Trace driven simulation of *rate adaptive* MPEG-4 video

Arne Lie

Q2S colloquium NTNU 02.12.05

Outline

- Motivation and problem
  - for rate adaptive video
  - how to simulate real rate adaptive media traffic
- Solution:
  - Evalvid tool-set modified: “Evalvid+”
  - “SVBR” modified
- How to use
- Early results
  - Subjective: perceptual quality
  - Objective: PSNR
- Summary

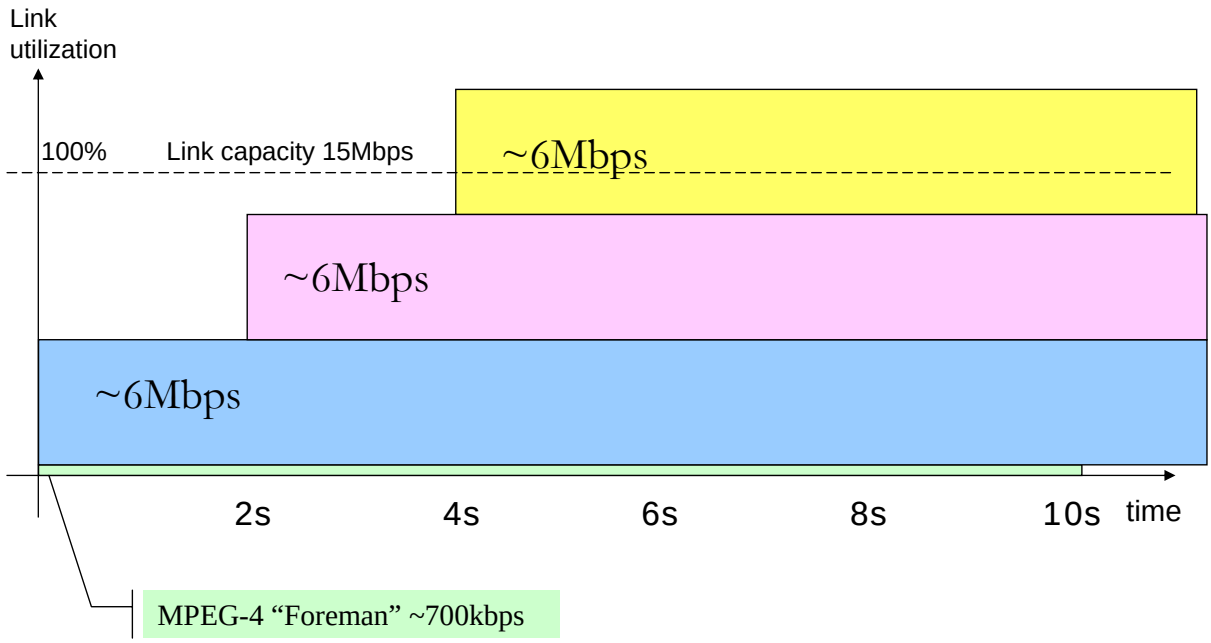
Congestion control

- Fast growth in usage of real-time applications
  - Interactive: VoIP, videoconferencing, gaming, …
  - One-way: Web-TV (live), VoD
- Congestion Collapse
  - TCP congestion control make Internet stability
  - Real-time applications currently are not scalable
- Avoiding congestion collapse
  - Admission Control (“call blocking” as in telephony)
  - Congestion control/avoidance
    - IETF drafts of DCCP (Datagram Congestion Control Protocol)
    - My proposal: “P-AQM + ECF”
    - Others…

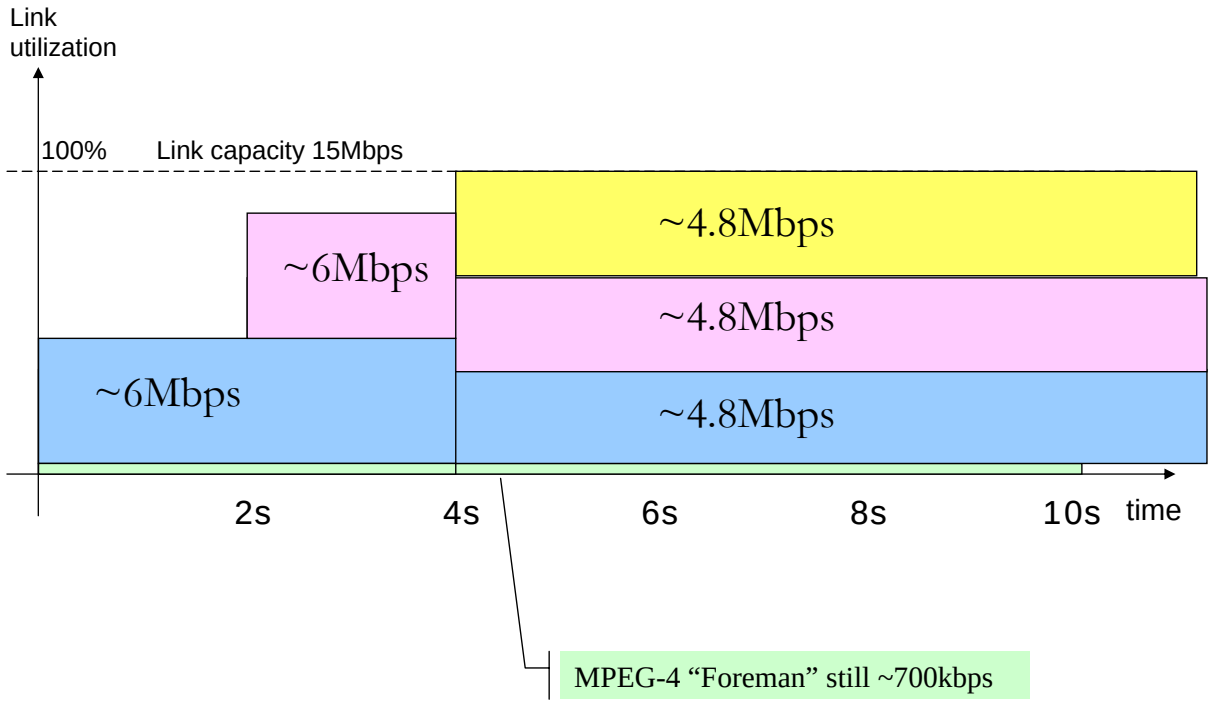
How to inspect media quality?

- CC objectives for media: high perceived quality!
  - Avoid persistent long queues
    - low latency
    - low drop probability
  - Bandwidth:
    - Fair bandwidth
    - Avoid unnecessary large rate reduction
    - Grab available excess bandwidth
- How can we do research by network simulation?
  - Run traffic with the right characteristics
    - Use source models
    - Use trace driven simulation
  - Perceived quality: need real media!
    - Evalvid tool-set
    - But we need “online” *rate adaptive* trace simulations

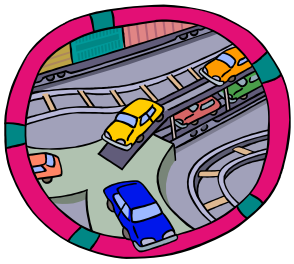
Scenario for original Evalvid tools



Competitive traffic is modeled…

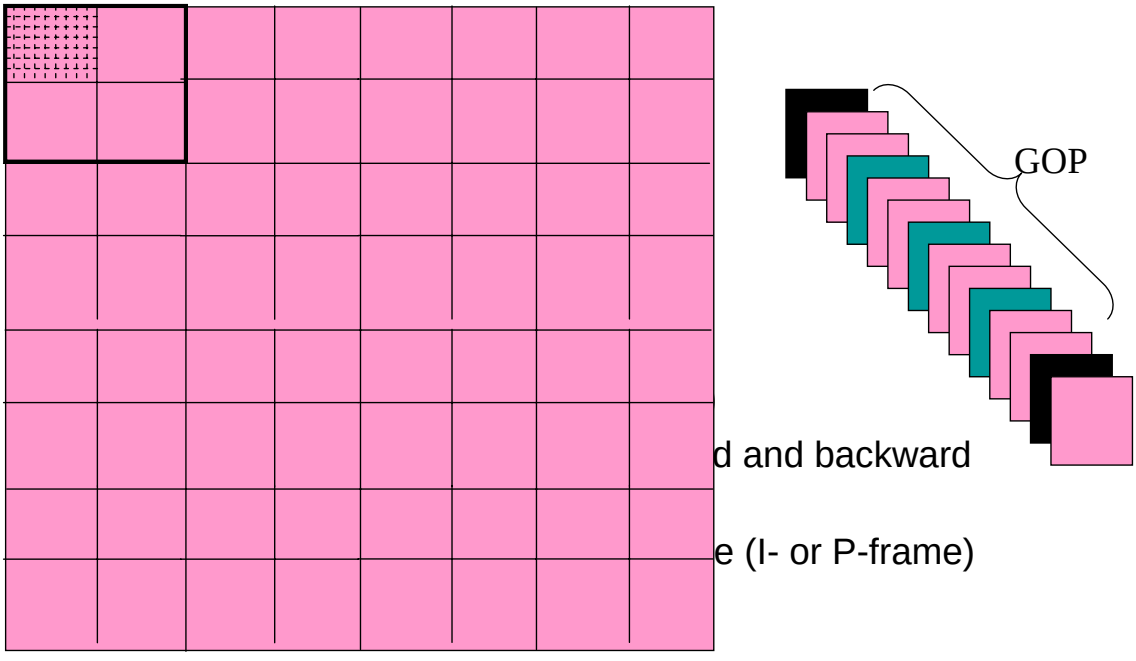


What is old



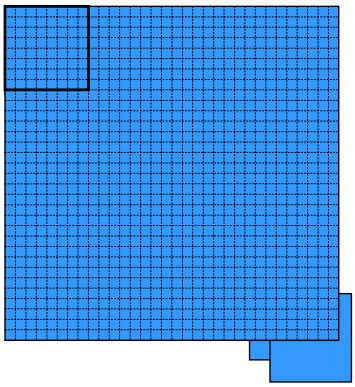
- Trace driven simulation need trace files from real sources, e.g.
  - http://www-tkn.ee.tu-berlin.de/research/trace/ltvt.html
  - Only the frame SIZES and timing is used, not the content
- or synthetic traffic that models real traffic very closely
  - e.g. GenSyn http://www.item.ntnu.no/~poulh/GenSyn/gensyn.html
- Evalvid tools from http://www.tkn.tu-berlin.de/research/evalvid/
  - real traces, and media is re-assembled after network simulation for visual inspection and PSNR calculation (Jirka Klaue)
- Evalvid interface to ns-2 (Ke Chih-Heng)
  - http://hpds.ee.ncku.edu.tw/~smallko/ns2/Evalvid_in_NS2.htm
- but *rate adaptive* media will change depending on network state…

Video coding in one slide



Scalable media

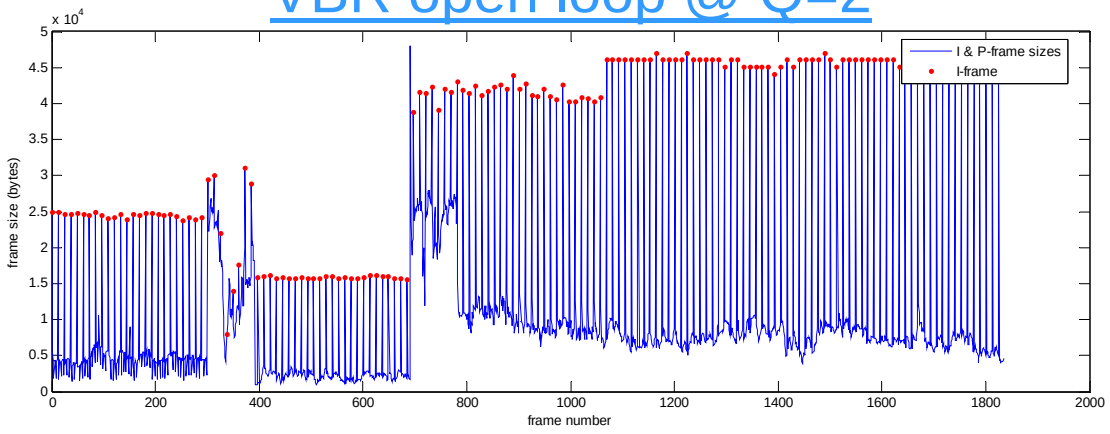
- change either (video / audio)
  - frame rate / sample rate (temporal)
  - frame size / sample size (spatial)
  - compression quantization Q (quality)
    - `=quantiser_scale` in MPEG-4
- or a combination
- Most players/decoders don't respond (correctly) to changes in frame size and frame rate
  - ➔ change of the Q-value (=quantiser_scale) is easiest
  - the Q-value actually normally change each frame, or even each macro block (video)
  - but how to avoid doing this "live" in the network simulation?

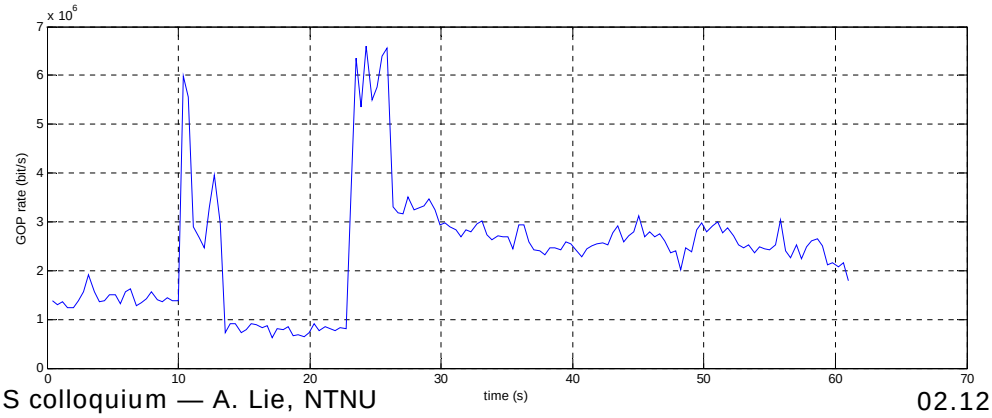


Rate controllers varies Q

- Adjust output rate according to a bit rate budget on time average and variability constraints
  - leaky bucket
- CBR: constant bit rate
  - each GOP has the same number of bits (or bit/s)
  - Q changes from macro block to macro block
  - Cost: algorithmic delay, variable quality
- VBR: variable bit rate
  - allows for more variability
  - Q changes less: more stable quality
- Quality based ("VBR open loop", constant Q)
  - rate totally dependent on content

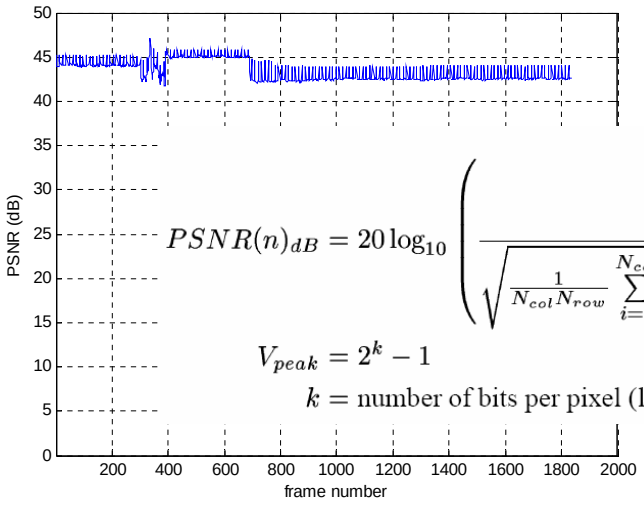
VBR open loop @ Q=2





Objective quality at Q=2: PSNR

- Constant Q gives ~ constant quality

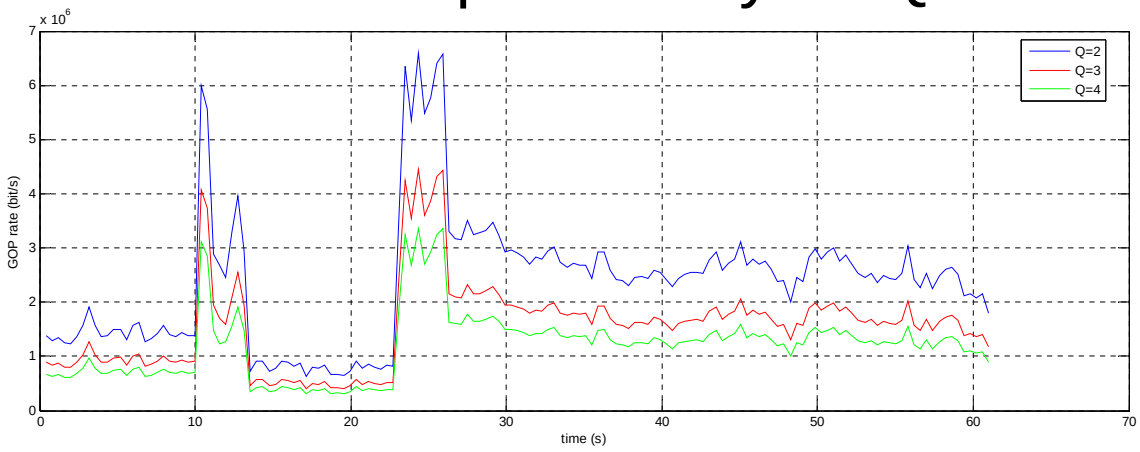


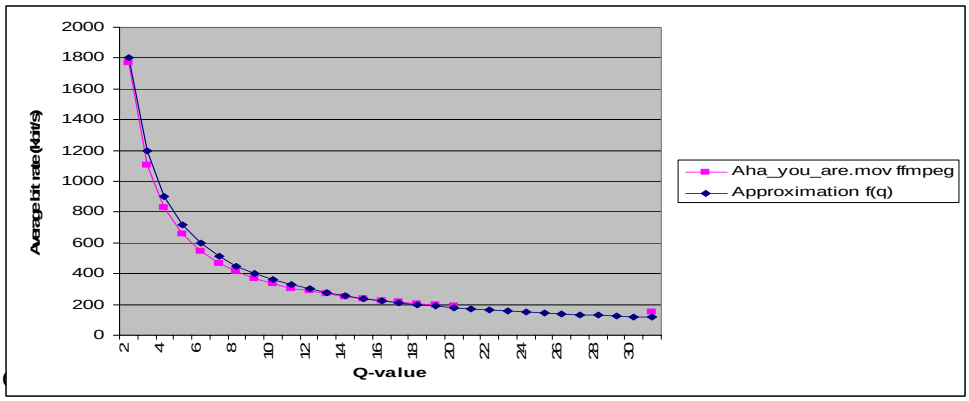
$$PSNR(n)_{dB} = 20\log_{10}\left(\frac{V_{peak}}{\sqrt{\frac{1}{N_{col}N_{row}}\sum_{i=0}^{N_{col}}\sum_{j=0}^{N_{row}}\left[Y_S(n,i,j) - Y_D(n,i,j)\right]^2}}\right)$$

$$V_{peak} = 2^k - 1$$

k = number of bits per pixel (luminance component)

Rate dependability on Q





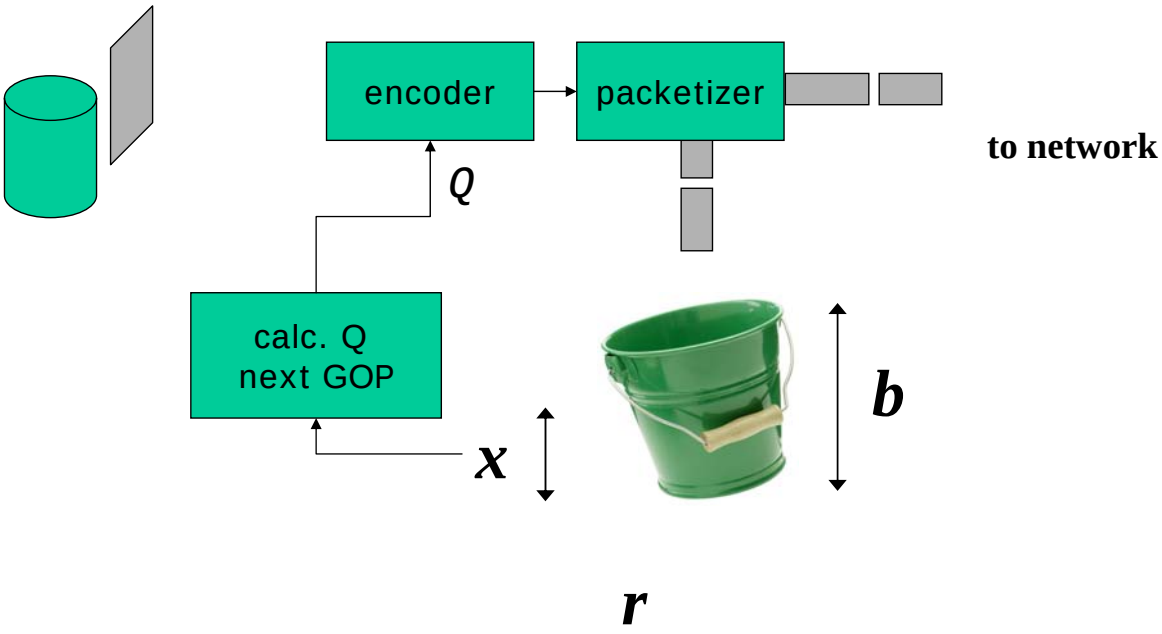
Rate controller objectives

- Limits the rate fluctuations & have an <u>average</u> rate constraint, by varying the quantization value Q
  - at each macro block
  - at each frame,
  - or at each GOP
- If Congestion Control is applied
  - the rate controller must have *adaptable* average rate constraint!
  - Problem: the rate controller must run at simulation time!

How to avoid having “online” encoder

- CBR changes Q at macro block granularity
  - too detailed for frame size trace files!
- VBR changes Q at frame or GOP granularity
  - Yes!
- “SVBR” (shaped VBR) by Hamdi/Roberts/Rolin ’97
  - change Q at GOP scale to constrain video to LB(*r*,*b*) constraint
    - *r*: average video rate (=leaky bucket rate)
    - *b*: bucket size (to allow variability)
  - very simple, no extra delay
  - my modification: variable *r*

Hamdi’s SVBR leaky bucket controller



How to map *r* to *Q*

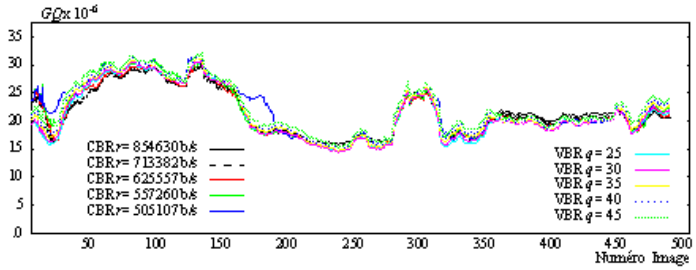


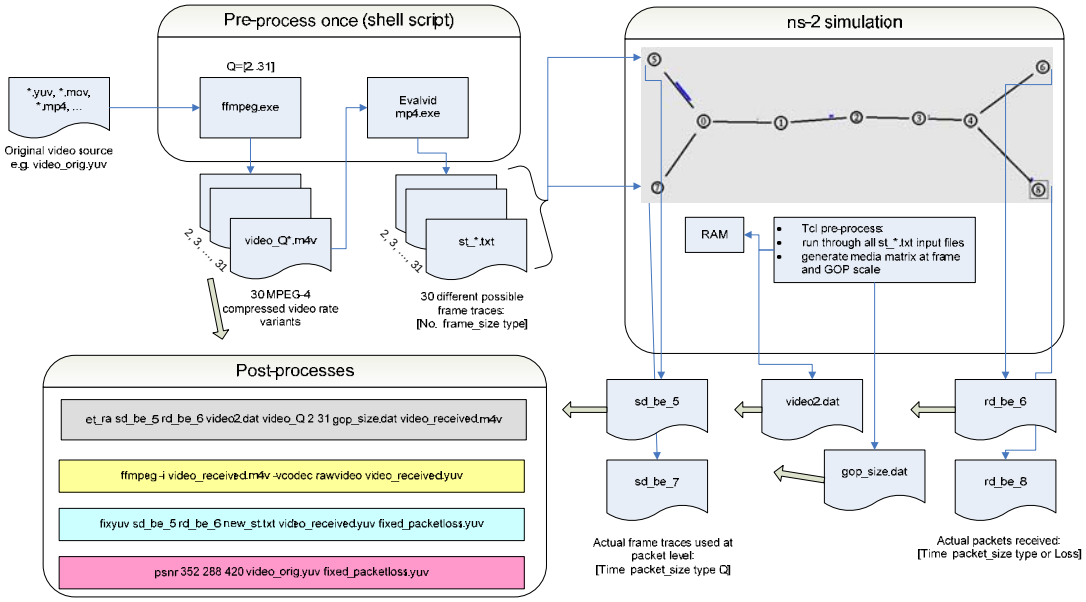
Figure 20 : Le produit *GQ*

- rate x Q product almost independent on Q
  - dependent only on content complexity
- ➔ at start of any new GOP, assuming complexity change smoothly from GOP to GOP
  - Q(k+1) ~ R(k)*Q(k)/R(k+1)
  - for stored media, next GOP complexity is known a priori

What is new in "Evalvid+"

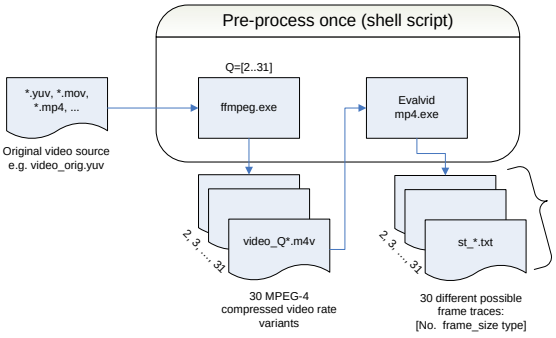
- Multiple trace files
  - one per Q-value
  - Q=[2,3,4,…,31] (ffmpeg)
- make SVBR calculate Q(k+1): select GOP(k+1) trace
- This requires fixed GOP sizes!
- LB(r,b) parameters change at feedback event
  - but the new Q-value is not used before start of next GOP
- Received video file must be assembled
  - using trace of actual Q(*i*)-values used, and
  - multiple *.m4v files

Solution: modified Evalvid tools "Evalvid+"



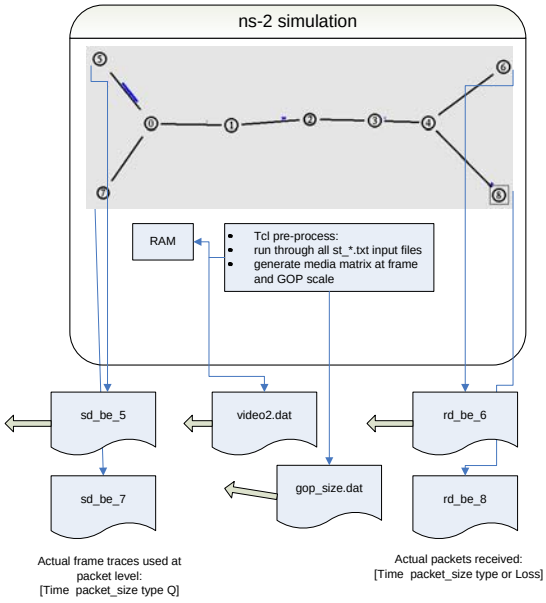
The tool-set overview Pre-process

- **ffmpeg** -s cif -r 30 -vframes 1836 -i video.yuv -vcodec mpeg4 -4mv -g 12 -sgop -sc_threshold 20000 -qscale 12 -s cif -r 30 -y video_Q12.m4v
- **mp4.exe** -send <IP address> <port #> <MTU> <fps> video_Q12.m4v > st_video_Q12.txt



The tool-set overview (cont.)
Network Simulation

- **ns-2**: evalvid_rateadapt.tcl
  - modified ns-2 interface to "Evalvid+" & adaptive SVBR responding to P-AQM feedback
  - Tcl init-function makes
    - video2.dat (frames all Q)
      - read into memory
      - used by all nodes sending the same media (different timing)
    - gop_size.dat (GOP size all Q)
      - used by et_ra.exe
  - sd_be_* stores e.g. actual Q used



The multi-rate trace file (multi Q)

- video2.dat:

<time μs> <bytes Q=2> <type> <MTU> <Q=3> <Q=4>… <Q=31>

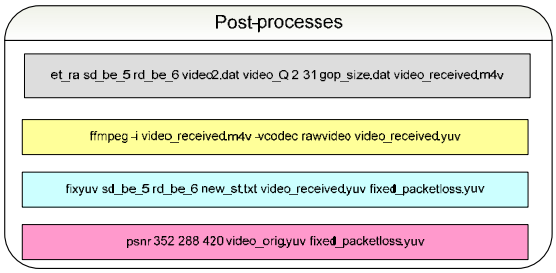
TextPad - [C:\cygwin\home\alie\Evalvid_in_NS2\Evalvid_in_NS2_May05\example\Aha\video2.dat]

File Edit Search View Tools Macros Configure Window Help

```
40000 62 0 1000 62 62 62 62 62 62 62 62 62 62 62 62 62 62 62 62 62 62 62 62 62 62 62 62 62 62 62 62 62
0 2239 1 1000 1442 1317 1263 1232 1205 1189 1172 1152 1133 1121 1124 1123 1098 1098 1098 1098 1119 1119 1115 1114 1100
40000 1556 2 1000 530 262 250 510 124 93 203 380 91 234 308 73 82 81 636 66 84 61 119 63 57 95 57 131 75 57 119 130 85
40000 1550 2 1000 615 356 251 669 192 118 165 422 153 205 341 57 92 66 786 61 127 64 99 58 58 143 96 120 69 57 99 113
40000 1501 2 1000 467 230 132 540 107 94 154 401 85 178 239 57 74 57 726 57 137 57 85 57 57 123 61 118 57 57 86 101 89
40000 1551 2 1000 437 213 143 597 85 62 131 413 76 233 222 57 66 92 733 66 66 57 85 57 57 111 57 105 64 57 150 147 896
40000 1561 2 1000 425 218 152 627 117 87 176 417 73 194 205 57 57 70 834 66 98 58 85 57 57 154 76 104 57 57 118 130 97
40000 1452 2 1000 409 202 161 548 113 80 204 351 117 192 202 57 79 90 757 61 110 61 86 57 57 145 70 132 65 57 160 147
40000 1579 2 1000 411 253 199 614 83 65 240 372 100 289 193 57 61 140 791 70 77 57 74 57 57 172 57 106 61 57 197 206 8
40000 1629 2 1000 448 237 153 650 133 91 199 301 96 304 163 62 66 139 841 70 86 57 65 57 57 175 77 80 57 57 213 215 97
40000 1570 2 1000 407 235 120 708 100 63 183 337 91 245 160 57 65 111 845 57 82 57 65 61 57 172 57 93 57 57 131 183 97
40000 1552 2 1000 366 187 111 553 65 67 150 299 93 238 148 57 70 139 706 57 78 57 69 64 57 126 57 122 57 57 244 174 85
40000 1566 2 1000 375 185 111 581 92 65 117 314 101 238 208 57 68 112 813 57 87 57 79 57 57 122 57 98 57 57 169 166 98
40000 62 0 1000 62 62 62 62 62 62 62 62 62 62 62 62 62 62 62 62 62 62 62 62 62 62 62 62 62 62 62 62 62
0 2312 1 1000 1457 1328 1259 1230 1198 1185 1162 1132 1111 1110 1123 1123 1099 1099 1098 1098 1116 1116 1111 1110 1104
40000 1661 2 1000 613 322 245 591 137 88 257 302 122 345 263 73 95 139 809 75 76 67 66 60 61 181 57 121 57 57 150 214
40000 1558 2 1000 513 280 203 631 131 77 220 297 88 350 176 57 96 105 854 61 91 76 92 58 74 174 74 98 57 57 152 219 96
40000 1589 2 1000 414 202 120 619 76 66 143 243 111 295 175 57 80 87 888 57 111 58 58 57 57 120 61 130 57 57 139 206 9
40000 1510 2 1000 388 197 125 472 74 66 145 276 105 271 143 57 72 112 702 61 86 57 73 57 57 111 57 115 57 57 212 173 8
40000 1538 2 1000 310 148 126 487 86 70 158 330 94 236 128 57 78 70 795 61 79 57 61 65 61 145 61 101 57 57 207 156 867
40000 1499 2 1000 386 181 129 646 80 57 138 336 119 272 173 57 61 78 820 61 114 57 57 57 57 118 57 114 57 57 102 169 9
40000 1533 2 1000 407 197 148 533 88 63 152 306 113 241 180 57 102 123 740 61 82 62 74 57 57 120 57 152 57 57 225 166
40000 1548 2 1000 430 195 127 648 71 57 131 354 90 253 203 57 70 128 804 57 83 57 65 57 61 125 57 85 57 57 172 197 985
40000 1529 2 1000 310 104 96 567 83 59 119 364 106 151 207 57 74 61 848 57 97 57 57 57 57 108 57 113 57 57 112 102 972
40000 1647 2 1000 450 217 191 561 67 71 200 282 128 249 157 66 107 114 727 61 81 66 113 65 57 142 61 159 74 57 216 195
40000 1705 2 1000 364 155 159 537 78 63 113 243 111 234 151 57 66 106 781 61 70 57 77 65 57 106 57 100 61 57 221 175 9
40000 62 0 1000 62 62 62 62 62 62 62 62 62 62 62 62 62 62 62 62 62 62 62 62 62 62 62 62 62 62 62 62 62
```

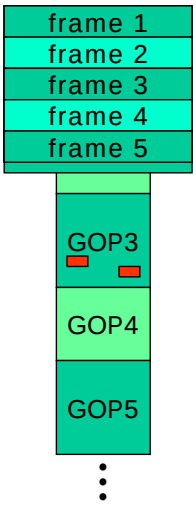
The tool-set overview (cont.)
Post-process

- **et_ra.exe**
  - modified Evalvid original et.exe
  - Reads packet Tx and Rx trace files
  - Finds used Q
  - Reads video2.dat for frame sizes and types
  - Reads gop_size.dat to assist assembling the resulting MPEG-4 file
- **ffmpeg** to decode to YUV
- **fixyuv.exe:** if packet losses
- **psnr.exe:** compare decoded YUV to original

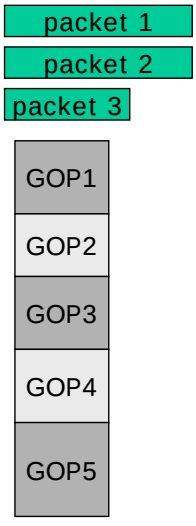


et_ra.exe (Evaluate Trace, rate Adaptive)

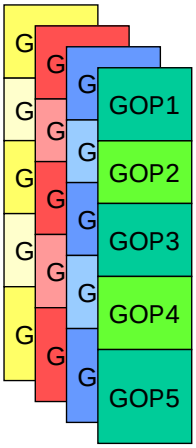
- original et.exe:



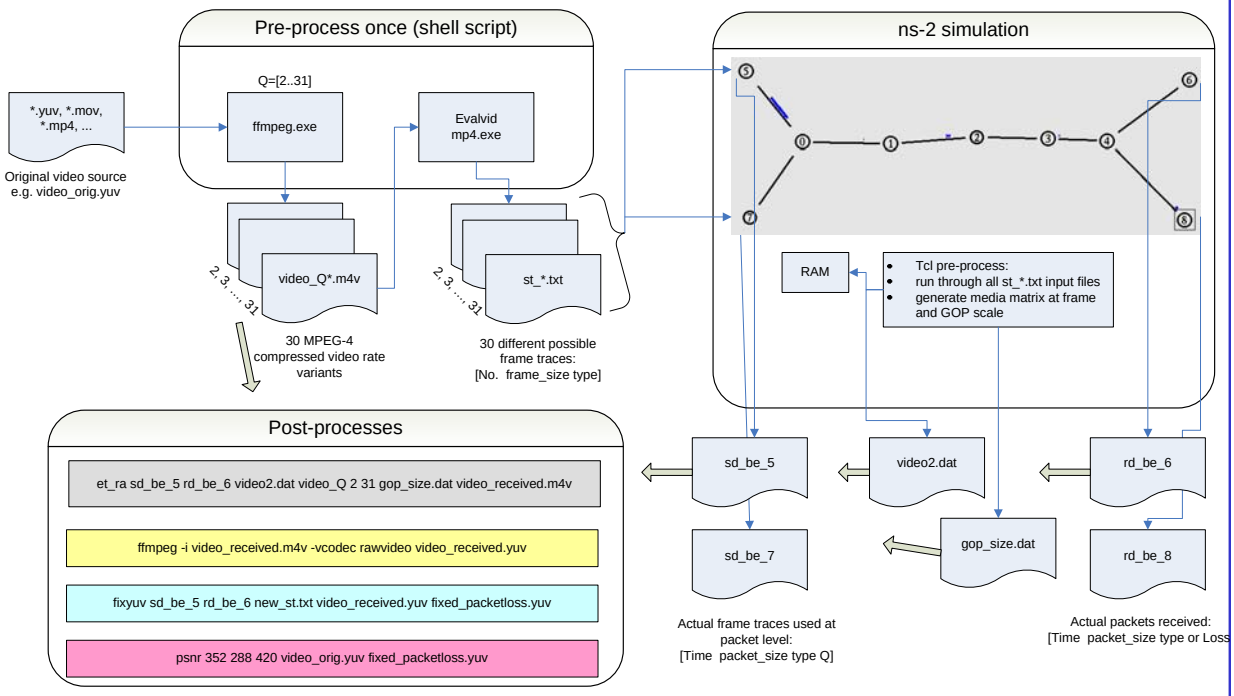
frame 1:



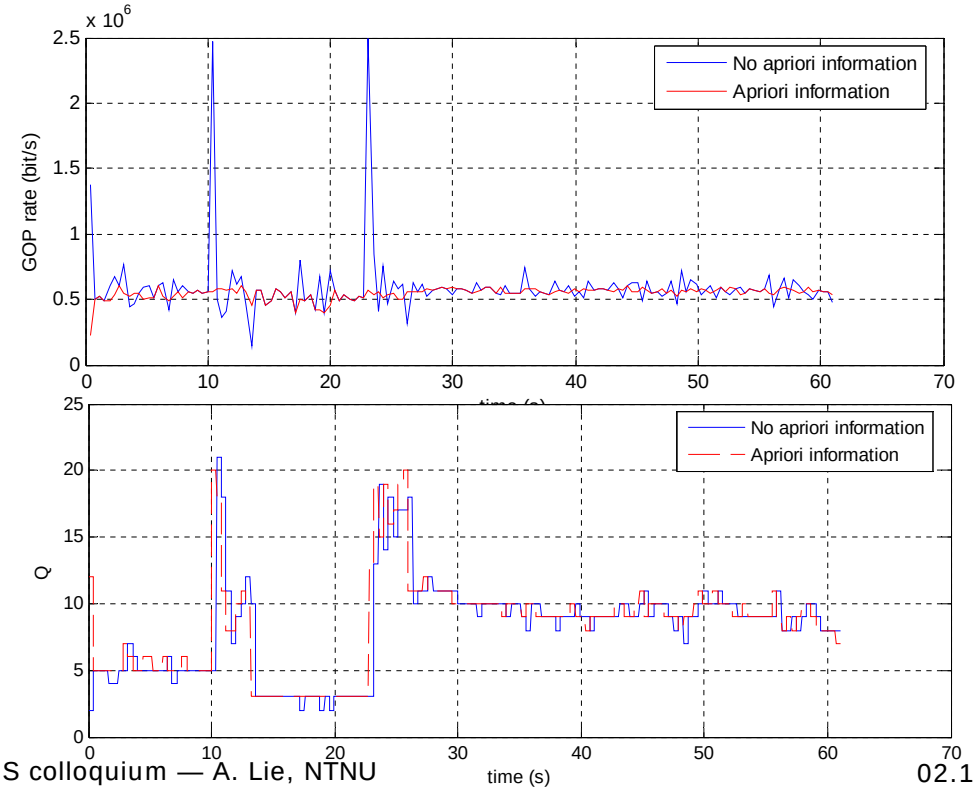
- et_ra.exe:



“Evalvid+”



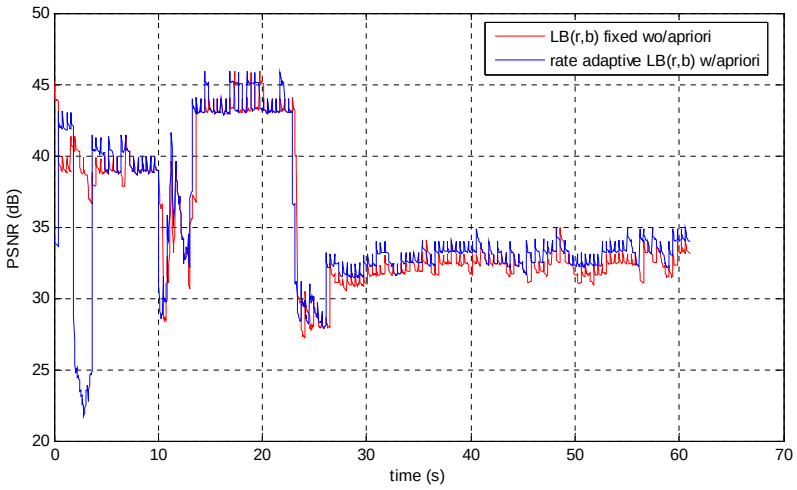
LB(r,b)=LB(600kbit/s,3.0*GOP-rate)



Simulation of 42 SVBR sources

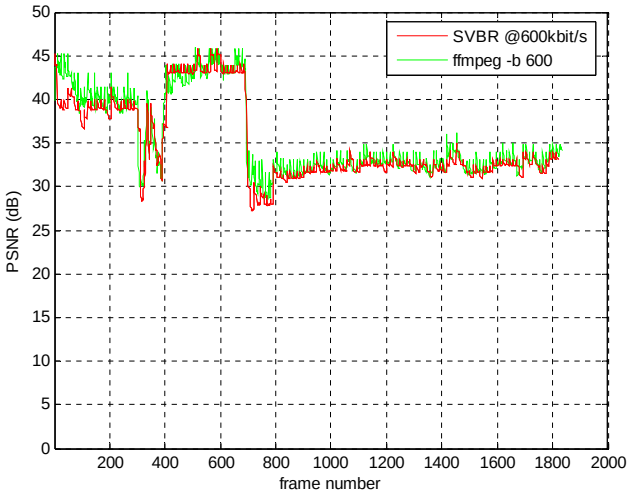
- over 25Mbit/s ➔ 25/42~600kbit/s
- All sources started within first 2s, at uniform dist. random time, and at random position within the same video sequence
  - using P-AQM nodes, utilization~95%, packet drop~0.5%
  - result
  - compare visually to static LB(r,b)

PSNR calculation



- PSNR original to static LB(r,b)
- PSNR original to rate adaptive LB(r,b)

PSNR comparison
SVBR vs. ffmpeg rate controller

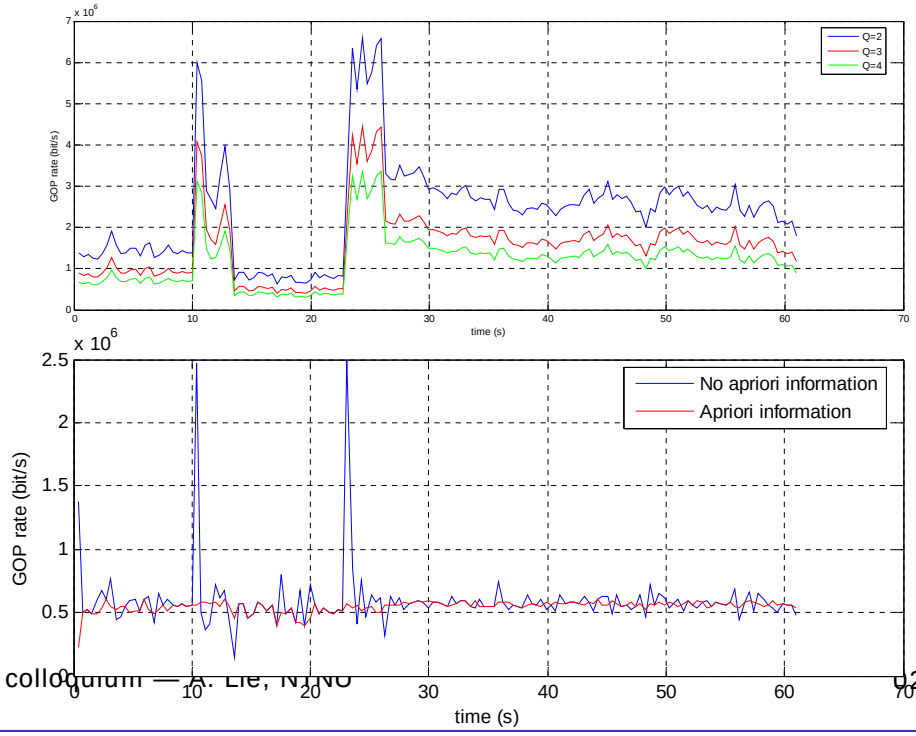


- PSNR original to static LB(r,b)
- PSNR using ffmpeg rate controller at same target bit rate

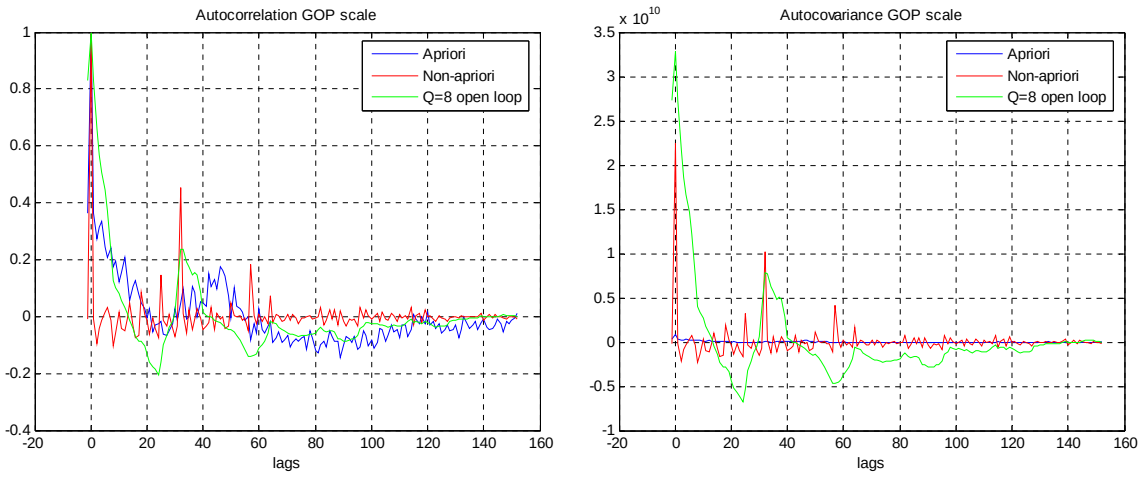
Long Range Dependence (LRD)

- Garrett & Willinger ’94: VBR video traffic is self-similar
  - the autocorrelation $\rho(k)$ function decays slowly at increasing lag
  - makes buffer dimensioning & high link utilization very difficult
  - The cause of LRD: scene complexity changes!
- many papers on video characterization (GOP scale, frame scale)
  - very little related to what kind of *rate controller* in use!
- Hamdi showed in his thesis that
  - a stream satisfying a LB(r,b) constraint, where r equals the traffic average rate, is not self similar

rate controllers limits the rate variance…

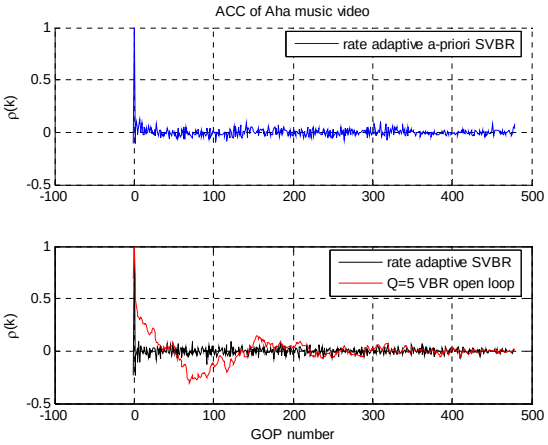


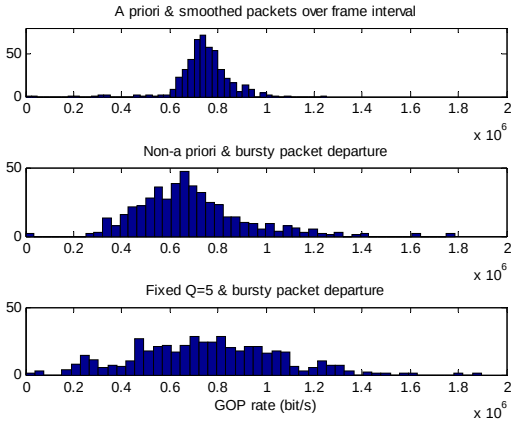
ρ(k) of "concatenated" video (GOP)



- Positive correlations at lag *k* poses long bursts of time duration *k*

Aha music video
(30 flows, ~750kbit/s fair share)

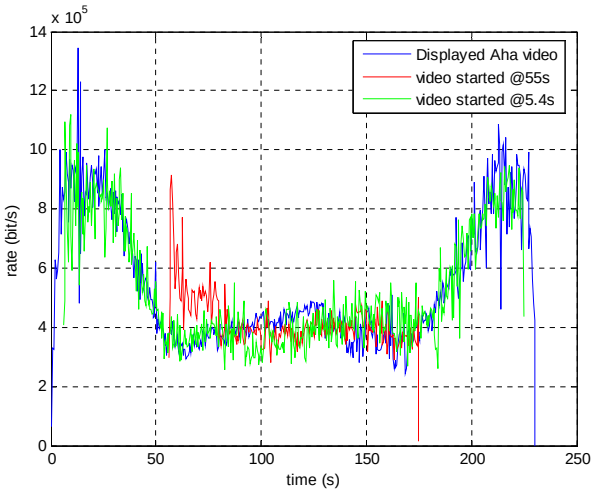


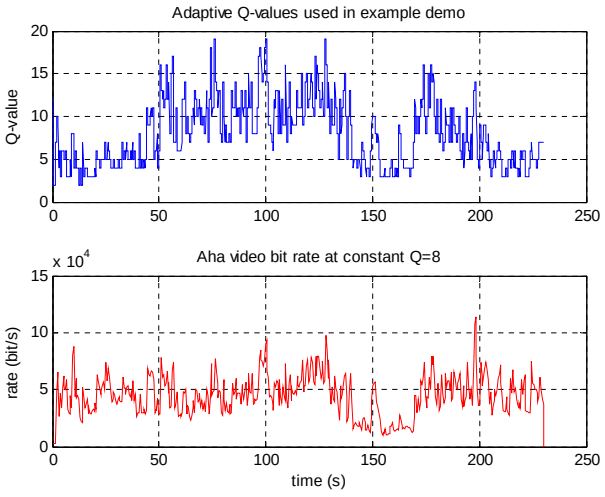


GOP size autocorrelation

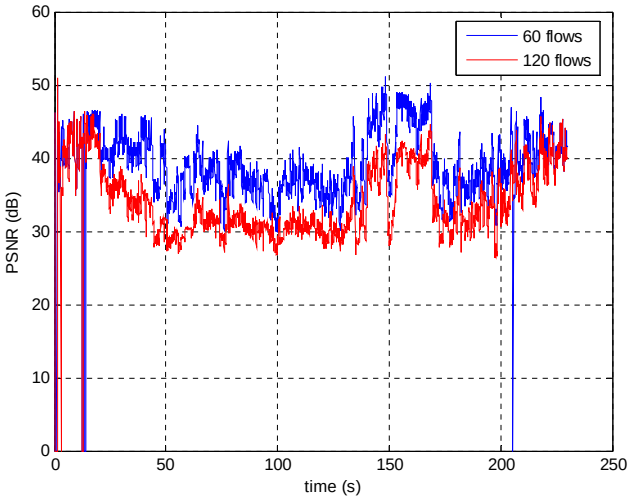
GOP size histograms

Aha video example
(dynamic case, 1-60 flows)





Aha example: PSNR values 60 and 120 flows



- Aha video, 120 flows

Limitations of this implementation

- GOP time scale rate adaptation
  - Hamdi confirms that SVBR could be modified to frame scale
- Fixed GOP size
  - live encoders could start a new GOP (i.e. next frame being I-frame) at a feedback event!
  - relaxation will make distortions?
- error concealment (packet loss)
  - FRAME mode vs. PACKET mode considerations
  - ffmpeg drops first frame after frame marked with “loss”
- No audio yet
  - limitation in mp4.exe tool

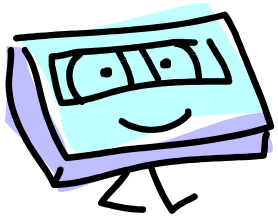
To-do

- Interface rate adaptive SVBR to TFRC/DCCP, …
- Publish
  - paper
- Polish
  - sw and upload

Usage of this tool-set

- Simulate many flows, coming from many sources, all of them rate adaptive
- Have different media sources, not only one
- Wireless rate adaptive multimedia
- Different congestion control algorithms
  - Self-limited sources and their actual bandwidth
  - friendliness (towards TCP, UDP, DCCP, etc.)
- Different queueing systems (FIFO, AQMs, QoS; DiffServ e.g.)
- Removal of LRD, or not?
- trade latency for loss (short queues)
- how to inject new flows
- new initiatives for rate adaptation incentive (Stian@Q2S)
- vary the sources rates
- …

Summary



- "Online" *rate adaptive* video codec at simulation time obtained by
  - fixed GOP size
  - pre-processed media compressed with static quantization values
  - "online" VBR adaptive rate control by "mixing" many constant quality encoded media
- Result
  - rate adaptive media traffic simulated (traffic, loss, latency)
  - received media can be re-assembled
  - visually played
  - PSNR calculated, values close to online rate controller