Lecture 25
Multimedia

Peter Steenkiste
Department of Computer Science and
Electrical and Computer Engineering
Carnegie Mellon University

15-441 Networking, Spring 2006
http://www.cs.cmu.edu/~prs/15-441

Overview

- Multimedia applications
- Current challenges of multimedia networking
- Media Encoding
- Methods of streaming stored media
- Case Studies:
  - » Real Time Voice over IP (VoIP)
  - » Video over IP
- Real Time Protocol (RTP)
- Real Time Control Protocol (RTCP)

Characterizing Media

- Continuous media **vs. multimedia**
  - » Video
  - » Audio
- Compare with web pages
  - » Web pages are definitely multimedia, but don't fit the characteristics of continuous media
- 3 + 1 classes of multimedia applications:
  - » Streaming Stored Media
  - » Unidirectional Real-Time
  - » Interactive Real-Time
  - » Others

Application Classes: Streaming Stored Media

- Pipeline reception of stored files over the network.
  - » No need to wait for entire file before viewing
- Interactive: Users can play, pause, rewind, etc.
  - » Delay: from client request until display start can be 1 to 10 seconds
  - » Time is used to buffer data to reduce chance of stalling
- Delay constraints on data delivery to maintain continuous playout.

Application Classes: Unidirectional Real-Time

- Similar to existing broadcast TV/radio, but delivery on the network.
  - » Data is generated on the fly
- Non-interactive: just view/listen
  - » By storing media: can pause, rewind
- As with streaming media, delay constraints on data delivery to maintain continuous playout
- Often used combined with multicast

Application Classes: Interactive Real-Time

- Phone conversation/video conference
- More stringent delay requirement than Streaming and Unidirectional because of real-time interactive nature
  - » Need to handle "switch over"
  - » There is a limit to how much people are willing to adapt to the delay
- Match audio and video
  - » Video: < 150 msec acceptable
  - » Audio: < 150 msec not perceived, <400 msec acceptable

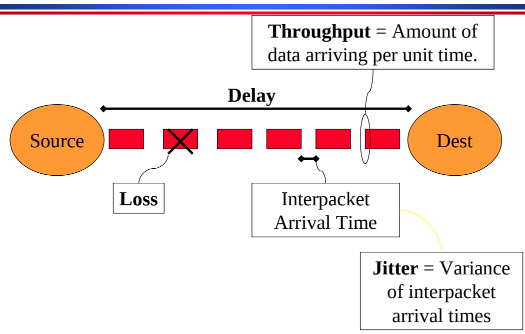
Application Classes: Others

- Multimedia sharing applications
  - » Download entire file and view later
  - » Eg. Napster, Gnutella, Scour, Freenet, LimeWire, BearShare, iMesh, Napigator,…
- Distance learning applications.
  - » Coordinate video, audio, and data (e.g. slides)
  - » Typically distributed on CDs

Overview

- Multimedia applications
- Challenges of multimedia networking
- Media Encoding
- Methods of streaming stored media
- Case Studies:
  - » Real Time Voice over IP (VoIP)
  - » Video over IP
- Real Time Protocol (RTP)
- Real Time Control Protocol (RTCP)

Review: Network Basics



Problems with Delay

- Interactive applications
  - » Video conferencing
    - – Human communication starts to break down at ~500 msec delays
  - » Feedback loops
  - » Delay affected by encoding/decoding
- Select low delay paths.
  - » Can be negotiated as part of service
- Longer term solution: Quality of Service (QoS), Diffserv
  - » Best effort service and end-to-end mechanisms provide little opportunity for controlling delay

Problems of Jitter

- Real-Time systems must deal with this
- Media jitter vs Packet jitter
- Jitter is hard to bound even with QoS
- Partial solution: Buffer management
  - » Buffers are used to smooth jitter, but how big should the buffer be?
  - » Unexpected jitter can lead to starvation or buffer overflow if buffer is too small
  - » But over-provisioning buffers increases delay

Problems with Throughput

- Estimating throughput is problematic
  - » Media coding may be tunable, but changes in quality are generally perceptually annoying
  - » Throughput will often vary faster than feedback delay which makes estimation hard
- Multicast makes this worse
- Possible Solutions:
  - » QoS mechanisms
  - » Scalable encoding combined with congestion control (I.e. dynamically adapt compression to bandwidth)

Problems of Loss

- Effect of loss is media and encoding specific
  - » Loss of one packet in a video stream may result in the loss of multiple frames
  - » Other packets received can be rendered useless
  - » Retransmission often not feasible (no time)
- Possible Solutions:
  - » Encoding that alleviates packet loss effects
  - » Forward Error Correction (FEC)

Overview

- Multimedia applications
- Current challenges of multimedia networking
- Media Encoding
- Methods of streaming stored media
- Case Studies:
  - » Real Time Voice over IP (VoIP)
  - » Video over IP
- Real Time Protocol (RTP)
- Real Time Control Protocol (RTCP)

Media Encoding

- Process has three steps on each side
- Digitize: represent information in bits
  - » Sample, quantize, (eg. PCM)
- Compress: reduce number of bits
  - » Audio: GSM, G.729, G.723.3, MP3, …
  - » Video: MPEG 1/2/4, H.261, …
- Send over the network
- Reverse process on receive side: uncompress, convert, play

Video Encoding

- Once frames are captured ("raw" video) resulting file is very large:
  - » 320 x 240 x 24-bit color = 230,400 bytes/frame
  - » 15 frames/second = 3,456,000 bytes/second
  - » 10 seconds takes around 30 Mbytes! (no audio)
- Commonly-used encodings: AVI, MPEG
  - » Per-frame versus inter-frame encoding
  - » Layered encoding

Image Encoding: JPEG

- Divide digitized image in 8 x 8 pixel blocks.
- The DCT phase converts the pixel block into a block of frequency coefficients.
  - » Discrete Cosine Transform – similar to FFT
- The quantization phase limits the precision of the frequency coefficient.
  - » This is based on a quantization table, which controls the degree of compression
- The encoding phase packs this information in a dense fashion.
  - » One block at a time – include the DC component

MPEG Encoding of Video

- MPEG uses inter-frame encoding.
  - » Exploits the similarity between consecutive frames
- Three frame types:
  - » I frame: independent encoding of the frame (JPEG)
  - » P frame: encodes difference relative to I-frame (predicted)
  - » B frame: encodes difference relative to interpolated frame
  - » Note that frames will have different sizes
- Complex encoding, e.g. motion of pixel blocks, scene changes, ..
  - » Decoding is easier than encoding.
- MPEG often uses fixed-rate encoding.

I	B	B	P	B	B	P	B	B	I	B	B	P	B	B

MPEG System Streams

- Combine one more MPEG video and audio streams in a single synchronized stream.
- Consists of a hierarchy with meta data at every level describing the data.
  - » System level contains synchronization information
  - » Video level is organized as a stream of group of pictures
  - » Group of pictures consists of pictures
  - » Pictures are organized in slices
  - » …

Audio Encoding

- Traditional telephone quality encoding: 8KHz samples of 8 bits each.
- CD quality encoding: 44.1 KHz of 16 bits.
  - » 1.41 Mbs uncompressed
- MP3 compression similar to MPEG.
  - » Frequency ranges that are divided in blocks, which are converted using DCT, quantized, and encoded

	Range	Ratio
Layer 1	384 kbps	4
Layer 2	192 kbps	8
Layer 3	128 kbps	12

Overview

- Multimedia applications
- Current challenges of multimedia networking
- Media Encoding
- Methods of streaming stored media
- Case Studies:
  - » Real Time Voice over IP (VoIP)
  - » Video over IP
- Real Time Protocol (RTP)
- Real Time Control Protocol (RTCP)

Streaming Protocols

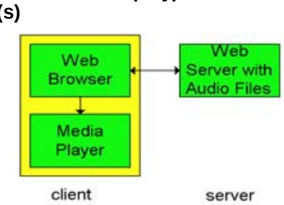
- Important and growing application due to:
  - » Reduction of storage costs
  - » Increase in high speed net access (eg. DSL)
  - » Enhancements to network caching of video using content distribution networks
- User interactive control
  - » Real Time Streaming Protocol (RTSP)
- Transport layer support for audio and video
  - » Real-Time Transport Protocol (RTP)

Streaming Helper Application

- Displays content, which is typically requested via a Web browser (e.g. RealPlayer)
- Typical functions:
  - » GUI for user control
  - » Decompression
  - » Jitter removal
  - » Error correction

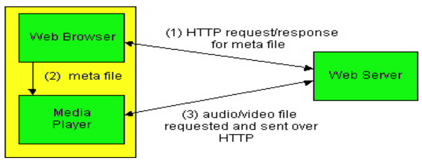
Streaming From Web Servers (1)

- Video (interleaved audio and images in one file, or two separate files and client synchronizes the display) sent as HTTP object(s)



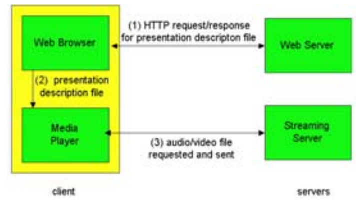
Streaming From Web Servers (2)

- Alternative: set up connection between server and player, then download

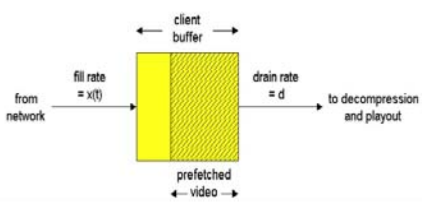


Using a Streaming Server

- Application layer protocol can be better tailored to streaming
  - » Gets us around HTTP, allows a choice of UDP vs. TCP

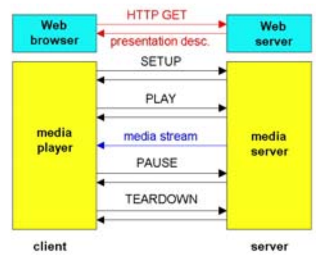


Options When Using a Streaming Server



Real Time Streaming Protocol (RTSP)

- Out-of-band protocol that allows users to control transmission of stream
  - » display: rewind, fast forward, pause, resume, etc…
- Think: remote control for your VCR



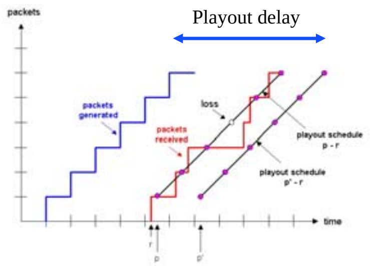
Overview

- Multimedia applications
- Current challenges of multimedia networking
- Media Encoding
- Methods of streaming stored media
- Case Studies:
  - » Real Time Voice over IP (VoIP)
  - » Video over IP
- Real Time Protocol (RTP)
- Real Time Control Protocol (RTCP)

Real-Time Voice Over IP (VoIP): How it works (simply)

- The coded voice information is encapsulated and sent out typically using UDP
- Delay
  - » End-to-end delays > 400 msec cannot be tolerated
  - » Packets that are that delayed are ignored (i.e. Lost)
- Jitter
  - » Use timestamps, sequence numbers
  - » Delay playout at receivers (fixed or adaptive)
- Packet loss
  - » Forward Error Correction (FEC)
  - » Interleaving
  - » Receiver-based repair

Compensating for Jitter: VoIP with Fixed Playout Delay



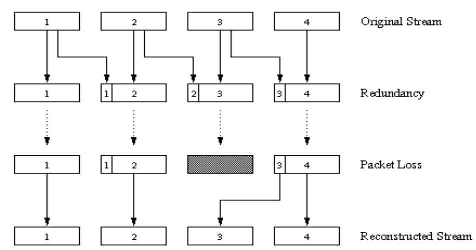
Compensating for Jitter: Adaptive Playout Delay

- Objective is to use a value for playout schedule that tracks the network delay performance as it varies
- Estimated average delay and deviation of average delay are computed in a manner similar to estimates of RTT and deviation in TCP
- The playout delay is then adjusted, typically during silent periods

Recovery from Packet Loss: Forward Error Correction (FEC)

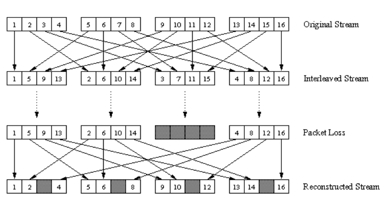
- Send redundant encoded chunk every *n* chunks (XOR original n chunks)
  - » If 1 packet in this group lost, can reconstruct
  - » If > 1 packet lost, cannot recover
- Disadvantages:
  - » The smaller the group size, larger the overhead
  - » Playout delay increased

Recovering from Packet Loss: Piggybacking Low-Fidelity Stream



Recovering from Packet Loss: Interleaving

- Divide 20 msec of audio data into smaller units of 5 msec each and interleave
- Upon loss, have a set of partially filled chunks

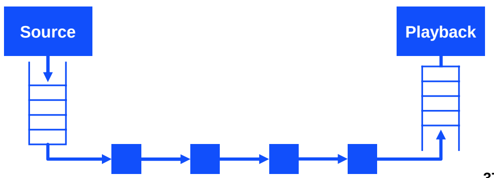


Recovering from Packet Loss: Receiver-based Repair

- Exploit temporal similarity
  - » Replicate packets
- Smart solutions
  - » Interpolation

Playback of Audio/Video

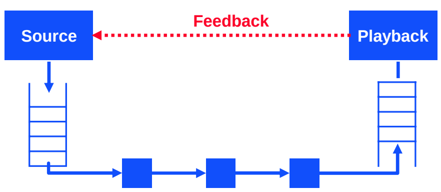
- Source attaches a timestamp to each sample.
- Receiver uses the timestamp to playback the sample at the right time.
- The playback point corresponds to the delay in the replay relative to the generation of the data.
  - » Packets that arrive too late have no value and are ignored
  - » Packets that arrive too early have to be buffered



Transport Protocol Properties

- Reliability.
  - » Some lost data is typically acceptable (impact depends on the type of encoding)
  - » Timeouts typically result in unacceptable delay, and there is often not enough time to retransmit data (live transfers)
- Congestion control.
  - » Nature of the flow fundamentally limits its bandwidth
  - » Reduction of rate in response to congestion should not be done by sending samples slower but by reduce data size
    - – E.g. change frame rate or frame size
- Flow control: natural pacing.
  - » Samples should be paced at the natural rate of the data
  - » Sending too slow --> underflow and missed deadlines
  - » Sending too fast --> buffer overflow and lost data

Adaptive Video Streaming



- Receiver reports dropped packets.
- Sender reduces frame size/rate if loss rate is too high.
  - » Also uses probing to detect additional bandwidth
- Similar to TCP, except that lost data is typically not retransmitted.

Overview

- Multimedia applications
- Current challenges of multimedia networking
- Media Encoding
- Methods of streaming stored media
- Case Studies:
  - » Real Time Voice over IP (VoIP)
  - » Video over IP
- Real Time Protocol (RTP)
- Real Time Control Protocol (RTCP)

Real Time Transport Protocol (RTP)

- Multimedia senders append header fields before passing to transport layer
  - » Format, sequence numbers, timestamps, ...
- RTP logically extends UDP.
  - » Sits between UDP and application
  - » Implemented as application library
- RTP does not guarantee timely data delivery.
  - » Simply helps applications with formatting and the collection of session information
  - » Guarantees have to be provided by lower level protocols
- The protocol has two parts.
  - » Real-time Transport Protocol: carry data
  - » Real-Time Control protocol: monitor quality, exchange information about participants, ..

RTP: What does it do?

- Framing
- Multiplexing
- “Real-time delivery”
- Synchronization
- Feedback

RTP Packet Format

- Source/Payload type
  - » Different formats assigned different codes
  - » Eg. GSM -> 3, MPEG Audio -> 14
- Sequence numbers
- Time stamps
- Synchronization source ID
- Miscellaneous fields

RTP Header

Timestamp vs. Sequence No.

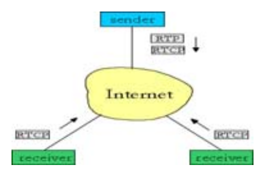
- Timestamp relates packet to real time
  - » Timestamp values sampled from a media specific clock
- Sequence number relates packet to other packets
- Allows many packets to have the same timestamp but different sequence numbers

Audio silence example

- Consider audio data type: what do you want to do during silence?
  - » Not send anything
- Why might this cause problems?
  - » Other side needs to distinguish between loss and silence
- How does the timestamp/seq. no mechanism help?
  - » After receiving no packets for a while, next packet received will reflect a big jump in timestamp, but have the correct next seq. no. Thus, receiver knows what happened

Real Time Control Protocol (RTCP)

- RTCP packets transmitted by each participant in RTP session to all others using multicast
- Distinct port number from RTP
- Reports on:
  - » Loss rate
  - » Inter-arrival jitter
  - » Identity of receivers
  - » Delay, (indirectly).



Control Bandwidth Scaling

- Needed for scalability
  - » Consider one speaker and 2 receivers as compared to one speaker with 20 or 200 receivers
- Recommendation:
  - » 5% of data bandwidth
  - » Sending delay should be randomized
  - » Should account for overhead of underlying network service (i.e., UDP and IP)

H.323: Standard for Real-Time Conferencing

- Components:
  - » Standalone endpoints: terminals
  - » Gateways: permit communication between endpoints and circuit-switched phones
  - » Gatekeepers: provide address translation, authorization, bandwidth management, …