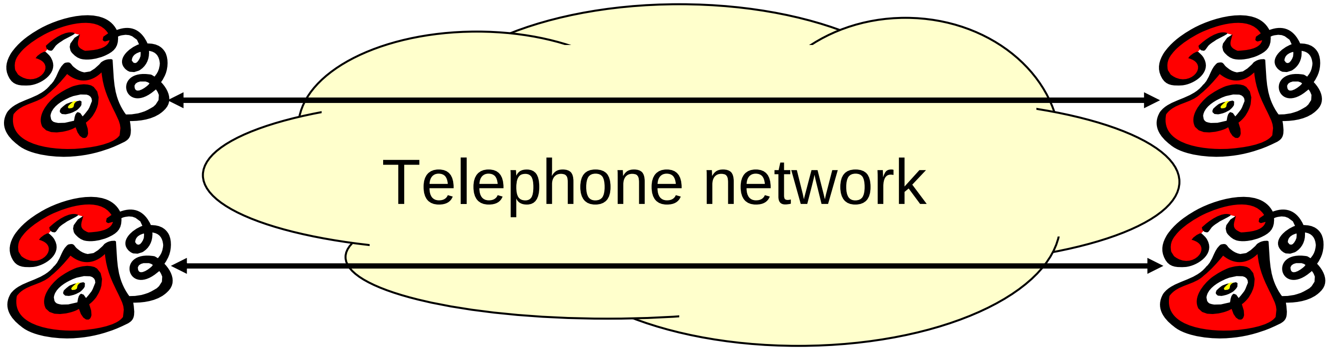
Introduction to Voice over IP

Introduction to technologies for transmitting voice over IP

Telephone network and circuit switching

- **Static circuit allocation**
  - 64Kbps full duplex
- **Characteristics**
  - No compression
  - No high quality communication (e.g. stereo, better codecs), if not in multiples of 64kbps
  - No pause suppression
  - No statistical multiplexing (static allocation of bandwidth)
  - Signalling procedure (call setup)



Data network and packet switching

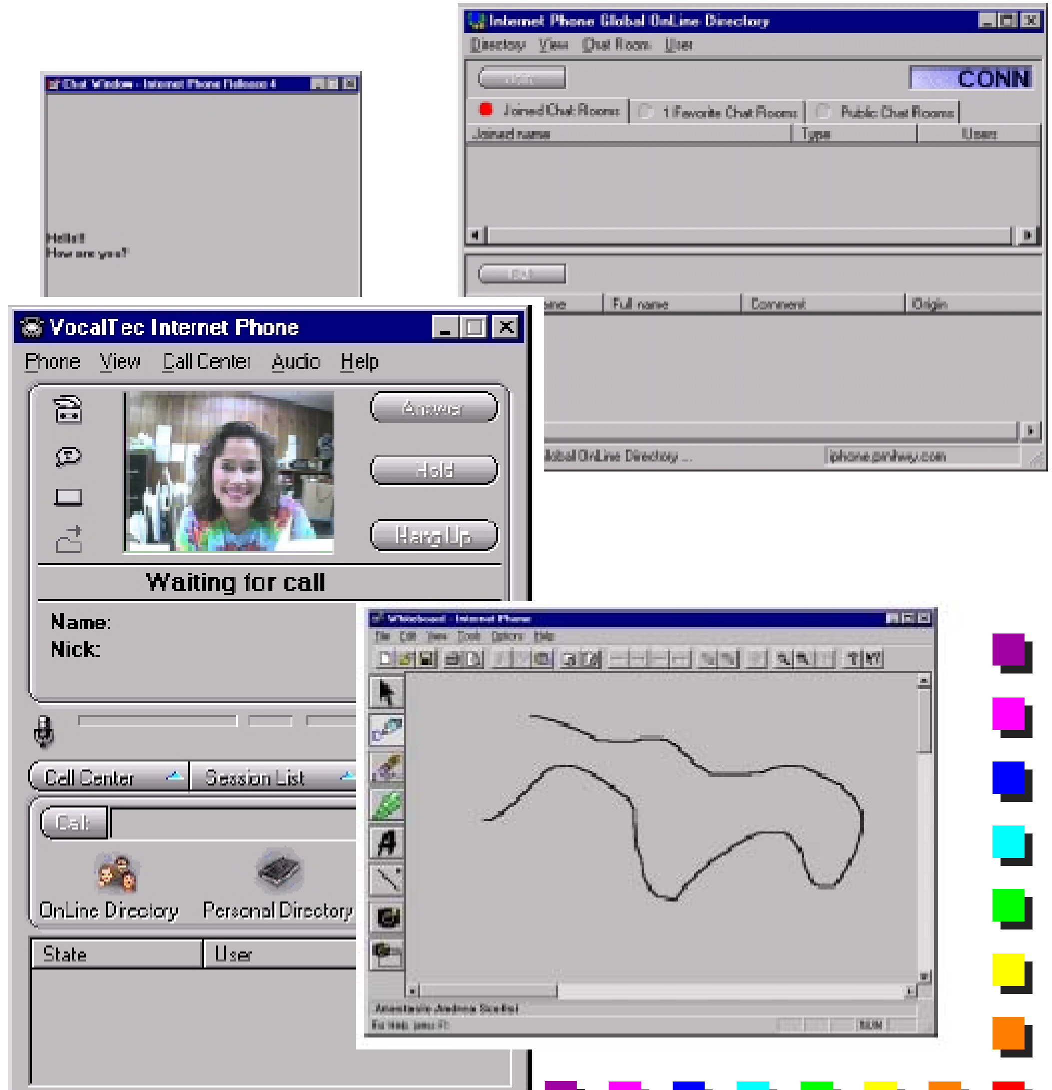
- **Solutions to previous problems**
  - Better compression
  - High quality communication
  - Pause suppression
  - Statistical multiplexing (flexible bandwidth allocation)
  - Signaling procedure (call setup)
- **New problem**
  - Quality of service management
    - Caused by lack of session setup in IP

Different perspectives for Voice over IP

- **Always the same basic technologies**
- **Different user groups have different interests in VoIP**
  - domestic user ("consumer" perspective)
  - telephone operator ("telecom" perspective)
  - corporate user ("enterprise" perspective)

VoIP "consumer" perspective

- **Phase 1**
  - Vocaltec Internet Phone, 1995
  - Microsoft NetMeeting, Microsoft Messenger
- **Phase 2**
  - domestic VoIP services



Pros and cons of the "consumer" perspective

- **Software phone**
  - Pros
    - Reduced costs
    - New services (video, white board, desktop sharing)
  - Problems
    - It is necessary to use a PC, which should be on and connected
    - Only PC-to-PC communication allowed
- **Hardware phone**
  - It is like a normal phone set, with reduced costs
    - IP phone, IP adaptor, USB phone, …
- **In both cases, mobile telephony is not considered**

The "telecom" perspective of VoIP: ToIP

- **Using IP technologies to transport phone calls**
  - PC is no longer an enabling element for VoIP
    - Traditional phone sets still used
- **VoIP**
  - Set of technologies to transport voice samples
  - Include also signaling operations
  - ToIP: set of technologies to transport voice over IP
    - They include VoIP technologies, but more is required
    - Intelligent network services
    - Integration services for integration with traditional telephone network (POTS)
      - SS#7 signaling over IP, translation between SS#7 and VoIP signaling, …

ToIP pros and cons

+ No change is required for the terminals at the network edge
+ Update required only for few devices, under operator control
- No change in user perception of the service
- No innovative services (voice/video/data integration)

Why to migrate toward ToIP?

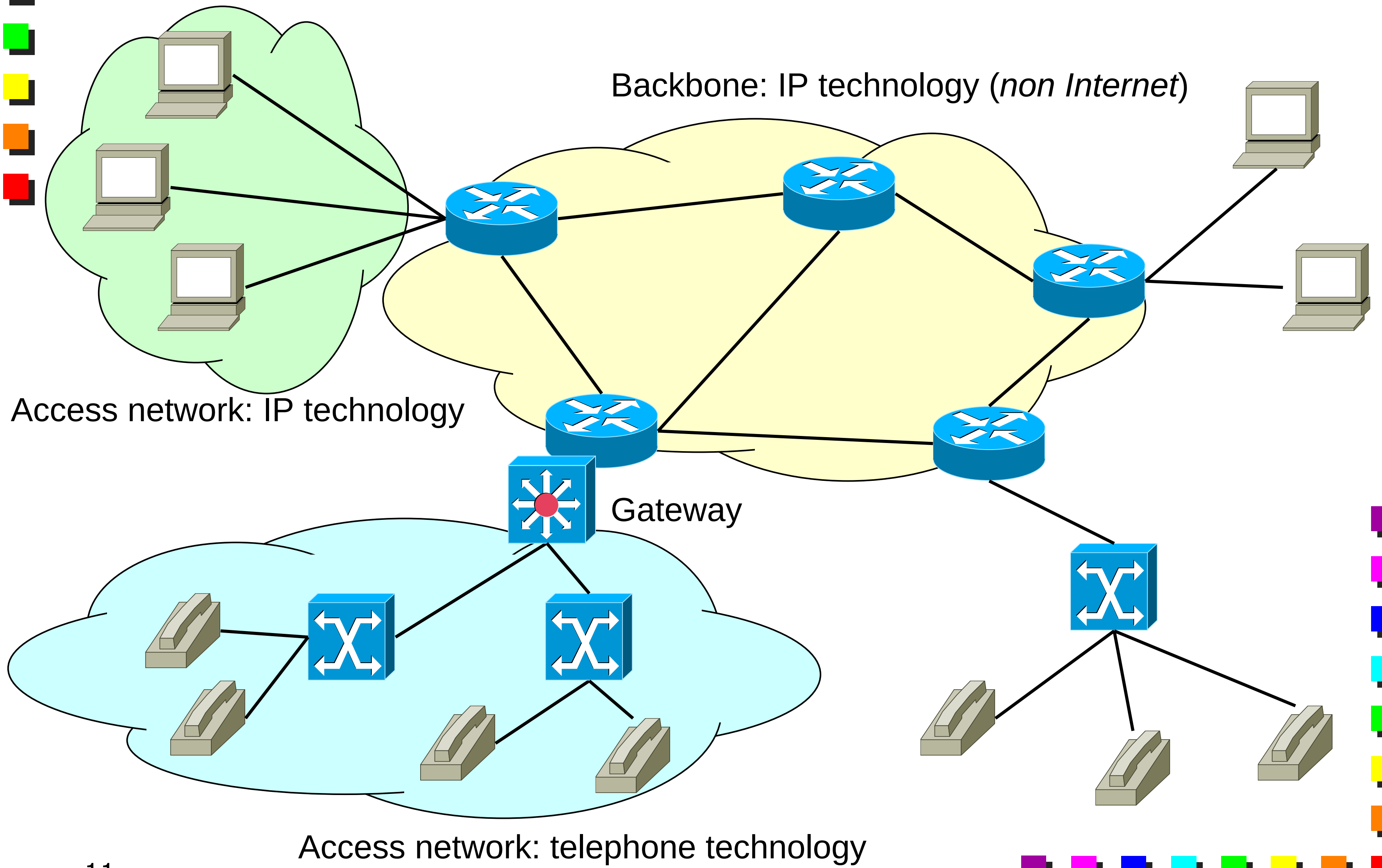
- **If ToIP does not offer innovative services, why to implement it?**
- **Economic and management issues**
  - Single network = lower costs
  - Before, telephone network used to transport all the traffic
  - Future trend: data network will transport all the traffic (including phone calls)
- **Evolution of the data network**
  - Only data, all equal
  - New applications with different requirements (delay, bandwidth,...)
    - The network should change to respond to new requirements
  - Network ready to transport non only data with different needs, but also differentiated services (multiservice network)
    - Distinct edge network for different services, same core

➔ *implementation of a single multi-service network*

Migrating toward a multi-service network

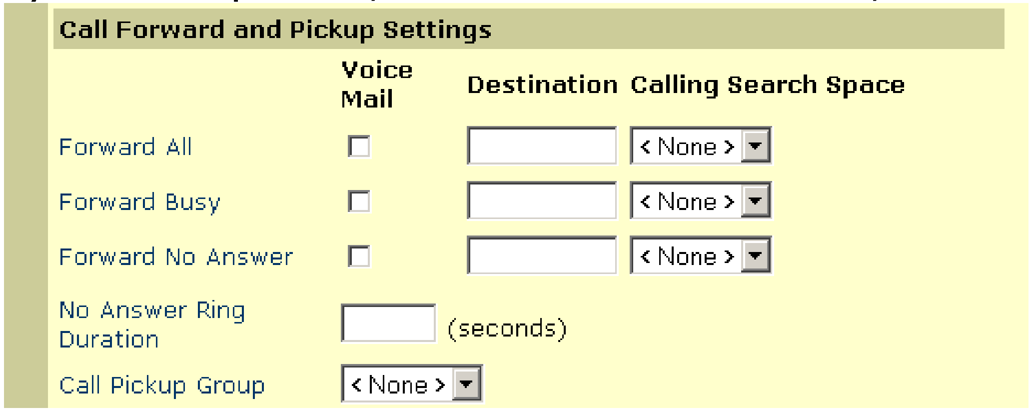
- **Often, it is immediate for new telecom operators**
- **“Tradizional” operators have more problems:**
  - Wide bandwidth in traditional telephone network is already installed
  - Personnel already trained on old technologies
  - Revenues for tlephone traffic still higher than for data traffic
  - Problems to switch to new technologies
    - Mature telephone technologies, while data technologies stil partially immature

Example of ToIP network



The "enterprise" perspective of VoIP (1)

- **Focused on value added services**
  - The economic motivation is less important
  - Integration between POTS and VoIP
- **First motivation: service personalization (often via web)**
  - Call forwarding over different channels according to several parametrs (time, caller/called identity, …)
  - Display of calls placed, chiamate unanswered, …

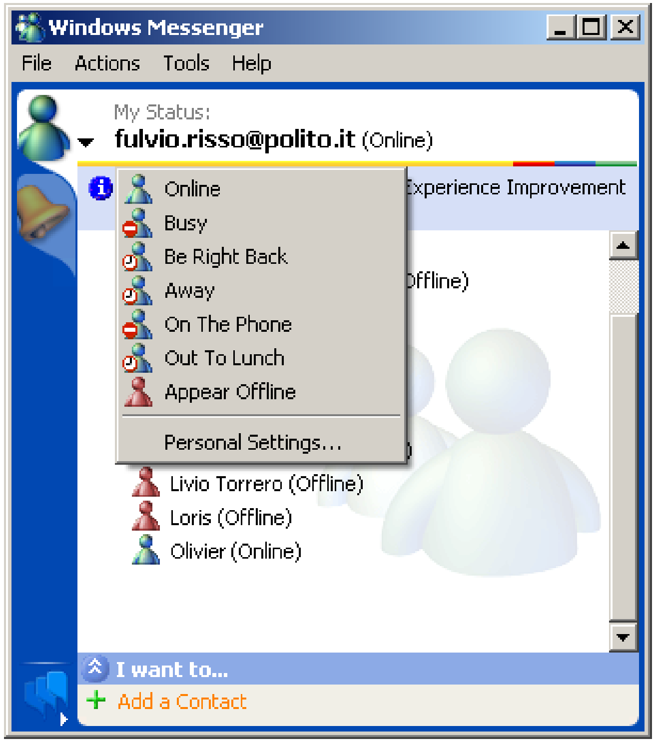


Call Forward and Pickup Settings

	Voice Mail	Destination	Calling Search Space
Forward All	☐		< None >
Forward Busy	☐		< None >
Forward No Answer	☐		< None >
No Answer Ring Duration	(seconds)		
Call Pickup Group	< None >		

The "enterprise" perspective of VoIP (2)

- **Seond motivation: integration with other applications**
  - E-presence and Instant Messaging
  - Videocalls, application sharing
  - File transfer
  - …

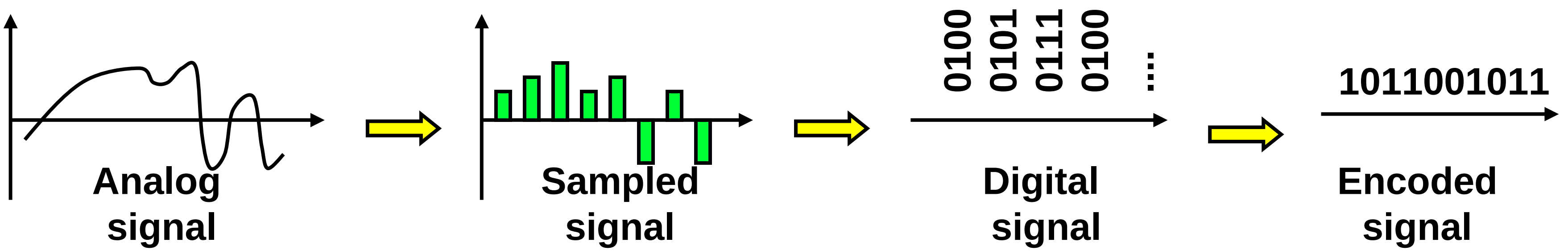


Creating a VoIP flow

- **Summarized in 9 phases**
  - Sampling
  - Encoding
  - Packetization
  - Queuing
  - Transmission
  - Propagation
  - De-jitter
  - Re-ordering
  - Decoding

Sampling and encoding

- **Sampling**
  - Digitalization of an analog signal
    - Sensibility (bit)
    - Sampling frequency (hertz)
    - Theoretical bit rate
- **Encoding**
  - Processing of digital samples
    - Compression factor
    - Actual bit rate
  - Delay is introduced (e.g. differential encoding)



Possible ancoding techniques

- **Main approaches:**
  - Differential encoding
  - Weighted encoding
  - Lossy encoding (problems with modems)
- **Pause suppression**
  - Often used in VoIP
  - The receiver introduces white noise during pauses
  - Problem: prompt recognition when the speaker resume talking
    - Loss of initial fragments of the signal
  - Several techniques can be combined together
- **Low rate does not imply low quality**
  - Aggressive Codecs may not work well with sources they are designed for (e.g. music)

Encoding problems

- **Complexity**
  - More effective techniques, more complex computations
  - Compression may be located in two places:
    - Terminal (phone set): difficult to update all
    - Gateway: large processing power (it should encode lot of conversations at once)
- **Delay, in particular for differential encoding**
  - MPEG uses differential encoding respect to both previous and following frame

Codec for telecom operators

- **Normally PCM64**
  - Works for both voice signal and other types
  - Processing poser required in terminals
- **One of VoIP promises is not fulfilled: lower bit rate**
- **Codec selection:**
  - Classical parameters: processing complexity, delay introduced, bandwidth required e quality of the encoded signal
  - "Logistic" parameters
    - Need to update terminals and computing power required in the VoIP gateway
  - Commercial parameters
    - Implement data services over the telephone network

Voice codecs

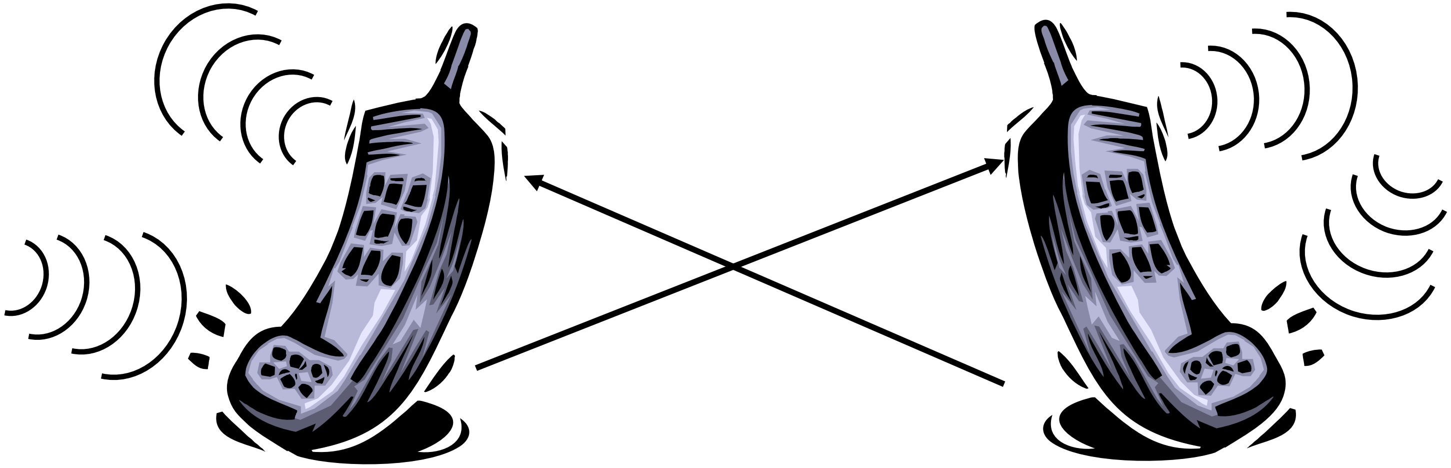
- **PCM family**
  - Standard sampling, one each 125 μs
  - G.711: 64 kbps
- **ADPCM family**
  - Adaptive encoding
  - G.726: 16 – 24 – 32 kbps
- **CELP family**
  - Interpolation encoding
  - G.728: 8 – 16 kbps
  - G.729: 8 kbps
    - CS-ACELP, very popular
- **Adaptive codecs**
  - G.723: 5.3 – 6.4 kbps
  - Very popular in PC-to-PC communication

Codec and silence suppression

- **Better transmission efficiency**
  - Conversations are normally “half duplex”
  - Pauses between syllables, words and phrases
- **Problems introduced**
  - It may be necessary to introduce artificial environmental moise, inorder to reproduce normal conditions
  - The encoder may introduce a delay in recognizing that the pause is terminated
    - Some old coders cut the first part of a word, when it was preceded by a pause

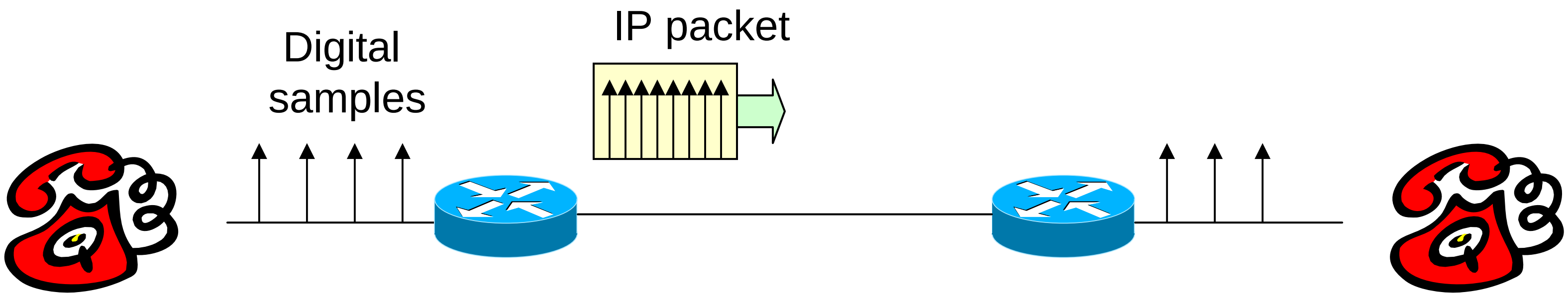
Codec and echo cancellation

- **Negligible if the round trip delay is small**
  - ~ 10 ms
- **VoIP network**
  - Delays of up 200ms (round trip)
  - Echo cancellation is required
    - Increase of the computing power



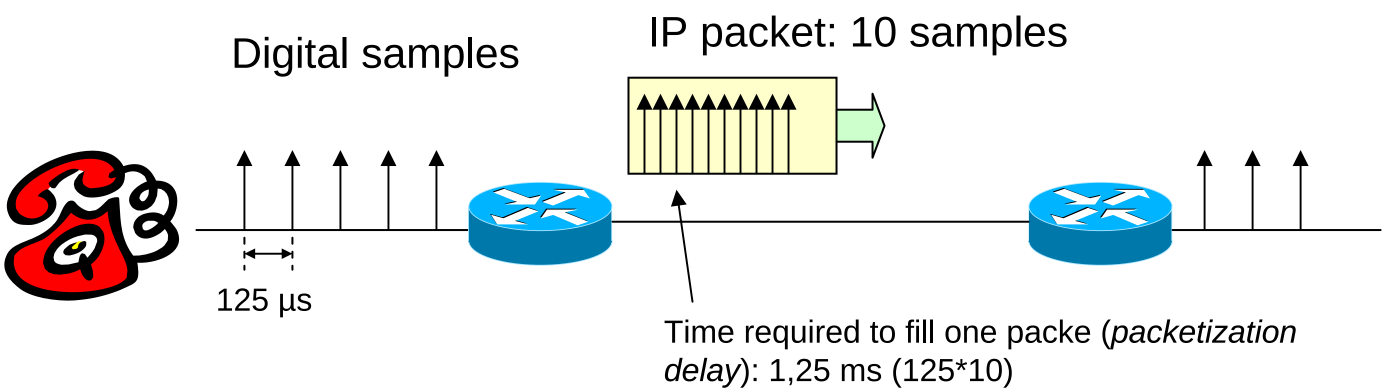
Packetization

- **First peculiar operation of a packet switched network**
- **Characteristics:**
  - Needed to lower header overheads
    - 64kbps, in 1byte/packet: 3.7Mbps!
  - Important delay introduced
  - Trade-off between delay and efficiency
    - Normal values between 20 and 40 ms



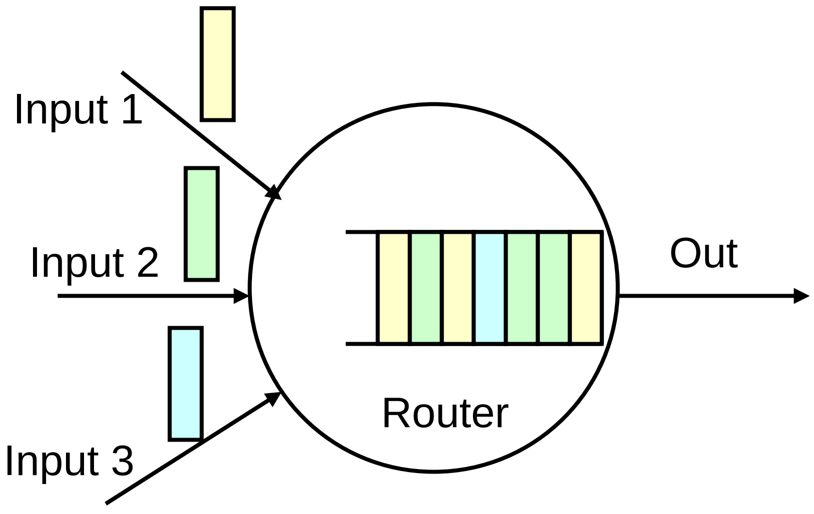
Packetization delay

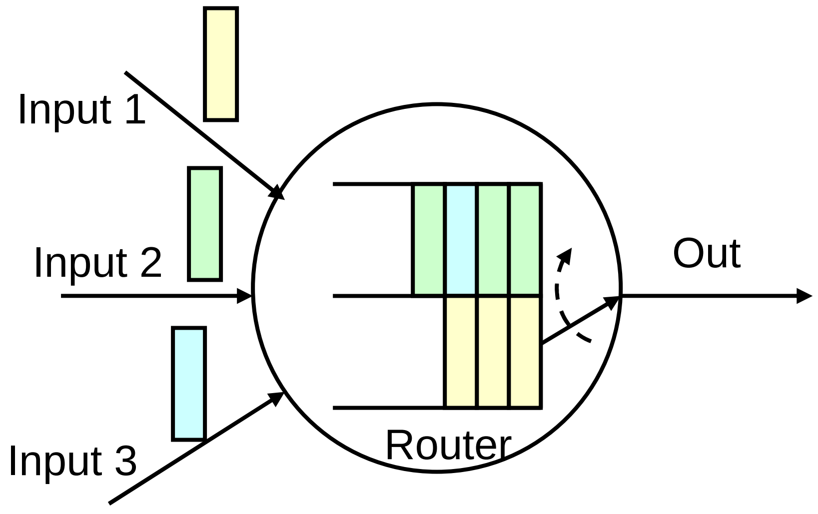
- **Packetization delay**
  - It depends on the number of samples per packet
- **Trade-off between dealy and efficiency**
  - Normal values between 20 and 40 ms



Queuing problems

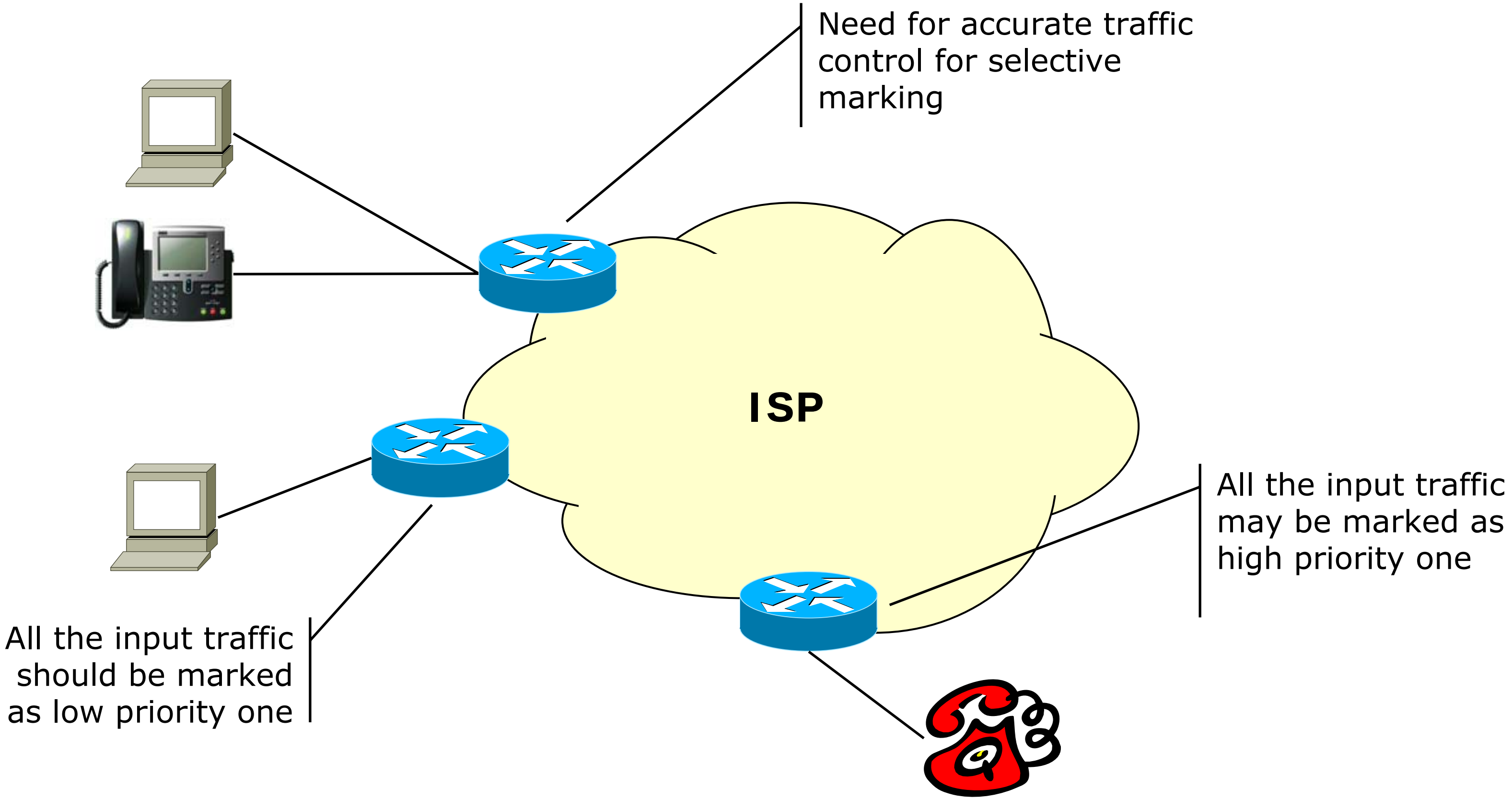
- **When input traffic is larger than the ouput link capacity (for some period of time)**
  - The router should store packets waiting for transmission (buffering)
  - Delay increases
- **Possible solution: priority queue management**





Priority queue management: marking

- **Need to control the amount of high priority traffic in the network**

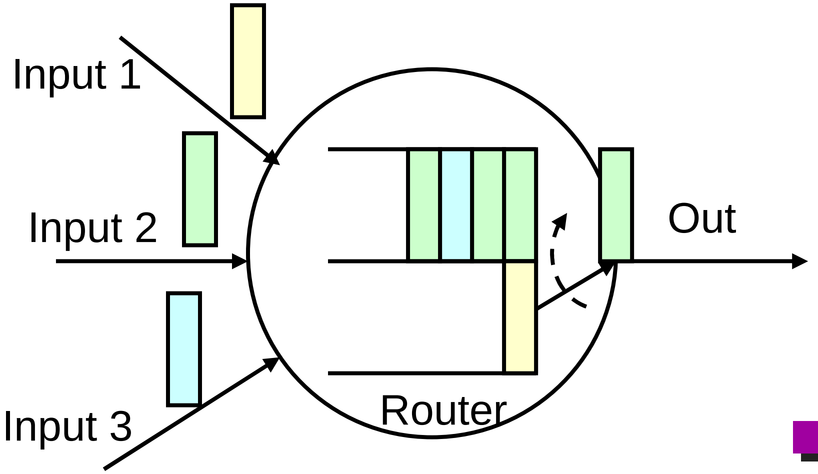


Transmission issues (1)

- **Finite size of the packets**
  - It is necessary to wait until the end of the current transmission, before starting the next one

$$T_{tx}(P) = L(P)/B + MTU/B$$

  - The time required to transmit a packet P ($T_{tx}(P)$) is proportional to its length L(P) + time required to transmit the largest packet in the network (whose size is given by MTU) (maximum time, without waiting line)

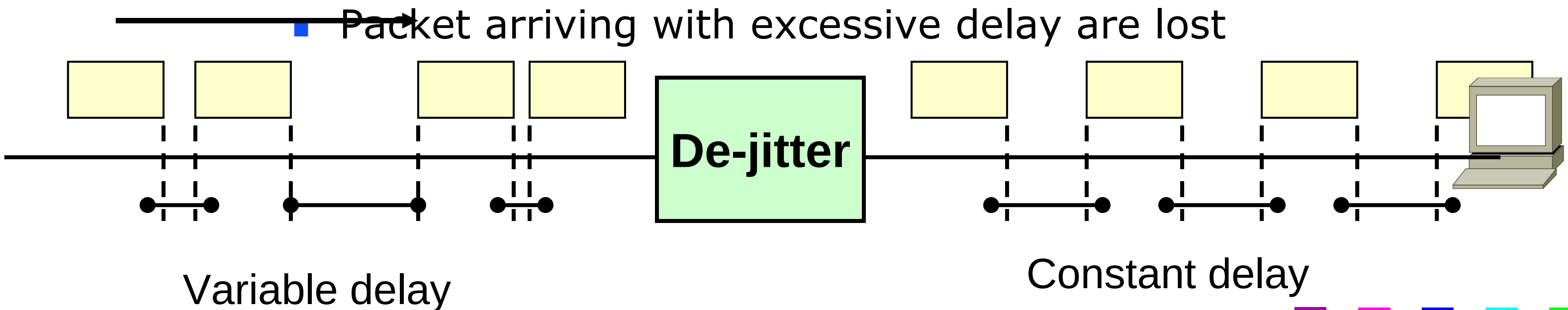


Transmission issues (2)

- **Priority Queuing**
  - Limits waiting times, but it cannot avoid transmission delays
- **Some figures**
  - ADSL (1 Mbps upload): $T_{tx,min}$= 1500/1 Mbps= 1,5 ms
  - In general, not all packets incur on this delay; however, jitter is increased
- **Solutions**
  - Use links with large bandwidth
  - PPP interleaving
  - Do not use other applications during voice calls

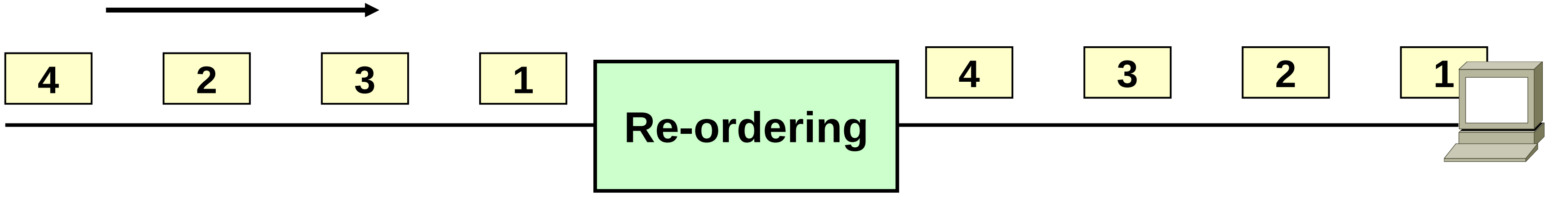
De-jitter

- **Problem**
  - Variable delay is introduced by the network for each packet
  - Voice samples in the packets should be played back at the same pace used to generate them
- **Solution**
  - De-jitter block
  - Buffer that allows the playback application to extract at constant pace the samples
  - Size: maximum jitter introduced by the network, or maximum delay allowed for one block
    - Packet arriving with excessive delay are lost



Packet re-ordering

- **The network can deliver out-of-order packets**
- **Solution**
  - The same as for de-jitter
  - Normally, the same blocks deals with both problems

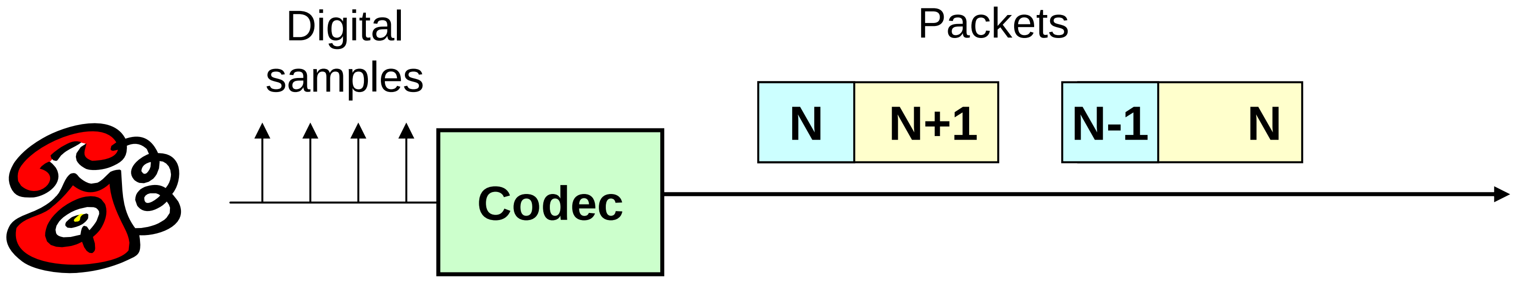


Decoding

- **Symmetric task respect to encoding**
- **Reconstruction of missing packets:**
  - Predictive techniques
  - Silence insertion
  - Replay of the samples in the last packet received
  - Some combination of the techniques listed above
- **Less complex (normally) than encoding**
  - decoding process is determined by the transmitted information
  - Encoding may require the selection between different options, to achieve better quality
  - Same delay characteristics as for encoding

Error correction techniques

- **Based on redundancy**
  - Information about sample N:
    - In the current packet, with high rate encoding
    - In the next packet, with lower rate encoding
  - Hierarchical encoding
- **Not very used, actually**
  - It is better to rely on the recovery features of the human ears



Parameters of a voice session

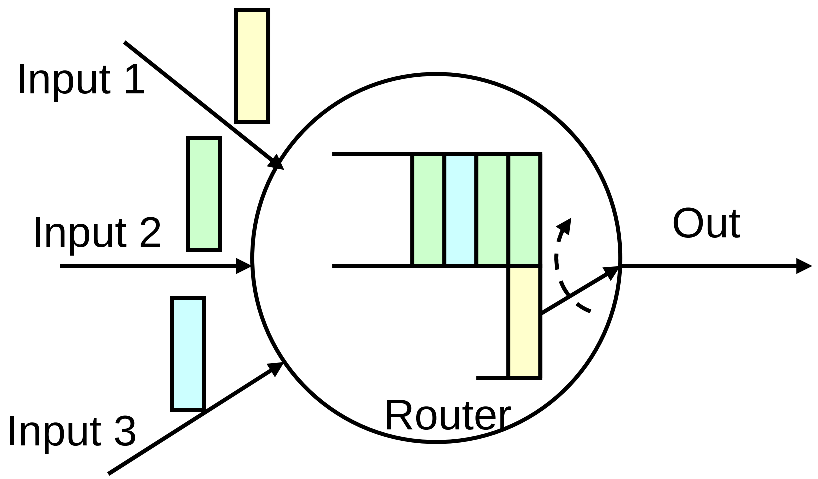
- **Delay**
  - The most important one
- **Bandwidth**
- **Loss rate**

Delay

- **Very important parameter for correct interaction**
- **End-to-end delay (reference values defined by ITU)**
  - 0 – 150 ms: acceptable
  - 150 – 400 ms: only for inter-continental calls
  - > 400 ms: not acceptable
    - *Talking overlap* harms conversation
- **Actual delay: round trip delay**

Bandwidth

- **Voice traffic: anelastic**
  - Packet flow cannot be delayed, even for short periods
  - Buffering within the network is not important
    - In the case of priority queuing, waiting line for voice packets may be very short
- **Data traffic: elastic**



Losses

- **Maximum tolerated percentage: 5%**
  - The human ear can tolerate without problems a certai number of missing packets
- **Quality of the conversation**
  - Round-trip delay is more important than data integrity
  - Re-ordering and de-jitter blocks are normally configured with reduced delay budget

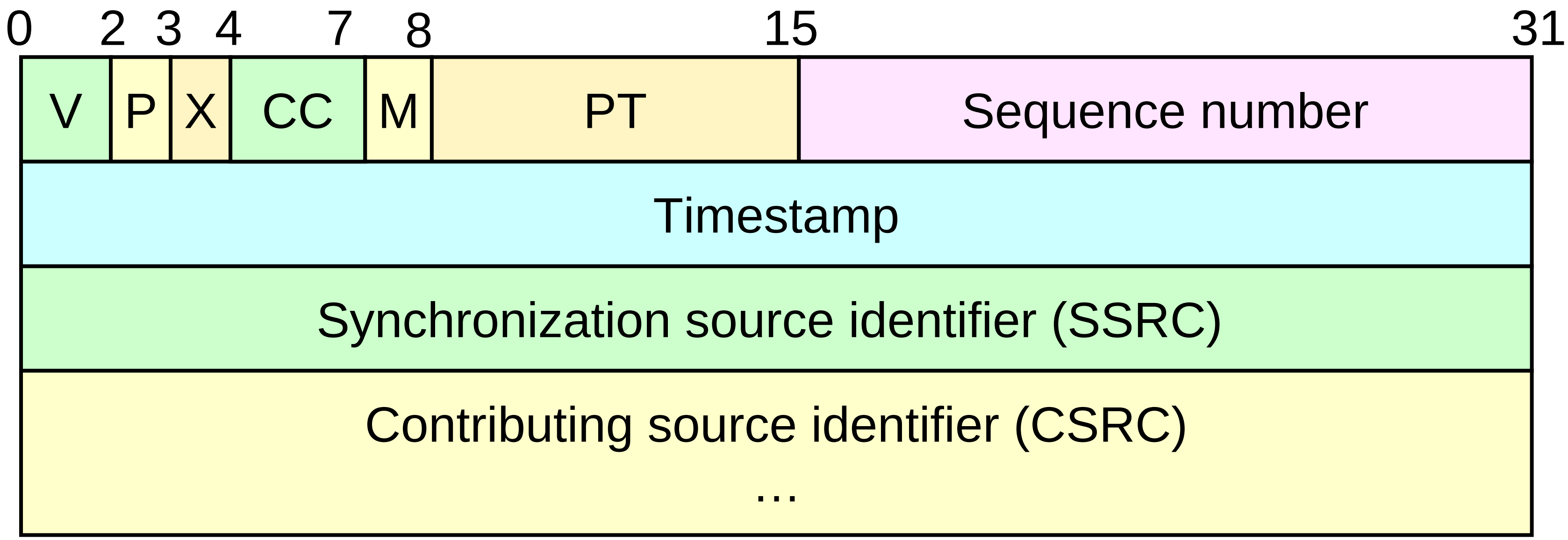
RTP (Real-Time Protocol), RFC 1889

- **General features**
  - Native multicast transmission
  - Not connected to a specific network (currently used only over IP/IPv6)
  - Packet fragmentation/re-assembly is not considered
    - It may implemented at lower layers
  - No error transmission detection (checksum)
    - If necessary, it should be provided by the underlying network
  - Data formats not specified
    - Specified in separte docuemnts (Audio Video Profiles)
    - Not connected to a specific codec
    - Able to use different “Payload Types”

RTP (2)

- **Real time data transport**
  - Packet sequencing
  - Time information (timestamp)
  - Only one flow per session
  - No lip-synch
    - It is possible to use an external block, all the required information is provided
- **RTCP (Real Time Control Protocol)**
  - Connection monitoring and control
  - Odd numbered UDP port following the one used by RTP
- **Difficult to detect (firewall, QoS)**
  - It does not sue standard ports
  - Several implementations use a static range of ports

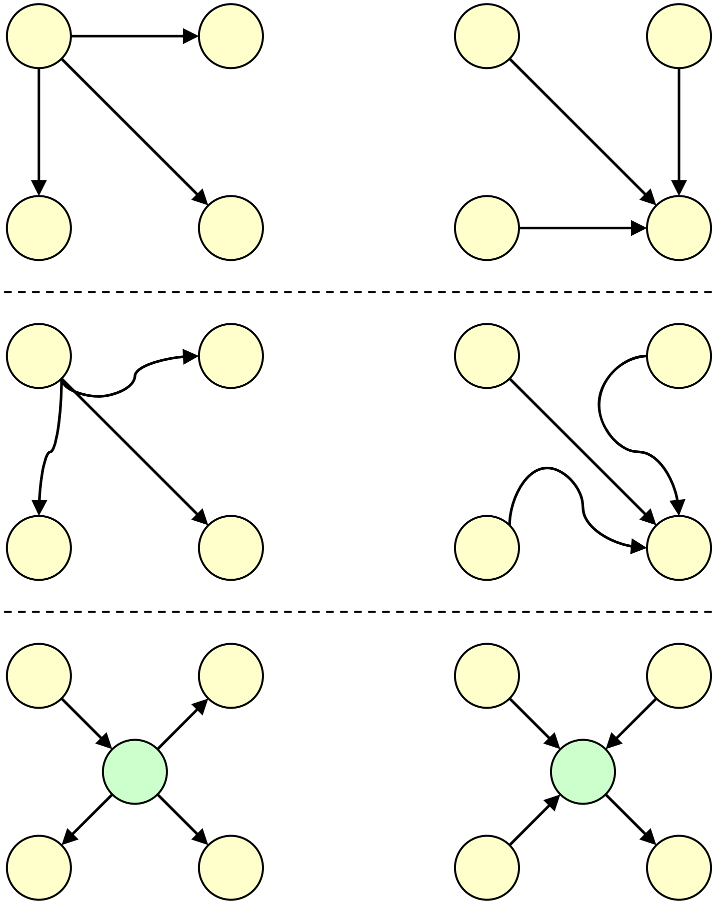
RTP packet format



RTP Mixer

- **Device able to manipulate RTP flows (e.g. mixing several flows)**
  - Transmission is transformed to a virtual hub topology
  - Useful for a session with several unicast users
  - Useful also in case of unicast/multicast users in the same session
  - The field CSRC is used to distinguish the original flows that have been merged into one
  - It is possible to do signal processing (e.g. suppression of non active audio channels)

RTP Mixer and Multicast



Unicast host:
Transmission: (N-1) flows
Reception: (N-1) flows

Multicast host:
Transmission: 1 flow
Reception: (N-1) flowsi

Unicast host with mixer:
Transmission: 1 flow
Reception: 1 flow

The mixer is always useful to save bandwidth, even when source may use multicast transmission
The processing load is not different from "traditional" case

RTP and dynamic ports

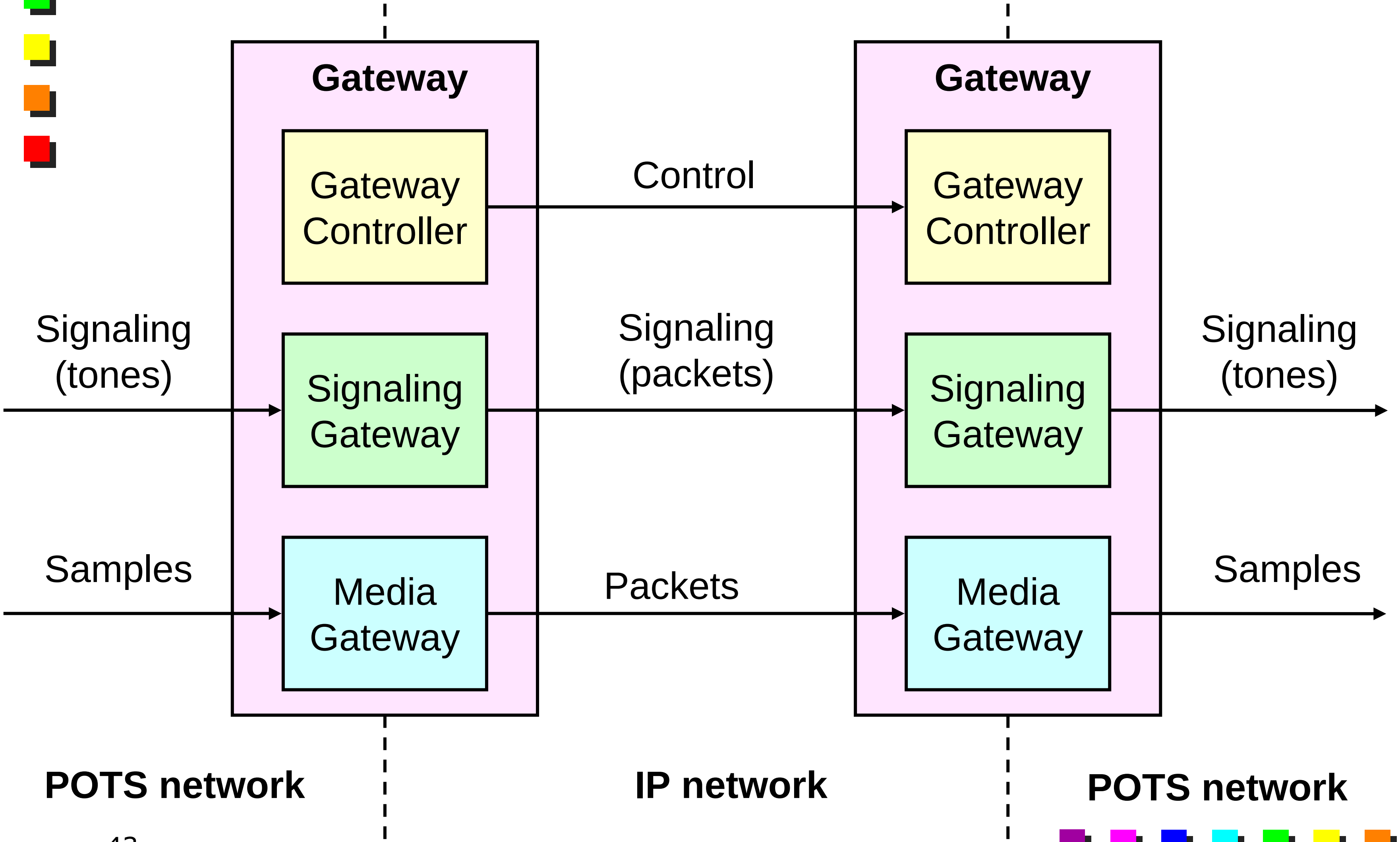
- **Each RTP session is dedicated to only ONE medium**
  - The PT field is used to discriminate among different payload types
    - It may change at each packet sent (e.g. change of codec)
    - It may convey a “neutral” code (*dynamically negotiated*)
    - Different media should use different RTP sessions
    - The number of sessions is not known a priori
    - Audio, video, white board, etc?

➔ *it is not possible to assign “well-known” ports*

Model for a VoIP network

- **Gateway between POTS and IP network**
  - Media Gateway
  - Signaling Gateway
  - Gateway Controller
- **Gateway in homogeneous networks**
- **Network architectures**
  - IP network as a backbone
  - Mixed network
  - IP network
  - IP-only network

Gateway between POTS and IP network



Media Gateway

- **Translation of the audio encoding**
  - E.g. between PCM@64kbps, popular in telephone network, and G.723@5.3kbps (and vice versa)
- **Included already in intelligent terminals**

Signaling Gateway

- **Signaling interface**
  - Dialing
  - Busy/ringing/idle tones
  - On/off-hook
  - Signaling within the network
    - Call setup with the correct end-point
  - Signaling in intelligent network
    - Call back when busy, caller ID, 3 party conversation, ...
- **The distinction between Media and Signaling Gateway is often not clear**
  - Generating busy/ringing tones: normal audio packets sent to the phone set

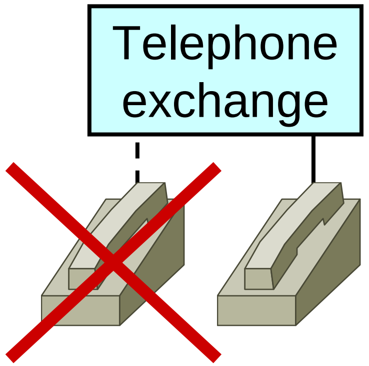
Gateway Controller

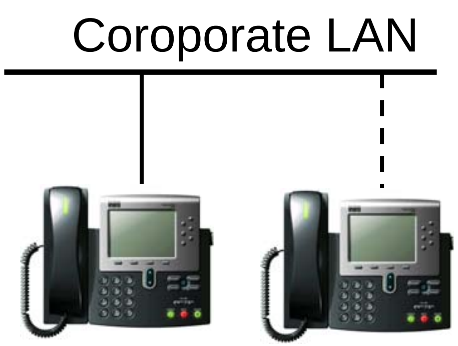
- **Supervision and monitoring of the whole gateway**
  - Control of traffic quality
    - Often, a maximum percentage of telephone traffic is allowed in a data network (otherwise the quality degrades)
  - Authorization
    - User authorized to place/receive calls
  - Authentication
    - E.g. billing to the right customer

Support server in homogeneous networks

- **Some functions cannot conglobated in the user terminal**
  - Complex functions
    - E.g. call forwarding, path preparation, etc
  - Reserved functions
    - Caller authentication/authorization
- **Gateway: still present in homogeneous networks**
  - Reduced functionalities: e.g. media gateway normally integrated in the user terminal

Impossible

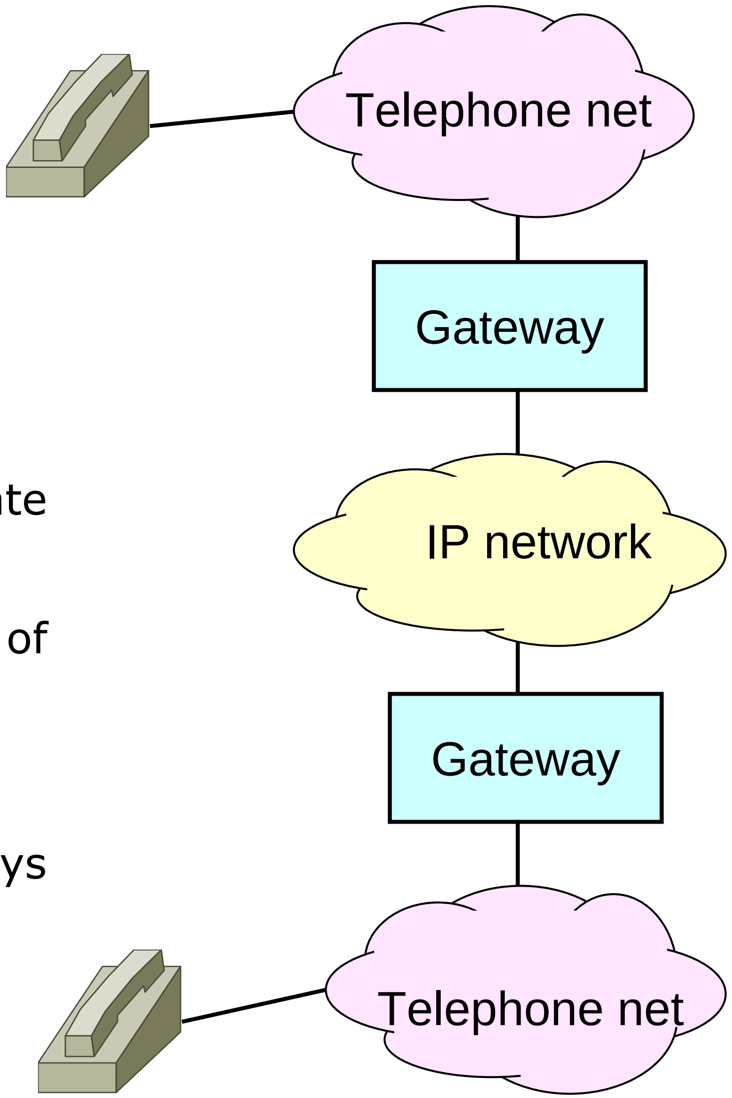




Possible

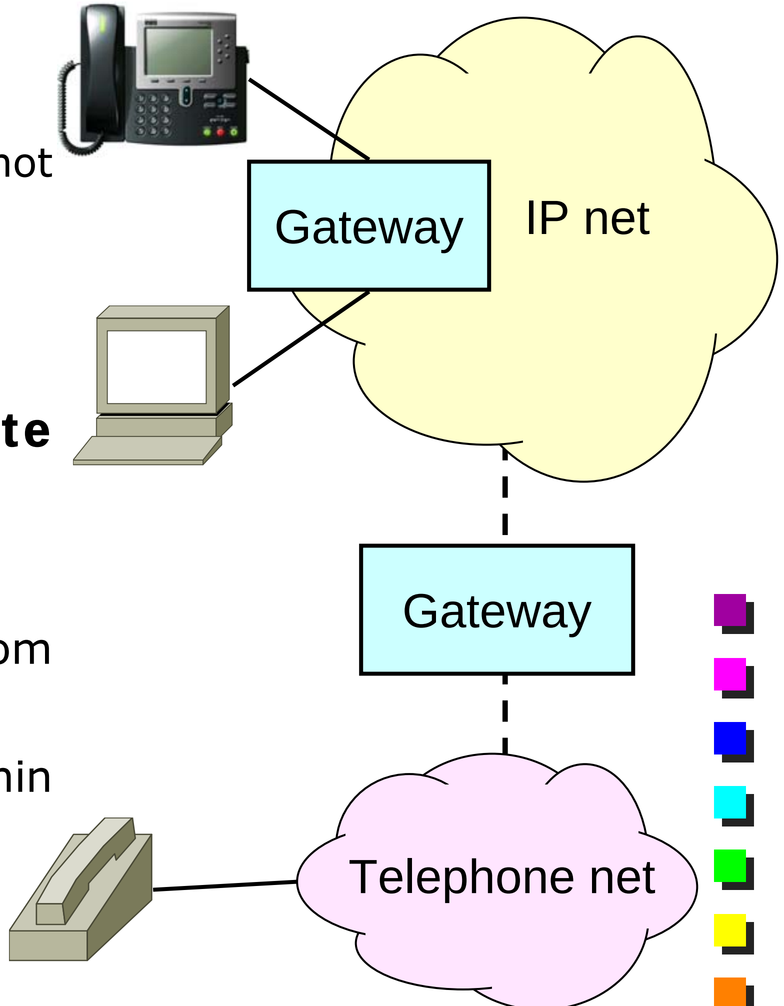
Telephone network, backbone IP

- **Traffic collection**
  - Traditional technology
- **Backbone**
  - IP technology
- **Migration process**
  - Similar to that used to migrate towards data network
  - Lower costs (smaller number of points to update)
- **Phone call**
  - Goes normally through 2 gateways (no gateway, for local calls)



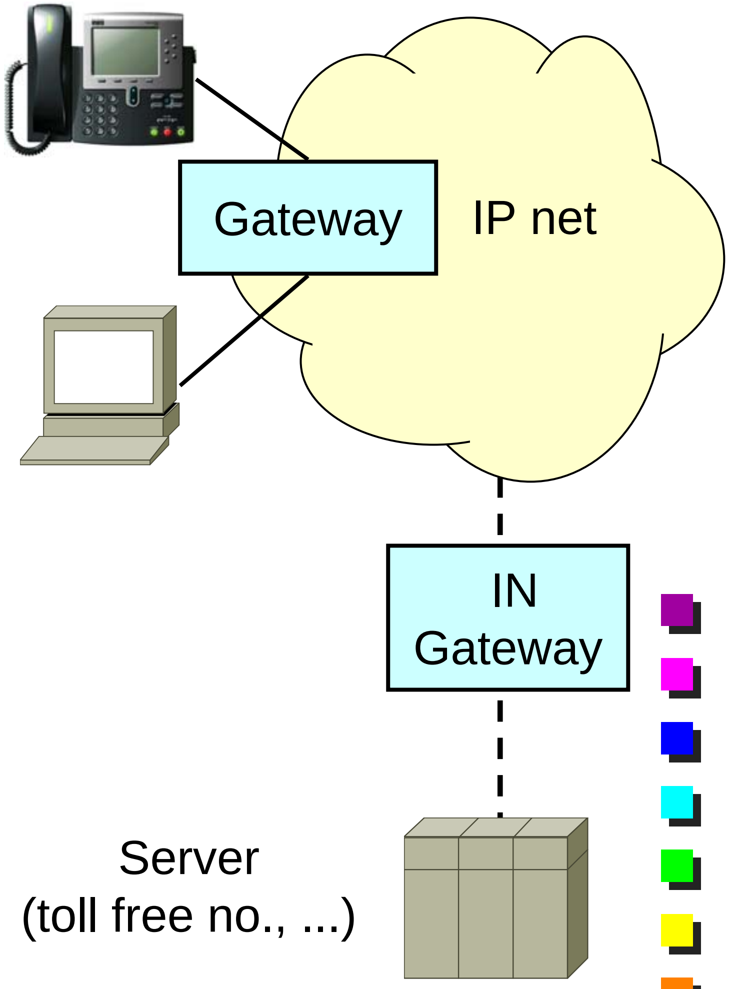
Mixed network

- **Use cases**
  - New provider
    - Pre-existent infrastructure is not available
  - Company with a new site
    - Unified data+voice network
- **Interfacing between corporate and external networks**
- **Characteristics**
  - Usually, VoIP phone set different from a PC
  - It is an example of a gateway within an IP network



IP network

- **Two successive steps**
  - Intelligent network services still with "telephone" interface
    - In particular, signaling
  - IP-only network



Most important signaling protocols

- **Goals**
  - Addressing
  - Data transport
  - Security
  - Intelligent network support
  - Simplicity and transparency
- **Main standards**
  - H.323, ITU
    - Several implementations exist
    - Complicated
      - It uses components defined for other purposes by ITU
  - SIP, IETF
    - More trendy solution