Internet Real-Time Audio Overview

Jonathan Lennox
Columbia University, New York
lennox@cs.columbia.edu

AES 14th International Conference
internetaudio.aes.org
Seattle, Washington
June 13, 1997

Internet Real-time Audio Overview

- Internet Standards Process
- Audio Compression
- Real-time Streams

IETF Standards Process

Not every RFC is an Internet Standard!

- 53 official standards; 2149 RFCs
- Best Current Practice RFC, “Experimental,” “Historic,” “Informational”

Stages of Internet Standards Process

- Proposed Standard
- Draft Standard
- Standard

Differences between Standards Bodies

	IETF	AES	ITU
Decision-making	Consensus	Voting	Voting
Participation	Open (any attendee)	Broad (any AES member)	Small (national PTTs and large companies)
Availability of Standards	On the Internet	Available for Purchase	Available for Purchase
Availability of Drafts	On the Internet	Committee only	Committee only
Speed	Rapid	Slow	Slow

Compression: motivations

Current Internet: about 20-30 kbps throughput per user (V.34 modems and typical share of backbone)

Local Networks, Internet II: better, but still not unlimited: Ethernet 10/100 Mbps; OC-12 622 Mbps

Uncompressed voice-grade audio: 64 kbps

Uncompressed CD-grade audio: 1.4 Mbps

Compression: tradeoffs

- compression ratios vs. audio quality vs. CPU requirements vs. delay
- real-time encoding (necessary for interactivity) vs. pre-encoding (a possibility for some applications, but reduces flexibility)
- dedicated hardware vs. software on general hardware

Compression: general techniques

- companding: non-linear quantization ➠ μ-law (G.711)
- simple algorithmic: ADPCM
- model: model voice, extract parameters
- subband: split signal into bands and code individually; make use of masking properties of human ear
- entropy reduction: exploit statistical correlations; pack losslessly. Necessary for professional work / “Golden Ears” …

Compression: open standards

coding	kb/s	use
PC-10	2.4	robotic, secure telephone
GSM	13.0	European mobile phone
G.729	8.0	mobile telephony
G.723.1	5.3/6.3	videophones
MPEG L3, MPEG 2, ...	$\geq$ 128.0	near-CD stereo / multichannel
AC-3	$\sim$ 384.0	DVD

Compression: proprietary standards

coding	company	kb/s	use
RealAudio	Progressive	10 / 20	“AM”/“FM” Radio
RT24	Voxware	2.4 kbps	pre-recorded speech

Motivations for Real-time Protocols

- A late packet is as bad as a lost packet for real-time steams
- Current Internet provides no loss or delay guarantees (made cheap, early, ubiquitous implementation possible)
- TCP - provides reliable transfer (eventually); not suitable for real-time data

Real-Time Transport Protocol — RTP

lightweight: specification and implementation

flexible: provide mechanism, don't dictate algorithms

protocol-neutral: UDP/IP, IPX, ATM-AALx, ...

scalable: unicast, multicast from 2 to $\sim$1000

separate control/data: some functions may be taken over by conference control protocol

secure: support for encryption, possibly authentication

RTP functions

- segmentation/reassembly done by UDP (or similar)
- resequencing (if needed)
- loss detection for quality estimation, recovery
- intra-media synchronization: remove delay jitter through playout buffer
- intra-media synchronization: drifting sampling clocks
- inter-media synchronization (lip sync between audio and video)
- quality-of-service feedback and rate adaptation
- source identification

Delay, Loss, and Jitter

- Very situation-dependent — “What is the delay like on the Internet?” is like “What is the weather like in the United States?”
- Can change rapidly due to behavior of distant hardware, software
- Can change due to your own traffic
- Measuring one-way delay is very difficult — delay is assymetric

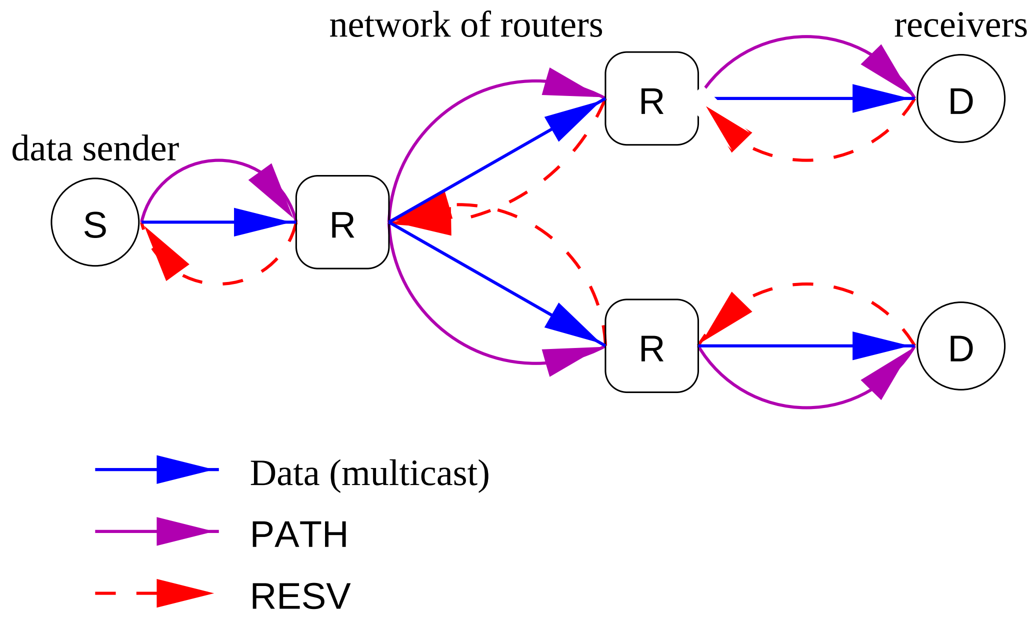
Multicasting — connections between multiple receivers

- one-to-many; few-to-many; all-to-all
- host joins group with IGMP (Internet Group Membership Protocol); broadcasts to router
- Routers maintain knowledge of all destinations; no hierarchical routing protocols yet
- MBone — most core Internet routers don’t support multicast today; “tunnel” multicast transmissions through unicast

Quality of Service

- Provide guaranteed data rates, bounds for delay
- Very hard problem — 100,000 streams through typical core routers
- Flow agglomeration — helps somewhat

RSVP — Resource Reservation Protocol — control QoS



- Receiver initiates QoS request
- Designed to work with Multicast
- Many questions about scalability

RTSP — Real-Time Streaming Protocol

- "Internet Remote Control" — control multimedia streams
- Media Servers
  - VOD servers (pre-recorded)
  - Live Feeds (concerts, TV, etc)
- Desirable to control media servers
  - Content descriptions
  - Start, stop, record, pause
- media stream and sessions defined by a RTSP URL

  `rtsp://media.example.com:554/twister/autrack`
- Suitable for professional applications
  - SMPTE timecode
  - Remote digital editing