



UNIVERSITÀ
DI PARMA

DIPARTIMENTO DI SCIENZE MATEMATICHE, FISICHE ED INFORMATICHE
Corso di Laurea in Informatica

Il Livello Applicativo – Parte D

Multimedia

RETI DI CALCOLATORI - a.a. 2022/2023

Roberto Alfieri

Livello Applicativo: sommario

PARTE A

- ▶ Applicativi UDP: TFTP e DNS

PARTE B

- ▶ I servizi di posta elettronica: SMTP, POP e IMAP.

PARTE C

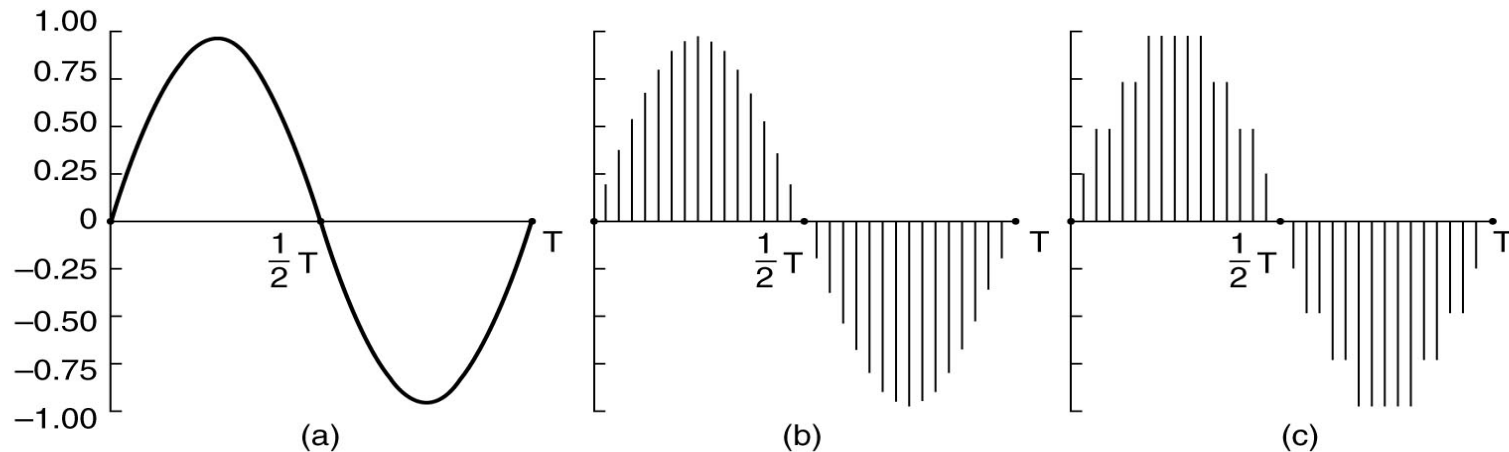
- ▶ Il World Wide Web

PARTE D

- ▶ **Multimedia**

Audio Analogico e Digitale

Onda sonora percepita dall'uomo: onda monodimensionale con frequenze comprese tra 20 e 20.000 Hz (voce sotto i 3KHz).

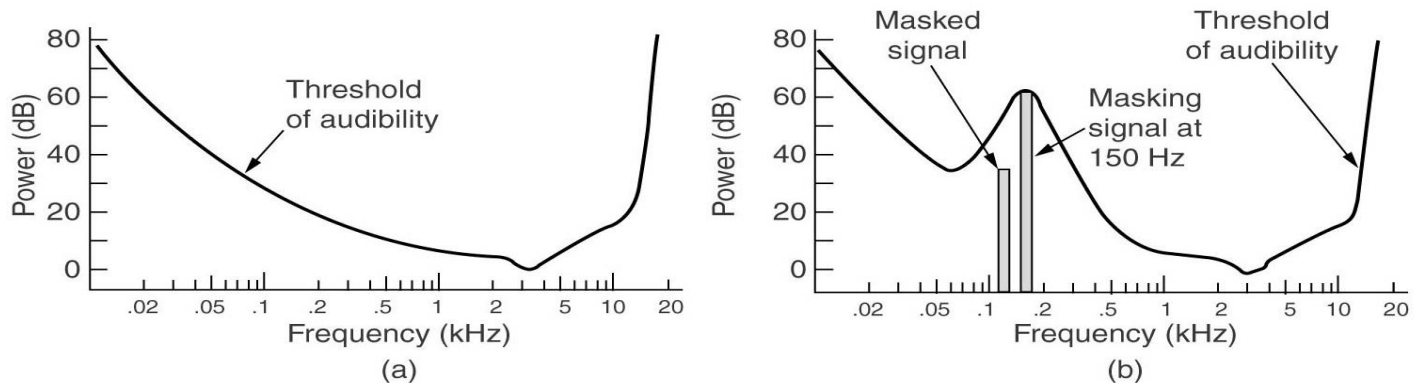


Conversione analogico/digitale:

- ▶ **La frequenza di campionamento è determinata dal teorema di Nyquist:**
- ▶ Freq. di campionamento = $2 * \text{Freq. Max del segnale da digitalizzare}$
- ▶ **quantizzazione:** numero discreto di valori possibili per le letture nel campionamento (8 bit = 256 valori, 16 bit = 65.536)
- ▶ Il sistema telefonico usa 8 bit per 8.000 volte al secondo (max 4KHz)
- ▶ CD audio: 44.100 campioni al secondo di 16 bit \Rightarrow 1,411Mbps (stereo)

Compressione Audio

- ▶ Per la trasmissione su internet è necessaria forte compressione.
- ▶ La compressione audio si può fare in 2 modi:
 - **Conversione per forma d'onda:** il segnale viene convertito, usando la trasformata di Fourier, nelle sue componenti nel dominio delle frequenze; ogni componente viene codificata con la minima quantità di bit
 - **Codifica percettiva:** sfrutta alcuni limiti del sistema uditivo umano (psicoacustica) per codificare il segnale in modo che sembri lo stesso ad un ascoltatore umano pur essendo diverso dal segnale originario, utilizzando la tecnica del mascheramento.



I formati MP3 e AAC eseguono la trasformata di Fourier, quindi codificano solo le frequenze non mascherabili.

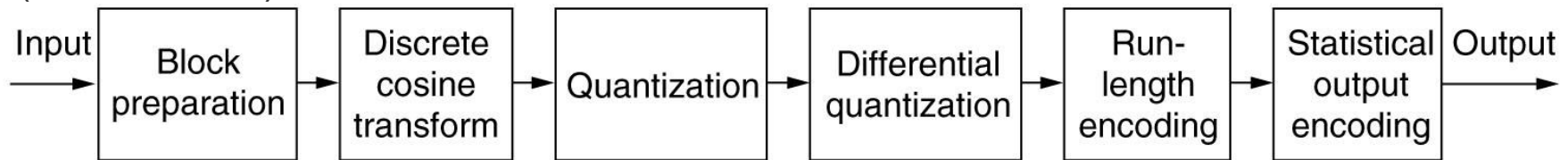
CD audio: si arriva a circa 100 Kbps (1.4Mbps non compresso, Compressione 14:1)

Video digitale

- ▶ Video digitale: sequenza di fotogrammi, ognuno dei quali composto da una griglia rettangolare di elementi detti **pixel**
- ▶ La geometria è quella del video analogico con la differenza che le linee di scansione vengono sostituite da righe di pixel discreti
- ▶ Per avere un *movimento fluido* servono almeno 25 fotogrammi al secondo
- ▶ Per un flusso video alla risoluzione di 640x480, con 24 bit per pixel e 30 fotogrammi al secondo serve una linea di comunicazione a 200 **Mbps**
- ▶ La compressione è l'unica possibilità per riuscire ad inviare filmati video su internet.
- ▶ Servono due algoritmi: uno di **codifica** per la compressione all'origine, uno di **decodifica** per la decompressione alla destinazione
- ▶ *Per i sistemi non in tempo reale* il sistema di codifica può essere lento e utilizzare hardware costoso a favore di un sistema di decodifica veloce ed economico
- ▶ *Il sistema di codifica/decodifica può non essere reversibile (lossy)*: a costo di una piccola **perdita di informazioni** si ottiene un fattore di compressione elevato

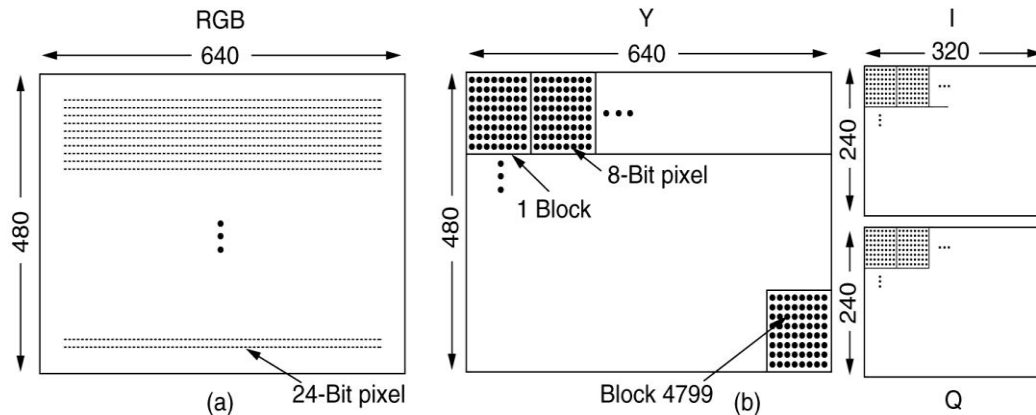
JPEG

- ▶ Un video è essenzialmente una sequenza di immagini con l'audio.
- ▶ un algoritmo per la codifica Video è la codifica in successione di ogni singola immagine, con o senza perdite di informazione
- ▶ Lo standard JPEG esegue la compressione di immagini statiche a toni continui (come le foto)



Supponiamo di lavorare con una immagine 640x480 con 24 bit per pixel

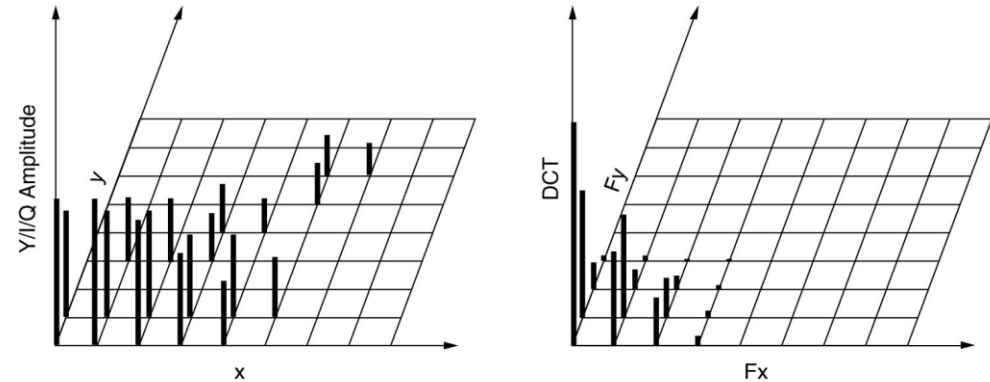
1) **Preparazione del blocco:** si passa dalle coordinate RGB in coordinate Y (luminanza), Cb e Cr (crominanza) attraverso una trasformazione lineare.



$$\begin{aligned} Y &= \text{Blu} + \text{Rosso} + \text{Verde} \\ Cb &= \text{Blu} - Y \\ Cr &= \text{Rosso} - Y \end{aligned}$$

JPEG

2) Si applica quindi una **DCT** ([Trasformazione Discreta del Coseno](#)) ad ognuno dei blocchi separatamente. DCT e' simile alla DFT (rappresenta una variabile temporale nello spazio delle frequenze), ma usa solo numeri reali.



3) **Quantizzazione:** vengono eliminati i coefficienti DCT meno importanti. Questo comporta una perdita e avviene dividendo ogni coefficiente per un peso estratto da una tabella.

Tale tabella non fa parte dello standard e ogni applicazione deve fornire i propri

DCT Coefficients

150	80	40	14	4	2	1	0
92	75	36	10	6	1	0	0
52	38	26	8	7	4	0	0
12	8	6	4	2	1	0	0
4	3	2	0	0	0	0	0
2	2	1	1	0	0	0	0
1	1	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Quantization table

1	1	2	4	8	16	32	64
1	1	2	4	8	16	32	64
2	2	2	4	8	16	32	64
4	4	4	4	8	16	32	64
8	8	8	8	8	16	32	64
16	16	16	16	16	16	32	64
32	32	32	32	32	32	32	64
64	64	64	64	64	64	64	64

Quantized coefficients

150	80	20	4	1	0	0	0
92	75	18	3	1	0	0	0
26	19	13	2	1	0	0	0
3	2	2	1	0	0	0	0
1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

JPEG

Quantizzazione differenziale: per ogni blocco, il valore (0,0) viene sostituito con la quantità di cui differisce dall'elemento corrispondente nel blocco precedente.

Linearizzazione dei 64 elementi e
applicazione della codifica **run-length**
(schema a zig-zag)

Codifica di Huffman: si assegnano ai numeri più
frequentemente codici più brevi di quelli meno frequenti

150	80	20	4	1	0	0	0
92	75	18	3	1	0	0	0
26	19	13	2	1	0	0	0
3	2	2	1	0	0	0	0
1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Fattore di compressione molto elevato: si può arrivare a rapporti di 20:1

Per la decodifica si applica l'esecuzione inversa dell'algoritmo, che richiede circa lo stesso tempo

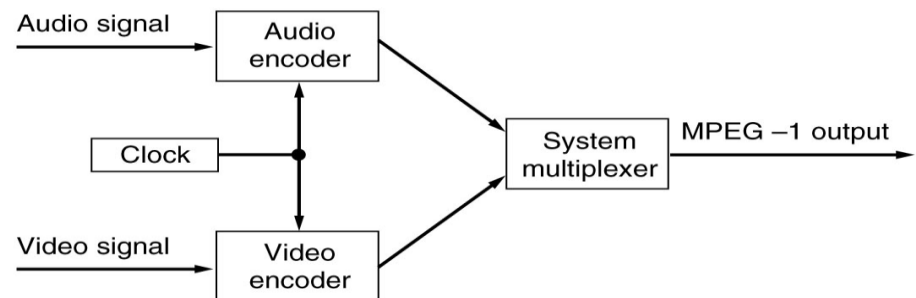
MPEG

E' un gruppo di lavoro (Motion Picture Experts Group) che dal 1993 si occupa di algoritmi e formati per le immagini in movimento.

- ▶ MPEG-1: pubblicato nel 1993, è ancora molto usato. Il suo scopo era quello di produrre un output di qualità simile a quella di un videoregistratore (352x240 per NTSC) usando un bit rate pari a 1,2Mbps con compressione 40:1 (richiederebbe circa 50,7Mbps non compressi)

- ▶ E' composto da tre parti:

- Audio
- Video
- Multiplexer di sistema

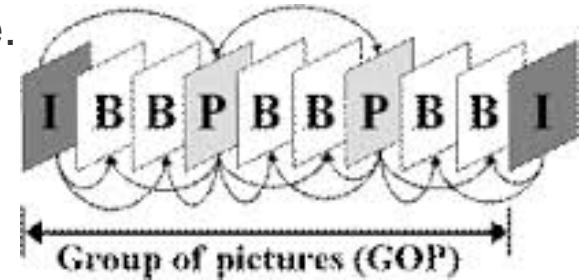


- ▶ Audio e video lavorano indipendentemente e sono sincronizzati usando un segnale comune a 90 KHz che emette il segnale del tempo corrente verso entrambi i codificatori
- ▶ **Ridondanza spaziale:** si usa la codifica JPEG per ogni singolo fotogramma. Utile soprattutto per accedere in modo casuale ad ogni singolo fotogramma
- ▶ **Ridondanza temporale:** si cerca di trarre vantaggio dal fatto che i fotogrammi consecutivi sono quasi identici

MPEG-1 e MPEG-2

MPEG-1 è composto da 3 tipi di fotogrammi:

- ▶ **I-frame** (Intracodificati): immagini statiche codificate in JPEG. Vengono inviate ad intervalli regolari (ad esempio ogni 10 frame).
- ▶ **P-frame** (Predittivi): E' una immagine che dipende dal frame precedente. Viene determinata calcolando le differenze blocco per blocco con l'ultimo fotogramma. Utile se non ci sono troppi cambiamenti rispetto alla precedente.
- ▶ **B-frame** (Bidirezionali): dipende sia dal precedente che dal successivo



MPEG-2: rilasciato nel 1996, è progettato per comprimere video con bit-rate compresi tra 4 e 6 Mbps, per essere inserito in trasmissioni NTSC e PAL per poi arrivare anche a risoluzioni superiori (HDTV)

- ▶ In prima approssimazione MPEG-2 è un superset di MPEG-1
- ▶ Utilizza fotogrammi I,P,B
- ▶ La trasformata DCT utilizza blocchi 10x10 invece di blocchi 8x8
- ▶ Supporta sia immagini progressive che interallacciate
- ▶ Supporta più livelli di risoluzione:

Low (352x240 compatibile MPEG-1), Main (720x480) High-1440 (1440x1152), High (1920x1080)

Dati multimediali in rete

Esistono diverse modalità di utilizzo del multimedia in internet, con diverse requisiti di rete:

Streaming di contenuti registrati (e.g. Youtube, Netflix)

- Numerosi flussi singoli. Riproduzione durante la ricezione.

Realtime streaming (e.g. radio Internet, dirette sportive)

- Ridurre il ritardo per minimizzare lo scostamento temporale rispetto alla trasmissione via etere.
- Centinaia o migliaia di utenti contemporanei

Conferenza in tempo reale (e.g. Skype)

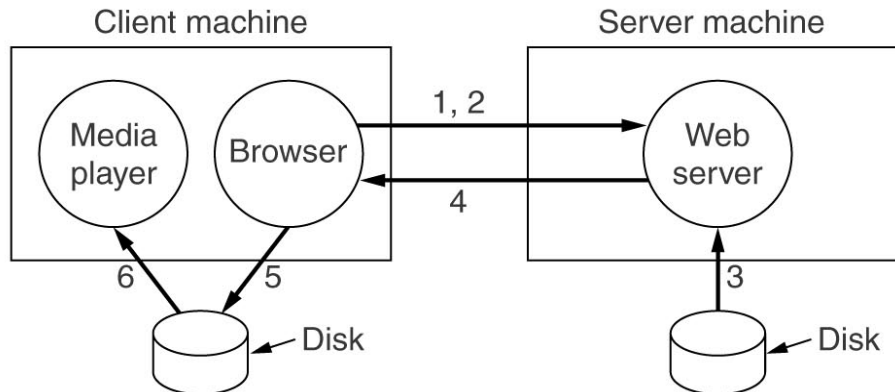
- L'interattività richiede latenze molto contenute.
- Numero di partecipanti tipicamente limitato ma in alcuni casi può essere elevato.

.

Streaming di contenuti registrati: Download and play

il file deve essere scaricato completamente prima di poter essere riprodotto utilizzando il sistema “classico” del MIME

(esempio Content-type: video/mp4) e delle helper applications



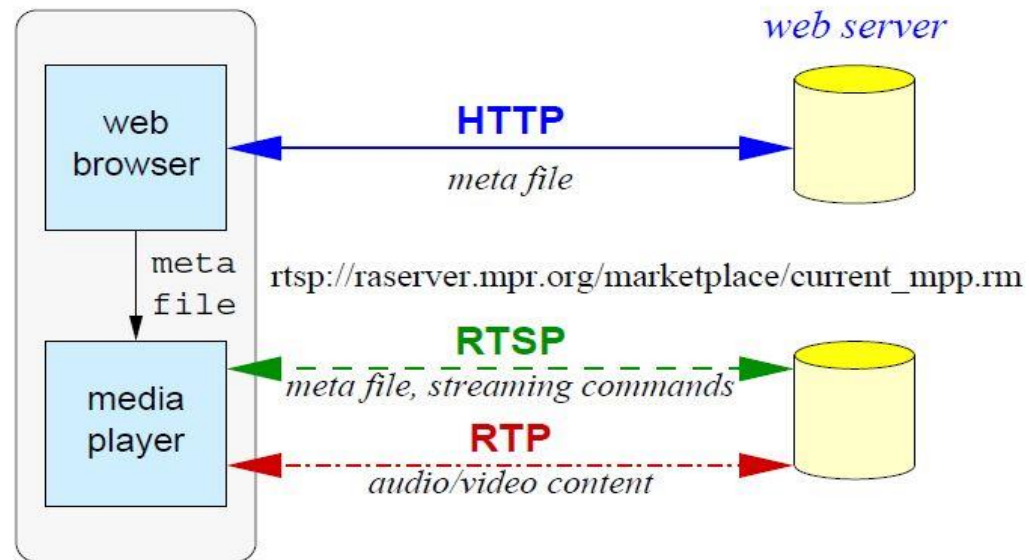
1. Establish TCP connection
2. Send HTTP GET request
3. Server gets file from disk
4. File sent back
5. Browser writes file to disk
6. Media player fetches file block by block and plays it

Sistema poco efficiente: si deve scaricare il file prima di iniziare l'ascolto.

Streaming di contenuti registrati:

Streaming on download

La soluzione generalmente adottata (Audio Streaming) è la seguente : il file collegato al titolo non è quello contenente l'audio ma un **metafile** che rimanda al file vero e proprio da ascoltare (il contenuto potrebbe essere la sola riga `rtsp://server/file.mp3`). Il browser, una volta passato il *metafile* all'applicazione esterna, non fa più parte del ciclo di comunicazione



L'accesso al file multimediale (e.g. `rtsp://server/file.mp3`) avviene mediante un protocollo per la gestione dell'interfaccia utente (Play/Record/Pause/ ecc) denominato Real Time Streaming Protocol (RTSP).

Il trasporto di flusso multimediale avviene su un canale separato basato sul protocollo RTP (Real-time Transport Protocol) o HTTP.

Real Time Streaming Protocol

- ▶ Il protocollo RTSP supporta un set di comandi che vengono inviati al server mediante un scambio testuale simile all'HTTP

Command	Server action
DESCRIBE	List media parameters
SETUP	Establish a logical channel between the player and the server
PLAY	Start sending data to the client
RECORD	Start accepting data from the client
PAUSE	Temporarily stop sending data
TEARDOWN	Release the logical channel

Vedi un esempio di dialogo: http://en.wikipedia.org/wiki/Real_Time_Streaming_Protocol

Nota: YouTube usa HTTP/TCP. RTSP è comunque utilizzata per la rete dei cellulari
<http://m.youtube.com> I video sono memorizzati nella CDN di Google.

Realtime Transport Protocol

RTP (RFC 3350) è un protocollo client/server per il trasporto di flussi anche multipli di dati audio e video. E' basato su UDP e può funzionare sia in unicast che in multicast.

I campi principali dell'intestazione RTP sono:

- Payload type: rende possibile l'identificazione del contenuto (esempi: 26 → JPEG, 32 → MPEG1, 33 → MPEG2)
- Sequence Number: numerazione progressiva per il riordino
- Timestamp: istante di campionamento del primo byte nel payload.
- Synchronized Source ID: identificatore della sorgente di stream, per distinguere diversi flussi contemporanei.

Con l'aiuto di un piccolo buffer il ricevente può ricostruire il flusso nella sequenza temporale corretta e rifasare eventuali flussi contemporanei.



RTP Header

Alcuni fornitori preferiscono implementare protocolli di trasporto brevettati.

Ad esempio i server di Realnetworks utilizzano il protocollo di trasporto [Real Data Transport \(RDT\)](#) di proprietà della RealNetworks stessa.

Realtime streaming

Requisiti:

- **Ridurre il ritardo** per minimizzare lo scostamento temporale rispetto alla trasmissione via etere.
- **Centinaia o migliaia di utenti contemporanei**

Soluzioni:

- Multicast con protocolli RTP/UDP
- Vedi [*FastWeb multicast*](#)

Buona scalabilità, ma sia multicasting che la porta RTP possono incontrare problemi nel supporto da parte dei provider.

- Dynamic adaptive streaming over HTTP, in unicast.

Dynamic Adaptive Streaming over HTTP

Dynamic Adaptive Streaming è una tecnica particolare, basata su HTTP, che adegua la qualità dello streaming video alle risorse disponibili nel dispositivo dell'utente, quali le condizioni della rete (bandwidth) o la capacità della CPU, avendo come risultato il miglioramento della QoS dell'utente.

Funziona bene se il numero di utenti è moderato. Necessario avere il server con buona connettività internet, eventualmente in CDN.

Principali tecnologie: MPEG-DASH e HLS

Vedi: https://en.wikipedia.org/wiki/Dynamic_Adaptive_Streaming_over_HTTP

<https://www.cloudflare.com/learning/video/what-is-mpeg-dash/>

Conferenze in tempo reale

Requisiti

- L'interattività richiede latenze molto contenute.
- Numero di partecipanti tipicamente limitato ma in alcuni casi può essere elevato.

.Latenza: La rete telefonica considera accettabili latenze in una direzione entro **150 ms**. Il ritardo di propagazione di una f.o. per 8000 Km (Seattle-Amsterdam) è di 40ms, a cui si aggiungono i ritardi introdotti dai router.

Ritardo dovuto al riempimento del pacchetto: un pacchetto da 1KB impiega 125ms per riempirsi a 64Kbps. Se si utilizzano pacchetti piccoli (160 bytes) è possibile ridurre la latenza, accettando un degrado di larghezza di banda.

Altra latenza è data dal software (compressione / decompressione), soprattutto per il video.

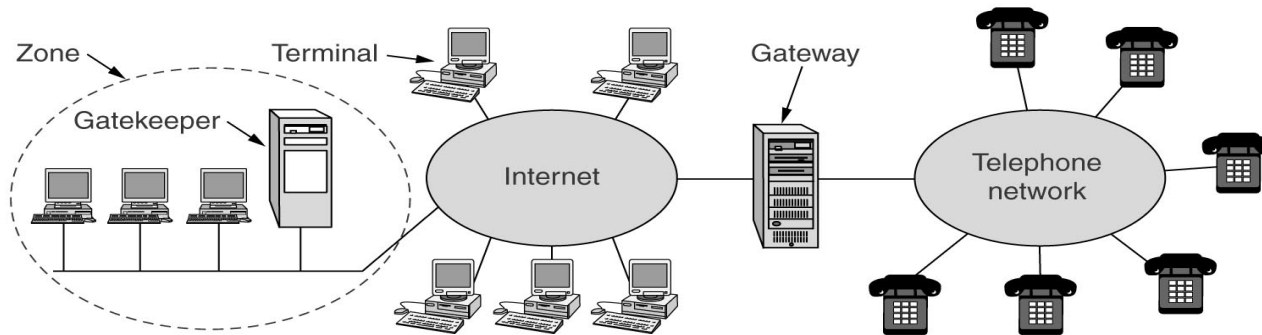
Buffer: è ancora necessario per evitare riproduzioni a scatti, ma deve essere piccolo per limitare la latenza. La presenza del buffer consente comunque la gestione dei pacchetti ricevuti fuori ordine mediante l'uso del numero di sequenza.

Protocolli: Occorre un protocollo per il trasporto dei dati, che generalmente è RTP/UDP. Il protocollo RTCP serve per lo scambio di messaggi di controllo, qualità del servizio, controllo di sessione, ecc.

VoIP con H323

H323 (ITU) rappresenta una panoramica architetture (dal punto di vista dell'industria telefonica) della telefonia Internet

Non emette proprie specifiche ma fa riferimento a diversi protocolli specializzati:
codifica del parlato, impostazione delle chiamate, segnalazione, trasporto di dati, etc.



Codifica e decodifica del parlato (almeno PCM **G.711**
 ma supporto anche di altri es. **G.723.1**)

Negoziante parametri di comunicazione: **H.245**

Protocollo per la gestione delle connessioni, per fornire i toni
 di composizione, gli squilli, etc. **Q.931**

Protocollo di comunicazione tra terminali e gatekeeper **H.225**

Speech	Control			
G.7xx	RTCP	H.225 (RAS)	Q.931 (Call signaling)	H.245 (Call control)
RTP				
UDP			TCP	
IP				
Data link protocol				
Physical layer protocol				

VoIP con SIP

SIP (Session Initiation Protocol) è la risposta di IETF (RFC 3261) ad H.323, considerato un prodotto tipico per telecomunicazioni (grande, complesso e poco flessibile), con un protocollo più semplice e modulare per la telefonia via Internet

Descrive come impostare le telefonate via internet, le videoconferenze e altre connessioni multimediali. Gestisce solamente l'impostazione, la gestione e la terminazione delle sessioni.

Altri protocolli si occupano del trasporto dei dati. Per il traffico multimediale si usa di preferenza RTP (RFC 3550, RFC 5506)

Riferimenti:

http://www.garr.it/eventiGARR/ws9/pdf/Sommani_pres.pdf

http://wiki.ekiga.org/index.php/Technical_details

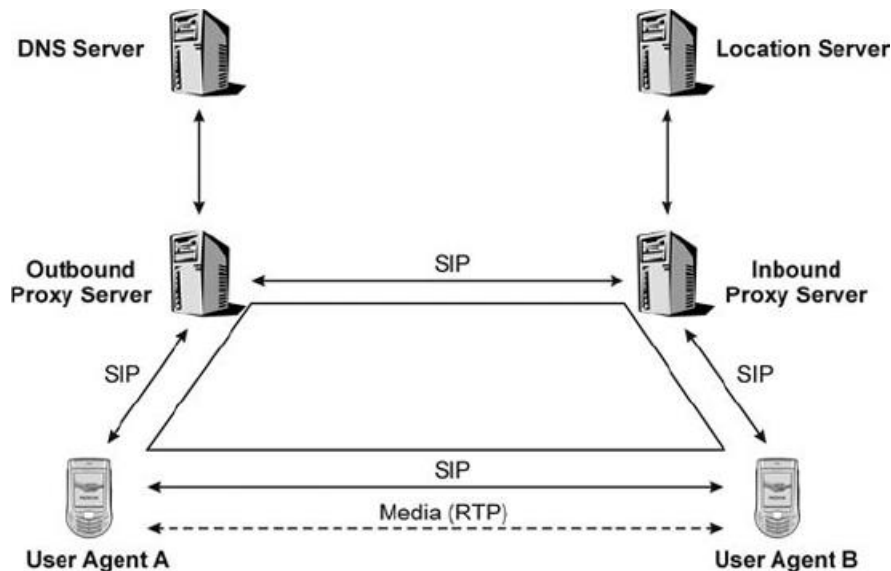
Architettura SIP

Principali elementi:

User Agent: è un end-point e può fungere da client o da server

Proxy server: è un server intermedio; può rispondere direttamente alle richieste oppure inoltrarle ad un client, ad un server o ad un ulteriore proxy. Un proxy server analizza i parametri di instradamento dei messaggi e "nasconde" la reale posizione del destinatario del messaggio - essendo quest'ultimo indirizzabile con un nome convenzionale del dominio di appartenenza.

Location Server: è un database contenente informazioni sull'utente, come il profilo, l'indirizzo IP, l'URL.



Protocollo SIP

Un SIP-URI (RFC 3261) rappresenta lo schema di indirizzamento SIP per chiamare un altro soggetto attraverso il protocollo SIP. In altre parole, un SIP URI è il recapito telefonico SIP di un utente. Il SIP URI assomiglia ad un indirizzo e-mail scritto nel seguente formato: {sip|sips}:[user-part@]domain-part[:port]

Esempio: *sip:alfierir@ekiga.net* Queste URI possono contenere indirizzi IPV4, IPV6 o numeri di telefono veri e propri: sips:+004437612234@sip-proxy.org:5062

Il protocollo è basato su UDP/5060 con transazioni richiesta/risposta in ASCII (simile ad HTTP). Una transazione inizia con una Request inviata da uno User Agent Client ad un Proxy e termina con una Final Response inviata in senso inverso.

Method	Description
INVITE	Request initiation of a session
ACK	Confirm that a session has been initiated
BYE	Request termination of a session
OPTIONS	Query a host about its capabilities
CANCEL	Cancel a pending request
REGISTER	Inform a redirection server about the user's current location

