Multimedia Over Networks

N. C. State University

CSC557 ♦ Multimedia Computing and Networking

Fall 2001

Lecture # 17

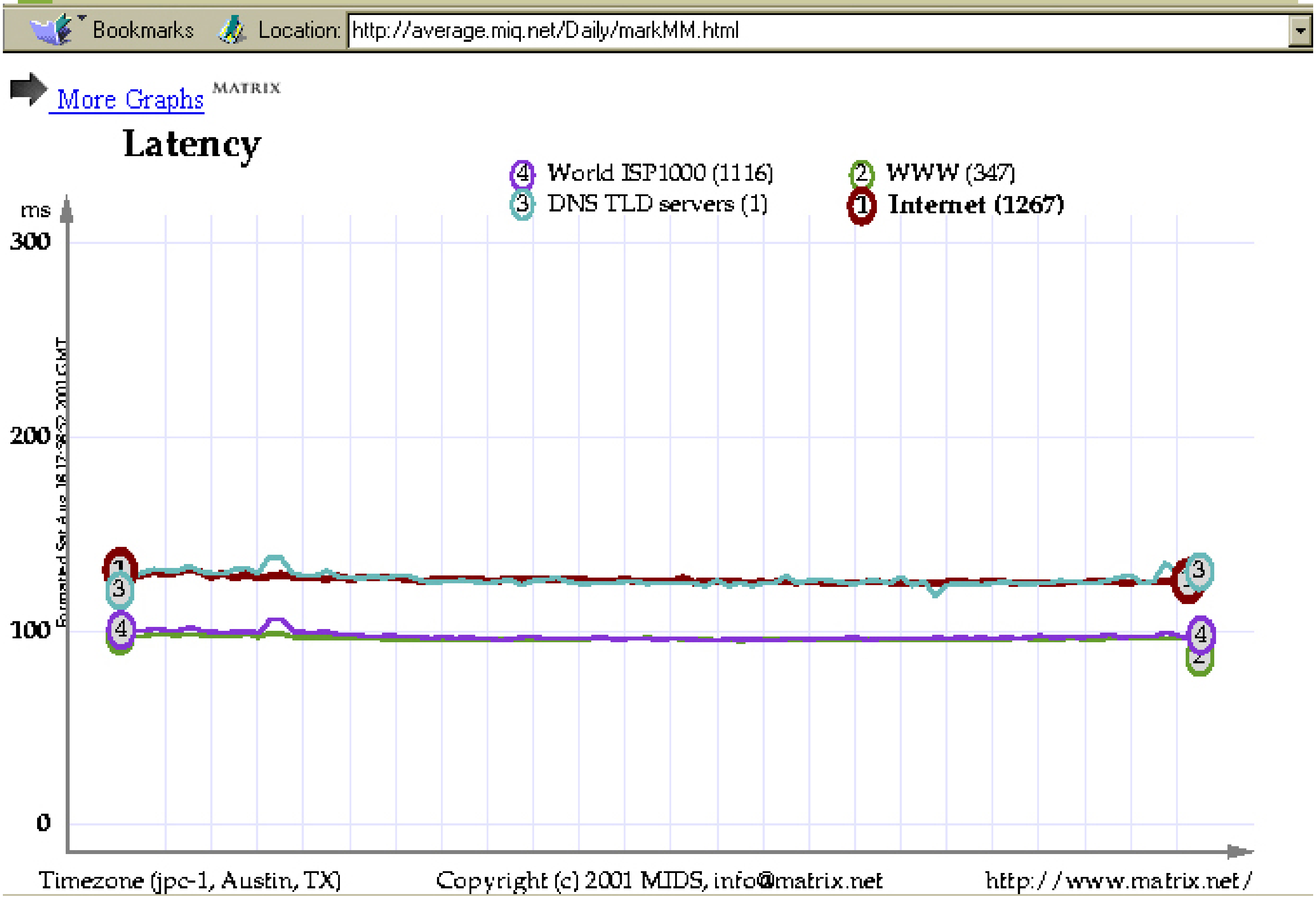
QoS Guarantees

- Deterministic (100%) guarantees
  – based on peak traffic rate
  – simple, predictable, conservative
- Statistical (< 100%) guarantees
  – based on peak and mean traffic rates
  – complex, less predictable, higher utilization
- No guarantees
  – the network performance is what it is
  – "best effort" service

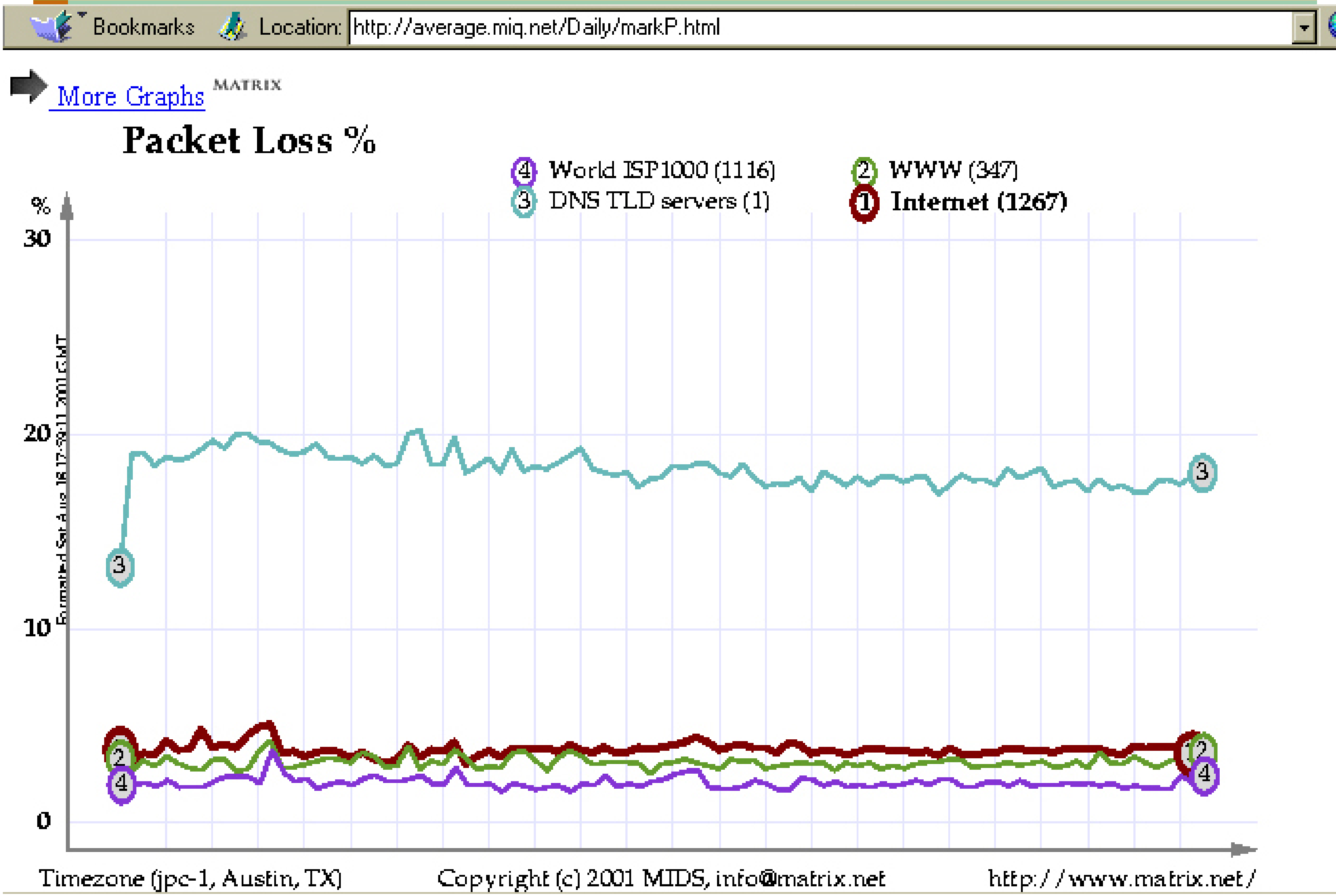
Problems

1. Less-than-ideal average delays and loss rates
2. Variations in traffic loads in the network
3. TCP's congestion control
4. Retransmission-based error recovery
5. Simplistic routing algorithms
6. "Burstiness" or variability of a single traffic source
   - peak rate, average rate, maximum burst size

Problem 1: Average Internet Latency



Problem 1: Average Internet Packet Loss

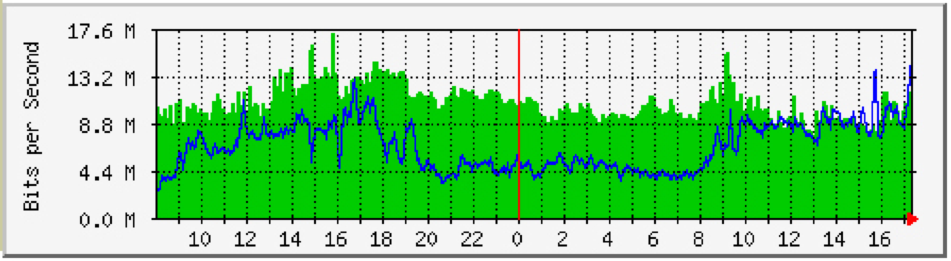


Problem 2: Variations in Network Traffic Load

Bookmarks Location: http://clio.ucs.ualberta.ca/canet2.gw.ualberta.ca_7.html

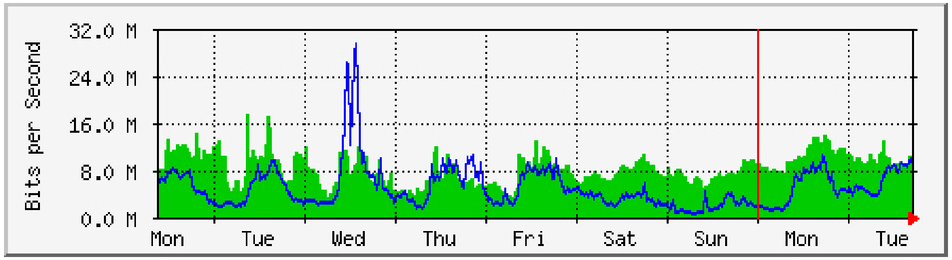
The statistics were last updated **Tuesday, 31 July 2001 at 17:20**, at which time **'canet2.gw'** had been up for **37 days, 13:44:04**.

`Daily' Graph (5 Minute Average)



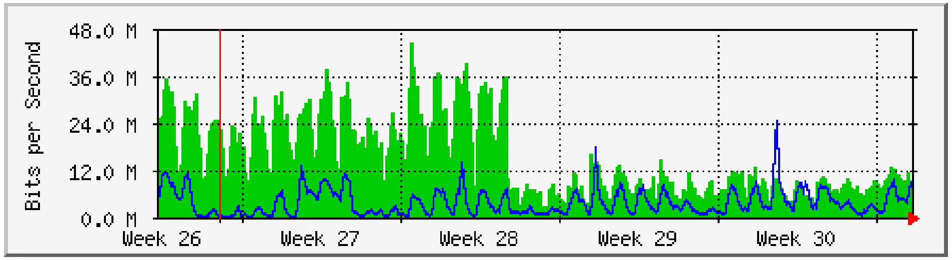
Max In: 17.4 Mb/s (17.4%)	Average In: 10.7 Mb/s (10.7%)	Current In: 11.6 Mb/s (11.6%)
Max Out: 14.2 Mb/s (14.2%)	Average Out: 6796.2 kb/s (6.8%)	Current Out: 12.2 Mb/s (12.2%)

`Weekly' Graph (30 Minute Average)



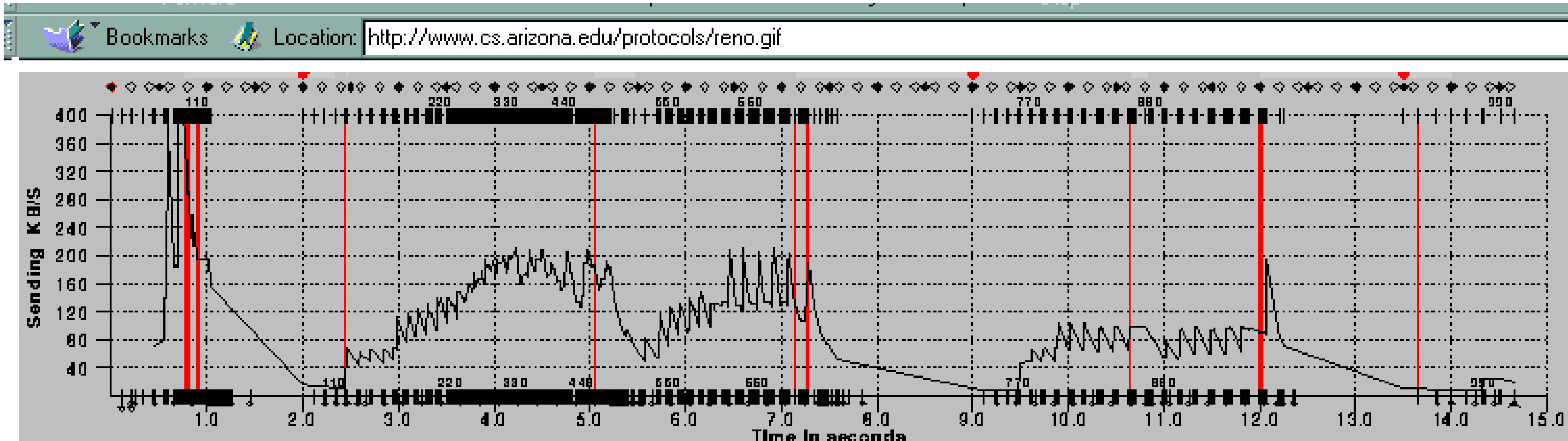
Max In: 17.7 Mb/s (17.7%)	Average In: 8487.0 kb/s (8.5%)	Current In: 10.1 Mb/s (10.1%)
Max Out: 29.6 Mb/s (29.6%)	Average Out: 5246.8 kb/s (5.2%)	Current Out: 9254.8 kb/s (9.3%)

`Monthly' Graph (2 Hour Average)

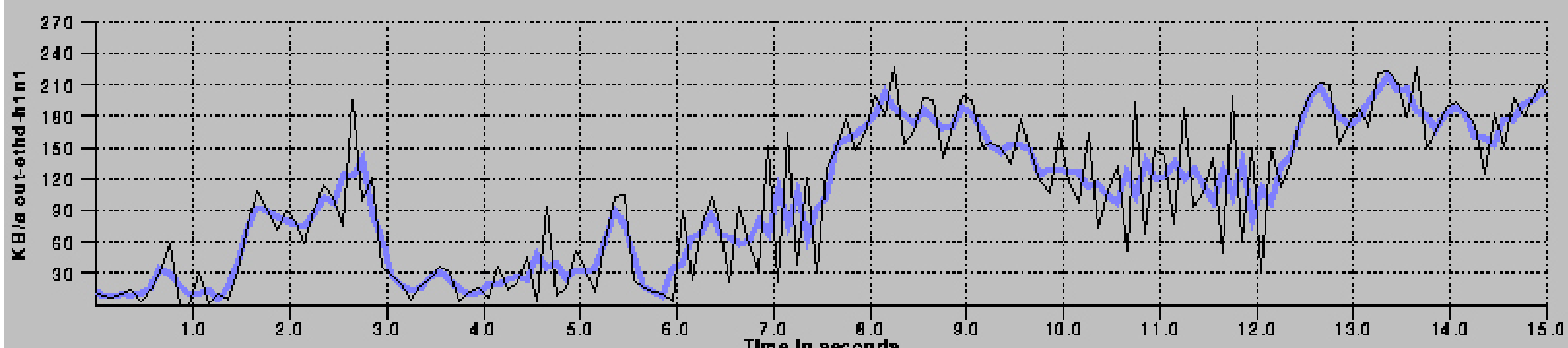


Max In: 44.7 Mb/s (44.7%)	Average In: 15.2 Mb/s (15.2%)	Current In: 9257.3 kb/s (9.3%)
Max Out: 24.6 Mb/s (24.6%)	Average Out: 4042.4 kb/s (4.0%)	Current Out: 8671.2 kb/s (8.7%)

Problem 3: TCP's Reactive Congestion Control



Throughput of one TCP flow



Background traffic load

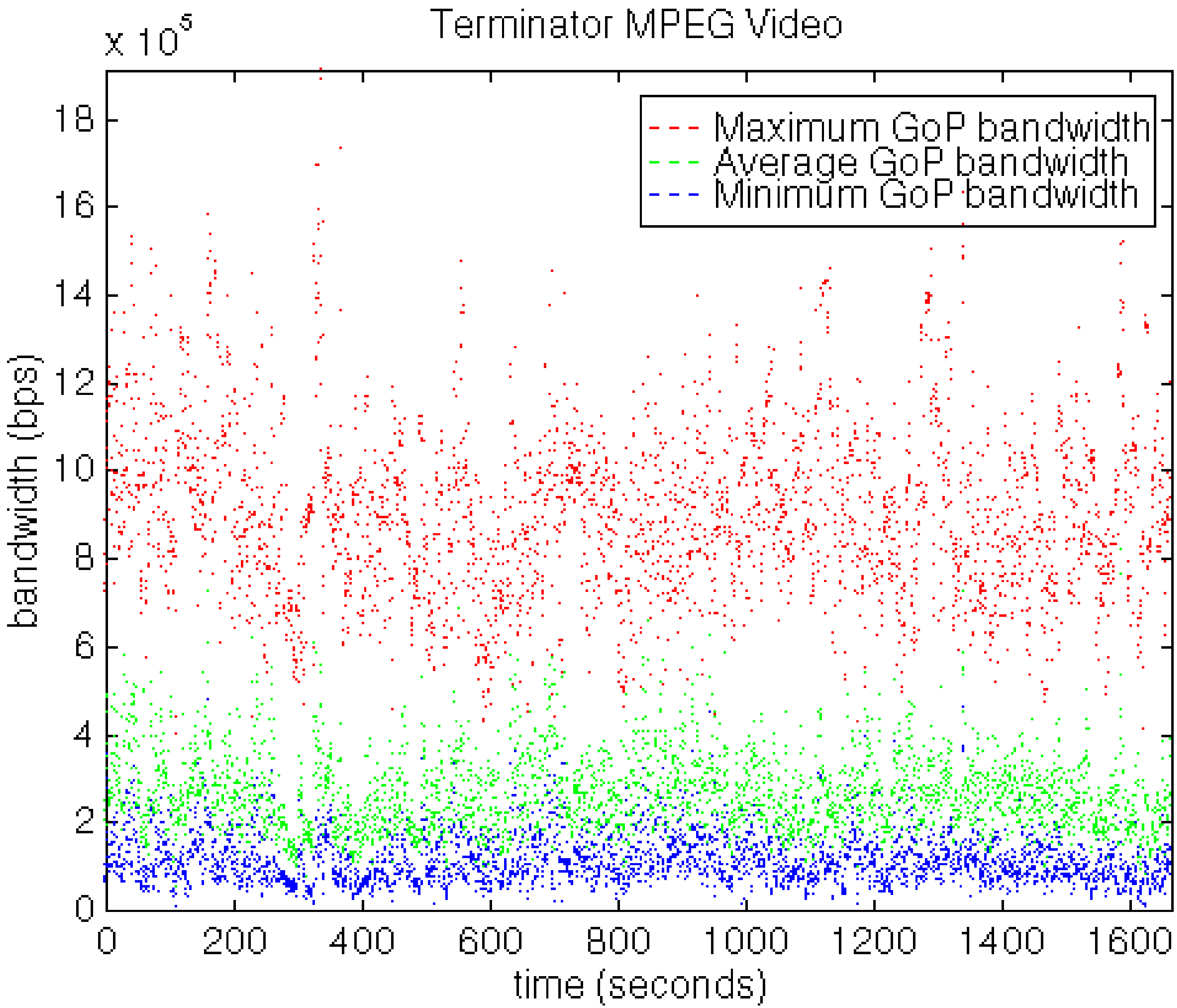
Retransmission-Based Error Recovery

- To recover from lost / corrupted packets you have to…
  - wait until you're sure it didn't arrive, or
  - wait until receiver asks you to retransmit specifically
  - then resend
- Problem: too much delay!
  - at least 3 one-way network delays retransmit and receive

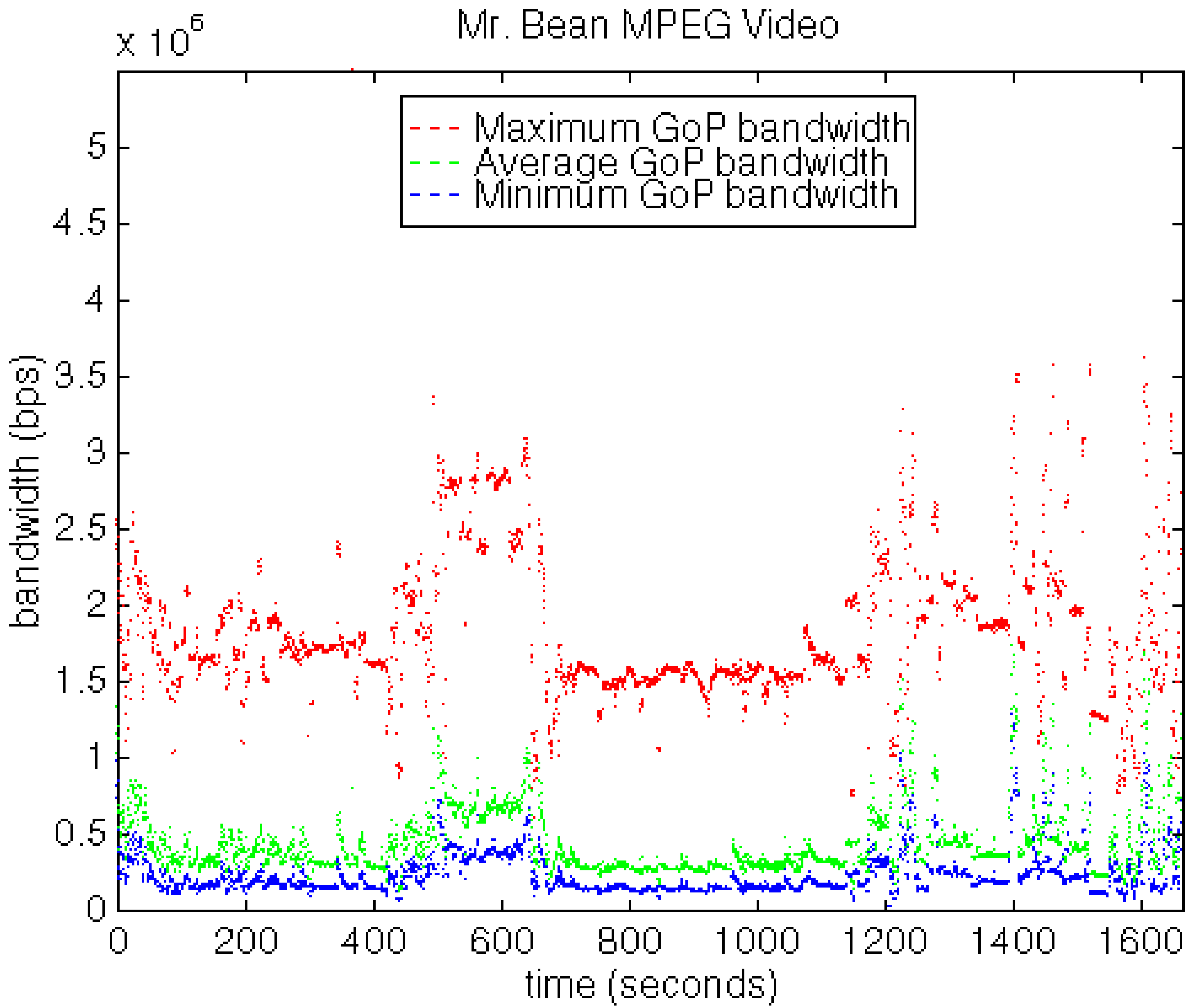
Problem 5: Simplistic Routing Algorithms

- Major routing criteria
  - within (intra-) a single network: minimum number of "hops"
  - between (inter-) networks: "friendliest neighbor"
- Where does "best available quality" come in???
  - neither of these considers congestion
  - inter-network may (usually doesn't) consider traffic type in making routing decisions

Problem 6: Burstiness of a Traffic Source (MPEG Video)



Problem 6: Burstiness (cont'd)



Which Multimedia is Being Networked?

- We will study voice transmission primarily
  - most important application currently
- Illustrates issues common to other real-time media types
- Very likely in near future
  - much more video transmission

What Contributes to End-to-End Delay in the Network?

- Packetization delay
- Access delay
- Serialization delay -- packet length / bandwidth (bits/sec)
  - Incurred at both sender and receiver
  - usually very small at high line speeds
- Propagation delay
  - around 20ms from east coast to west coast
- Handling (queueing and switching) delay
  - queueing delays go down as line speeds go up, if queue lengths stay roughly constant

Other Sources of Delay

- Operating system overhead and granularity
  - e.g., Windows timer interrupts only every 60 ms
- Jitter buffer delay
  - only way to smooth deliver is to add delay
- Voice codecs process speech in frames
  - time to accumulate a frame's worth of data for compression
  - time to do the compression (sender)
  - time to decompress (receiver)
  - shorter frame = less delay, longer frame = better compression
- Time for error recovery
  - time to receive redundant packet information
  - delay increased if redundancy spread out to avoid correlated losses

Example of Handling Delay

- Assume 1000 bit packets
- Assume average queue length = 1000 packets (10^6 bits)
- Then average queueing delay at 1 Mb/s = 1 second
- Or average queueing delay at 10 Gb/s = .1 millisecond

Compression Delays Example

- G.729 generates one block of samples every 10ms
- Two blocks per packet = 20ms
- Lookahead interval for G.729 = 5ms

Delay Example

5ms coder lookahead delay

20ms coder delay (two frames/packet)

(0ms packetization -- included in coder delay)

5ms low-speed trunk queueing delay

3ms low-speed trunk serialization delay

30ms propagation delay

? ms wide-area network queueing and switching delay

50ms jitter buffer at receiver

———

113+ ms total, not including WAN or decoder delay (if any)

Voice Quality

- The E-Model
  - Evaluates subjective model of voice quality
  - Impairments due to echo, jitter, compression lossiness, etc.
  - Scores: 90+ = very good, …
- Mean Opinion Score (MOS)
  - subjective measure of quality
  - 5 (excellent), 4 (good), 3 (fair), 2 (poor), 1 (bad)
- Codec MOSs
  - G.711 (mulaw): 4.2
  - G.726 (ADPCM): 4.2
  - G.728 / G.729 (CELP): 4.0
  - GSM 6.10 (cell phone): 3.8—3.8
  - USA CDMA (cell phone): 3.3—3.5

Echo

- Echos occur electrically (termination) and acoustically (speaker to microphone)
  - when one-way delay is 25 ms or greater, echo is annoying!
- Echo cancellation: detect and remove echo when it occurs
  - Can be done electrically (analog), or in software (digital)

Packet Loss Effect on Voice Quality

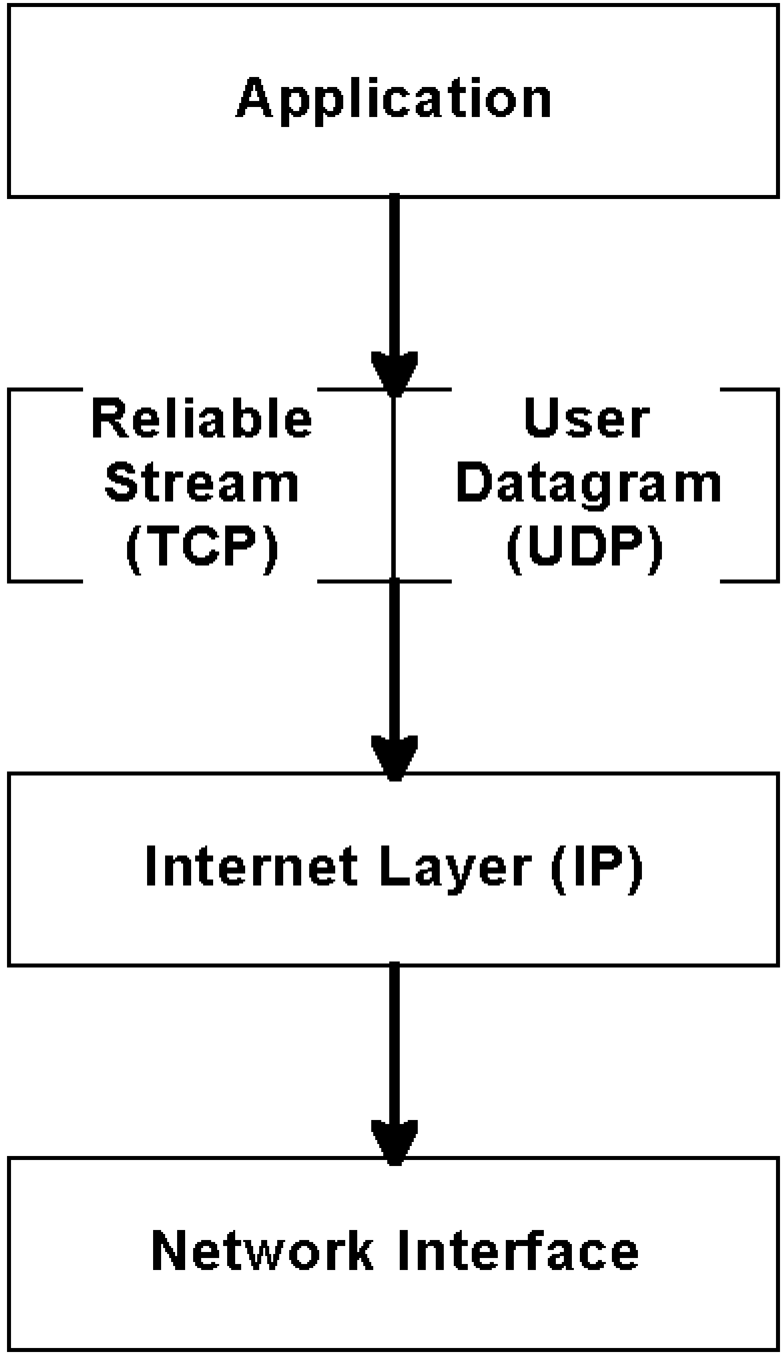
- 1-3% of lost packets is tolerable
  - allowable level depends on voice codec, latency, etc.
  - can use "concealment" technique to compensate for missing packets
- Burst (clustered) packet losses will cause noticeable problems
  - how spread them out?

Voice Activity Detection (VAD)

- Don't transmit bits when no speech activity
  - examples: listening to audio recording, letting other person talk, voice conference call, ...
- Monitor voice signal, if goes below a threshold for x sec., stop transmitting until it rises above threshold
  - example: x = .2 sec.
- Delay in detecting onset of voice can lead to “front end” clipping
- Problem: background (room, or ambient) noise

Internetworking

- On top of ethernet, phone lines, leased lines, frame-relay, ATM, etc.
- Connectionless inside network; varying routes and delays
- No time bounds, bandwidth restrictions, or priority support


Application
Reliable Stream (TCP)
User Datagram (UDP)
Internet Layer (IP)
Network Interface

IP

- Provides packet formatting, segmentation, and reassembly
- Identifies source and destination hosts for routing
- Routing
  - shortest paths
  - dynamic route updates
  - Consideration of delay / congestion?

Transport Protocols

- UDP
  - simplex (one-way) communication without acknowledgement
  - unreliable delivery of datagrams: lost, duplicated, out-of-order delivery possible!
  - good for multicasting, non-critical applications, and/or applications requiring low delay
- TCP
  - connection-oriented transport service
  - duplex (two-way) connections
  - reliable delivery of segments
    - in-order
    - recovery from losses/errors
    - no duplicates

UDP vs. TCP

- Multimedia (particularly voice transmission)
  - can tolerate a modest amount of packet loss
  - but needs low delays
- TCP has
  - 0% packet loss
  - but can have high delays (retransmission time)
- UDP is the transport method of choice

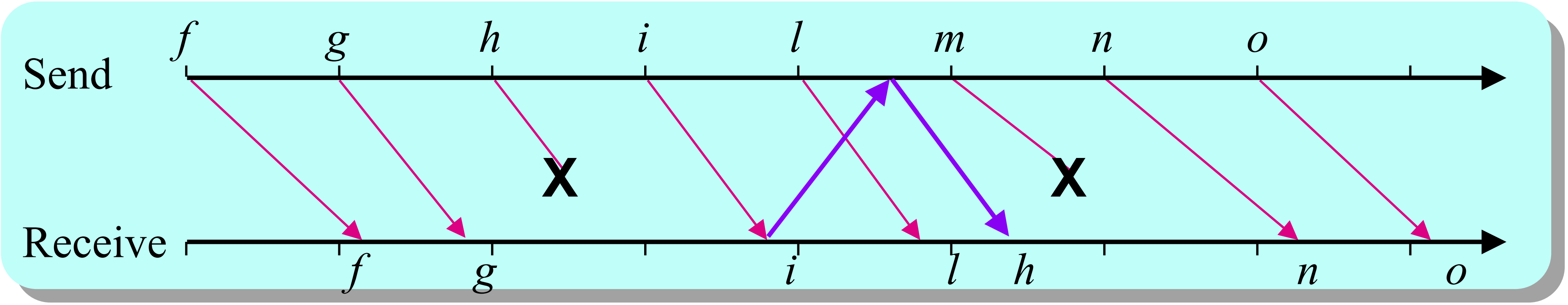
Handling Packet Loss: Basic Approaches

- Traditional "reactive" approach
  - receiver acknowledges transmissions
  - if sender fails to receive acknowledgment, retransmits packet
    - Assumption: packet must have been lost in transmission
    - "Automatic Repeat Request" (ARQ)

- Two proactive approaches
  1. introduce redundancy into streams to enable reconstruction of lost media samples
     - "Forward error correction" (FEC)
  2. dynamically adapt streams to the bandwidth perceived to be available at the current time
     - congestion control
     - media scaling

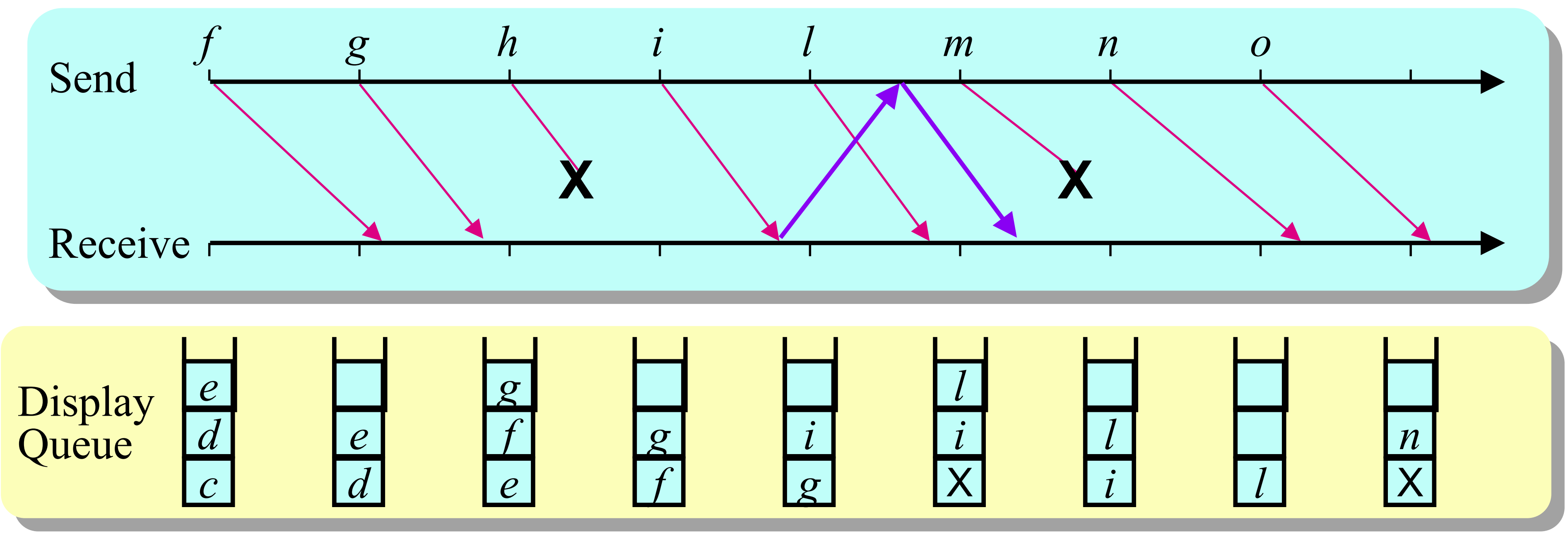
Retransmission-Based Error Correction
Conventional wisdom

- "Retransmission is silly…"
  - by the time you realize something is lost, it's too late to retransmit it
  - retransmission techniques do not work well with multicast applications



Retransmission-Based Error Correction The retransmission process

- 1. Loss is detected
- 2. A retransmission request is issued
- 3. The requested packet is retransmitted



Retransmission-Based Error Correction
The retransmission “budget”

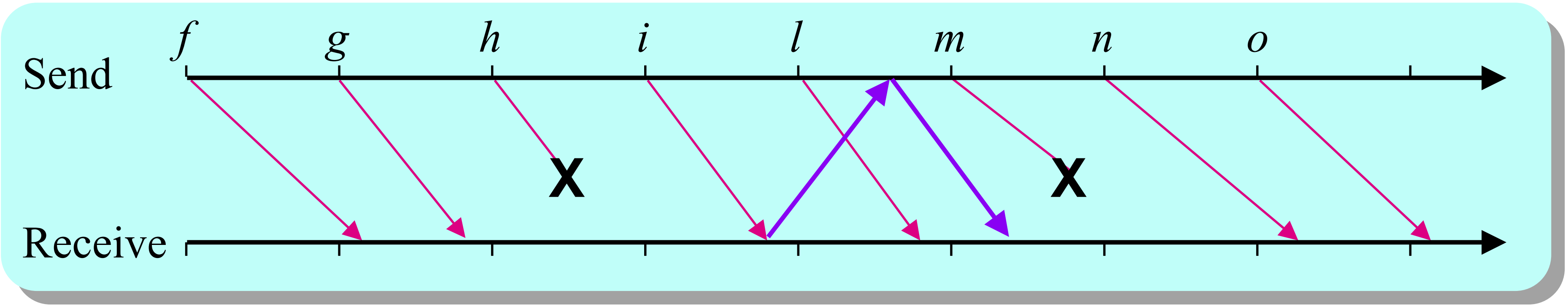
- Retransmission is potentially beneficial...
  - Since data is buffered at the receiver to ameliorate the effects of jitter, provide intermedia synchronization, etc., retransmission may work!

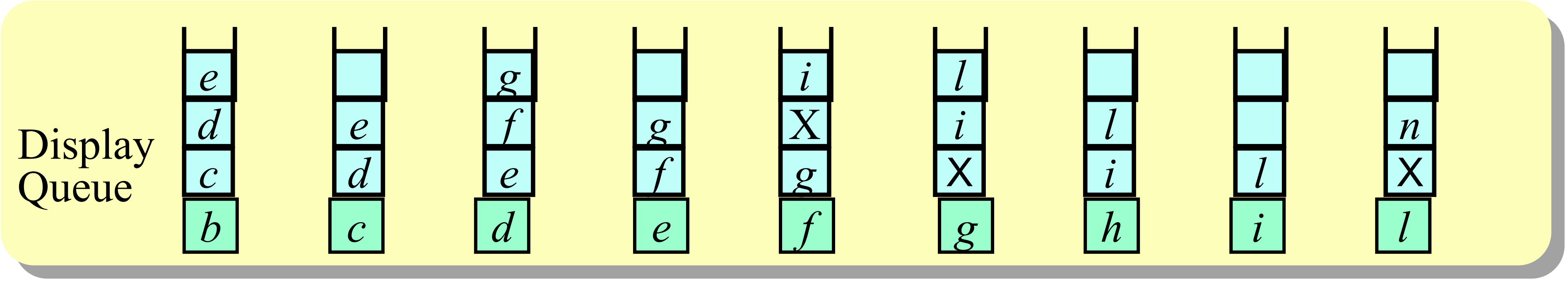
- If:

$$\textit{Interframe Interval} + (3 \times \textit{one-way transmission time}) < \textit{playout latency}$$

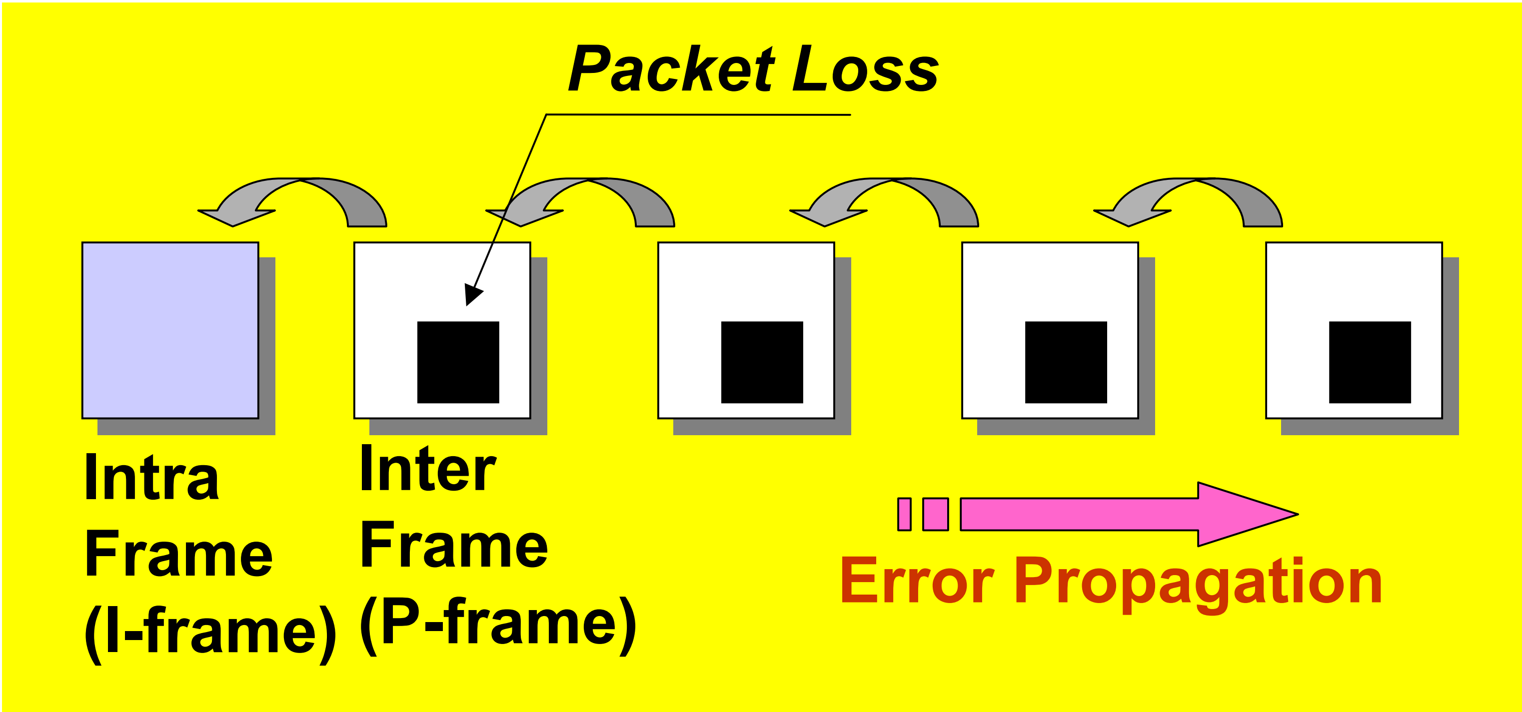
then retransmission is a possibility

Example


Send
f
g
h
i
l
m
n
o
X
X
Receive

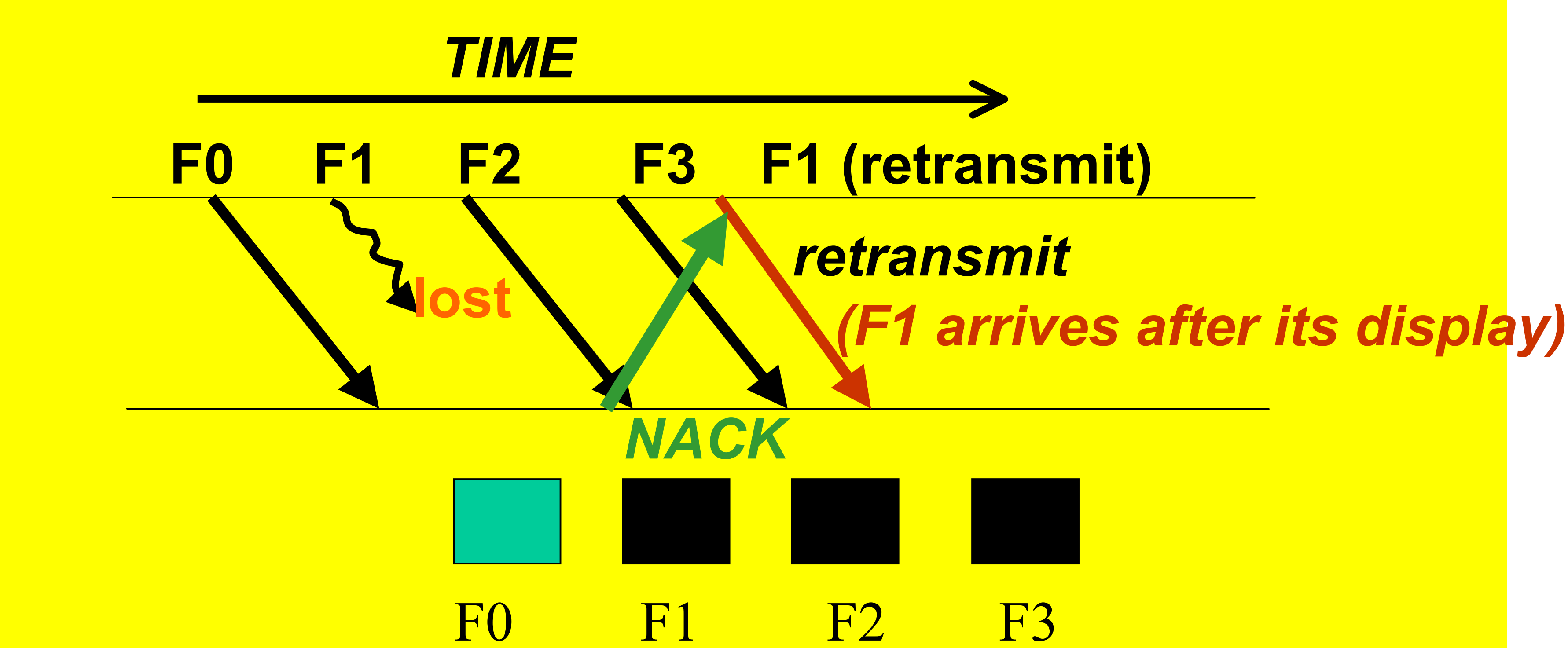

Display Queue
e d c b
e d c
g f e d
g f e
i X g f
l i X g
l i h
l i
n X l

Packet Loss for Videos



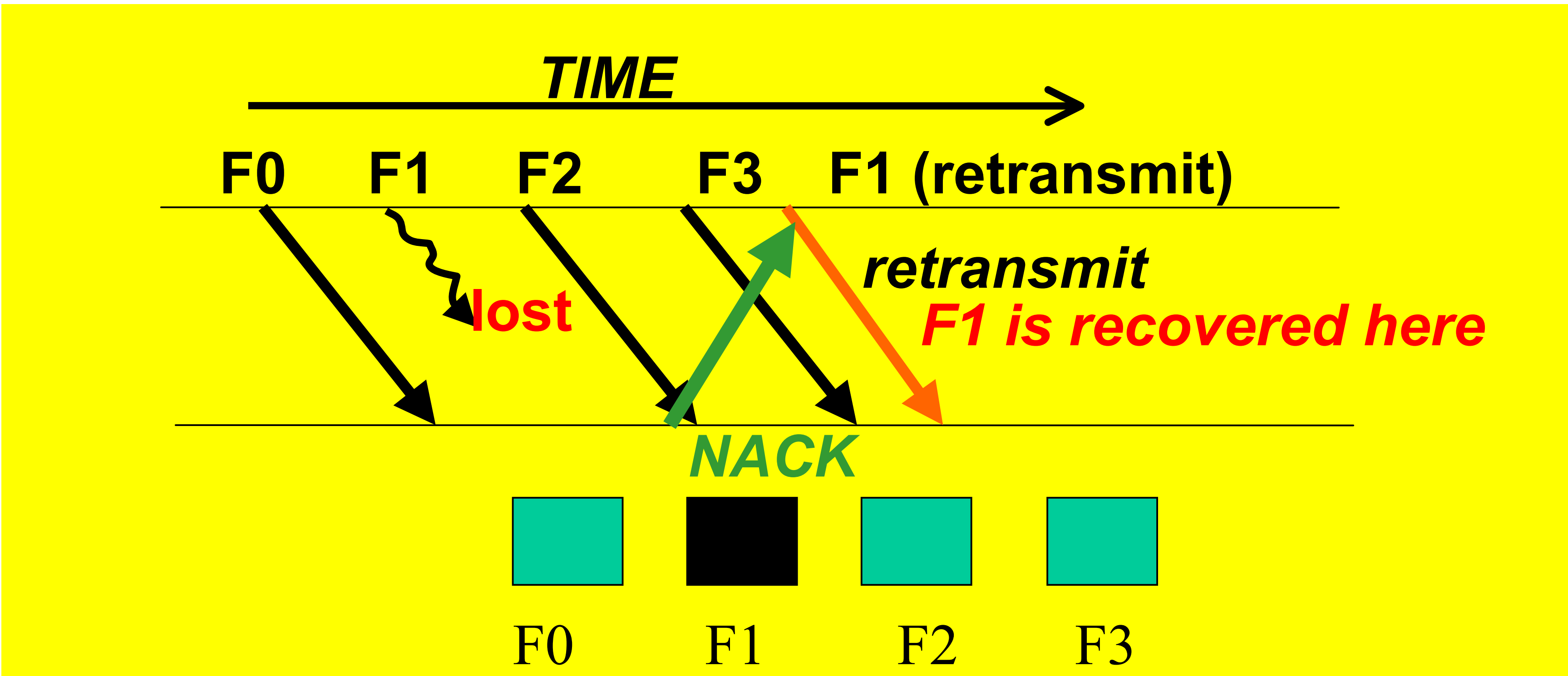
- Problem: packet losses in P-frames propagate
- Conventional solution: transmit more I frames
  - drawback: compression ratio goes down

Retransmission Not Useful for Interactive Video?



- Problem: frame information arrives too late to be displayed

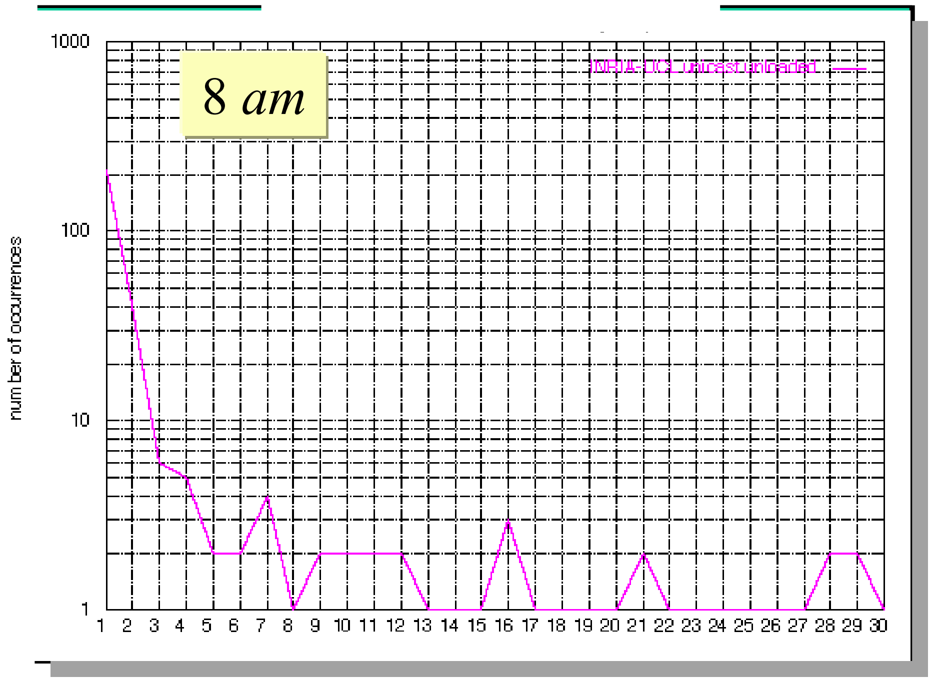
Retransmission Can Stop Error Propagation

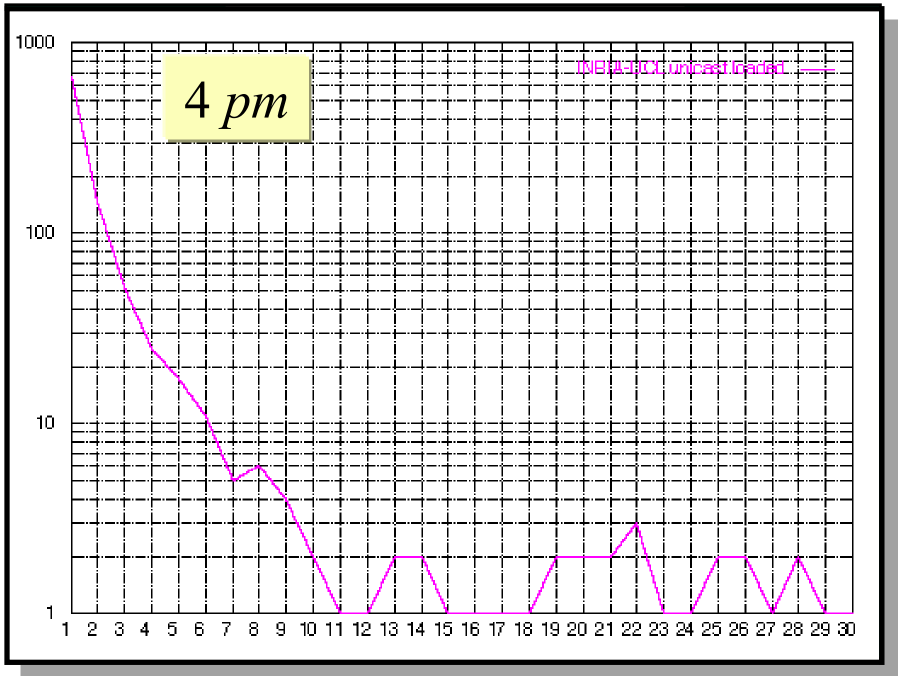


- Don’t discard late (retransmitted) packets!
  - use them to recover the reference frames

Consecutive Packet Losses; Common?

- Frequency distribution of consecutive packet losses from INRIA to UCL





number of occurences of n consecutively lost packets v. n

Sources of Info

- Recommended
  - Hersent, *IP Telephony*, 2000 (chapter 4)
- Optional
  - Douskalis, *IP Telephony*, 2000 (chapter 2)
  - Davidson, *Voice over IP Fundamentals*, 2000 (chapters 8 and 9)
- Web sites
  - Quality of Service Technical White Paper
  - Quality of Service glossary of terms
  - White paper on QoS Protocols and Architectures