Lecture 11: Media Adaptation

Scalable Coding, Dealing with Errors

Some slides, images were from http://ip.hhi.de/imagecom_G1/savce/index.htm
and John G. Apostolopoulos http://www.mit.edu/~6.344/Spring2004 video_streaming2_2004.pdf

1

Recommended Readings

- **The Scalable Video Coding Amendment of the H.264/ AVC Standard**
  - http://ip.hhi.de/imagecom_G1/savce/index.htm
- **Video Communication and Video Streaming II: Error Resilient Video Coding, by John G. Apostolopoulos**
  - http://www.mit.edu/~6.344/Spring2004 **video**_streaming2_2004.pdf

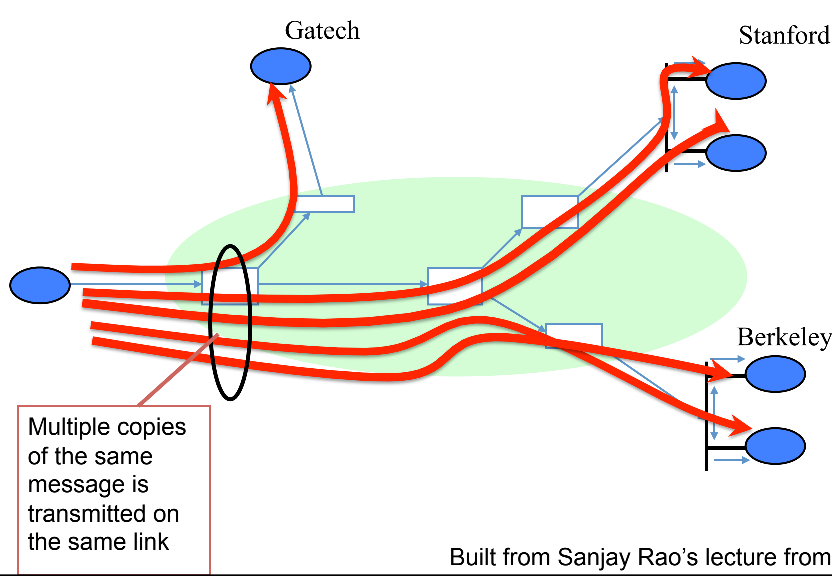
2

Previously

- Ways to distribute video online
  - Client-server
  - IP Multicast
  - Content delivery networks
  - P2P media streaming design
- YouTube Study

3

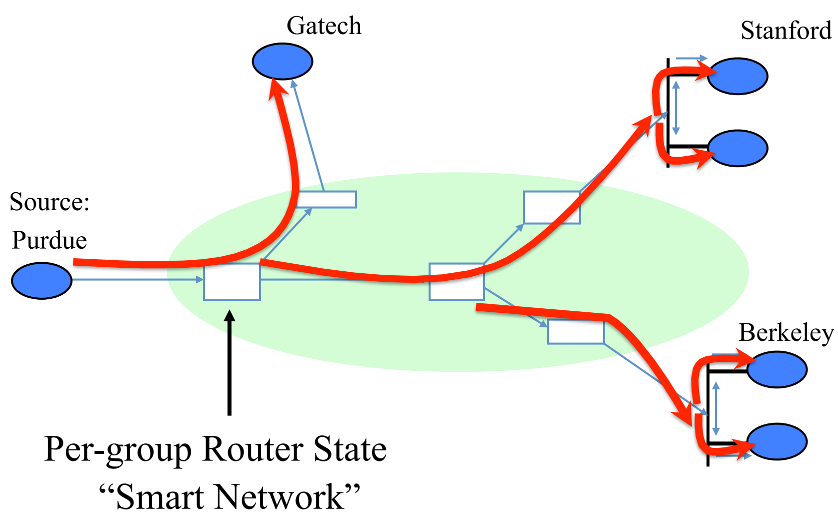
Client-Server



Built from Sanjay Rao's lecture from Purdue

4

IP Multicast



Built from Sanjay Rao's lecture from Purdue

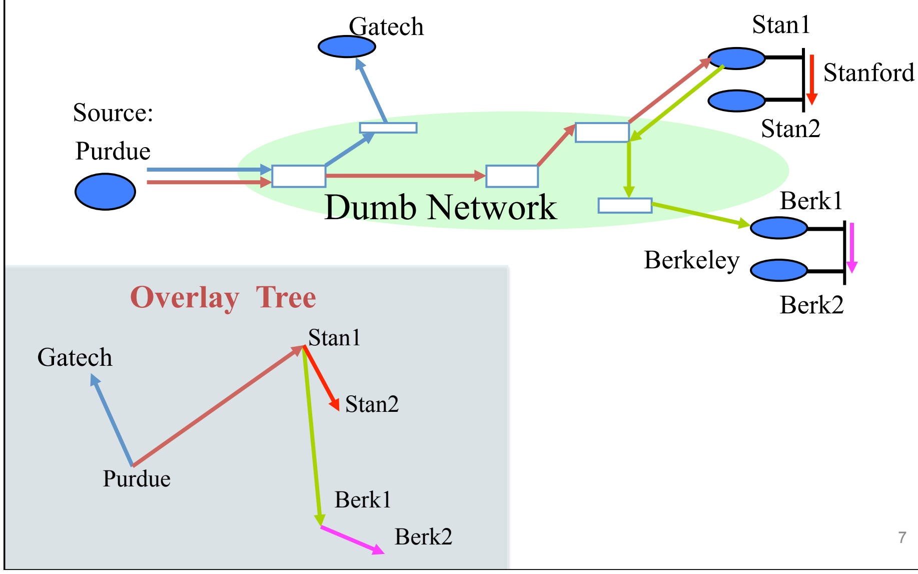
5

Content Delivery Networks

- Strategically located **replicas** **unicast** content to nearby clients
- Clients can access the content on the replica servers nearest to them or the one with less current load.
  - Reduces burden on primary server
  - Improves perceived performance at client

6

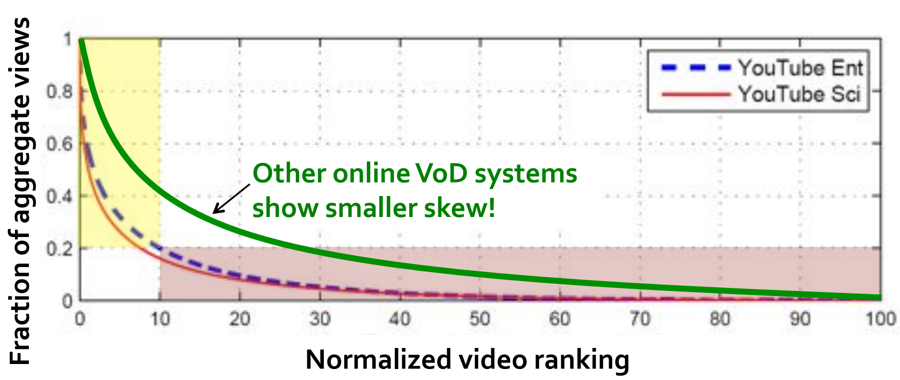
P2P Multicast



7

Pareto Principle

- 10% popular videos account for 80% total views



8

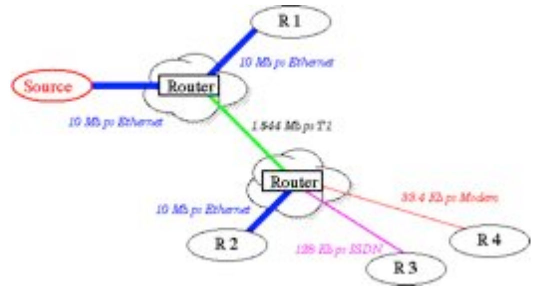
TODAY: WHAT CAN MEDIA DO TO COPE WITH NETWORK ARTIFACTS

9

Network Artifacts

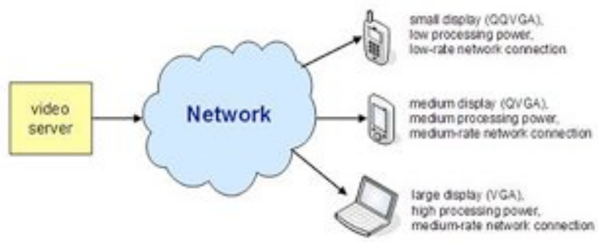
- Network congestion →
  - Time-varying network bandwidth
  - Multiple users (IP multicast) with different bandwidth
  - Packet losses
- Transmission errors →
  - Corrupted bits
  - Common in wireless

10



Artifact #1 Time-varying Network Bandwidth

Artifact #2 Heterogeneous Receivers



11

The Concept of Media Scaling

- Dynamically adapt the media encoding to accommodate changes in available network bandwidth
  - Not that easy!
- **Need to address the following:**
  - Estimate the bandwidth
  - Figure out what can be adjusted
  - How to adapt encoding, and encoding rate
  - What if both audio and video are present? How to partition bandwidth between the two?
  - What if the media is pre-encoded, i.e. MPEG-1.
    - Decode and then re-encode?

12

Estimating Bandwidth

- Requires feedback.
  - How often?
  - In what form?
  - Recall: RTP + RTCP
- TCP-friendliness
  - Current thought: act like TCP.
  - Padhye's equation:

$$B_{TCP} = \frac{M}{t_{RTT}\sqrt{\frac{2Dl}{3}} + t_{out}\min\left(1, 3\sqrt{\frac{3Dl}{8}}\right) l\left(1 + 32l^2\right)}$$

M = packet size

l = loss fraction

D = number of packets ACK'ed by one ACK

13

Feedback

- How often?
  - Want to limit this to once per RTT.
  - Otherwise, rate of feedback is a design decision.
    - Tradeoff: responsiveness vs. feedback bandwidth
    - Should be related to when the source can use the info.

14

Feedback

- In what form?
  - Media independent information:
    - Loss rate and delay
    - RTCP reports from RTCP layer
  - Media specific information:
    - Quality of reconstructed media.
    - Not necessarily well related to loss and delay.
    - User preferences
      - For example: Regions of interest.
    - Very application dependent.

15

Scaling Dimensions

- Different media can be scaled along different dimensions.
- Two major considerations:
  - Encoding support.
    - The media representation often restricts the range and type of scalability.
  - Perceptual quality.
    - Non-linear subjective response to changes.
    - Think of frame size changes.

16

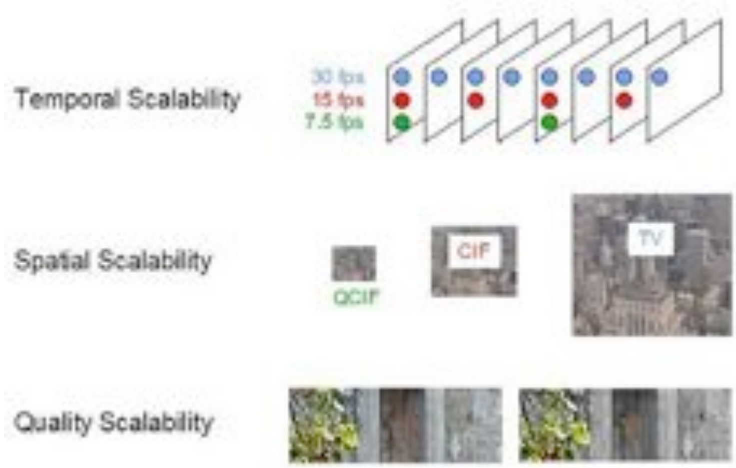
SVC: Scalable Video Coding

- **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** — Quality 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

An excellent read! http://ip.hhi.de/imagecom_G1/savce/index.htm

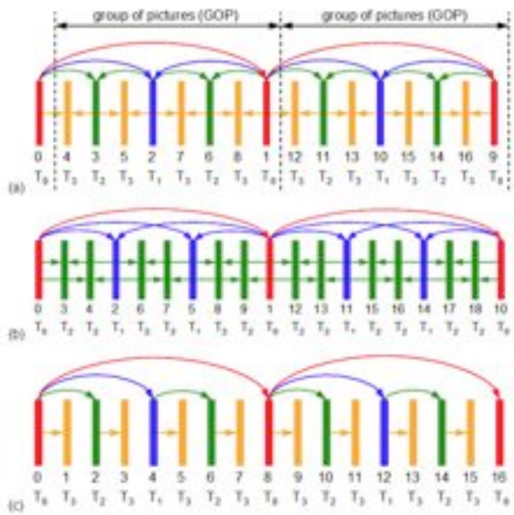
17

SVC (2)



18

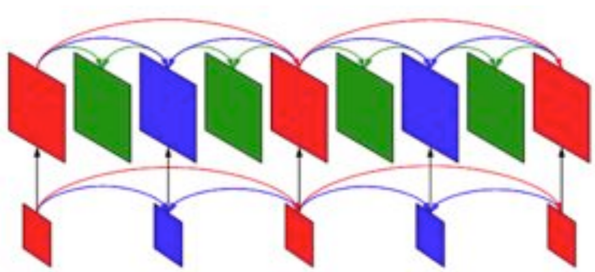
Temporal Scalability



Easy implementation: choose reference frames for motion compensation

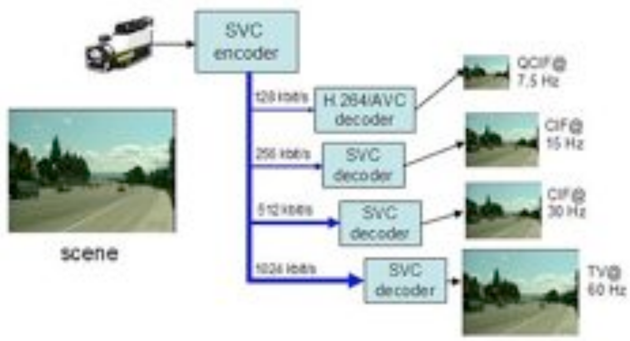
Minimum coding/decoding delay

Spatial Scalability



Multi-layer structure with additional inter-layer prediction (black arrows).

The Result



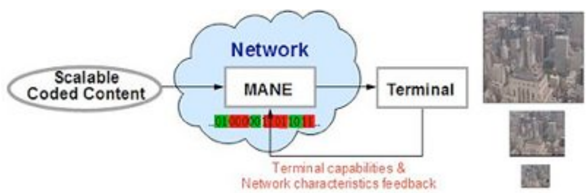
21

Generalizing Scalability → Layered coding

22

Layered Representations

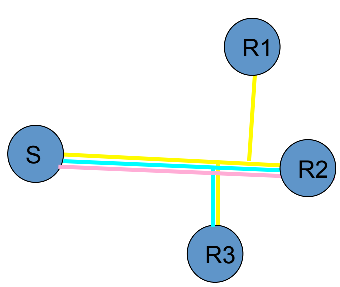
- An encoding specifically designed to produce multiple representations.
- Characteristics:
  - **Additive**: the more layers you get, the better the media
  - **Efficient**: the sum of the layers is only slightly greater than the best rep. at that quality.

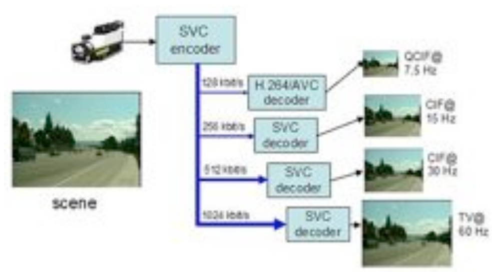


23

Layered Solution

- Use a layered representation.
- Receivers decide
  - Layers added and dropped to adjust to appropriate target rate.





24

Layering Mechanisms

- Strictly Additive Layering
- Independent Layering
  - Multi-description Coding

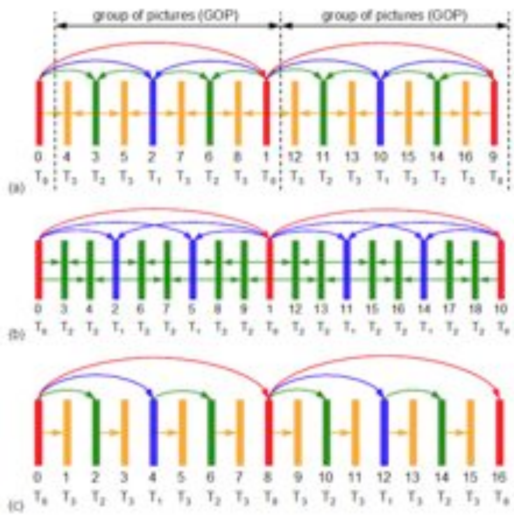
25

Strictly Additive Layering

- Layering split into:
  - Base Layer
  - Enhancement Layers
- Each layer depends on all data in lower layers.
- Advantages: increased compression
- Disadvantages: packet loss in lower layers makes packets in higher layers useless.

26

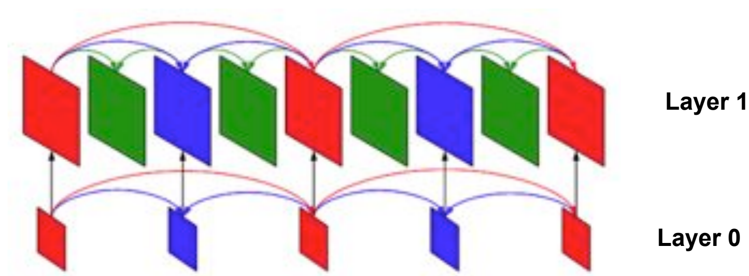
Examples: Temporal Scalability



Need T_i to decode T_j if i<j

27

Examples: Spatial Scalability



Multi-layer structure with additional inter-layer prediction (black arrows).

Need T_i to decode T_j if i<j

28

Independent Layering

- Every packet from every layer improves quality.
- No ordering or dependency between layers.
- Advantages: Good ADU properties
- Disadvantages:
  - Hard to construct
  - Compression suffers
- Most layering schemes fall between these two extremes in some hybrid fashion.

29

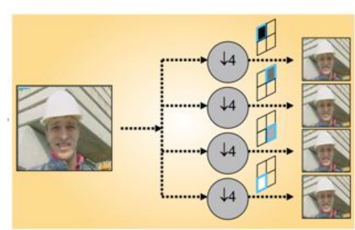
Examples

- Temporal layering
  - Video: Odd frames vs. Even frames.
    - Mostly independent layers.
    - Packet loss affects smoothness.
  - Audio: Multiple, offset, lower-rate streams.
    - Inter-sample compression compromised.
    - Not all combinations equally pleasing.
- Spatial layering
  - Each layer improves video size/resolution.
  - Many of the issues as in temporal.

30

Examples

- Multiple description coding (MDC)
  - Partition a single media stream into n-substreams ($n \geq 2$) referred to as descriptions.
  - Packets of each description are routed over multiple, (partially) disjoint paths.
  - In order to decode the media stream, any description can be used
  - The quality improves with the number of descriptions received in parallel.



http://en.wikipedia.org/wiki/Multiple_description_coding

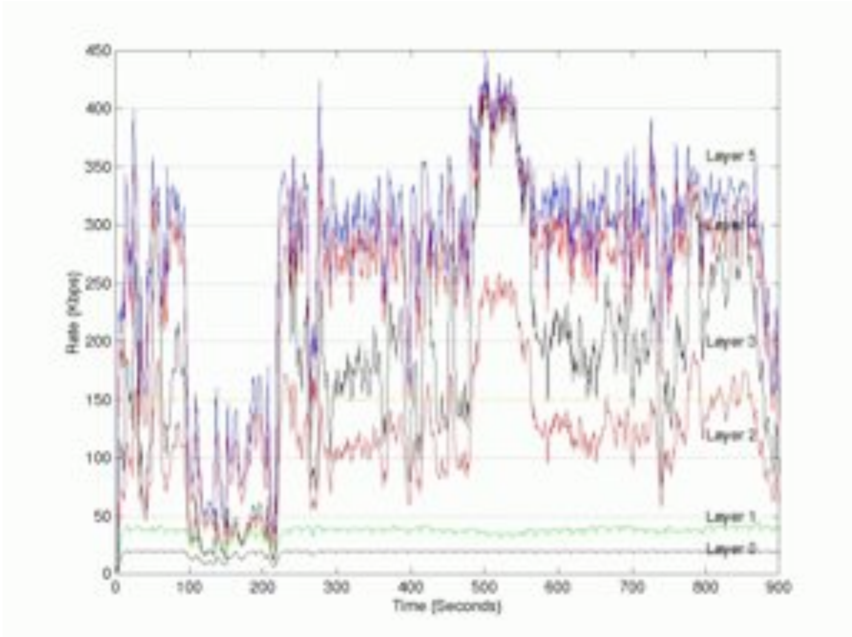
V. K. Goyal, "Multiple Description Coding: Compression Meets the Network," IEEE Signal Processing Magazine, vol. 18, no. 5, pp. 74–94, Sept. 2001.

31

Question

- SNR layering
  - Layers contain different DCT coefficient ranges.
    - DC and first few AC
    - Low AC
    - Middle AC
    - High AC
  - **Is this an independent layering?**

32



33

Layered Video Codec Example: LayerStream

- video stream → a base layer + multiple enhancement layers
- The encoder == a lightweight transcoder
  - Takes high quality video as input
  - Transcodes it to a layered stream in real-time.
  - The layers are subsequently packetized and sent to the network
- The decoder : packets are reassembled by the depacketizer and transcoded back into a valid MPEG stream.
- Some enhancement layers may have been dropped by the network due to congestion.
  - This causes the signal-to-noise ratio to decrease.
  - But the client will never experience freezes or artifacts, as real-time delivery of the lower layers is guaranteed by the network at all times.

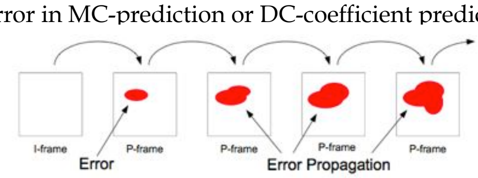
34

Artifact #3 Packet Losses + Corruptions

Video is vulnerable to losses

Assuming conventional MPEG-like system: MC-prediction, Block-DCT, run length and Huffman coding

- Losses create two types of problems
  - **Loss of bit stream synchronization**: Decoder does not know what bits correspond to what parameters
    - E.g. error in Huffman codeword
  - **Incorrect state and error propagation**: Decoder's state is different from encoder's, leading to incorrect predictions and error propagation
    - E.g. error in MC-prediction or DC-coefficient prediction



Measurement-based Study

Effects of Packet Loss on MPEG

Jill Boyce and Robert Gaglianello
ACM Multimedia 1998

37

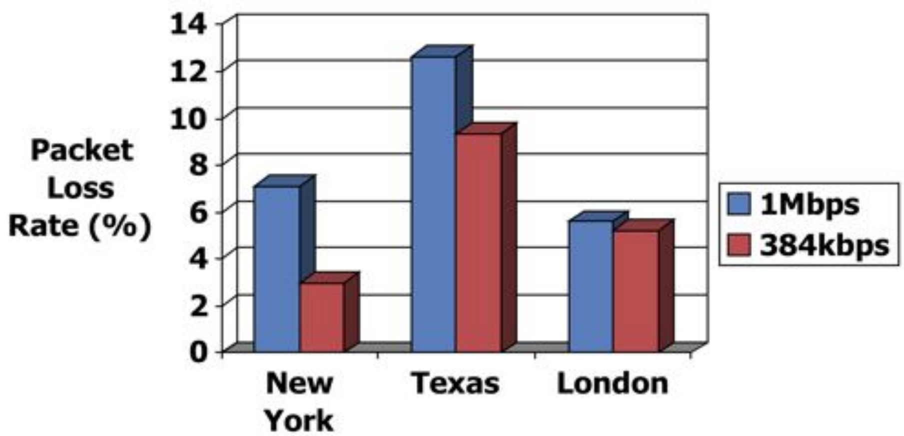
Measurements

Two 5-mins MPEG videos
30 fps
384 kbps and 1 Mbps
IBBPBBPBBPBBPBB
QSIF 176x112 and SIF 352x240
One row per slice

- **From**: NYC , Austin, London
- **To**: Holmdel, NJ

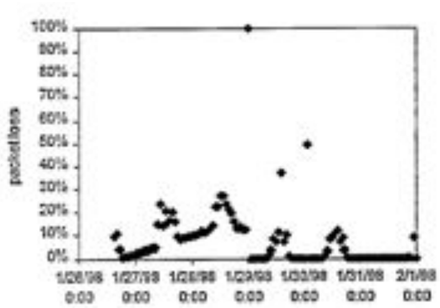
38

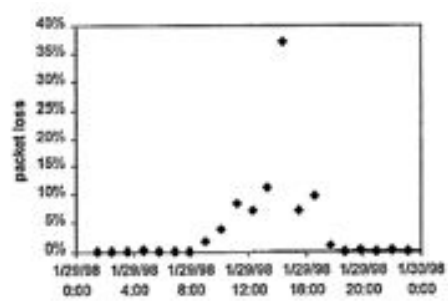
Average Packet Loss



39

Packet Loss vs Time





40

How to Fill Packets

- **B:** 1 frame → 1 packet
- **P:** 1 slice → 1 packet
- **I:** 1 slice → 1 packet

Packet Size Distributions

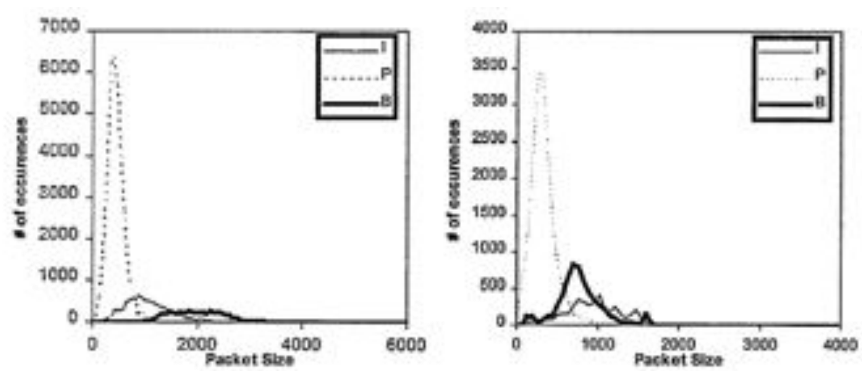


Figure 5. 1Mbps Packet size distribution **Figure 6. 384 kbps Packet size distribution**

Packet Size vs Loss Rate

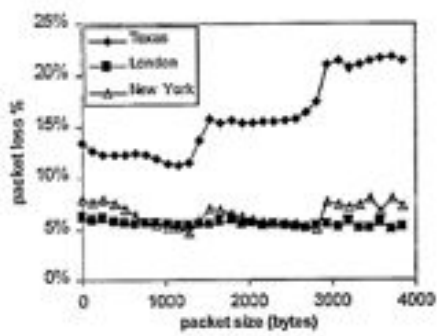


Figure 7. 1 Mbps Packet size vs. loss

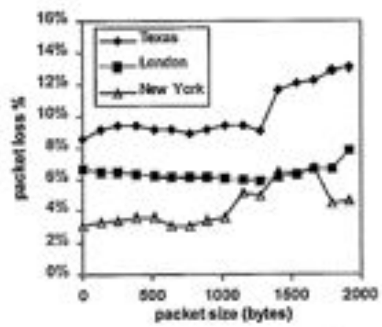


Figure 8. 384 kbps Packet size vs. loss

43

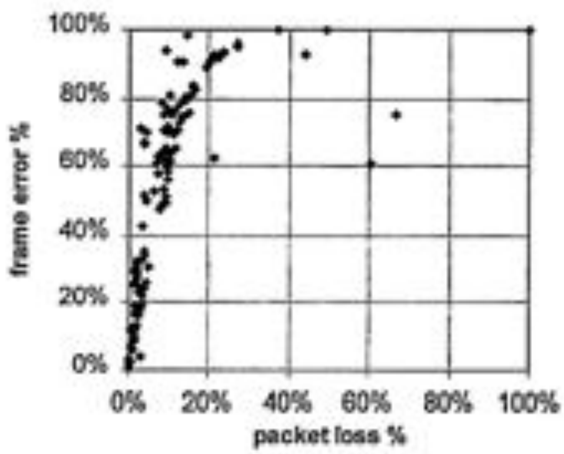
Lesson

- 1 slice per packet even if larger than MTU
- If smaller packets means higher loss rate: fill packet until MTU
  - Else one slice per MTU

- Alt: Change size of slice

44

Frame Effected by Errors



Correlation of Packet Loss

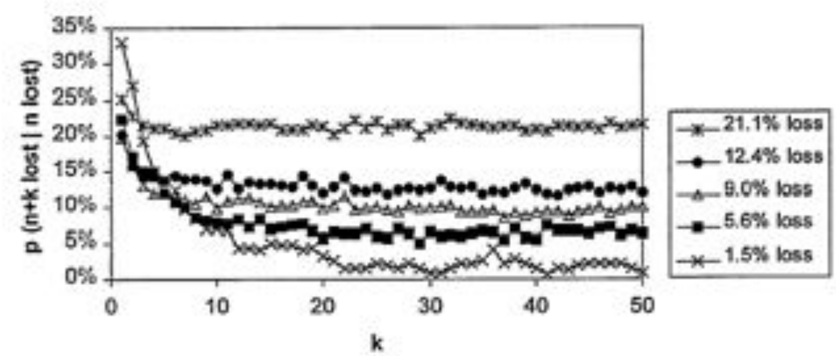
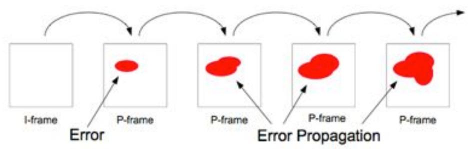


Figure 12. Conditional Probabilities of Packet Loss

Video is vulnerable to losses

Assuming conventional MPEG-like system: MC-prediction, Block-DCT, run length and Huffman coding

- Losses create two types of problems
  - Loss of bit stream synchronization: Decoder does not know what bits correspond to what parameters
    - E.g. error in Huffman codeword
  - Incorrect state and error propagation: Decoder's state is different from encoder's, leading to incorrect predictions and error propagation
    - E.g. error in MC-prediction or DC-coefficient prediction



47

Some Ways to Deal with Packet Losses

- Network-driven (done by the sender)
  - Retransmission
    - Happen infrequently
  - FEC
    - Unequal error protection + layered coding
  - Multiple-description coding
- **Error Concealment (done by the receiver)**
- **Limiting Error Propagation (done by both)**
  - **Reference Frame Selection**

48

Yes We Can Do This

http://www.vidyo.com/resources/videos.php#vidyo_vs_legacy

49

Decoding: Reference Frame Selection

- If a I-frame gets lost, how to decode subsequent P, B-frames?
  - Select the next I-frame and use it as the reference I-frame

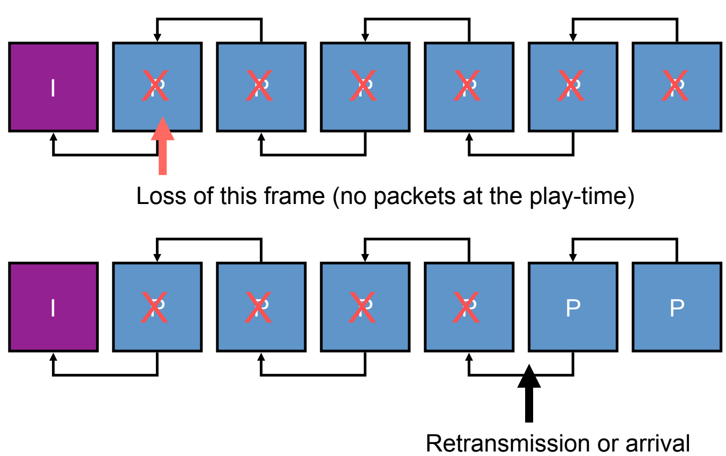
50

Decoding: Use Late Packets

- Packets can be dropped by routers due to congestion
- They can also be dropped by receivers because they arrive late

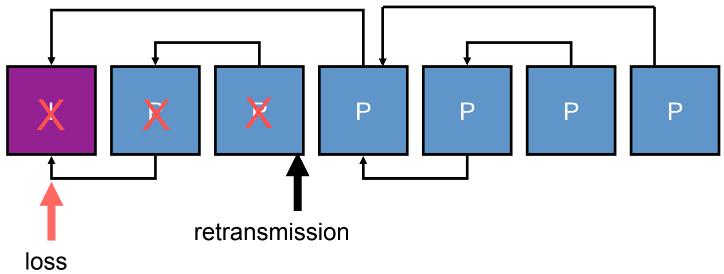
51

Illustrative Examples



52

Encoding: Changing Temporal Structure



Need to protect I-frame strongly, adding more FEC or allow retransmissions

Another solution: use all I-frames? No error propagation but BAD compression

53

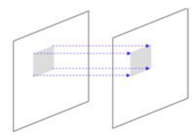
Decoding: Error Concealment for Videos

- Estimate the lost information in order to conceal the fact that an error has occurred
- Performed at the decoder
- Possible because:
  - Video exhibits a significant amount of correlation along the spatial and temporal dimensions
- Basic approach:
  - Perform some form of spatial/temporal interpolation to estimate the lost information from correctly received data

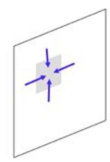
54

Error Concealment for Video

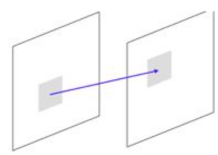
- **Repeat** pixels from previous frame
  - Effective when there is no motion, potential problems when there is motion



- **Interpolate** pixels from neighboring region
  - Correctly recovering missing pixels is extremely difficult, however even correctly estimating the DC (average) value is very helpful



- **Interpolate** motion vectors from previous frame
  - Can use coded motion vector, neighboring motion vector, or compute new motion vector



55

Impact of Error Concealment

Figure 1
On the left, errors are not concealed. On the right, state-of-the-art concealment has been applied

56

www.ebu.ch/en/technical/trev/trev_312-vitali_streaming.pdf

An excellent tutorial

- Video Communication and Video Streaming II: Error Resilient Video Coding by John G. Apostolopoulos

*http://www.mit.edu/~6.344/Spring2004/**video**_streaming2_2004.pdf*

57

Summary of Today's Learning

- Media coding meets networking
  - How can media address various networking artifacts
- Scalable, error-resilient coding
  - Layered coding (dependent vs. independent layering)
  - Reference frame selection
- Error recovery/concealment

58