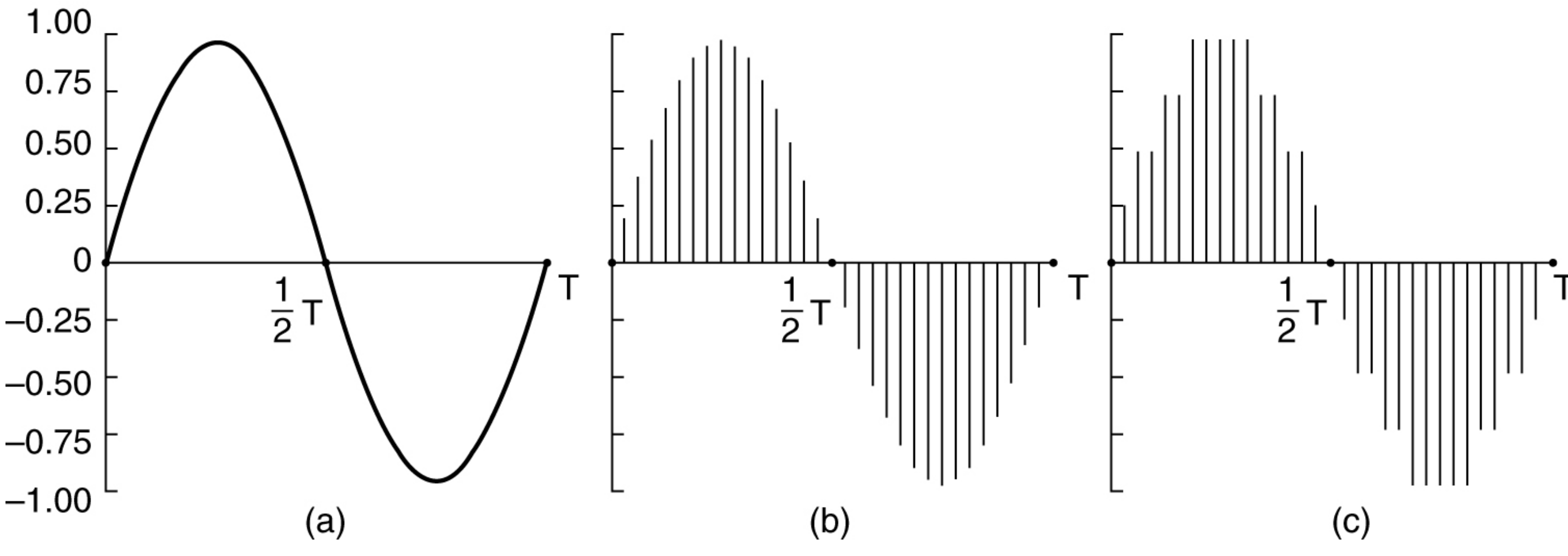
Introduction to LAN/WAN

Application Layer 4

Multimedia

- Multimedia: Audio + video
- Human ear: 20Hz – 20kHz, Dogs hear higher freqs
- DAC converts audio waves to digital
- E.g PCM uses 8-bit samples 8000 times per sec



Audio Compression

- Audio CDs:
  - 44,100 samples/sec (22 kHz)
  - Require 1.411 Mbps to transmit real-time
- Audio compression
  - reduces bandwidth,
  - Internet transmission more practical
  - MP3 (MPEG audio layer 3) most powerful, best known

Audio Compression

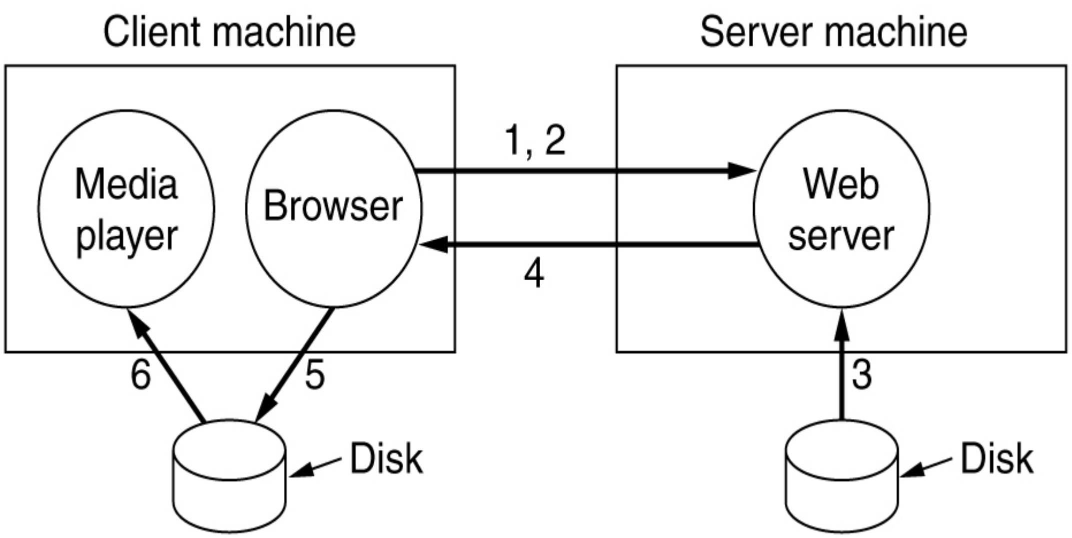
- Two types of audio compression
  - ***Waveform coding:*** fourier transform, then encode frequency components
  - Perceptual coding:
    - Exploit flaws of human ear
    - Encoded form is different but **sounds** same to human ear
- MP3 based on perceptual coding

MP3

- ***Frequency masking:*** Sound in freq A can mask softer sound in freq B (Suppress freq B sound)
- ***Temporal masking:*** Soft sound B is still not heard for few seconds even after freq A sound stops

Streaming Audio

- Naïve implementation: download entire mp3 file, play
- Disadvantage: Latency! All of file must be downloaded before it can be played



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

Streaming Audio

- Link *metafile* (just name) not actual mp3 file
  rtsp://joes-audio-server/song-0025.mp3
- Browser writes name to disk, launches media player as helper app
- Media player receives streamed mp3 file, browser not involved
- RTSP (Real Time Streaming Protocol) used, not HTTP

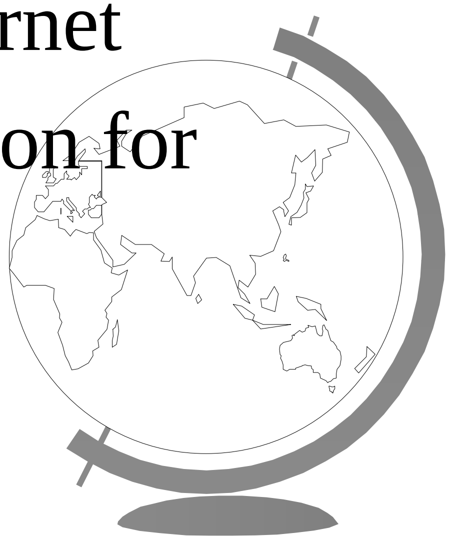
Media Player

☞ Four major functions

- Manage user interface: skins, user choices
- Transmission errors: User RTP, interleaving
- Decompress music
- Eliminate jitter: Playback buffer,
  - ◆ Pre-download 10-15 secs of music
  - ◆ Try to download new blocks at same rate as playback

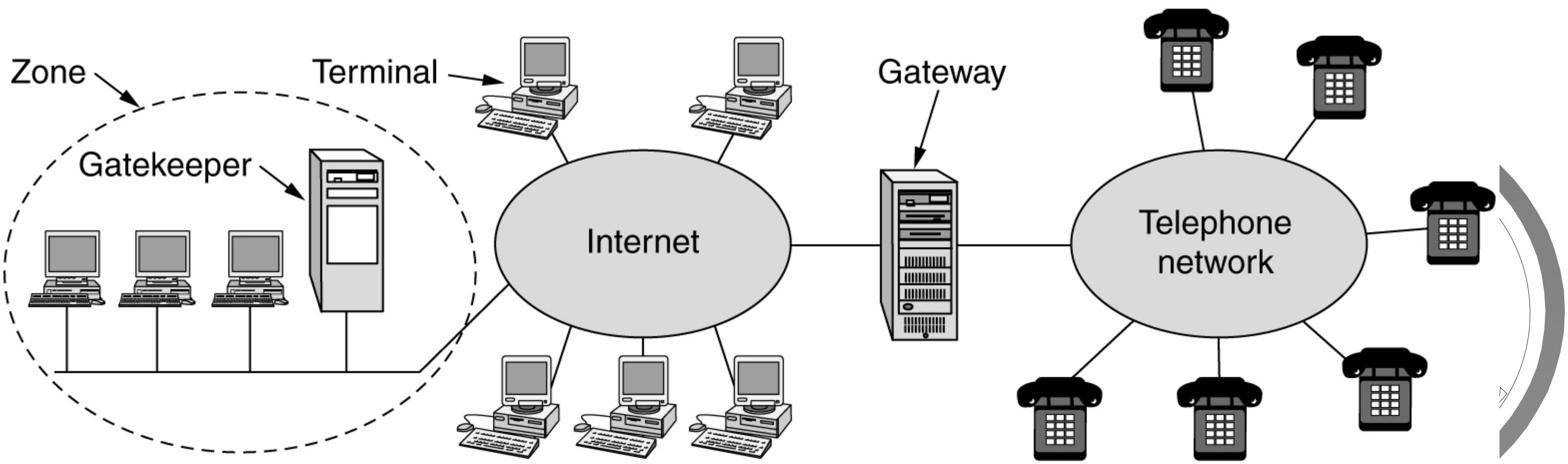
Voice over IP

- Telephone network initially carried only voice
- Data traffic grew and equalled voice by 1999
- Much more data than voice by 2002
- Large data numbers made packet-switched network operators interested in voice
- Av person has higher phone bill than Internet
- Data network operators can easily provision for voice and make more money



H.323

- H.323: VoIP standard, an architectural overview
- References specific protocols for speech coding, call setup, signalling, data transport, etc
- Features gateway, terminals, gatekeeper & zones



The H323 architectural model for Internet telephony.

Session Initiation Protocol (SIP)

- H.323 created by telcos: large, complex
- IETF set up committee to design better VoIP standard
- SIP resulted, RFC 3261
  - How to setup Internet phone calls, video conferences
  - Unlike H.323 single module, not protocol suite
  - Integrates well with existing Internet applications
  - Phone numbers are URI, on web page same as *mailto*

SIP

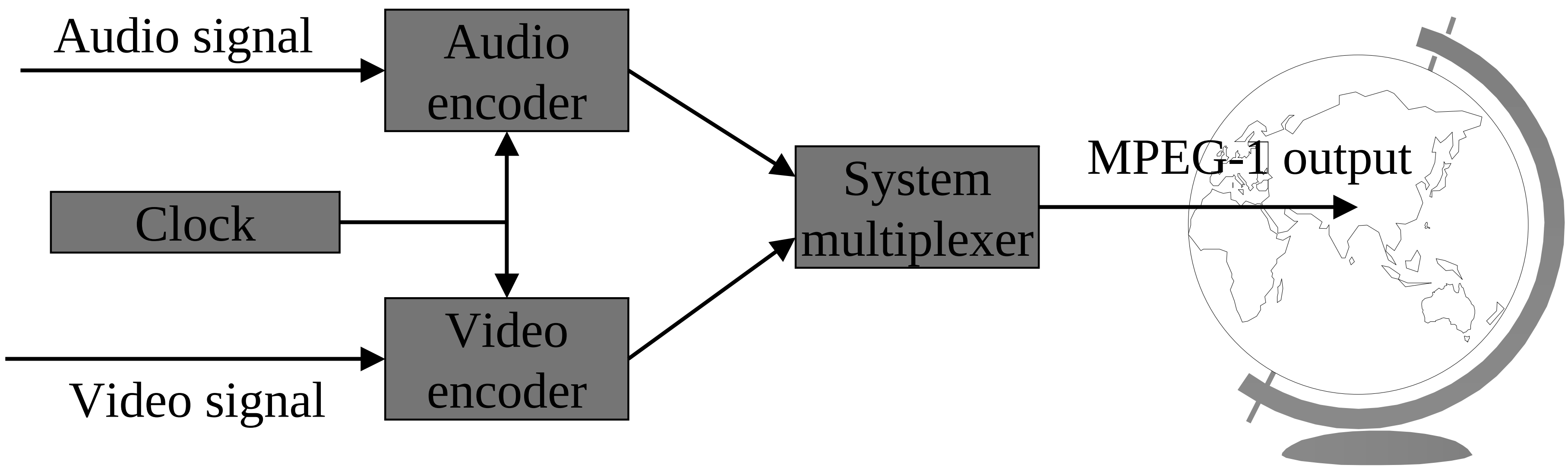
- ☞ Application layer protocol, can use UDP or TCP
- ☞ Text-based, modelled on HTTP
- ☞ Can establish 2-party, multparty calls
- ☞ Sessions may contain audio, video and data
- ☞ Services
  - – Locate callee, may be away from home machine
  - – Determine callee’s capabilities
  - – Handles call setup and termination mechanics

MPEG-1

- MPEG (Motion Picture Experts Group) standard for compressing video files since 1993
- Movies contain sound: MPEG can compress both audio and video
- Different generations of MPEG
- MPEG-1:
  - Goal: video-recorder quality (352 x 240 for NTSC) using a bit rate of 1.2Mbps
  - Uncompressed at 24 bits per pixel requires 50.7 Mbps
  - Compression ratio of 40 required to reduce to 1.2 Mbps
- Notes:
  - NTSC is video standard in US
  - PAL is standard in Europe

MPEG

- MPEG-2:
  - designed for compressing broadcast-quality video into 4-6 Mbps (to fit into NTSC and PAL broadcast)
  - Also forms basis for DVD and digital satellite TV
- MPEG-1 and 2 are similar: MPEG-2 almost superset of MPEG-1
- MPEG-1: audio and video streams encoded separately, uses same 90-KHz clock for synchronization purposes



MPEG

- Compression techniques usually take out redundancies
- MPEG compresses using **spatial** and **temporal** redundancies in movies
- Think of streaming movie as sequence of still (JPEG) images
- Spatial coherency is redundancy within 1 still image (each JPEG)
- Temporal redundancy
  - exploits the fact that consecutive frames are almost identical
  - reduced in new scenes in a movie, etc
  - Increased for slow-moving objects, stationary camera/background
- Every run of 75 similar concurrent frames can be compressed

MPEG

- MPEG-1 output consists of four kinds of frames:
  - **I (Intracoded)** frames:
    - self-contained JPEG-encoded still pictures
    - Act as reference, in case packets have errors, are lost or stream fast forwarded, etc
  - **P (Predictive)** frames:
    - Block-by-block difference with last frame
    - Encodes differences between this block and last frame
  - **B (Bi-directional)** frames:
    - Difference between the last or next frame
    - Similar to P frames, but can use either previous or next frame as reference
  - **D (DC-coded)** frames:
    - Encodes average values of entire block
    - Allows low-res image to be displayed on fast-forward

MPEG

- MPEG-2:
  - I, P, B frames supported
  - D frames NOT supported
  - Supports both progressive and interlaced images
  - Encodes smaller blocks to improve output
  - Also supports multiple resolutions

Mobile MPEG

- Mobile multimedia apps: **indoor** or **outdoor**
- Indoor applications have low mobility, high bandwidth (e.g. on WPI wireless LAN)
- Outdoor applications have higher mobility, low bandwidth (e.g. on Sprint PCS cellular network)
- Conflict:
  - Low bandwidths argue for more efficient encoding/compression, less redundancy
  - High wireless error argue for more redundancy to recover
- Conclusion: careful with what redundancy you take out

MPEG

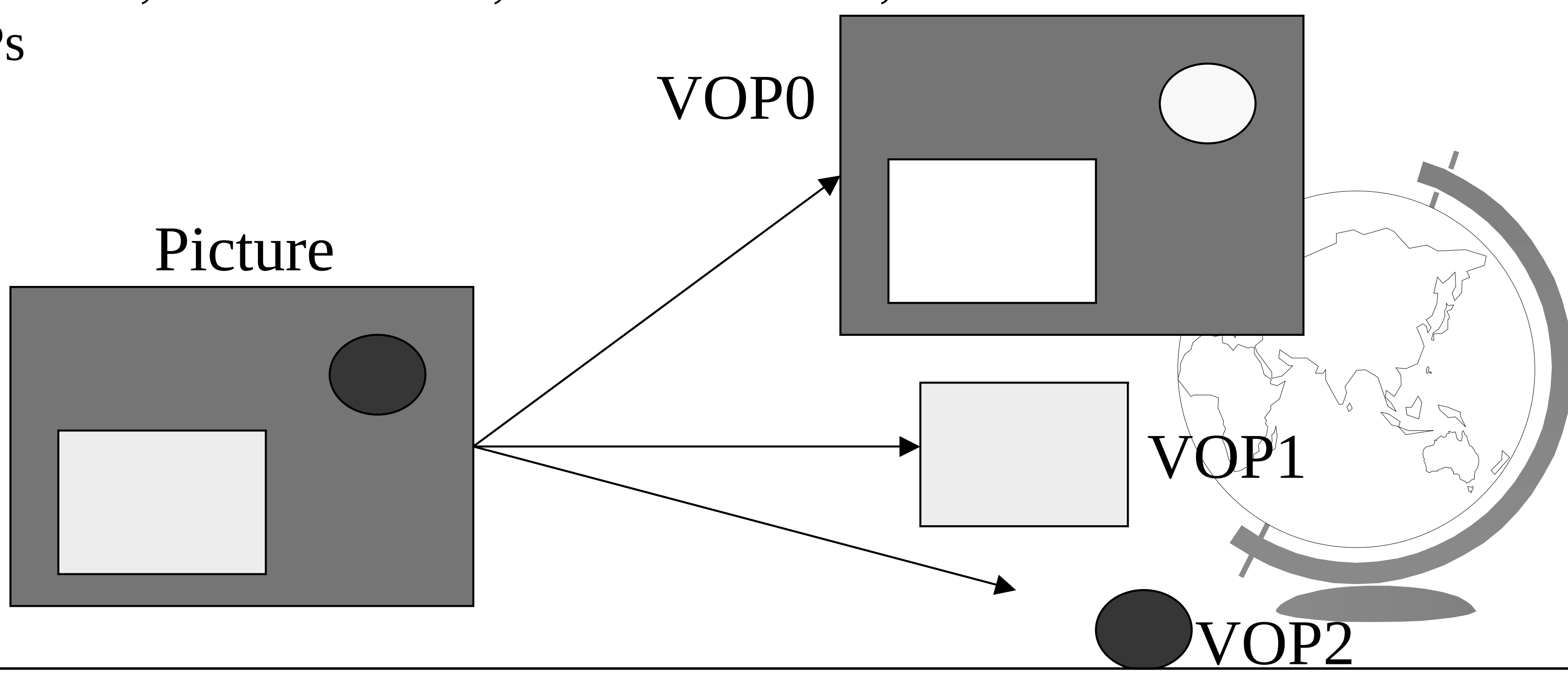
- MPEG-4:
  - In addition to previous audio, video encoding and multiplexing, also has
    - coding of text/graphics and synthetic images
    - Representation of audio-visual scene and composition
  - Has some wireless features
  - New features considered important included robustness to errors and coding efficiency
  - Example applications:
    - Internet and Intranet video
    - Wireless video
    - Video databases
    - Interactive home shopping
    - Video e-mail, home movies
    - Virtual reality games, simulation and training

MPEG-4

- MPEG-4 specific wireless-friendly standards requirements:
  - **Universal access:** *"Robustness in error prone environments: The capability to allow robust access to applications over a variety of wireless and wired networks and storage media. Sufficient robustness is required, especially for low bit-rate applications under severe error conditions"*
  - **Compression:** *"Improved coding efficiency: The ability to provide subjectively better audio-visual quality at bit-rates compared to existing or emerging coding standards"*
- Formal tests to verify these requirements with:
  - high random Bit Error Rate (BER) of 10^{-3}
  - Multiple burst errors

MPEG-4 Video Basics

- Input video sequence = series of related snapshots/pictures
- Elements of a picture = Video Object (VO)
- Video Objects are changed by translations, rotations, scaling, brightness, color, etc
- Several MPEG-4 functions access these VOs not pictures
- Video Object Planes (VOPs) described by texture variations
- Similar to I, B and P frames, there are I-VOPs, B-VOPs and P-VOPs



MPEG-4

- Other features such as:
  - sprite coding for games,
  - scalable video coding for variable video quality
  - robust video coding
- Robust video coding including:
  - Object priorities: lost low priority objects have little effect
  - Resynchronization: errors don’t accumulate
  - Data partitioning:
  - Reversible VLCs
  - Intra update and scalable coding
  - Correction and concealment strategies (not specified due to channel-specific nature). E.g. addition of FEC bits