An Empirical Study of Real Audio Traffic

A. Mena and J. Heidemann
USC/Information Sciences Institute

In *Proceedings of IEEE Infocom*
Tel-Aviv, Israel
March 2000

Introduction

- Internet is growing
- Web facilitates integration of streaming audio
  - Radio juke boxes
  - Broadcast radio
  - Live concerts
- Has been some work on measuring the Web and the Internet, little audio
- This study begins to redress this

→ Study RealNetwork's RealAudio traffic from a "Major" source

Contributions to Understanding

- Majority of data (60%-80%) is UDP
  - Limited congestion control
- RealAudio is CBR at 10s of seconds, but at single seconds is bursty on/off
- RealAudio can use 2 flows, one for control and one for data
  - Most use 2, using TCP for control and UDP for data
  - Those that use 1 use TCP
- User arrivals correlated with time of day
  - Like Web traffic
- Session lengths are long (mean 78 minutes)
  - Unlike Web traffic

Identifying Audio Traffic

- Audio data is mostly unidirectional (from server), ratio 50:1
- UDP RealAudio flow can be identified by packet length and packet inter-departure time
- So, this paper describes how to use results to simulate audio users

Outline

- Introduction (done)
- Methodology ←
- Results
- Simulation
- Future Work
- Conclusions

Methodology

- Capture 5 long traces from `broadcast.com`
  - (Bought by Yahoo! (see link on Web page))
- Trace 1 and 2 using sniffer, 3-5 with `tcpdump`
- Via CISCO Ethernet switch that replicated traffic
  - (*Why replicate and not just measure*?)
  - No packets dropped
  - Saved 98 bytes to get audio header, too

Trace	1	2	3	4	5
Date	Mar 99	Mar 99	Jun 99	Jun 99	Jun 99
Start time, GMT	N/A	N/A	16:02	13:32	13:38
Duration	83 sec	141 sec	5.5 hr	10.5 hr	18.2 hr
Packets	134 K	284 K	5.5 M	1.6 M	5.9M
Bytes	38 M	63 M	1.3 G	0.4 G	1.3 G

(Trace 1 and 2 not used, much … too short)

Terms used in Results

- IP address of receiver is *client*
  - Could be proxy or behind firewall
- IP address of sender is *server*
  - One for this study, fixed
- User initiates *session*, with one or more *flows*
  - With one or more flows (*source-destination* pairs)
  - Control flow: authenticate, start-stop, …
    - Always TCP (defined by RTSP standard)
  - Data flow: encoded audio information
- Inbound traffic – received by server
- Outbound traffic – sent by server

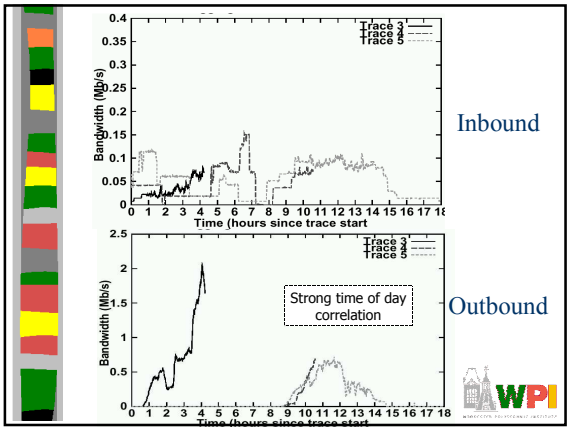
Outline

- Introduction (done)
- Methodology (done)
- Results ←
  - Overview
  - Aggregate
  - Users
  - Flows
- Simulation
- Future Work
- Conclusions

Distribution of Traffic

	Trace 3	Trace 4	Trace 5
Inbound Metrics	1.8M pkts 50 MB	0.9M pkts 216 MB	2.9M pkts 425 MB
Outbound Metrics	3.6M pkts 1,202 MB	0.6M pkts 198 MB	2.9M pkts 866 MB

- Inbound to Outbound
  → 1:24 to 1:1
- But 1:1 are mostly when upload from codec
  – Omit from further analysis
- Other Inbound is primarily acks + feedback



Summary of Traffic Traces

	Trace 3	Trace 4	Trace 5
Audio Data	1,160 MB 3.7M packets	403 MB 1.2M packets	1,268 MB 4.6M packets
Control Data	41.3 MB 1.7M packets	10.3 MB 0.3M packets	22.67 MB 1.2M packets
Other Packets	1.0 MB 90K packets	0.8 MB 44K packets	1.0 MB 98K packets

- Control only small portion of bandwidth
  – 1% - 3% of bytes
- Overall 99% of all bytes, 98% of all packets are audio
- Other is remnants of old flows, connections by admin

Aggregate Traffic

	Trace 3	Trace 4	Trace 5
Bytes			
UDP	723 M (60 %)	415 M (79%)	955 M (74 %)
TCP/Non-HTTP	432 M (36 %)	68 M (17%)	304 M (24%)
TCP/ HTTP	47 M (3.9 %)	18 M (4%)	36 M (2%)
Multicast	0	0	0
Packets			
UDP	3.68 M (67 %)	1.26 M (80%)	4.52 M (77%)
TCP/Non-HTTP	1.66 M (30 %)	0.26 M (17%)	1.21 M (21%)
HTTP/TCP	0.14 M (3 %)	0.05 M (3%)	0.12 M (2%)
Multicast	0	0	0

- 1/3 TCP, most not HTTP
- 2/3 UDP
- No multicast

Outline

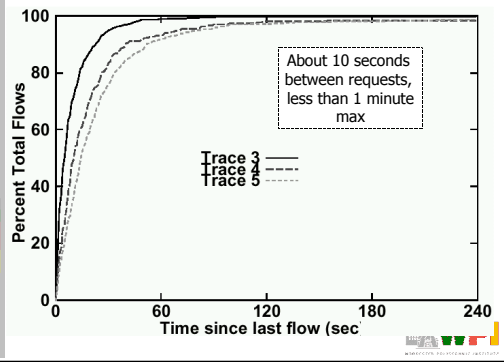
- Introduction (done)
- Methodology (done)
- Results
  - Overview (done)
  - Aggregate ←
  - Users
  - Flows
- Simulation
- Future Work
- Conclusions

Summary of Audio Flows

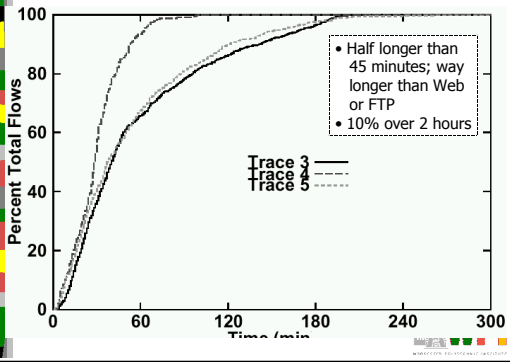
- Identified audio flows by sending 100K or more
  - Port numbers unreliable since negotiated by RTSP (*what is this*?)
  - Identified about 90% of flows

	Trace 3	Trace 4	Trace 5
Audio Data Flows	1460	324	837
Inbound	14	20	42
Outbound	1446	304	795
UDP flows	1165 (81 %)	217 (71%)	611 (77%)
TCP flows	281 (19 %)	87 (29%)	184 (23%)

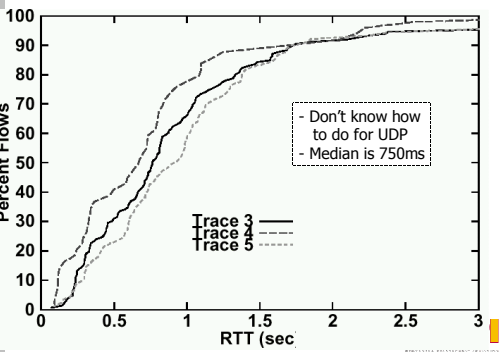
Audio Flow Interarrivals



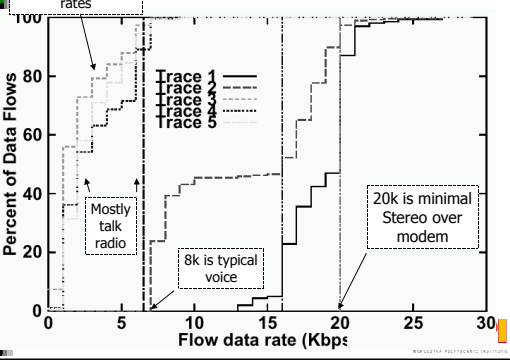
Flow Duration



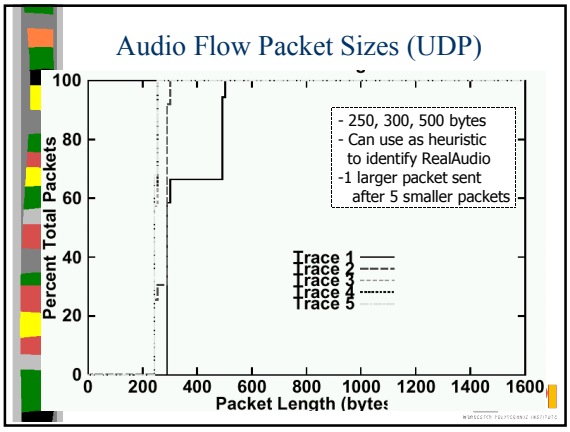
Audio Flow Round Trip Time (TCP)



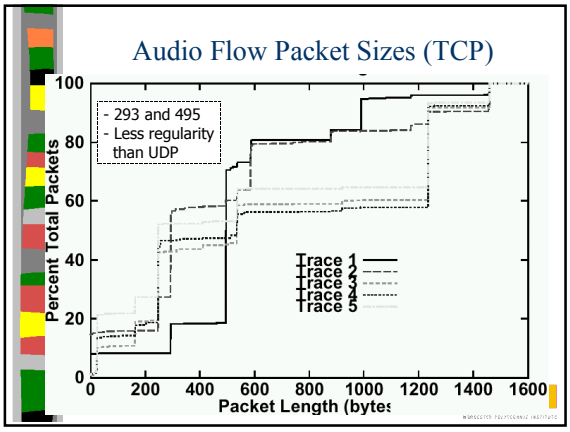
Audio Flow Data Rates



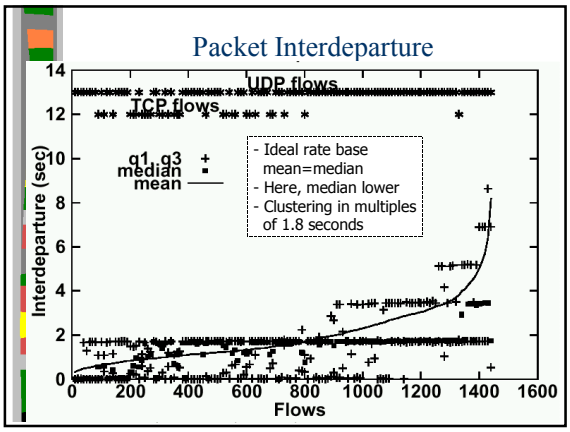
Audio Flow Packet Sizes (UDP)



Audio Flow Packet Sizes (TCP)



Packet Interdeparture

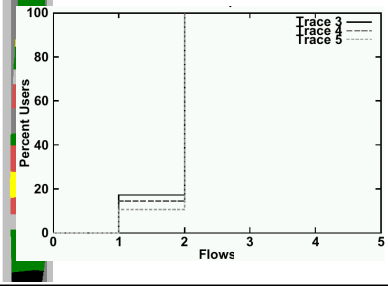


Outline

- Introduction (done)
- Methodology (done)
- Results
  - Overview (done)
  - Aggregate (done)
  - Users ←
  - Flows
- Simulation
- Future Work
- Conclusions

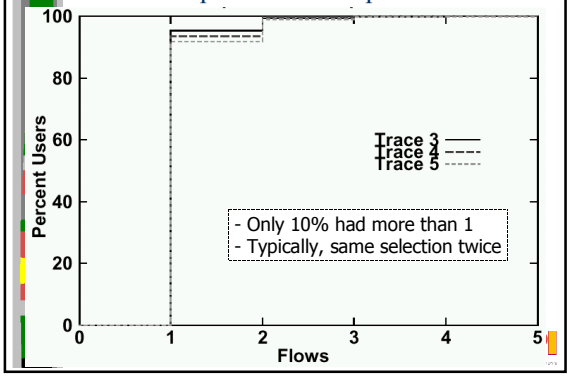
Aggregate Users

	Trace 3	Trace 4	Trace 5
Active outbound audio flows	1460	324	837
Active outbound audio users	1384	288	728
Mean Number of Flows per user	1.06	1.13	1.15

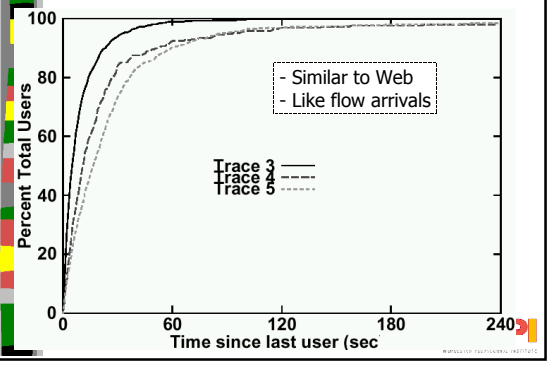


- Users can have more than one flow
  - Probably a proxy
- Not significant

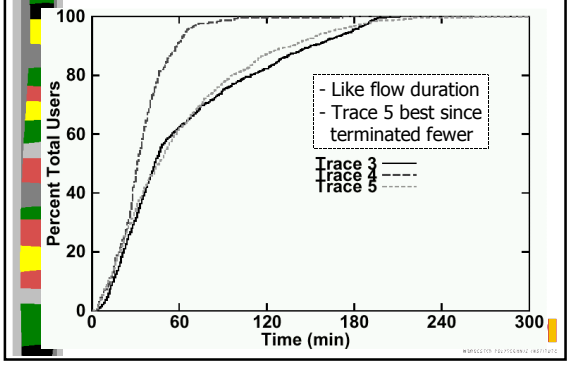
Sequential Flows per User



User Arrivals



User Duration



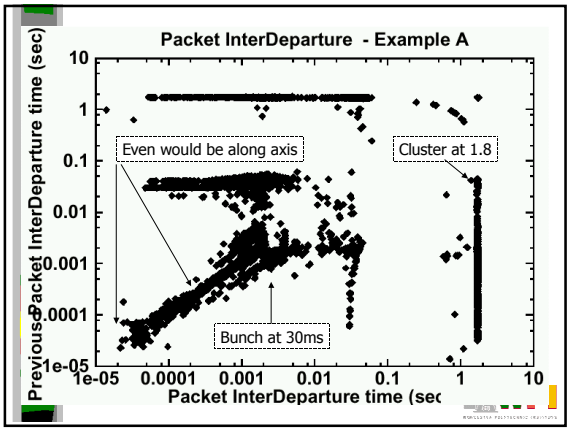
Outline

- Introduction (done)
- Methodology (done)
- Results
  - Overview (done)
  - Aggregate (done)
  - Users (done)
  - Flows ←
- Simulation
- Future Work
- Conclusions

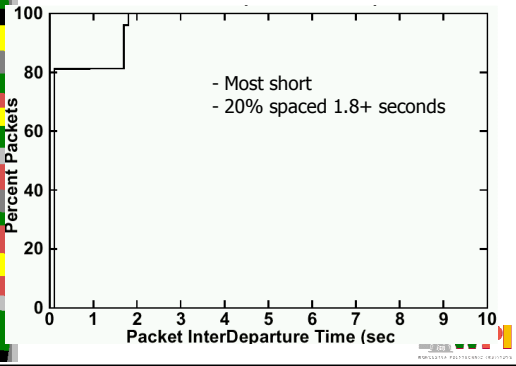
Packet Departure for Individual Flows

Example flow	A	B
Trace	3	3
Transport	Udp	udp
Mean bandwidth	6.7Kbp/s	2.2Kb/s
Mean interdeparture	0.324 sec	4.124 sec
Median interdeparture	0.002 sec	1.730 sec

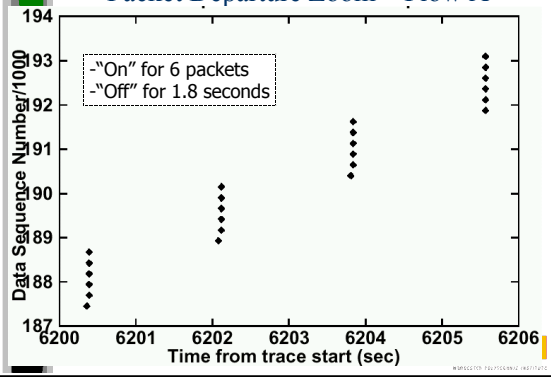
- Given packet sent times: t_0, t_1, t_2
- Compute $\delta_1 = t_1 - t_0$, $\delta_2 = t_2 - t_1$
- Graph (δ_1, δ_2)

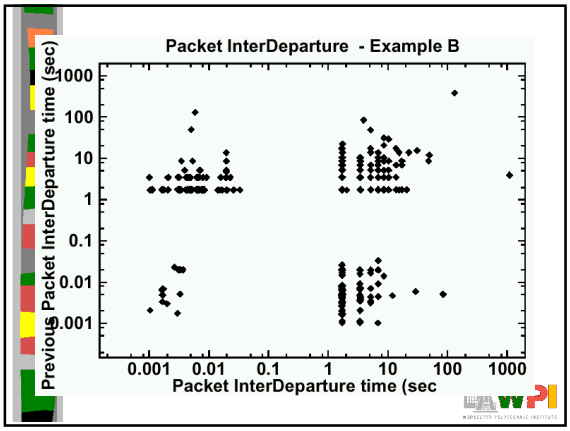


Packet Interdeparture CDF – Flow A

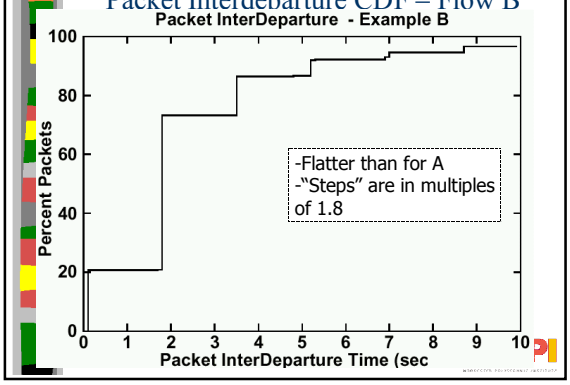


Packet Departure Zoom – Flow A





Packet Interdeparture CDF – Flow B



Outline

- Introduction (done)
- Methodology (done)
- Results (done)
- Simulation ←
- Future Work
- Conclusions

Simulation of Audio Flows

- Place server in network. Assume busy or not based on time of day.
- Pick RTT from CDF (Figure 5 in paper)
- Pick audio flow duration (Figure 4)
- Pack audio data rate (Figure 6)
  - Select appropriate packet length
- Send data at appropriate rate (Figure 14)
  - On/Off process

Future Work

- Data sources
  - Other kinds: individual songs, conferencing …
  - Packet traces closer to client
- Audio congestion control
  - For UDP flows (topic in class)
- Multimedia flow identification
  - Source-Dest or Packet Length or Port …

Conclusions

- Measured RealAudio flows to better understand
- Different than FTP, HTTP or Telnet
  - audio longer duration
- 60-70% use UDP
- Regular packet length, bit rates and interarrival times
- CBR over long time scales, but not short ones
- Overall, MM is growing and it does not look like typical traffic