Modern Internet architecture, technology & philosophy

Advanced Internet Services
Dept. of Computer Science
Columbia University

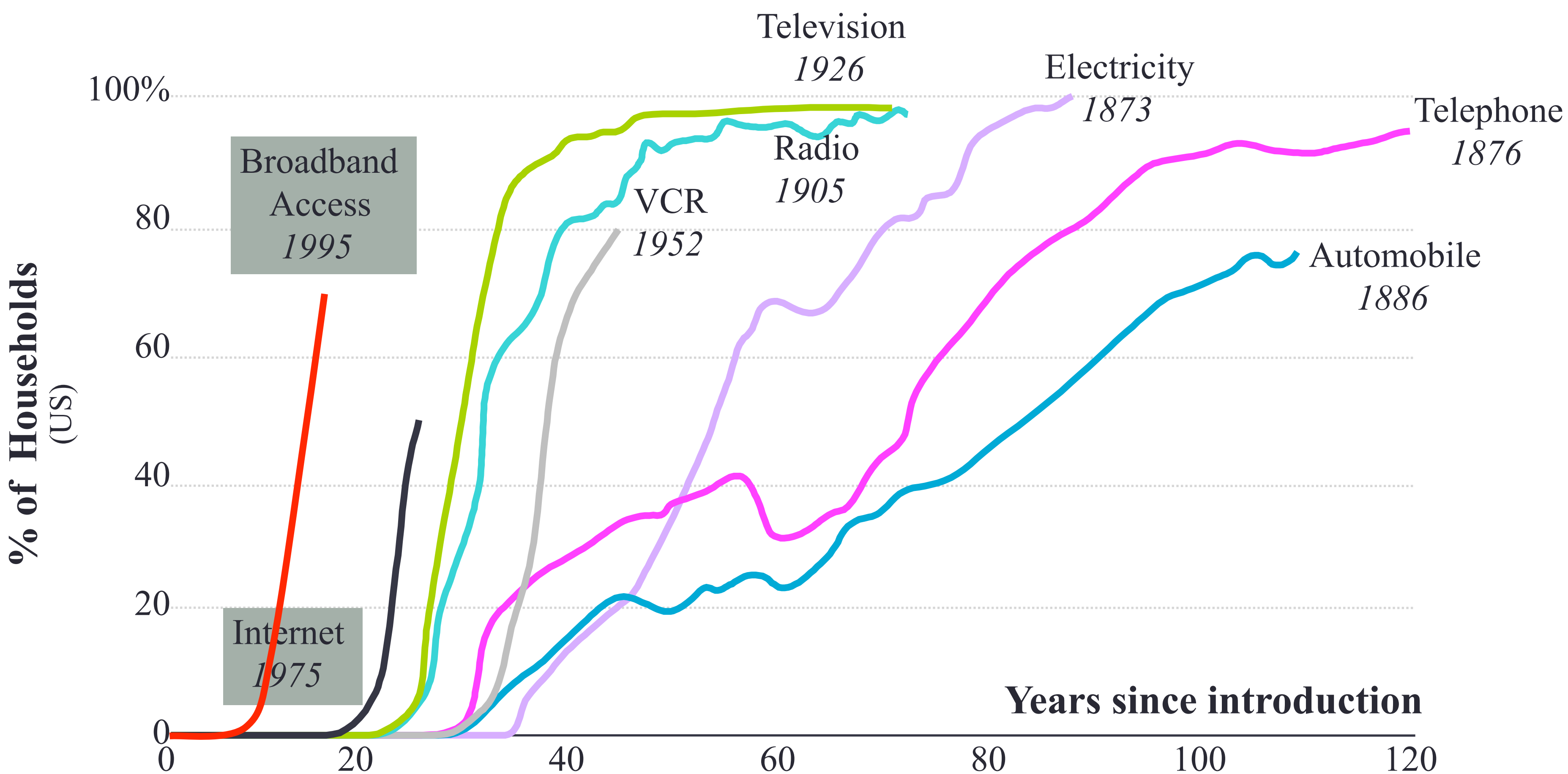
Henning Schulzrinne
Spring 2015
02/23/2015

NETWORK EVOLUTION & RESEARCH

Networking is getting into middle years

	idea	current
IP	1969, 1980?	1981 (RFC 791)
TCP	1974 (RFC 675)	1981 (RFC 793)
telnet	1969 (RFC15)	1983 (RFC 854)
ftp	1971 (RFC 114)	1985 (RFC 959)

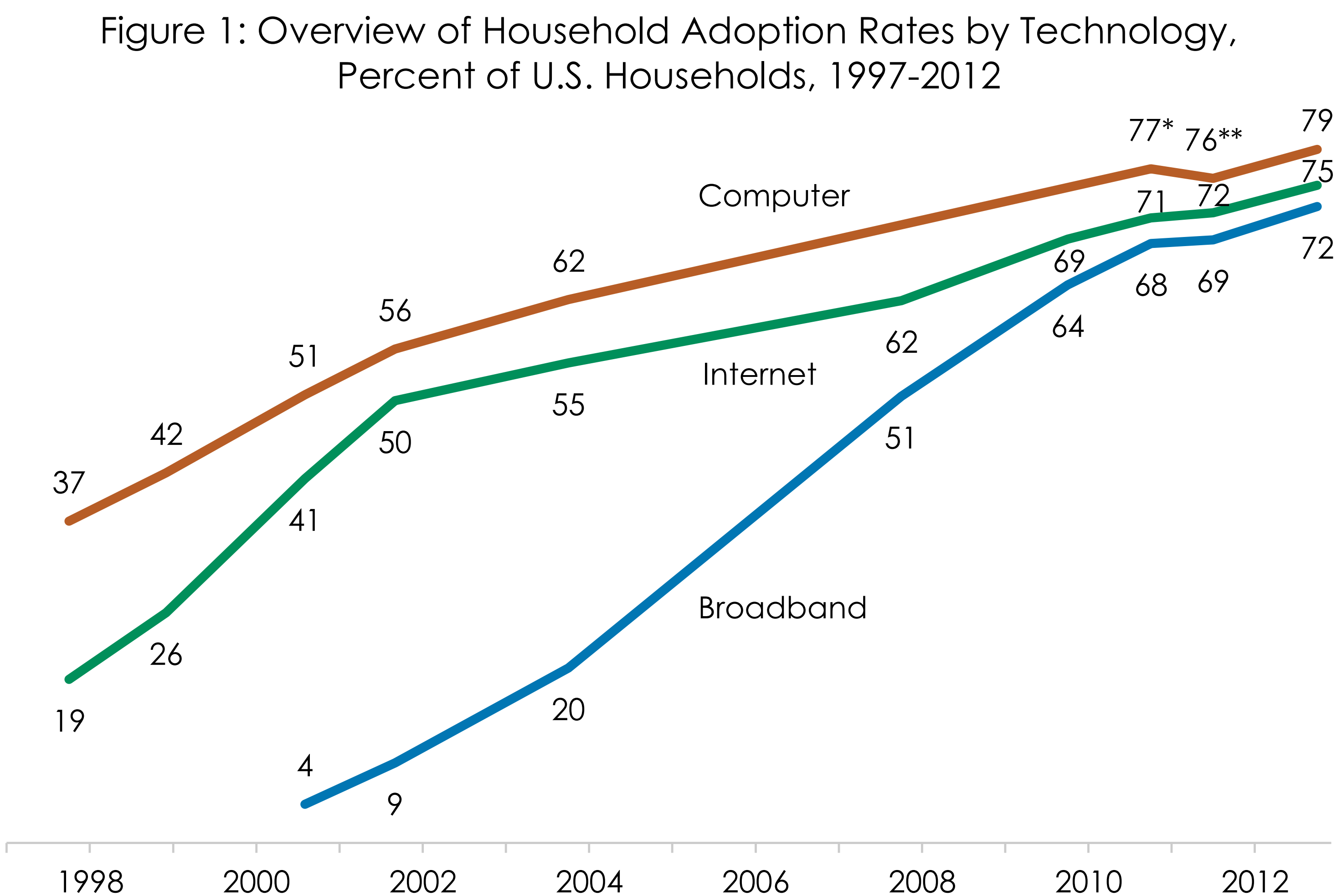
Internet/broadband: one of the fastest applications ever introduced



2005 = 30% broadband / 2010 = 70% broadband estimate

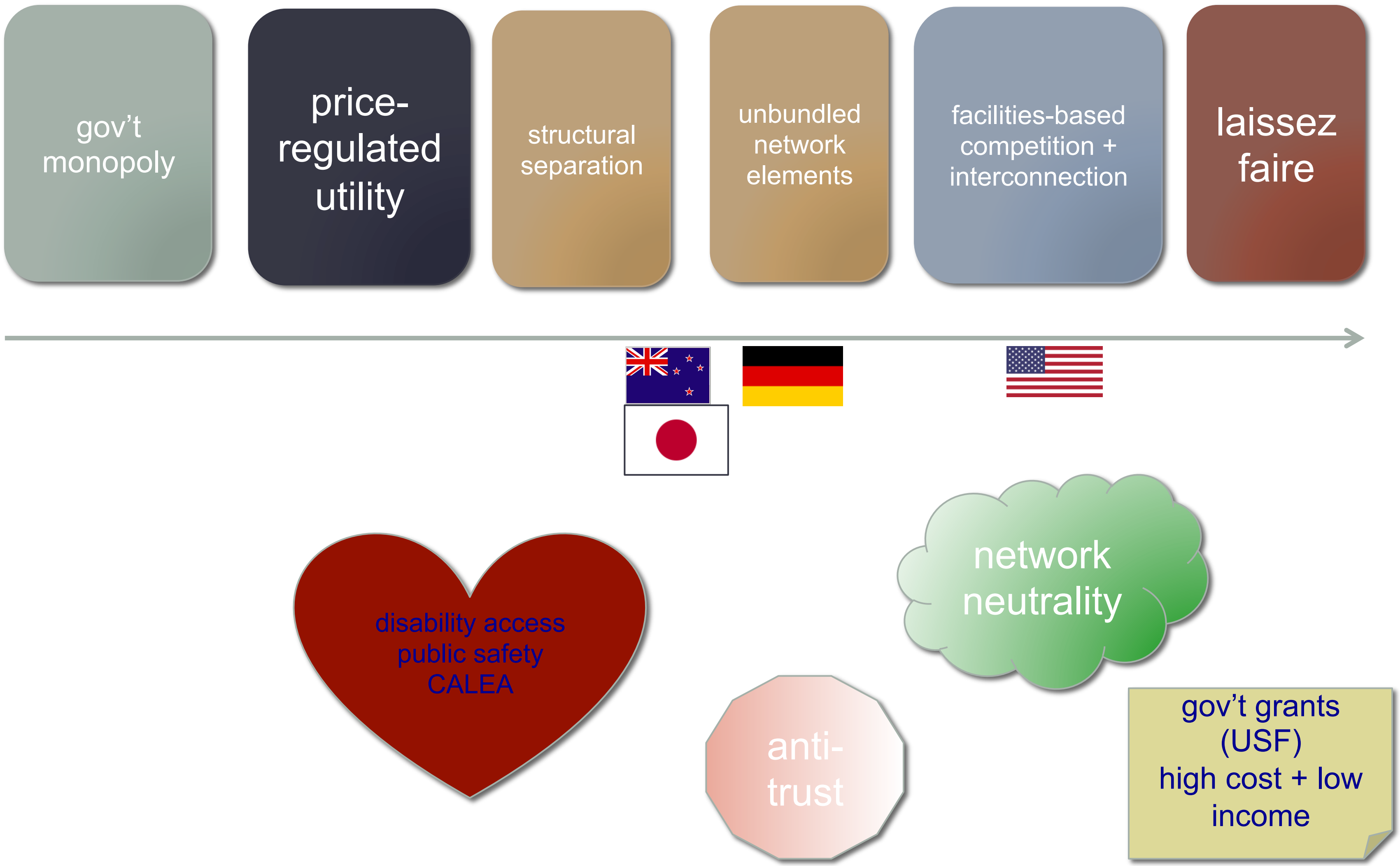
Source: Michael Fox and Forbes Magazine, Morgan Stanley

US broadband adoption



NTIA 2014

Telecom policy tool kit



Standardization

- Oscillate: convergence → divergence
  - continued convergence clearly at physical layer
  - connectivity trumps functionality
  - niches larger → support separate networks
- Two facets of standardization:

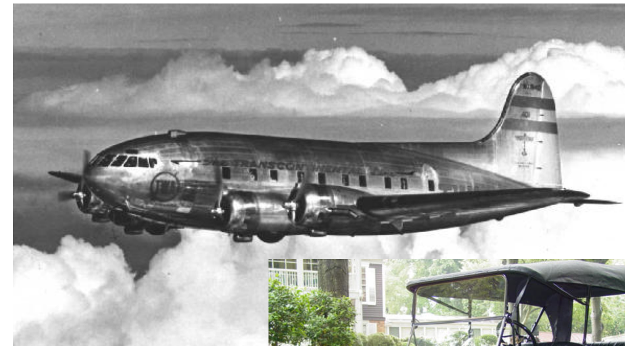
1. *public, interoperable description of protocol, but possibly many* (Tanenbaum)
2. reduction to 1-3 common technologies
   - L2: Arcnet, tokenring, DECnet, ATM, FDDI, DQDB, SONET … → Ethernet
   - L3: IP, IPX, X.25, OSI → IP
   - OS: dozens → Windows, MacOS, Linux

Standardization

- Have reached phase 2 in most cases
  - RPC (SOAP, REST) and presentation layer (XML, JSON) most recent 'conversions'
- Often, non-standardized technologies can be deployed faster
  - single (dominant) vendor
    - Skype vs. SIP and H.323
    - AOL IM and XMPP (Jabber) vs. SIMPLE
    - SMB vs. NFS vs. WebDAV
- → Standardization *after* success?
- IETF one-protocol-for-application vs. everything-is-RPC
  - not enough network experts → standardization scales better
  - see OASIS, OMA standardization groups

Technologies at ~30 years

- Other technologies at similar maturity level:
  - air planes: 1903 – 1938 (Stratoliner)
  - cars: 1876 – 1908 (Model T)
  - analog telephones: 1876 – 1915 (transcontinental telephone)
  - railroad: 1820s – 1860s (transcontinental railroad)

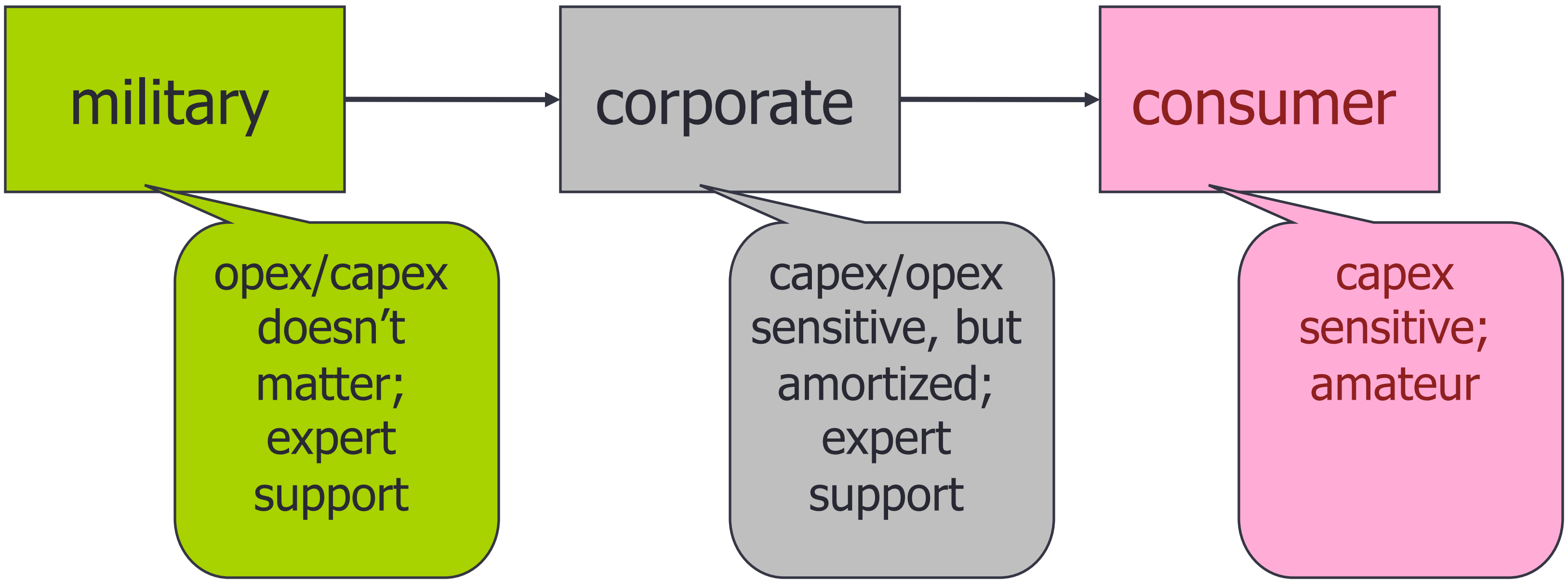


Observations on progress

- 1960s: military → professional → consumer
    - now, often reversed
- Communications technologies rarely disappear (as long as operational cost is low):
    - exceptions:
        - telex, telegram, semaphores → fax, email
        - X.25 + OSI, X.400 → IP, SMTP
    - analog cell phones
    - → thus, NGN (post-IP, future Internet) discussions likely academic

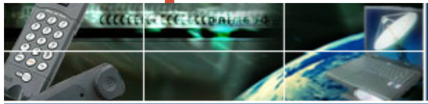
Lifecycle of technologies

traditional technology propagation:



Can it be done? Can I afford it? Can my mother use it?

Example: Telex



Sidereal Corporation

About Us Contact Us EDI/FACNET Products Services Support Home Intranet

Telex Services

Telex remains one of the most reliable, universally available forms of electronic data communication, with approximately one million users worldwide. Today, many thousands of Telex messages are exchanged hourly, securing business transactions throughout the globe. By operating within its own physical network, Telex provides a legal and secure transmission standard, regardless of origin or destination. Still leading the way in the Global Telex market are the Finance, Shipping, Oil and Gas industries.

Within finance, Telex still remains a medium for transferring and confirming trading instructions. Although the primary medium for electronic data communication within the banking community is now SWIFT, there are still many banks that are are not connected to the SWIFT network. Furthermore, in the event of a disaster normal trading can still continue via Telex if it is kept as a back up to SWIFT.

Shipping companies still use Telex as a method of communication with vessels at sea and for various methods of general trading. Although Email and Fax is used widely in the shipping industry, there are still many shipping organizations where Telex is primary means of communications.

Sidereal offers many connection options to the Telex network for sending and receiving, including dial-up and ip.

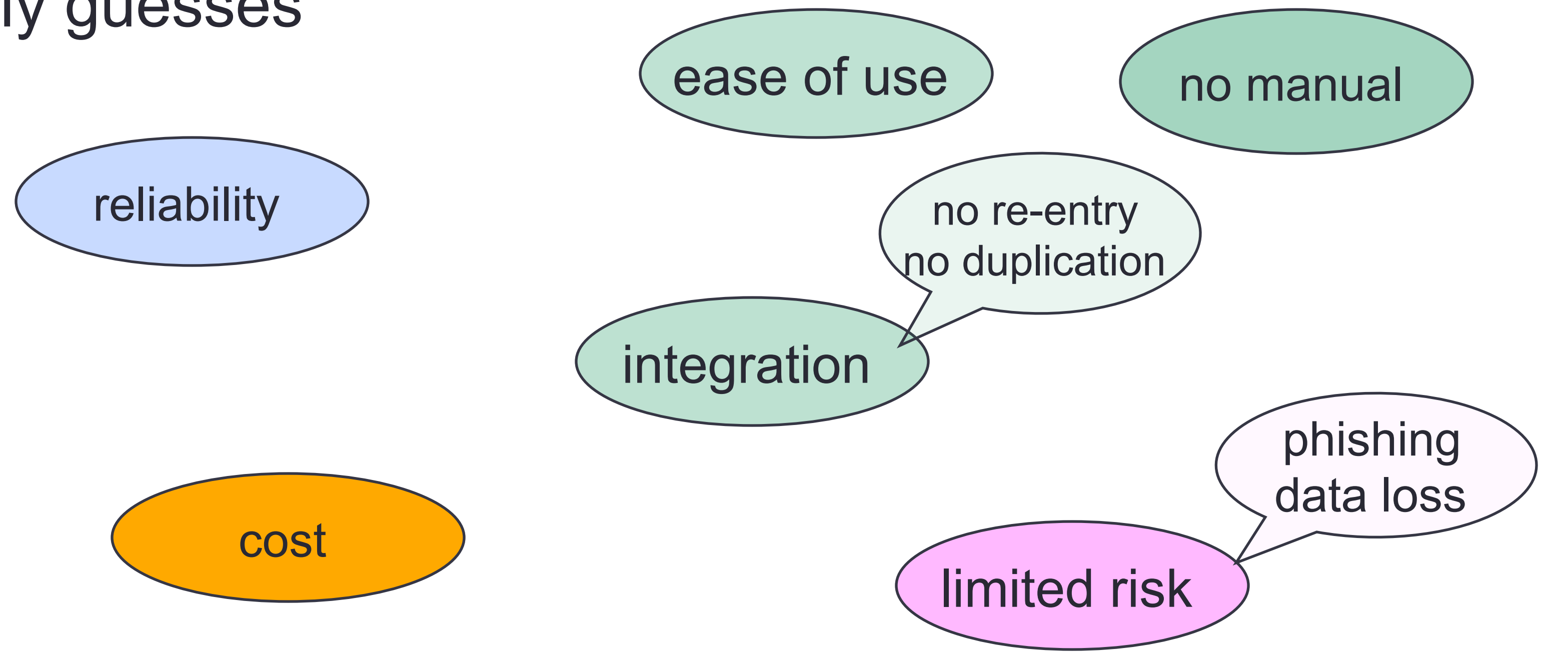
Click here to see the Telex rate and country schedule.

Transition of networking

- Maturity → cost dominates
  - can get any number of bits anywhere, but at considerable cost and complexity
  - casually usable bit density still very low
- Specialized → commodity
  - OPEX (= people) dominates
  - installed and run by 'amateurs'
  - need low complexity, high reliability

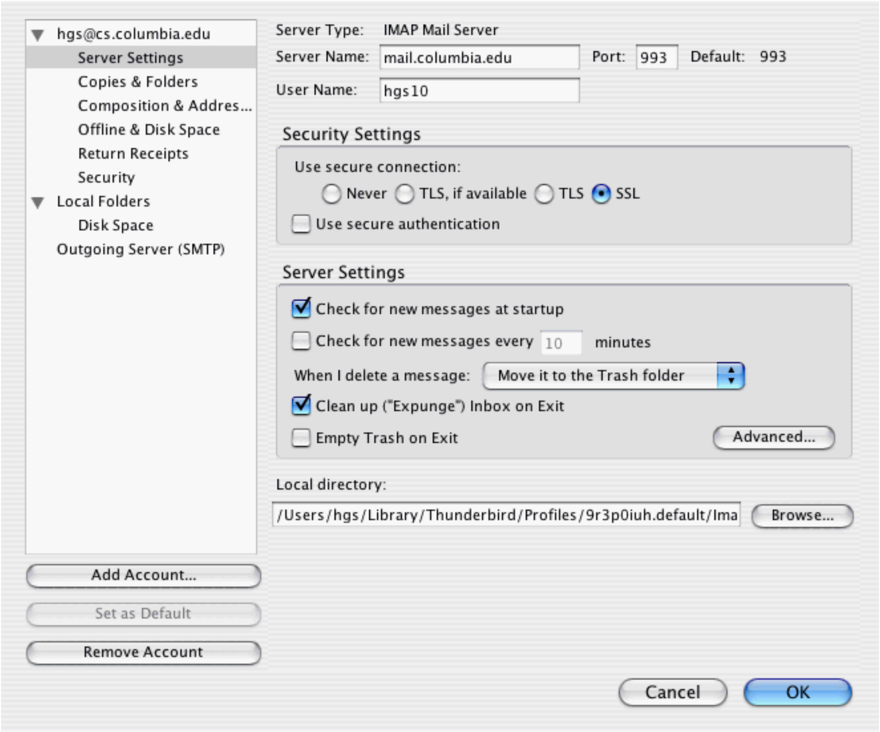
User challenges vs. research challenges

- Are we addressing real user needs?
    - Engineering vs. sports scoring
- My guesses



Example: Email configuration

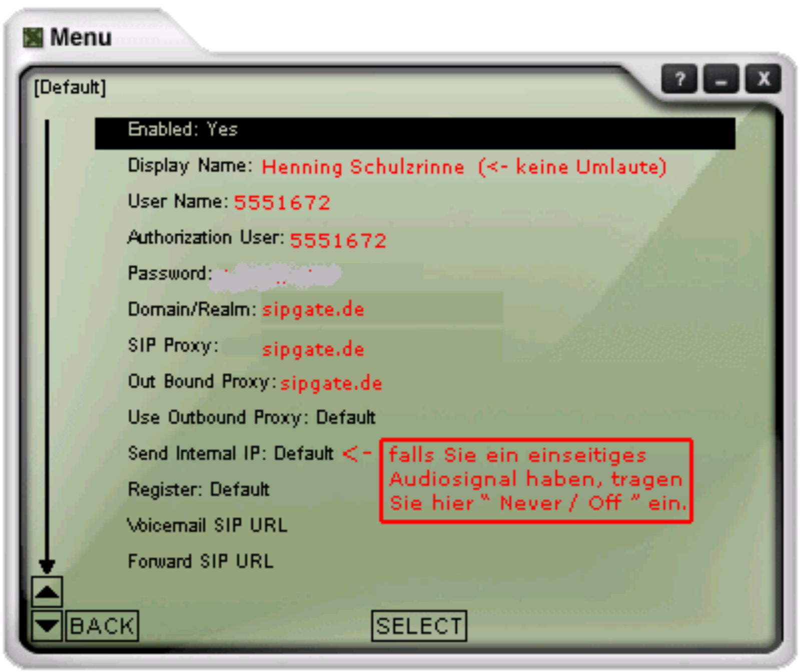
- Application configuration for (mobile) devices painful
- SMTP port 25 vs. 587
- IMAP vs. POP
- TLS vs. SSL vs. “secure authentication”
- Worse for SIP...



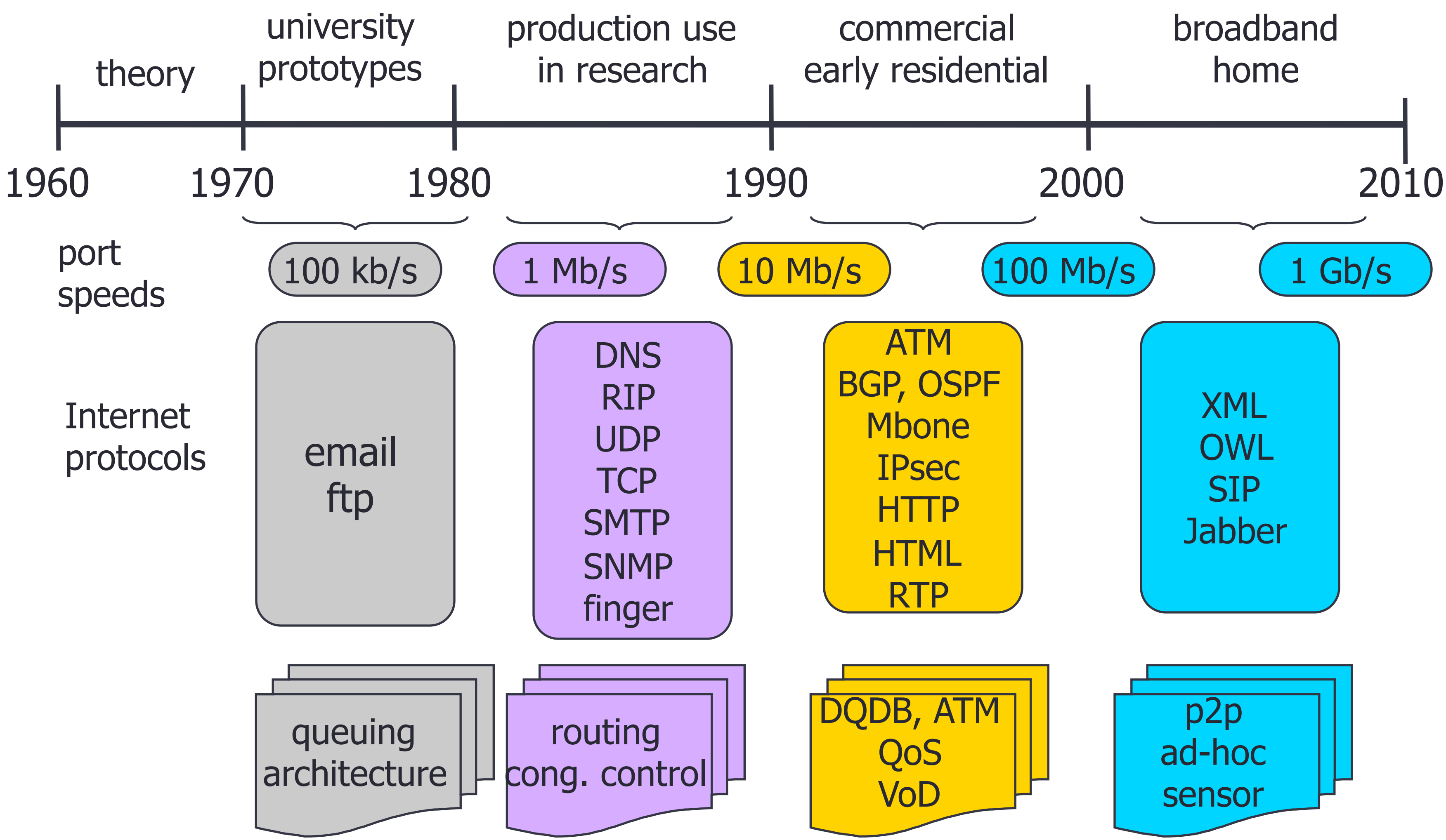
Example: SIP configuration

partially explains

- highly technical parameters, with differing names
- inconsistent conventions for user and realm
- made worse by limited end systems (configure by multi-tap)
- usually fails with some cryptic error message and no indication which parameter
- out-of-box experience not good



Internet and networks timeline



Cause of death for the next big thing

	QoS	multi-cast	mobile IP	active networks	IPsec	IPv6
not manageable across competing domains	✝	✝	✝	✝		
not configurable by normal users (or apps writers)	✝			✝	✝	
no business model for ISPs	✝	✝	✝	✝	✝	✝
no initial gain	✝	✝	✝	✝		✝
80% solution in existing system	✝	✝	✝	✝	✝	✝ (NAT)
increase system vulnerability	✝	✝	✝	✝		

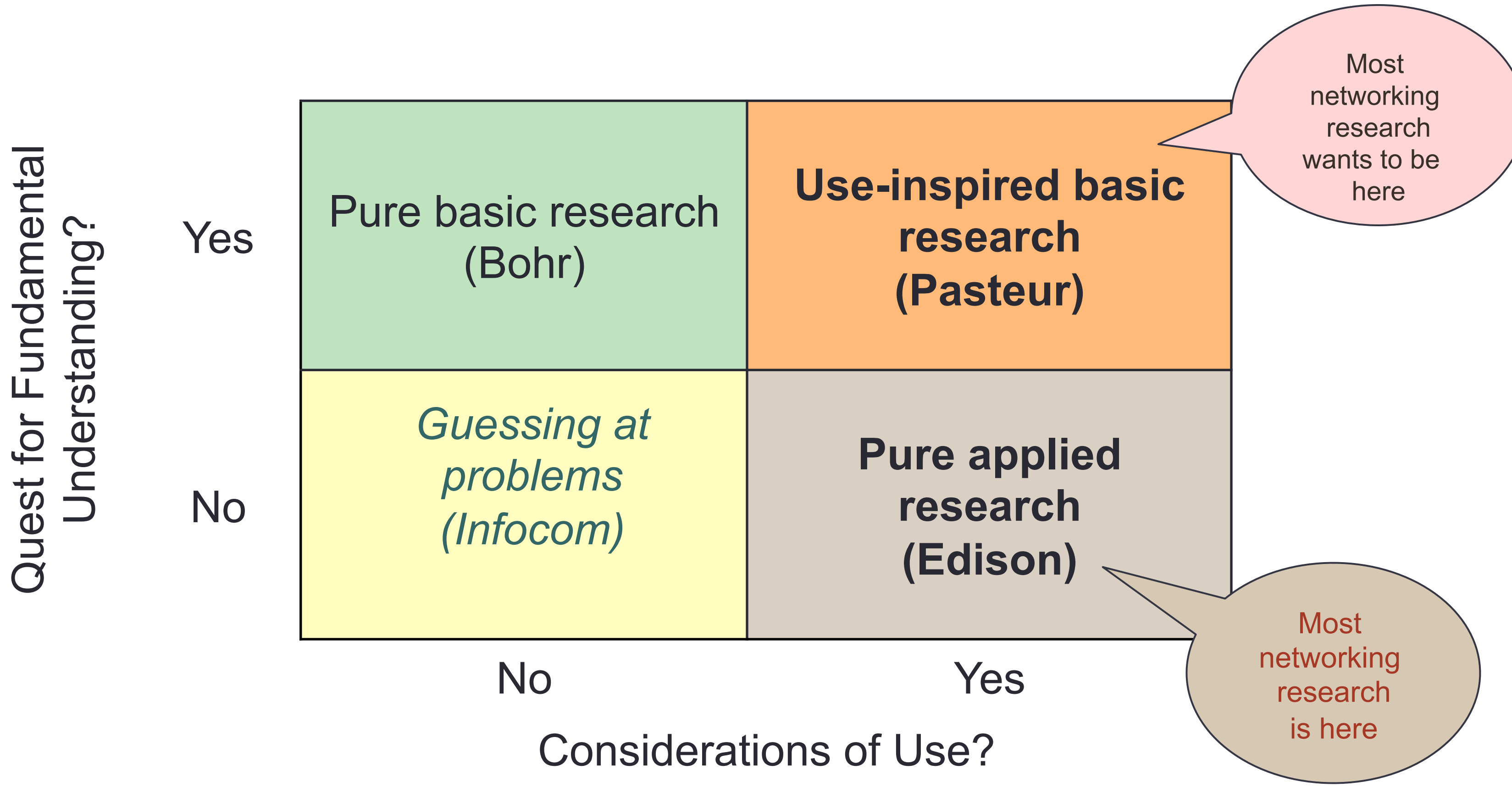
Why do good ideas fail?

- Research: O(.), CPU overhead
  - “per-flow reservation (RSVP) doesn’t scale” → not the problem
    - at least now -- routinely handle O(50,000) routing states
- Reality:
  - deployment costs of any new L3 technology is probably billions of $
- Cost of failure:
  - conservative estimate (1 grad student year = 2 papers)
  - 10,000 QoS papers @ $20,000/paper → $200 million

Research: Network evolution

- Only three modes, now thoroughly explored:
    - packet/cell-based
    - message-based (application data units)
    - session-based (circuits)
- Replace specialized networks
    - left to do: embedded systems
        - need cost(CPU + network) < $10
        - cars
        - industrial (manufacturing) control
        - commercial buildings (lighting, HVAC, security; now LONworks)
        - remote controls, light switches
        - keys replaced by biometrics

Research: Pasteur’s quadrant



Pasteur’s Quadrant: Basic Science and Technological Innovation, Stokes 1997 (modified)

Maturing network research

- Old questions:
  - Can we make X work over packet networks?
    - All major dedicated network applications (flight reservations, embedded systems, radio, TV, telephone, fax, messaging, …) are now available on IP
  - Can we get M/G/T bits/s to the end user?
  - Raw bits everywhere: "any media, anytime, anywhere"
- New questions:
  - Dependency on communications → Can we make the network reliable?
  - Can non-technical users use networks without becoming amateur sys-admins? → auto/zeroconfiguration, autonomous computing, self-healing networks, …
  - Can we make networks affordable to everyone?
  - Can we prevent social and financial damage inflicted through networks (viruses, spam, DOS, identity theft, privacy violations, …)?

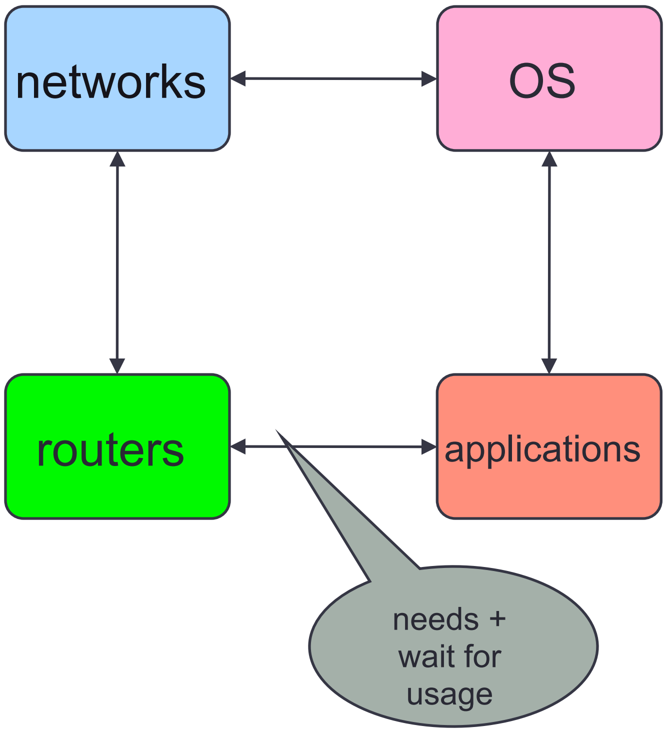
New applications

- New bandwidth-intensive applications
  - Reality-based networking
  - (security) cameras → "ambient video"
- New bandwidth-**ex**tensive applications
  - communicate infrequently → setup overhead
  - SIGFOX network
- Distributed games often require only low-bandwidth control information
  - current game traffic ~ VoIP
  - 4G, 5G → low latency
- Computation vs. storage vs. communications
  - communications cost has decreased less rapidly than storage costs

SIGFOX (902 MHz, 100 bps) is a connectivity solution that focuses on low throughput devices. On SIGFOX you can send between 0 and 140 messages per day and each message can be up to 12 bytes of actual payload data.

Change is hard

- No new network services deployed since 1980s
  - universal upgrade
  - chicken/egg (network/OS) problem
- "Innovation at edges"
- Applications easier, as long as
  - TCP-based
  - client-server
  - … but there are exceptions (p2p)

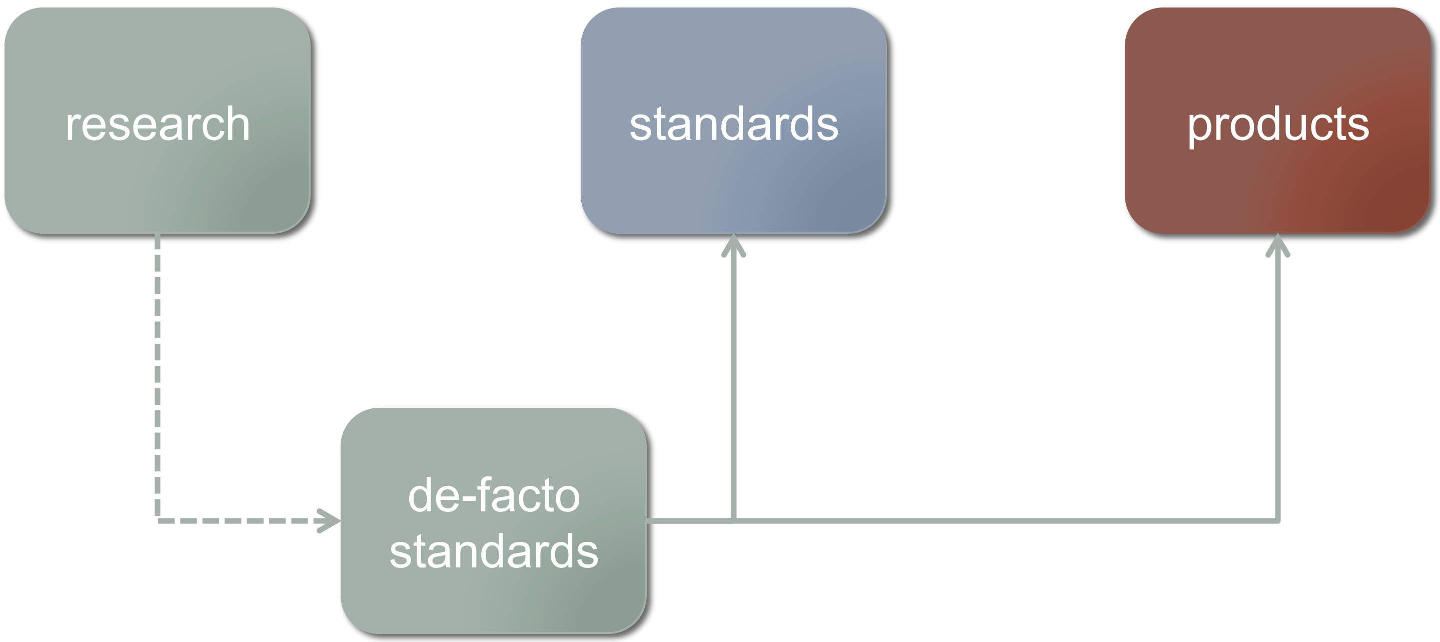


Time of transition

Old	New
IPv4	IPv6
circuit-switched voice	VoIP
separate mobile voice & data	LTE + LTE-VoIP
911, 112	NG911, NG112
digital cable (QAM)	IPTV
analog & digital radio	Pandora, Internet radio, satellite radio
credit cards, keys	NFC
end system, peers	client-server v2 aka cloud

all the energy into transition → little new technology

Technology transition



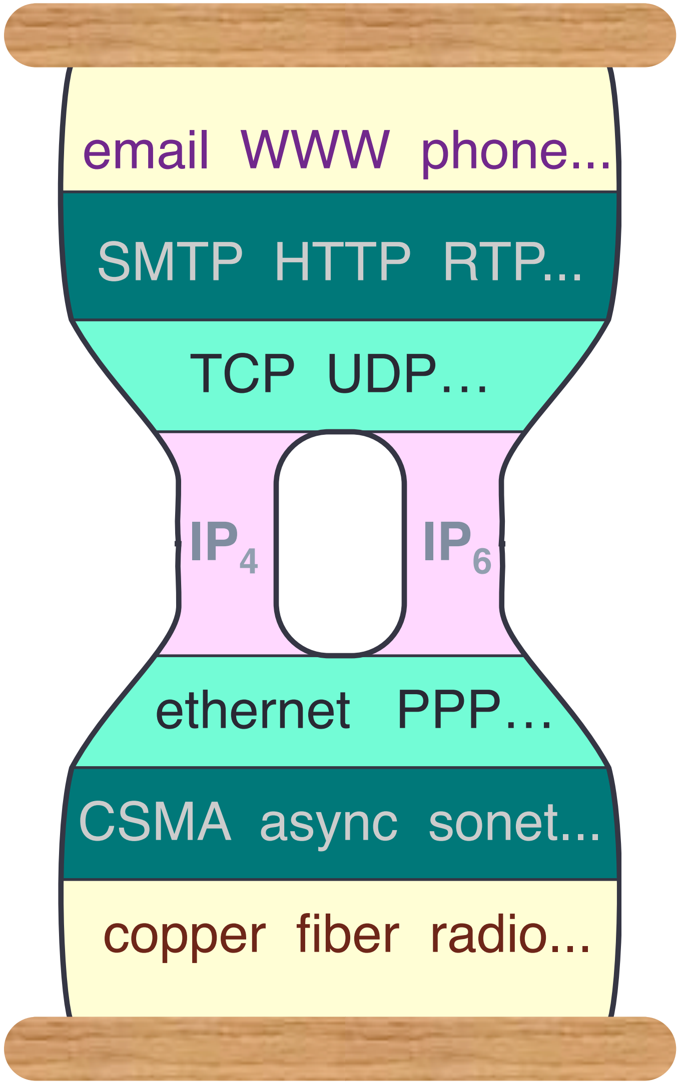
protocols vs. algorithms!

Internet challenges

- IP address depletion
- NAT, middle boxes and the loss of transparency
- Routing infrastructure
- Quality of service
- Security
- DNS scaling
- Dealing with privatization
- Interplanetary Internet

COMPLEXITY

Mid-Life Crisis



- doubles number of service interfaces
- requires changes above & below
- major interoperability issues

"Why architectural complexity is like body fat"

- You naturally tend to gain it while you grow older
- Very easy to gain and very hard to get rid of
  - Designing complex solutions and protocols easier than designing simple ones.
- Healthy to have some, but not too much
- Having it on waist may be worse than elsewhere
- Younger and slimmer will eventually beat you
  - Architectural complexity → reduced agility → younger and less complex systems eventually replace older and more complex system.
- *Sometimes* surgery is a good way to start
- Long term results require constant exercise

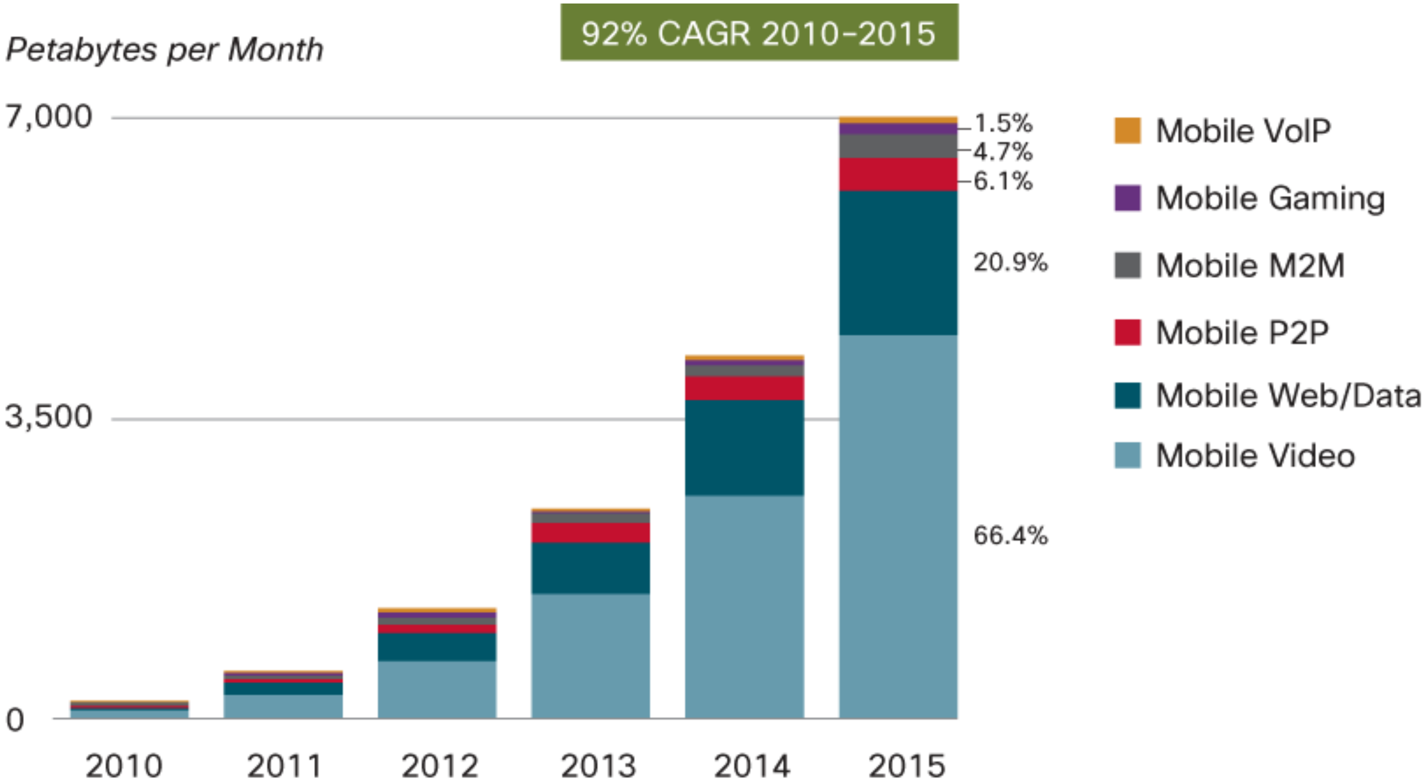
http://www.tml.tkk.fi/~pnr/FAT/

Causes of complexity

- Complexity: implementation vs. run-time
  - system vs. protocol
- After-the-fact enhancements:
  - security
  - NAT traversal
  - mobility
  - internationalization (e.g., DNS)
- Wrong layer for function
  - multicast? IP security?
- Options
  - e.g., multiple transport protocols, IPv4 & IPv6
- Lots of special protocols
  - e.g., IMAP, POP, SMTP
- Manual configuration

NETWORK TRAFFIC & ECONOMICS

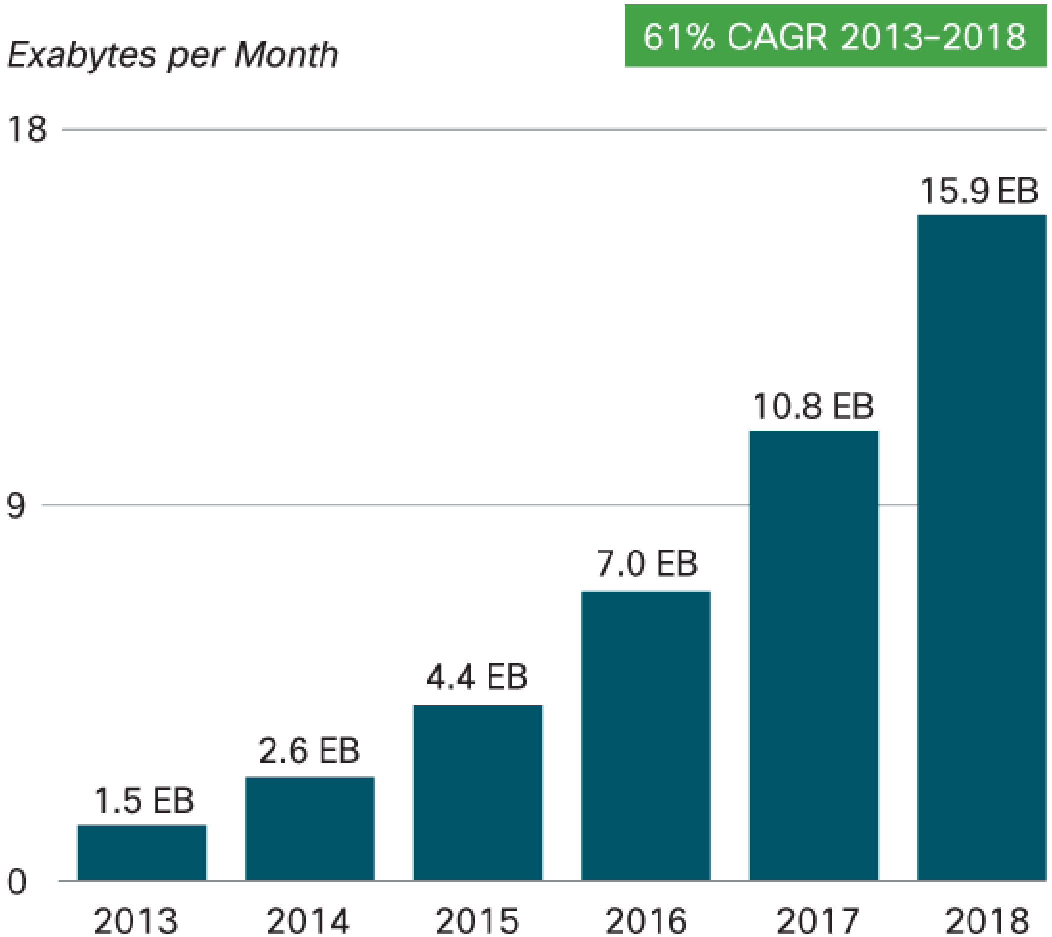
Mobile traffic distribution – 2011 prediction



VoIP traffic forecasted to be 0.4% of all mobile data traffic in 2015.

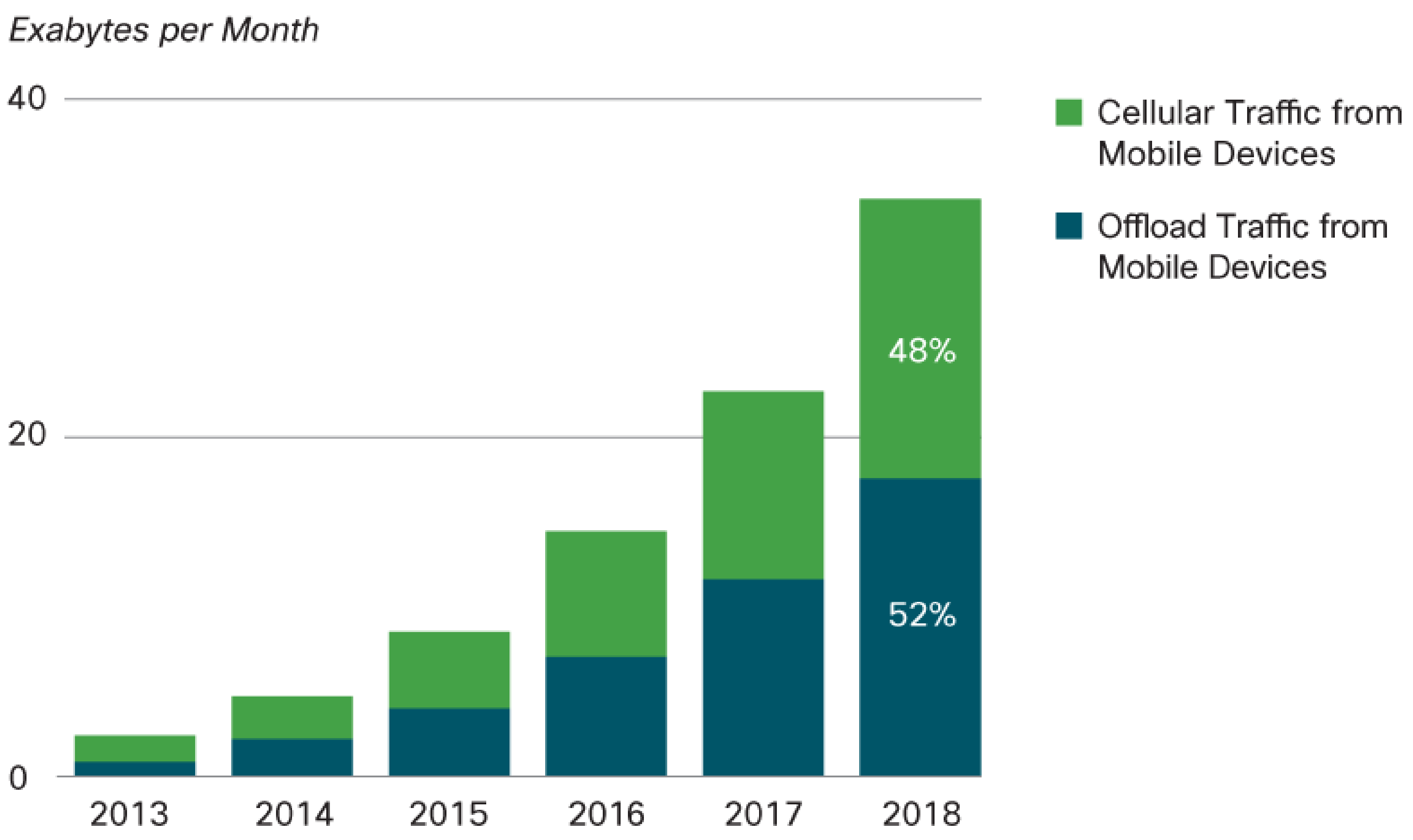
Source: Cisco VNI Mobile, 2011

Mobile traffic distribution – 2014 prediction



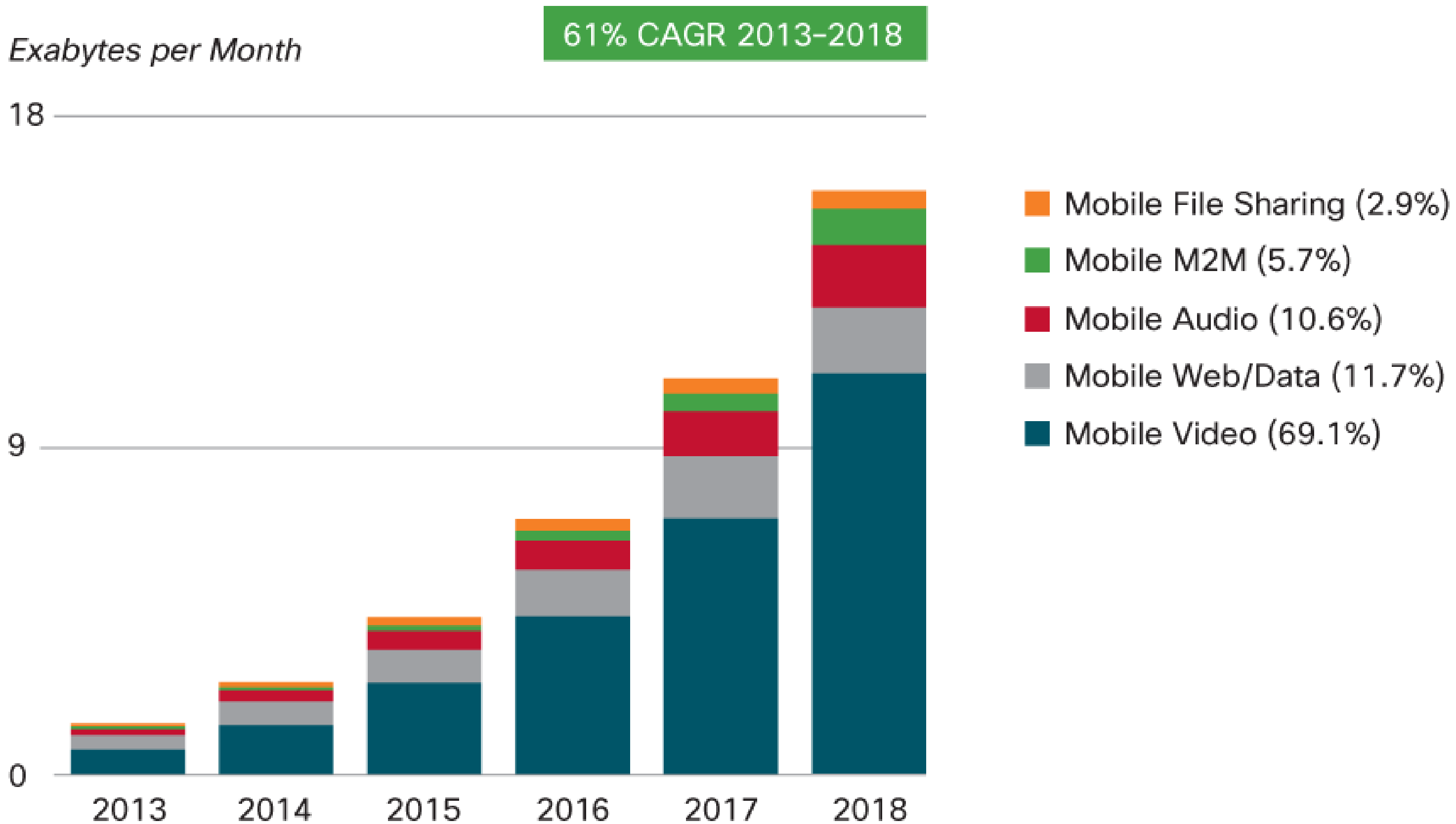
Source: Cisco VNI Mobile, 2014

Mobile traffic is mostly Wi-Fi



Source: Cisco VNI Mobile, 2014

Mobile traffic



Figures in parentheses refer to traffic share in 2018.

Source: Cisco VNI Mobile, 2014

Monthly Consumption (fixed)

North America	Mean	Median	Mean : Median
Upstream	8.5 GB	1.8 GB	4.7
	5.8	11.3	
Downstream	48.9 GB	20.4 GB	2.4
Aggregate	**57.4 GB**	**22.5 GB**	**2.6**

- top 1% →
  - 49.7% of upstream traffic
  - 25% of downstream traffic

Europe	Mean	Median	Mean : Median
Upstream	5.1 GB	1.5 GB	3.4
	4.5	5.8	
Downstream	23.1 GB	8.7 GB	2.7
Aggregate	**28.2 GB**	**10.1 GB**	**2.8**

The value of bits

- Technologist: A bit is a bit is a bit
- Economist: Some bits are more valuable than other bits
  - e.g., $(email) >> $(video)

Application	Volume	Cost per unit	Cost / MB	Cost / TB
Voice (13 kb/s GSM)	97.5 kB/minute	10c	$1.02	$1M
Mobile data	5 GB	$40	$0.008	$8,000
MMS (pictures)	< 300 KB, avg. 50 kB	25c	$5.00	$5M
SMS	160 B	10c	$625	$625M

Video, video and more video

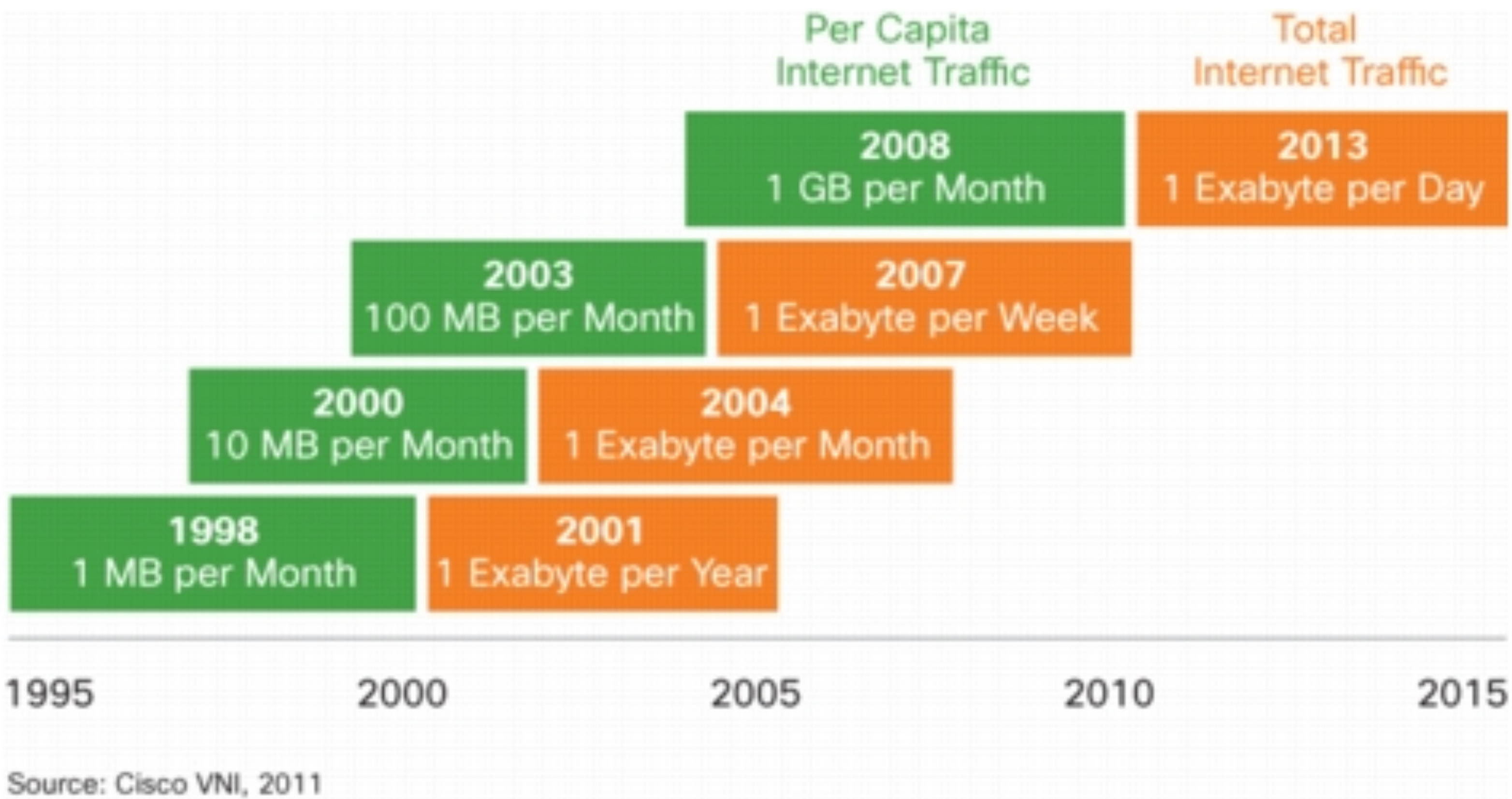
Upstream		Downstream		Aggregate	
BitTorrent	52.01	Netflix	29.70%	Netflix	24.71%
HTTP	8.31%	HTTP	18.36%	BitTorrent	17.23%
Skype	3.81%	YouTube	11.04%	HTTP	17.18%
Netflix	3.59%	BitTorrent	10.37%	YouTube	9.85%
PPStream	2.92%	Flash Video	4.88%	Flash Video	3.62%
MGCP	2.89%	iTunes	3.25%	iTunes	3.01%
RTP	2.85%	RTMP	2.92%	RTMP	2.46%
SSL	2.75%	Facebook	1.91%	Facebook	1.86%
Gnutella	2.12%	SSL	1.43%	SSL	1.68%
Facebook	2.00%	Hulu	1.09%	Skype	1.29%
Top 10	83.25%	Top 10	84.95%	Top 10	82.89%

Average monthly usage

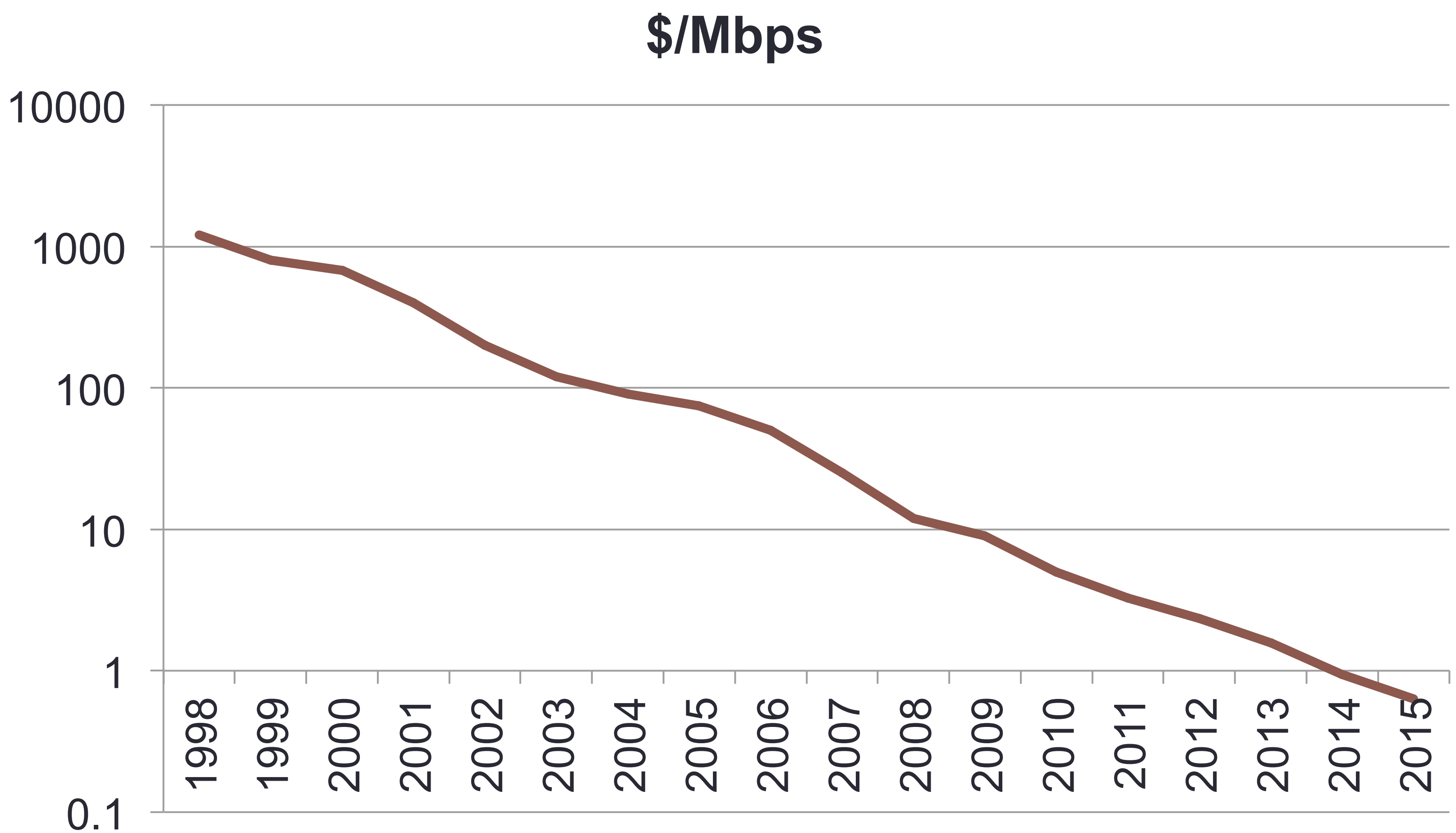
- Average monthly TV consumption (US): 154 hours
- Netflix: 1 GB/hour (SD) … 2.3 GB/hour (HD)
  - → 300 GB/month
  - more if people in household watch different content

monthly usage	overage cost (AT&T Uverse)	2010	2012	2015
> 50 GB	$0	9.4%	14.1%	21.5%
> 100 GB	$0	5.3%	8.2%	15.3%
> 200 GB	$10	1.4%	4.4%	8.8%
> 500 GB	$50	0.4%	0.8%	2.6%
> 1 TB	$150	0.0%	0.2%	0.7%

Bandwidth generations



Transit prices



http://drpeering.net/white-papers/Internet-Transit-Pricing-Historical-And-Projected.php

Cost of bandwidth

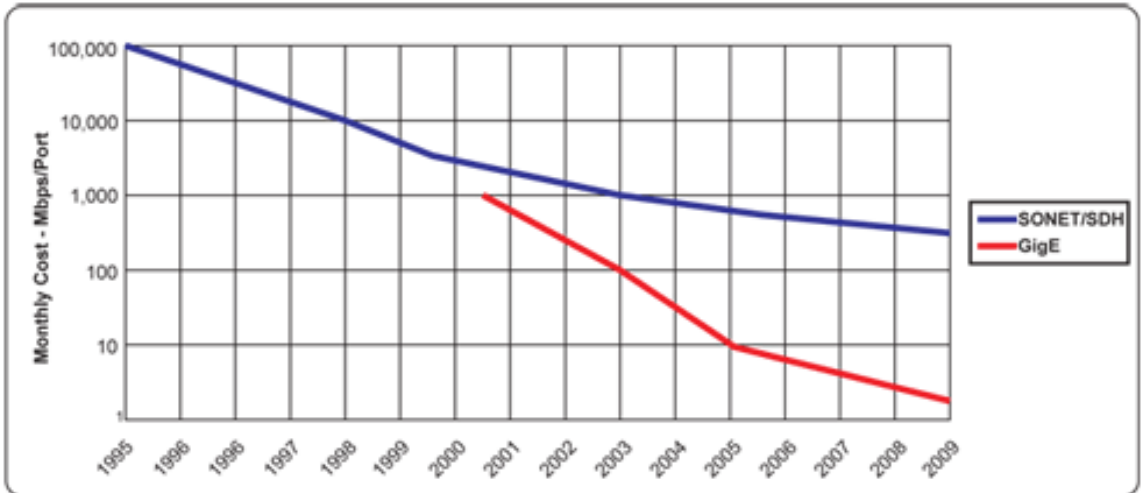
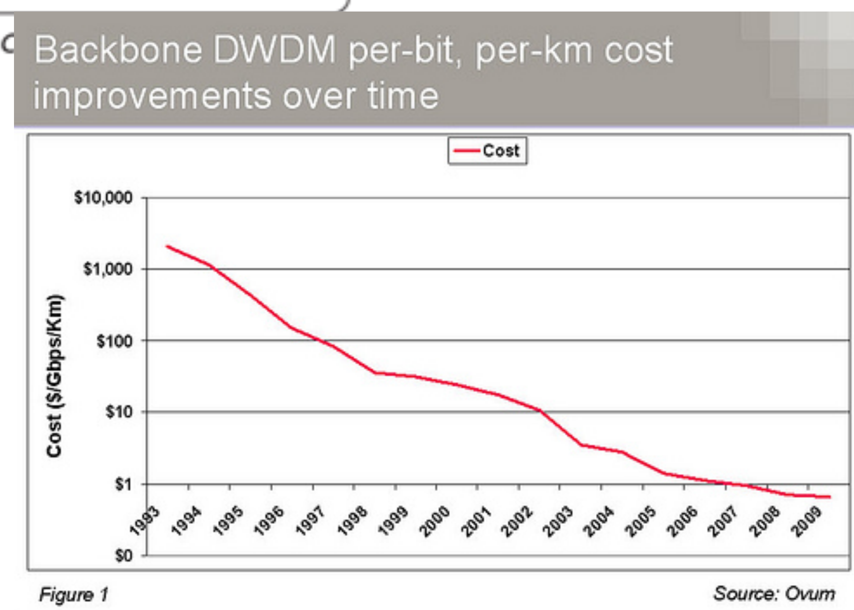


Fig. 5: SONET/SDH vs. Ethernet Monthly C



Bandwidth costs

- Amazon EC2
  - $50 - $120/TB out, $0/TB in
- CDN (Internet radio)
  - $600/TB (2007)
  - $7-20/TB (Q1 2014 – CDNpricing.com)
- NetFlix (7 GB DVD)
  - postage $0.70 round-trip → $100/TB
- FedEx – 2 lb disk
  - 5 business days: $6.55
  - Standard overnight: $43.68
  - Barracuda disk: $91 - $116/TB
- DVD-R (7 GB)
  - $0.25/disk → $35/TB

Cost of bandwidth (2011 & 2015)

Service	Speed (Mb/s)	Average price/ month 2015 (2011)	$/Mb/s
DS1 (T1)	1.54	$295 ($450)	$197 ($292)
DS3	45	$1950 ($5,000)	$43 ($111)
Ethernet over Copper	10	$310 ($950)	$31 ($95)
Fast Ethernet	100	$1,800 ($5,000)	$18 ($50)
Gigabit Ethernet	1000	$4,000 ($25,000)	$4 ($25)

Bandwidth costs

- Amazon EC2
  - $100/TB in, $100/TB out
- CDN (Internet radio, Hulu, Netflix, …)
  - $600/TB (2007)
  - $100/TB (Q1 2009 – CDNpricing.com)
  - $15/TB to $50/TB (Q4 2010, 500 TB/month)
- NetFlix (7 GB DVD)
  - postage $0.70 round-trip → $100/TB
- FedEx – 2 lb disk from NY to San Diego
  - 5 business days: $9.08
  - Standard overnight: $66.99
    - 18 hours → 0.25 Gb/s
  - Hitachi disk: $34/TB

NETWORK REALITY

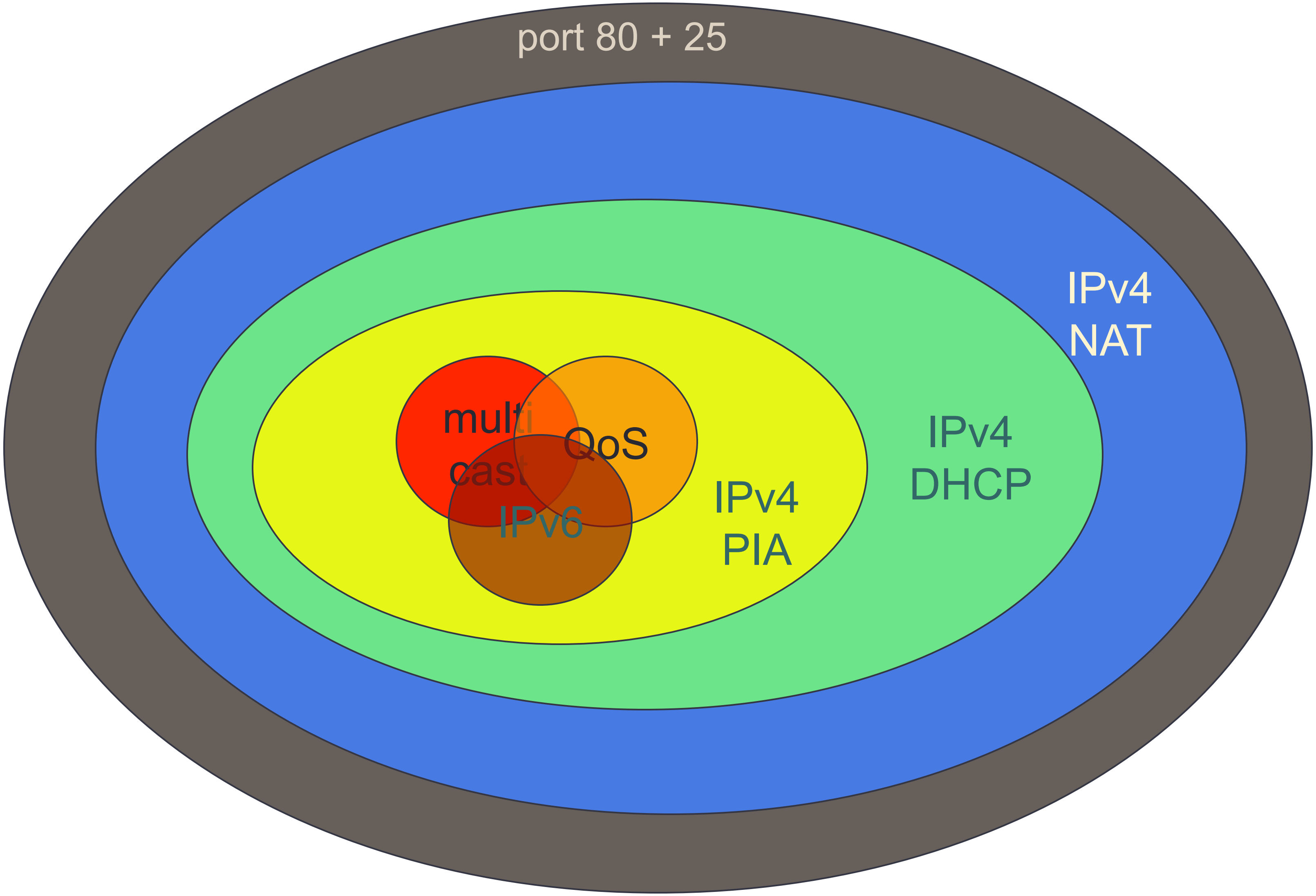
Textbook Internet vs. real Internet

Ideal	Reality
end-to-end (application only in 2 places)	middle boxes (proxies, ALGs, …)
permanent interface identifier (IP address)	time-varying (DHCP)
globally unique and routable	network address translation (NAT)
multitude of L2 protocols (ATM, ARCnet, Ethernet, FDDI, modems, …)	dominance of Ethernet, but also L2's not designed for networks (1394 Firewire, Fibre Channel, MPEG2, …)

Textbook Internet vs. real Internet

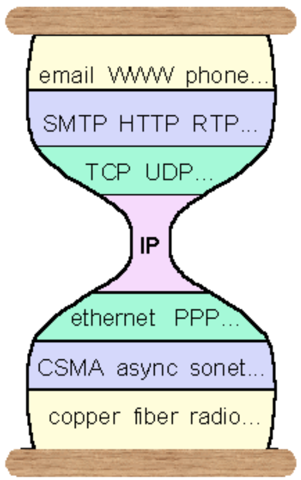
mostly trusted end users	hackers, spammers, con artists, pornographers, …
small number of manufacturers, making expensive boxes	Linksys, Dlink, Netgear, …, available at Walmart
technical users, excited about new technology	grandma, frustrated if email doesn't work
4 layers (link, network, transport, application)	layer splits
transparent network	firewalls, L7 filters, "transparent proxies"

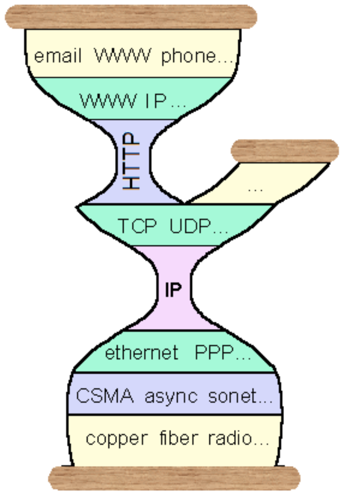
Which Internet are you connected to?



The two-port Internet

- Many public access systems only allow port 80 (HTTP) and maybe 25 (SMTP)
  - e.g., public libraries
- Everything tunneled over HTTP
  - Web-based email
  - Flash video delivery (e.g., YouTube)
  - HTTP CONNECT for remote login





Dave Thaler

Causes

- Link-layer technologies
  - satellite, DSL
  - NBMA
- Network-layer technologies
  - security: broken by design vs. broken by accident?
  - NATs
  - Ill-defined meaning of IP addresses and names
    - theoretically, single network interface
    - practically, often more than that
      - virtualization
      - multi-homing
      - fail-over

Private Internet -- challenges

- *Public* Internet = collection of *privately*-owned (mostly) for-profit networks
- Incentives for greedy behavior
- Special-purpose networks
  - VoIP networks
  - 3GPP, NGN, … ➔ "walled garden"
  - sub-applications large enough to support own infrastructure
- Private protocols
  - e.g., most IM protocols
- Patent encumbrances
  - see https://datatracker.ietf.org/public/ipr_disclosure.cgi
- D. Clark, J. Wroclawski, K. Sollins, R. Braden, "Tussle in Cyberspace: Defining Tomorrow's Internet", ToN, June 2005

Tussle in Cyberspace

- Traditional view: design technology to make choices
- Tussle view: design technology to allow choices
  - “we are designing the social contract that the Internet embodies”
  - not a final outcome, on-going process → lawyers vs. engineers
- Multiple competing interests
  - application value capture
    - high value content looks the same to ISP
  - traffic price differentiation
    - willingness to pay
  - investment in infrastructure vs. open interfaces
    - sunk costs
  - greed (local traffic optimization vs. social optimum)
  - privacy and anonymity vs. societal goals
    - CALEA, network resource protection, spam, DRM
- → Allow multiple outcomes, but give users choice (competition)
  - user-selected routes and servers

see also http://www.aarnet.edu.au/engineering/wgs/video/presentations/2004Feb/clark.ppt

Other network models

- Interplanetary networks
  - Extremely long round-trip times, large feedback delays
  - Protocols designed with terrestrial timeout parameters
  - See Vint Cerf’s web page and Delay-Tolerant Networking Research Group (DTNRG)
- Disconnected or delay-tolerant operation
  - K. Fall, “A Delay Tolerant Networking Architecture for Challenged Internets”, SIGCOMM 2003
  - “store-and-forward” at the content level
- Sensor networks
  - Extremely lossy links
- Content-based networks

Other network types

network	partially disconnected	mobile end systems	wireless links	mobile routers	energy optimization	node computation
“classical” Internet	caching, sync.	fixed nomadic mobile	last hop			
mesh networks			all links	slowly		
MANET		only	only	fast		
delay-tolerant networks	possibly	planets space craft	only			
sensor networks	some systems	yes	common	some	crucial	common