SMTP, http
- transport: host-host data transfer
  - TCP, UDP
- network: routing of datagrams from source to destination
  - IP, routing protocols
- link: data transfer between neighboring network elements
  - PPP, Ethernet
- physical: bits “on the wire”

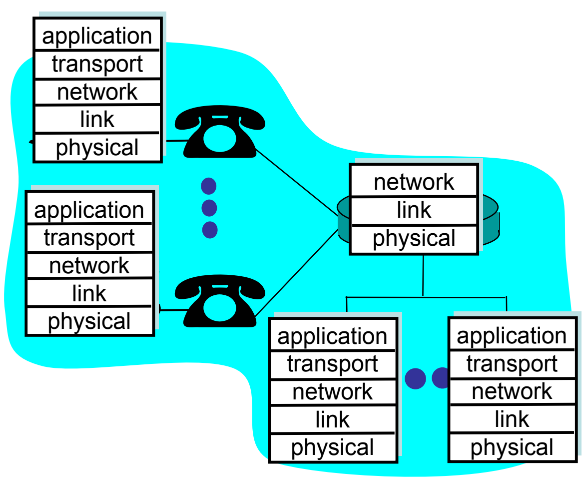


ISO/OSI Reference Model

- 7 layers instead
  - Application, Presentation, Session, Transport, Network, Data Link, Physical
  - Presentation: Syntax and semantics of information transmitted
  - Session: Long-Term transport, such as checkpointing
- 3 central concepts
  - Service: Tells what the layer does
  - Interface: Tells the process above how to access the layer
  - Protocol: How the service is performed; the layer‘s own business.
- In this course, we use the Internet reference model

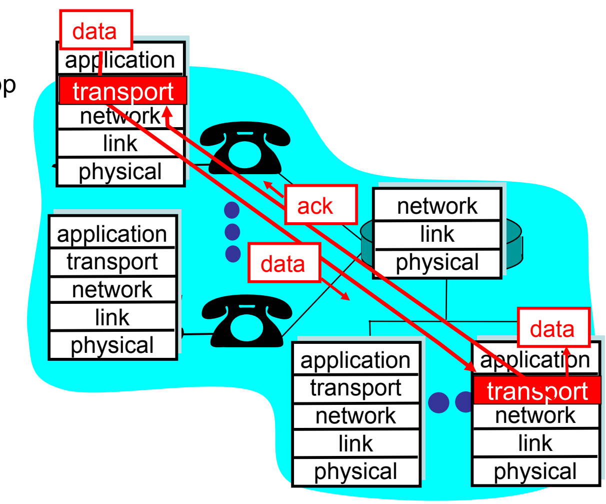
Layering: logical communication

Each layer
- distributed
- “entities” implement layer functions at each node
- entities perform actions, exchange messages with peers

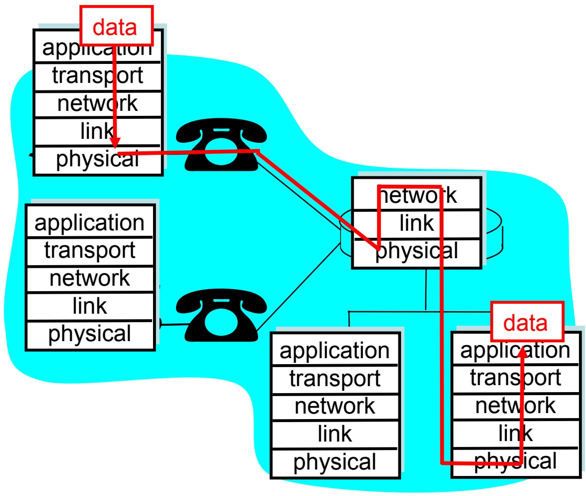


Layering: *logical* communication

Example: transport
- take data from app
- add addressing, reliability check info to form “datagram”
- send datagram to peer
- wait for peer to ack receipt
- Analogy: post office

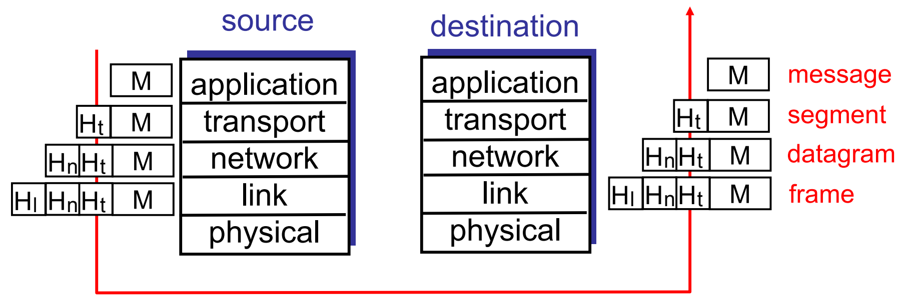


Layering: *physical* communication



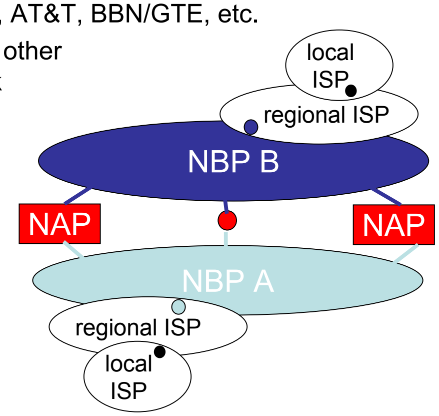
Protocol layering and data

- Each layer takes data from above
  - adds header information to create new data unit
  - passes new data unit to layer below

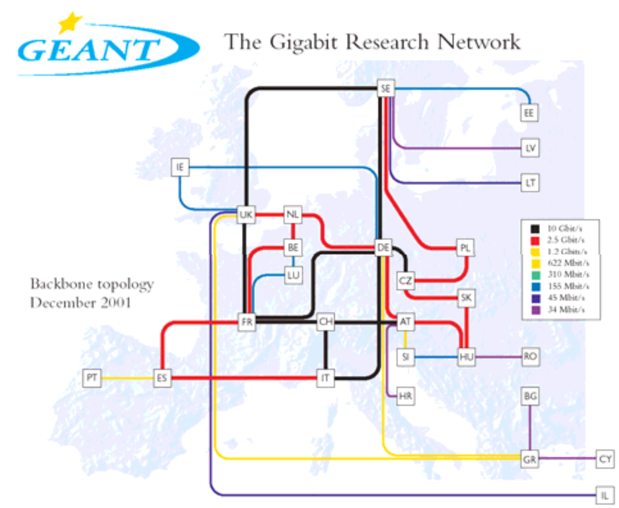


Internet structure: network of networks

- roughly hierarchical
- national/international backbone providers (NBPs), a.k.a. "tier 1"
  - e.g. UUNet, Sprint, Abovenet, AT&T, BBN/GTE, etc.
  - interconnect (peer) with each other privately, or at public Network Access Point (NAP)
- regional ISPs
  - connect into NBPs
- local ISP, company
  - connect into regional ISPs

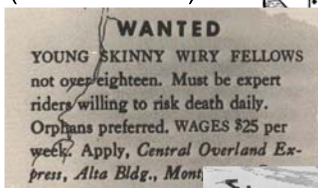


Network of typical backbone provider



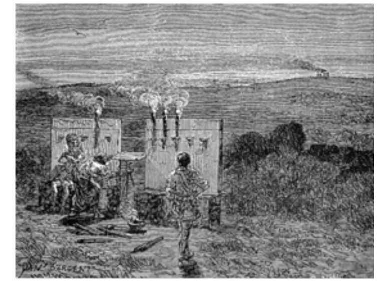
Zur Geschichte der Kommunikation

- Tontäfelchen (3000 v.u.Z)
- Fackeltelegraphie
  - bereits im 5. Jh. v.u.Z. (Griechenland)
- Brieftauben
  - Spätestens Mittelalter
- Reiterboten
  - Ab 1860
- Trommeln, Spiegel, Flaggen, …
- Optische Telegraphen
  - Claude Chappe (Frankreich, 1791)
  - Schweiz: ab 1850



Protokoll von Polybius (2. Jhd. v.u.Z, Griechenland)

- Alphabet als 5 Gruppen zu 5 oder 4 Zeichen
- 2 Gruppen mit je 5 Fackeln
- Verbindungsaufbau
  1. Sendeabsicht: Heben von 2 Fackeln
  2. Empfangsbereitsschaft: Heben von 2 Fackeln
  3. Senken der Fackeln
- Datenübertragung für jedes Zeichen
  1. Linke Fackelngruppe: Zeichengruppe anzeigen
  2. Senken der Fackeln
  3. Rechte Fackelngruppe: Zeichen anzeigen
  4. Senken der Fackeln



Protokoll bei Optischen Telegraphen

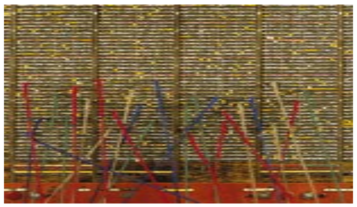
- Regeln für korrekten Nachrichten-austausch
- Typischerweise synchrones Protokoll, d.h. sendende Station muss Symbol so lange zeigen, bis es von der empfangenden Station bestätigt wird.
- Es gab ein Fehlersignal, mit dem man wie bei “backspace” das letzte Zeichen löschen konnte.
- Dieses Protokoll erinnert stark an moderne Protokolle.

Elektrische Telegraphen

- 1774: 26 Drähte (unpraktisch)
- 1837: Elektrischer Zeigertelegraph
  - Cooke und Wheatstone
  - 5 Magnetnadeln, jeweils 2 werden ab-gelenkt und zeigen auf 1 von 20(!) Zeichen
- Man erreicht ca. 25 Zeichen pro Minute
- 1837: Samuel Morse
- 1851: Paris – London
- 1852: 6400km Kabel in England
- 1866: London – New York
  - 20 Wörter kosten $100
- Eigenständige Industrie

Telefon

- Reiss (1863), Bell (1876), Edison (1877), Siemens (1878)
- “This ‘phone’ has way to many shortcomings to consider it as a serious way of communicating. The unit is worthless to us.” [Aktenvermerk Western Union, 1876]
- Ab 1880: Öffentliche Telefonnetze
  – Zuerst maximal 30km Ausdehnung



[New York 1895]

Wireless Transmission

- 1895: Guglielmo Marconi (1874 – 1937)
  – first demonstration of wireless telegraphy (digital!)
  – long wave transmission, high transmission power necessary (> 200kW)
  – Nobel Prize in Physics 1909
- 1901: First transatlantic connection
- 1906 (Xmas): First radio broadcast
- 1907: Commercial transatlantic connections
  – huge base stations (30 100m high antennas)
- 1920: Discovery of short waves by Marconi
- 1928: First TV broadcast
  – Atlantic, color TV

Weitere historische Meilensteine

- 1964: Nachrichtensatelliten
- 1966: Glasfaser
- 1958 : Erste Analoge Handynetze: Deutsches A-Netz
  – Vergleich PTT (Swisscom) NATEL: 1978 – 1995
- 1982 : Start der GSM Standardisierung
- 1997: Wireless LAN
- …

Internet History 1961-72: Early packet-switching principles

- 1961: [Kleinrock] queuing theory shows effectiveness of packet-switching
- 1964: [Baran] packet-switching in military nets
- 1967: ARPAnet conceived by Advanced Research Projects Agency
- 1969: first ARPAnet node operational, first network with 4 nodes
- 1972
  – ARPAnet demonstrated publicly
  – NCP (Network Control Protocol) first host-host protocol
  – first e-mail program
  – ARPAnet has 15 nodes

1972-80: Internetworking, new and proprietary nets

- 1970: ALOHAnet satellite network in Hawaii
- 1973: Metcalfe's PhD thesis proposes Ethernet
- 1974: [Cerf and Kahn] architecture for interconnecting networks
- Late 70's:
  - proprietary architectures: DECnet, SNA, XNS
  - switching fixed length packets (ATM precursor)
- 1979: ARPAnet has 200 nodes
- Vinton G. Cerf and Robert E. Kahn's (Ehrendoktoren der ETH seit 1998) internetworking principles:
  - minimalism
  - autonomy
  - no internal changes required to interconnect networks
  - best effort service model
  - stateless routers
  - decentralized control
  - ➢ define today's Internet architecture

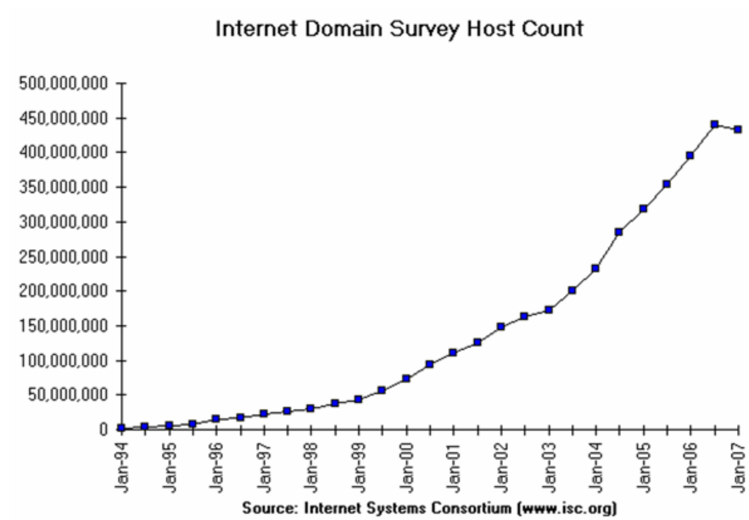
1980-90: new protocols, a proliferation of networks

- 1983: deployment of TCP/IP
- 1982: SMTP e-mail protocol defined
- 1983: DNS defined for name-to-IP-address translation
- 1985: FTP protocol defined
- 1988: TCP congestion control
- new national networks: NSFnet, CSNET, BITnet, Minitel
- 100,000 hosts connected to confederation of networks

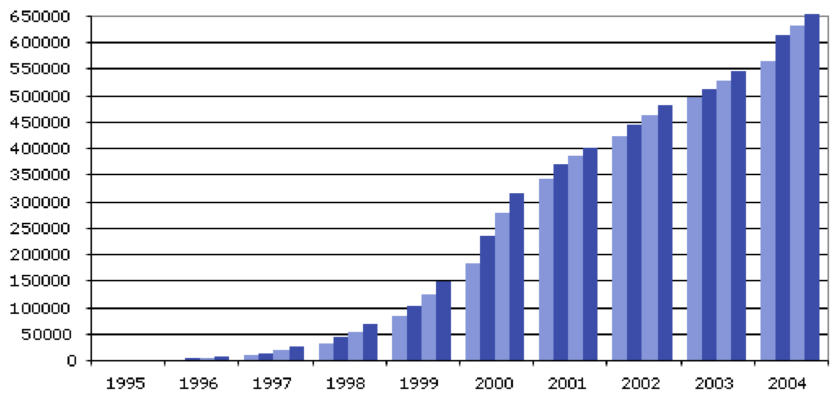
1990's: Commercialization, WWW

- Early 1990's: ARPAnet decommissioned
- 1991: NSF lifts restrictions on commercial use of NSFnet (decommissioned, 1995)
- early 1990s: WWW
  - hypertext [Bush 1945, Nelson 1960's]
  - HTML, http: Berners-Lee
  - 1994: Mosaic, later Netscape
  - late 1990's commercialization of the WWW
- Late 1990's
  - est. 50 million computers on Internet
  - est. 100 million+ users
  - backbone links running at 1 Gbps

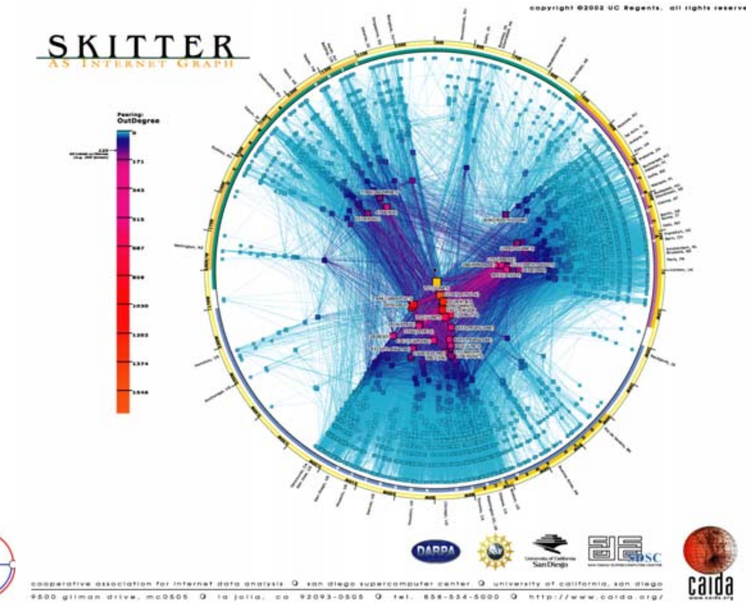
Number of hosts in the Internet (lower bound)



Domain Names ending in .ch

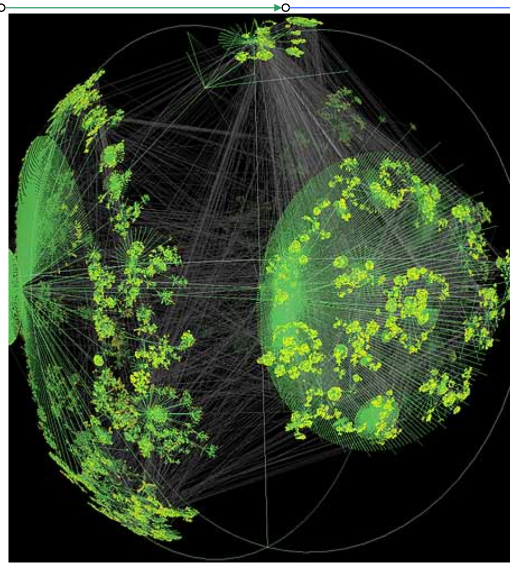


Internet Providers by “size” and “region”



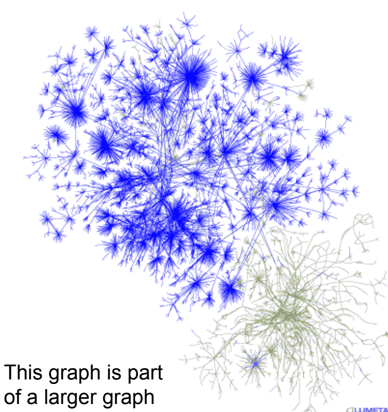
Internet Topology

The image depicts the Internet topology. It shows 535,000-odd Internet nodes and over 600,000 links. The nodes, represented by the yellow dots, are a large sample of computers from across the whole range of Internet addresses.

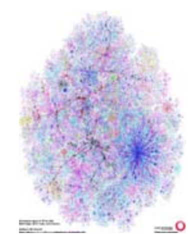


[http://www.cybergeography.org/atlas/topology.html]

More Internet Topology



This graph is part of a larger graph and shows the portion of a corporate Intranet that is 'leaking' with the Internet.

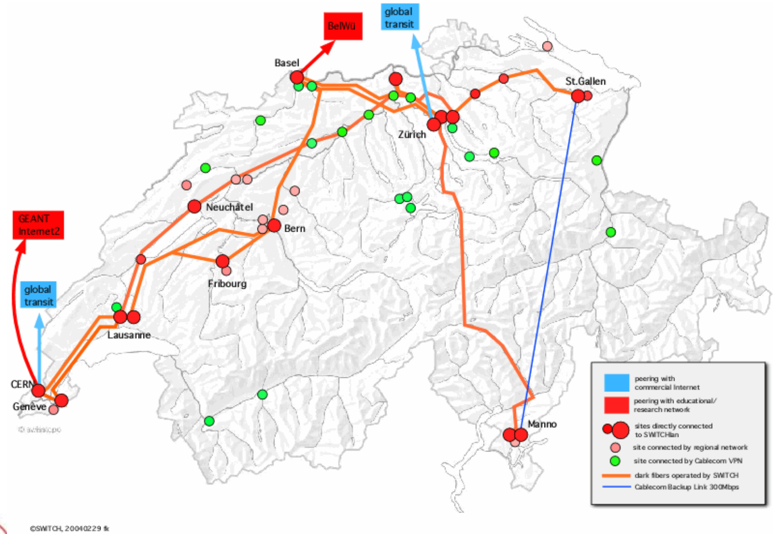


This graph shows the router level connectivity of the Internet.

A topology map of a core network of a medium-sized ISP.

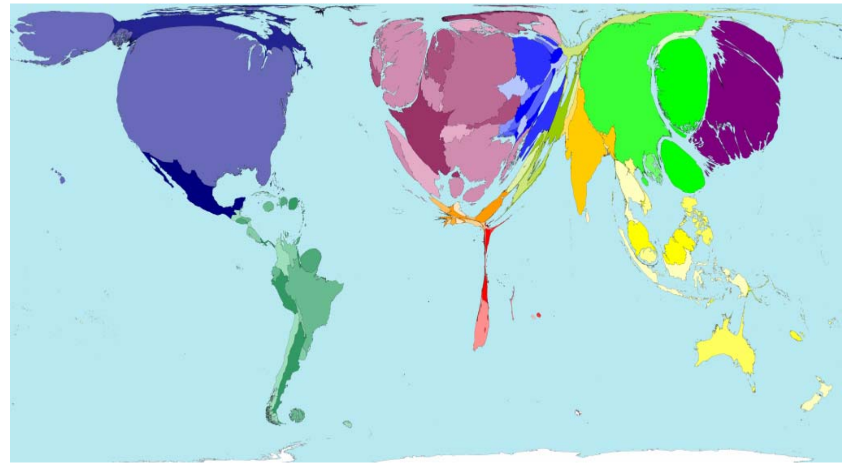
[http://www.cybergeography.org/atlas/topology.html]

The SWITCH network



[www.switch.ch]

Internet Users Worldwide

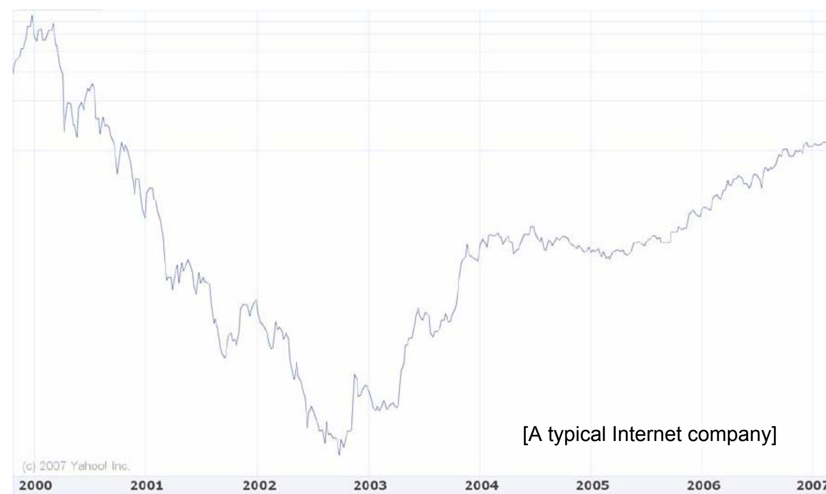


[http://www.worldmapper.org]

Favorite Web Sites

- Switzerland
  - bluewin (world rank 498), ricardo, pctipp, bluewin, bluewin, bluewin, libellules, sunrise, 20min, ubs, gameswelt, sunrise...
- World
  - google.com, yahoo, google.de, google (https), google.co.uk, google.fr, mail.google, microsoft, bbc, bbc, foxnews, google.ca

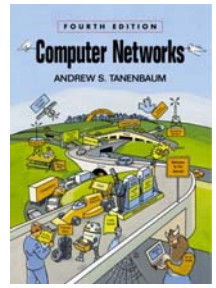
The “Dot-Com Bubble”



Course overview

Introduction Overview	*Intro*
Applications: Email, WWW, etc. More Applications and Sockets	*Layer 5*
Transport Layer: UDP and TCP Advanced Transport Layer	*Layer 4*
Network Layer: Routing Basics Advanced Network Layer	*Layer 3*
Link Layer: Aloha, etc. Link Layer: Ethernet, Hubs, etc.	*Layer 2*
Physical Layer, Wireless	*Layer 1*
Peer-to-Peer Computing Network Security Distributed Systems Network Management Etc.	*Special Topics*

Literature



Course book
Andrew S. Tanenbaum
Computer Networks
Fourth Edition

German version
also available

There are alternatives, for example Kurose/Ross

Other Courses, Master in Distributed Systems

Vernetzte Systeme

Verteilte Systeme

- Enterprise Application Integration – Alonso
- Parallel and Distr. Databases – Alonso
- Ubiquitous Computing – Mattern
- Distributed Algorithms – Mattern
- Ad Hoc and Sensor Networks – Wattenhofer
- Principles of Distributed Computing – Wattenhofer
- Web Algorithms – Wattenhofer & Widmayer
- More: See www.{dcg, tik.ee, pc.inf}.ethz.ch.