II.1 End-to-End QoS for Video Delivery

Prof. Chia-Wen Lin
Department of CS
National Chung Cheng University
886-5-272-0411 ext. 33120
cwlin@cs.ccu.edu.tw

Outline

- Introduction
- Network-Centric QoS Support
- End-to-End QoS Control for Video streaming
  - Congestion control
  - Error control
  - Video multicasting
- Video Steaming Properties for Network Use
- Conclusions

QoS Problems in Current Internet Infrastructure

- No QoS Guaranteed for current Network
  - No bandwidth reservation;
  - No delay guarantee;
  - No packet loss guarantee
- Heterogeneity: (multicast)
  - network: different users, different packet loss / delay
  - receiver: different latencies / visual quality requirements / processing powers / display formats

Technical Challenges for QoS Support

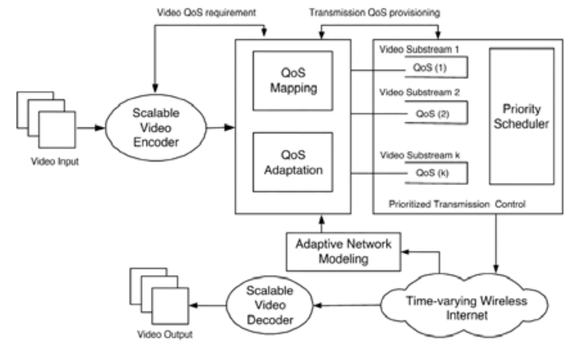
- To support end-to-end QoS for video delivery over wireless Internet, there are several fundamental challenges:
  - QoS support encompasses a wide range of technological aspects
  - Different applications have very diverse QoS requirements in terms of data rates, delay bounds, and packet-loss rates
  - Different types of networks inherently have different characteristics (network heterogeneity)
  - Dramatic heterogeneity among end users

QoS Control for Internet Video streaming

- Network Centric
  --- next generation network providing QoS support
    - Link layer: the probability of buffer overflow or delay violation
    - Application layer: MSE, PSNR
- End System-based
  --- compatible with current network structure
    - Congestion control
    - Error control
    - Power control

Components for End-to-End QoS Support

Network-Centric QoS Control: General Framework



Network QoS Provisioning

- IETF QoS Provisioning Approaches
  - IntServ: per-flow-based
    - Impractical for lack of scalability and difficulty in resource reservation
  - DiffServ: per-aggregate-based
    - Provides a scalable and manageable network with service differentiation capability
- Service Differentiation
  - QoS delay & packet loss
- QoS Control Mechanisms
  - Packet scheduling, queue management algorithms, etc.
- Theories
  - Network calculus, effective bandwidth, etc.

Network QoS Provisioning for Wireless Networks

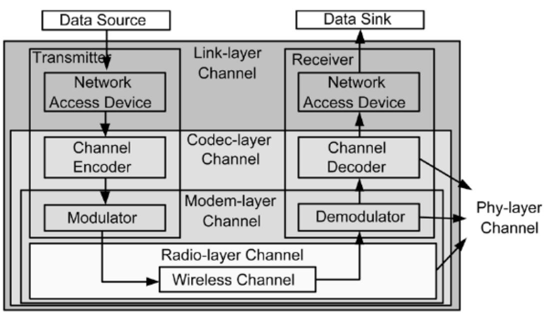
- 3GPP has defined four different UMTS QoS classes according to delay sensitivity
  - Conversational, streaming, interactive, and background classes
- IEEE 802.11e enhanced communication modes for providing QoS support
  - Distributed Coordination Function (DCF)
    - ➔ Enhanced Distribution Coordination Function (EDCF)
  - Point Coordination Function (PCF)
    - ➔ Hybrid Coordination Function (HCF)
- Wireless Multimedia Enhancements (WME) has also proposed to provide an interim QoS solution for 802.11

Cross-Layer QoS Support for Video Delivery over Wireless Internet

1. Wireless Network Modeling
   - Physical layer
     - Radio-layer channel: large-scale loss and small-scale fading
     - Modem-layer channel: modeled by a finite-state Markov chain whose states being characterized by different BERs
     - Codec-layer channel: modeled by a finite-state Markov chain whose states being characterized by different data-rates, or a symbol being error-free/in-error, or a channel being good/bad
   - Link layer
     - Effective Capacity (EC) model: capture the effect of channel fading for the link queuing behavior

Cross-Layer QoS Support for Video Delivery over Wireless Internet

Different Channel Models



Cross-Layer QoS Support for Video Delivery over Wireless Internet

2. Prioritized Transmission Control
   - Requires a class-based buffering and scheduling mechanism
     - Each QoS priority class can obtain a certain level of statistical QoS guarantees in terms of probability of **packet loss** and **packet delay**
     - Translate the statistical QoS guarantees of multiple priority classes into **rate constraints** based on the effective capacity
     - The rate constraints can be derived according to the guaranteed packet loss probabilities and different buffer sizes of each priority class

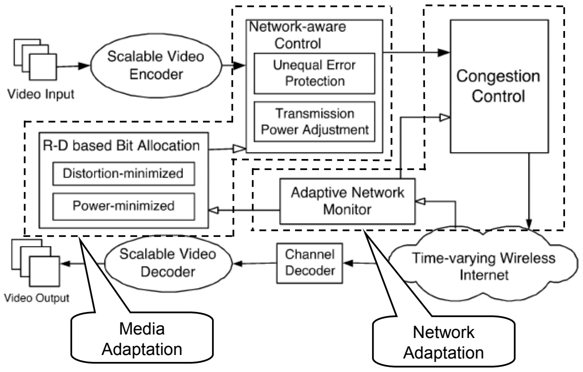
Cross-Layer QoS Support for Video Delivery over Wireless Internet

3. QoS Mapping and QoS Adaptation
   - Application-specific
   - The common approach is to partition multimedia data into smaller units and then map these units to different classes for prioritized transmission.
   - The partitioned multimedia units are prioritized based on its contribution to the expected quality at the end user.

Adaptation for End-System Centric QoS Control

- Network adaptation
  - Design an adaptive media transport protocol to determine the network resources (e.g., bandwidth and battery power) for video delivery
- Media adaptation
  - control the bit rate of the video stream based on the estimated available bandwidth
  - adjust error and power control behaviors according to the varying wireless Internet conditions

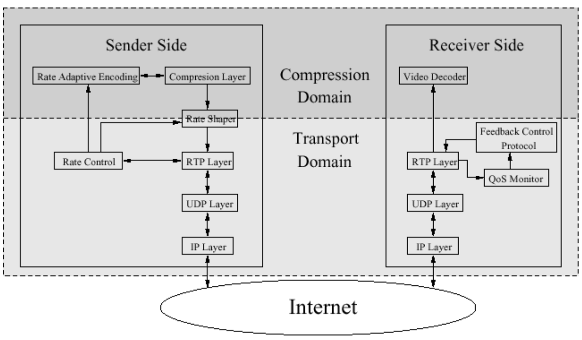
General Framework for End-to-End QoS Provisioning for Wireless Video



QoS Control for Internet Video streaming

- Network Adaptive Congestion Control
  - To reduce packet loss and delay
  - Rate control, rate adaptive encoding and rate shaping
- Adaptive Error Control
  - To handle video quality when packet loss happens
  - FEC, retransmission, error resilience and error concealment

Congestion Control (1/8)



Congestion Control (2/8)

- Rate control:
  - UDP replaces TCP for delay reason
  - no congestion control for QoS in UDP
  - rate-based control is usually employed (source based, receiver based and hybrid)

Congestion Control (3/8)

- Rate control – rate-based control
  - – source based: sender regulates video stream applied to unicast & multicast
  - – receiver based: each receiver regulates the receiving rate; typically for multicast
  - – hybrid

Congestion Control (4/8)

- TCP-friendly flow control – source-based
  - – Probe-based
    - AIMD (Additive Increase Multiplicative Decrease)
    - MIMD (Multiplicative Increase Multiplicative Decrease)
  - – Model-based

$$\lambda = \frac{1.22 \times MTU}{RTT \times \sqrt{p}}$$

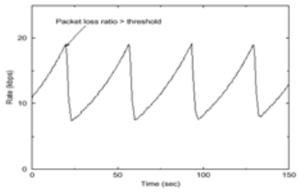


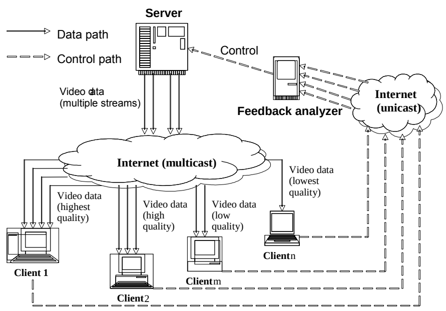
Fig. 5. Source rate behavior under the AIMD rate control.

Congestion Control (5/8)

- Rate control – receiver-based
  - – for solving the heterogeneity in multicast
  - – probe-based approach
  - – model-based approach
  - – joint-leaving for large number of receiver
    -- congestion
  - – shared learning or synchronization control

Congestion Control (6/8)

Receiver-Driven Multicasting



Congestion Control (7/8)

- <u>Rate-adaptive Video Encoding</u>

  to maximize the perceptual quality under a given rate.
- Scalable Rate Control in MPEG-4
  - – second-order R-D model for target bit allocation
  - – sliding-window to smooth the scene change effect
  - – adaptive data points selection for model updating
  - – adaptive threshold shape control
  - – dynamically bit-rate allocation among VOs

Congestion Control (8/8)

- <u>Rate Shaping</u>
  - adapt the video rate to target network rate constraint
  - Server selective frame discard
  - Selective DCT coefficient discard

Congestion Control (8/8)

- End-to-end packet loss differentiation and estimation
  - Split connection method: places an agent at the edge of wired and wireless networks
  - End-to-end method: uses inter-arrival time or packet pair
- Available bandwidth estimation
  - RTT, packet loss ratio
  - Receiver Based Packet Pair (RBPP)

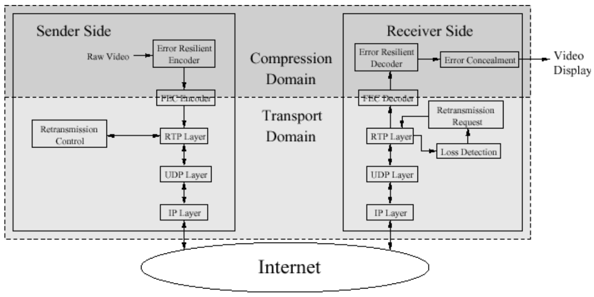
Error Control (1/4)

- To prevent packet loss by matching the rate of video streams to the available bandwidth in the network.
- packet loss is unavoidable
- other mechanisms to maximize the video presentation quality

Error Control (2/4)

- FEC
- Retransmission
- Error resilience coding
- Error concealment

Error Control (3/4)



Error Control (4/4)

- <u>FEC</u> – channel coding
- Unequal Error Protection and Equal Error Protection
  - increase transmission rate
  - increase delay: long block or interleaving
  - Not adaptive to varying loss characteristic

QoS for Video Multicasting

- FEC (Forward Error Correction)
  - Not suitable for bursty error network condition
- ARQ (Automatic Repeat reQuest)
  - May cause feedback implosion
- Pseudo-ARQ
  - Solves feedback implosion
  - To the server, it looks like ordinary multicast
  - To the receiver, it looks like ordinary ARQ

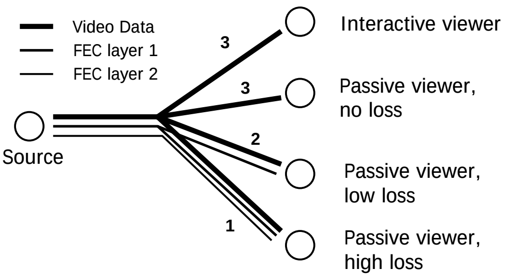
Layered FEC for Video Multicasting (1/2)

- More source layer, higher video quality
- More FEC layer, higher protection level

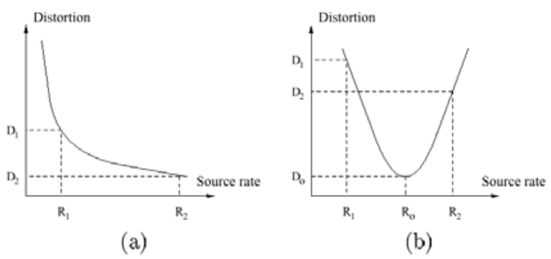
FEC Layer 2			FEC 0-2	FEC 1-2
FEC Layer 1		FEC 0-1	FEC 1-1	FEC 2-1
Video Layer 3	GOP 0-3	GOP 1-3	GOP 2-3	GOP 3-3
Video Layer 2	GOP 0-2	GOP 1-2	GOP 2-2	GOP 3-2
Video Layer 1	GOP 0-1	GOP 1-1	GOP 2-1	GOP 3-1
	Time = 0	Time = Δ	Time = 2Δ	Time = 3Δ

from "Multicast Transmission of Scalable Video using Receiver-driven Hierarchical FEC"

Layered FEC for Video Multicasting (2/2)



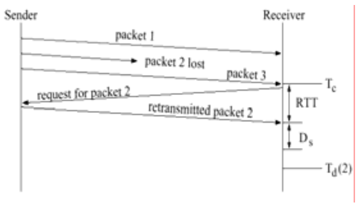
Joint Source-Channel Coding



Retransmission: Unicast

- <u>Delay-constrained Retransmission</u>

  unicast:



When the receiver detects the loss of packet N:
if $(T_c + RTT + D_s < T_d(N))$
send the request for retransmission of packet N to the sender;

Retransmission: Multicasting (1/2)

- Delay-constrained Retransmission

  multicast
- Restricted within closely located multicast members -- local recovery ;
- Feedback implosion;
- Receiver buffer to absorb delay jitter and to receive re-transmitted packet.

Retransmission: Multicasting (2/2)

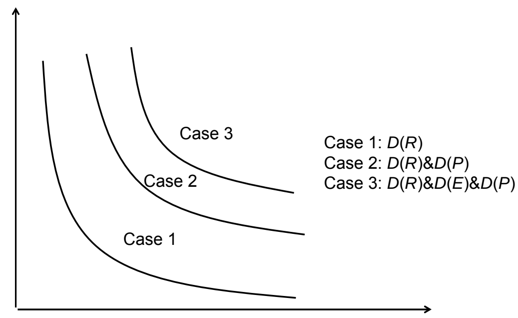
Pseudo ARQ for Video Multicasting

S_i	S_{i+1}	S_{i+2}	S_{i+3}	S_{i+4}	S_{i+5}	°°°°°°	Source layer
S_{i-1}	S_i	S_{i+1}	S_{i+2}	S_{i+3}	S_{i+4}	°°°°°°	1st delayed source layer
S_{i-2}	S_{i-1}	S_i	S_{i+1}	S_{i+2}	S_{i+3}	°°°°°°	2nd delayed source layer

Joint Power Control and Error Control

- Multipath fading and multiple access interference (MAI) in wireless networks necessitate the use of high transmission power
- More sophisticated coding scheme and powerful channel coding can be applied to decrease transmission power while maintaining a desired video quality
- Three cases in joint power control and error control for video communication
  - Case 1: $D = D(R)$
  - Case 2: $D = D(R)$ & $D(P)$
  - Case 3: $D = D(R)$ & $D(E)$ & $D(P)$

Joint Power Control and Error Control



Rate-Distortion-Based Bit-Allocation

- The resource allocation problem can be formulated as follows:

$$\text{Min} \quad D_T(D_s, D_c) \quad \text{s.t.} \quad R_T \le R_0$$

D_T: the expected end-to-end distortion

D_s: the source distortion

D_c: the channel distortion

R_T: the total bandwidth

R_0: the total bandwidth budget

$$\text{Min} \quad P_T(P_s, P_c, P_t) \quad \text{s.t.} \quad R_T \le R_0 \quad \text{and} \quad D_T \le D_0$$

Video Compression for Internet Video Streaming

- Scalable and non-scalable coding
- Requirements upon streaming video codec:
  - Bandwidth
  - Delay
  - Loss
  - VCR like functionality
  - Decoding complexity

Video Steaming Properties for Network Use

- Natural breakpoints for packetization
- Adjustable packet sizes
- No bit level shifts during packetization
- Well defined high-priority information
- Flexible rate control
- Ease of transcoding
- Layered coding
- Resilience to error propagation