The MBone —

Interactive Multimedia on the Internet

Van Jacobson

Lawrence Berkeley Laboratory
Berkeley, CA 94720

UC Berkeley
Feb 17, 1995

A widely believed myth:

Real-time traffic like audio and video requires a connection-oriented, virtual circuit network. Datagram networks can’t be used for real-time traffic because:

- they don’t have the state necessary to meet real-time scheduling and delivery constraints and
- IP delivery is ‘best effort’ so transit times are unbounded and vary wildly.

But . . .

There has been world-wide, IP-based, real-time conferencing over the Internet (via the MBone) in daily use for the past three years.

20,000 users on 1500 networks in 30 countries may be disappointed to learn that their meetings, seminars, social events, etc., are a mass delusion.

Over the past 20 years there have been many attempts at connection-oriented approaches to conferencing.

- Most have been dismal failures.
- None have worked well.

There are reasons for this:

- Large user burden — have to know other participants and details of network topology.
- Intolerant — difficult to join in-progress conference.
- Poor scaling — if everyone can talk and listen there are $O(n^2)$ connections for n participants.
- Unreliable — conference fails if any of the n^2 connections fail.

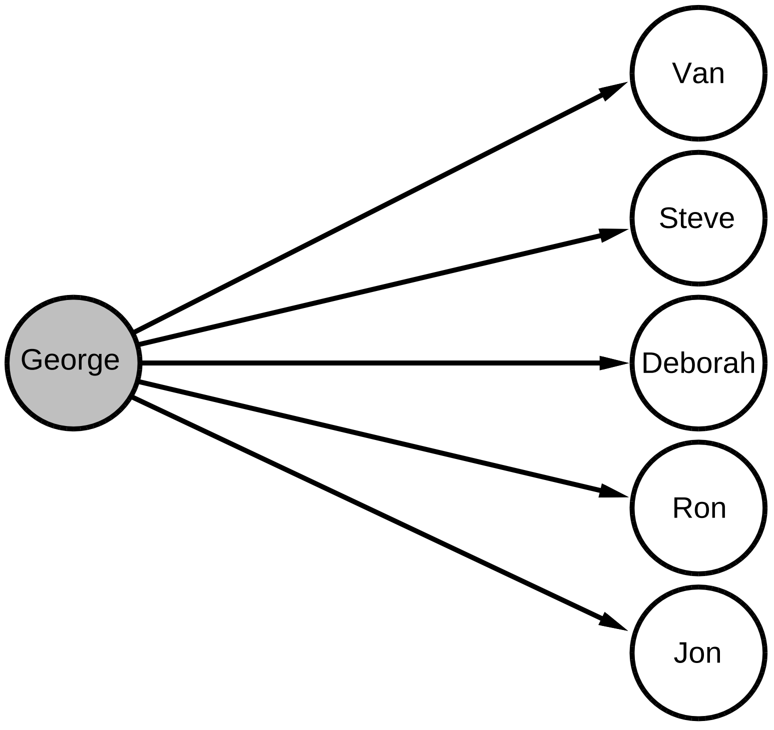
IP scales and works well because it has a very clear separation of roles:

- End-nodes know nothing about topology.
- Routers know nothing about ‘conversations’.

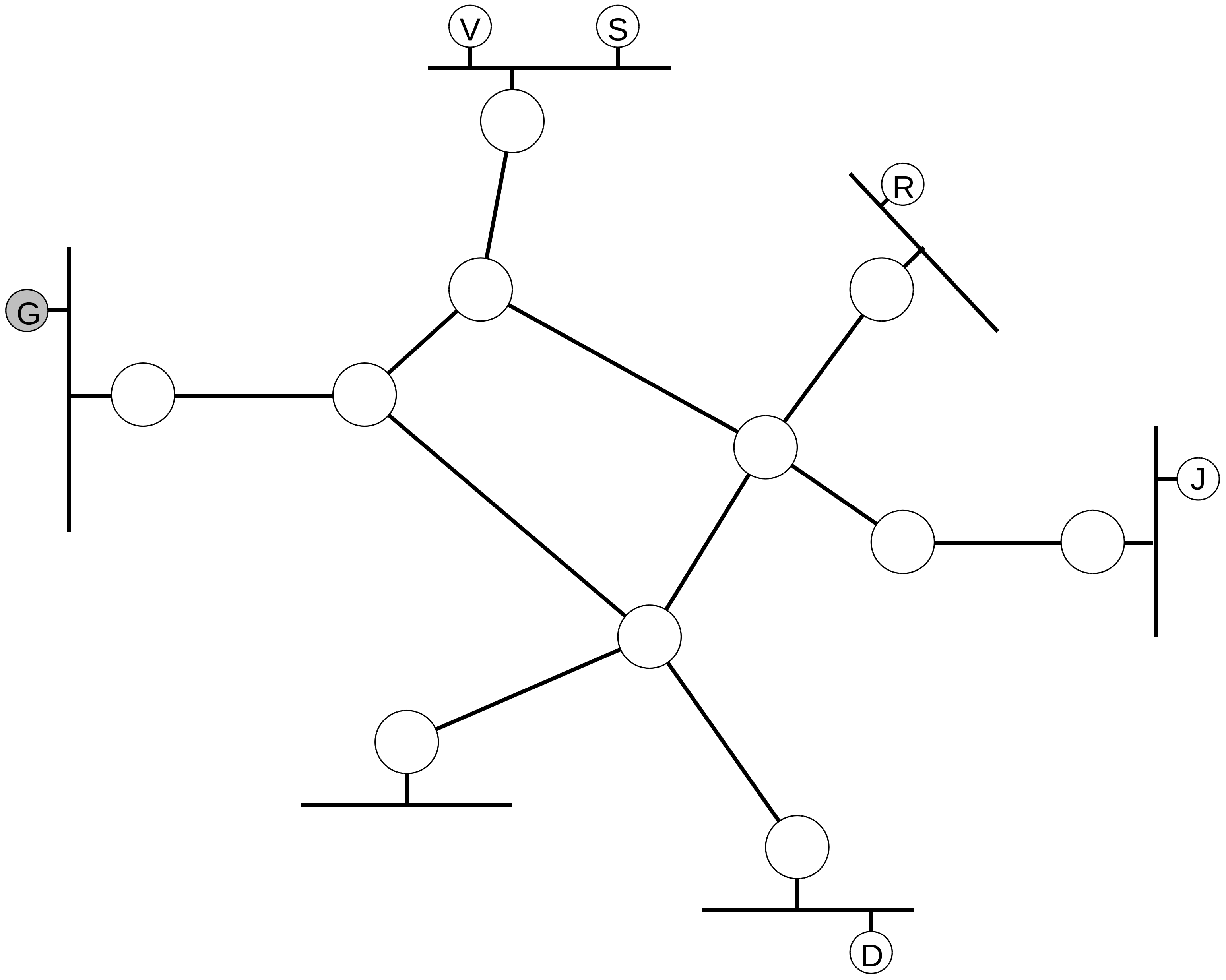
First item allows net to dynamically change delivery topology.

Second means that changes can be done without global coordination (no end-to-end state to move around).

But separation of roles makes multipoint delivery hard. I.e., sender thinks the world looks like this:



When in fact the topology is:



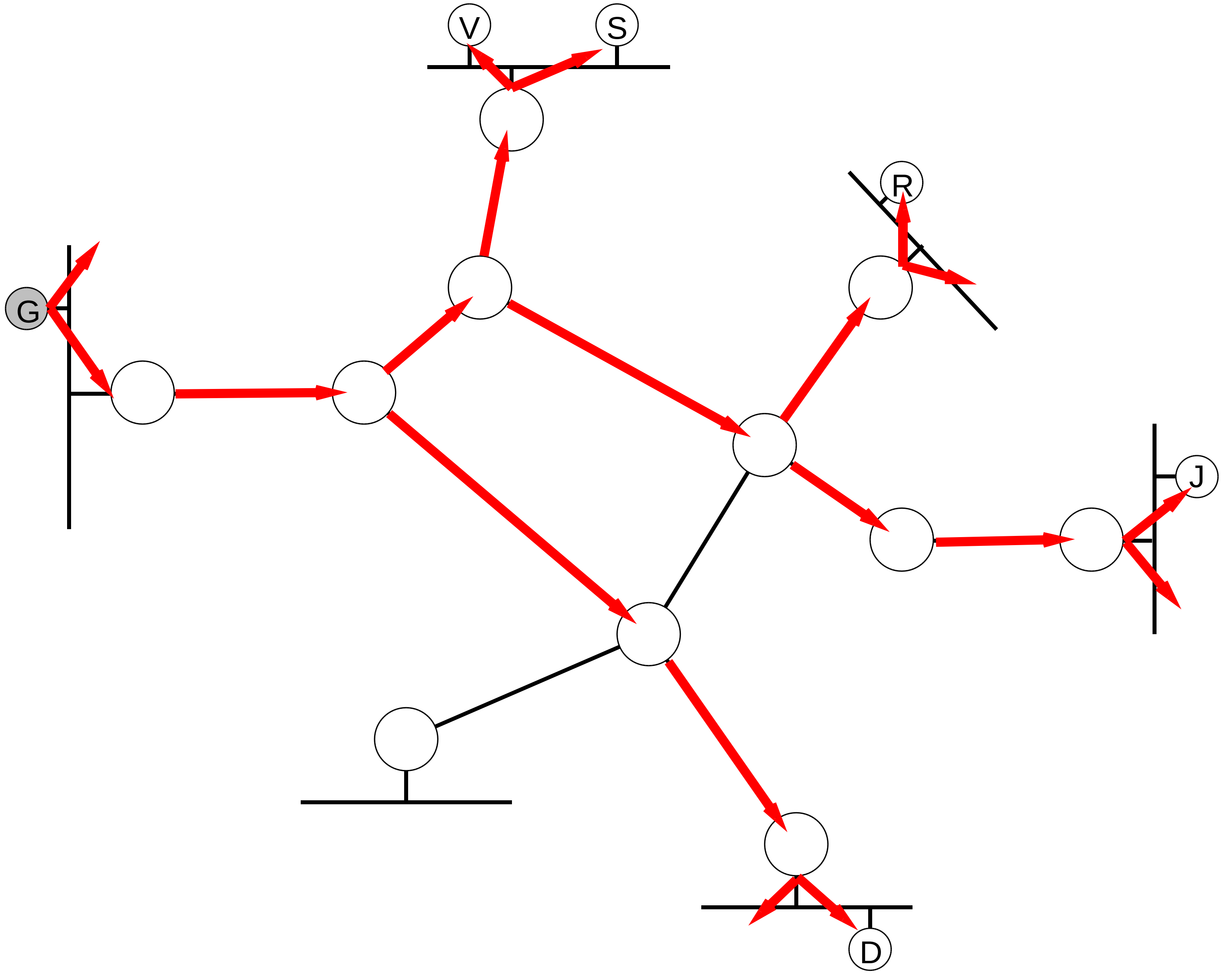
Core problem: How to do efficient multipoint distribution (i.e., at most one copy of a packet crossing any particular link) without exposing topology to end-nodes.

Solution: Deering’s IP Multicast

Dynamically constructs efficient delivery trees from sender(s) to receiver(s).

Very simple service model:

- Receivers announce interest in some multicast address
- Senders just send to that address
- Routers conspire to deliver sender’s data to all interested receivers.


V
S
G
R
J
D

Note that IP Multicast is more than a efficient, simple, robust, delivery mechanism. It also greatly simplifies the conferencing problem for applications and users.

I.e., if we associate a ‘conference’ with a multicast address, then

- Users can join the conference without enumerating (or even knowing) other participants.
- Users can join and leave at any time.

Since the conference has a network visible identity, the net takes care of all the hard problems that plague connection-oriented conferencing (rendezvous, efficient distribution and dynamic membership).

This means we preserve the robustness of IP and gain the simplicity and group-size-independent scaling of IP multicast.

Some MBone Chronology

Nov 88 Small group (MIT, BBN, UDel, ISI, SRI, PARC, LBL) led by Bob Braden of USC/ISI proposes testbed net to DARPA. This becomes DARTNET (DARPA Research Testbed Net).

Mar 90 Routers and T1 lines in place.

Nov 90 Routers and T1 lines start to work.

Feb 91 First packet audio conference (using ISI’s vt).

Apr 91 First multicast audio conference.

Jun 91 Regular weekly conferences.

Aug 91 McCanne writes first version of vat.

Sep 91 First audio + video conference (hardware codec).

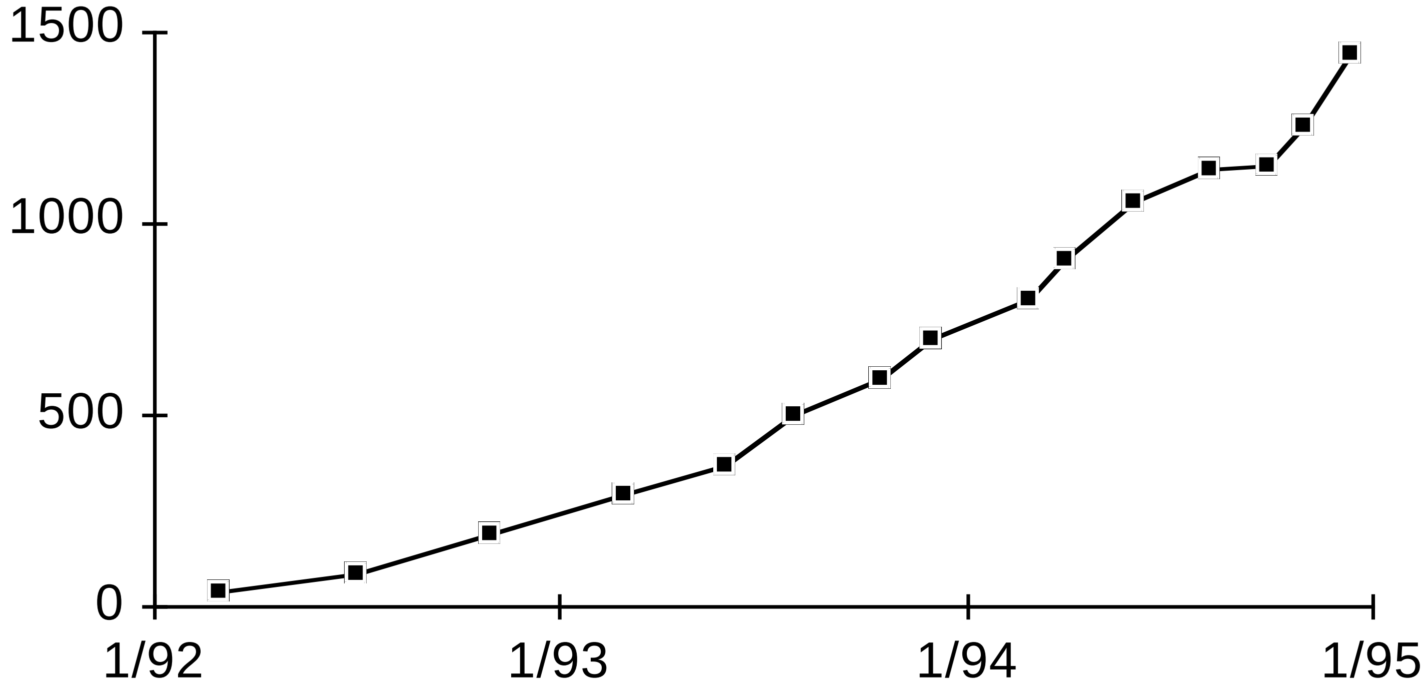
Mar 92 Deering & Casner broadcast San Diego IETF to 32 sites in 4 countries.

Oct 92 McCanne writes first version of wb.

Dec 92 first nv (from Ron Frederick), first sd.
Washington DC IETF – four channels of audio and video to 195 watchers in 12 countries.

Jan 93 MBone events go from one every 4 months to several a day.

Growth of the MBone
(number of subnets)



The MBone is growing exponentially, like the Internet, but almost twice as fast (a doubling time of around 8 months).

But, since the MBone is still an ‘overlay’ on the Internet (multicast routers are distinct from normal, unicast routers), it’s not trivial to get hooked up and requires cooperation from local and regional network people.

All that is about to change. Most major router vendors now support IP multicast (e.g., cisco just released it with v10.2).

In the next year we should see the MBone go away as a distinct entity as we evolve to ubiquitous multicast support throughout the Internet.

That will mean that anyone hooked to the Internet can participate on conferences with everyone else hooked to the Internet.

Resource Reservation

There are three different reasons for doing resource reservation:

- Interarrival variance reduction / jitter control (Does network preserve detailed time structure of sender’s traffic?)

- Capacity assignment / admission control (Is there enough bandwidth available, on average?)

- Resource allocation (Who gets bandwidth?)

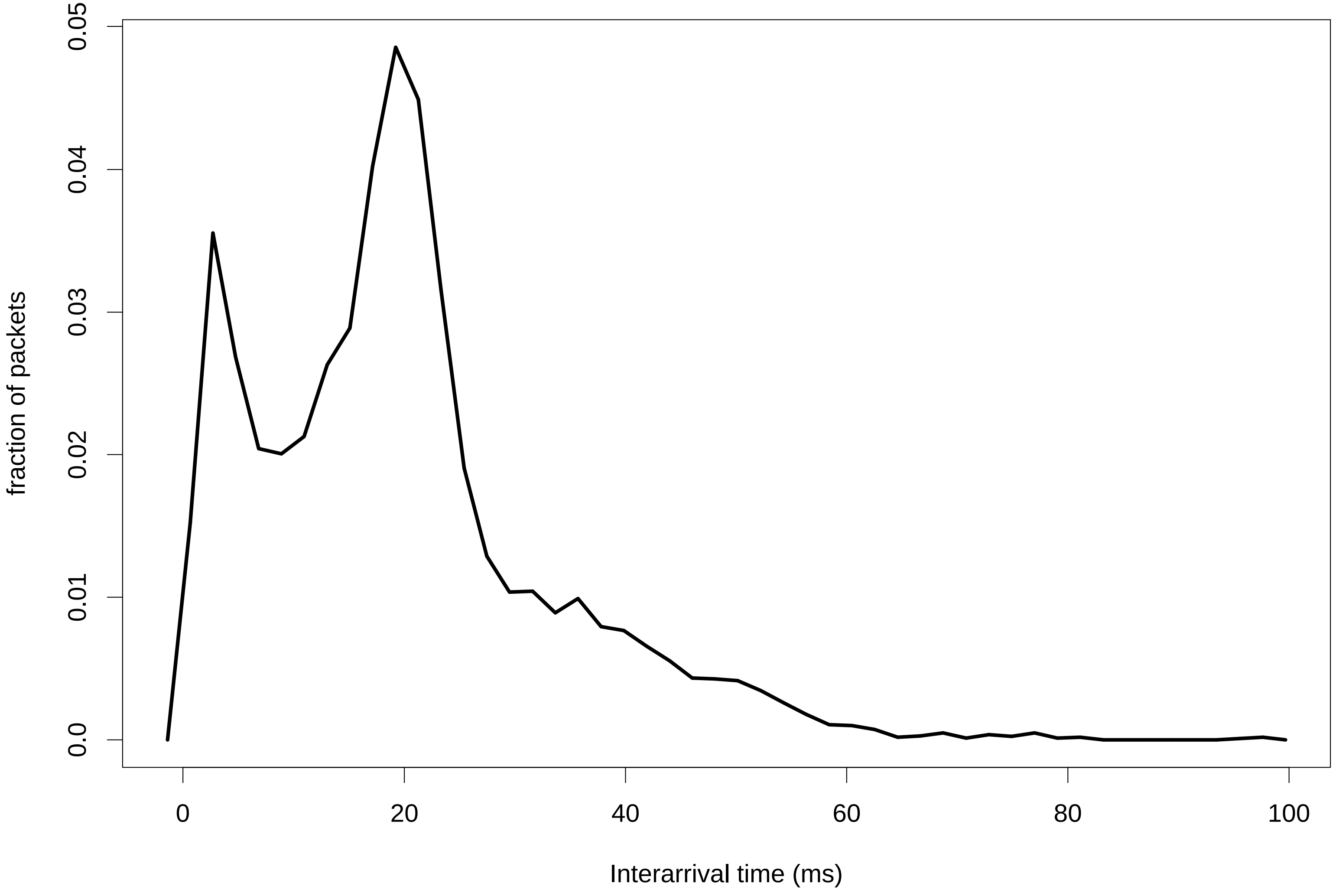
Jitter Control

- If network has enough capacity for conversation, average departure rate at sender must equal average arrival rate at receiver.

- Only contribution to jitter is queue waits due to competing traffic.

- Queue waits can be at most amount of competing data in transit and total amount of in-transit data should be at most round-trip propagation time (e.g., 100ms for transcontinental path).

Typical audio packet interarrival time distribution



You can take out jitter

- by scheduling packets so they don’t interfere, or
- with a buffer at receiver.

Distributed scheduling problems are NP-hard and system won’t work unless everything in path adheres to schedule.

For transcontinental voice conversation (64kbits/sec PCM) worst case jitter is on the order of 8kB/s $\times .1$sec $=$ 800 bytes.

Is it really necessary to solve NP-hard scheduling problem then rebuild network from scratch just to save 800 bytes of buffer?

Capacity Assignment

Model is that end-nodes ask net if there’s enough free bandwidth for average rate of conversation. Can get ‘yes’ or ‘no’ (busy signal) answer.

Deering points out that the answer must almost always be ‘yes’ or service won’t be used. It shouldn’t require complex machinery to almost never say ‘no’.

Why would net answer ‘no’? I.e., what limits available bandwidth?

Economics of bandwidth

At turn of the century there were approximately 50 transcontinental wires available for long distance calls, one call per wire.

Complex technical, economic and regulatory systems were evolved to manage this precious bandwidth resource.

Since the 1960s, the technology has changed radically, the economics slightly and the regulations (e.g., tariffs) hardly at all.

Technology of bandwidth

The US has about 126 million phones. One phone conversation uses 64Kbits/sec so total potential demand is 4×10^{12}b/s.

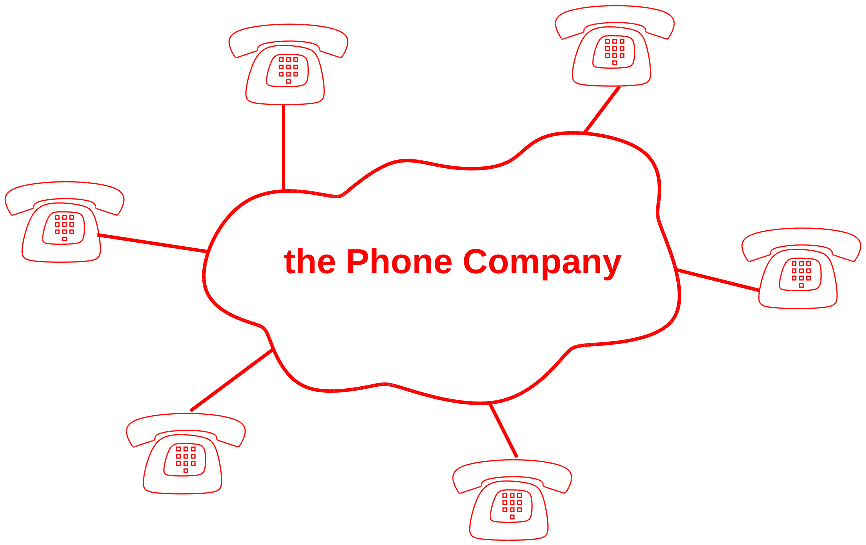
One optical fiber has a bandwidth of 25×10^{12}b/s.

There are well over 1000 transcontinental fibers and new fiber being installed at a rate of several miles per day.

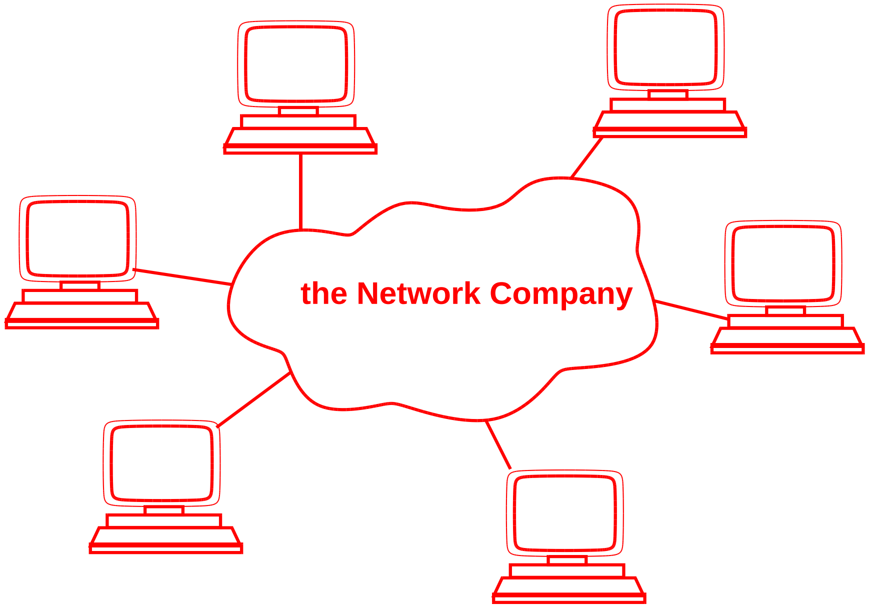
At some point in the future, bandwidth has to stop being a problem.

Resource Allocation

Most virtual circuit (e.g., ATM, ST-II) systems modeled on pre-divestiture telco model. All transactions are a simple, bilateral communication user → telco or telco → user.



When the world is neatly divided into ‘the users’ and ‘the net’, it’s easy to claim the resource allocation problem is just putting some sort of limiter in place to ‘keep bad guys from swamping the net’.



There isn’t ‘the net’ — there’s an Internet with thousands of nets, all run by different organizations.

Communication isn’t between ‘users’ and ‘the net’, it’s between talkers and listeners.

Since a multicast talker can send nothing unless there’s at least one receiver, ‘stopping bad guys’ is already part of the architecture — don’t listen and they go away.

Since multicast traffic is **always** associated with an act of communication (someone talking plus someone listening) the resource management problem is not about stopping abuse but about choosing who gets to communicate.

This is a difficult social problem and we have no solution for it.

When (if) we have a model for how to deal with the social problem, the associated technical problem is trivial.