CERFnet Seminar:

MBONE—the Multicast Backbone

Videoconferencing Over the Internet

March 3, 1993
San Diego Supercomputer Center

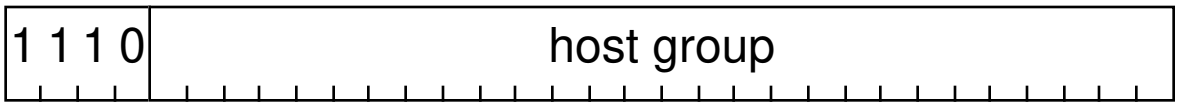
Stephen Deering
Xerox Palo Alto Research Center
(deering@parc.xerox.com)

Copyright © 1993 Xerox Corporation
All Rights Reserved

IP Multicasting — What Is It?

An extension of IP to allow datagrams to be sent to groups of hosts.

- group members may be anywhere in internet, may join and leave at will
- groups identified by Class D IP addresses (permanent or transient):



- senders need not be group members
- senders can limit scope of transmission using IP TTL, e.g., for “expanding-ring search”

(except for scope control, it is the internet analog of Ethernet multicast)

IP Multicasting — What's It Good For?

(1) more efficient, lower delay delivery of same data to multiple destinations (compared to multiple unicasts).

e.g. text or multimedia conferencing

mailing-list / news / software distribution

updates to replicated file systems or databases

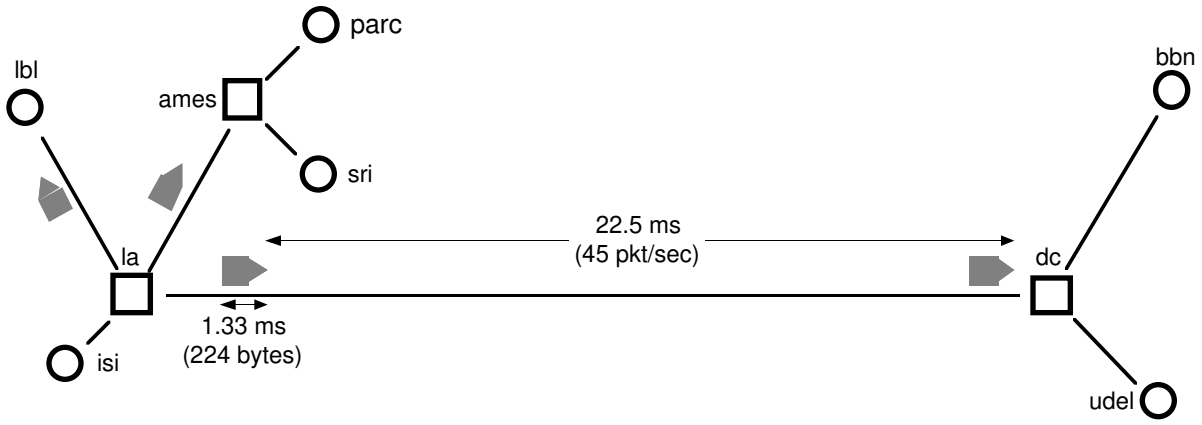
(2) simpler, more robust delivery to unknown or frequently-changing destinations (compared to directory services or configuration files).

e.g. soliciting or advertising services

reporting to changing set of monitoring stations

tracking migrating processes

Animated Example of IP Multicast Audio
(on videotape)



trace-driven animation

- program by Steve McCanne & Van Jacobson, LBL
- traces recorded by Van Jacobson on DARTnet, August '91

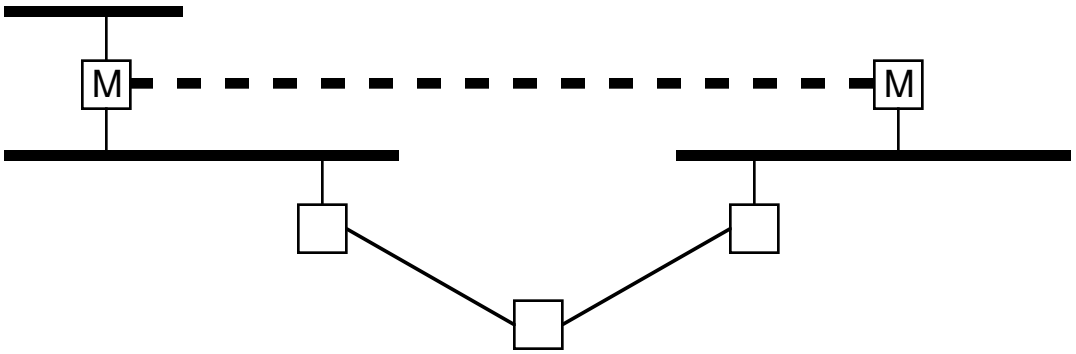
T1 links, Sparcstation routers
(local hosts and Ethernets not shown)

link length proportional to propagation time
pkt length proportional to duration of pkt on wire

animation 1: 100x slowdown of a 1.3 second interval
animation 2: 250x slowdown of a 0.3 second interval

in one example, multicast consumes 1/8 of T1 bandwidth;
unicast would consume 1/2 of T1 bandwidth

Multicast Tunneling



normal multicast packet

original source	multicast dest.	

source-route tunneling

original source	tunnel dest.	LS opt	tunnel source	multicast dest.	

encapsulation tunneling

tunnel source	tunnel dest.		original source	multicast dest.	

MBONE — the Multicast Backbone

A “virtual internet” overlaid on existing Internet, to carry IP multicasts

- consists of subnets that have attached multicast routers, plus interconnecting tunnels.
- to become one-to-one with physical Internet, eventually (we hope!)
- multicast routers (Unix workstations) provided both by IP service providers and by user sites

Evolution of the MBONE (see maps at end of notes):

March ’92 — San Diego IETF
40 subnets, 4 countries

July ’92 — Boston IETF
90 subnets, 10 countries

November ’92 — Washington, DC IETF
190 subnets, 12 countries

today?

Metrics, Thresholds, and TTLs

Each link (i.e., subnet or tunnel) in the MBONE is assigned:

a **metric**

- represents the “cost” or “length” of the link
- multicast routing algorithms compute delivery paths with minimum sum-of-metrics

a **threshold**

- the minimum TTL needed for a packet to be forwarded on the link
- used (as a hack) to enforce administrative or resource boundaries

Each multicast packet carries:

an IP **time-to-live (TTL)**

- decremented by one at each multicast router (normal IP processing)
- if less than a link’s threshold, packet is dropped

Metrics, Thresholds, and TTLs (cont.)

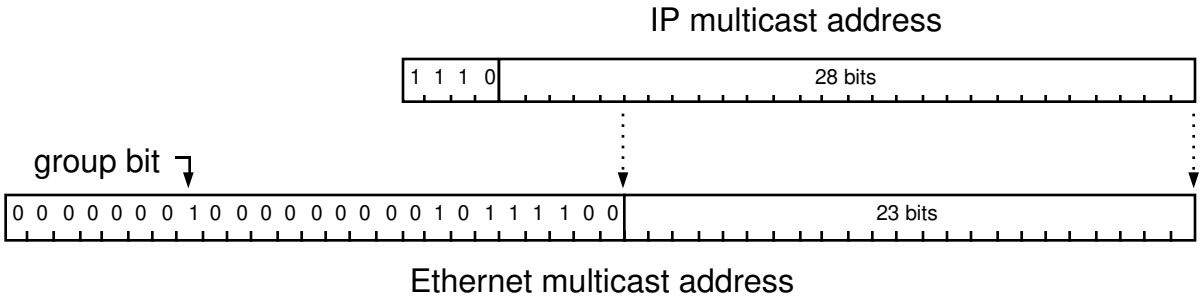
Table of TTLs and Thresholds for Nov. '92 IETF

	peak rate	TTL	thresh
IETF chan 1 audio, GSM	16 kbps	255	224
IETF chan 2 audio, GSM	16 kbps	223	192
IETF chan 1 audio, PCM	70 kbps	191	160
IETF chan 2 audio, PCM	70 kbps	159	128
IETF chan 1 video, nv or ivs	128 kbps	127	96
IETF chan 2 video, nv or ivs	128 kbps	95	64
local event audio, PCM	70 kbps	63	32
local event video, nv or ivs	128 kbps	31	1

Multicast Address Mapping

IP multicast addresses algorithmically mapped to link-layer multicast addresses.

e.g., Ethernet mapping:



note: many-to-one mapping; collisions possible

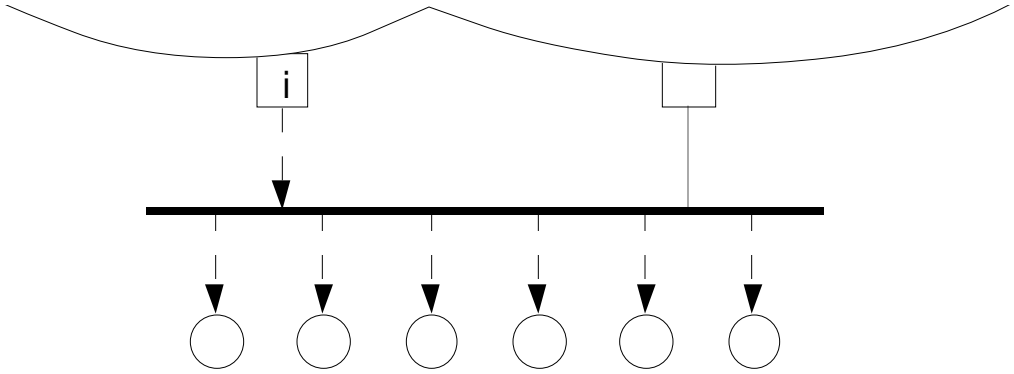
no mapping required for point-to-point links

unicast-only, multi-access subnets treated as p-to-p links by multicast routers; not support by hosts, currently

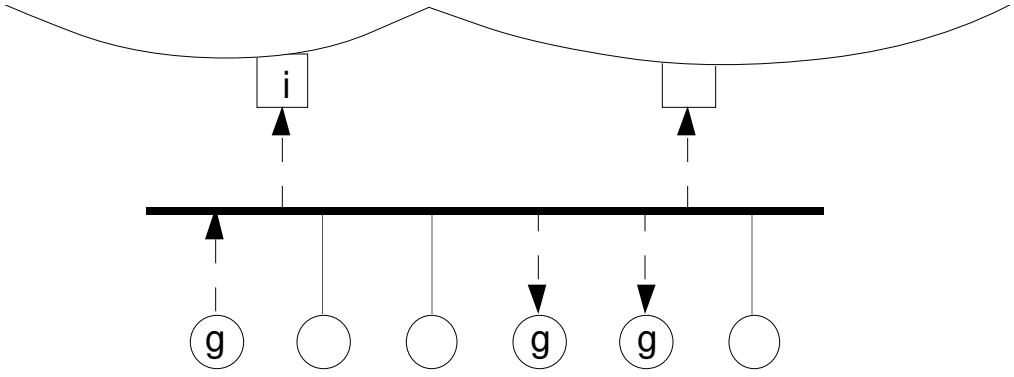
Internet Group Management Protocol (IGMP) RFC-1112

Used by multicast routers to learn what groups are present on attached subnets:

- one router elected “interrogator”, periodically sends Host Membership Query to all-hosts group (TTL 1)



- hosts start random timer for each group membership
- when first timer for group G expires, host sends a Host Membership Report to group G, with TTL of 1



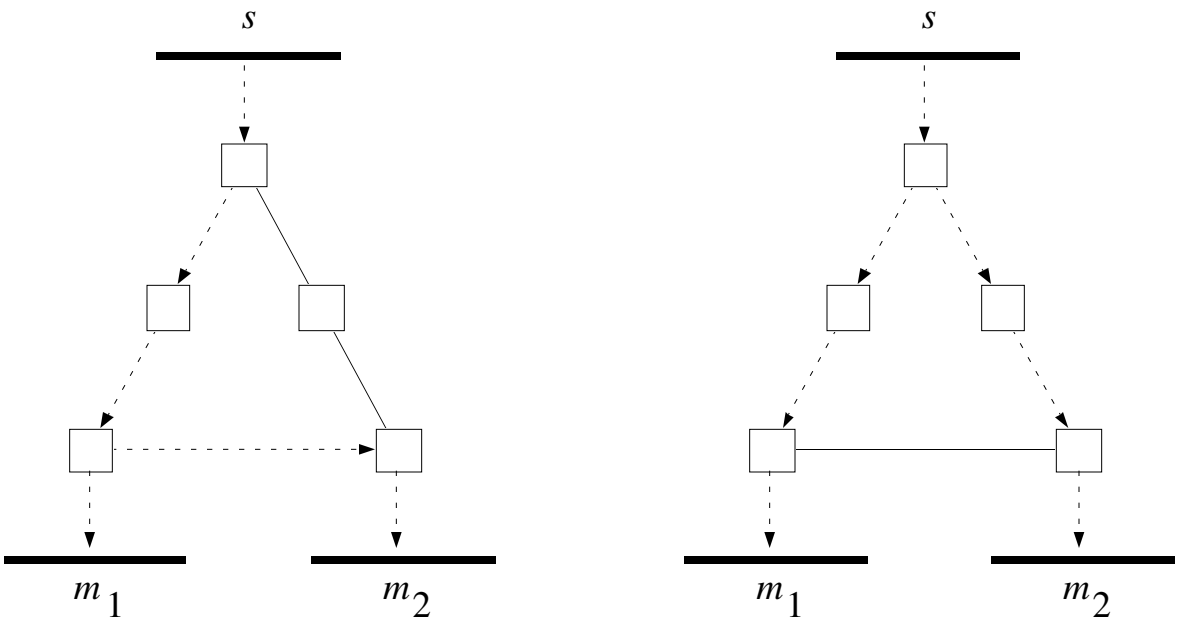
- other members of G hear the report, stop their timers
- routers hear all reports, and time out non-responding groups

Multicast Routing

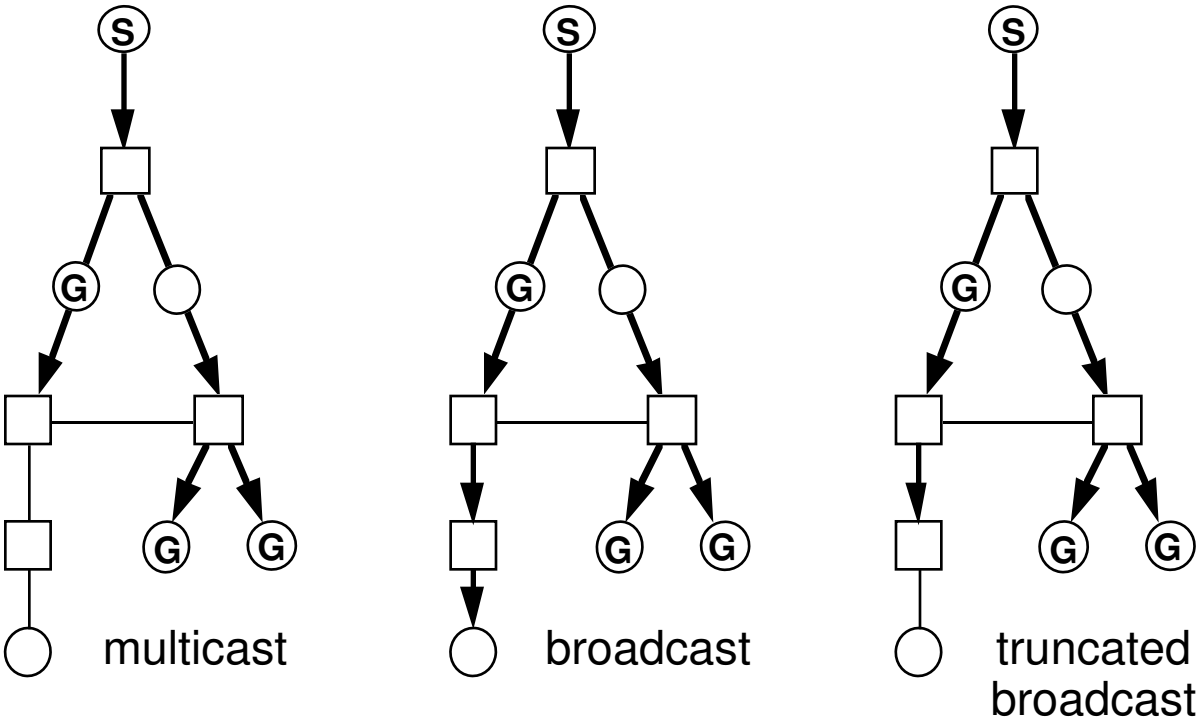
IP Multicast service model & IGMP are compatible with many approaches to multicast routing

e.g.:

- single spanning tree (like LAN bridges)
- minimum spanning tree
- shortest-path trees



Different Multicast Delivery Schemes



Each is capable of delivering multicast packets

Tradeoff is between control traffic and excess data traffic; best choice depends on:

- size & distribution of groups
  (sparse vs. pervasive groups)
- frequency & scope of multicast packets
  (local vs. expanding vs. global transmissions)
- topology

Distance-Vector Multicast Routing Protocol (DVMRP)

Current routing protocol of the MBONE

Earlier version specified in RFC-1075; significantly changed in current implementation ("mrouted")

Based on **reverse-path broadcast** algorithm of Dalal & Metcalfe, with enhancements:

- duplicates eliminated by "parent" election algorithm
- leaf links detected by snooping DV routing updates (assumes "poisoned reverse" form of split horizon)
- unnecessary branches removed on-demand by "prune" messages; restored by "graft" messages

prunes and grafts are *not implemented yet* — currently does a **truncated broadcast**

Multicast Extensions to OSPF (MOSPF)

Developed by MOSPF working group; to be submitted as a Proposed Standard soon

Available now from Proteon;
“gated” implementation underway at Cornell

Straighforward extension to link-state routing algorithm:

- list of groups present on a link is part of link’s state; changes are flooded like any other state change
- when router gets first multicast packet from S to G, computes SPF tree, and caches (incoming-link, outgoing-link-set) for forwarding subsequent pkts
- parts of cache flushed on membership changes, topology changes

MOSPF–DVMRP interoperation to be tested soon, using DVMRP to interconnect MOSPF “islands”

Other Multicast Routing Protocols

Multicast Extensions to BGP

for inter-domain multicast routing (to replace DVMRP)

under development by Scott Brim (Cornell) and the BGP working group

same approach as DVMRP:

- reverse-path forwarding
- on-demand pruning and grafting

Core-Based Trees (CBT)

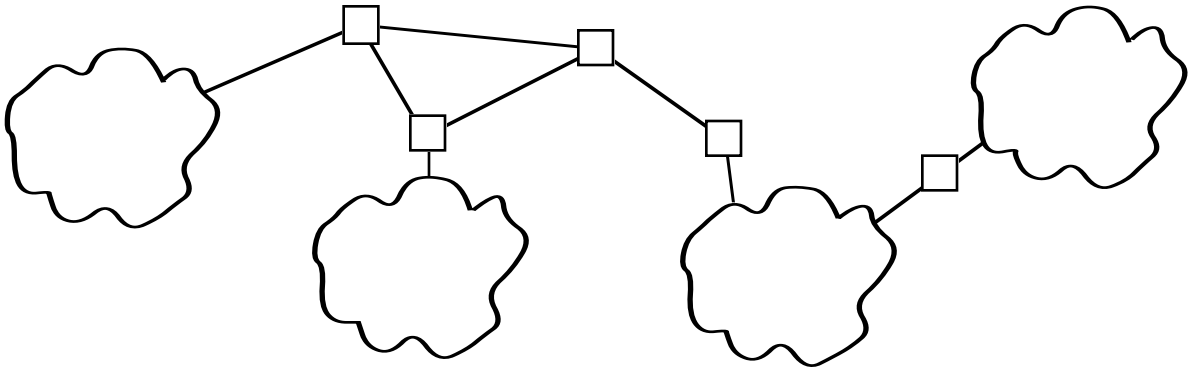
under development by Crowcroft, Ballardie, and Tsuchiya (UCL)

uses single spanning-tree approach, for reduced control traffic for “sparse” groups

Scalability of Multicast Routing

Routing overhead is proportional to number of active (source, group) pairs

- within a single domain, source = subnet
- if each subnet talks to ~ same number of groups, scales the same as unicast routing
- between domains, source = domain



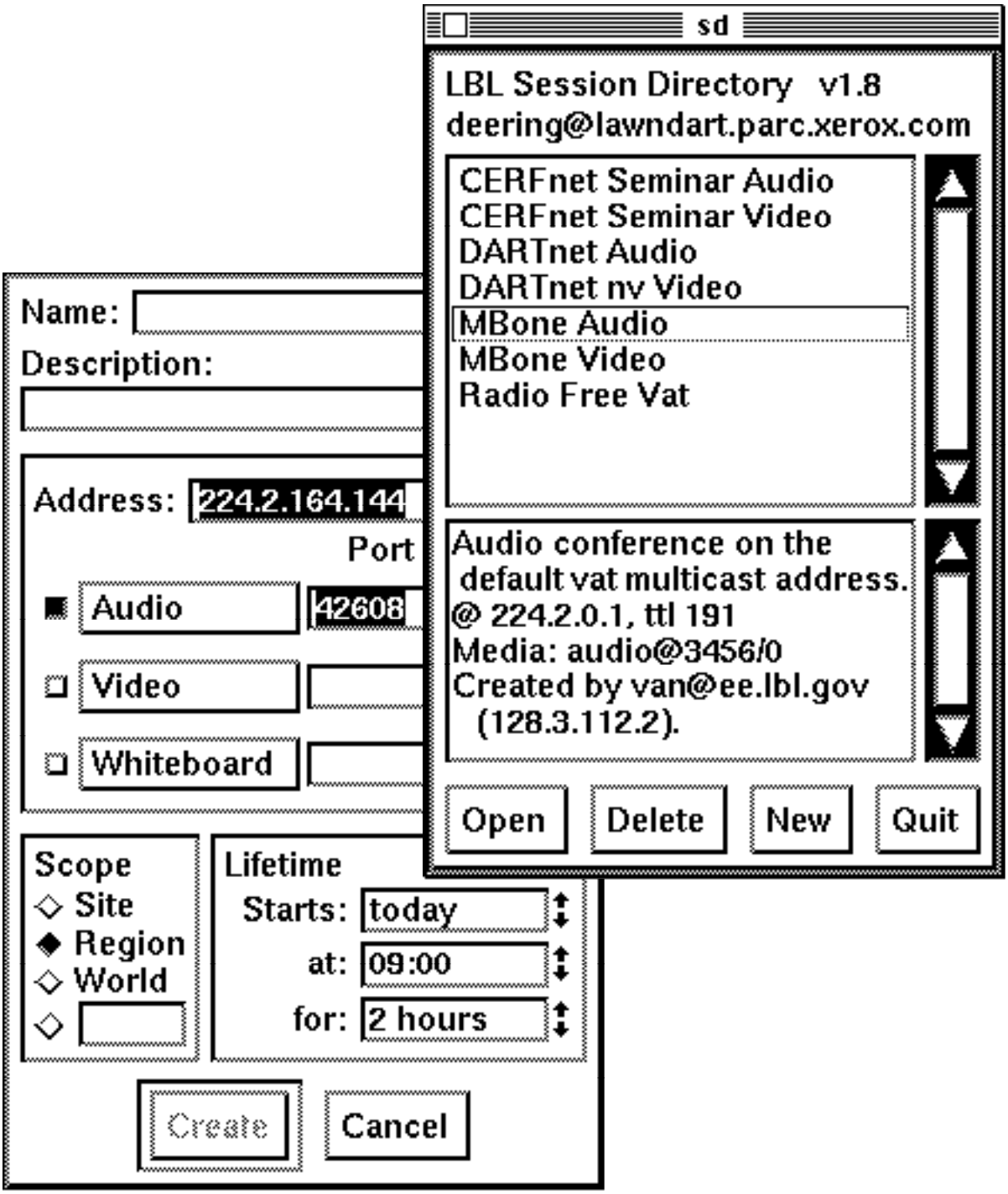
- growth by adding domains scales the same as unicast routing
- growth within domains may increase number of active destination groups per domain

Future IP Multicast Work

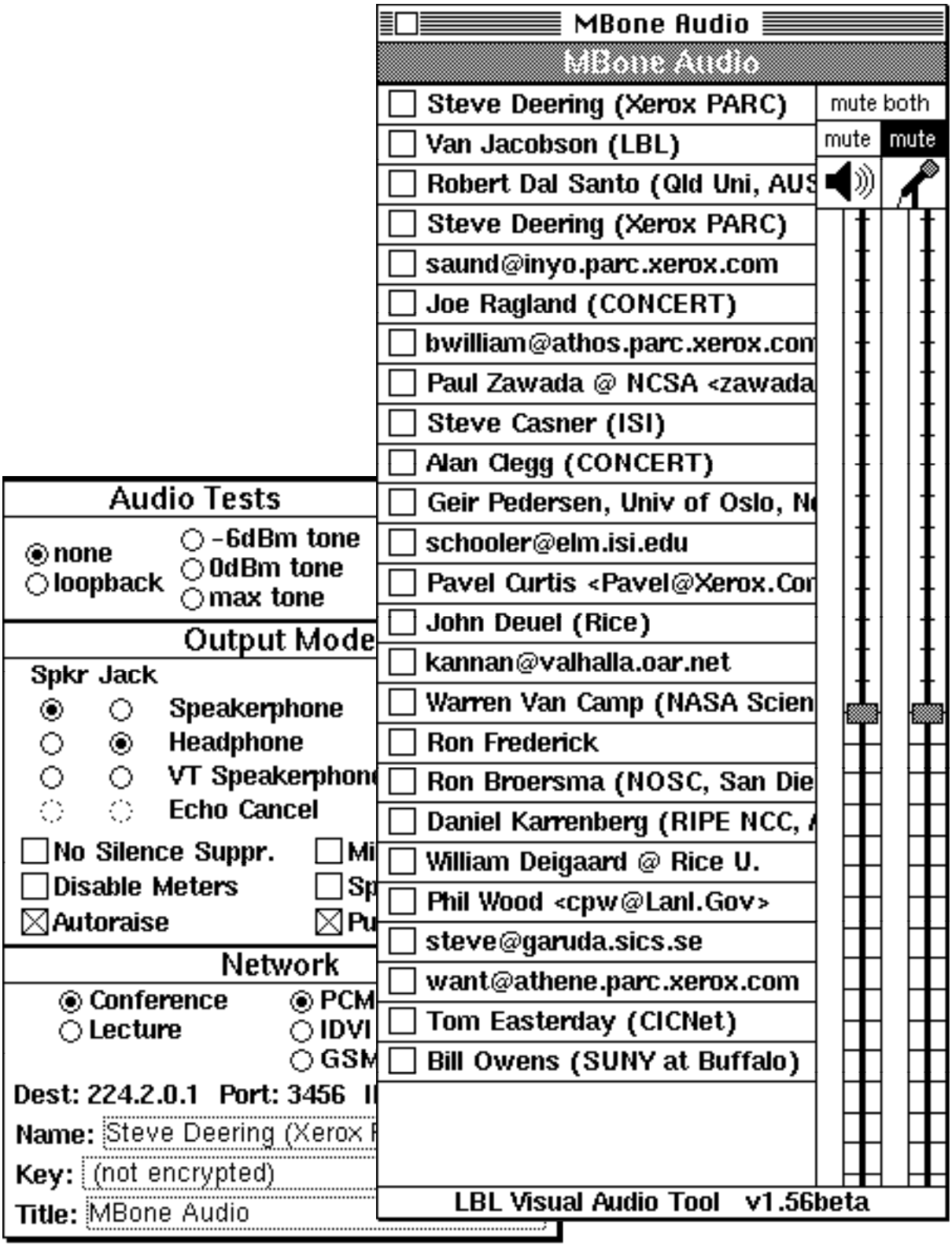
- scoped addresses for better scaling
- encapsulating tunnels for better performance
- DVMRP tree pruning and grafting
- hierarchical DVMRP
- better debugging/management tools, e.g. multicast *traceroute*
- bandwidth limiter and “drop priority” on tunnels and links, for rate control
- lower-overhead routing for sparse groups
- resource reservations for real-time traffic

Multicast Applications

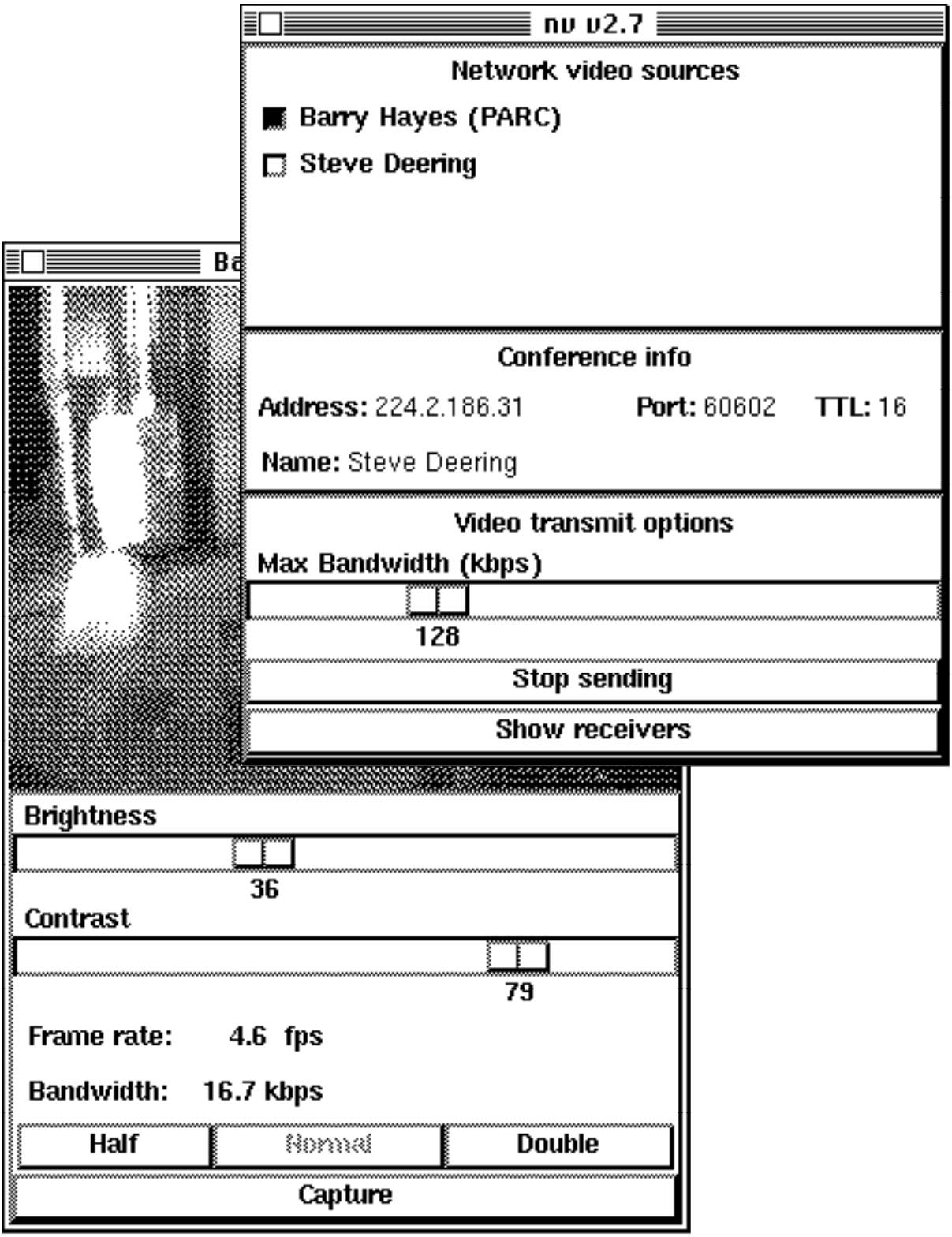
LBL Session Directory Tool (sd)



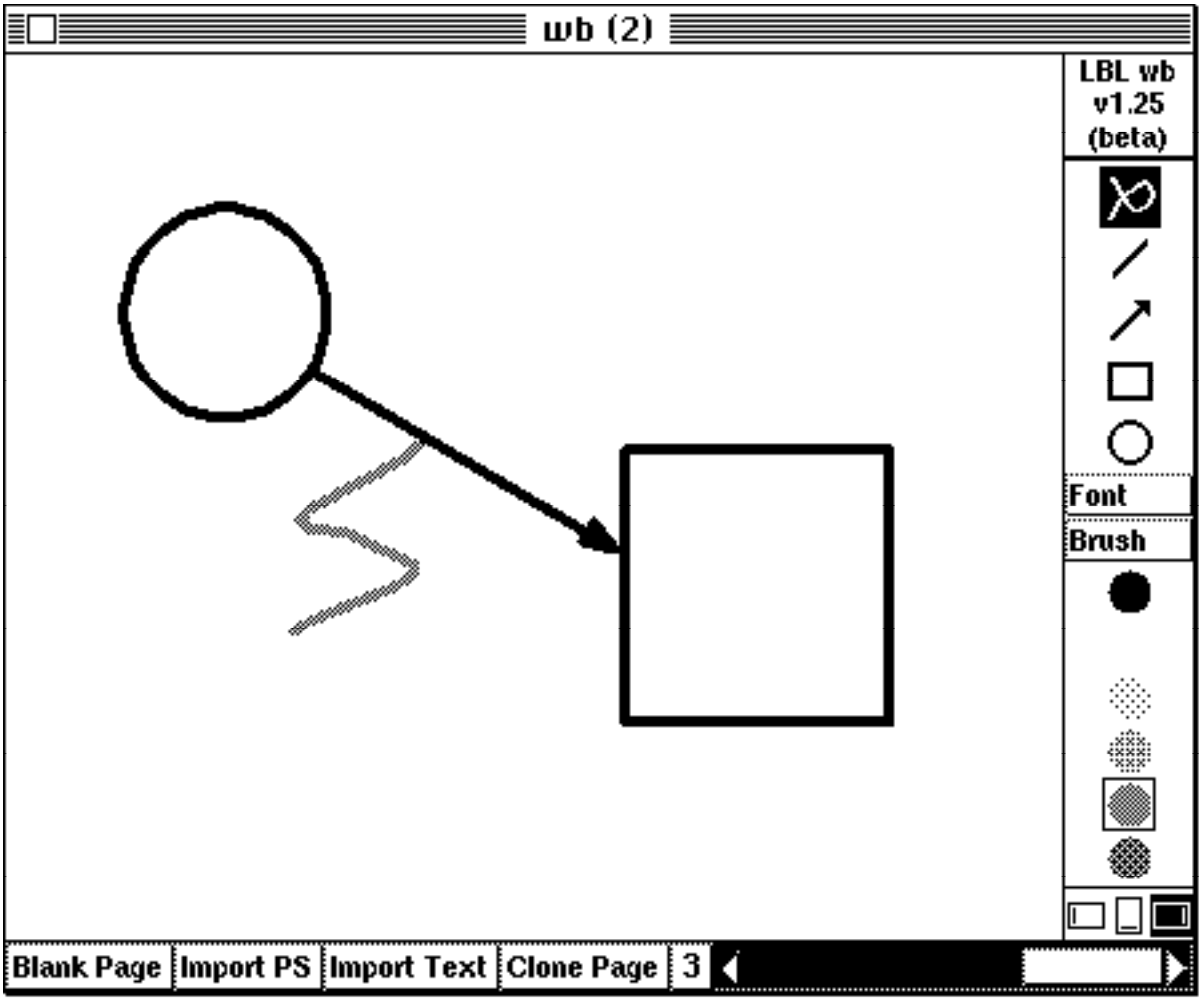
LBL Visual Audio Tool (vat)



PARC Network Video Tool (nv)



LBL Shared Whiteboard Tool (wb)



Note: not released for general distribution yet.

Unlike other tools, wb implements a *reliable* multicast.

Supports drawing, typing, and import of PostScript pages (can annotate imported pages)

Other Applications
(existing, planned, or possible)

mazewar, spacewar, dogfight
("distributed real-time simulation")

NTP (for both autoconfig. and time dissemination)

scientific data collection & dissemination

resource discovery/advertisment

network management

• • •