

Multimedia Networking and Quality-of-Service Support

Kai Shen

11/24/2014 CSC 257/457 - Fall 2014 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.

11/24/2014 CSC 257/457 - Fall 2014 2

Stored Multimedia

- Media stored at source (possibly produced a lone time ago)
- Transmitted to client when requested

11/24/2014 CSC 257/457 - Fall 2014 3

Stored Multimedia over the Internet: The basic approach

```

    graph TD
        subgraph Client
            WB[Web Browser]
            MP[Media Player]
            WB --> MP
        end
        WS[Web Server with Audio Files]
        WS <--> WB
    
```

- Audio or video stored in files
- Files transferred as HTTP objects
 - received in entirety at client
 - then passed to player
- Media player
 - decompression depending on media format
 - user interface allow interactive control
- Long delays until playback!

11/24/2014 CSC 257/457 - Fall 2014 4

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 playback may begin before all data has arrived

11/24/2014 CSC 257/457 - Fall 2014 5

Streaming from a Streaming Server

- This architecture allows for special-purpose (non-HTTP) protocol between server and media player

11/24/2014 CSC 257/457 - Fall 2014 6

Streaming Multimedia Performance

- Timing constraint for still-to-be transmitted data: in time for playback
- Sensitive to delay jitter: variability of packet delays within the stream
- Solution:** playback delay and client-side buffering

11/24/2014 CSC 257/457 - Fall 2014 7

Streaming Multimedia: TCP or UDP?

TCP

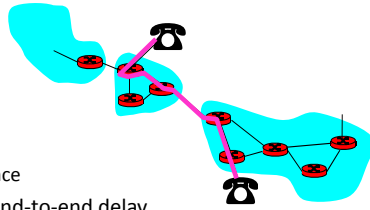
- send at maximum possible rate under TCP
 - rate fluctuates due to TCP congestion control
 - hard to control

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
- incompatible with HTTP ⇔ need non-HTTP streaming client/server

11/24/2014 CSC 257/457 - Fall 2014 8

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

11/24/2014

CSC 257/457 - Fall 2014

9

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 msec

11/24/2014

CSC 257/457 - Fall 2014

10

Internet Phone: Packet Loss

- **Network loss:**
 - packet dropped by router (due to network congestion)
 - delay loss
- Loss tolerance: depending on voice encoding, packet loss rates below 10% can be tolerated
- Beyond that?
 - small tolerable playout delay prohibits retransmission
 - we need loss concealments – recover from packet loss without retransmission

11/24/2014

CSC 257/457 - Fall 2014

11

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

Impact of a larger n :

- less bandwidth waste
- longer playout delay (playout delay needs to be long enough for the client receiving all $n+1$ packets)
- higher probability that 2 or more chunks will be lost

11/24/2014

CSC 257/457 - Fall 2014

12

Recovery from Packet Loss (2)

Original Stream

Redundancy

Packet Loss

Reconstructed Stream

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?

11/24/2014 CSC 257/457 - Fall 2014 13

Recovery from Packet Loss (3)

Original Stream

Interleaved Stream

Packet Loss

Reconstructed Stream

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

11/24/2014 CSC 257/457 - Fall 2014 14

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.

- RTP normally runs on UDP, but can run on TCP.

11/24/2014 CSC 257/457 - Fall 2014 15

Multimedia Networking Summary

Fundamental characteristics:

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

Application techniques for multimedia networking:

- Use UDP 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

11/24/2014 CSC 257/457 - Fall 2014 16

Improving QoS inside the Internet

Thus far (practically doable now): applying application-level end-to-end techniques to enhance the “best effort” service of the Internet.

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

[Specific goals and mechanisms](#)

[Overall Internet architectures](#)

11/24/2014

CSC 257/457 - Fall 2014

17

QoS Goals and Mechanisms

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

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

11/24/2014

CSC 257/457 - Fall 2014

18

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
- **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 $\leq (r t + b)$
 - For a conforming traffic, if available bandwidth $R \geq r$, then the queue never exceeds b packets, therefore each packet's delay $\leq b/R$.

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

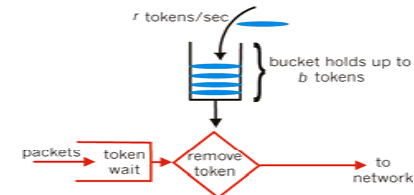
11/24/2014

CSC 257/457 - Fall 2014

19

Policing Mechanism: 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 $(r t + b)$
- Limit input to specified average rate and burst size

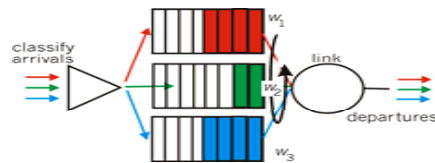
11/24/2014

CSC 257/457 - Fall 2014

20

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 to reach a certain target



11/24/2014

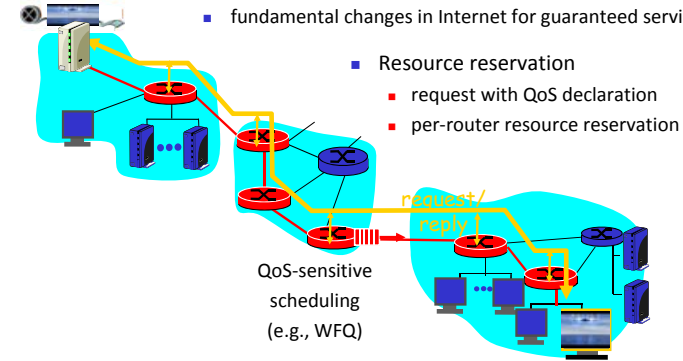
CSC 257/457 - Fall 2014

21

Internet QoS Architectures

Integrated services (Intserv) approach:

- fundamental changes in Internet for guaranteed services



11/24/2014

CSC 257/457 - Fall 2014

22

Differentiated Services

Concerns with Intserv:

- **scalability & overhead:** signaling, maintaining per-flow router state difficult with large number of flows
- **deployment challenge:** requires new complexity in hosts & routers

Differentiated services (Diffserv) approach:

- provide differentiated services to different classes, but may not give firm guarantee on delay and bandwidth
- lower overhead and better scalability
- simpler functions in network routers

11/24/2014

CSC 257/457 - Fall 2014

23

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
- Proposed Internet QoS architectures: Intserv, Diffserv

11/24/2014

CSC 257/457 - Fall 2014

24



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).