Multimedia over the Internet

François Fluckiger
CERN, Geneva

francois.fluckiger@cern.ch

Who am I?

François Fluckiger

CERN

Geneva, Switzerland

Birth place of the World-Wide Web

Head, Internet Infrastructure

Part-time Professor of Computer Science at the

University of Geneva

For further reading ...

- **Reference book:**

  "Understanding Networked Multimedia"

  Prentice Hall ***ISBN 0-13-190992-4***

  by F.Fluckiger

Where to find more in the Book?

See pointer at the bottom

See UNM book, p. xyw

Part 1

Requirements of Multimedia over the Internet

- **Key requirements**
- **Bit rate requirements**
  - **Audio requirements**
  - **Video requirements**
- **Delay requirements**
  - **Jitter**
  - **Inter-media synchronization**
- **On compression ...**

Multimedia characteristics

- **Information types**
  - Text
  - Images (pixel matrix)
  - Graphics (logical objects)
  - **Animation** (moving graphics)
  - **Video** (moving images)
  - **Audio**

Text, Images, Graphics: ***discrete media***

Animation, Video, Audio: ***continuous media***

- ***continuous media*** (time is part of their semantics)

Audio/video network requirements

- **Key requirements**
  - Bit rates
  - Transit delay **variation**
  - Multicasting capabilities (for distribution)

- **Other requirements**
  - Transit delay
  - Error rate

- Key requirements
- Bit rate requirements
  - Audio requirements
  - Video requirements
- Delay requirements
  - Jitter
  - Inter-media synchronization
- On compression ...

Types of applications

- **Traditional real-time applications** e.g. PABXs

  constant bit rate (CBR)

- **Traditional bulk data applications** e.g. file transfer, email

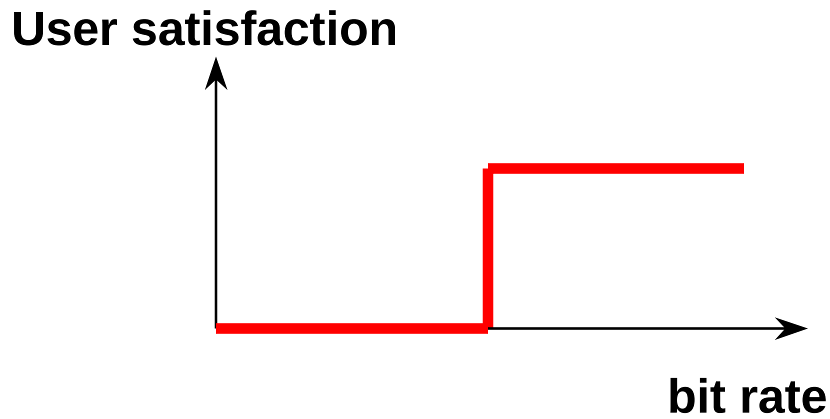
  available bit rate (ABR)

- **Modern real-time applications** e.g compressed audio, video

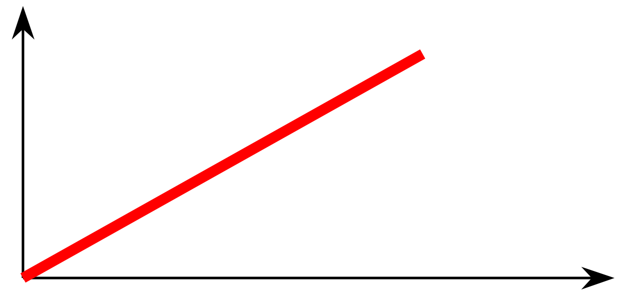
  variable bit rate (VBR)

Quality of Service and bit rate

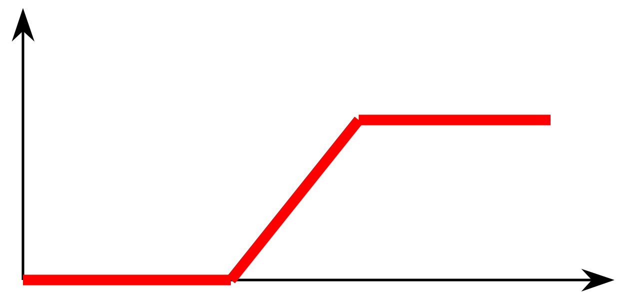
- **CBR applications**



- **ABR applications**



- **VBR applications**



From S.Shenker,
Fundamental Design Issues for the Future Internet, 1995

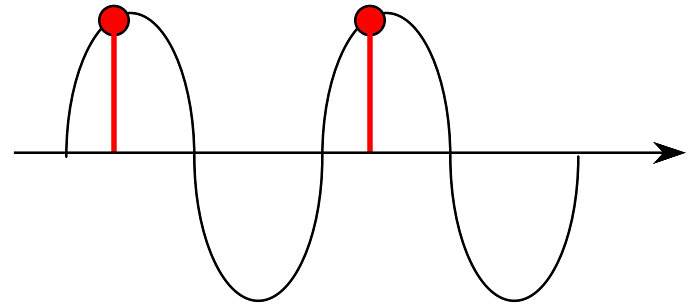
Principle (or platitude)

The grass is always greener on the other side of the hill ...

- **Key requirements**
- **Bit rate requirements**
  - **Audio requirements**
  - **Video requirements**
- **Delay requirements**
  - **Jitter**
  - **Inter-media synchronization**
- **On compression ...**

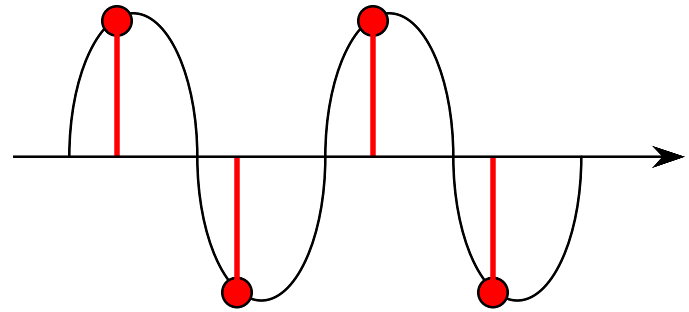
Nyquist theorem

- **To faithfully represent an analog signal**
  - if maximum frequency ***f***
  - sampling rate at least ***2f***



Sampling at *f* :
impossible to reconstruct

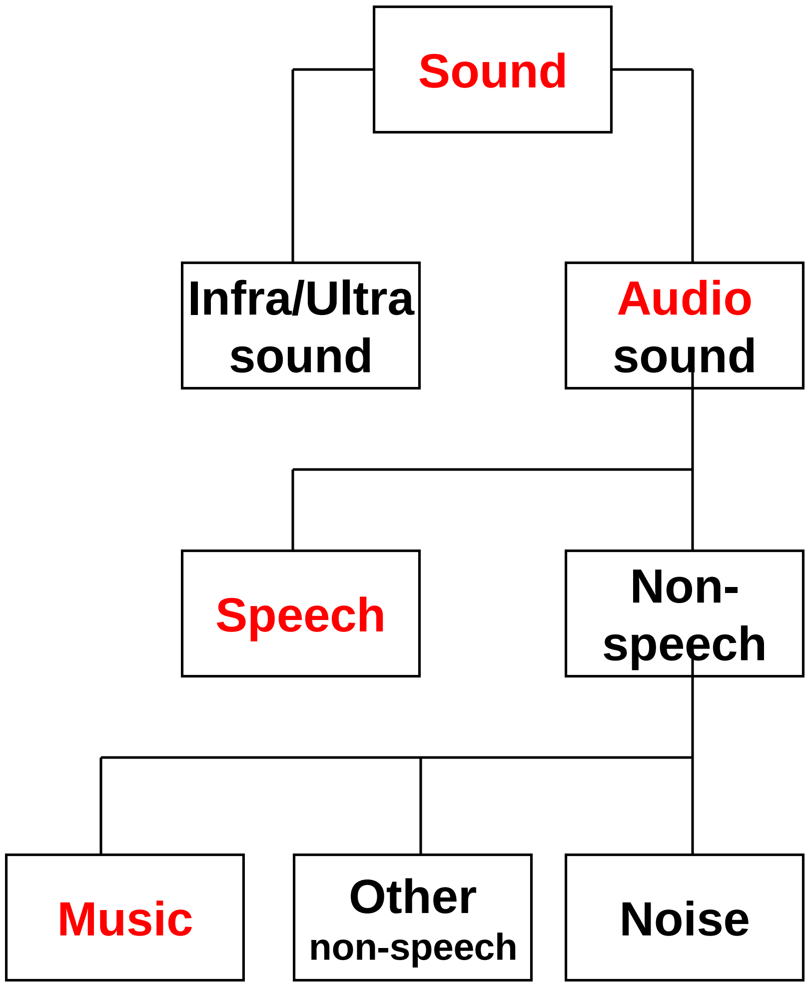
- **Application to audio**
  - if sampling rate is **8 kHz**
  - bandwidth is **3.4 kHz**



Sampling at *2f* :
easier to reconstruct

Sound, Audio, Speech, ...

- ***Sound***: vibration of matter
- ***Audio***: audible sound (by humans)
  - human audible spectrum:
    20 Hz - 20kHz
- ***Speech***: **a particular type of sound**
  - we hear better than we talk
  - speech spectrum:
    50-10 kHz
- ***Music***: a particular case of non-speech sound



Audio bit rate requirements

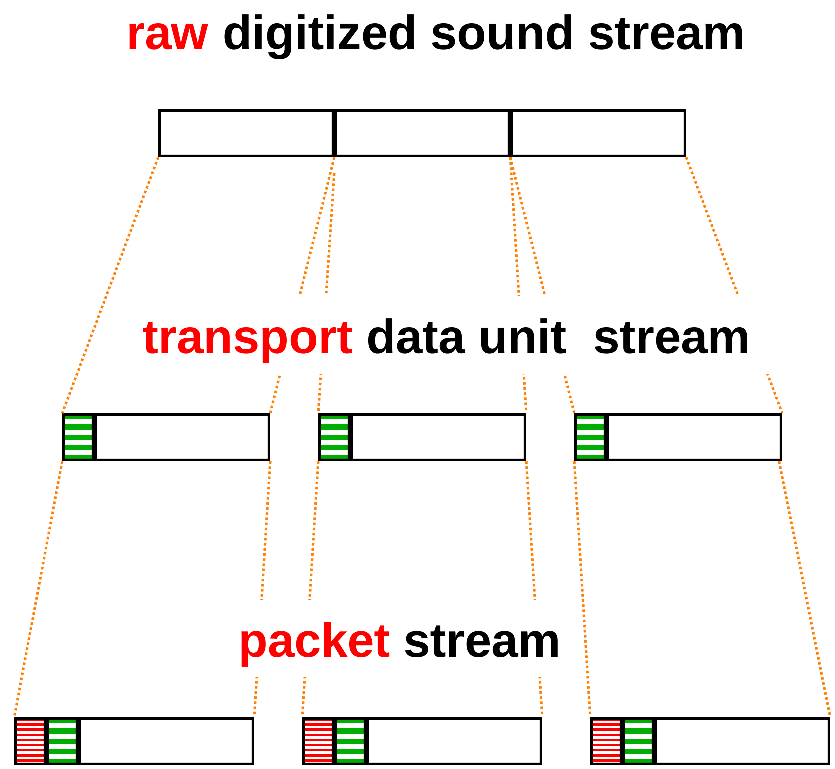
Quality	Technique or standard	Kbps	Compr.
Telephone quality			
Standard	G.711 PCM	64	
Standard	G.721 ADPCM	32	Y
Lower	G.728 LD-CELP	16	Y
Lower	GSM	13	Y
Standard-	**G.729 LD-CELP**	**8**	**Y**
Lower+	CELP	5-7	Y
CD Quality			
Consumer CD-audio	CD-DA	1441 (stereo)	
Consumer CD-audio	MPEG with FFT	192-256	Y
Sound studio quality	MPEG with FFT	384	Y
Consumer CD-audio (MP3)	MPEG2.5 Layer III	128 (stereo)	Y

Reproduced from "Understanding Networked Multimedia" by François Fluckiger, Prentice Hall 1995

Which bit rate is actually needed?

- **Network overheads incl.:**
  - RTP header (12 bytes)
  - Transport Protocol header (usually UDP, 8 bytes)
  - IP header (20 bytes)
  - Example:

    raw G.711 **64** Kbps requires from **68** to 80 Kbps
- ***However, speech contains silence***



Silences in speech

- **Monologue**
  - typically **15%** silence

- **Bi-party telephone conversation**
  - **20%** silence for overall conversation
  - **60%** silence for each party
  - If silence suppressed, required bit rate is in effect **<40%** of nominal raw bit rate

Observations, Trends

Audio does not eat bandwidth

Voice packets will swim in an ocean of data packets

- Key requirements
- Bit rate requirements
  - Audio requirements
  - Video requirements
- Delay requirements
  - Jitter
  - Inter-media synchronization
- On compression ...

Video bit rate requirements

Quality	Technique or standard	Mbps	Compr.
Video conf. quality	**H.261**	**0.1**	**Y**
VCR quality	**MPEG-1**	**1.2.**	**Y**
Broadcast quality	**MPEG-2**	**2-4** (1)	**Y**
Studio-quality digital TV			
Uncompressed	**ITU-R 601**	**166**	
Compressed	**MPEG-2**	**3 to 6** (2)	**Y**
HDTV			
Uncompressed	**CD-DA**	**2000**	
Compressed	**MPEG-2**	**25 to 34**	**Y**

(1): future; current implementations: 4 to 7

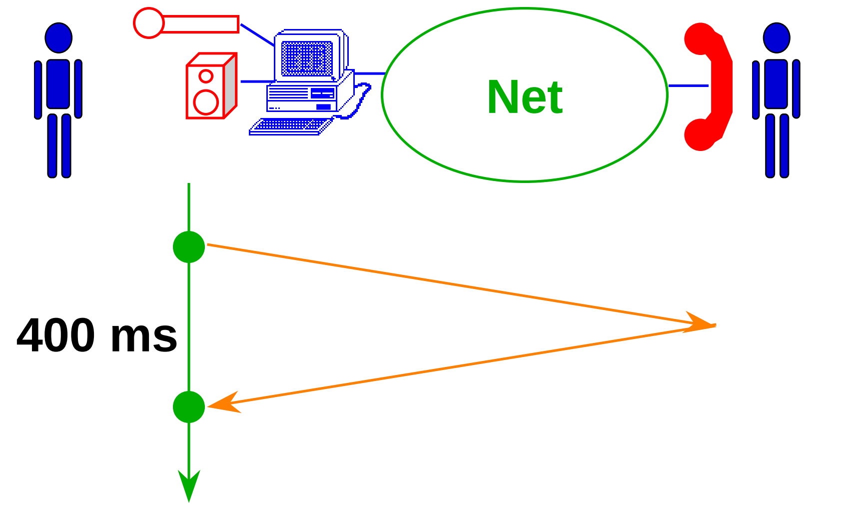
(2): future; current implementations: 6 to 10

Reproduced from "Understanding Networked Multimedia" by François Fluckiger, Prentice Hall 1995

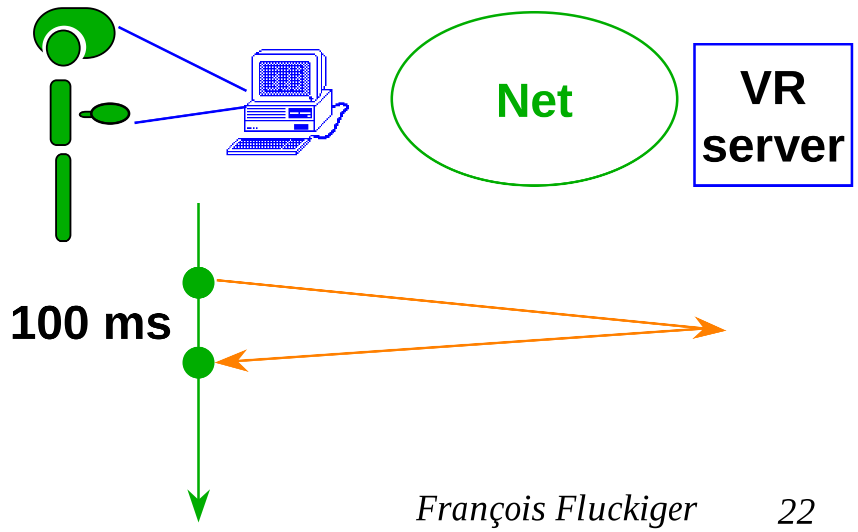
- **Key requirements**
- **Bit rate requirements**
  - **Audio requirements**
  - **Video requirements**
- **Delay requirements**
  - **Jitter**
  - **Inter-media synchronization**
- **On compression ...**

Network Transit Delay

- **Telephone conversation:**
  - Round-trip delay **< 400** ms
  - for natural conversation

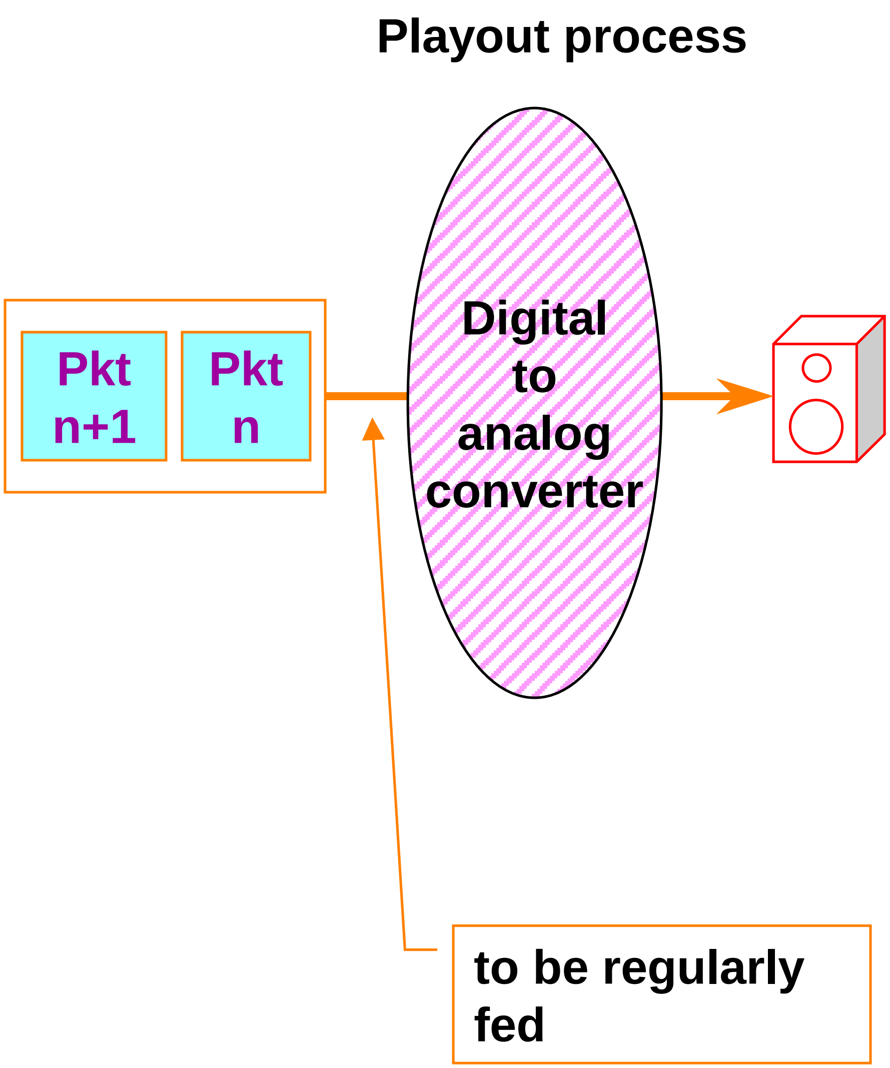


- **Virtual reality**
  - Round-trip delay **< 100** ms
  - for impression of immersion



Transit delay variation (Jitter)

- Receiver to wait a ***delay offset*** before playout
- **Called *delay equalization***
- **Increases overall end-to-end latency**



- Key requirements
- Bit rate requirements
  - Audio requirements
  - Video requirements
- Delay requirements
  - Jitter
  - Inter-media synchronization
- On compression ...

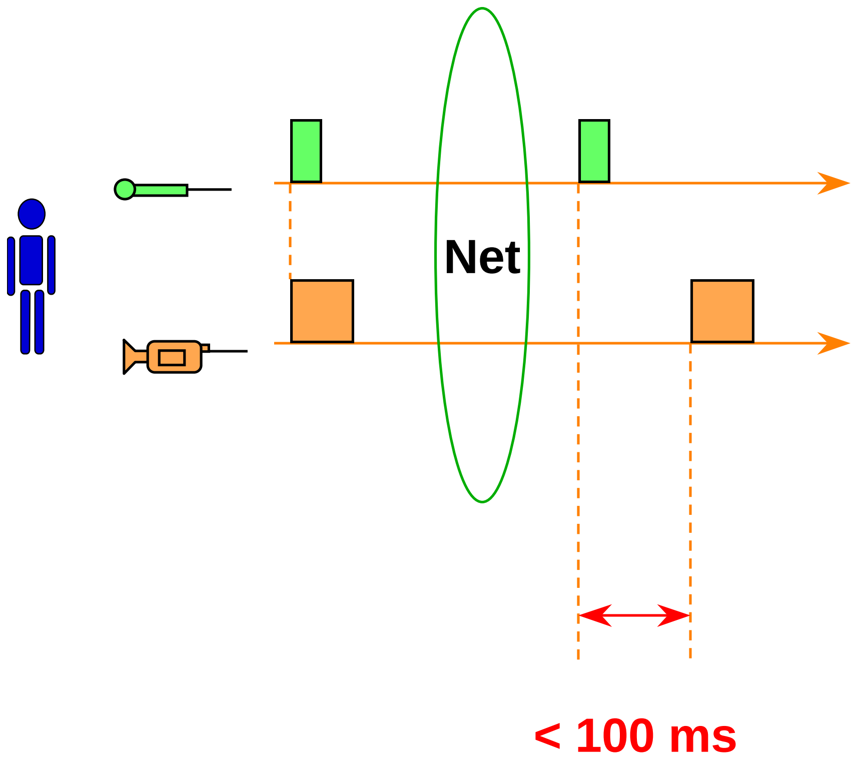
Inter-media synchronization

- **Called *orchestration***

- **Particular case:**

  lip synchronization

  - A skew of **80-100 ms** is **generally tolerated**



Audio/video relative priorities

- The ear behaves as a ***differentiator***
- The eye behaves as an ***integrator***

- Toleration of transmission errors affecting sound much lower than for video

- **When audio and video streams part of the same application competing for network resources, audio stream should have priority**

See UNM book, p. 347-348

- **Key requirements**
- **Bit rate requirements**
  - **Audio requirements**
  - **Video requirements**
- **Delay requirements**
  - **Jitter**
  - **Inter-media synchronization**
- **On compression ...**

Audio-compression techniques

Encoding techniques

- **DPCM, Delta**
- **ADPCM**

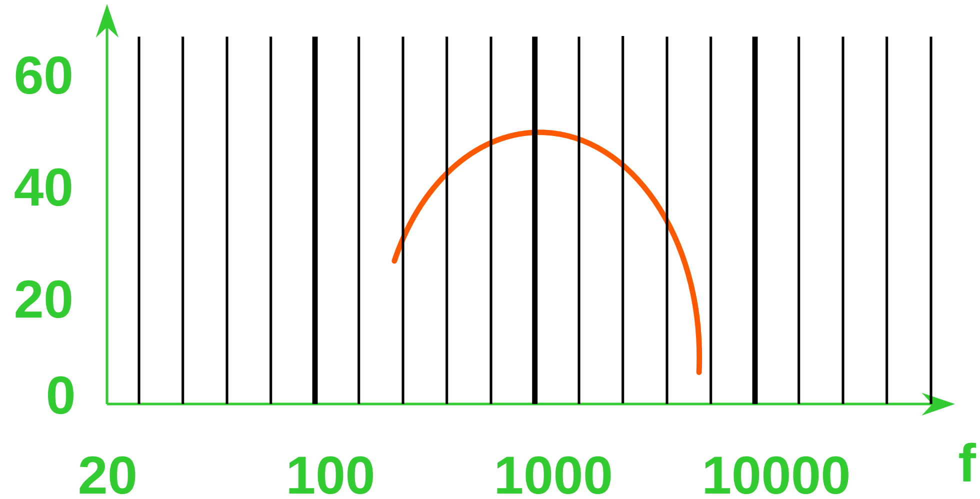
Source compression techniques

Based on psycho-acoustic model

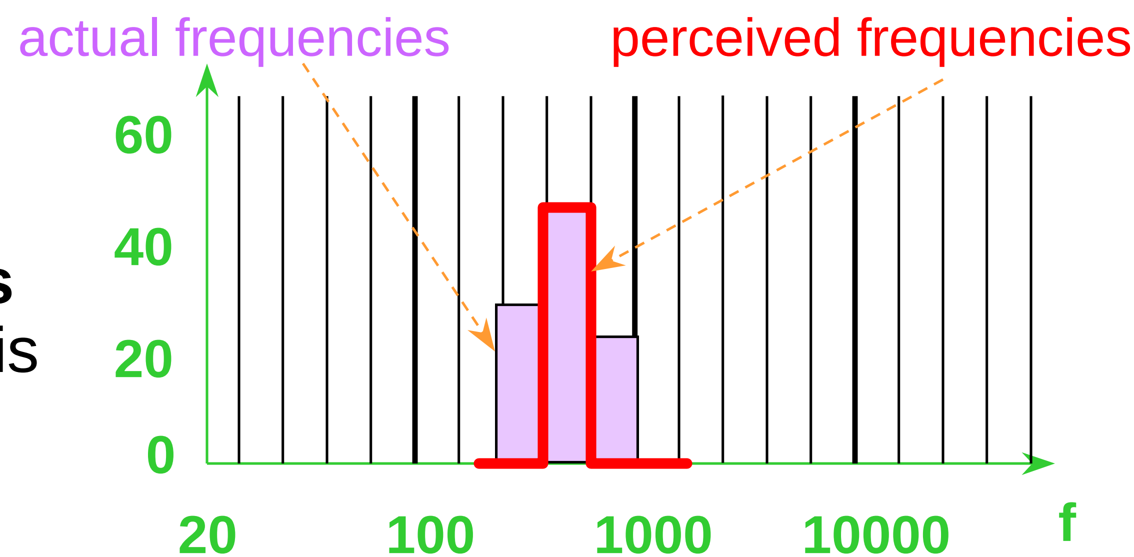
- **Transform encoding (all sounds)**
  - Fast Fourier Transform (FFT)
  - Wavelet transform
- **Source modeling/synthesis coding (for speech)**

Psycho-acoustic and Masking

- **Response of ear to frequency:**
  - **ear most sensitive between 2 - 5 kHz**

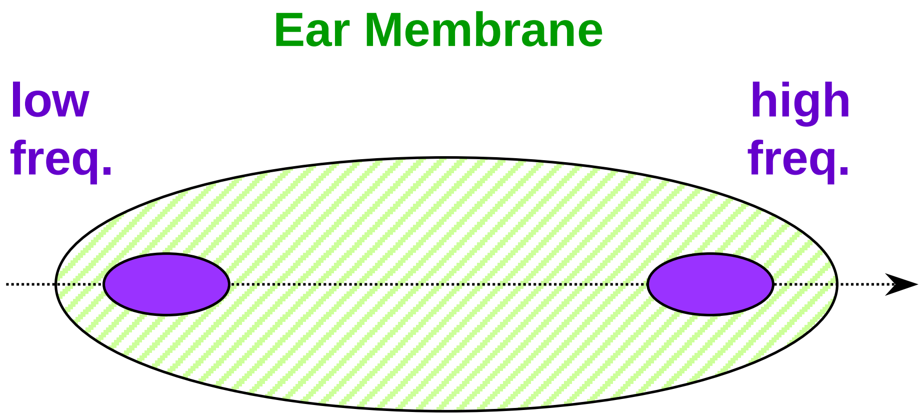


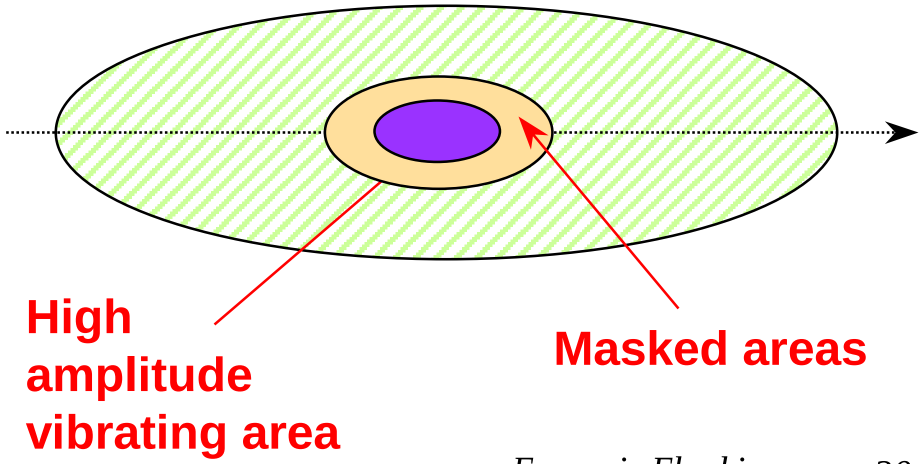
- **Masking:**
  - ear does not register energy in **some frequencies band,** when there is more energy in a nearby band



Physiology and masking

- **Ear membrane vibrates as a function of frequency**
- **High frequencies:**
  - at one end
- **Low frequencies:**
  - at opposite end
- **Vibration of a area forces close areas to vibrate at the same frequency, and not at their own**





Voice modeling techniques

- **Human Vocal system model relies**
  - on a **set of cylinders** of differing diameters

    (e.g. 10 in LPC-10)
  - excited by a signal at a certain frequency

- **Operates over 20 ms, on standard PCM samples**

Principle (or platitude)

This is what we perceive that count,

not what the physical reality is!

or

The Reality is what we perceive

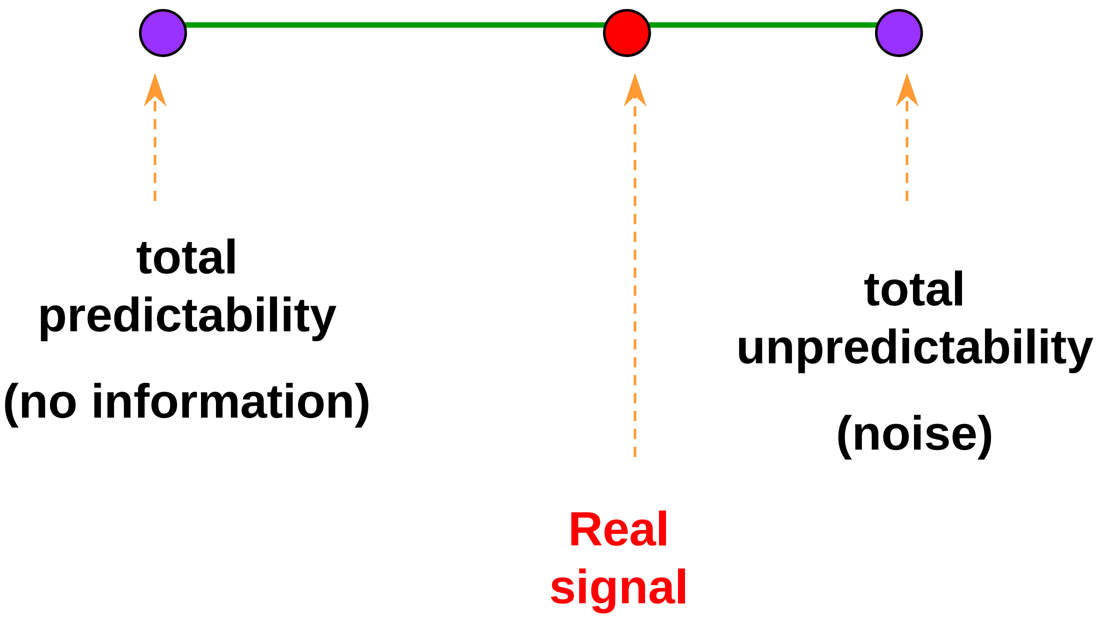
Information rate, bit rate, entropy

- **Information rate is different from bit rate**
- **Information content or *entropy* of a sample:**
  - a function of how different it is from the predicted value
- **Shannon's theory:**
  - any signal which is **totally predictable**: carries **no information**
    - (e.g. a sine wave)
  - **noise** is completely **unpredictable:** high entropy

Real, contentless signals, noise

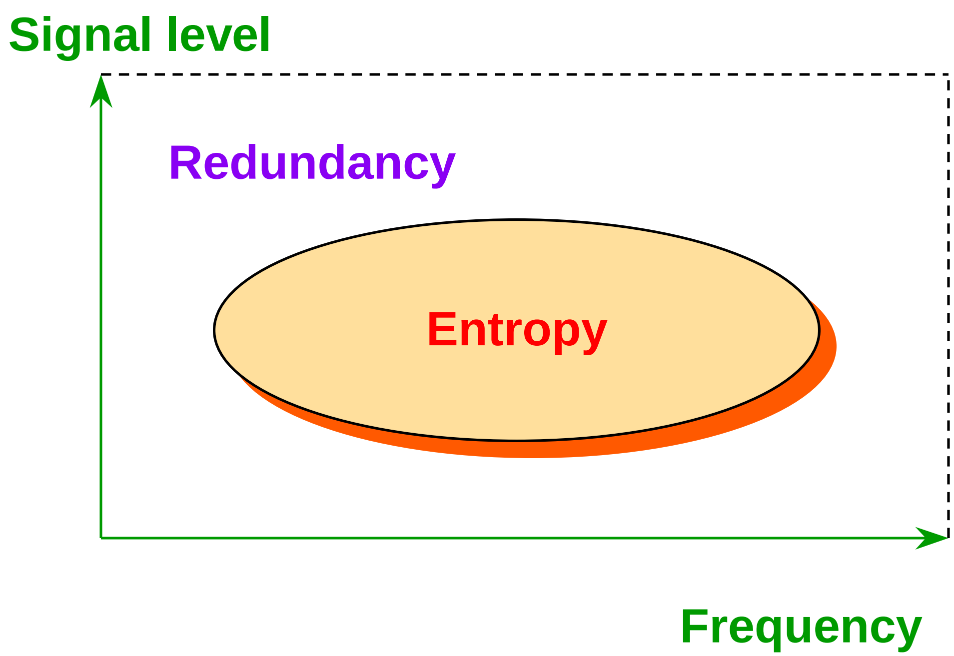
Real signals

lie somewhere between the two extremes



Entropy, redundancy, compression

- **Real signals only occupy a portion of the Signal level - Frequency plane**
- **Entropy area: where the information lies**
- **Redundancy area adds no information**



Compression: removes (part of) the redundancy

The Effect of compression

Compression removes redundancy ... *but*

Principle (or platitude)

Redundancy

is essential for resistance to

errors

The Effect of compression

Compression removes redundancy

- **Redundancy essential for resistance to errors**
- **Compressed data more sensitive to errors**

End of

Part 1

Requirements of Multimedia over the Internet

Part 2

Transporting Multimedia over the Internet

*Who has a **good** or **fair** understanding of the difference between …*

FTP *and* **HTTP**

HTTP *and* **TCP**

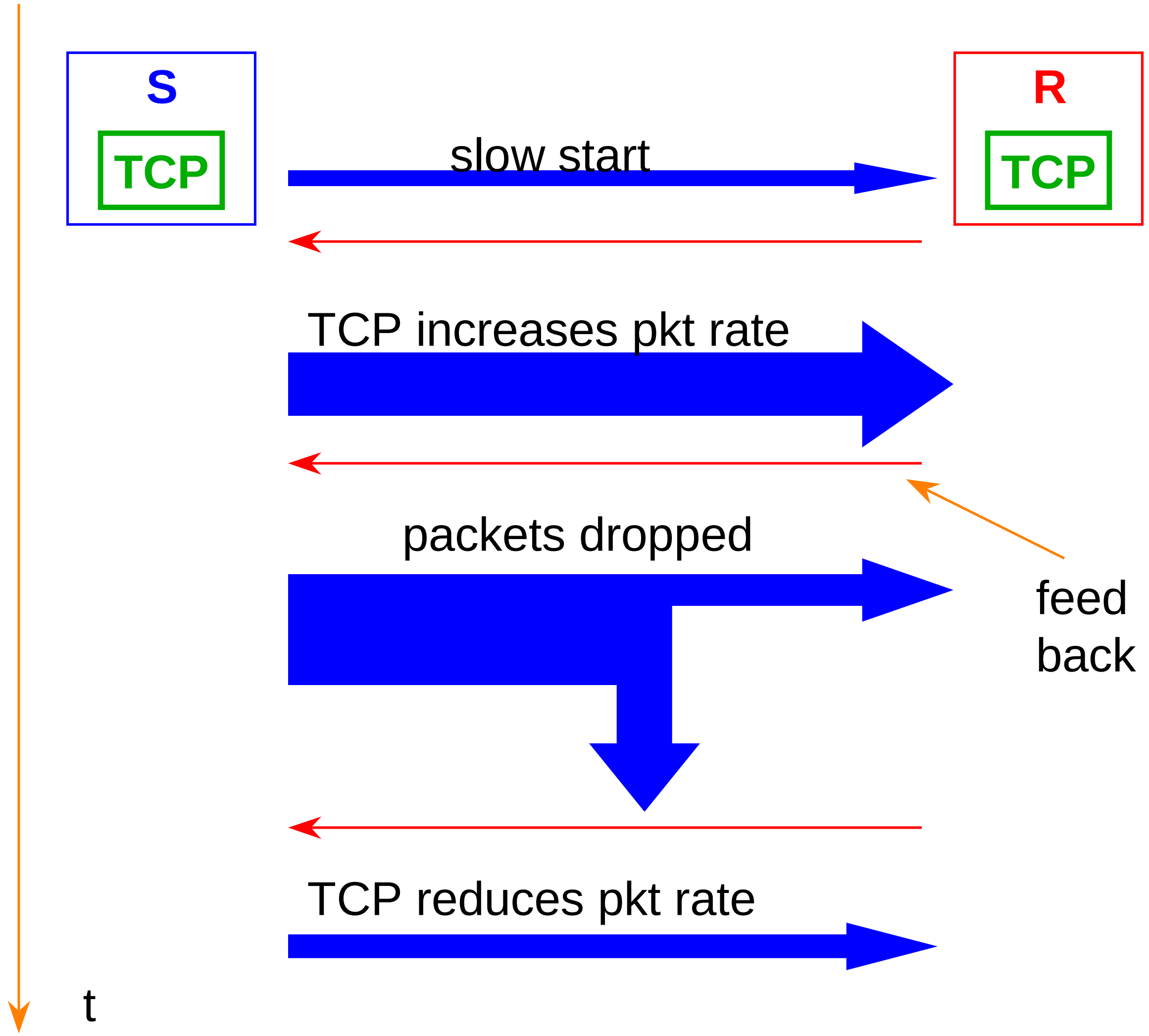
TCP *and* **UDP**

UDP *and* **RTP**

RTP *and* **RTCP**

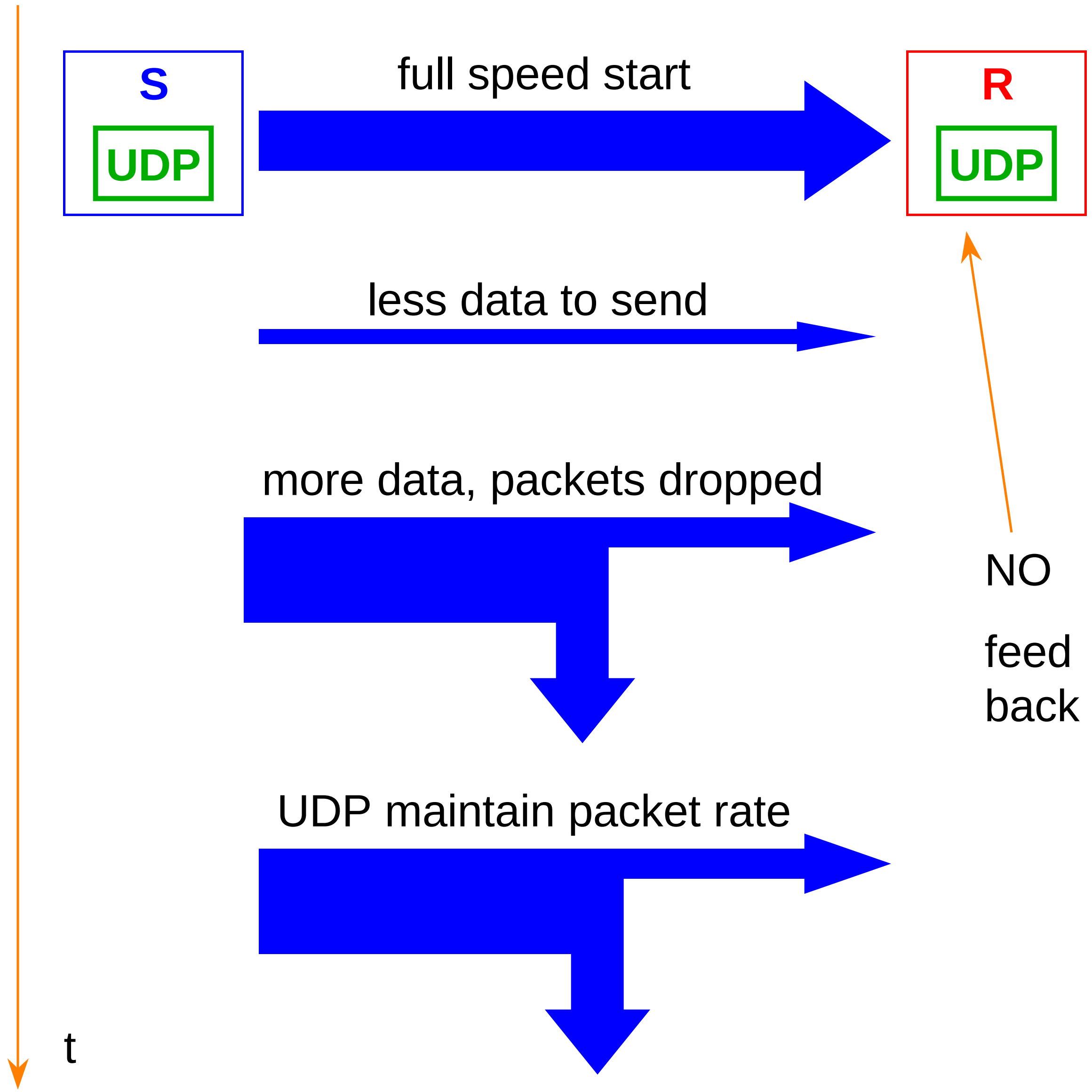
TCP behaviour

- **Slow start**
- **Sender aware of packets dropped**
- **Sender decreases bit rate when packet dropped**



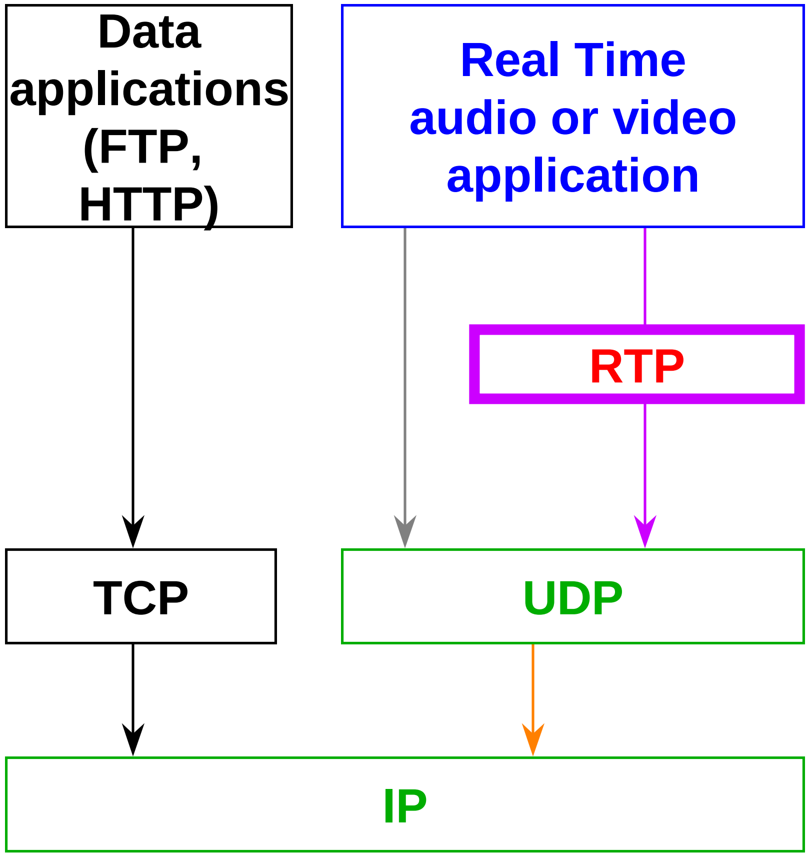
UDP behaviour

- **UDP sends blindly to a receiver**
- **No feedback from the receiver**
- **Sender unaware whether packets are dropped/lost**



Protocols for real-time audio and video

- **Audio/video applications cannot operate over TCP**
- **They use UDP**
    - which has no timestamp, feedback, ...
- **All applications use RTP (Real-Time Transport Protocol)**
    - time-stamp
    - packet loss detection



Real-Time Protocol

- **RTP: an Internet IETF standard**

- **Supports**
  - timing reconstruction: timestamp (4 bytes)
  - loss detection: sequence number (2 bytes)

- **Lighter than TCP**
  - no retransmission, no flow control
  - TCP header: **20 bytes**; RTP header: **12 bytes**

Real-Time Protocol services

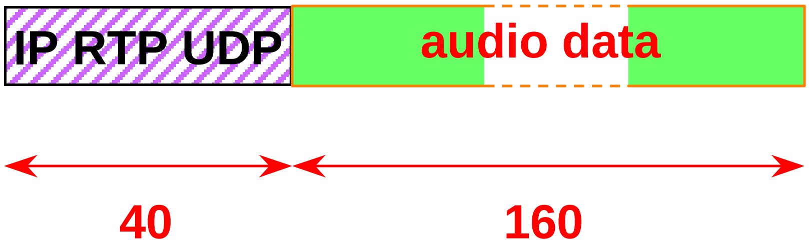
- **Two parts in RTP**
  - ***RTP*** per se: for carrying data
  - ***RTCP*:** to identify participants, monitor the quality of the service

- **Session control (*RTCP*)**
  - Receivers send periodically “reports”
  - “Reports” indicate how good the reception is

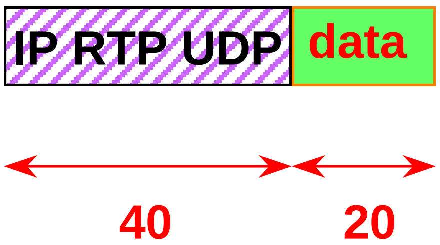
- RTP
  - Needs and Principles
  - Header overhead
- End-systems improvements
  - Redundancy coding
  - Error concealment
- Quality of Service
- Unfair competition ...

On header overhead

- **IP+UDP+RTP headers =** ***40 bytes***
- **At 64 Kbps PCM**
  - **20** ms = **160** Bytes
  - overall rate **= 80** Kbps



- **At 8 Kbps** (e.g. with G.729) (e.g. over modem lines)
  - **20** ms = **20** Bytes



IP, UDP, RTP compression

- **IP/UDP/RTP compression specified by**
  - ***Robust Header Compression*** **(*ROHC*) IETF draft**

- Can reduce to ***1 byte*** (best case)

- Operates on a ***link-by-link*** basis

Basic principles

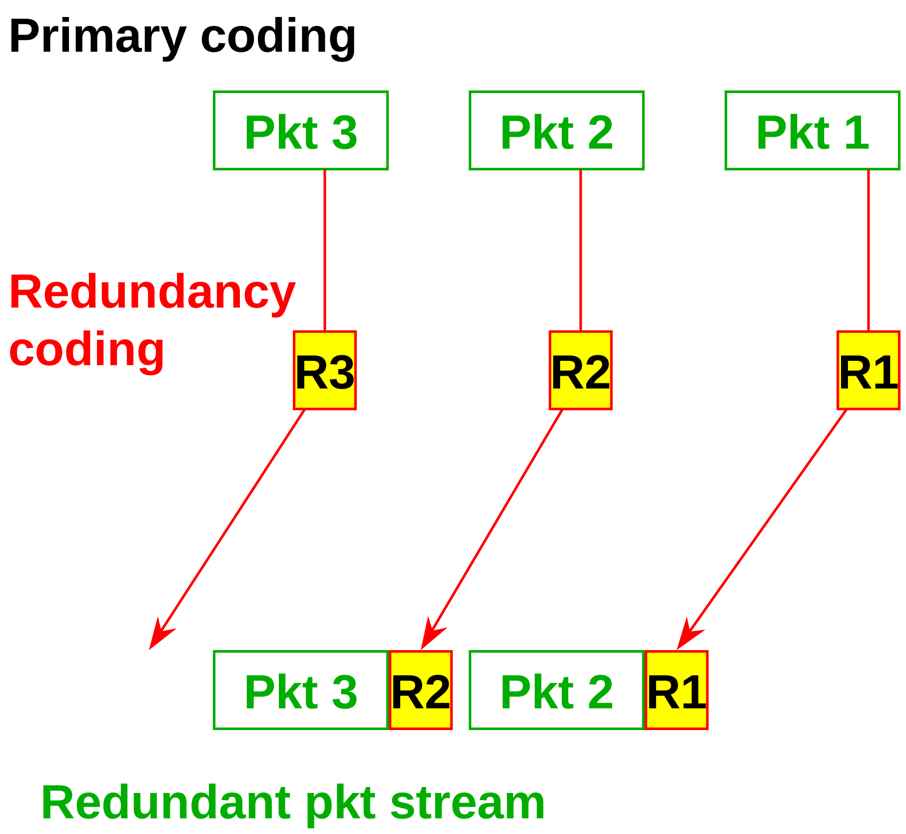
- **Fixed fields removal**
  - parts of the headers remain unchanged between pkts
- **Differential encoding**
  - some fields vary in a predictive, monotonic way
- **Re-coding combinations of fields**
  - some fields may be combined and hash coded

- RTP
  - Needs and Principles
  - Header overhead
- **End-systems improvements**
  - Redundancy coding
  - Error concealment
- Quality of Service
- Unfair competition ...

Low-bit rate redundancy

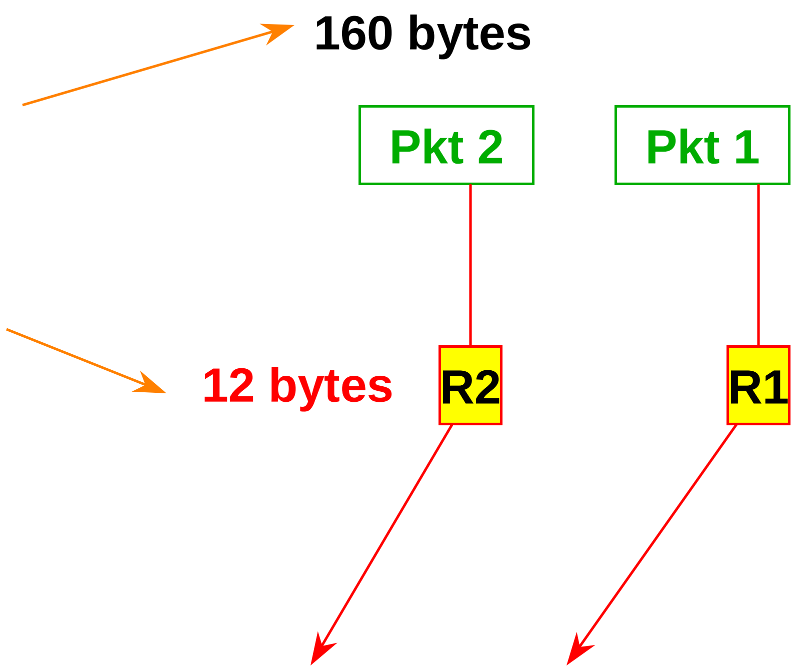
Compression aims at removing redundancies ... but redundancies improve resistance to data errors

- re-code each packet at lower resolution
- insert re-coded packet into one subsequent pkt(s)



Example of redundancy: RAT (UCL)

- Primary coding = 64 Kbps
- Redundancy coding = 4.8 Kbps
- **Experiment shows reasonable repair with high loss rate (40%)**



- RTP
  - Needs and Principles
  - Header overhead
- End-systems improvements
  - Redundancy coding
  - Error concealment
- Quality of Service
- Unfair competition ...

Error concealment (audio example)

Replace missing packet with

- **silence**
  - ‘‘OK’’ if pkt<16ms, loss rate<1%; beyond, clipping effect (1)
- **white noise**
  - (better than silence)

(1) ‘‘OK’’ means tolerable; does not mean unnoticed

Phemic Restoration

- brain uses ***phemic restoration*:**

  ***“**the ability of the brain to subconsciously repair a missing segment of speech with the correct sound”*

- phonemic restoration
  - occurs better when missing segment replaced by ***noise*** instead of silence

- RTP
  - Needs and Principles
  - Header overhead
- End-systems improvements
  - Redundancy coding
  - Error concealment
- Quality of Service
- Unfair competition ...

Improving QoS

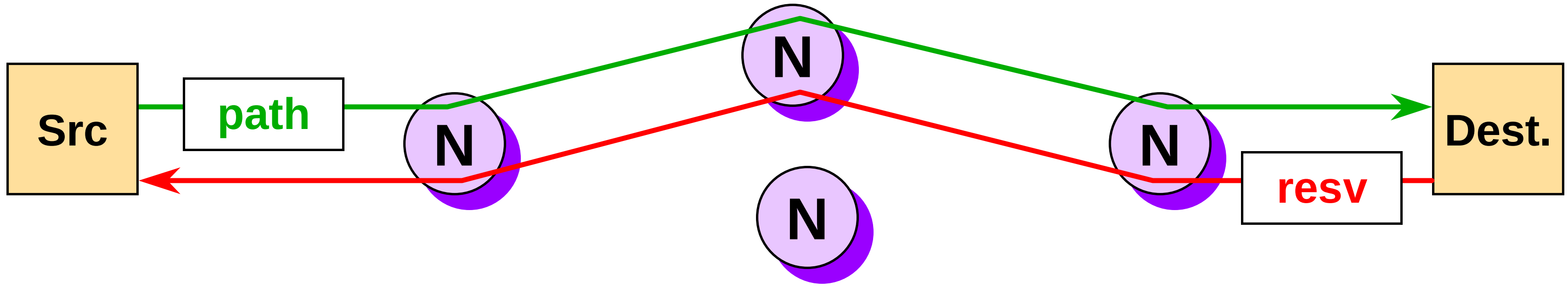
Combination of solutions

- ***Integrated Services*** **: RSVP Protocol**
- ***Differentiated Services***

IntServ principles

- ***Resource reservation*** **is necessary**
- **Reservations on a per-flow basis**
- **Routers have to maintain flow-specific *states***

RSVP protocol (simplified)



- "***path***" control message sent periodically by source
- "***path***" establish an RSVP route in intermediary routers
- sink replies with a "***resv***" message, according to its capabilities
- "***resv***" reserve resources in node on the route back
- if "***path***" not repeated after time-out, resources released
- "***path***" and "***resv***" are carried by ordinary best-effort datagrams

See UNM book, p. 441

Concern: Scalability

- **Problem**
  - how many soft-states can network handle at a time?
  - problem of **granularity** of the flows

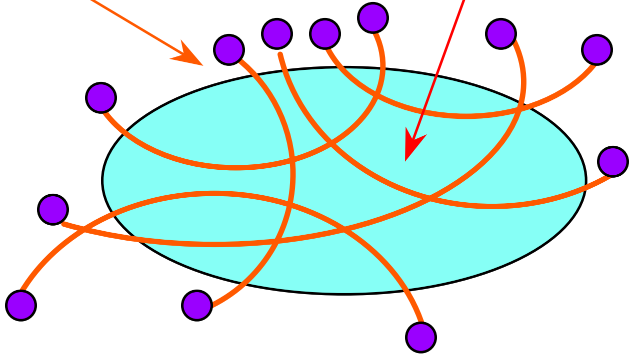
RSVP scalability

- **Millions per-flow RSVP reservations in high speed backbone?**

  - *Problem 1:*

    Signaling overhead

    - CPU**: PATH/RESV processing**
    - memory**: states**



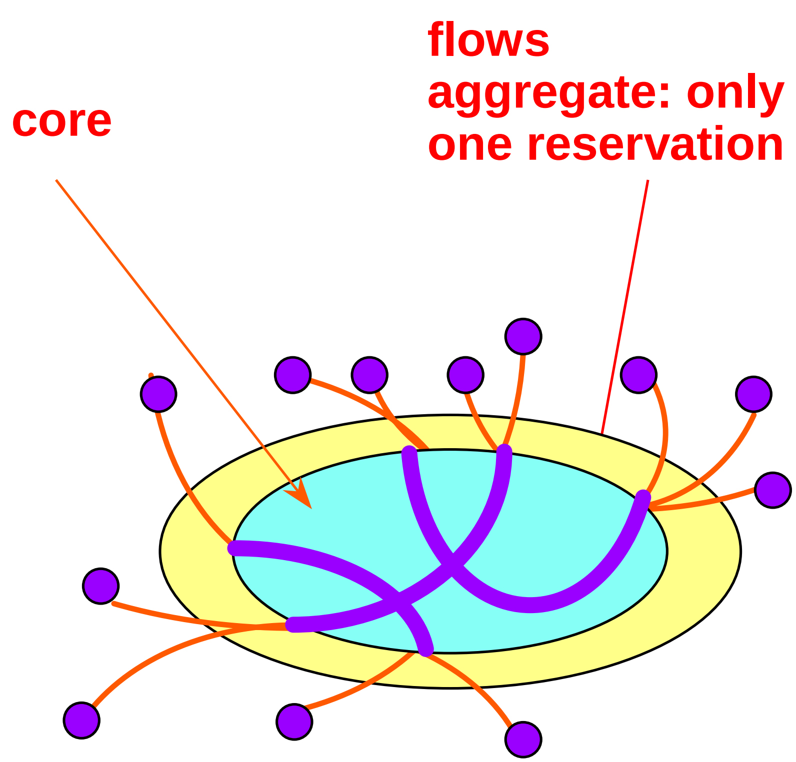
  - *Problem 2:*

    Data pkt overhead:

    - CPU**: Multi-field classification**

Solution to *problem 1* *(Signaling Overhead)*

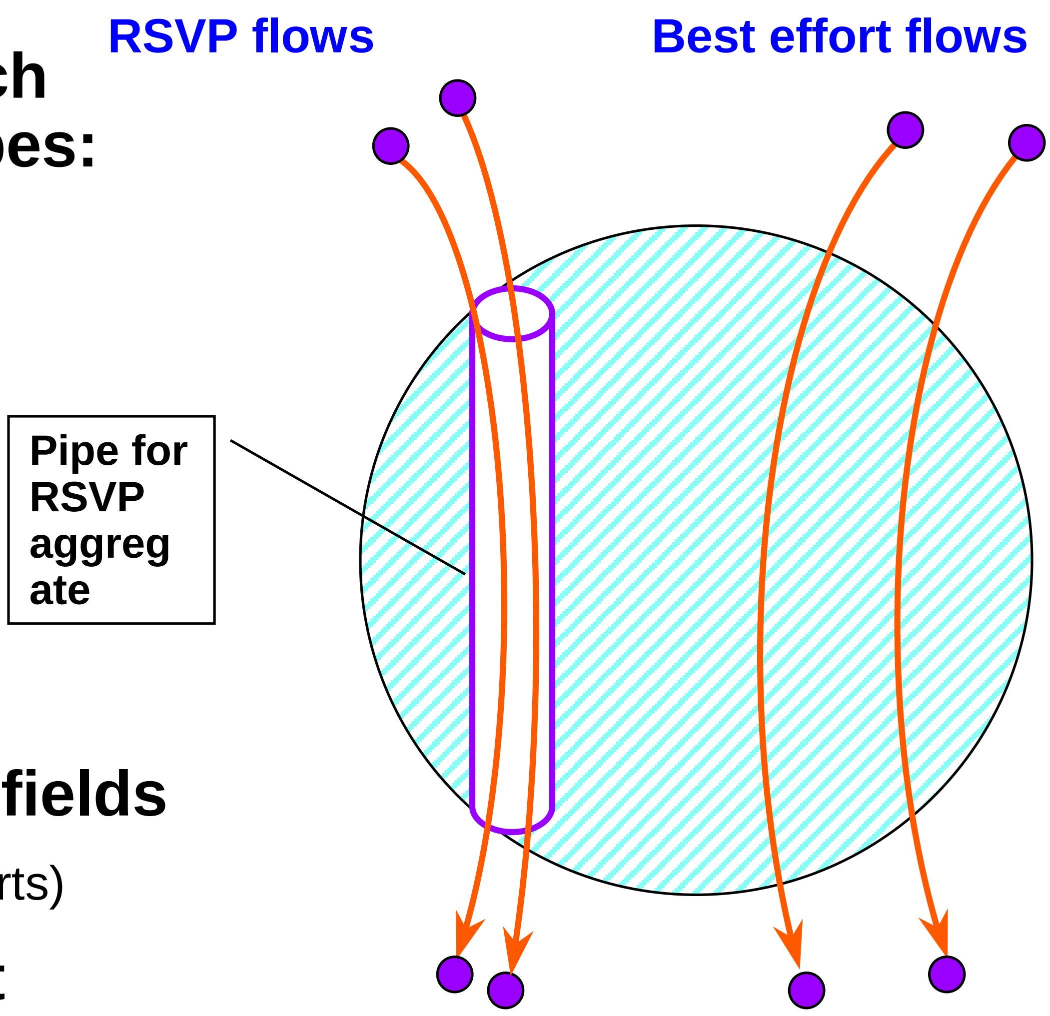
- ***RSVP aggregation*** **in core**
- **Only one reservation between ingress/egress pairs or routers**
  - If N boundary routers,
  - N^2-N RSVP reservations
- **Size of the aggregate does not need change for every new flow request**



Problem 2: overheard of classification

- **Recognizing pkts which belong to reserved pipes:**

  heavy process!



- **Requires**
  - **examining multiple fields**

    (source, dest addresses, ports)
  - **match them against "filters"**

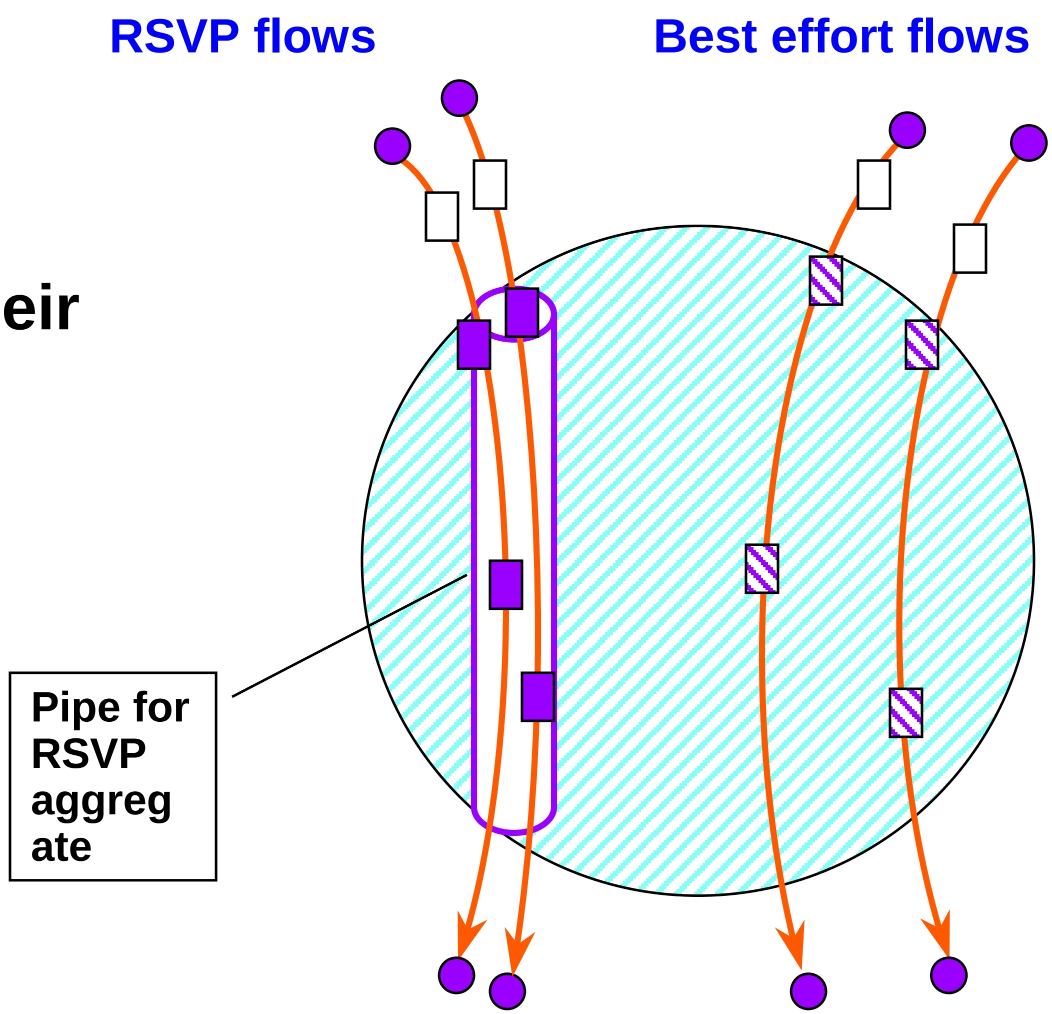
Differentiated Services : principles

- **Rationale**
  - Knowing pkt priority needs heavy classification process
  - Classification may be faster if packets are "***marked***" (1) with a priority

(1) "Marking" pkts also called "***coloring***"

Differentiated Services Packet marking

- **Pkts carry explicitly their priority**



High Priority

Low Priority

Principle (or platitude)

- **Systems with no reservation (e.g. connectionless networks) scale well, but are poor at QoS guarantees**

  Too bad for IP, Ethernet

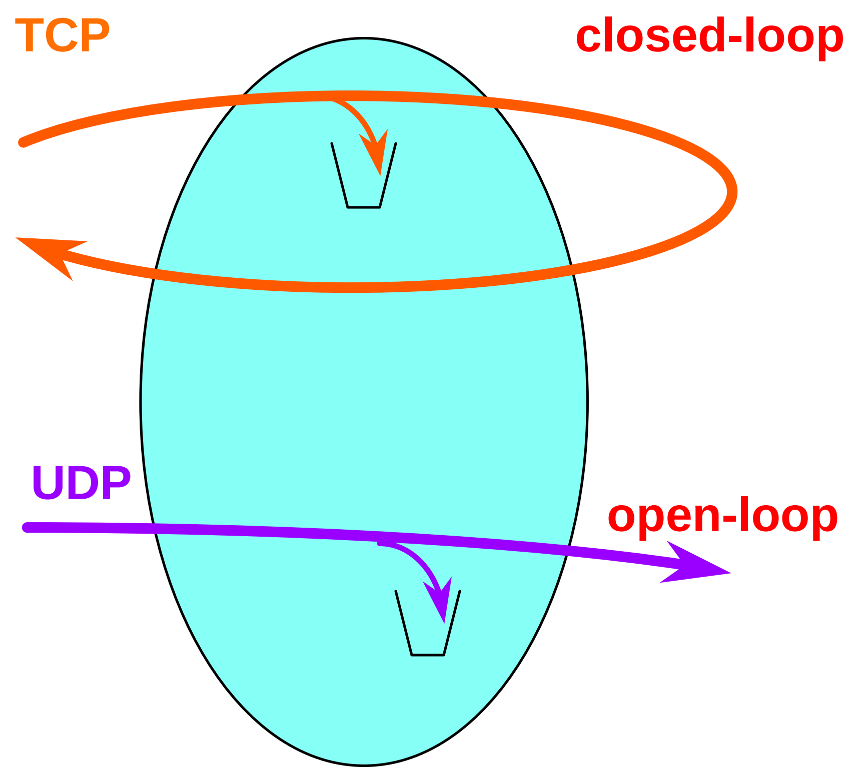
- **Systems with reservations (e.g. connection-oriented networks) are good at QoS guarantees and poor at scaling**

  Too bad for RSVP, ATM

- RTP
  - Needs and Principles
  - Header overhead
- End-systems improvements
  - Redundancy coding
  - Error concealment
- Quality of Service
- Unfair competition ...

Unresponsive flows

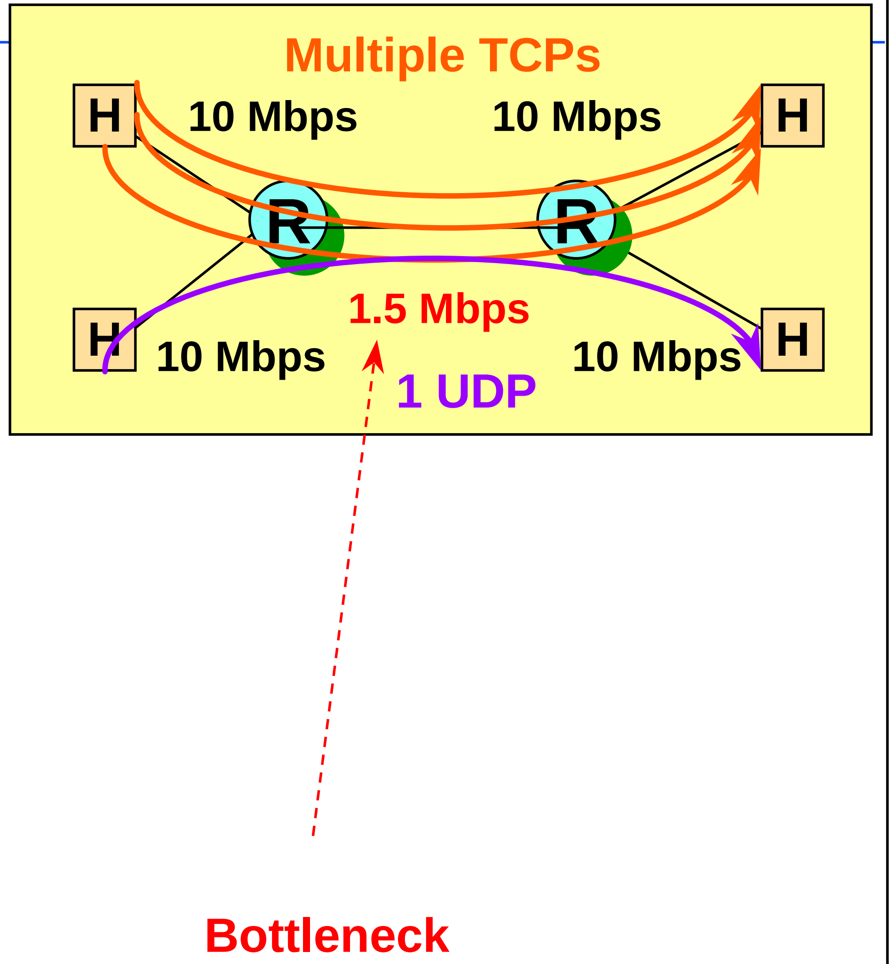
- **Unresponsive flows do not react to congestion indication (pkt loss)**



- **Can create**
  - bandwidth **starvation** inflicted to well-behaved responsive traffic
  - Congestion **collapse** (network busy transmitting pkts that will never reach dest.)

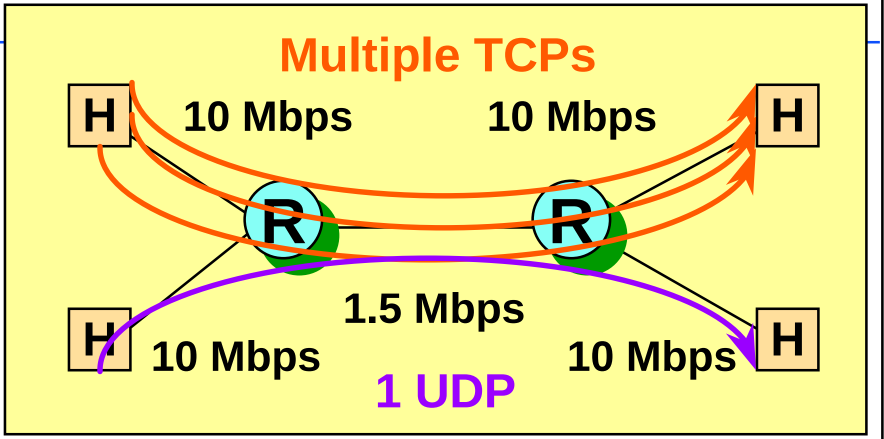
Unfair competition

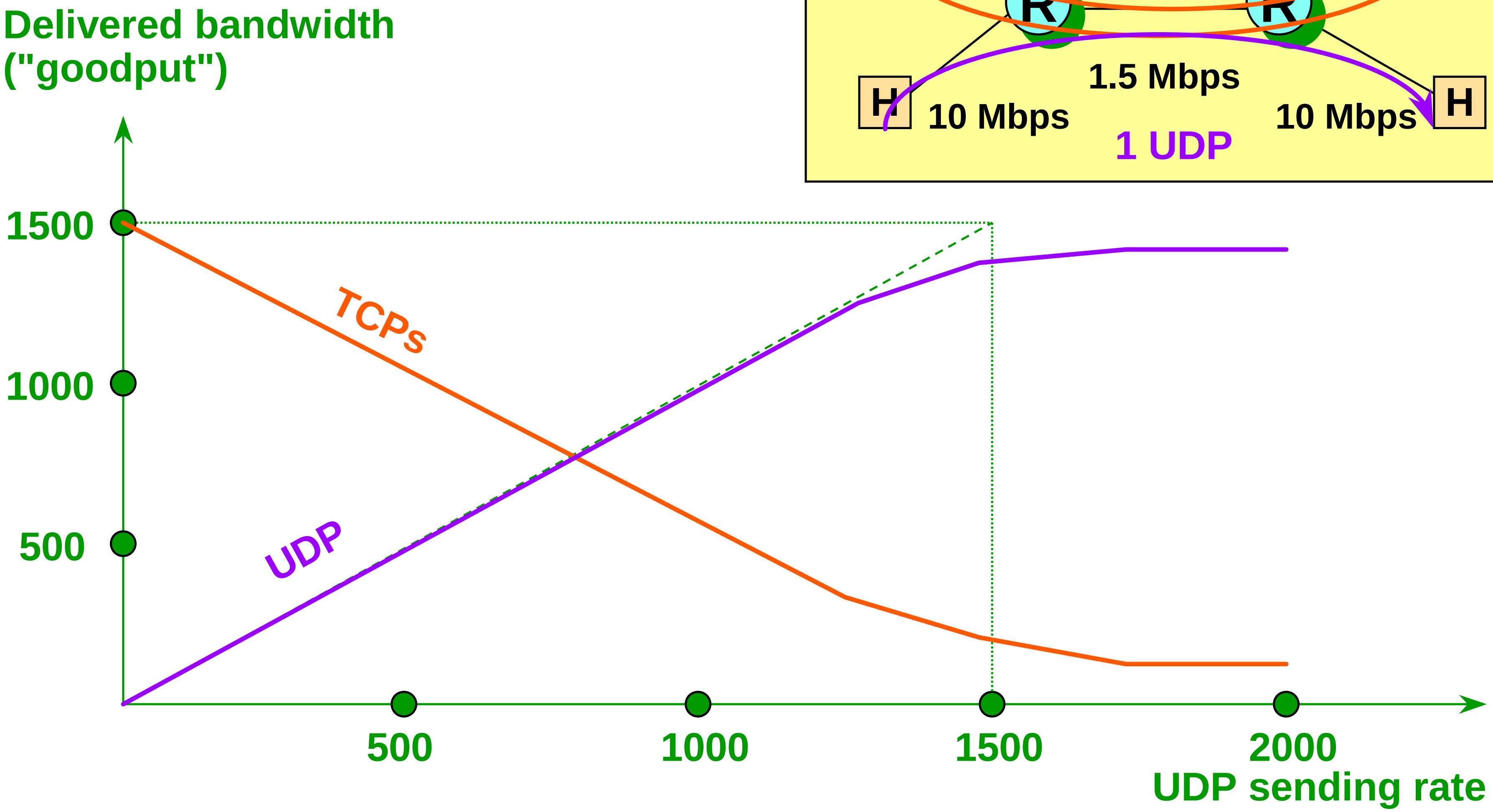
- **Case 1:**
  - **2 LANs (10 Mbps)**
  - **interconnected with T1 and a pair or routers**
- **Competition between**
  - **3 TCP connections and**
  - **1 UDP connection**



From S.Floyd et al, February 97

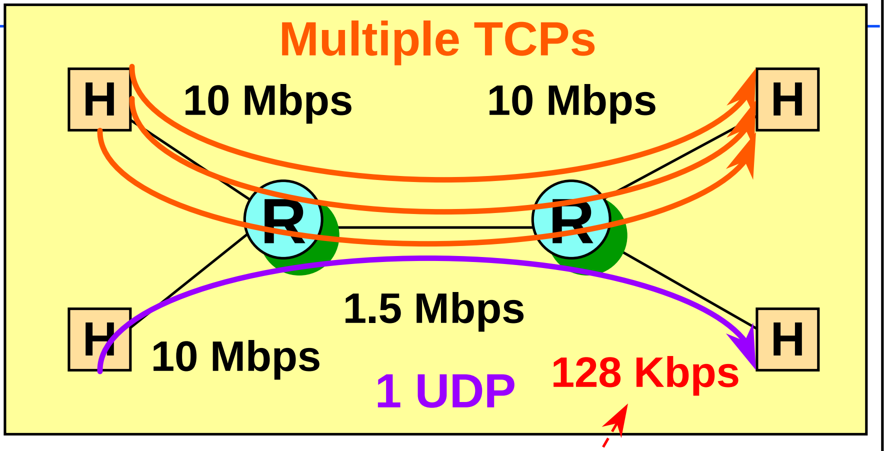
Unfair competition





From S.Floyd et al, February 97

Congestion Collapse

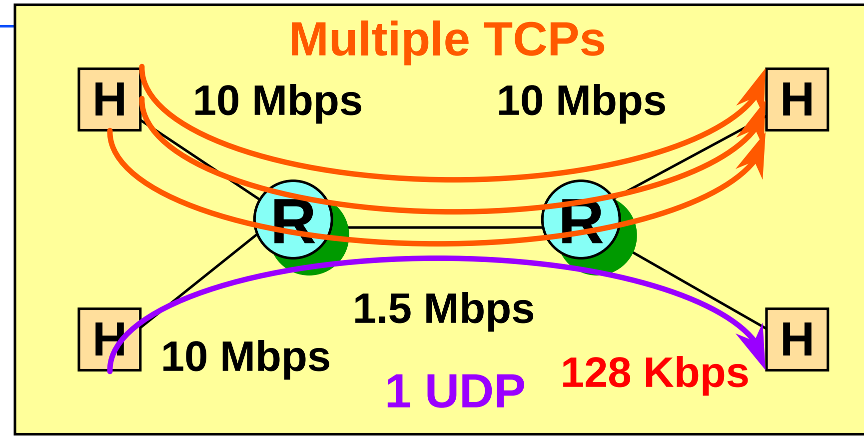


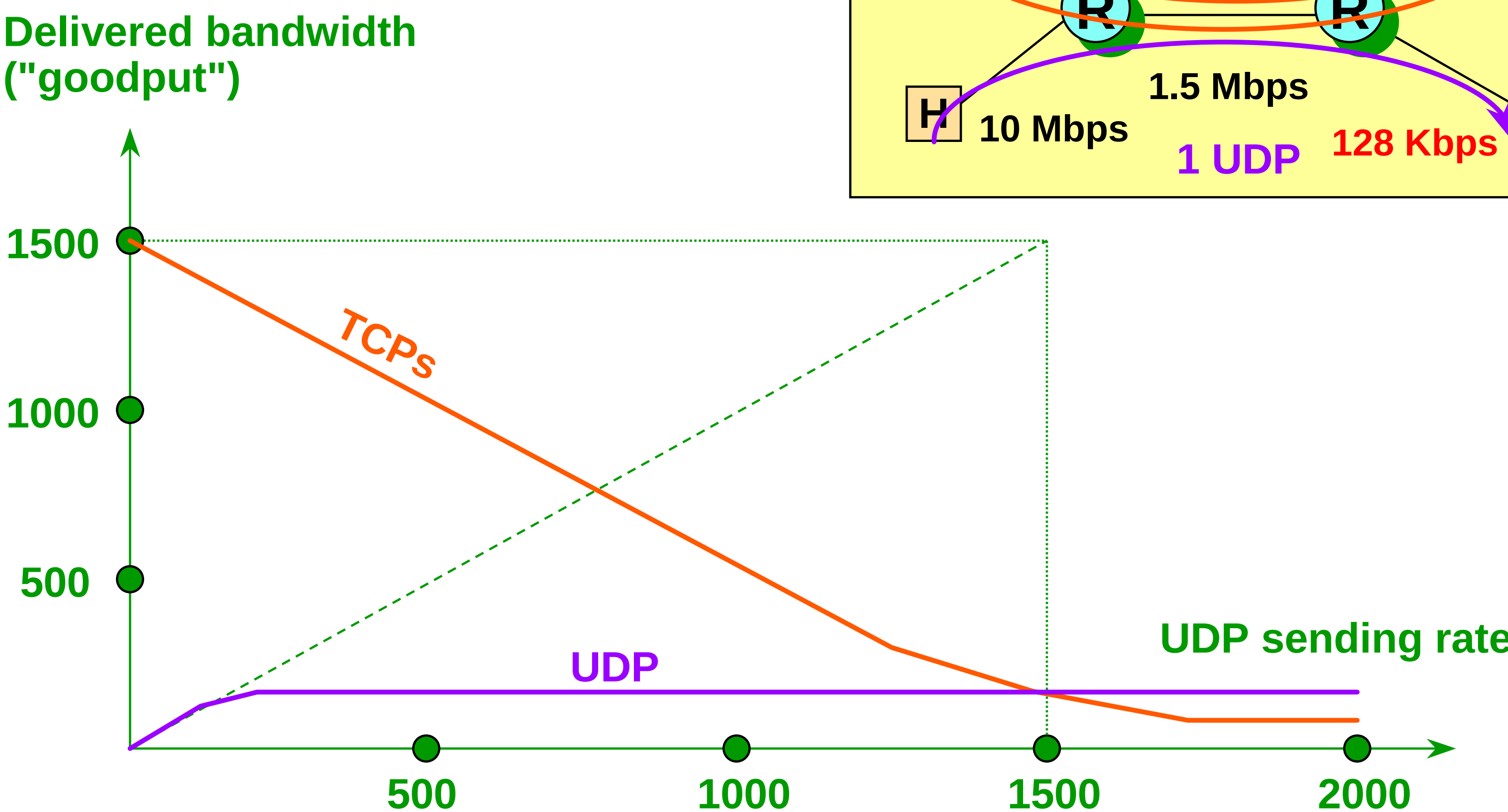
- Case 2:
  - 2 LANs (10 Mbps)
  - interconnected with T1 and a pair or routers
  - UDP receiver connected via ISDN (dual)
- Competition between
  - 3 TCP connections and
  - 1 UDP connection

Bottleneck

From S.Floyd et al, February 97

Congestion Collapse





From S.Floyd et al, February 97

Future or non-TCP applications

- **In the future, UDP traffic may be penalized in public Internets**
- **Real-time Applications will have to behave "more responsively" to congestion indications**
  - adaptive compression levels
  - hierarchical encoding/compression
  - frame, block rate reduction for video
  - ...

Future or non-TCP applications

- **In the future, UDP traffic may be penalized in public Internets**
- **Real-time Applications will have to behave "more responsively" to congestion indications**
  - adaptive compression levels
  - hierarchical encoding/compression
  - frame, block rate reduction for video
  - ...

Epilog

“640KB should be enough for anybody”

Bill Gates, 1981

"We'll have infinite bandwidth in a decade's time."

Bill Gates, 1994

Predictions is a difficult art,
in particular when ...

... predicting the future

W.Allen

Principle (or platitude)

When you have

no memory of the past,

you cannot

predict the future

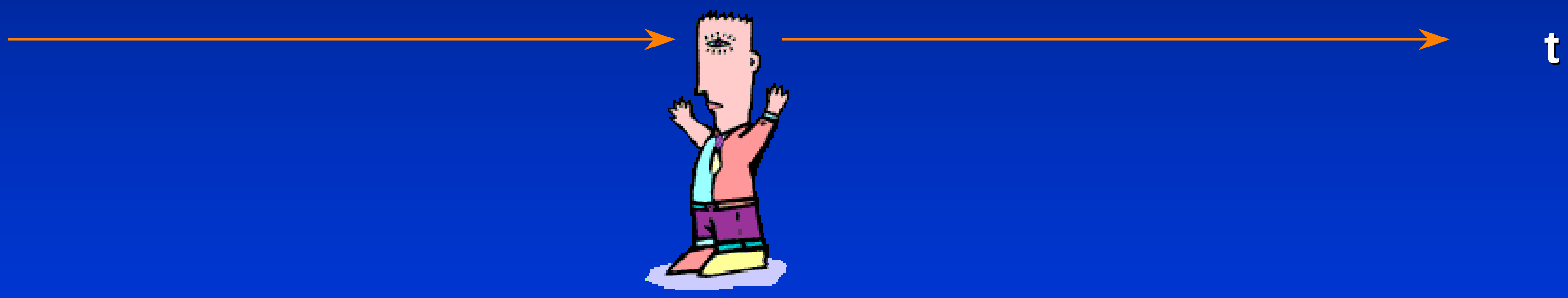
Where the future is?

- Dr Raphael Nunez
  - Prof. Cognitive Psychology and Ethno-Mathematics, Uni of Friburg,Swizerland

Studied how a people in South America see mentally "the future"

- Spatial Localization of the future becomes a phenomenon of "naturalization"

Where the future is?



Thank you