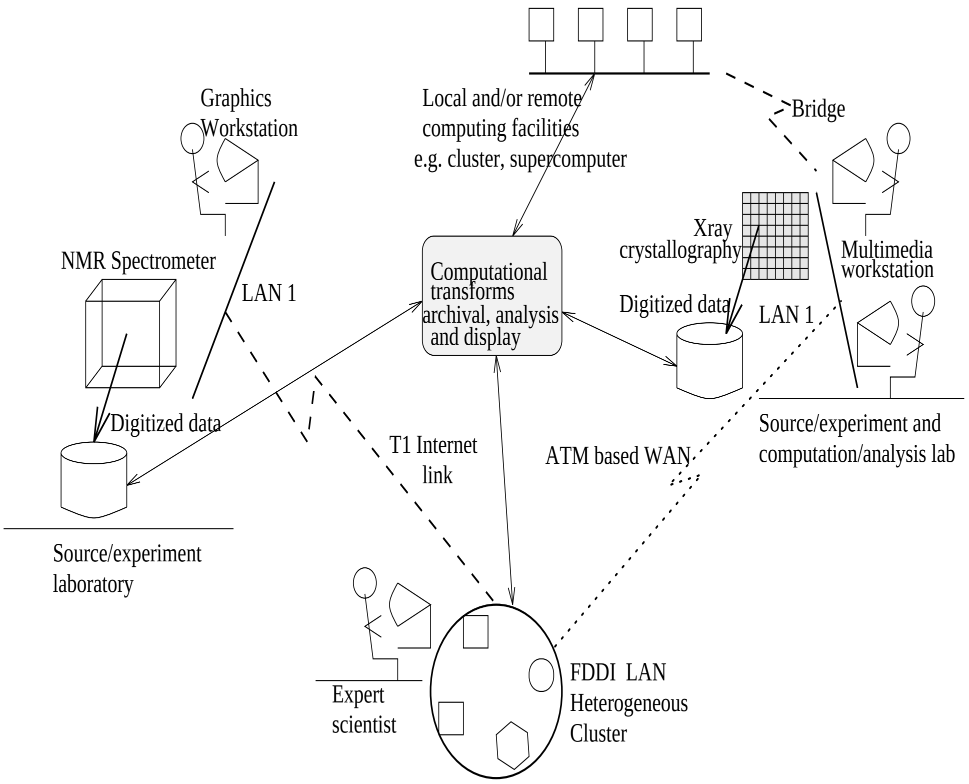
Design and Implementation of a Multicast Audio Conferencing Tool for a Collaborative Computing Framework

Sarah E. Chodrow
Michael D. Hirsch
Injong Rhee
Shun Yan Cheung

Department of Math. & Comp. Science
Emory University
Atlanta, GA 30322

Collaborative Computing Framework (CCF)



- Cooperation via
  - shared virtual desktop (application sharing tools)
  - clear/white boards
  - audio/video desktop conferencing
  - shared data space, launching applications
- Tools should be integrated

Audio in Collaboration

Human Audio Communication:

- Typically carries the critical component for group discussion
- Highly sensitive to distortions

Requirements:

- Very high audio quality
- Multi-party conferencing (audio mixing)
- Without MBone (limited availability to researchers in non-CS discipline)
- "Hands free" operation
- Record/Playback facility
- Integration into CCF Virtual Desktop and other CCF collaboration tools, e.g., laboratory note book

Related Work

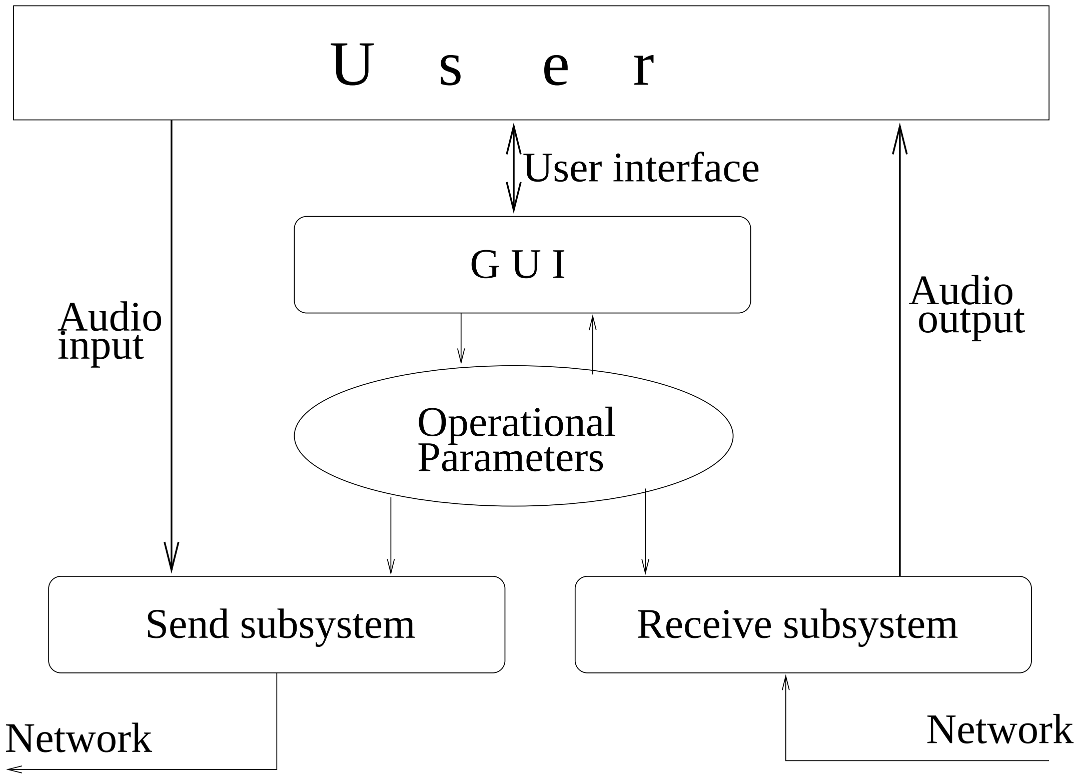
Audio Conferencing Tools:

- vat: first and currently most popular audio tool
- rat: redundancy to combat packet loss
- fphone: generalized redundancy & flow control
- SpeakFreely
- LinuxPhone
- ShowMe (SUN)
- ...

Why YAA:

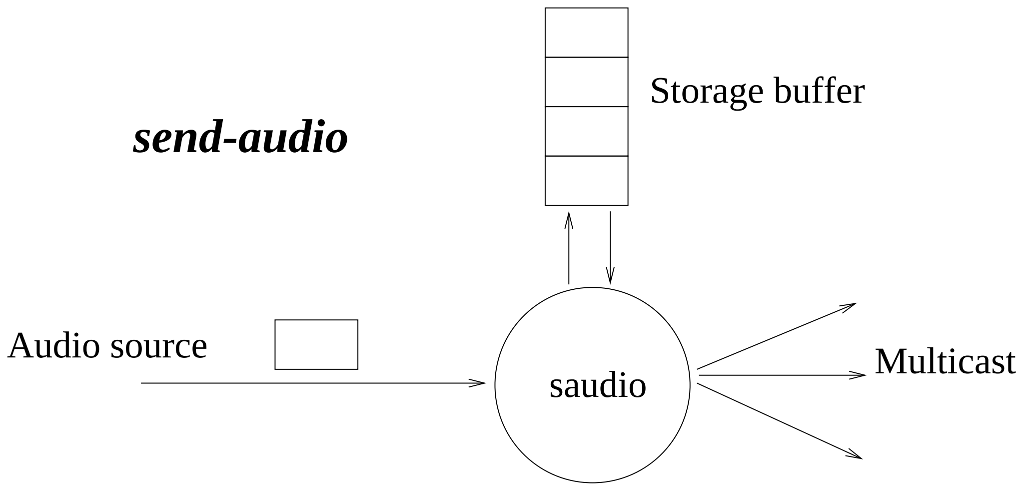
- Multicast relies on MBone
- Integrated with CCF Virtual Desktop and with other CCF tools
- Better audio quality

CCFAudio Architecture



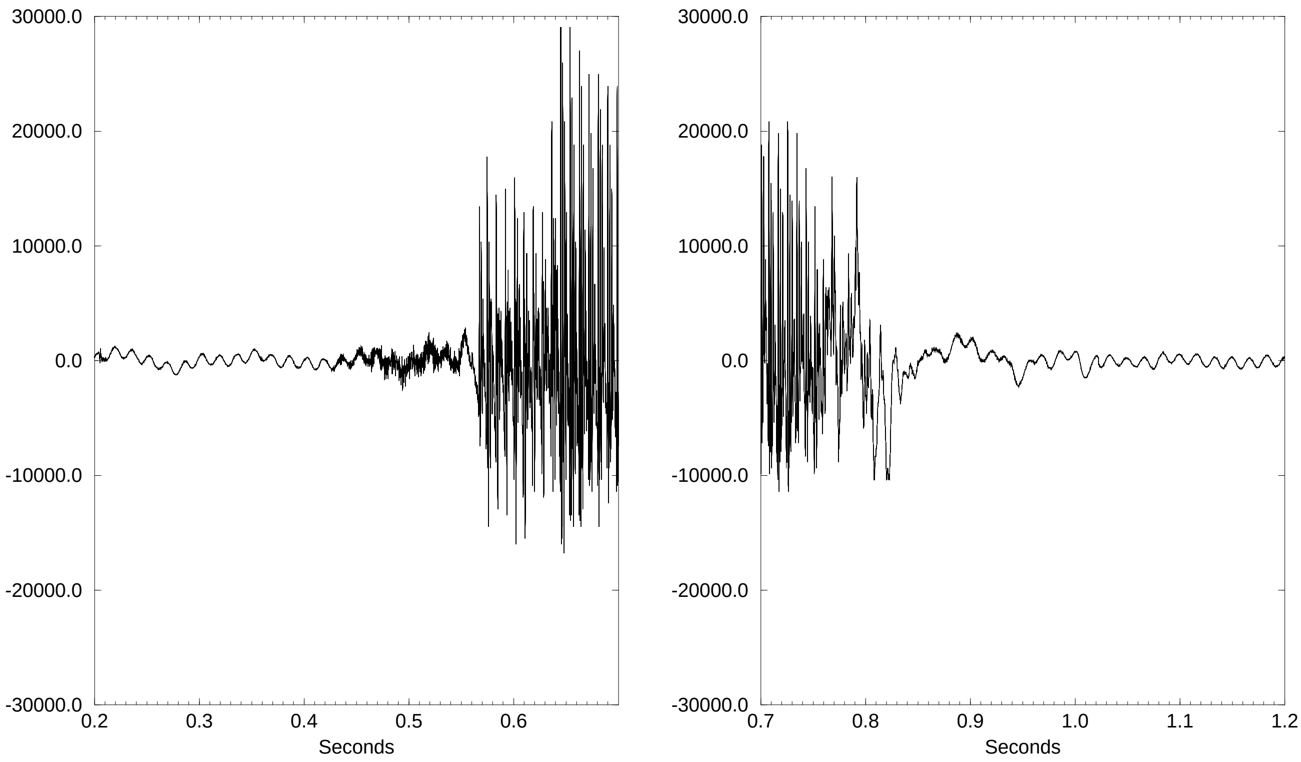
- Three independent subsystems:
  - GUI
  - Send
  - Receive
- One or more threads per subsystem

Audio Send Subsystem



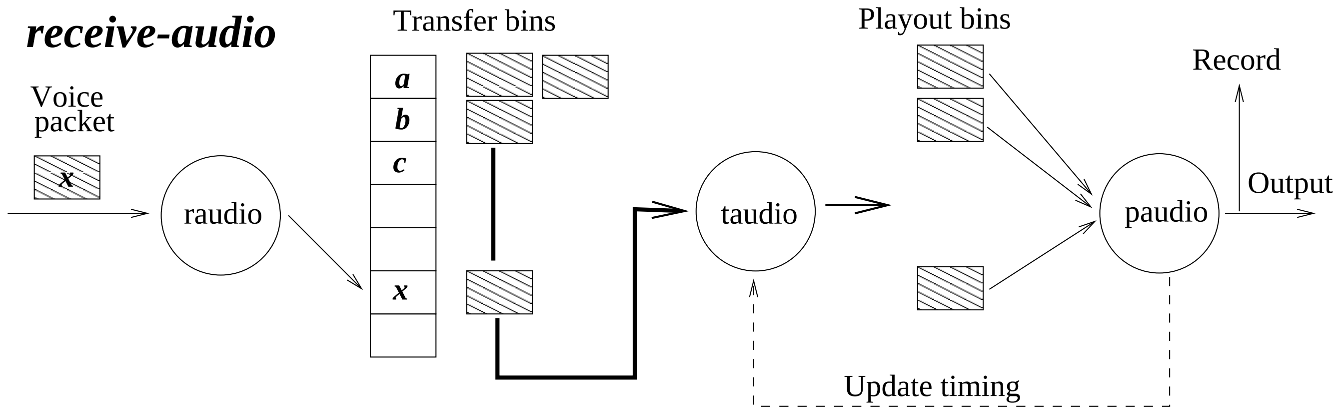
- Audio source can be mike or audio file
- Hands-free operation through voice/silence detection with threshold τ
- Circular storage buffer holds all (voice and silence) audio clips
- Transmits upto k_1 silence clips before talk spurt
- Transmits upto k_2 silence clips after talk spurt
- τ, k_1 and k_2 adjustable through GUI
- Simple and efficient redundancy scheme (Hardman et al): same primary and secondary encoding

Hands-free high quality audio



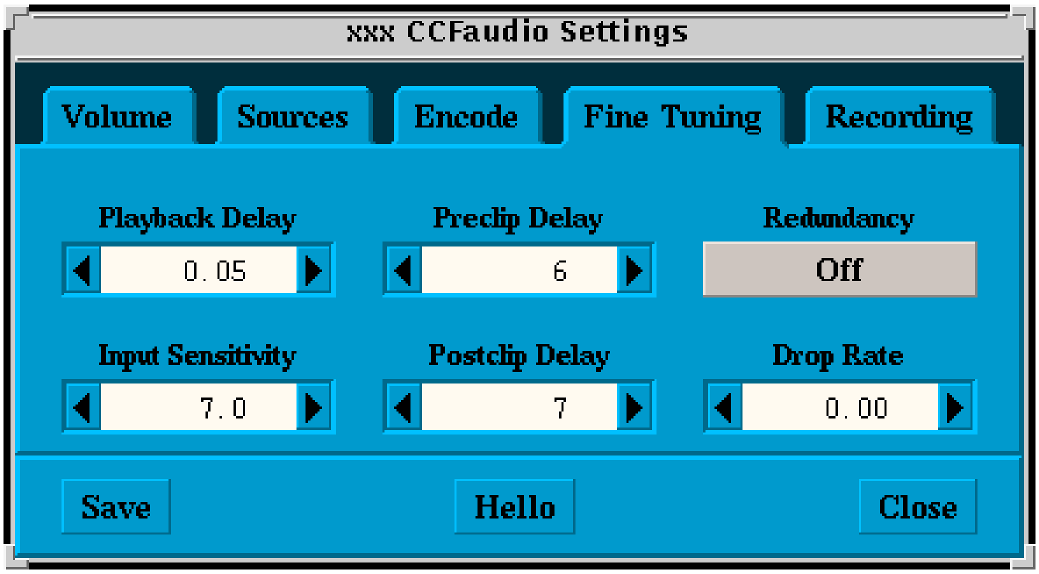
- Voice resounds approximately 1 second before it can reliably detected
- Resounds until approximately 1 second after voice spurt ended
- Need “pre-clips” and “post-clips” to achieve high quality audio
- Clip length depends on articulation
  $\Rightarrow$ # pre-clips and post-clips adjustable

Audio Receive Subsystem



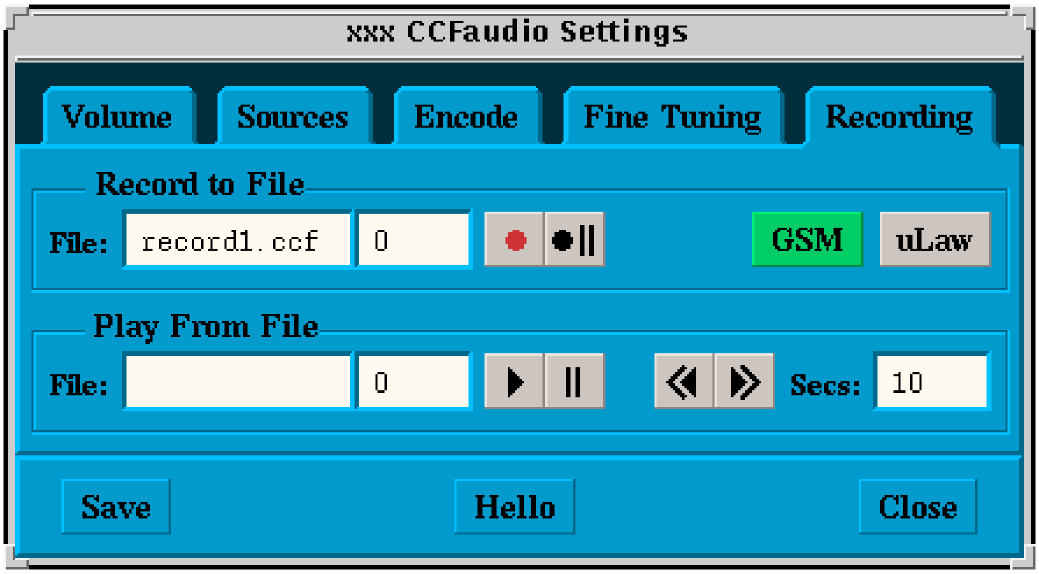
- Producer/consumer design with 3 threads
- raudio: speedy receive thread
- taudio: jitter control, audio decoding & recovery
  - Jitter control: accumulate ℓ packets before transfering (ℓ adjustable)
  - Larger ℓ reduces likelihood of late arrivals
  - Recovery with simple and efficient redundancy scheme (Hardman et al)
- paudio: audio mixing, playout, recording & end time estimation
- End time estimate used by taudio to wait for late packets without causing audio underflow

GUI: Fine Tuning Parameter



- Send Audio Subsystem:
  - Sensitivity: voice detection level
  - Preclip Delay: length before talk spurt
  - Postclip Delay: length after talk spurt
  - Redundancy: Hardman-like redundancy
- Receive Audio Subsystem:
  - Playback Delay: Jitter control

GUI: Recording and Playback



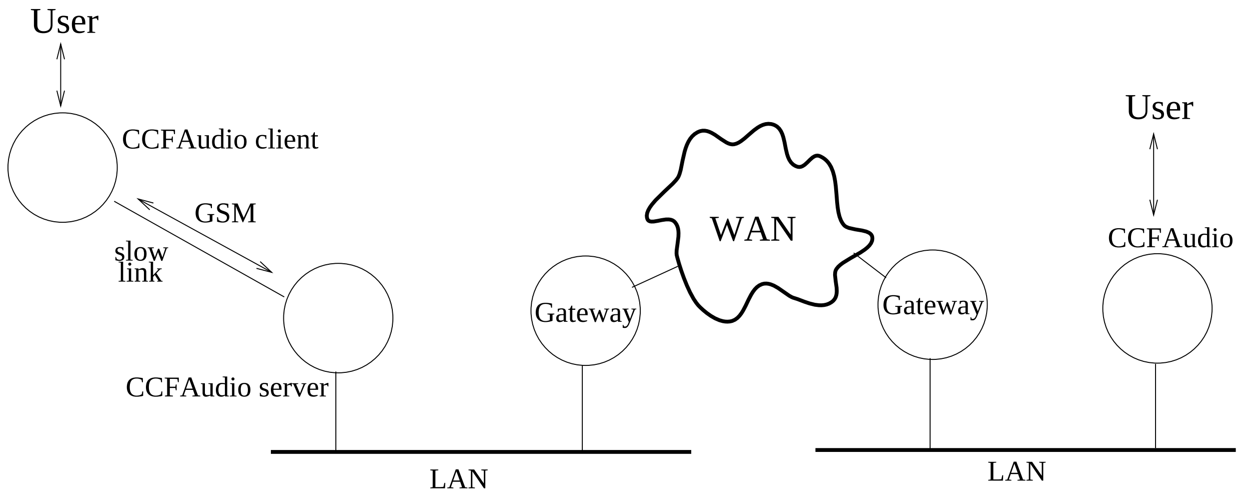
Recording:

- Start/Stop
- Pause

Playback:

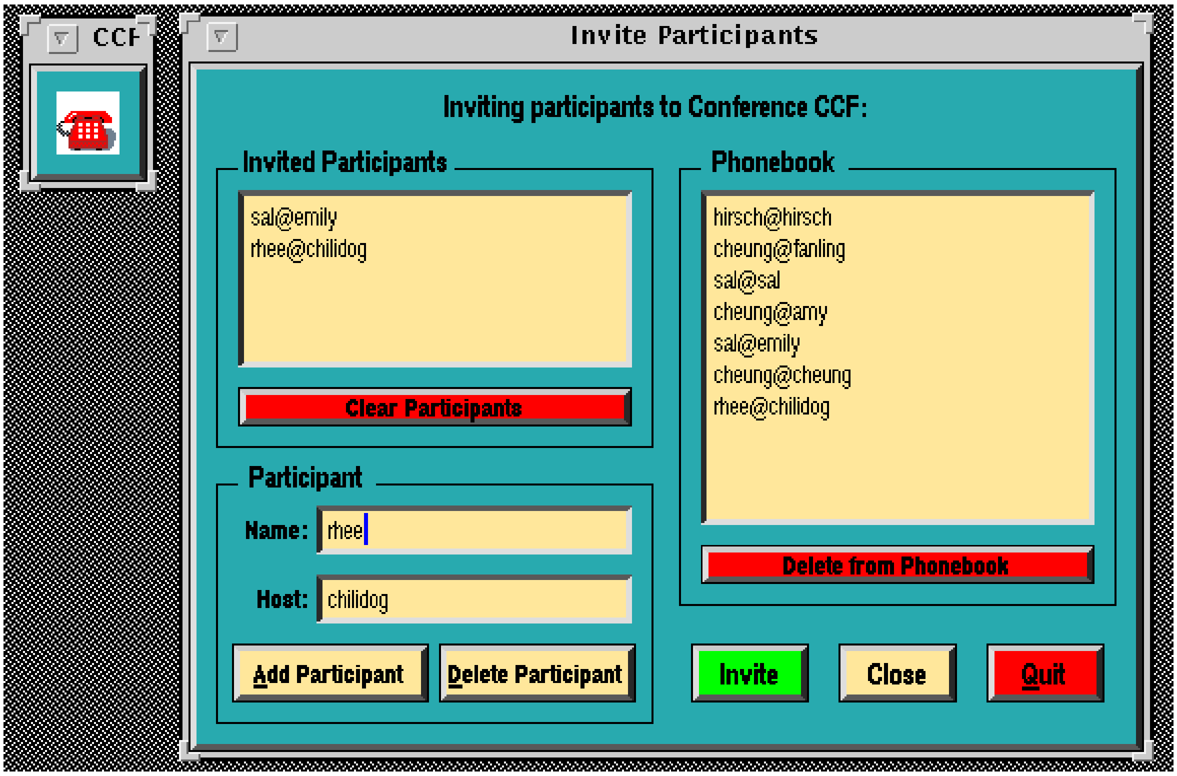
- Start/Stop
- Pause
- Fast Forward & Rewind (# seconds)

Low Bandwidth Client/Server Operation Mode



- CCFAudio server on Internet
- CCFAudio client on workstation at home
- Low bandwidth (GSM) connection between client and server
- Server:
  - receives audio from client and multicasts
  - receives audio from other participents, mixes, compresses and then forward mixed audio stream to client

Ringer Companion Tool



- Used to invite participants into an audio session
- Starts CCF Audio Tool automatically in correct session
- Future work: generalize for inviting into CCF session (virtual desktop)
- Bonus: ringer/CCFAudio combo provide answering machine functionality

Performance data from controlled experiment

Experiment 1: GSM, without redundancy

Drop Rate	Total Sent	Total Lost	Perceived Quality
9%	1152	100	A/B
20%	1321	266	C
29%	1037	425	F

Experiment 2: GSM, with redundancy

Drop Rate	Total Sent	Total Lost	Total Rep'd	Eff. D. Rate	Perc'ved Qual
9%	1147	100	91	.01%	A
20%	1317	266	209	4%	A/B
29%	1866	545	373	9%	B
39%	1669	652	386	16%	B/C
49%	1438	699	347	24%	D

Rating:

- A: Perfectly understandable
- A/B: Very understandable
- B: Understandable
- B/C: Understandable with minor effort
- C: Understandable with effort
- D: Understandable with major effort
- F: Barely understandable

CCFAudio usage

- Regularly used by CCF team members over departmental LAN to debug CSCW apps
- Occasional use: long distance audio calls to friends
  - Atlanta – Toronto
  - Atlanta – South Carilina
- Tested on a recent visit to Japan by a Chemistry collegue.

  A comment after this test:

  ```
          There was an almost neglible delay;
  we didn't even notice it.  The quality was
  very clear on our end.  It was probably a
  little better than an international call in
  relation to both quality and delay (we don't
  have a lot of experience with international
  calls, but this went very well).
  ```

Conclusion

CCFAudio:

- Efficient threaded design
- Intergrated with CCF virtual desktop
- Stand-alone version also available
- Achieves high quality audio
- Includes features tailored for collaboration in natural and other sciences
  - do not need MBone feed
  - playback of previously recorded dicussion
  - hands-free operation
- Will soon be available on a Web site near you: http://ccf.mathcs.emory.edu