ELEC S347F Multimedia Technologies

Networked Multimedia

Classification

- Based on Delivery
  - Real-time delivery
  - Soft-real-time delivery
- Based on Application
  - Streaming Stored Multimedia
  - Streaming Live Multimedia
  - Conversational Multimedia

Networked Multimedia

- Video Distribution
  - Streaming stored audio/video
    - Support random access, pause/resume, thumbnails
  - Streaming live audio/video
  - E.g. Spotify, Netflix, Disney+, YouTube, Facebook Live
- Video Conversation
  - Voice over IP, Video Telephony/Conferencing
  - Everyone is a sender and a receiver
  - Support call invitation, acceptation, rejection, etc
  - E.g. Skype, WeChat

Real-time Delivery

- Media data must be delivered from source and presented at destination within a given delay budget
- Applications involve interactions between users
- Real-time delivery requirement may conflict with requirements for data integrity and presentation timing integrity
- e.g. Internet phone:
  - Long delays will lead to talking collision
  - Both sides trying to speak at the same time
  - May allow data loss and playback jitter (in order to meet the given delay budget)

Soft-real-time Delivery

- Unlike real-time delivery, no delay budget given
- However, must preserve data integrity and presentation timing integrity
- Reduce delay as much as possible
- e.g. Video-on-Demand (VoD)
  - Take longer start-up delays (buffering) as long as smooth playback is maintained after playback has started

Streaming Stored Multimedia

- Streaming == download?
  - What if we have an ultra high speed network?
- Download
  - Start decoding and playback soon after obtaining a full copy of the multimedia data
- Streaming
  - Start decoding and playback soon after receiving the initial portion of the multimedia data while at the same time receiving the later portions

Streaming vs. Download

- Advantages of Streaming
  - Shorter playback delay
  - Require less buffer memory for receivers (not necessary to hold a full copy, but a portion only)
  - Allow more concurrent clients (since the download rate is as large as the playback rate)
  - Able to support live content
- Disadvantages of Streaming
  - Traffic expensive for random access, replay
  - Require support of partial decoding (special designed codec)
    - e.g. streaming audio (e.g. MP3), video (e.g. AVC), rich text (e.g. PDF), etc

Characteristics of Multimedia Traffic

- Bidirectional symmetry
  - Video-on-Demand: asymmetric
  - Video conference: symmetric
- Throughput variation with time
  - Constant bit rate traffic vs. variable bit rate traffic
  - At least as large at the data rate of the multimedia data
  - Internet: transmission rate is usually non-constant and not guaranteed
  - Solution: reserve bandwidth for streaming (how?)

Characteristics of Multimedia Traffic

- Delay Sensitive
  - Stored multimedia: User can tolerant an initial playback delay
  - Live multimedia: at most as large as the perceptual limit
    - For voice, delays smaller than 150 ms are not perceived by a human listener
    - Delays between 150 ms and 400 ms are acceptable
    - Delays exceeding 400 ms can result in frustrating voice conversation
- Loss Tolerant
  - Occasional loss only causes occasional glitches in playback
  - These loss can often be partially concealed

Streaming Multimedia

- Property of the Internet
  - Best effort (to move each packet from source to destination as quick as possible)
  - Good for file transfer and web browsing
  - However, multimedia applications typically have quality of service (QoS) requirement in terms of throughput, delay, error rate, etc

Application	Throughput (bps)	Delay Bound (ms)	Error Rate
Voice	9.6 k – 64 k	< 150	$<10^{-2}$
Video Conference	128 k – 6 M	< 150	$<10^{-3}$
Video (non-real-time)	1 M - 10 M	Large	$<10^{-6}$
File Transfer	1 M – 100 M	Large	$<10^{-8}$
Web Browsing	1 M – 10 M	Large	$<10^{-8}$

Quality of Service (QoS)

- There are 4 Principles for a network to support quality of service (QoS)
  - Packet Classification
    - Marking of packet is needed for router to distinguish between different classes
  - Isolation
    - Provide protection (i.e. isolation) for one class from other classes so that one flow is not adversely affected by another misbehaving flow
  - High Utilization
    - While providing isolation, it is desirable to use resources as efficiently as possible
  - Call Admission
    - Need a call admission process
    - The network may block call if it cannot satisfy the needs

Networked Multimedia

- New Requirements and Design Issues
  - Bandwidth
    - Throughput
    - Scalability
  - Latency
    - End-to-end Delay
    - Delay Jitter
  - Reliability
    - Packet Loss Rate
    - Bit Error Rate
    - Out of Order Reception

Bandwidth Issue

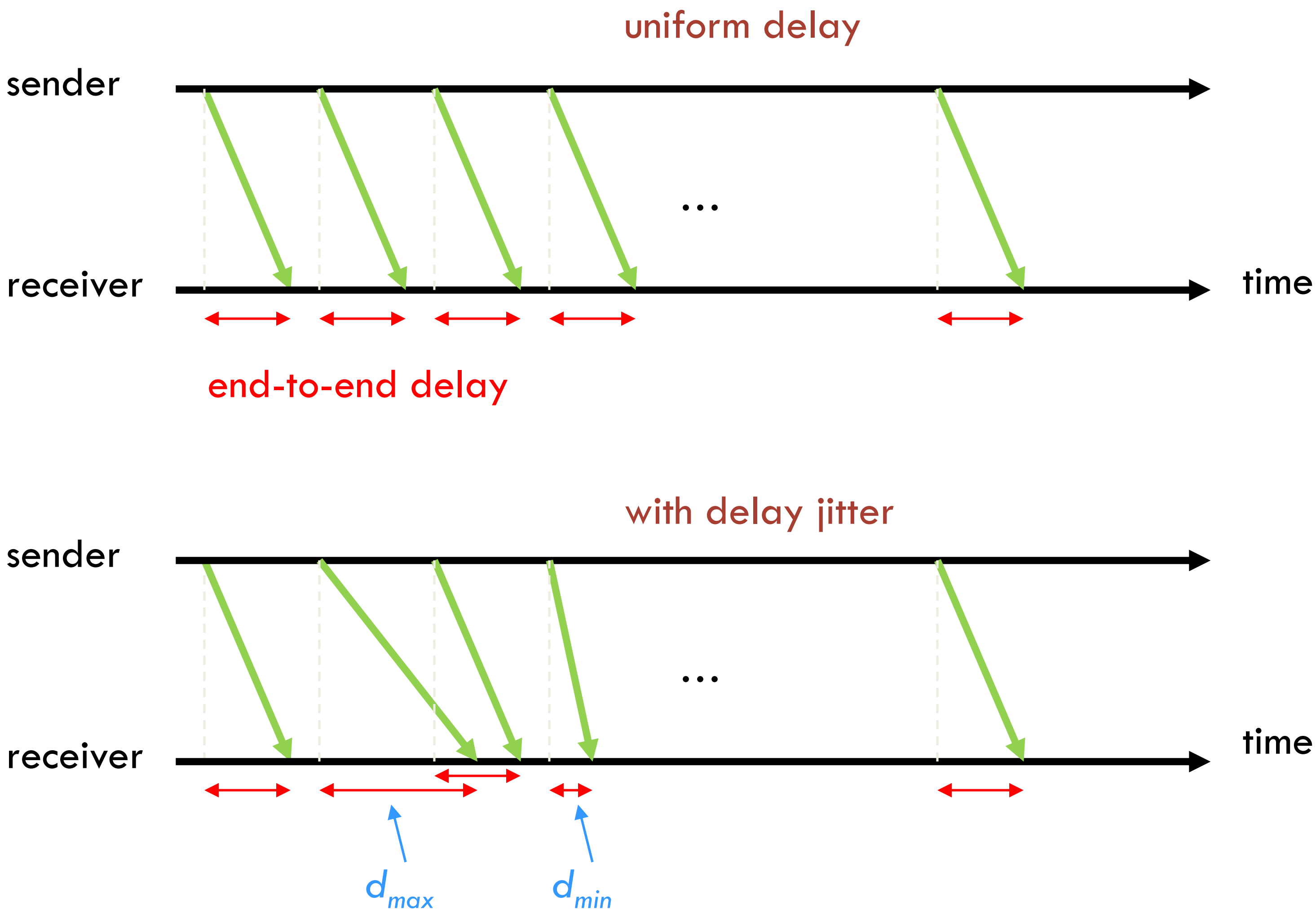
- Throughput
  - The throughput of the network must be at least as large as the data rate of the multimedia
  - Or to encode the data to the rate lower than the network throughput
- Scalability
  - Provide multiple quality level
    - For users with various device resolution
    - For users with various bandwidth (varying in time or with users)
  - Require special support from coding
    - E.g. MPEG-2 Layered Video, MPEG-4 Scalable Lossless Coding (SLS)

Latency Issue

Delay Jitter

- End-to-end Delay
  - Transmission delay from sender to receiver
- Delay Jitter
  - Different packets may experience different end-to-end delays
  - The variation in end-to-end delays is called delay jitter
- If the delay, d, is bounded, $delay\ jitter$ is also bounded
  - e.g., if $d_{min} \leq d \leq d_{max}$,
  - then $delay\ jitter \leq d_{max} - d_{min}$

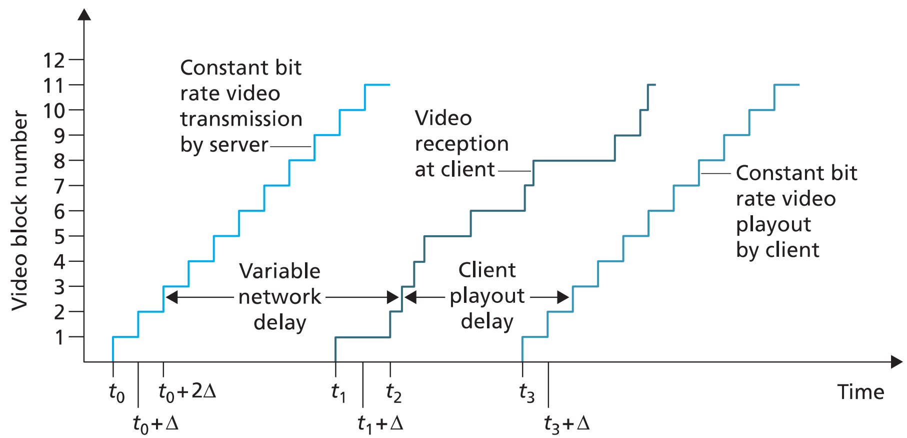
Delay Jitter



Delay Jitter

- End-to-end delay is not the critical issue
  - Users can tolerate a small initial delay before starting playback
  - e.g. for streaming video, $d_{max} = 10$ s may be acceptable
- However, the end-to-end delays are varying
  - Users may experience frame freezing (when users wait for the delayed frames) or frame skipping (when users skip over delayed frames)
  - A delay jitter of 5 s may require a buffer requirement of several MBs, which may not be acceptable
- Buffering
  - When the video frames arrived at the client, the playback is not started immediately, but instead build up a reserve of video in a buffer
  - To mitigate the effects of delay jitter

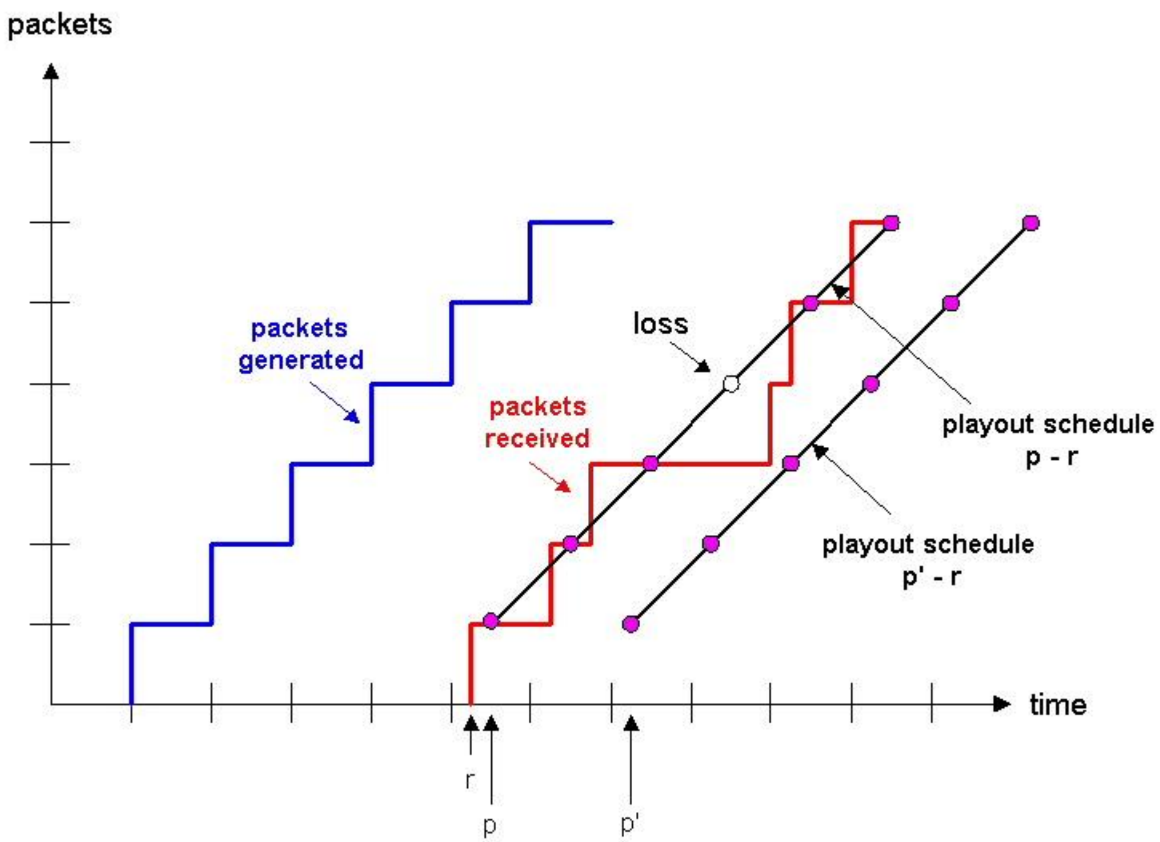
Buffering



- Suppose video is encoded in a fixed data rate
- Each video block contains video frames that are to be played out over some fixed amount of time Δ
- The server transmits the first video block at t_0, the second block at $t_0 + \Delta$, the third block at $t_0 + 2\Delta$, and so on

- The client receives the first video block at t_1, and expects to receive the second block at $t_1 + \Delta$
- Because of the variance of end-to-end network delays, the second block arrives at t_2 but not $t_1 + \Delta$
- If the client begins the playback as soon as the arrival of the first video block (t_1), the video playback would either have to stall (wait for the second block to arrive) or to skip the second block
- Both resulting in undesirable playback impairments
- Instead, if the client delay the start of playback until t_3, playback can proceed smoothly
- What is the appropriate value of t_3?
  - If t_3 is too small, playback would be unsmooth
  - If t_3 is too large, playback would be smooth, but result in long initial playback delay (and require a large memory buffer)

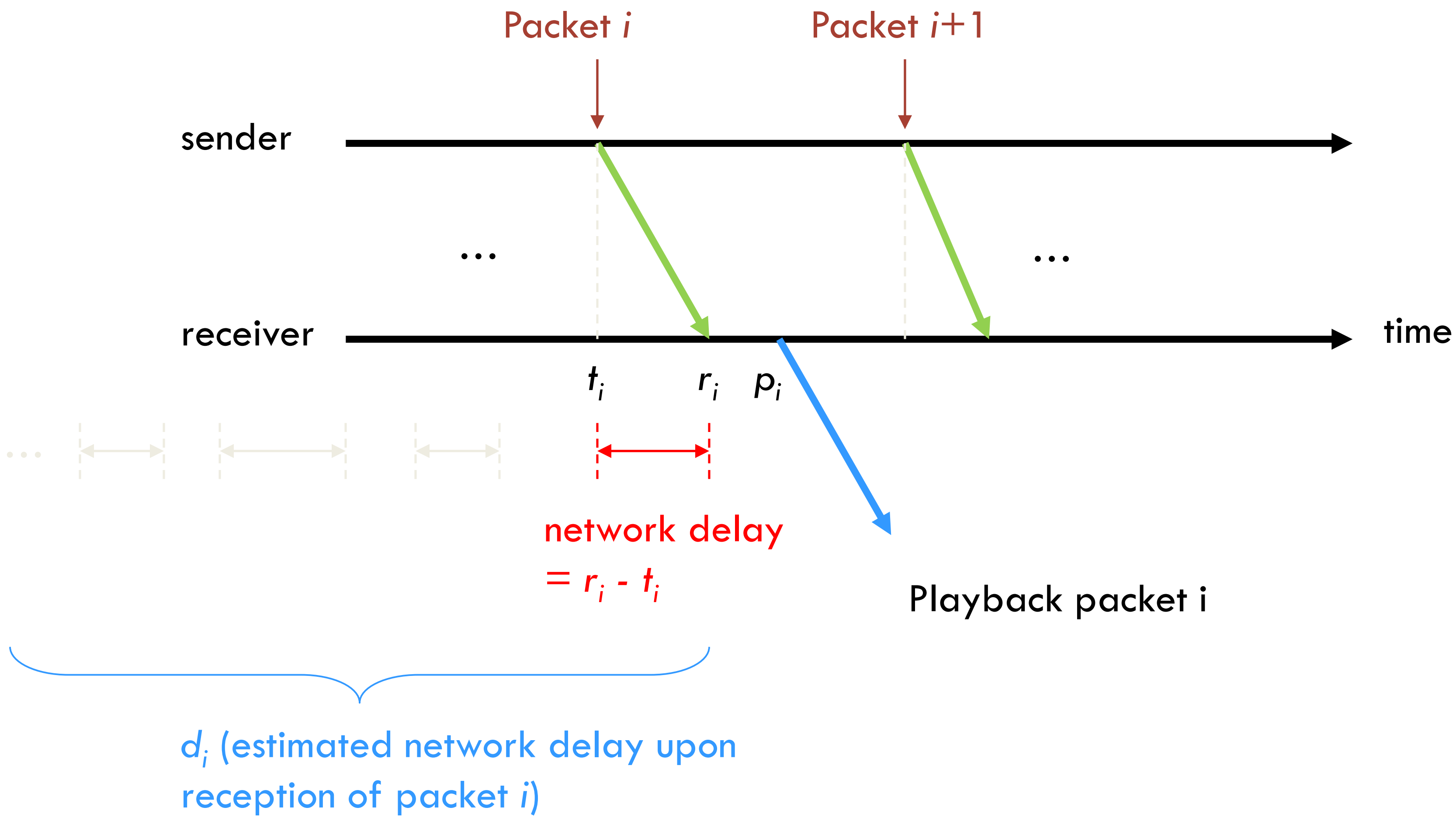
Fixed Playout Delay



Adaptive Playback Delay

- Estimate network delay and delay variation in real time
- Adjust the playback delay accordingly
  - t_i = timestamp of packet i
  - r_i = the time packet i is received
  - p_i = the time packet i is played at receiver
  - Then $r_i - t_i$ is the network delay of packet i
  - Let d_i be the estimated network delay upon reception of packet i
  - Let v_i be the estimated delay variation upon reception of packet i

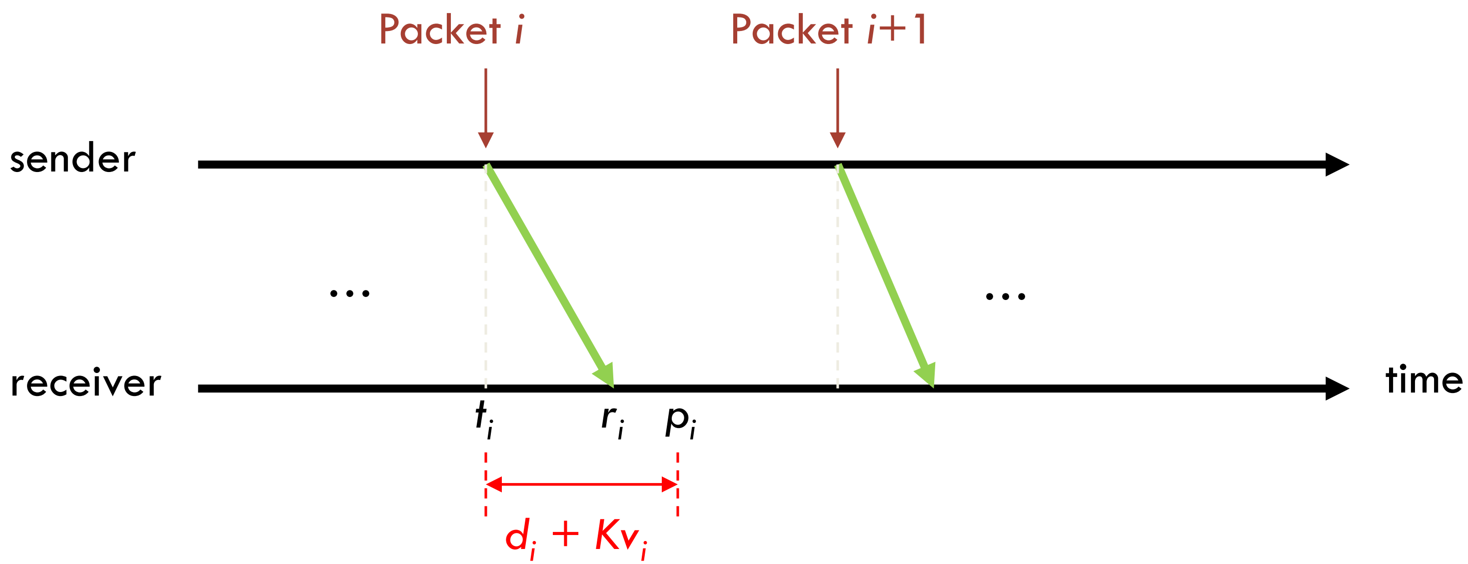
How to Estimate?



Adaptive Playback Delay

- Exponential Averaging
  - $d_i = (1-u)d_{i-1} + u(r_i - t_i)$
  - $v_i = (1-u)v_{i-1} + u|r_i - t_i - d_i|$
  - where u is a fixed constant between 0 and 1 inclusive (e.g. 0.01)
- The playback time of packet i is computed as
  - $p_i = t_i + d_i + Kv_i$
  - where K is a positive constant (e.g. 4)

Adaptive Playback Delay



Reliability Issue

Error Models

- Bit Error
  - One or more bits in a packet flipped
  - Solved by channel coding
    - Adding redundancy to the encoded data
    - Decoders can fully recover the data if there are errors (if not serious) in reception
    - e.g. Reed-Solomon (RS) code for CD, Cyclic Redundancy Check (CRC) for packet
- Packet Error
  - Out-of-order Reception
    - Each packet may experience different delay in arrival due to network congestion
    - How to solve? Require a larger buffer at client application
  - Packet Lost
    - Arrive too late or never arrive due to buffer overflow in clients/routers
    - How to solve? Try to recover the lost packet

Lost Packet Recovery

- Forward Error Correction
- Interleaving
- Error concealment: produce a replacement for a lost packet
  - Piggybacking
  - Packet repetition
  - Interpolation
- Or a mix of the above methods

Forward Error Correction

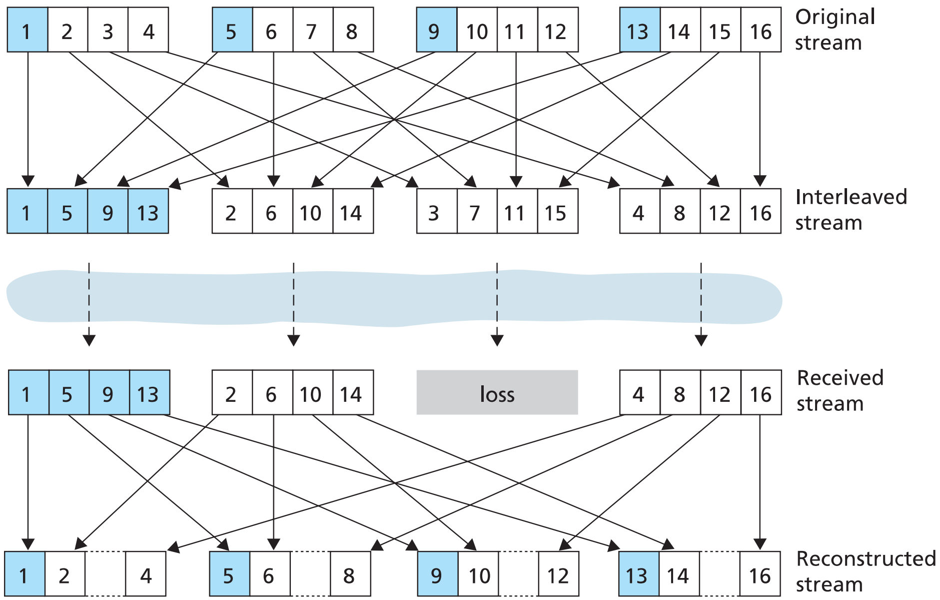
- Forward Error Correction
  - Apply channel coding in packet level
  - Pros: no knowledge to the coding of multimedia
  - Cons: extra bandwidth required and extra playback delay introduced

Interleaving

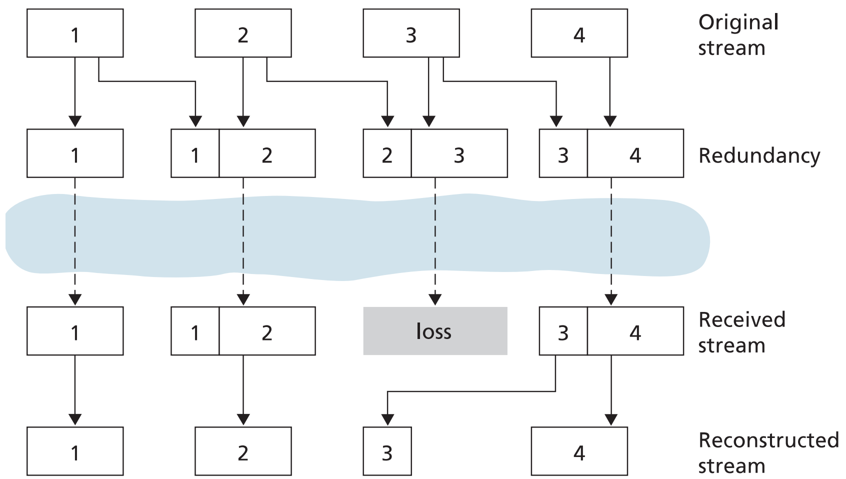
- Interleaving: the sender resequences the data to be sent
  - The display order is different from the transmission order
  - To minimize the effect of packet losses
- Advantage:
  - No extra bandwidth
  - No knowledge is required of the multimedia data
  - Also work for bursty packet loss
- Disadvantage
  - Introduce extra playback delay
  - Not good for real-time applications

Interleaving and Piggybacking

Interleaving



Piggybacking



Piggybacking

- Attach a lower-resolution stream to the next original packet stream
  - Once a packet is lost, client can reconstruct a lower quality version from the succeeding packet
- Advantage
  - Small playback delay (1 packet delay if attaching one lower version to the next packet)
- Disadvantages
  - Does not work for burst packet loss
  - Require extra bandwidth
  - Require knowledge of the multimedia (require to code the data into two versions)

Packet Repetition and Interpolation

- Packet Repetition
  - Simple
- Interpolation
  - Computation intensive but generally better performance than packet repetition
  - Require knowledge of the multimedia

- Or a mix of the above methods