End-to-end Delay of Videoconferencing over Packet Switched Networks

Mario Baldi

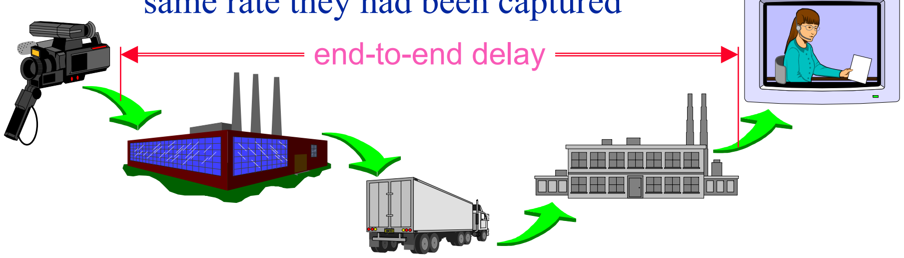
Technical University of Torino
Computer Engineering Department

(Joint work with Dr Yoram Ofek)

Nota di Copyright

Videoconferencing Requirements

- Bound on end-to-end delay
  - 100 ms
- Synchronization
  - the receiver continuously shows pictures at the same rate they had been captured



Goals

Identify components of the end-to-end delay

Find out which configurations of the videoconferencing system allow the end-to-end delay to be kept below the 100 ms bound

Components of End-to-end Delay

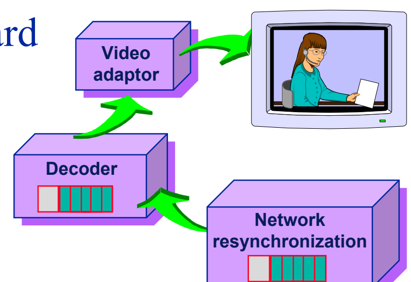
- *Processing* delay
  - e.g., encoding
- *Network* delay
  - e.g., shaping, propagation, queueing
- Resynchronization
  - *processing resynchronization* delay
    - e.g., from constant bit rate to constant frame rate
  - *network resynchronization* delay
    - e.g., jitter compensation

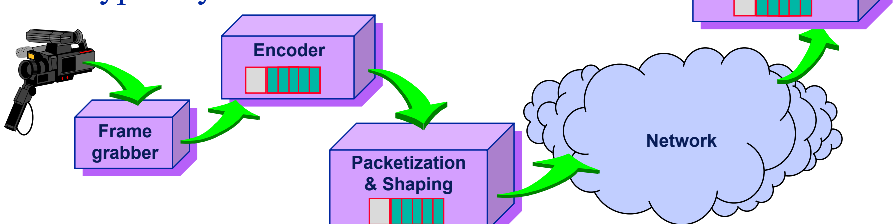
Configurations

	Dedicated link	Circuit switching	Statistical multiplexing	Time driven priority
Raw video	$\frac{F_r}{C} + P + P_d$ 1	$\frac{F_r}{B} + Sw + P + P_d$ 2	$\frac{F_r}{C} + P + Q_M + E_r + P_d$ 4 $S_n + \frac{P_s}{C} + P + Q_M + E_r + P_d$	$L \cdot T_f + P_d$ 3
CBR MPEG	$S_c + P + D + P_d$	$S_c + Sw + P + D + P_d$ 5	$S_c + P + \frac{P_s}{C} +$ 6 $+ Q_M + E_r + D + P_d$	$S_c + L \cdot T_f + D + P_d$
VBR MPEG	$C_M + P + D + P_d$		$C_M + S_n + P + \frac{P_s}{C} +$ 7 $+ Q_M + E_r + D + P_d$	$C_M + L \cdot T_f + D + P_d$ 8

System model

- Frame grabber/ Capture card
  - small delay (3 ms)
- Video adaptor
  - *presentation delay*
  - typically less than 17 ms





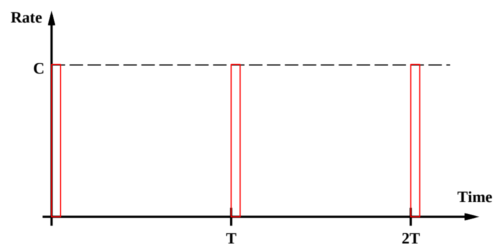
Road Map

	Dedicated link	Circuit switching	Statistical multiplexing	Time driven priority
Raw video	1	$\frac{F_r}{B} + Sw + P + P_d$ 2	$\frac{F_r}{C} + P + Q_M + E_r + P_d$ 4 $S_n + \frac{P_s}{C} + P + Q_M + E_r + P_d$	$L \cdot T_f + P_d$ 3
CBR MPEG	$S_c + P + D + P_d$	$S_c + Sw + P + D + P_d$ 5	$S_c + P + \frac{P_s}{C} +$ 6 $+ Q_M + E_r + D + P_d$	$S_c + L \cdot T_f + D + P_d$
VBR MPEG	$C_M + P + D + P_d$		$C_M + S_n + P + \frac{P_s}{C} +$ 7 $+ Q_M + E_r + D + P_d$	$C_M + L \cdot T_f + D + P_d$ 8

Dedicated link

$$\Delta_{raw}^{ded} = \frac{F_r}{C} + P + P_d$$

- P propagation delay
- C link capacity
- F_r picture dimension
- P_d presentation delay
  - synchronize adaptor and capture card



Dedicated link

- For example C = 100 Mb/s
  - QCIF: F_r = 176x144 = 198 kb → $\frac{F_r}{C} = 1.98$ ms
  - HDTV: F_r = 1920x1080 = 16200 kb → $\frac{F_r}{C} = 162$ ms
- For real-time video ⇒ $\frac{F_r}{C} \leq T$
  - HDTV (30 fps) C > 486 Mb/s → need for compression
- ☹ Short delay ⇒ large capacity ⇒ low utilization
  - QCIF example: 3%

Road Map

	Dedicated link	Circuit switching	Statistical multiplexing	Time driven priority
Raw video	$\frac{F_r}{C}+P+P_d$	2	$\frac{F_r}{C}+P+Q_M+E_r+P_d$ 4 $S_n+\frac{P_s}{C}+P+Q_M+E_r+P_d$	$L \cdot T_f + P_d$ 3
CBR MPEG	$S_c+P+D+P_d$	$S_c+Sw+P+D+P_d$ 5	$S_c+P+\frac{P_s}{C}+$ 6 $+Q_M+E_r+D+P_d$	$S_c+L \cdot T_f+D+P_d$
VBR MPEG	$C_M+P+D+P_d$		$C_M+S_n+P+\frac{P_s}{C}+$ 7 $+Q_M+E_r+D+P_d$	$C_M+L \cdot T_f+D+P_d$ 8

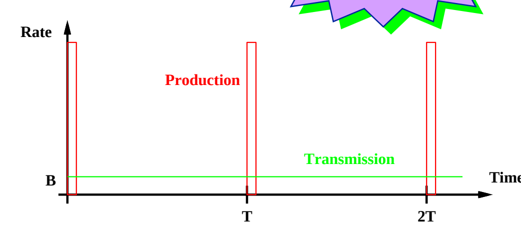
Circuit Switching

$$\Delta_{raw}^{CS} = \frac{F_r}{B} + Sw + P + P_d$$

- P propagation delay
- Sw switching delay
- B circuit bandwidth
- F_r picture dimension

See it

$$\frac{F_r}{B} = T$$

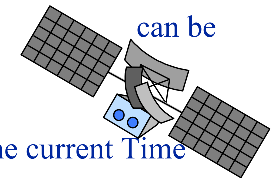


Road Map

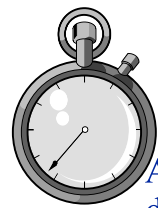
	Dedicated link	Circuit switching	Statistical multiplexing	Time driven priority
Raw video	$\frac{F_r}{C} + P + P_d$ N	$\frac{F_r}{B} + Sw + P + P_d$ N	$\frac{F_r}{C} + P + Q_M + E_r + P_d$ 4 $S_n + \frac{P_s}{C} + P + Q_M + E_r + P_d$	3
CBR MPEG	$S_c + P + D + P_d$	$S_c + Sw + P + D + P_d$ 5	$S_c + P + \frac{P_s}{C} + Q_M + E_r + D + P_d$ 6	$S_c + L \cdot T_f + D + P_d$
VBR MPEG	$C_M + P + D + P_d$		$C_M + S_n + P + \frac{P_s}{C} + Q_M + E_r + D + P_d$ 7	$C_M + L \cdot T_f + D + P_d$ 8

Time Driven Priority

- Nodes share a global timing reference
  - external reference (e.g., GPS) can be used



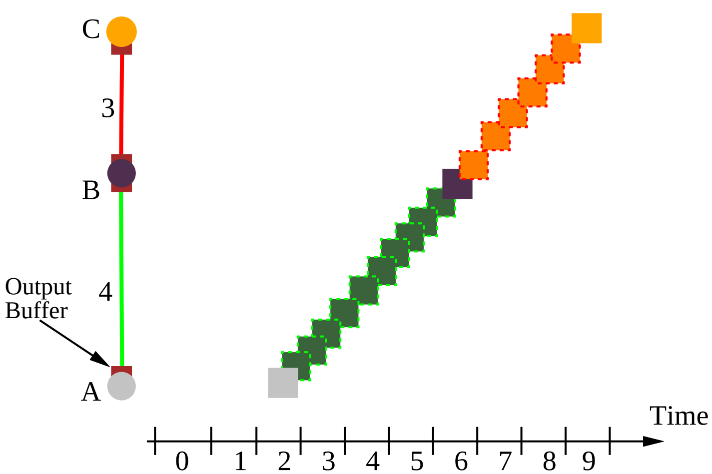
- Time is divided in *Time Frames*
  - each node has the same notion of the current Time Frame
    - beginning and end
  - typical duration $T_f = 125$ µs



- A fixed amount of bits $T_f \cdot C$ can be sent on a link during a Time Frame

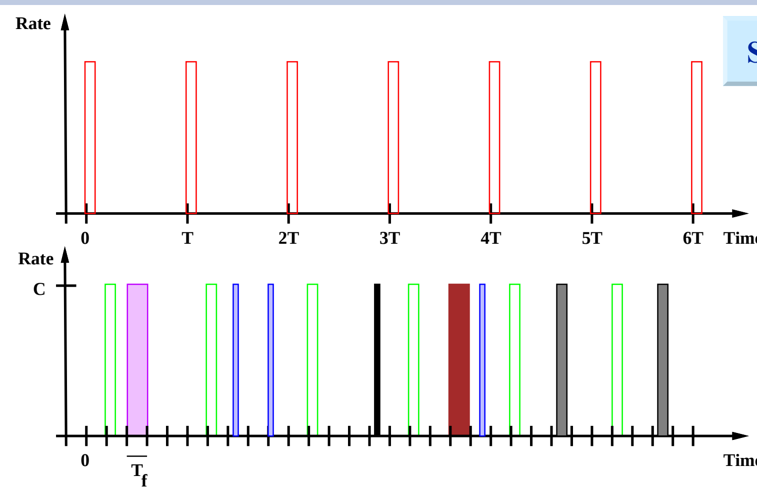
See it

RISC-like forwarding of packets



Traffic Multiplexing

See it



End-to-end Delay

$$\Delta^{TDP}_{raw} = L \cdot T_f + P_d$$

N

- ➔ L depends on number of hops
- ➔ Network jitter $2 \cdot T_f$
  - ➔ no need for resynchronization
- ➔ P_d presentation delay

NR

smaller delay than circuit switching

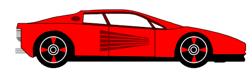
Comparison with Dedicated Link

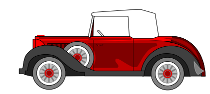
System parameters

- Capacity C = 100 Mb/s
- $L = 3$, $P = 0$, $Sw = 0$
- QCIF at 15 fps

$\Delta^{ded}_{raw} = 1.98$ ms $\Delta^{TDP}_{raw} = 2.175$ ms $\Delta^{CS}_{raw} = 66.67$ ms

97 % of dedicated link capacity unused





Road Map

	Dedicated link	Circuit switching	Statistical multiplexing	Time driven priority
Raw video	$\frac{F_r}{C} + P + P_d$ N	$\frac{F_r}{B} + Sw + P + P_d$ N	4	$L \cdot T_f + P_d$ N
CBR MPEG	$S_c + P + D + P_d$	$S_c + Sw + P + D + P_d$ 5	$S_c + P + \frac{P_s}{C} + Q_M + E_r + D + P_d$ 6	$S_c + L \cdot T_f + D + P_d$
VBR MPEG	$C_M + P + D + P_d$		$C_M + S_n + P + \frac{P_s}{C} + Q_M + E_r + D + P_d$ 7	$C_M + L \cdot T_f + D + P_d$ 8

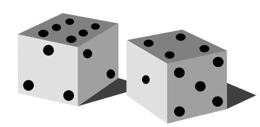
Network Delay

N

- Fixed transmission and propagation delay
- Variable queueing delay
  - queueing policies
  - network load

NR

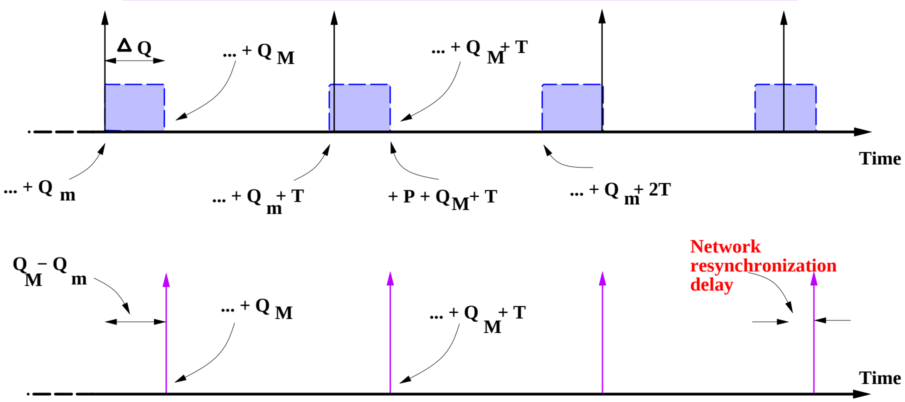
Non deterministic behavior



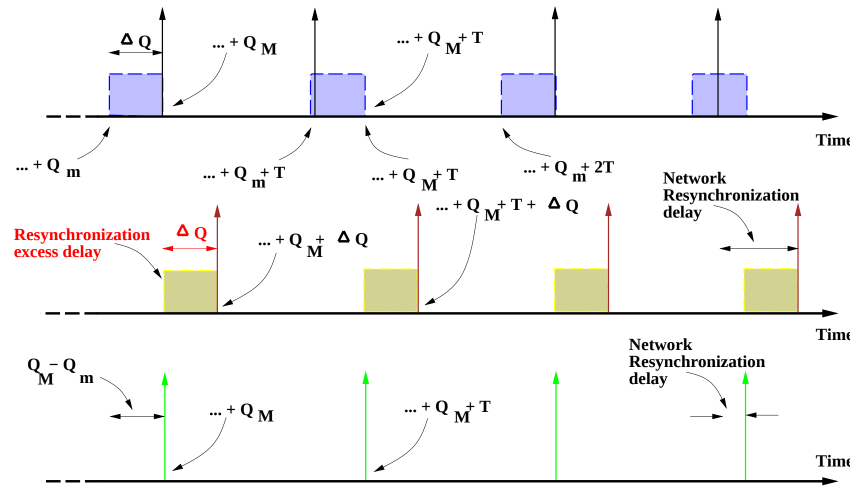
Network delay is not bound deterministically

Network Resynchronization

Use a **guessed bound** Q_M on network delay



Resynchronization Excess Delay



End-to-end Delay

$$\Delta_{raw}^{bursty} = \frac{F_r}{C} + P + Q_M + E_r + P_d$$

- $E_r \in [0, \Delta Q]$ resynchronization excess delay
  - constant during the videoconference call
- $\Delta Q = Q_M - Q_m$ maximum jitter
- Q_M (**guess on**) maximum queueing delay
- Q_m minimum queuing delay
- P propagation delay
- C capacity of links

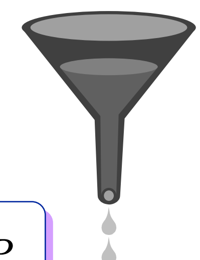
→ F_r raw picture dimension

→ P_d presentation delay

See it

Traffic Shaping

- For example, *leaky bucket*
  - token generation rate B
  - token bucket size A



$$\Delta_{raw}^{TS} = S_n + \frac{P_s}{C} + P + Q_M + E_r + P_d$$

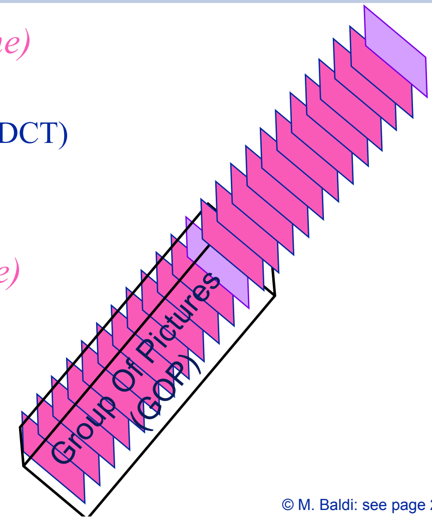
- *network shaping delay* $S_n = \frac{F_r - A}{B}$
- P_s packet size

Road Map

	Dedicated link	Circuit switching	Statistical multiplexing	Time driven priority
Raw video	$\frac{F_r}{C} + P + P_d$ N	$\frac{F_r}{B} + Sw + P + P_d$ N	$\frac{F_r}{C} +$ NR $E_r + P_d$ $S_n + \frac{P_s}{C}$ N $E_r + P_d$	$L \cdot T_f + P_d$ N
CBR MPEG	$S_c + P + D + P_d$	$S_c + Sw + P + D + P_d$ 5	$S_c + P + \frac{P_s}{C} +$ 6 $+ Q_M + E_r + D + P_d$	$S_c + L \cdot T_f + D + P_d$
VBR MPEG	$C_M + P + D + P_d$		$C_M + S_n + P + \frac{P_s}{C} +$ 7 $+ Q_M + E_r + D + P_d$	$C_M + L \cdot T_f + D + P_d$ 8

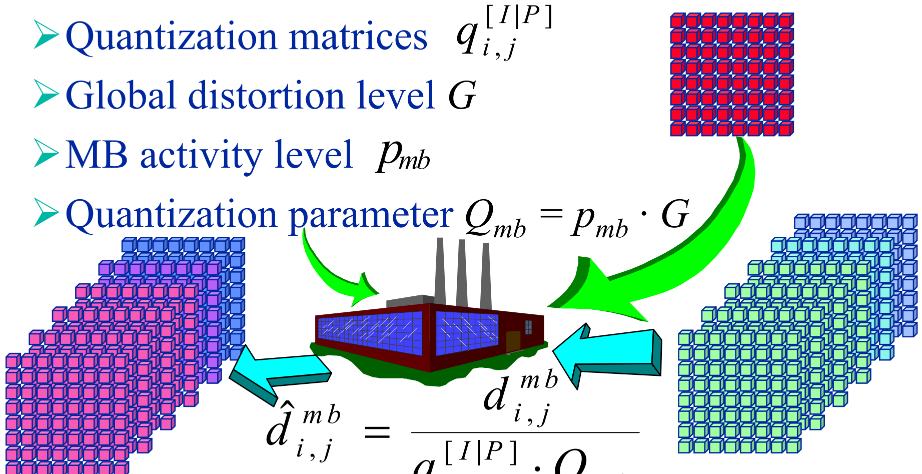
MPEG Compression Standard

- *Intra-frame coding (I-Frame)*
  - 8x8 blocks
  - Discrete Cosine Transform (DCT)
  - Quantization
  - Encoding
- *Predictive coding (P-Frame)*
  - MacroBlock (MB)
  - motion estimation
  - motion compensation



Quantization

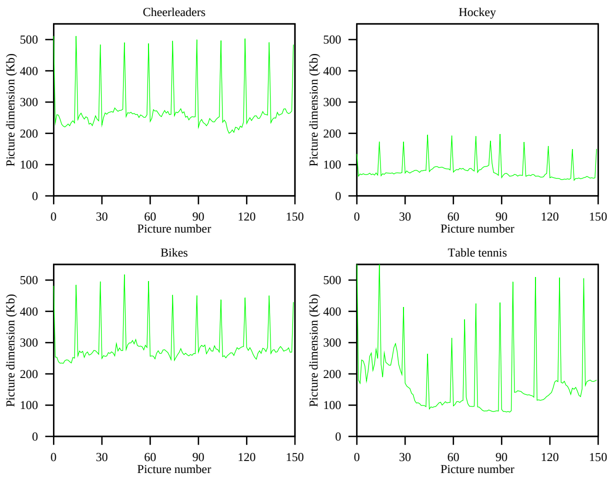
- Quantization matrices $q_{i,j}^{[I|P]}$
- Global distortion level G
- MB activity level p_{mb}
- Quantization parameter $Q_{mb} = p_{mb} \cdot G$



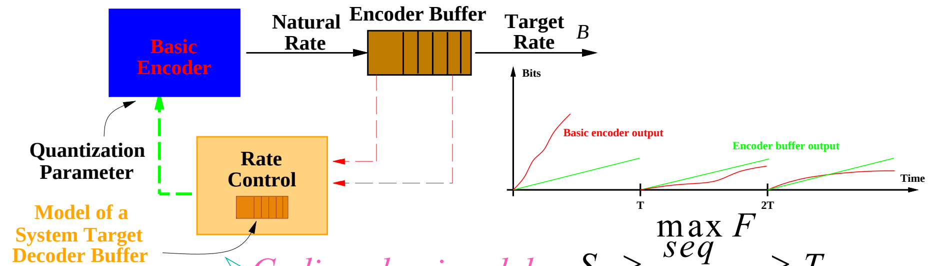
$$\hat{d}_{i,j}^{mb} = \frac{d_{i,j}^{mb}}{q_{i,j}^{[I|P]} \cdot Q_{mb}}$$

Cheerleaders Scene

Natural MPEG Bit Rate



CBR MPEG Encoder

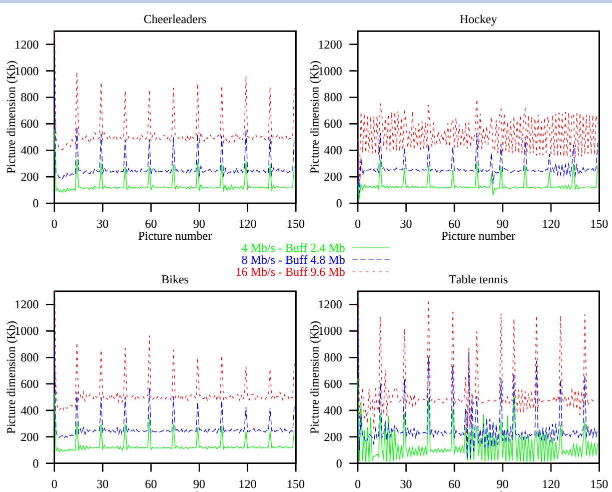


- *Coding shaping delay* $S_c \ge \frac{\max_{seq} F}{B} \ge T$

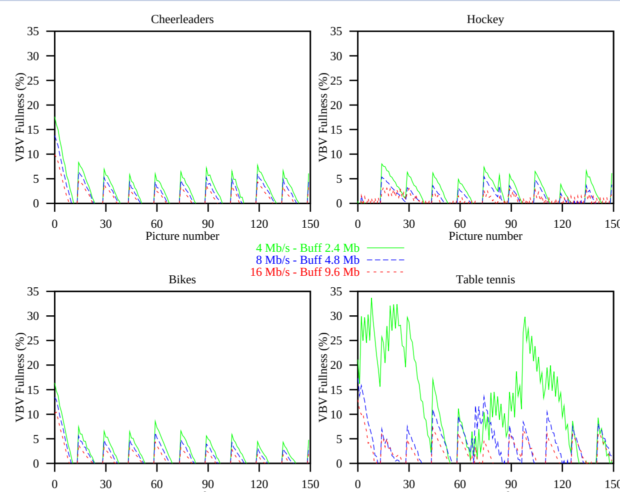
- variable delay introduced by encoder buffer
- processing resynchronization delay introduced by the decoder

The delay is larger than the video frame period

Dimension of Pictures



Video Buffer Verifier Fullness



Video Buffer Verifier and Picture Quality

- V_s Video Buffer Verifier (VBV) size determines
  - variability of picture dimension

$$\max_{seq} F \le V_s \qquad \min_{seq} F \ge 2 \cdot B \cdot T - V_s$$

  - visual quality of encoded video

High and uniform quality ⇒ large VBV

Up to ***GOP size*** for static scenes

Video Buffer Verifier and Delay

$$S_c \ge \frac{\max_{seq} F}{B}$$

- $\max_{seq} F$ is not known when starting encoding
- dimension the system using an upper bound (V_s)

$$S_c \ge \frac{V_s}{B}$$

High picture quality ⇒ large delay

Up to *GOP period* for static scenes

Road Map

	Dedicated link	Circuit switching	Statistical multiplexing	Time driven priority
Raw video	$\frac{F_r}{C} + P + P_d$ N	$\frac{F_r}{B} + Sw + P + P_d$ N	$\frac{F_r}{C} +$ NR $E_r + P_d$; $S_n + \frac{P_s}{C}$ N $E_r + P_d$	$L \cdot T_f + P_d$ N
CBR MPEG	$S_c + P + D + P_d$	5	$S_c + P + \frac{P_s}{C} +$ 6 $+ Q_M + E_r + D + P_d$	$S_c + L \cdot T_f + D + P_d$
VBR MPEG	$C_M + P + D + P_d$		$C_M + S_n + P + \frac{P_s}{C} +$ 7 $+ Q_M + E_r + D + P_d$	$C_M + L \cdot T_f + D + P_d$ 8

Circuit Switching

$$\Delta^{CS}_{CBR} = S_c + Sw + P + D + P_d$$

- S_c coding shaping delay
- D *decoding delay*
- Sw switching delay
- P propagation delay
- P_d presentation delay

Road Map

	Dedicated link	Circuit switching	Statistical multiplexing	Time driven priority
Raw video	$\frac{F_r}{C}+P+P_d$ N	$\frac{F_r}{B}+Sw+P+P_d$ N	$\frac{F_r}{C}+$ NR E_r+P_d $S_n+\frac{P_s}{C}$ N E_r+P_d	$L\cdot T_f+P_d$ N
CBR MPEG	$S_c+P+D+P_d$	S_c PR $P+D+P_d$ P N	6	$S_c+L\cdot T_f+D+P_d$
VBR MPEG	$C_M+P+D+P_d$		$C_M+S_n+P+\frac{P_s}{C}+$ 7 $+Q_M+E_r+D+P_d$	$C_M+L\cdot T_f+D+P_d$ 8

Packet Switching with Statistical Multiplexing

$$\Delta_{CBR}^{SM}=S_c+\frac{P_s}{C}+P+Q_M+E_r+D+P_d$$

- $E_r \in [0, \Delta Q]$ resynchronization excess delay
- $\Delta Q = Q_M - Q_m$ maximum jitter
- Q_M (**guess on**) maximum queueing delay
- Q_m minimum queueing delay
- P propagation delay
- P_s packet size
- C link capacity

NR N

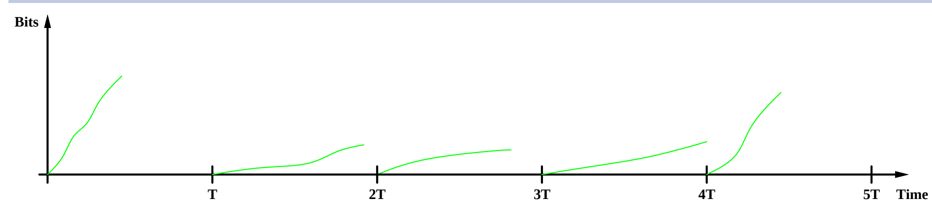
- S_c coding shaping delay
- D decoding delay
- P_d presentation delay

PR P

Road Map

	Dedicated link	Circuit switching	Statistical multiplexing	Time driven priority
Raw video	$\frac{F_r}{C}+P+P_d$ N	$\frac{F_r}{B}+Sw+P+P_d$ N	$\frac{F_r}{C}+$ NR E_r+P_d $S_n+\frac{P_s}{C}$ N E_r+P_d	$L\cdot T_f+P_d$ N
CBR MPEG	$S_c+P+D+P_d$	S_c PR $+D+P_d$ P N	PR NR $+$ P N P_d	$S_c+L\cdot T_f+D+P_d$
VBR MPEG	$C_M+P+D+P_d$		7	$C_M+L\cdot T_f+D+P_d$ 8

VBR MPEG Encoding



- C_M *maximum coding delay*
  - the decoder buffer compensates variations of coding delay
    - processing resynchronization delay

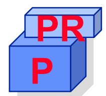
PR P

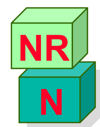
$C_M \leq T$

Packet Switching with Statistical Multiplexing

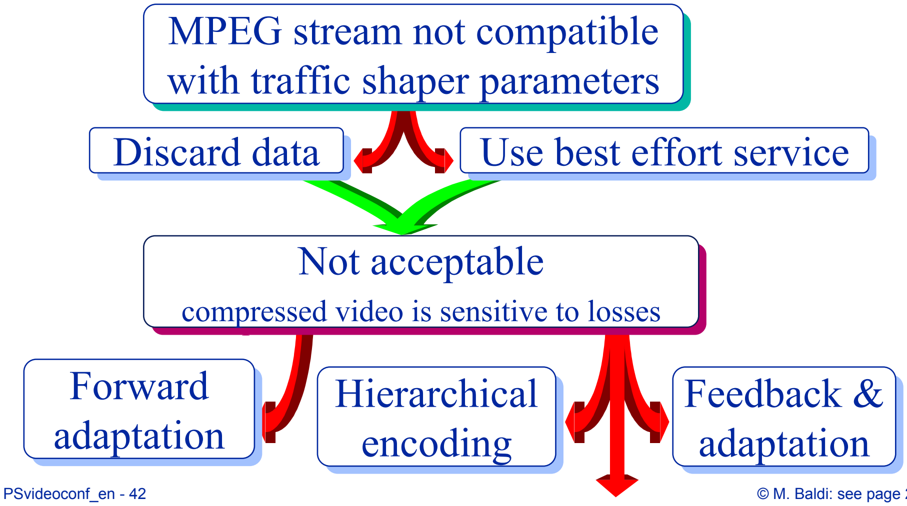
$$\Delta^{TS}_{VBR} = C_M + S_n + \frac{P_s}{C} + P + Q_M + E_r + D + P_d$$

- C_M maximum coding delay
- S_n network shaping delay
- Q_M (**guess on**) maximum queueing delay
- $E_r \in [0, \Delta Q]$ resynchronization excess delay
- P_s packet size





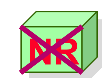
Problems with VBR MPEG and Statistical Multiplexing



Road Map

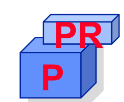
	Dedicated link	Circuit switching	Statistical multiplexing	Time driven priority
Raw video	$\frac{F_r}{C} + P + P_d$ N	$\frac{F_r}{B} + Sw + P + P_d$ N	$\frac{F_r}{C}$ + … $E_r + P_d$; $S_n + \frac{P_s}{C}$ … $E_r + P_d$ NR N	$L \cdot T_f + P_d$ N
CBR MPEG	$S_c + P + D + P_d$	S_c … $+ D + P_d$ PR P N	… + … P_d PR NR P N	$S_c + L \cdot T_f + D + P_d$
VBR MPEG	$C_M + P + D + P_d$		$C_M + S$ … + … P_d PR NR P N	8

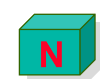
Packet Switching with Time Driven Priority



$$\Delta^{TDP}_{VBR} = C_M + L \cdot T_f + D + P_d$$

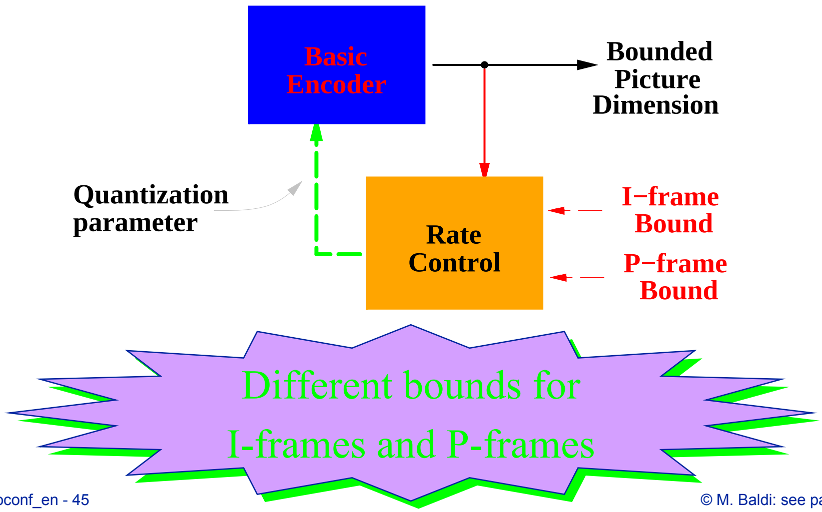
- C_M maximum coding delay
- L depends on number of hops
- P_d presentation delay





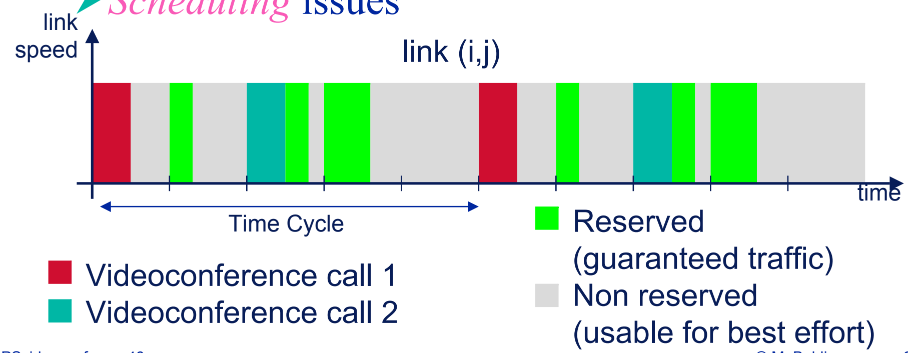
Picture dimension must be bound

Bounding Picture Dimension



Resource Allocation

- *Time cycle* equal to video frame period ***T***
- Reserve 1 time frame per time cycle
- *Scheduling* issues



Network Shaping Delay

$$S_n + L \cdot T_f + P_d \pm T_f$$

$$S_n = S_t \in [0, T]$$

- S_n *network shaping delay*
- $S_t = 0$ if the capture card is synchronized with network interface

$$C \ge \frac{F_r}{T}$$

  - QCIF C > 1.5 Gb/s
  - HDTV C > 130 Gb/s

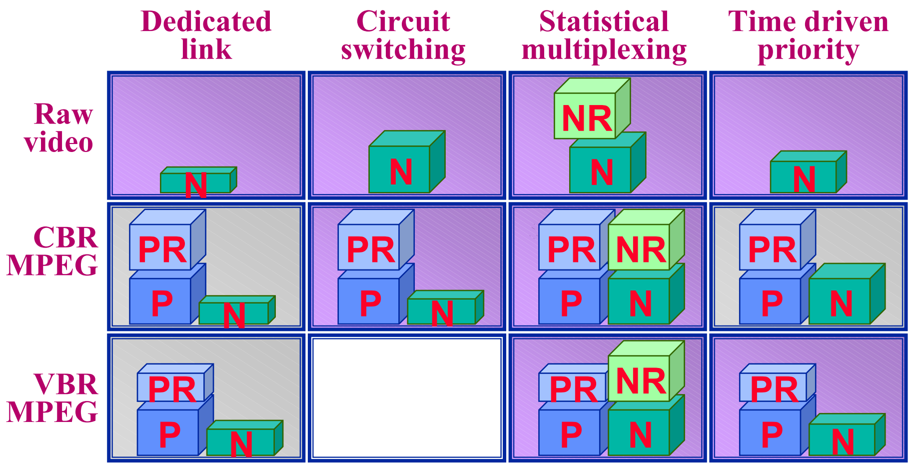
Network Shaping Delay

$$S_n = S_t + (N_r - 1)$$

$$N_r \ge \left\lceil \frac{F_r}{T_f \cdot C} \right\rceil$$

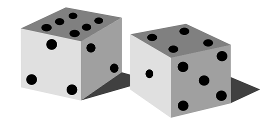
- N_r depends on scheduling
  - constant
  - fixed at reservation time

The Complete Picture



Conclusions

Statistical Multiplexing
- non deterministically bound delay
- large guessed bound



CBR MPEG Encoding
- long coding shaping delay
  up to GOP period

Conclusions

Time driven priority
- strict bound on jitter (250 μs)
- VBR MPEG encoder

The *end-to-end* delay can be *less* than a video frame period *T*