

Adaptive Congestion Control

Kevin Jeffay

Department of Computer Science
University of North Carolina at Chapel Hill

jeffay@cs.unc.edu

November 23, 1999

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

1

Best-Effort Multimedia Networking

Outline

- ◆ IP message delivery semantics
 - » The four common Internet pathologies
- ◆ Ameliorating the effects of delay-jitter
 - » “60 ways to queue & play your media samples”
- ◆ Ameliorating the effects of packet loss
 - » Recovery of lost samples through retransmission
 - » Recovery of lost samples through the addition of redundant information
- ◆ Congestion control
 - » Adaptive media scaling and packaging

2

Best-Effort Multimedia Networking

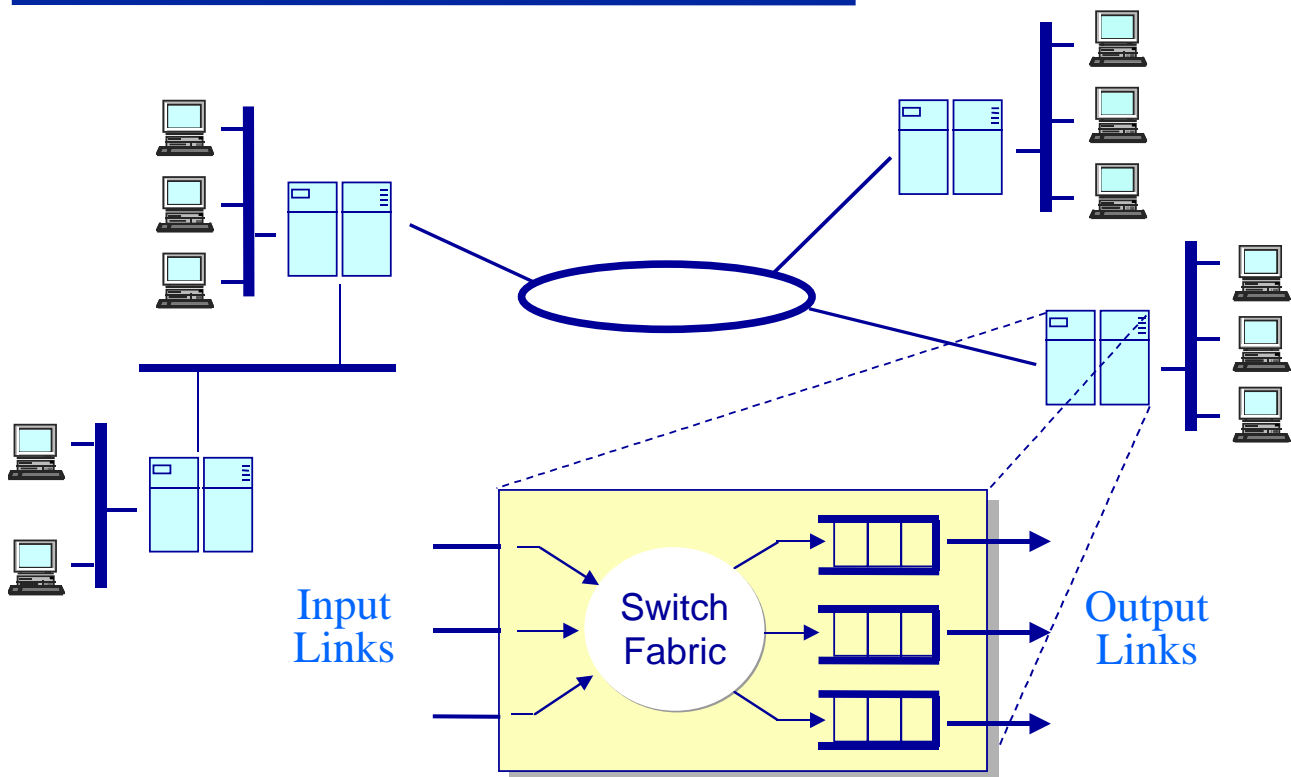
Congestion control

- ◆ Delay-jitter buffering, retransmission, and forward error correction *ameliorate* the effects of variation in end-to-end delay and packet loss
 - » They do not attempt to address the root cause
- ◆ Congestion control aims to eliminate or reduce these effects

3

Congestion Control

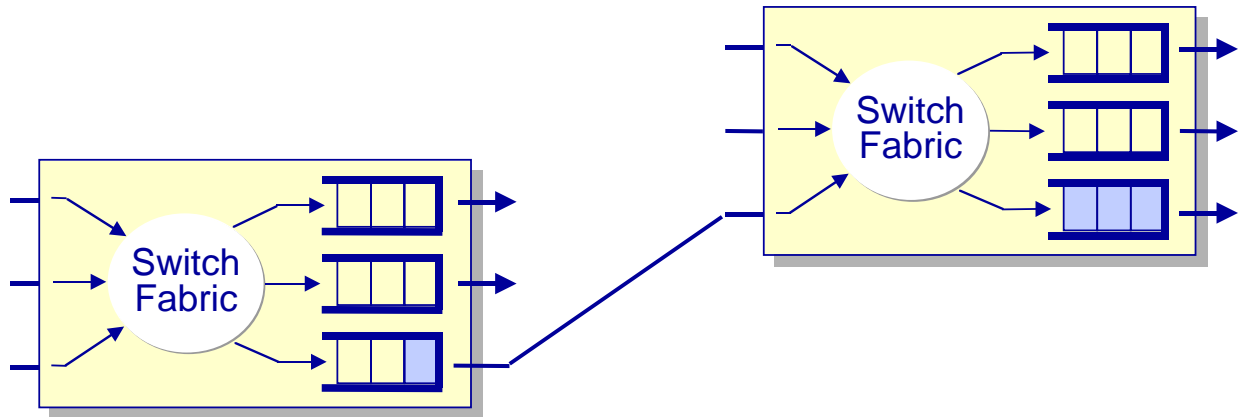
What is congestion?



4

Congestion Control

The nature of congestion



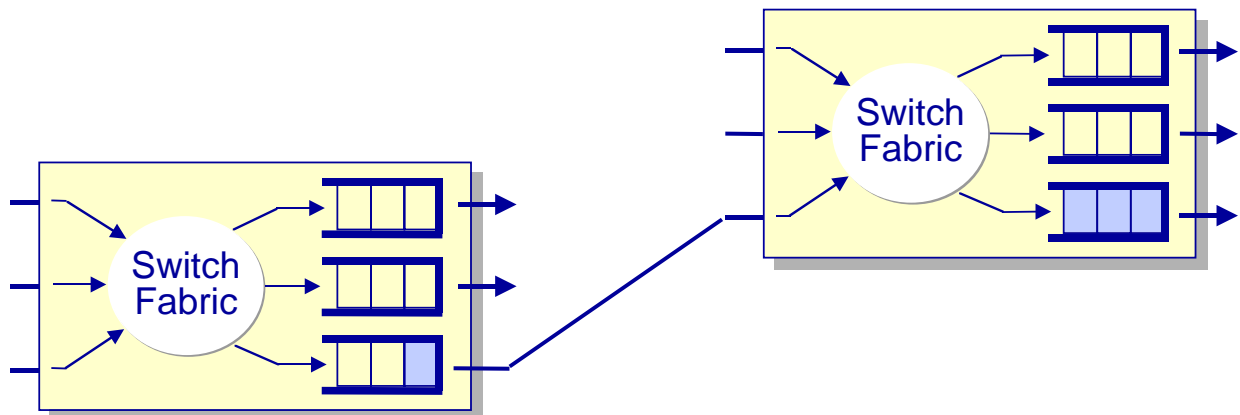
◆ What causes congestion?

- » Did our multimedia stream(s) cause the network to be congested?
- » Are there simply too many connections competing for too little bandwidth?

5

Congestion Control

The adaptive, best-effort, congestion control problem



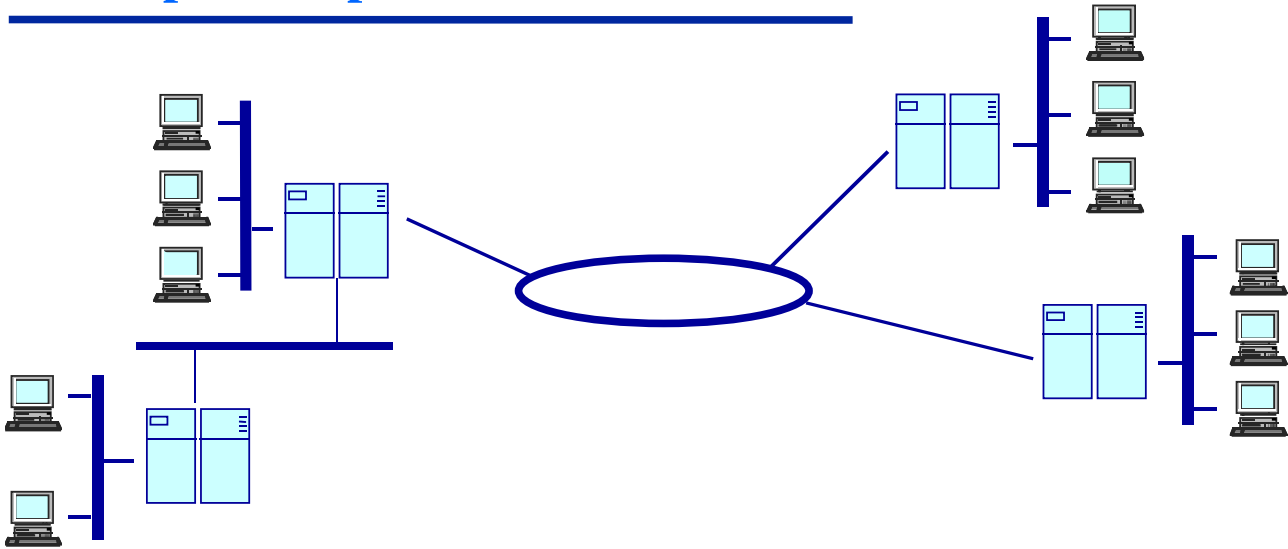
◆ How can we make the best use of the (time varying) bandwidth that is available to our streams?

- » How can we determine what this bandwidth is?
- » How can we track how it changes over time?
- » How can we match our codec(s)'s output the bandwidth "available" to our application?

6

Adaptive, Best-Effort Congestion Control

Principles of operation



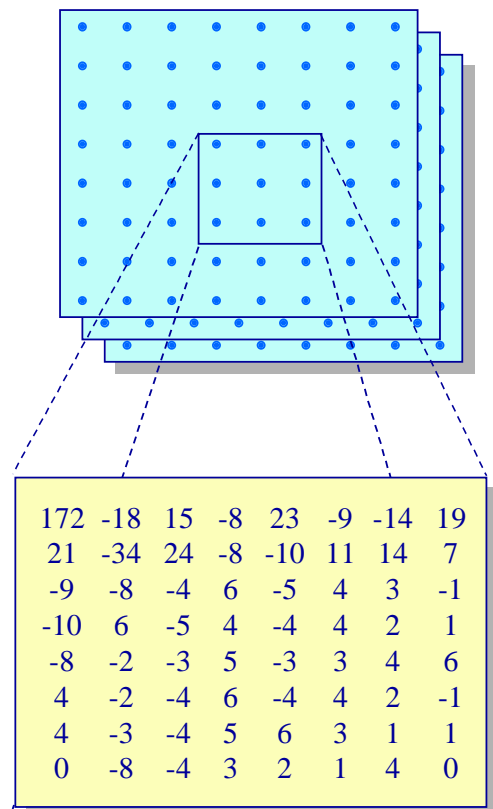
- ◆ Receivers periodically report throughput & loss statistics
- ◆ Sender adapts to match the bandwidth available
 - » Assume sufficient bandwidth exists for some useful execution of the system

7

Canonical Adaptive Congestion Control

Video bit-rate scaling

- ◆ Temporal scaling
 - » Reduce the resolution of the stream by reducing the frame rate
- ◆ Spatial scaling
 - » Reduce the number of pixels in an image
- ◆ Frequency scaling
 - » Reduce the number of DCT coefficients used in compression
- ◆ Amplitude scaling
 - » Reduce the color depth of each pixel in the image
- ◆ Color space scaling
 - » Reduce the number of colors available for displaying the image

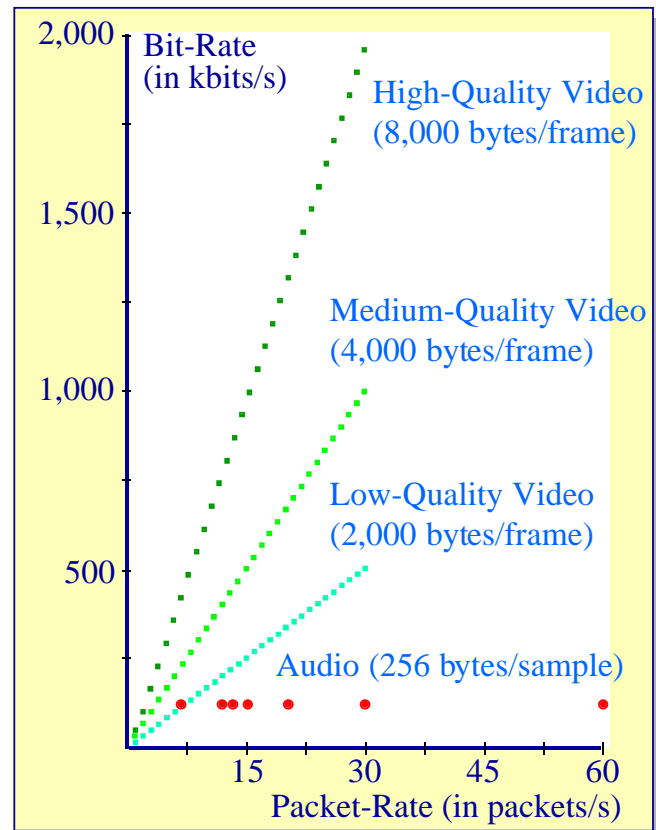


8

UNC Adaptive Congestion Control

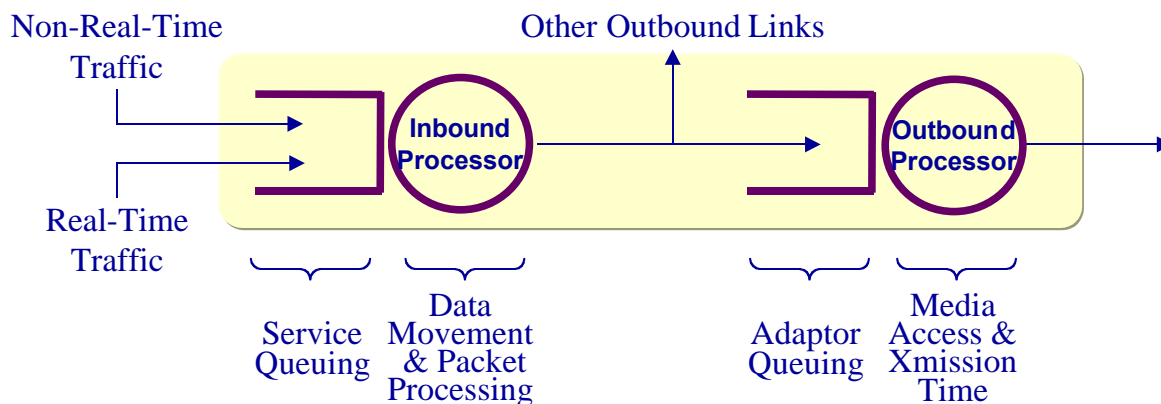
2-Dimensional media scaling

- ◆ Canonical approach to congestion
 - » Reduce (video) bit-rate
- ◆ Alternate approach
 - » View congestion control as a search of a 2-dimensional *bit-rate* \times *packet-rate* space
 - » Scale bit- and packet-rates simultaneously to find a sustainable *operating point*



Bit- and Packet-Rate Scaling

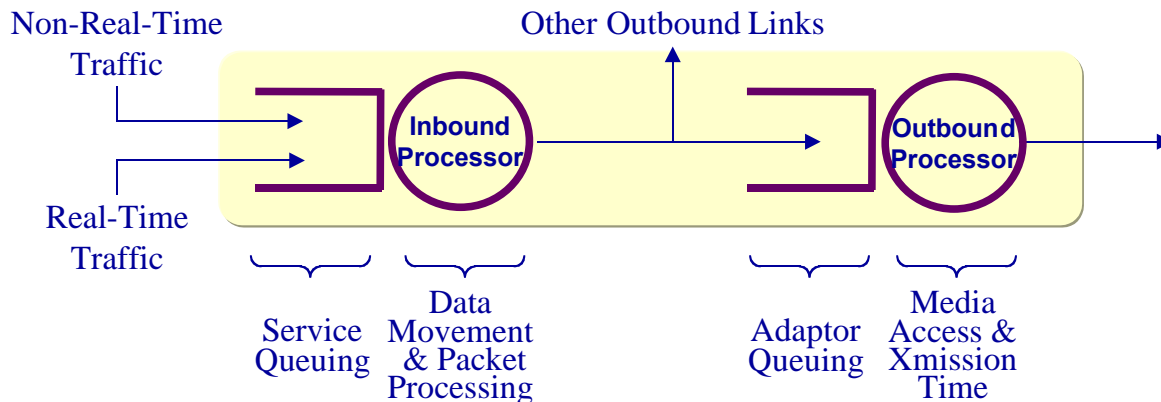
An analytic model of media scaling



- ◆ Capacity constraints
 - » the network is incapable of supporting the desired bit rate in any form
- ◆ Access constraints
 - » the network can not support the desired bit rate with the current packaging scheme

Two Types of Congestion Constraints

Two dimensions of adaptation



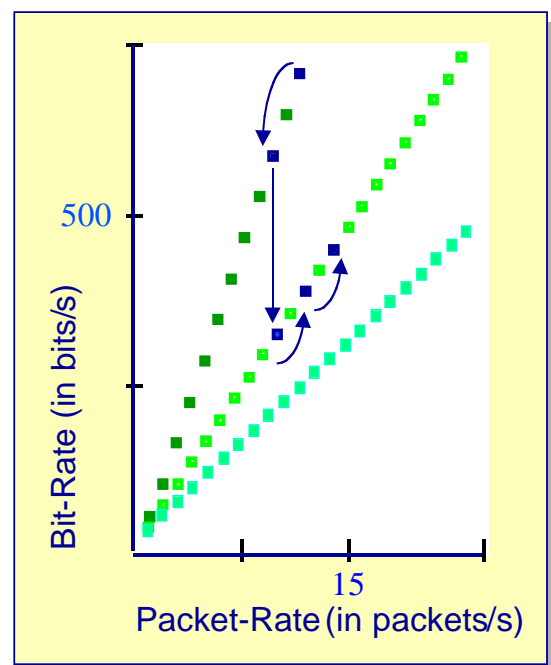
- ◆ Reduce the packet-rate to adapt to an access constraint
 - » Change the packaging or send fewer video frames
 - » Primary Trade-off: higher latency (potentially)
- ◆ Reduce the bit-rate to adapt to a capacity constraint
 - » Send fewer video frames or fewer bits per video frame
 - » Primary Trade-off: lower fidelity

11

2-Dimensional Scaling Example

The “Recent Success” heuristic

- ◆ Initial operating point:
(*high quality, 12 fps*)
- ◆ First adaptation:
(*high quality, 10 fps*)
 - » congestion persists
- ◆ Second adaptation:
(*medium quality, 10 fps*)
 - » congestion relieved
- ◆ First probe:
(*medium quality, 12 fps*)
- ◆ Second probe:
(*medium quality, 14 fps*)

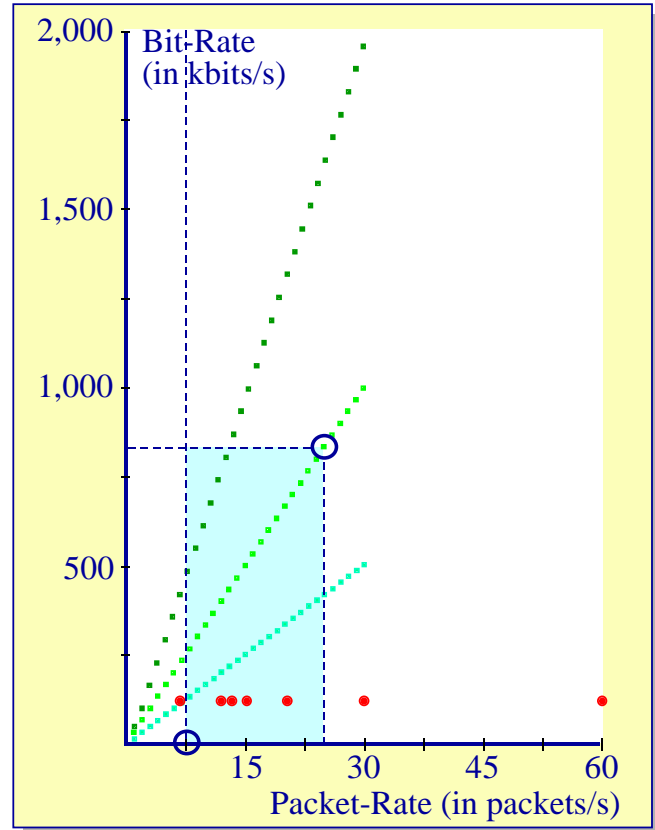


12

2-Dimensional Media Scaling

Finding a sustainable operating point

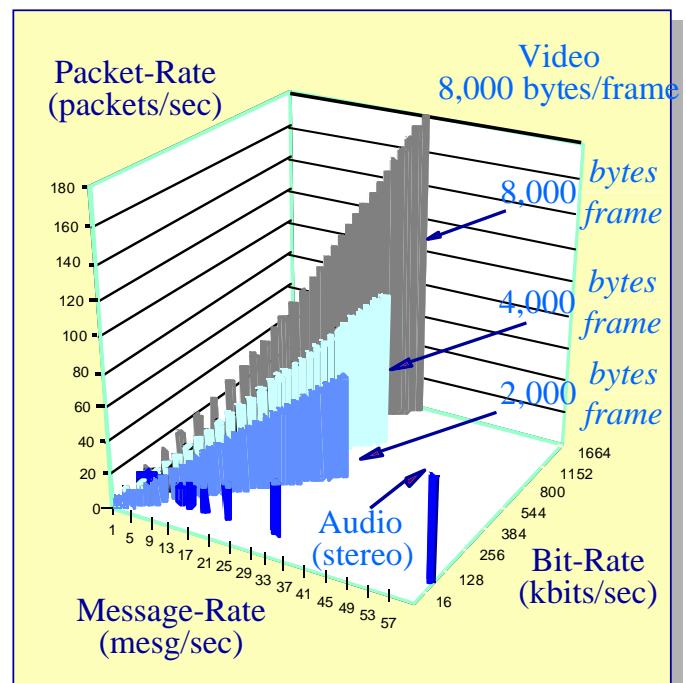
- ◆ The search space can be pruned by eliminating:
 - » Points that inherently lead to high latency
 - » Points that lead to high latency given the state of the network



2-Dimensional Media Scaling

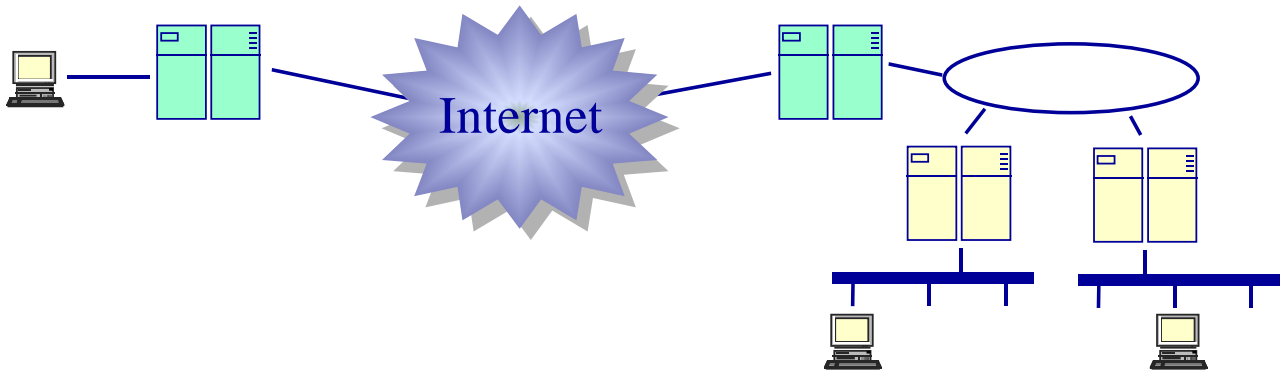
Dealing with effects of fragmentation

- ◆ The problem
 - » A sender can only (directly) effect the *message rate*, not the *packet rate*
- ◆ Does fragmentation render message-rate scaling obsolete?



Adaptive, 2-Dimensional Media Scaling

Does it work?

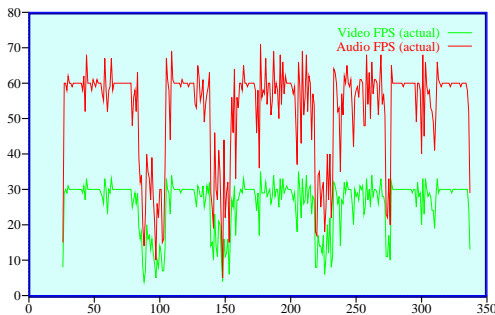


- ◆ Campus-sized internets — yes!
 - » It “solves” the first-mile/last-mile problem

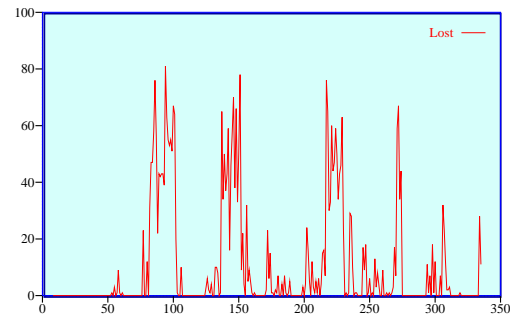
15

Media Scaling Evaluation on the UNC Campus

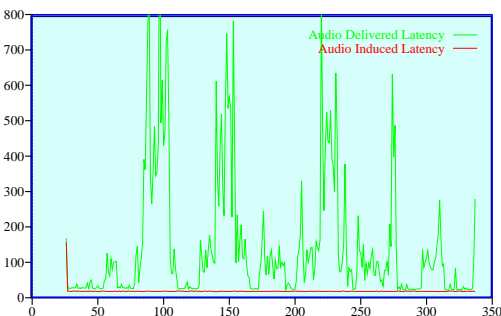
Performance with no media scaling



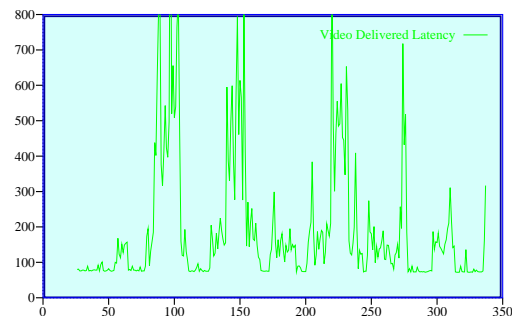
Throughput (frames/sec)



Packet Loss



Audio Latency (ms)

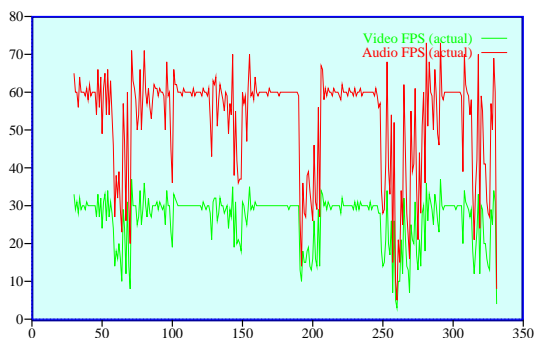


Video Latency (ms)

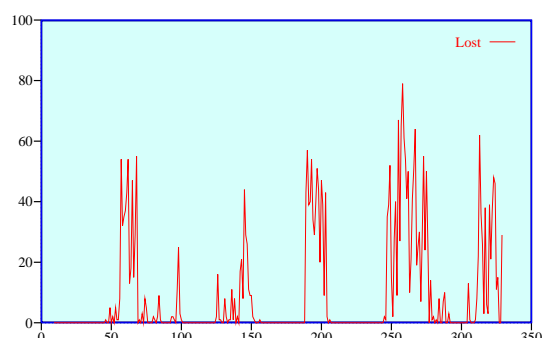
16

Media Scaling Evaluation on the UNC Campus

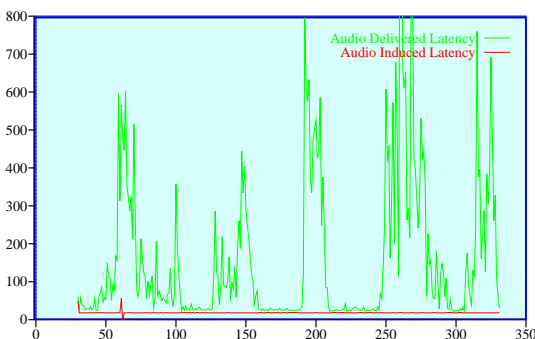
Performance with video scaling only



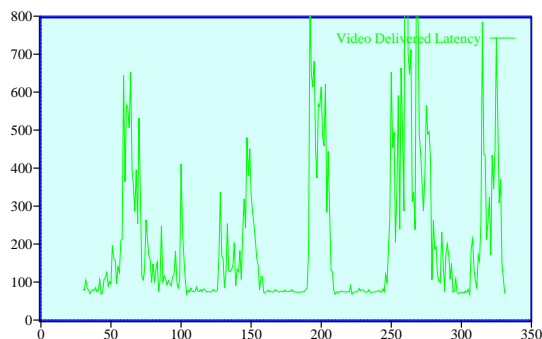
Throughput (frames/sec)



Packet Loss



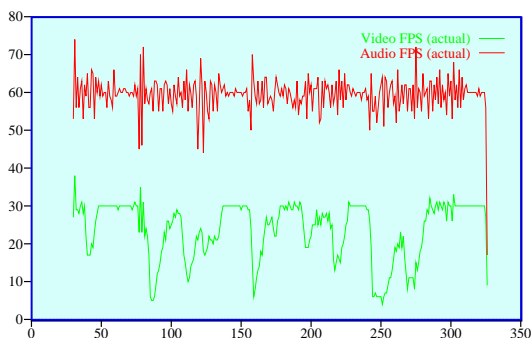
Audio Latency (ms)



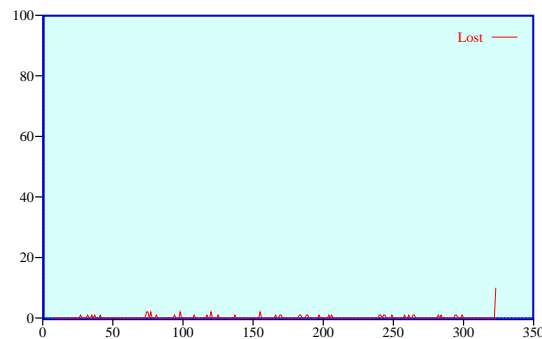
Video Latency (ms)

Media Scaling Evaluation on the UNC Campus

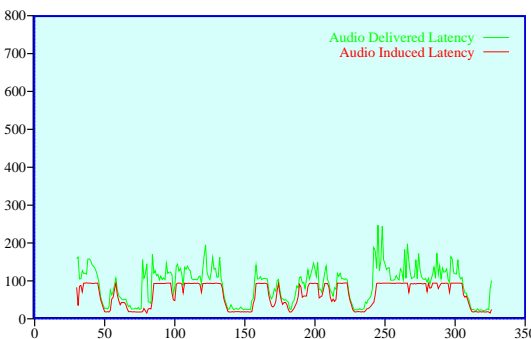
Performance with 2-dimensional scaling



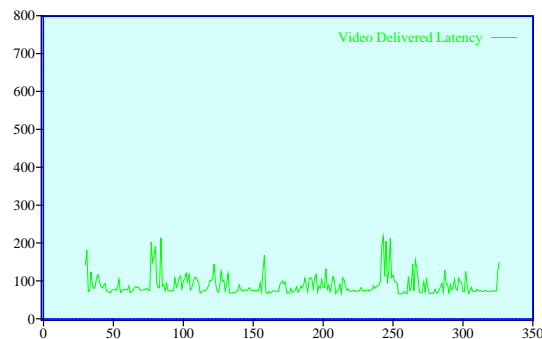
Throughput (frames/sec)



Packet Loss



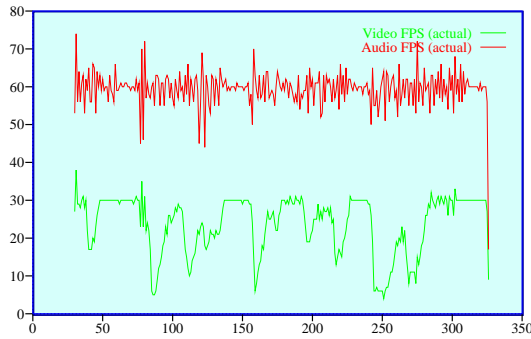
Audio Latency (ms)



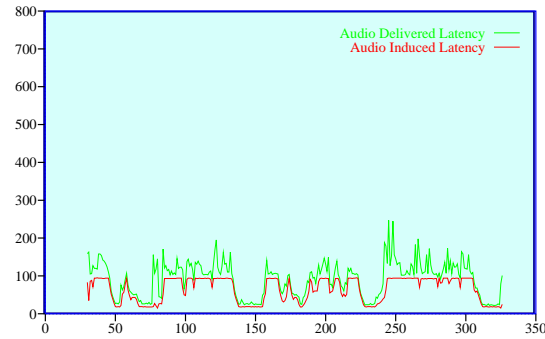
Video Latency (ms)

Media Scaling Evaluation on the UNC Campus

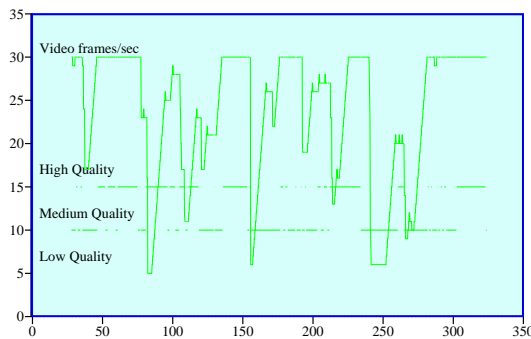
2-dimensional adaptation over time



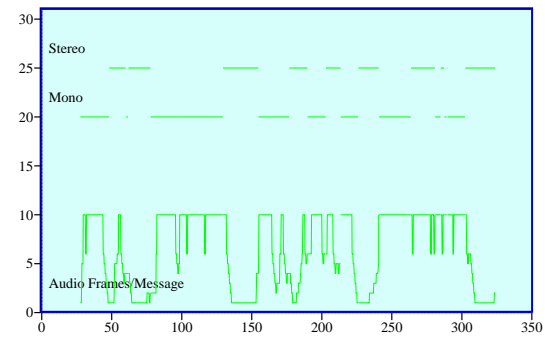
Throughput (frames/sec)



Audio Latency (ms)



Video Adaptation Over Time

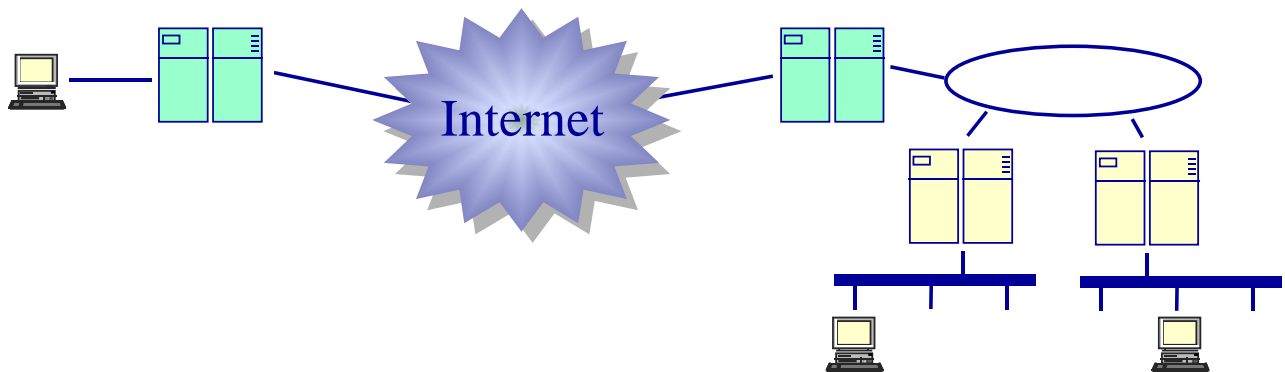


Audio Adaptation Over Time

19

Adaptive, 2-Dimensional Media Scaling

Does it work?

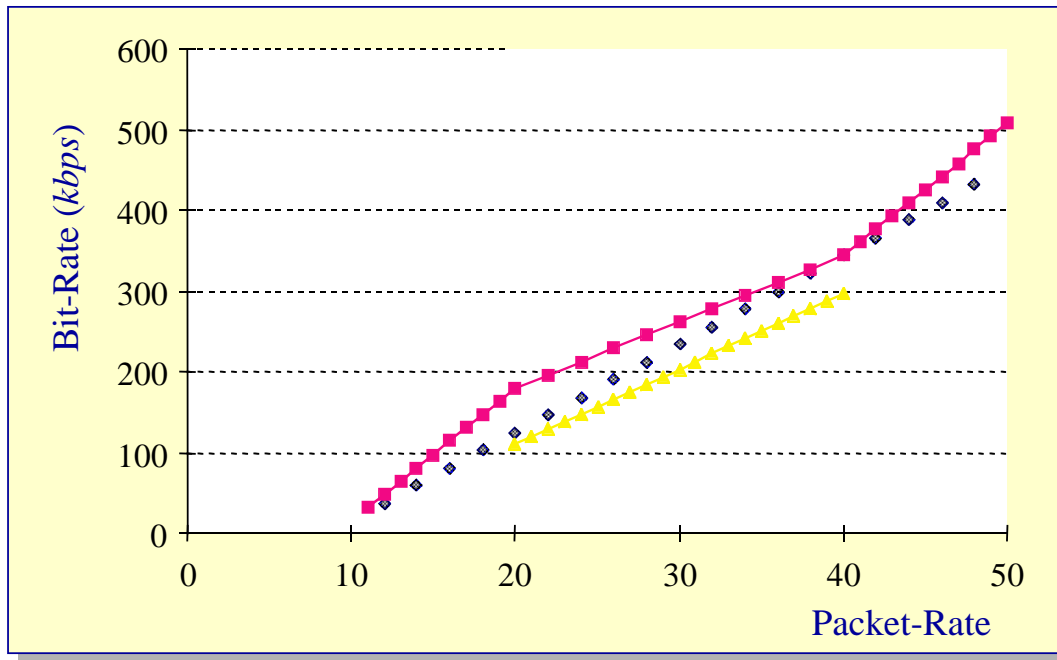


- ◆ Campus-sized internets — yes!
 - » It “solves” the first-mile/last-mile problem
- ◆ The Internet? — *well...*
 - » Does our necessary condition for success hold?
 - » Does it hold often enough to be useful?
 - » How much “room” is there for 2-D scaling in most codecs?

20

Media Scaling Evaluation on the Internet

Media scaling in Intel's ProShare™ codec

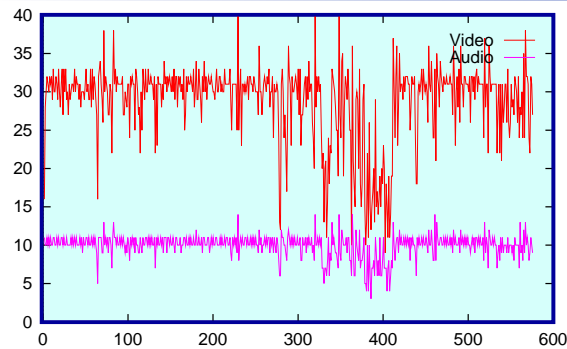


Proshare operating points

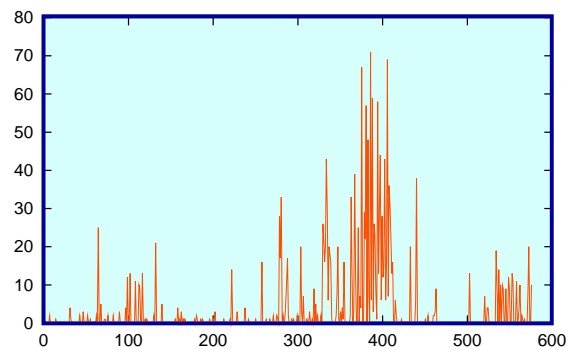
21

Media Scaling Evaluation on the Internet

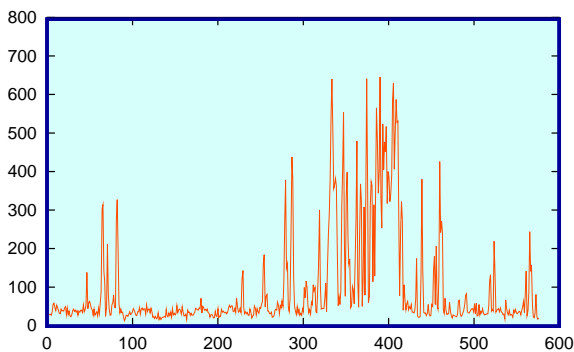
ProShare with no media scaling



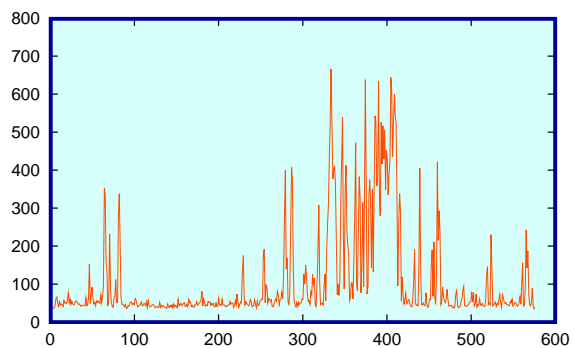
Throughput (frames/sec)



Packet Loss



Audio Latency (ms)

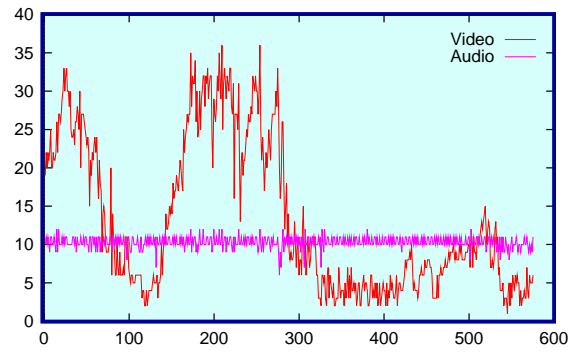


Video Latency (ms)

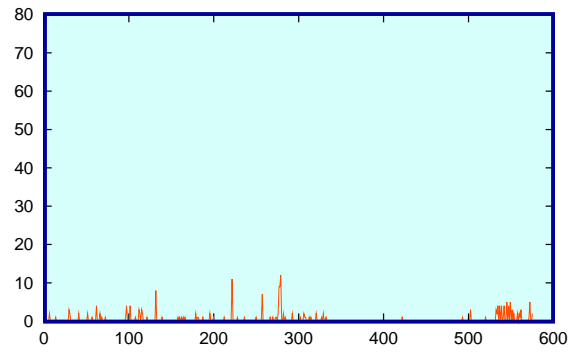
22

Media Scaling Evaluation on the Internet

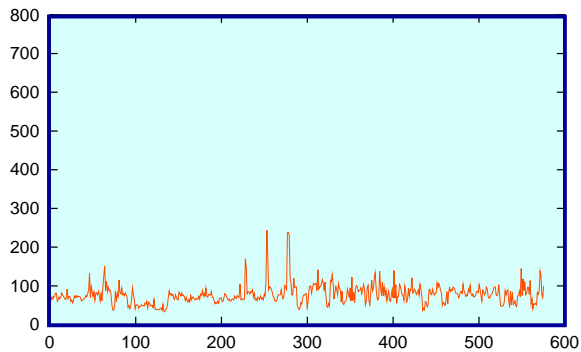
ProShare with 2-dimensional media scaling



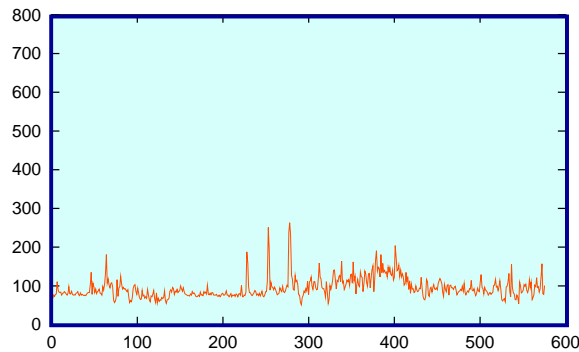
Throughput (frames/sec)



Packet Loss



Audio Latency (ms)

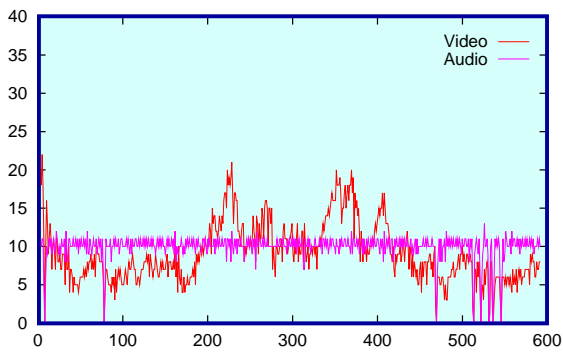


Video Latency (ms)

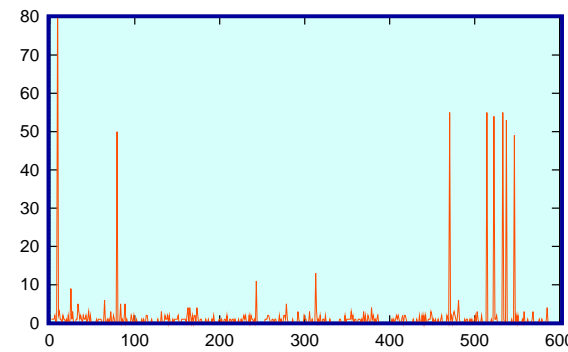
23

Media Scaling Evaluation on the Internet

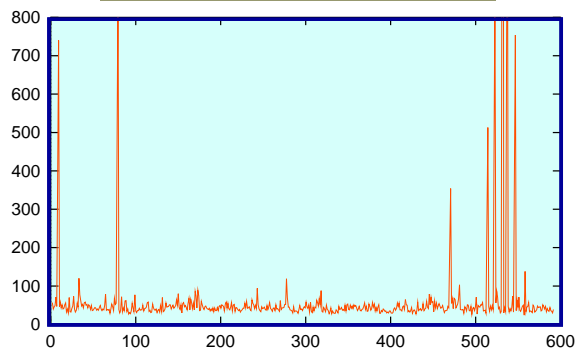
ProShare with video scaling only



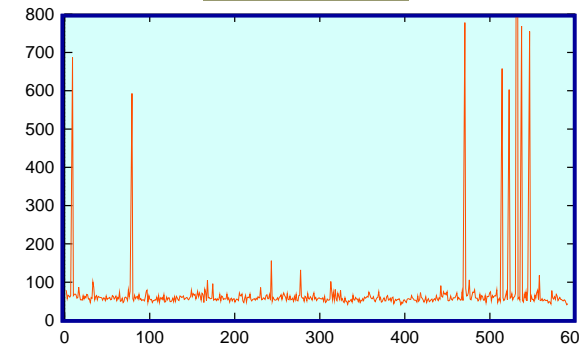
Throughput (frames/sec)



Packet Loss



Audio Latency (ms)

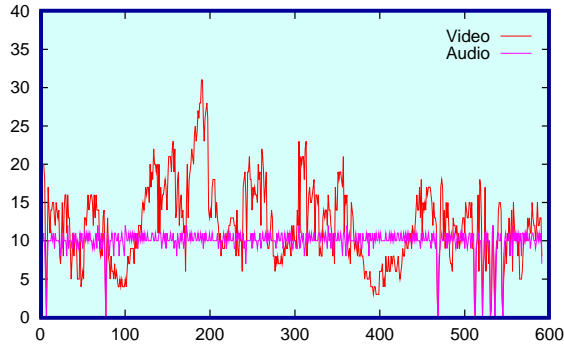


Video Latency (ms)

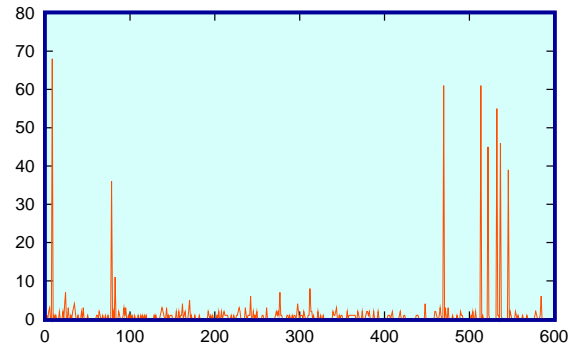
24

Media Scaling Evaluation on the Internet

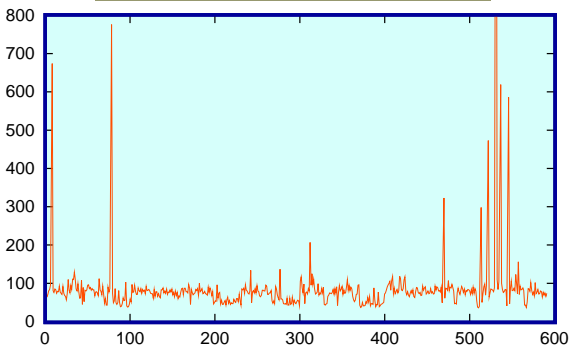
ProShare with 2-dimensional media scaling



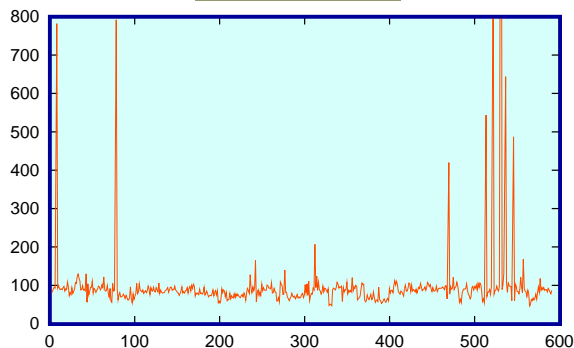
Throughput (frames/sec)



Packet Loss



Audio Latency (ms)



Video Latency (ms)

25

Sustainability Results

Adaptive methods on the Internet

- ◆ Results of an Internet performance study from UNC to UVA
 - » Repeated trials from 10 am to 7 PM weekdays
 - » Trials separated by at least two hours
 - » Scattered over three months

Time Slot	Sustainable	Not Sustainable	% Sustainable
10:00-12:00	6	3	67%
12:00-14:00	4	4	50%
14:00-16:00	1	11	8%
16:00-18:00	3	9	25%
18:00-20:00	4	5	44%
Percentage	36%	64%	