

Multimedia Applications
&
User Requirements

Kevin Jeffay

Department of Computer Science
University of North Carolina at Chapel Hill

jeffay@cs.unc.edu

August 24, 1999

<http://www.cs.unc.edu/~jeffay/courses/comp249f99>

1

Course Outline

Introduction to multimedia

- ◆ Distributed multimedia applications
 - » What are they and why do we want them?
 - » User and application requirements

- ◆ Audio and video as a data type
 - » Audio & video coding and compression technologies
 - » Audio/Video “protocols”

2

Distributed Multimedia Applications

Examples

- ◆ Entertainment
 - » Video-on-demand
 - » Multi-player games
- ◆ Collaborative work
 - » Remote consultation
- ◆ Distance learning
 - » Interactive television
 - » Content-on-demand
- ◆ Communication
 - » Internet telephony & videoconferencing

3

Interactive Multimedia Applications

Why do we want them?

- ◆ Who wouldn't love videoconferencing?
 - » Reduces the need for travel
 - » Reduces the need for physical proximity of co-workers
- ◆ Surely it's better than a phone call?
 - » Enhances audio communication with non-verbal cues
 - » Contributes to team and consensus building
 - » Enables remote demonstrations and "show and tells"

4

Interactive Multimedia Applications

Why do we want them?

- ◆ Numerous user studies have found:
 - » Groups are *neither* more effective *nor* efficient at problem solving when using videoconferencing than when using the telephone

- ◆ Numerous commercial products have either failed or at best not succeeded
 - » AT&T Picturephone
 - » Intel ProShare system
 - » IBM Person-to-Person system
 - » ...

5

Interactive Multimedia Applications

The Picturephone debacle

- ◆ A product of AT&T from 1964-1978
- ◆ Excellent quality
 - » 5.5 x 5 inch, black and white picture
 - » 250 scan lines (TV quality)
 - » 30 frames/s (perfect motion perception)
- ◆ Use
 - » Used handset or speakerphone for audio
 - » Display could show either local or remote user
- ◆ Cost
 - » \$75/month
 - » \$0.15/minute (first 30 minutes free)



6

The Picturephone Debacle

The marketing guys said...

- ◆ By 1975: 50,000 units
 - » Deployed in 25 cities
- ◆ By 1980: 1,000,000 units
 - » In 3% of all businesses and 1% of all homes
- ◆ By 1985: 2,000,000 units

Before the turn of the century, Picturephone will displace today's means of communication, and in addition will make many of today's trips unnecessary

— J.P. Molnar

Bell Labs executive vice president, 1969

7

The Picturephone debacle

Do we really want video-phones?

- ◆ Public picturephone booths were deployed in New York, Chicago, and Washington DC (1964-1970)
 - » 3 calls were placed in 1969
 - » 0 calls in 1970
- ◆ Local picturephone service deployed in Chicago & Pittsburgh (1971-1977)
 - » less than 100 users in Chicago in 1973
 - » 9 users in 1977
- ◆ Specialized test markets & free promotions
 - » public room teleconferencing
 - » numerous large corporations
 - » hospitals

all failed to create a market

8

The Picturephone debacle

What went wrong?

- ◆ No *compelling* uses of the technology
- ◆ Lack of inertia — need other picturephones to talk to
- ◆ Users didn't like it
 - » People were too uncomfortable & self-conscious on camera
 - » It didn't "add enough" to audio-only communications
 - » Wasn't close enough to face-to-face communication

9

The Picturephone debacle

Would the situation be different on the Internet today?

- ◆ Are the applications for this technology here?
 - ◆ Is the inertia problem any better?
 - » Much better for simple Internet telephony
 - ◆ Will post-MTV generation users accept video more today?
 - » Not likely — non-verbal cues still missing or wrong
 - ❖ no eye contact
 - ❖ when eye contact exists, it's not directed to any one participant
 - ❖ inflexible proximity & orientation
- and wrong cues are harder to compensate for than missing cues!*
- ◆ And Internet multimedia applications are not likely to perform well!

10

Interactive Multimedia Applications

Can interactive audio and video provide *any* benefits?

- ◆ They do not yet impact the *product* but they can impact the *process*
 - » Video helps management of the mechanics of conversation
 - » Potentially makes more efficient use of the time spent communicating
- ◆ But these benefits are realized *only if* critical performance criteria are met:
 - » Throughput
 - » End-to-end latency
 - » Delay-jitter
 - » Loss

11

Interactive Multimedia Applications

Performance requirements

- ◆ Latency — the duration between acquisition of a signal and its display

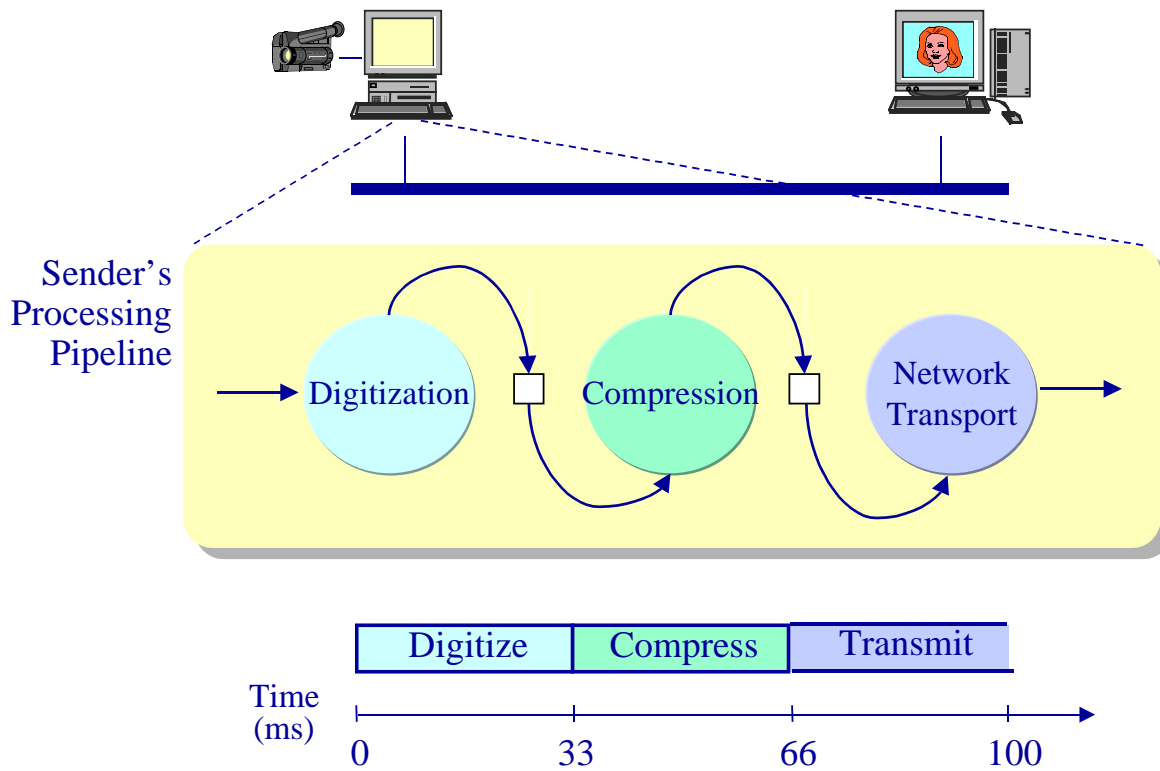


- ◆ Videoconferencing latency requirements
 - » telephony literature — 100 ms roundtrip
 - » multimedia networking literature — 250 ms one-way
 - » CSCW literature — tolerance of latency as high as 400 ms

12

Latency in Computer-Based Video Systems

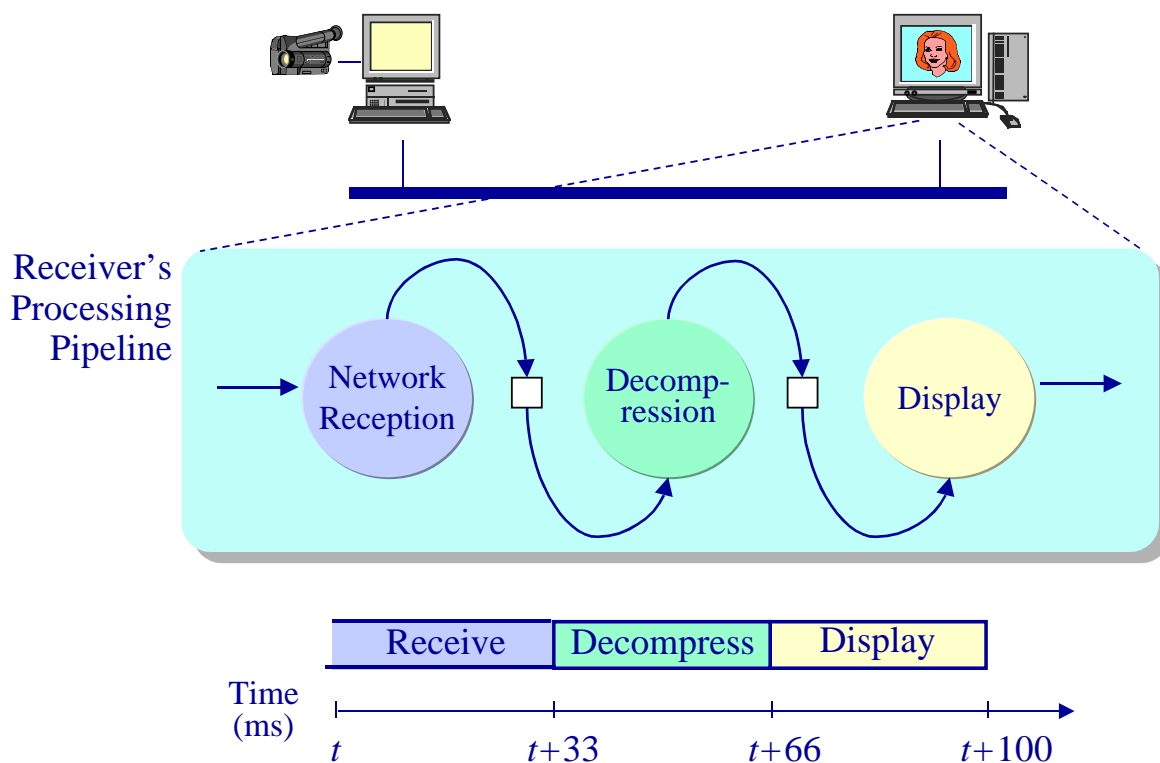
Canonical application structures



13

Latency in Computer-Based Video Systems

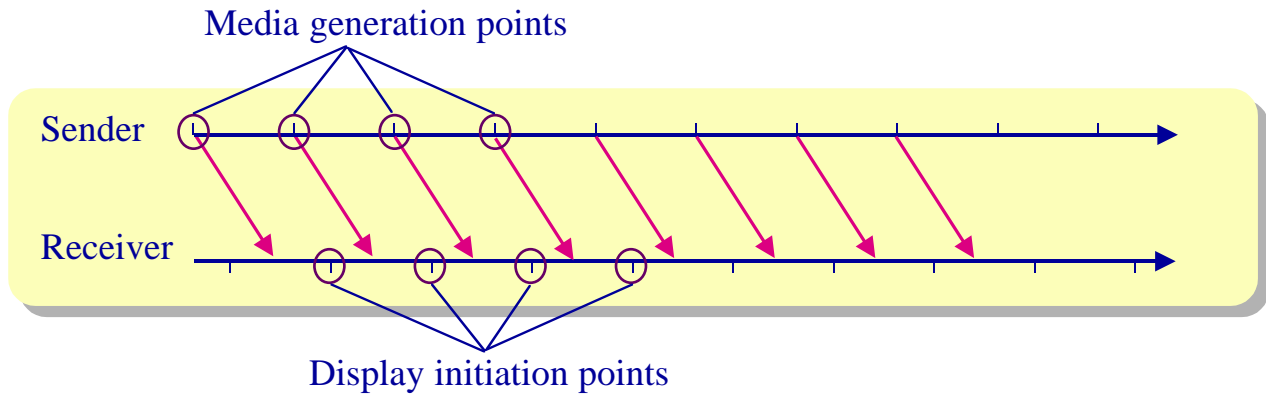
Canonical application structures



14

Latency in Computer-Based Video Systems

Receiver synchronization

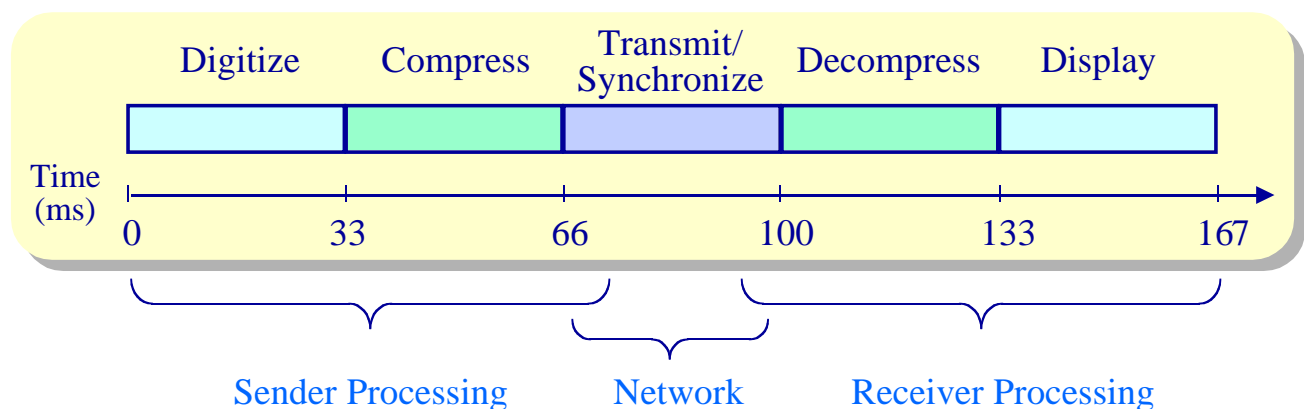


- ◆ In general, acquisition and display clocks are not synchronized...
- ◆ Therefore a buffer must be present at the receiver to adjust for phase-shift in sender's & receiver's media clocks
 - » Greatest lower bound on buffer residency is $33 - \epsilon$ ms

15

Latency in Computer-Based Video Systems

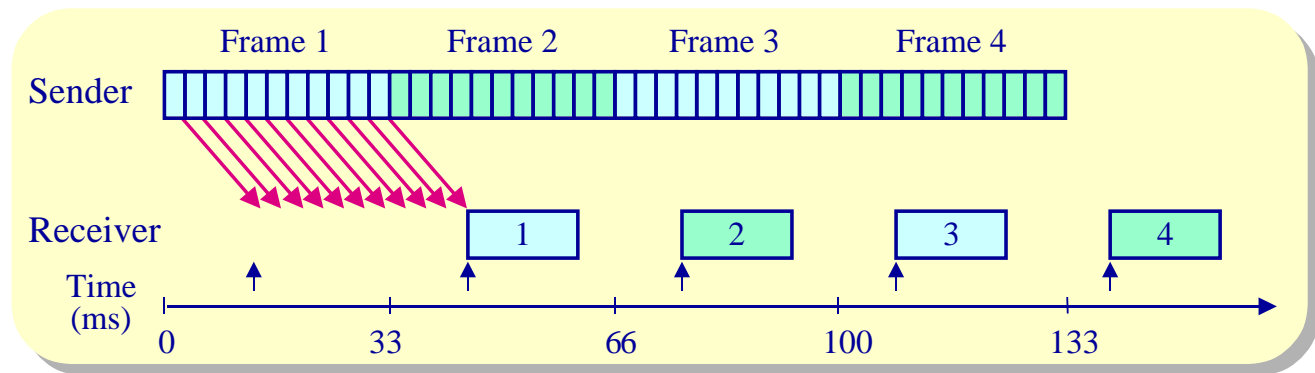
Best case end-to-end latency



16

Latency in Computer-Based Video Systems

Best case end-to-end latency (hypothetical)



◆ Assuming...

- » Compression/decompression times go to 0...
- » Displays with high-refresh rates...
- » “Early” transmission of digitized scan lines...

$$\begin{aligned} \text{latency}_{GLB} &= \text{digitization delay} + \text{network transmission delay} + \text{buffering/synchronization delay} + \text{display delay} \\ &= 16 \text{ ms} + x \text{ ms} + 33 \text{ ms} + 16 \text{ ms} \end{aligned}$$

17

Latency in Computer-Based Audio Systems

How bad can audio latency be?

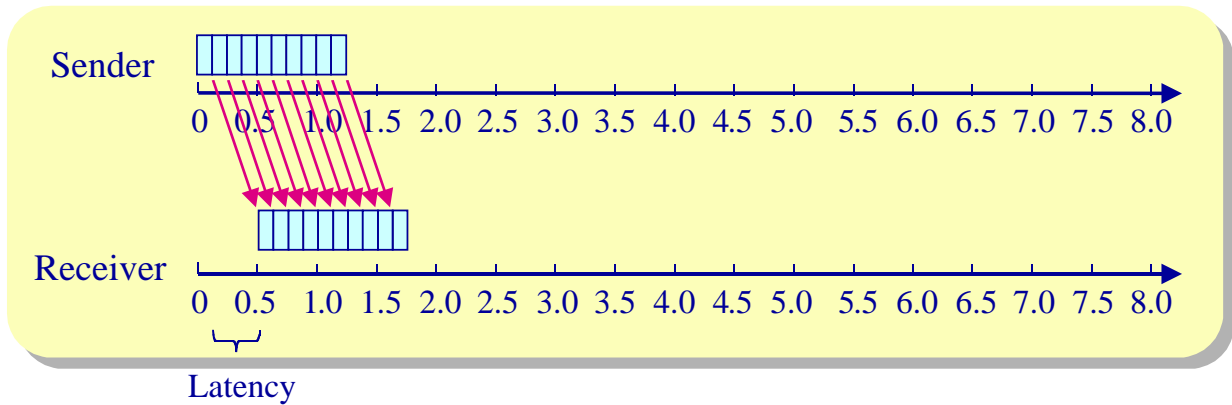
- ◆ Just as bad as video if lip-synchronization is required
 - » (More on this later...)
- ◆ Otherwise, it depends on how one manages the network interface
 - » Video frames are typically too large to fit into a single network packet
 - » Multiple audio samples can be transmitted together hence transmission buffering at the sender effects latency

18

Latency in Computer-Based Audio Systems

Sender buffering & packaging example

- ◆ Consider an audio codec generating 1 byte samples of data every $125 \mu\text{s}$
 - » Transmit 1 sample per packet
(No additional latency introduced at the sender)

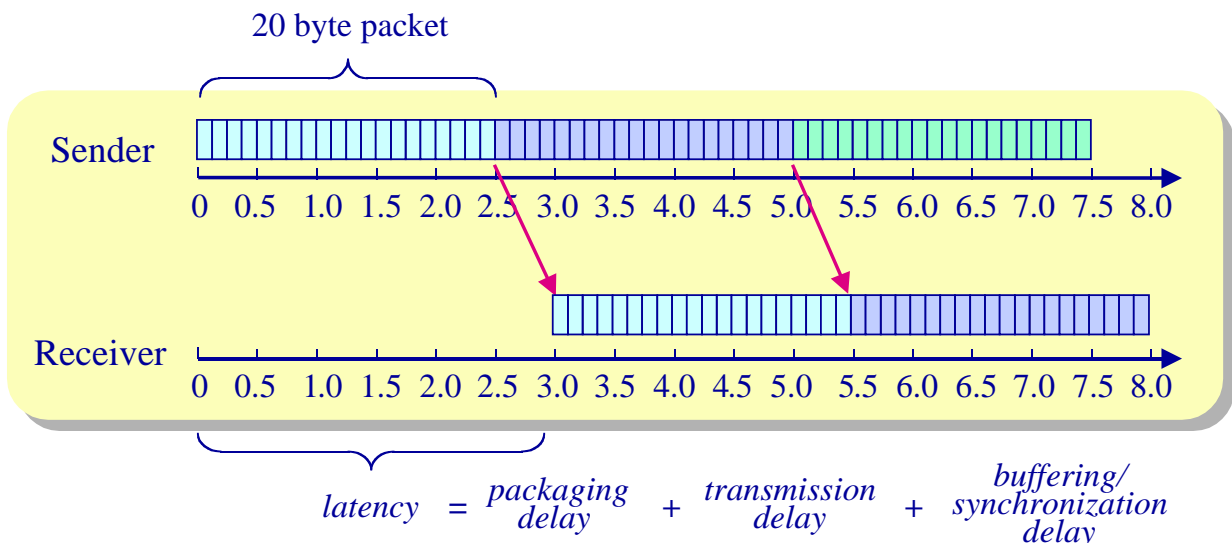


19

Latency in Computer-Based Audio Systems

Sender buffering & packaging example

- ◆ Consider an audio codec generating 1 byte samples of data every $125 \mu\text{s}$
 - » Transmit 20 samples per packet



20

Latency in Computer-Based Audio Systems

Sender buffering & packaging

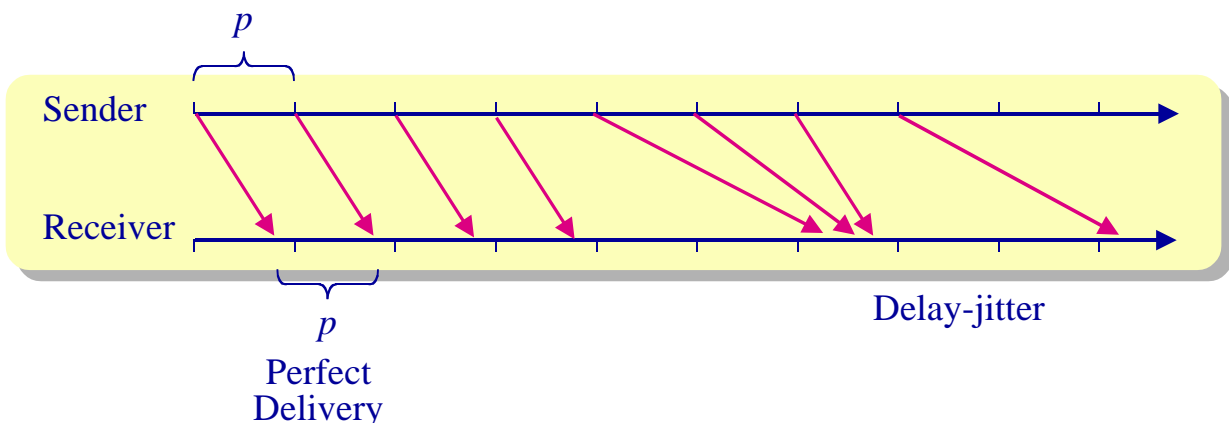
- ◆ Consider an audio codec generating 1 byte samples of data every $125 \mu\text{s}$
- ◆ More realistic examples
 - » Building 500 byte packets requires 62.5 ms
 - » Building 1,500 byte packets requires 187.5 ms
- ◆ A canonical Internet audio packet typically contains 20-30 ms worth of audio data
 - » The actual size of the packet depends on the coding/compression scheme employed

21

Performance Requirements

Delay-jitter

- ◆ Latency
 - » 250 ms one-way
- ◆ Delay-jitter — Variation in end-to-end latency

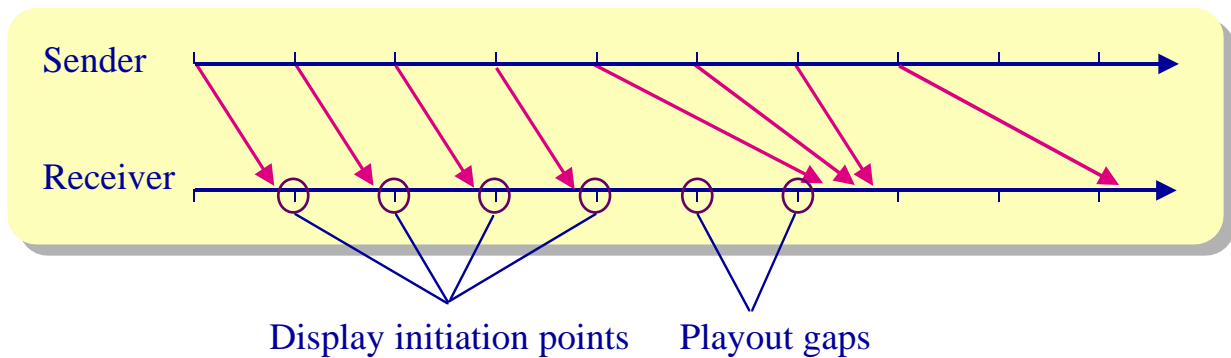


22

Performance requirements

The impact of delay-jitter

- ◆ Delay-jitter leads to “gaps” in the playout of media and increases playout latency

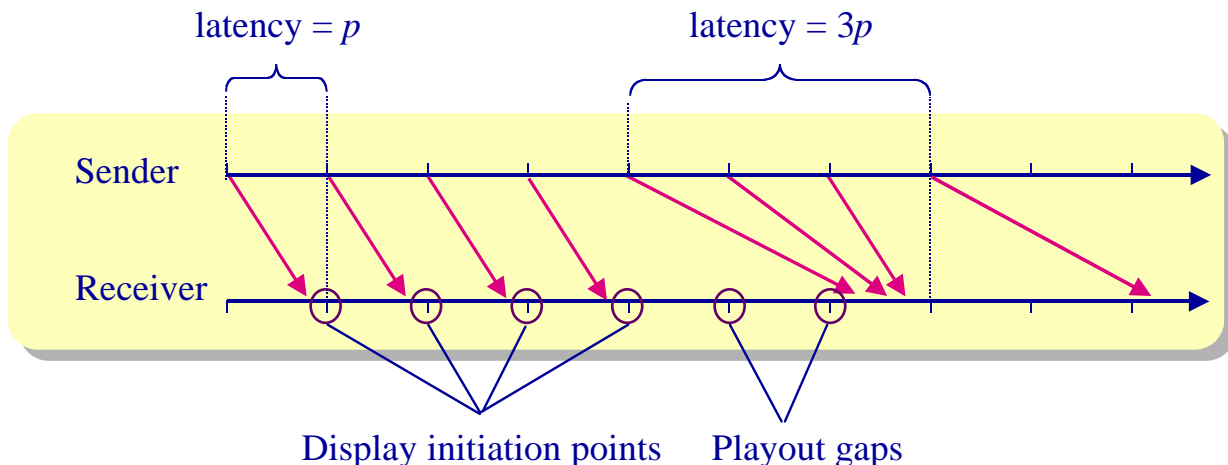


23

Performance requirements

The impact of delay-jitter

- ◆ Delay-jitter increases playout latency



24

Performance Requirements

Throughput and loss

- ◆ Latency
 - » 250 ms one-way
- ◆ Delay-jitter
- ◆ Throughput —the effective delivered frame or sample rate
 - » For video the issue is *motion perception*
 - » For audio the issue is comprehension

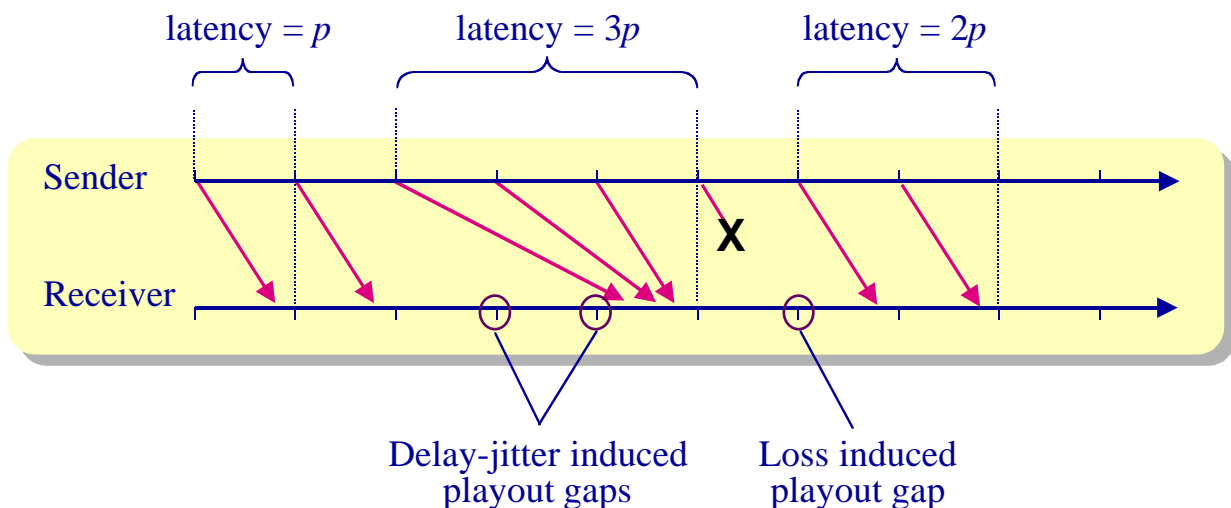
throughput requirements are a (complex) function of coding and compression schemes
- ◆ Loss — the complement of throughput

25

Performance requirements

Loss

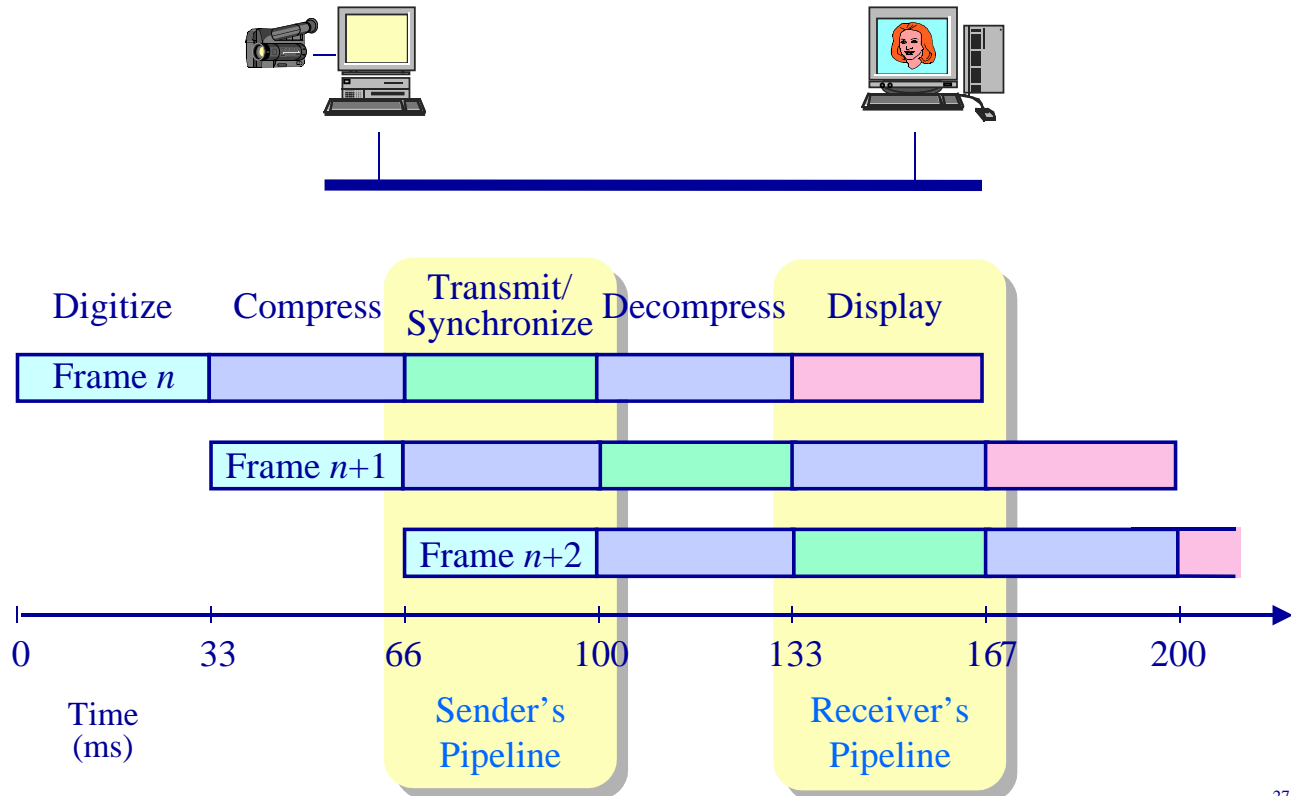
- ◆ Loss has the same effect as delay-jitter: *playout gaps*
 - » With a potentially beneficial effect of potentially lower latency



26

Avoiding Loss in the End System

Real-time management of a processing pipeline



27

Performance requirements

Loss requirements

- ◆ Audio — 1-2% sample loss
 - » individual sample losses (depending on sample size) are noticeable
 - » 5-10 lost samples per minute are tolerable
(the distribution of loss and size of sample are critical parameters)
- ◆ Video — 10-15 frames/s required for minimal motion perception
 - » highly application dependent
 - » video loss raise issues of “network citizenship”

28

Performance Requirements

Lip synchronization

- ◆ Latency
 - » 250 ms one-way
- ◆ Delay-jitter
- ◆ Throughput —the effective delivered frame or sample rate
 - » For video the issue is *motion perception*
 - » For audio the issue is comprehension
- ◆ Loss

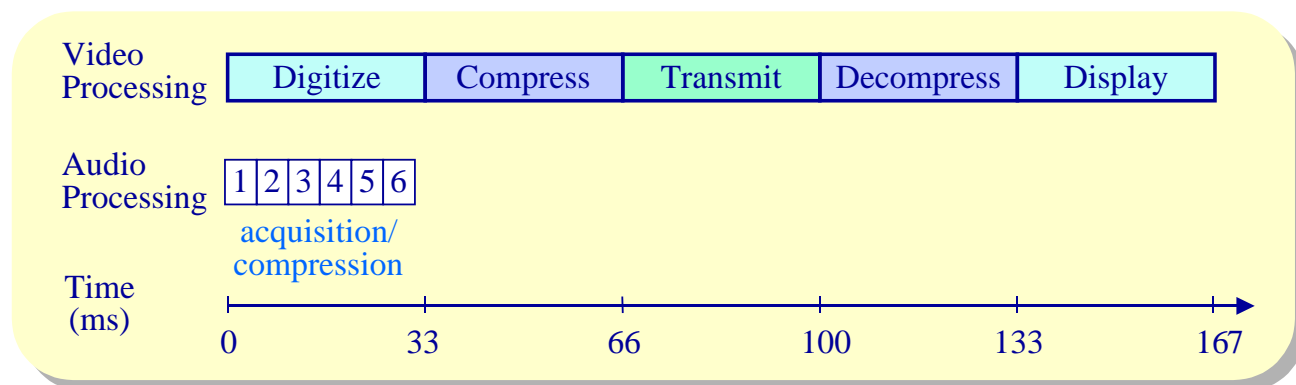
- ◆ Lip synchronization
 - » The temporal relationship between an audio and video stream representing a human speaking

29

Performance Requirements

Lip synchronization

- ◆ Perfect lip synchronization requires audio playout for sample 1 at time _____

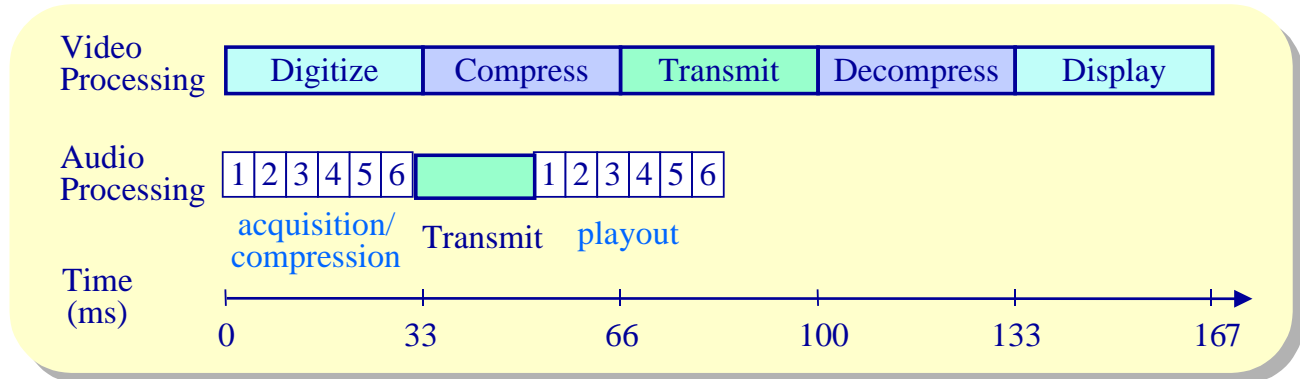


30

Performance Requirements

Lip synchronization

- ◆ Varying lip sync can be an effective technique in mitigating high video latency



- ◆ But... this is fundamentally unnatural!

31

Interactive Multimedia Applications

Performance requirements



- ◆ No more than 250 ms end-to-end, one-way latency
- ◆ Continuous audio
- ◆ Minimum of 10 frames per second video throughput
- ◆ “Loosely synchronized” playout — ± 80 ms skew

32

Multimedia Applications & the Internet

Summary

- ◆ The case for pervasive deployment of distributed, interactive, multimedia applications has yet to be made
- ◆ The Internet is inherently ill-suited to meeting the real-time requirements of real-time multimedia applications
 - » The problems are accidental
- ◆ But there's cool technology and science to be studied!