Quality of Service - 20 years old and ready to get a job?

Henning Schulzrinne
Dept. of Computer Science
Columbia University
New York, New York
(sip:)schulzrinne@cs.columbia.edu

IWQoS (Karlsruhe, Germany)

June 6th, 2001

Overview

- why has QoS failed in the Internet?
- do we still need it?
- some thoughts on future traffic
- musings on TCP fairness
- what's left to do?

QoS

- the bad news
  - hardly anybody uses QoS techniques
  - talking about the same issues year after year
- the good news
  - more than 3,000 papers since 1984
  - limited TOS support in end systems emerging
  - still need it – but rarely

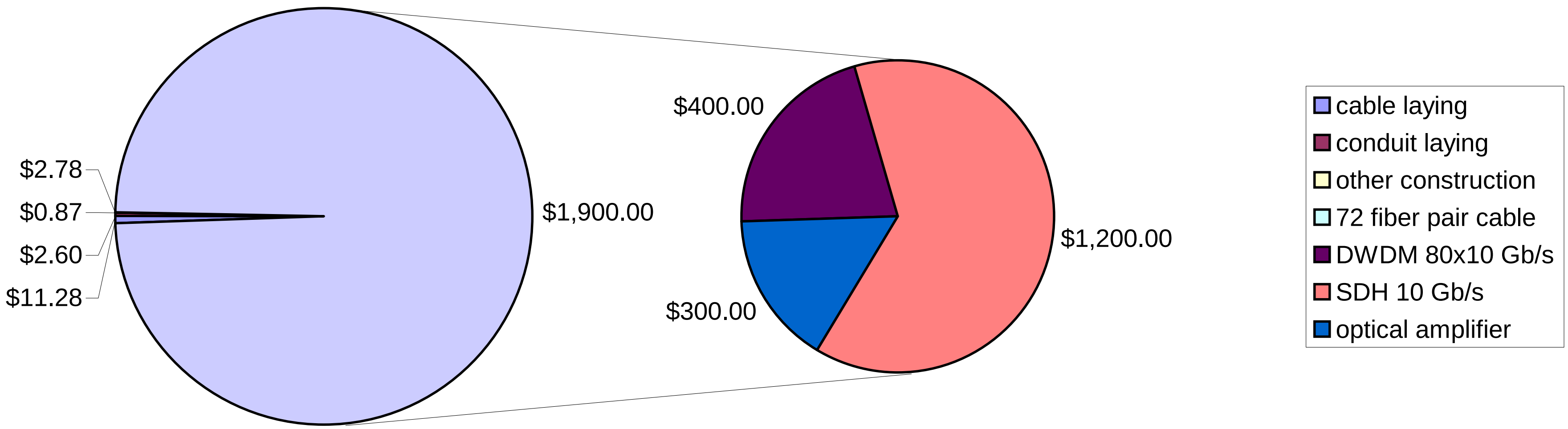
Why is QoS unpopular?

- need to admit failure – “bandwidth too cheap to meter”
- undemocratic: some traffic is more equal than other
- reminds you of your mom: no, you can’t have that 10 Mb/s now
- socialist: administer scarcity - we like SUVs (or to drive 100 mph)!
- “risky scheme”: security
- only displacement applications (such as telephony) need QoS
- requires cooperation: edge-ISP, transit ISPs, end systems
- snake oil: add QoS, lose half your bandwidth

Why is QoS hard?

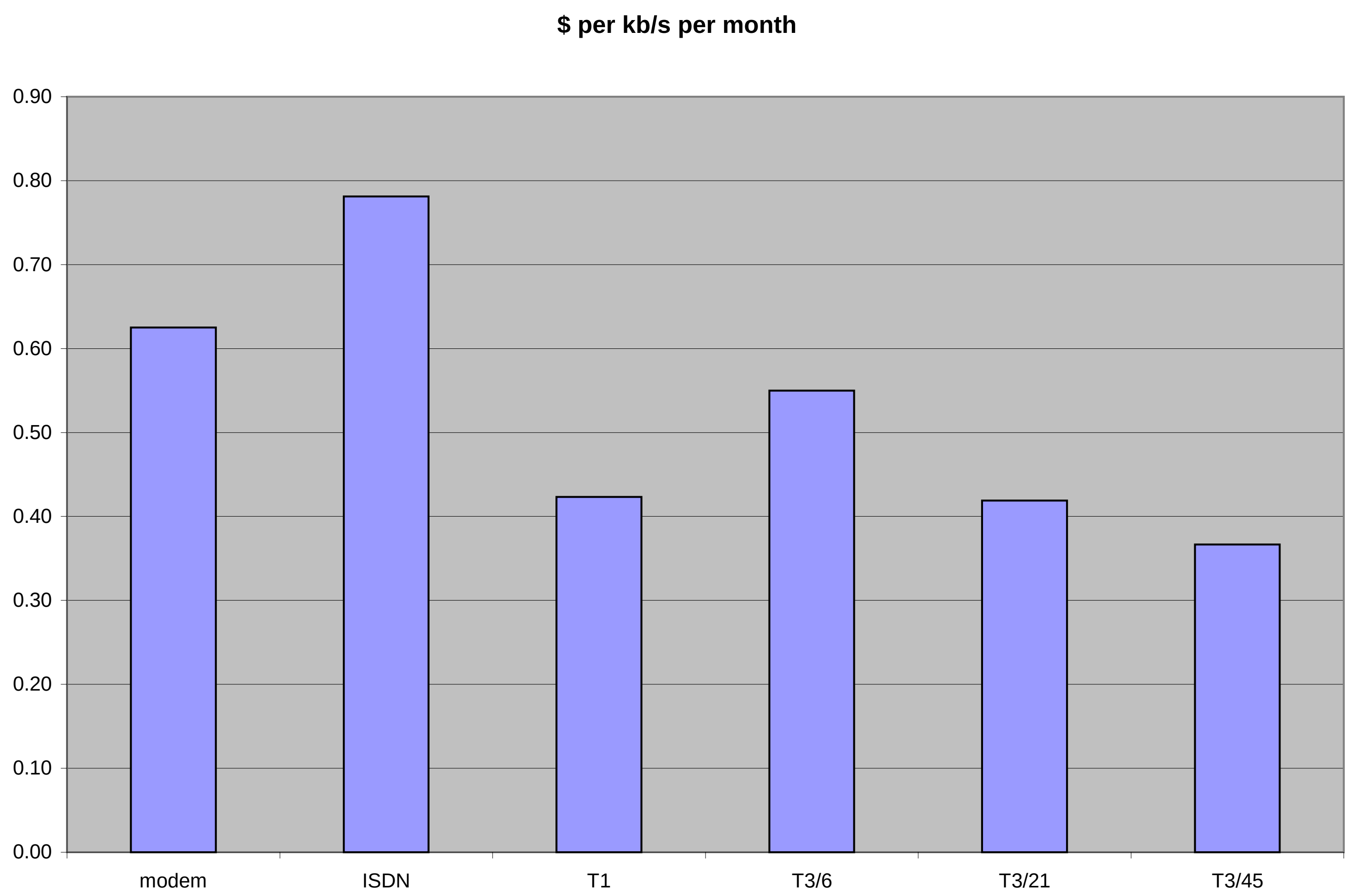
- dishonesty: we only talk about the beneficiaries
- network has become harder to evolve:
  - network address translation
  - firewalls
  - high packetization overhead (VPNs, IPv6)
- to be useful, has to be nearly universally supported ("no, you can't make calls to AS 123")
- network QoS vs. business class model: "coach is empty, please refund fare"
- currently, the ISP interface is IP and BGP – adding a third one is a big deal
- new Internet service model: TCP client (inside) – server (outside)
  - exception: peer-to-peer on college campuses
- network to host: you first, no, you first
- failure of IP QoS $\longrightarrow$ MPLS

Cost of a λ



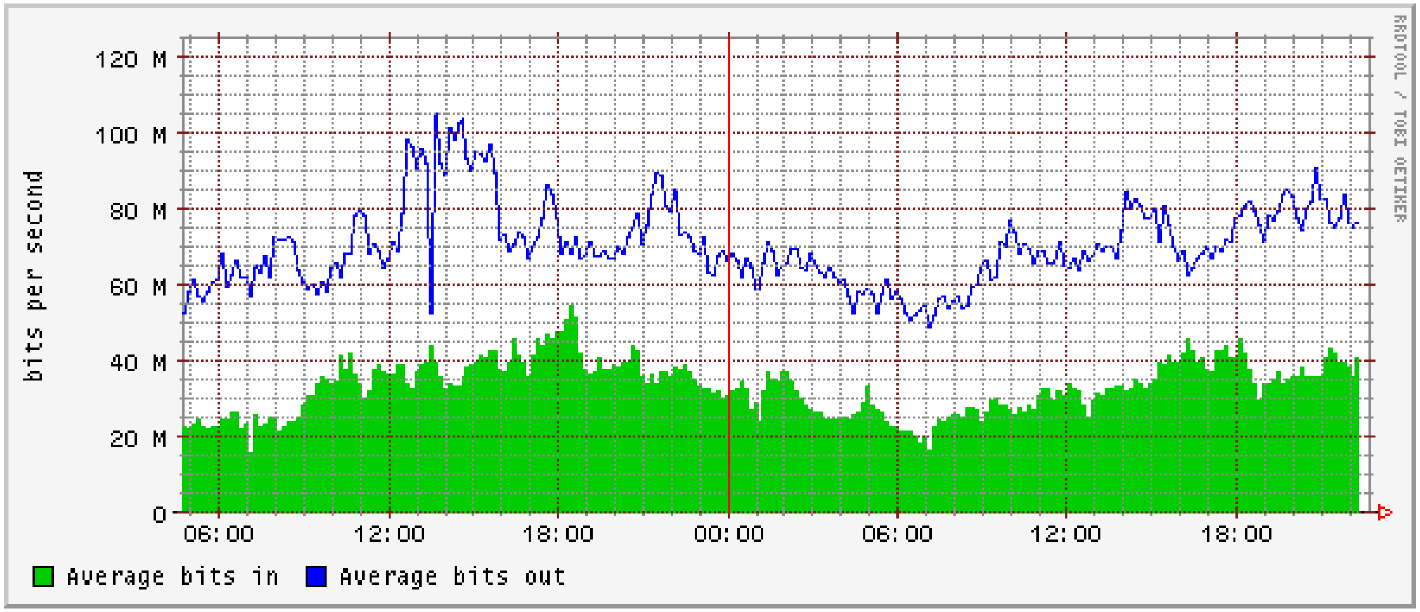
(Data: Telegeography)

Retail cost of bandwidth

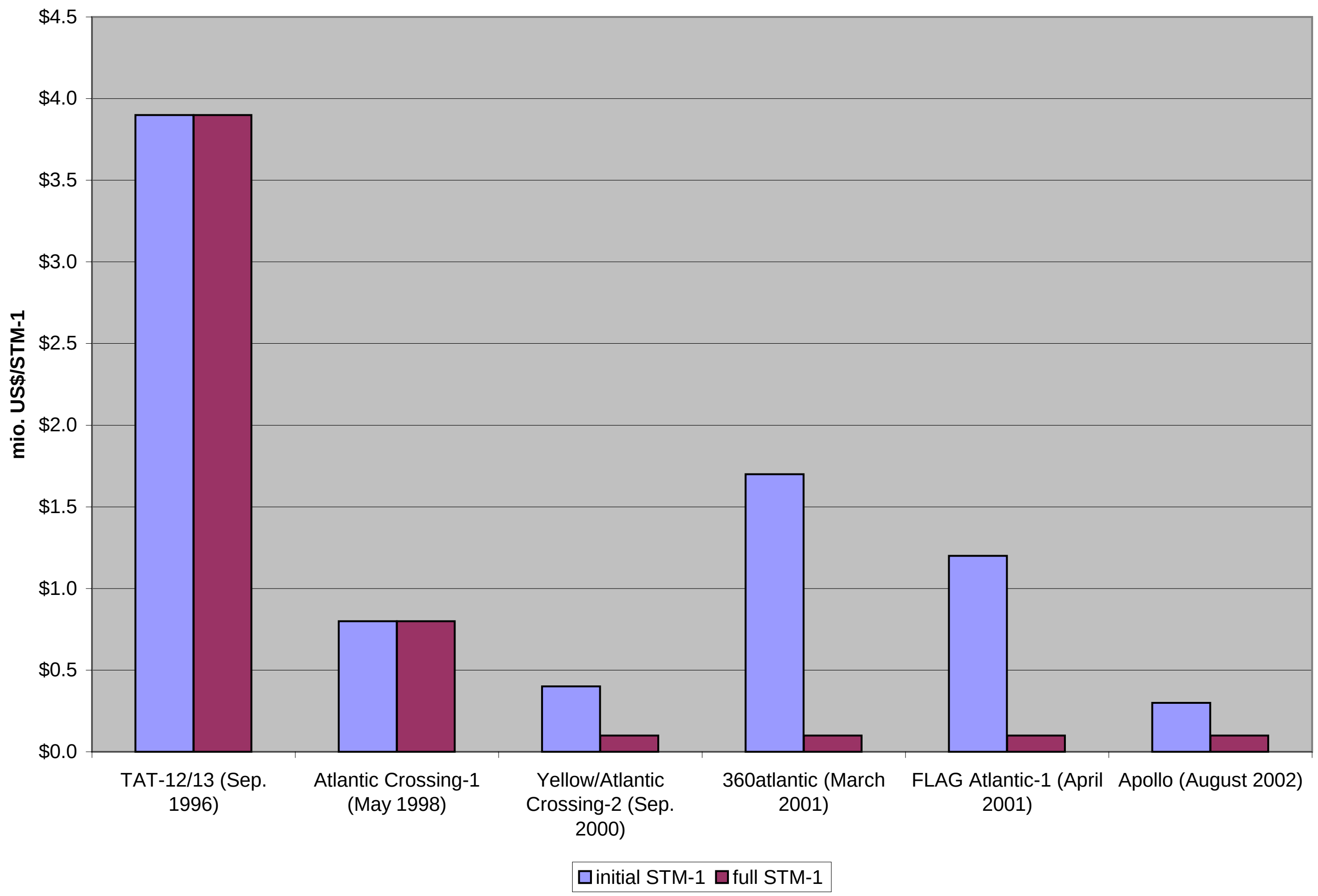


Bandwidth is cheap

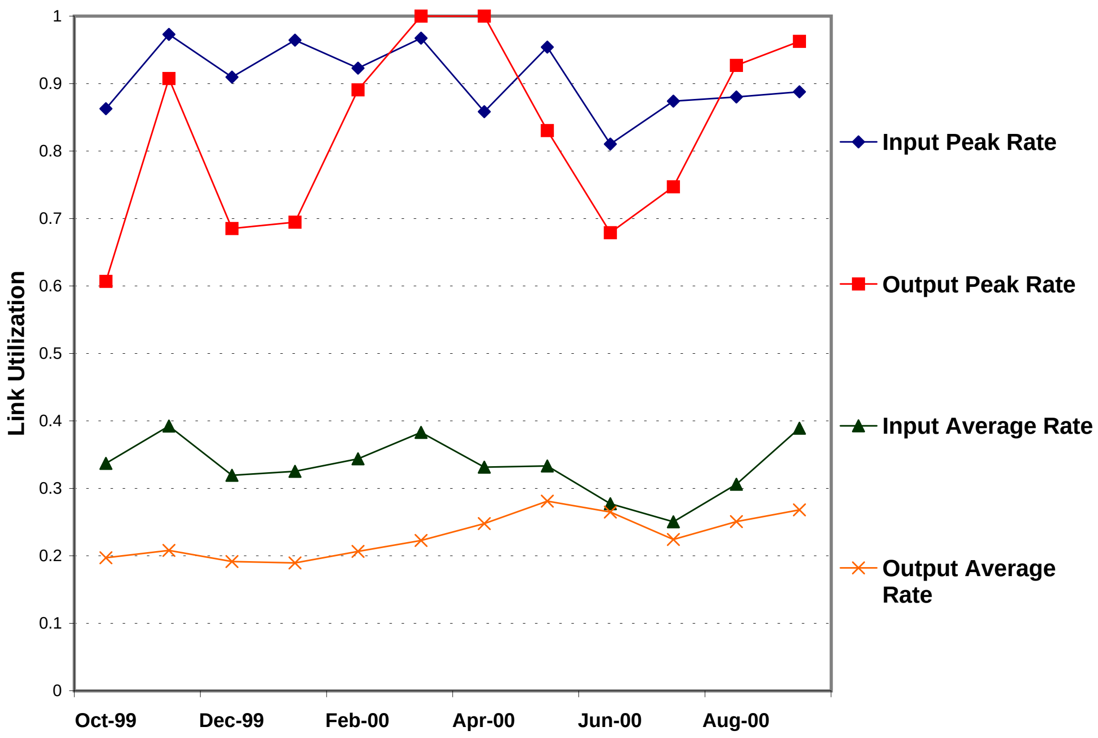
- T3: approx. $25k/month
- OC-n hardware cost rule: step = 4x bandwidth, 2x cost
- assume 50% average utilization
- cost: roughly, 0.2c/64 kb/s channel (vs. 5c/minute for US long distance)



Transatlantic bandwidth



But bottlenecks are plenty…



(NORDUNet)

Voice-over-IP as the savior of QoS?

- streaming media: looked promising, now mostly TCP
- 90-97% of bytes are TCP
- definitely not TCP: VoIP
- 2001: 6.2 billion minutes of international traffic (out of 140 billion total)
- most IP phones already support TOS bytes
- for reservation: embedded systems ➠ simple, self-configuring protocols

Example: Pingtel SIP phone

Example: Cisco and 3Com SIP phones

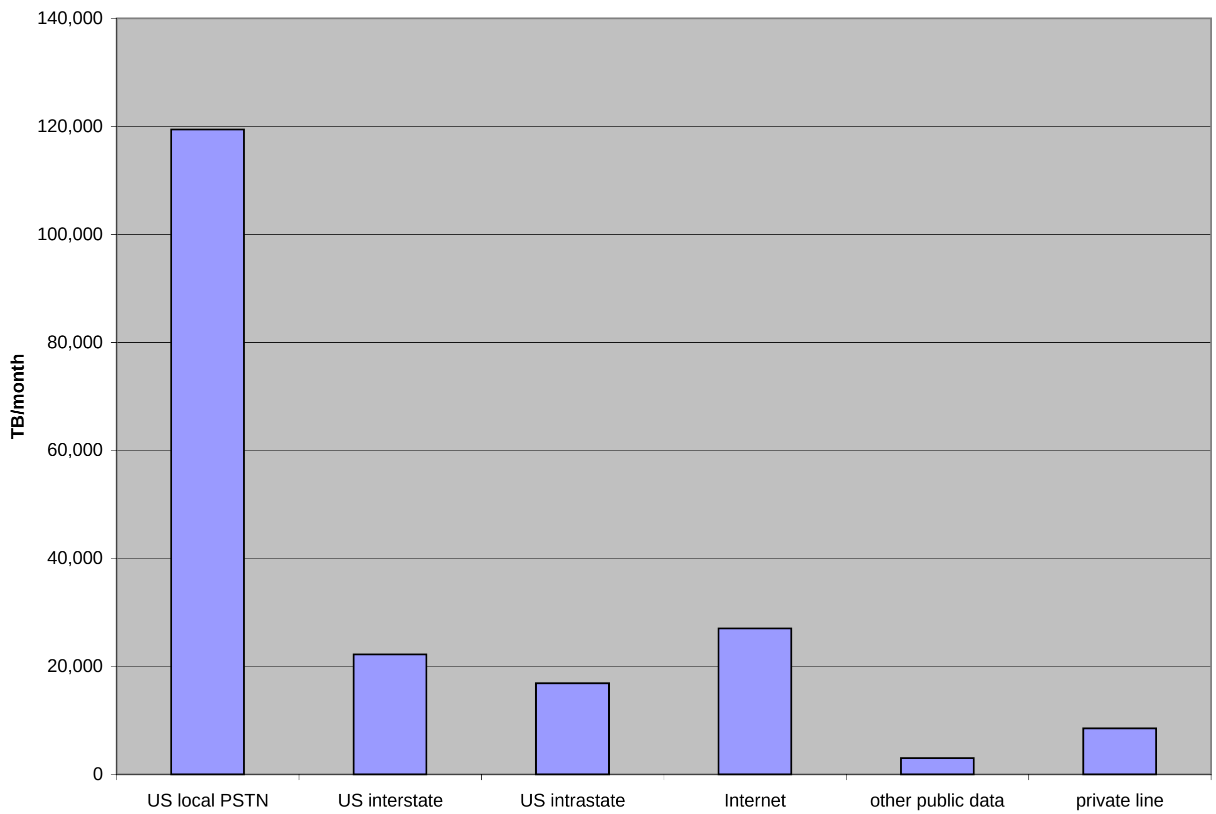
Cisco

3Com

QoS and VoIP

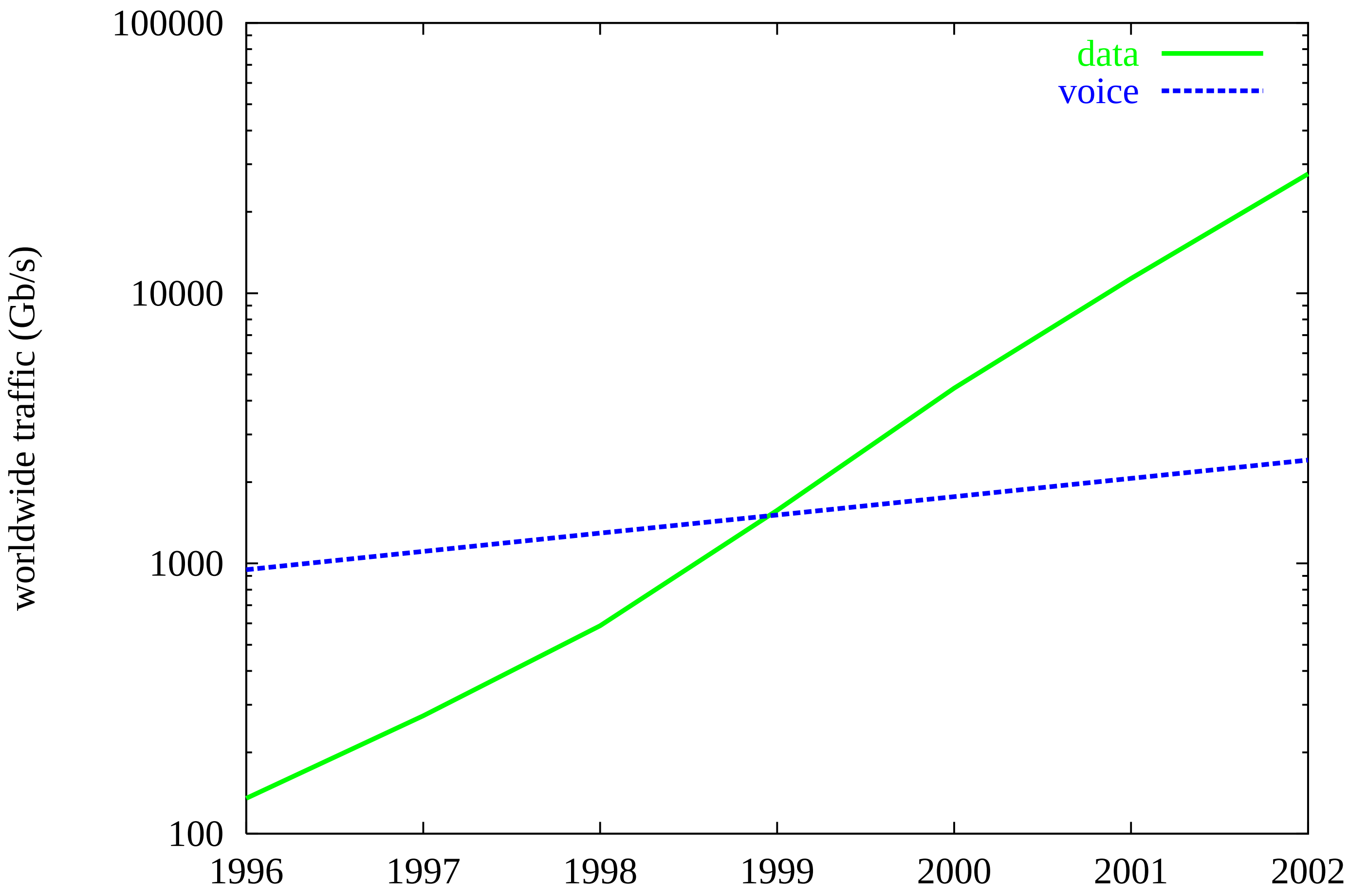
- users are accustomed to paying
- predictable traffic
- suitable for ABE (low delay or high throughput)
- harder: don't have simple LEC–LD–LEC model
- primary consideration: reliability ($<$ 1% call failure/blocking)

Traffic



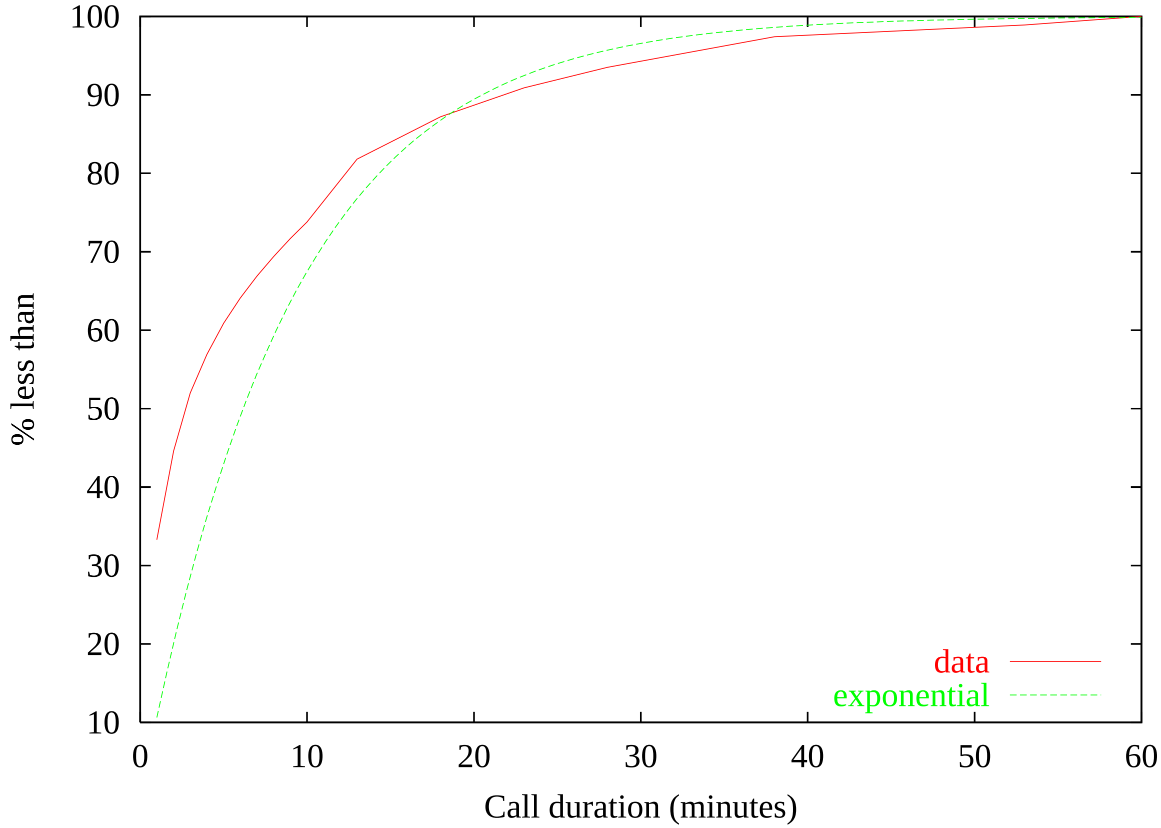
A. Odlyzko and FCC. Data estimates for 2000; PSTN for 1998.

Voice and data traffic

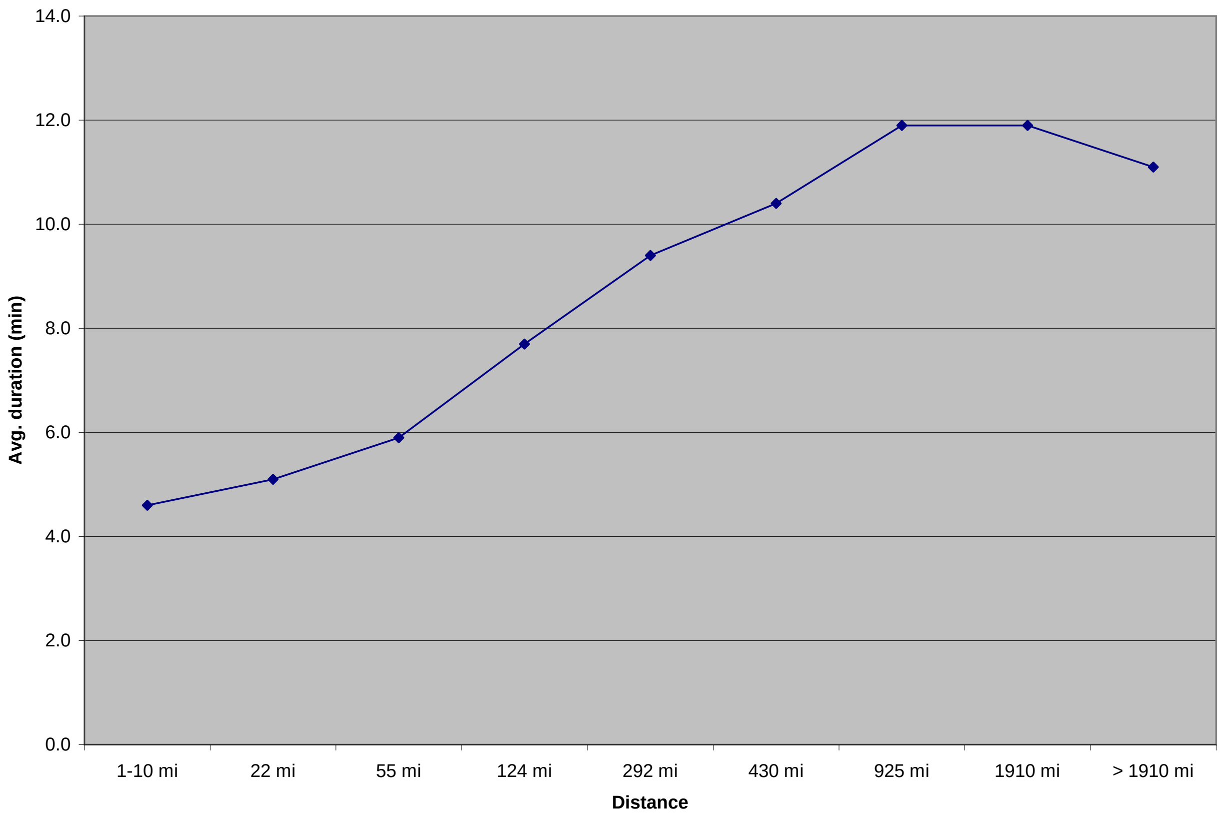


The three-minute myth

Local calls are about 2.4 minutes on average, but long distance calls are much *longer*, about 8.9 minutes:



Calls get longer with distance



average	310 mi	7.8 min.
median	60 mi	3.0 min.

What made other services successful?

VoIP is always happening next year . . .

email: available within self-contained community (CS, EE)

web: initially used for local information

IM: instantly available for all of AOL

All of these . . .

- work with bare-bones and variable connectivity ($\geq$ 14.4 kb/s)
- had few problems with firewalls and NATs
- don't require a reliable network
- work well with PCs

Future traffic

- voice growing at 10%/year, Internet traffic at 100%
- with provisioning cycle, means that you need to run your network at half capacity
- TCP (and kin) is likely to continue to dominate:
  - dominance of dial-up: only 13% DSL and cable modem ➠ applications are tailored for 34 kb/s
  - data on mobile devices (not just web browsing)
  - video P2P
  - movement of private-line networks to Internet VPNs
- interest is in *asynchronous* communication, not interactive
- exception: video games?

Lessons learned

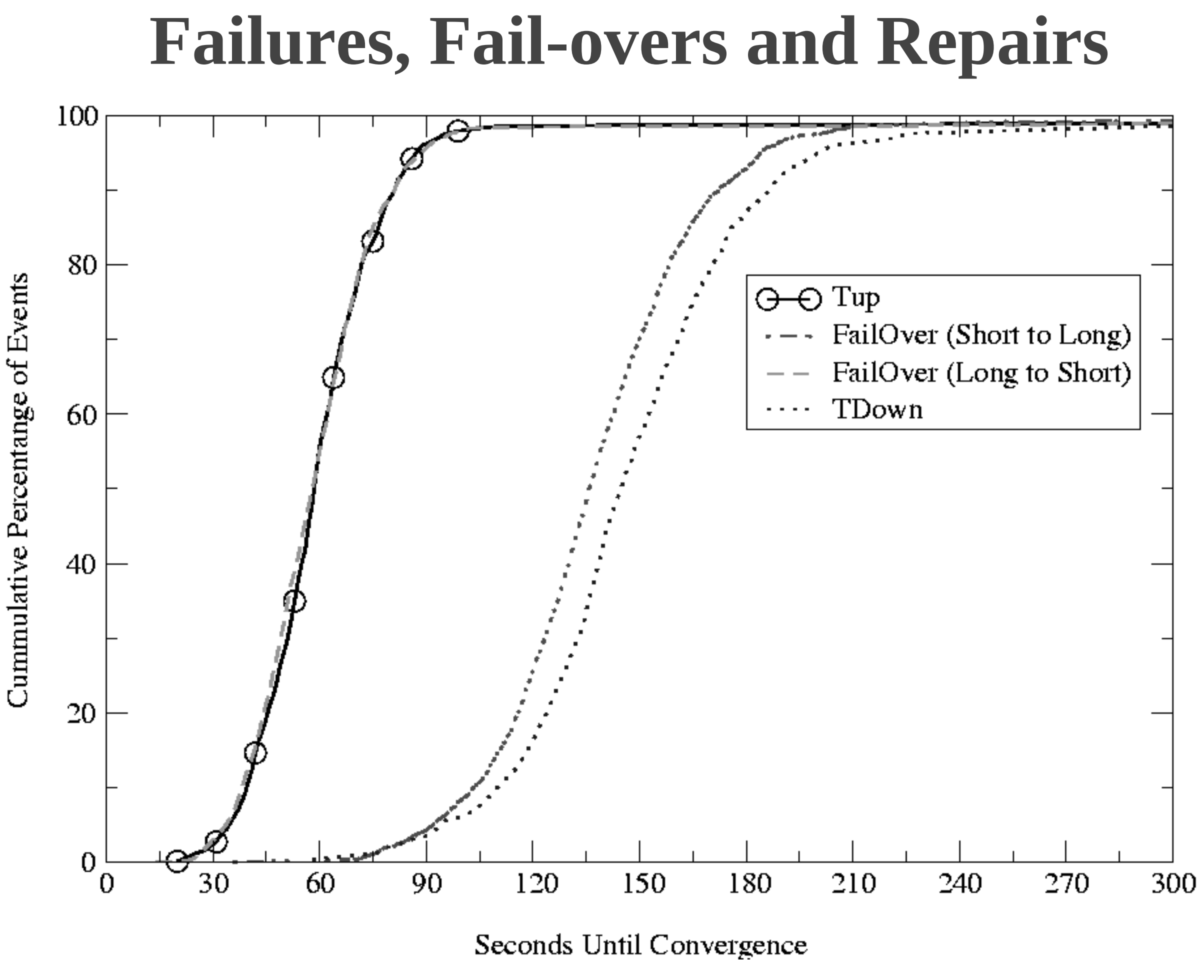
- queueing delay doesn't matter if you have OC-3 links
- generality can hurt:
  - receiver diversity: complexity, not that useful in practice
  - multicast orientation $\longrightarrow$ very little multicast
  - diversity of signaling protocols can be beneficial
- try to support niche applications, e.g., high-quality video conferencing (ISDN replacement)

QoS is about reliability

- typical utilization: 30% – but not during restoration
- can’t sell premium service that’s unavailable one day a year
- cable **will** be cut once per year/200 km
- consistent 5% packet loss is much better than 5% probability that network is unavailable for seconds
- who are you going to ask for a refund?

BGP Convergence Times

(From Abha Ahuja's IETF50 plenary talk and Geoff Huston's talk)



Reliability Issues

- $-$: software updates require “scheduled downtime”
- $+$: but signaling servers can be made redundant much easier than SS7 SCPs
- BGP convergence times of several *minutes*: 2 minutes to withdraw routes, 30 minutes to advertise routes
- “80% of withdraws take more than a minute”
- no clear IP reliability definition – reachability of any node? some large subset? “local calls”?

Have we explored the design space?

		flow	aggregate
signaling	flow	IntServ	–
	aggr.	?	DiffServ + RNAP/BGRP/. . . ?

Congestion pricing is back

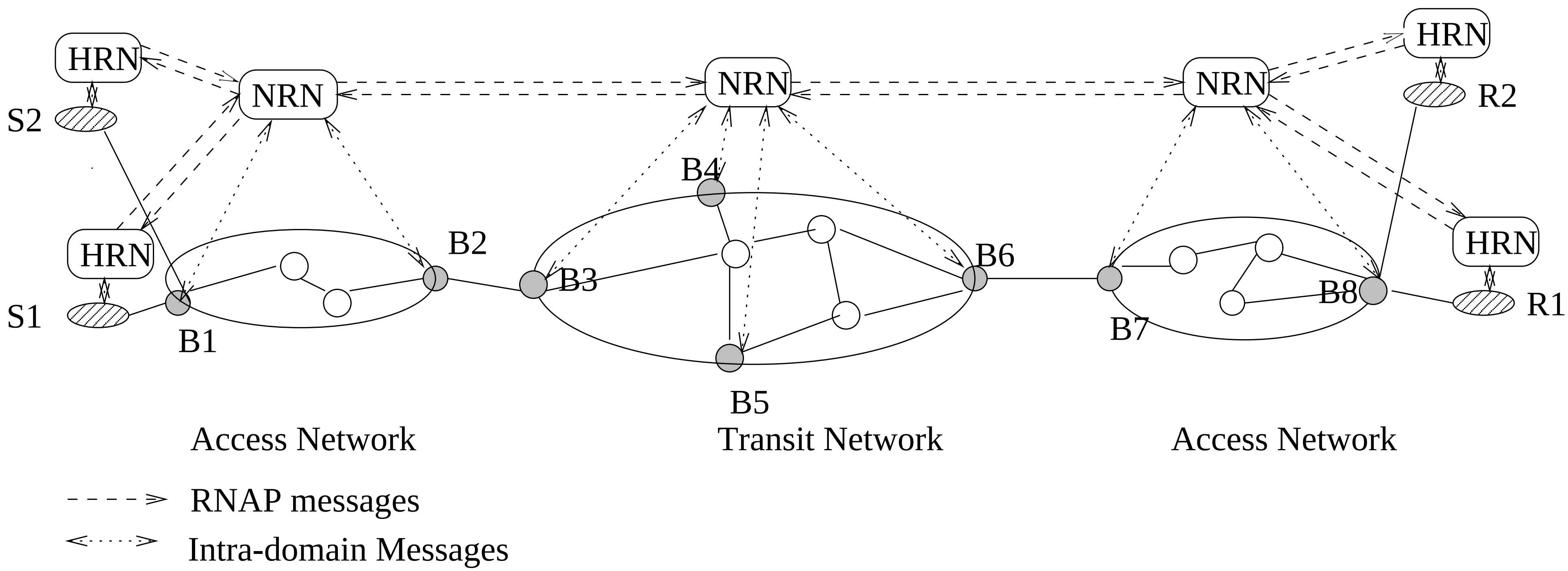
Myth: “people want fixed prices” – but congestion pricing is common:

- Traffic:

> Thanks to advances in technology, it’s becoming easier to use tolls to tame congestion. Solo drivers on Interstate 15 in San Diego, for example, can purchase the privilege of driving on the road’s high-occupancy-vehicle lanes. The price changes as often as every six minutes and depends on the amount of congestion on the HOV lanes. When traffic is light, the toll, which is collected using vehicle transponders and overhead antennas, goes as low as 50 cents; when it’s heavy, it may cost as much as $8.00. If you carpool, the ride is free. (USN&WR)

- electricity, airline and hotel pricing
- many T3’s are priced on 95th percentile of five-minute intervals
- not for 10c items – 1 Mb/s video conferences, not phone calls
- More correctly: people want to know what they’re being charged

Example: RNAP (Resource Negotiation and Pricing)



RNAP

- many applications have a wide operating range
- per-bit value decreases with bandwidth
- provide incentive to use only resources needed
- compute price by tatônnement or M-bid auction
- optimize user perceived value = utility $-$ price
- take “holding costs” into account for multiple classes
- allow range of commitment times, typically minutes
- allow ISP “buy back” if advance purchase price was low

What’s needed?

- forget scalability – most successful applications don’t scale (e.g., AOL IM)
- likely to be only for tiny fraction of traffic (but much larger revenue fraction. . .)
- high-value streams (conferences, VPNs) ➠ advance reservation needed
- prime QoS determinant is path reliability – zero bandwidth is not
- useful measurements – not just monthly averages
- reservations for one or two bottlenecks:
  - access links
  - wireless links
- single sign-on (e.g., Microsoft Passport)

The myth of TCP fairness

- = non-TCP flow should not send more than a TCP flow under similar network conditions
- over what time scale – short vs. long connections
- *users* pay for network access, not flows – longer-term user benefit equalization?
- should a multicast stream benefiting a million users get the same throughput as a single TCP stream?
- parallel connections to one or several web servers
- quirks: TCP throughput decreases with RTT
- very little UDP traffic is TCP-friendly

(see also Christophe Diot, J.Y. Le Boudec)

Challenges

- optimizing components $\longrightarrow$ buildable system
- roughly standardized service classes – allow settlements and advertising
- separation intra-domain vs. inter-domain
- security – new opportunity for massive DOS
- integrated services ➠ use second channel (PSTN)
- useful busy-hour measurements, not just monthly averages
- reliability measurements
- simple, manageable signaling with advance reservation

Conclusion

- from central issue ➠ niche application
- QoS is primarily reliability, then packet loss, then queueing delay
- sell reliability and predictability, not delay jitter
- need BGP fix plus “nailed down” circuits:
  - routes must stay in QoS-enabled network
  - faster recovery – try backup route on failure
  - doesn’t have to scale
  - two types of carriers: “classical IP” vs. “voice heritage”?

Cost of Bandwidth (per km)

cable laying	$65,000
conduit laying	$15,000
conduits, manholes	$ 5,000
24 fiber pair cable	$ 7,000
72 fiber pair cable	$16,000
DWDM, 80x10 Gb/s (per λ)	$ 550
10 Gb/s OC-192 terminal (per λ)	$ 1,200
optical amplifier	$ 300

(Source: Telegeography)