

## Multimedia Networking and Quality-of-Service Support

Kai Shen

11/30/2009

CSC 257/457 - Fall 2009

1

## Multimedia and Quality of Service

Multimedia applications:  
network audio and video  
("continuous media")

QoS support  
network provides application  
with level of performance  
needed for application to  
function.

11/30/2009

CSC 257/457 - Fall 2009

2

## Stored Multimedia

- media stored at source
- transmitted to client
- client playout may begin before all data has arrived (*streaming*)
  - timing constraint for still-to-be transmitted data: in time for playout
  - sensitive to delay jitter: variability of packet delays within the same packet stream

11/30/2009

CSC 257/457 - Fall 2009

3

## Real-Time Interactive Multimedia

- applications:
  - IP telephony, video conference
- sensitive to both jitters and end-to-end delay
- end-end delay requirements:
  - audio: < 150msec good, < 400msec OK
  - higher delays noticeable, impair interactivity
  - includes application-level (compression/decompression) and network delays

11/30/2009

CSC 257/457 - Fall 2009

4

## Multimedia Networking

### Classes of MM applications:

- Stored audio and video
- Real-time interactive audio and video

### Fundamental characteristics:

- Typically **delay sensitive**
  - delay jitter
  - end-to-end delay
- But **loss tolerant**: infrequent losses cause minor glitches

### QoS support for multimedia networking:

- No direct support from the Internet (best effort, no delay guarantee)
- Employ application techniques to mitigate problems

11/30/2009

CSC 257/457 - Fall 2009

5

## Outline

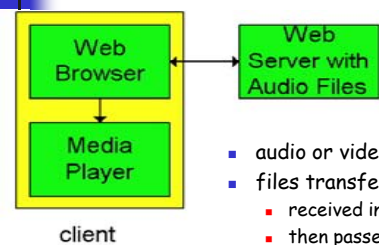
- Multimedia applications
- **Stored multimedia**
- Real-time interactive multimedia
- RTP (Real-time Transport Protocol)

11/30/2009

CSC 257/457 - Fall 2009

6

## Stored Multimedia over the Internet: the basic approach



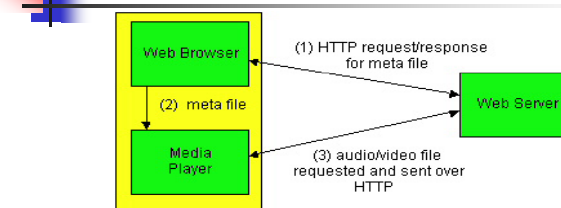
- audio or video stored in files
- files transferred as HTTP object
  - received in entirety at client
  - then passed to player
- media player
  - decompression depending on media format
  - user interface allow interactive control
- long delays until payout!

11/30/2009

CSC 257/457 - Fall 2009

7

## Stored Multimedia over the Internet: streaming approach

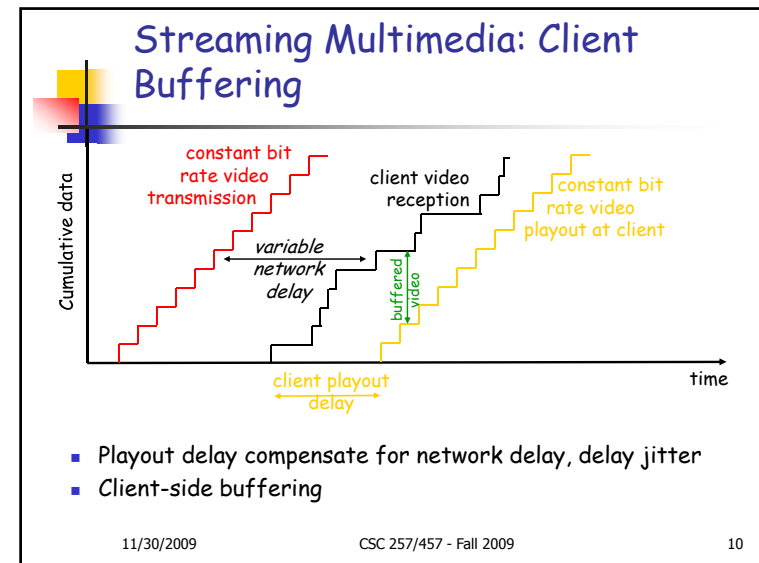
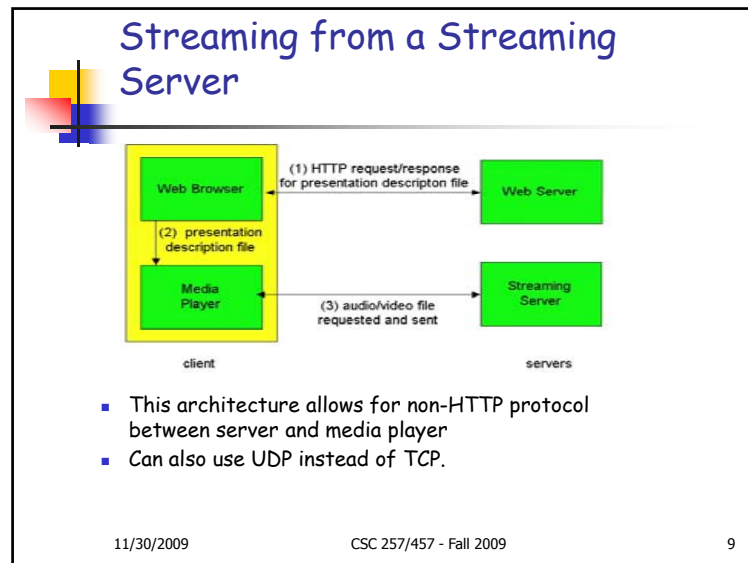


- browser GETs **metafile**
  - location of the media file, type of encoding, ...
- browser launches the media player, passing metafile
- the media player contacts server and the server **streams** audio/video to player
  - client payout may begin before all data has arrived

11/30/2009

CSC 257/457 - Fall 2009

8



### Streaming Multimedia: TCP or UDP?

**TCP**

- send at maximum possible rate under TCP
  - rate fluctuates due to TCP congestion control
  - hard to control

**UDP**

- server sends at rate appropriate for client (oblivious to network congestion!)
- retransmission of lost packets (at the application-layer) when timing constraints permit
- network unfriendly, unfair to other TCP streams employing congestion control

11/30/2009 CSC 257/457 - Fall 2009 11

### Outline

- Multimedia applications
- Stored multimedia
- **Real-time interactive multimedia**
- RTP (Real-time Transport Protocol)

11/30/2009 CSC 257/457 - Fall 2009 12

## Real-time Interactive Multimedia: Internet Phone

- speaker's audio:
  - alternating talk spurts, silent periods (no data)
- during talk spurts:
  - every 20 msec, a 160-byte chunk is generated (64 kbps);
  - application-layer header added to each chunk;
  - chunk+header encapsulated into UDP segment and sent out.

11/30/2009

CSC 257/457 - Fall 2009

13

## Real-time Transport Protocol (RTP)

- Real-time transport protocol (RTP)
  - specifies a packet structure for carrying audio/video data
  - Interoperability: If two Internet phone applications run RTP, then they can understand each other's data

...	Payload type	Sequence number	Timestamp	...
-----	--------------	-----------------	-----------	-----

- Payload type: indicates type of encoding.
  - Sequence#: helps detect loss and restore packet sequence.
  - Timestamp: reflects the data sampling/playout time.
- Where RTP sits in the layered network architecture?

11/30/2009

CSC 257/457 - Fall 2009

14

## Internet Phone: Playout Delay

- Receiver attempts to playout each chunk exactly  $q$  msec after chunk was generated.
  - chunk has time stamp  $t$ : play out chunk at  $t+q$
  - chunk arrives after  $t+q$ : data arrives too late for playout  $\Rightarrow$  delay loss
- Tradeoff for  $q$ :
  - large  $q$ : less packet delay loss
  - small  $q$ : better interactive experience
  - typical maximum tolerable delay: 400 ms

11/30/2009

CSC 257/457 - Fall 2009

15

## Internet Phone: Packet Loss

- **network loss:**
  - IP packet lost due to network congestion (router buffer overflow)
- loss tolerance: depending on voice encoding, packet loss rates between 1% and 10% can be tolerated
- beyond that?
  - small tolerable playout delay prohibits retransmission
  - we need loss concealments - recover from packet loss without retransmission

11/30/2009

CSC 257/457 - Fall 2009

16

### Recovery from Packet Loss (1)

**Forward Error Correction (FEC):**

- for every group of  $n$  chunks create a redundant chunk by exclusive OR-ing the  $n$  original chunks
- send out  $n+1$  chunks, increasing the bandwidth need by  $1/n$
- can reconstruct the original  $n$  chunks if there is at most one lost chunk from the  $n+1$  chunks
- playout delay needs to be long enough for the client receiving all  $n+1$  packets

**Impact of a larger  $n$ :**

- less bandwidth waste
- longer playout delay
- higher probability that 2 or more chunks will be lost

11/30/2009 CSC 257/457 - Fall 2009 17

### Recovery from Packet Loss (2)

**piggyback lower quality stream:**

- send lower resolution audio stream as the redundant information
- the receiver can conceal non-consecutive loss
- What is the playout delay?

11/30/2009 CSC 257/457 - Fall 2009 18

### Recovery from Packet Loss (3)

**Interleaving**

- chunks are broken up into smaller units (e.g. four 5 msec units per chunk)
- Packet contains small units from different chunks
- if packet is lost, still have most of every chunk
- has no redundancy overhead, but adds to playout delay

11/30/2009 CSC 257/457 - Fall 2009 19

### Streaming Multimedia on the Internet: Summary

- use UDP to avoid TCP congestion control (delays) for time-sensitive traffic
- compensate delay jitters:
  - Playout delay and client-side buffering
- error recovery (on top of UDP)
  - retransmissions, time permitting (for stored multimedia)
  - conceal errors (for real-time interactive multimedia): FEC, piggybacking, interleaving

11/30/2009 CSC 257/457 - Fall 2009 20

## Improving QoS inside the Internet

**Thus far:** applying application-level end-to-end techniques to enhance the "best effort" service of the Internet.

**Future:** incorporating quality-of-service techniques into Internet routers.

[Goals and Mechanisms](#)

[Proposed architectures](#)

11/30/2009

CSC 257/457 - Fall 2009

21

## QoS Goals and Mechanisms

- Packet classification - packet marking at network entrance
- Packet isolation - policing
- Efficiency and fairness - router scheduling
- Admit only when there is enough resource - reservation

Let's next look at policing and scheduling mechanisms ...

11/30/2009

CSC 257/457 - Fall 2009

22

## Policing Mechanism

**Goal:** limit traffic to not exceed specified traffic pattern (at the entrance of the network).

Characteristics of traffic pattern (in terms of bandwidth usage):

- **Average rate:** how many packets can be sent per unit of time
- **Peak rate:** the maximum or 95 percentile rate
  - what is the interval length: 100 packets per sec or 6000 packets per min?
- **Burst size:** maximum number of burst packets sent beyond the average rate over any time period
  - over any interval of length  $t$ : number of packets  $\leq (r \cdot t + b)$

How to enforce a traffic pattern with an average rate and a burst size?

11/30/2009

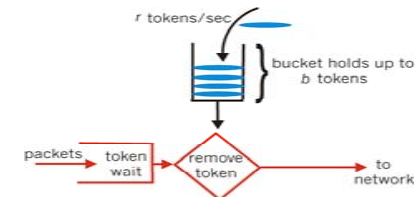
CSC 257/457 - Fall 2009

23

## Policing Mechanism: Leaky Bucket

### Leaky bucket

- bucket can hold  $b$  tokens
- tokens generated at rate  $r$  tokens/sec unless bucket full
- a token must accompany each admitted packet



- over any interval of length  $t$ : number of packets admitted less than or equal to  $(r \cdot t + b)$
- limit input to specified average rate and burst size

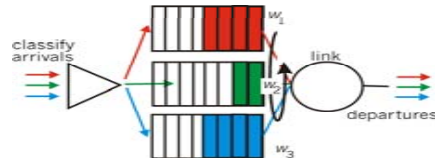
11/30/2009

CSC 257/457 - Fall 2009

24

## Packet Scheduling at Routers

- FIFO (first in first out) scheduling:** send in order of arrival
  - multiple classes
    - Round robin scheduling:** cyclically scan class queues, serving one from each class (if available)
    - Priority scheduling:** transmit packet in the highest priority queue
    - Weighted fair queueing:** each class gets weighted amount of service in each round robin cycle



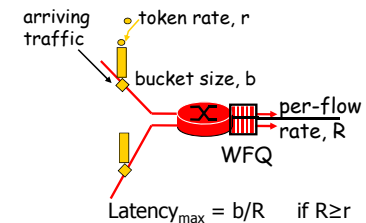
11/30/2009

CSC 257/457 - Fall 2009

25

## Combined Policing and Scheduling

Leaky bucket, WFQ combined to provide guaranteed upper bound on latency!



11/30/2009

CSC 257/457 - Fall 2009

26

## How should the Internet evolve to better support QoS for multimedia apps?

### No major changes:

- apply only application-level end-to-end techniques

### Integrated services (Intserv) philosophy:

- fundamental changes in Internet so that apps can reserve end-to-end bandwidth
- requires new, complex software in hosts & routers

### Differentiated services philosophy:

- Fewer changes to Internet infrastructure, yet provide differentiated services to different classes
- may not give firm guarantee on delay and bandwidth

11/30/2009

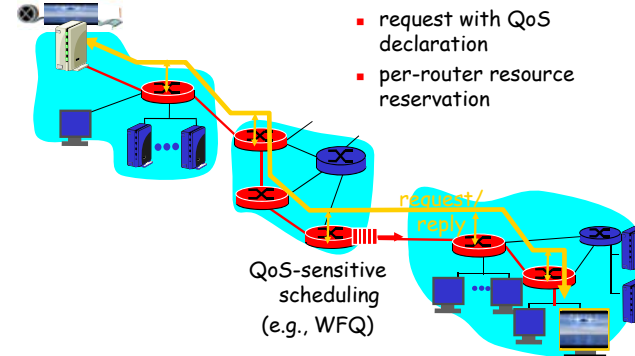
CSC 257/457 - Fall 2009

27

## Intserv: An Illustration

### Resource reservation

- request with QoS declaration
- per-router resource reservation



11/30/2009

CSC 257/457 - Fall 2009

28



## Differentiated Services

### Concerns with Intserv:

- **Scalability & overhead:** signaling, maintaining per-flow router state difficult with large number of flows

### Diffserv approach:

- simple functions in network core, relatively complex functions at edge routers (or hosts)
- provide differentiated services to different classes, may not give firm guarantee on delay and bandwidth

11/30/2009

CSC 257/457 - Fall 2009

29



## Multimedia Networking and Quality of Service Support: Summary

- multimedia applications and their QoS requirements
- applying application-level end-to-end techniques to enhance the "best effort" service of the Internet
- QoS support within the Internet: policing and scheduling, reservation mechanisms
- next generation Internet QoS architecture: Intserv, Diffserv

11/30/2009

CSC 257/457 - Fall 2009

30



## Disclaimer

- Parts of the lecture slides contain original work of James Kurose, Larry Peterson, and Keith Ross. The slides are intended for the sole purpose of instruction of computer networks at the University of Rochester. All copyrighted materials belong to their original owner(s).

11/30/2009

CSC 257/457 - Fall 2009

31