8 Streaming Architectures

Literature:

David Austerberry: The Technology of Video & Audio Streaming, Focal Press 2002

Gregory C. Demetriades: Streaming Media, Wiley 2003

Tobias Künkel: Streaming Media – Technologien, Standards, Anwendungen, Addison-Wesley 2001

Outline

1.	Introduction and Motivation	
2.	Media on the Web	Part I: Web Technologies for Interactive MM
3.	Interactive Web Applications	
4.	Communities, the Web, and Multimedia	
5.	Digital Rights Management	Part II: Content-Oriented Base Technologies
6.	Cryptographic Techniques	
7.	Multimedia Content Description	
8.	Streaming Architectures	Part III: Multimedia Distribution Services
9.	Web Radio, Web TV and IPTV	
10.	Electronic Books and Magazines	
11.	Multimedia Content Production and Management	
12.	Multimedia Conferencing	Part IV: Conversational Multimedia Services
13.	Signaling Protocols for Multimedia Communication	
14.	Visions and Outlook	

Streaming, Streaming Media

- *Streaming media* means real-time delivery of moving images, moving text and sound, over the internet.
- Delivery types for audio and video content:
  - *Download and Play:* Content must be downloaded completely to the client before it can be played
  - *Progressive Download:* Playback is started while download is still in progress. Download rate independent of program bit rate.
  - *True Streaming:* Delivered media is viewed/listened in “real-time”.
    - » Playback takes place with roughly the same rate as delivery of data
    - » Delay between send and receive event of data packet kept small
- Subtypes of True Streaming:
  - *Static File Streaming:* Delivery of pre-recorded media files. Often also called *on-demand* delivery (e.g. *Video on Demand)*
  - *Live Streaming:* Source of delivered media is not pre-recorded but fed into the server in real-time. Examples: *Webcast* of live events, Web radio
  - Static file streaming enables a higher degree of interactivity

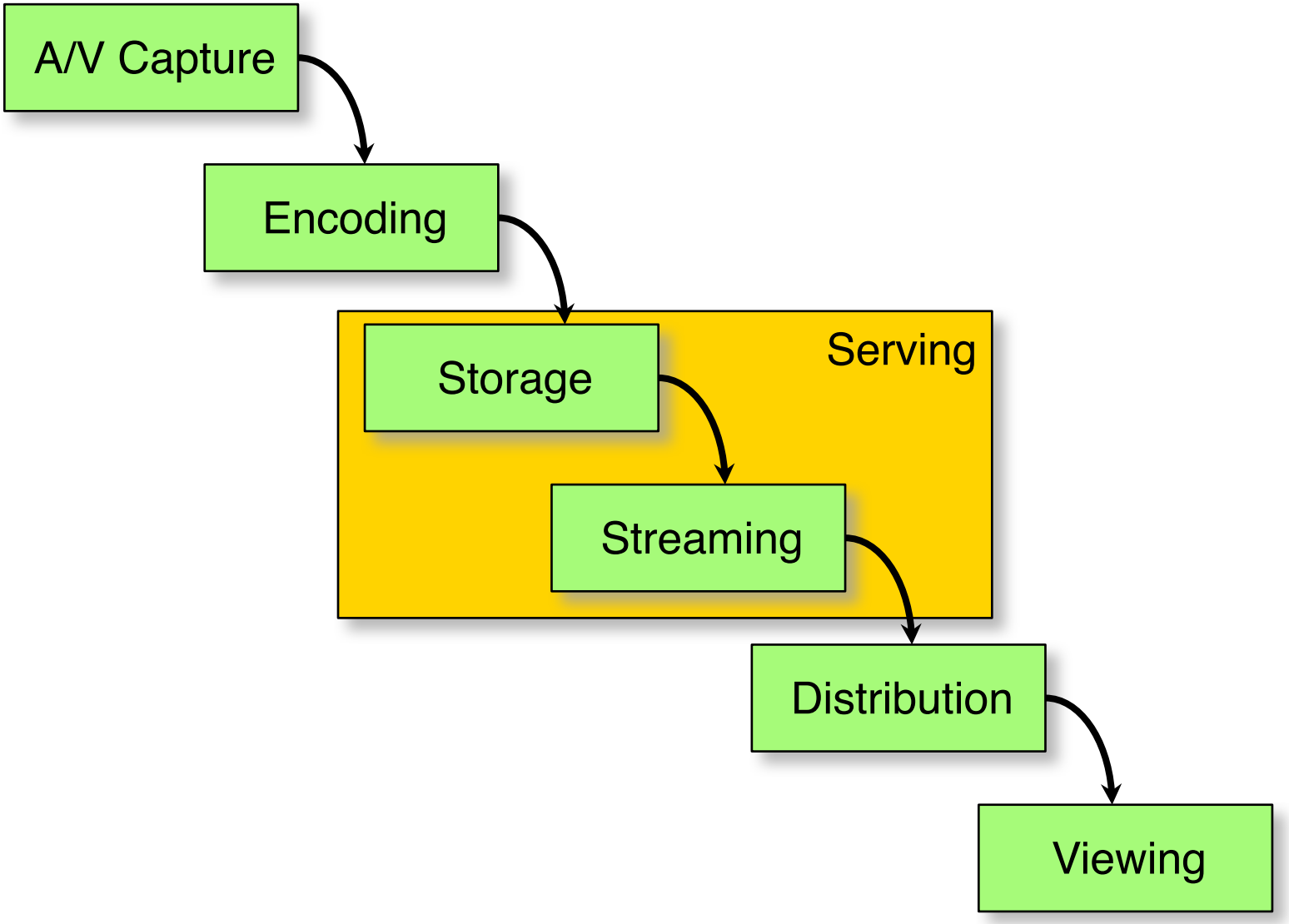
Based on material from www.streamingalliance.org

Session

- A *session* is an association between communicating parties, which
  - Persists over a limited time span
  - Incorporates at least two parties
  - May comprise a large number of communication connections of different characteristics
- Examples of sessions:
  - Movie streamed to consumer, consisting of audio and video parts
  - Multimedia conference among five participants, consisting of audio and video source from each of of the participants (plus possibly some global information)

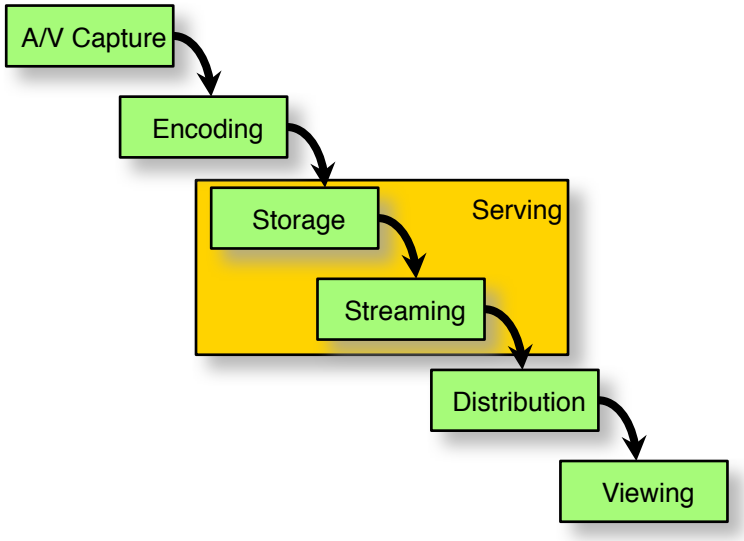
- Session awareness at which levels?
  - At application level: unavoidable
  - At network level: possible
    - » Requires specific protocols

Streaming Delivery Chain

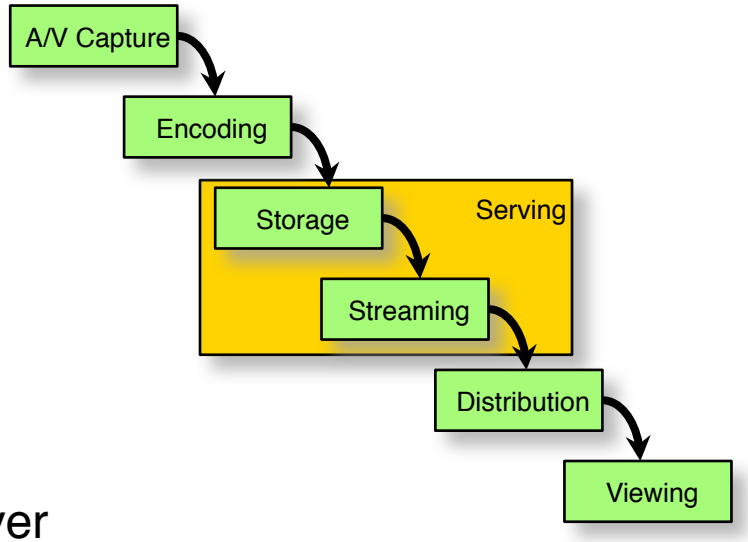


Encoding

- Format conversions
  - E.g. analog/digital conversion
  - E.g. downscaling of picture size
- Compression
  - Adequate for player capabilities and typical transmission bandwidth
- Indexing
  - Analyzing internal structure
- Metadata creation
  - Possibly including digital rights specification

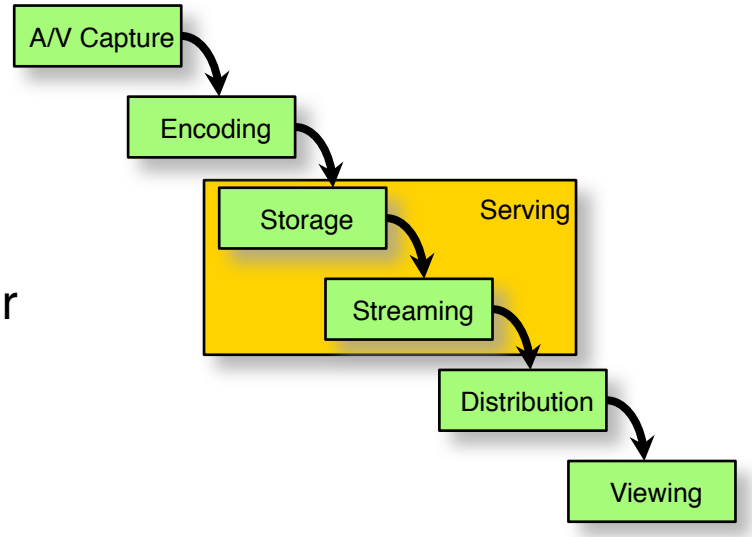


Serving



- Storage
  - Live: only buffering and archiving
  - Static files: archive management, retrieval
- Streaming
  - Request-response driven similar to Web server
- Interactivity
  - In static files:
    - » VCR-like control (PLAY, STOP, PAUSE, FFWD, REW)
    - » Random access based on various criteria
  - In structured (mostly static) material archives:
    - » Hyperlinks in A/V material (“hypervideo”)
    - » Web-like technology: Video-Web, links to arbitrary streaming sources

Distribution



- Key topic: Quality of Service (QoS)
  - Determining realizable bandwidth, delay, jitter

- Key concepts:
  - Overprovisioning
  - Detailed reservations (“Integrated Services”, reservation protocol RSVP)
    - » Difficult to scale to large numbers of users
  - Traffic classes (“Differentiated Services”)
    - » Difficult to control access to privileges
  - Resource management layer
  - Technology-specific solutions
    - » E.g. ATM (Asynchronous Transfer Mode)

8 Streaming Architectures

Literature:

David Austerberry: The Technology of Video & Audio Streaming, Focal Press 2002

Stephan Rupp, Gerd Siegmund, Wolfgang Lautenschlager: SIP – Multimediale Dienste im Internet. dpunkt 2002

IP and TCP

- Internet Protocol
  - Network communication protocol (ISO layer 3)
  - Packets transferred from address to address (through routers)
  - Main problems:
    - » Variable network latency
    - » Packet order on arrival may be different than on sending
    - » Packets may be lost
- Transport Control Protocol (TCP)
  - Connection establishment (by “three-way handshake”)
    - » Connections are sequences of associated IP packets
  - Sequencing of bytes with forwarding acknowledgement number
  - Non-acknowledged bytes are re-transmitted after a defined time period
  - Flow control
- For audio/video streaming:
  - Retransmissions (and associated delays) are harmful
  - Lost packets can be tolerated to some extent
  - TCP not adequate for data transport

UDP

- User Datagram Protocol (UDP)
- Extremely simple transport protocol over IP
  - Connectionless (TCP: connection-oriented)
  - Unreliable (TCP: reliable)
  - No flow control (TCP: has flow control)
- Contents of a UDP datagram:
  - Ports used by application program
  - Checksum
- Basically adequate for media data transport
  - Very efficient, protocol overhead of TCP avoided
  - Flow control and handling of packet loss have to be handled by higher protocol layer

Streaming with HTTP?

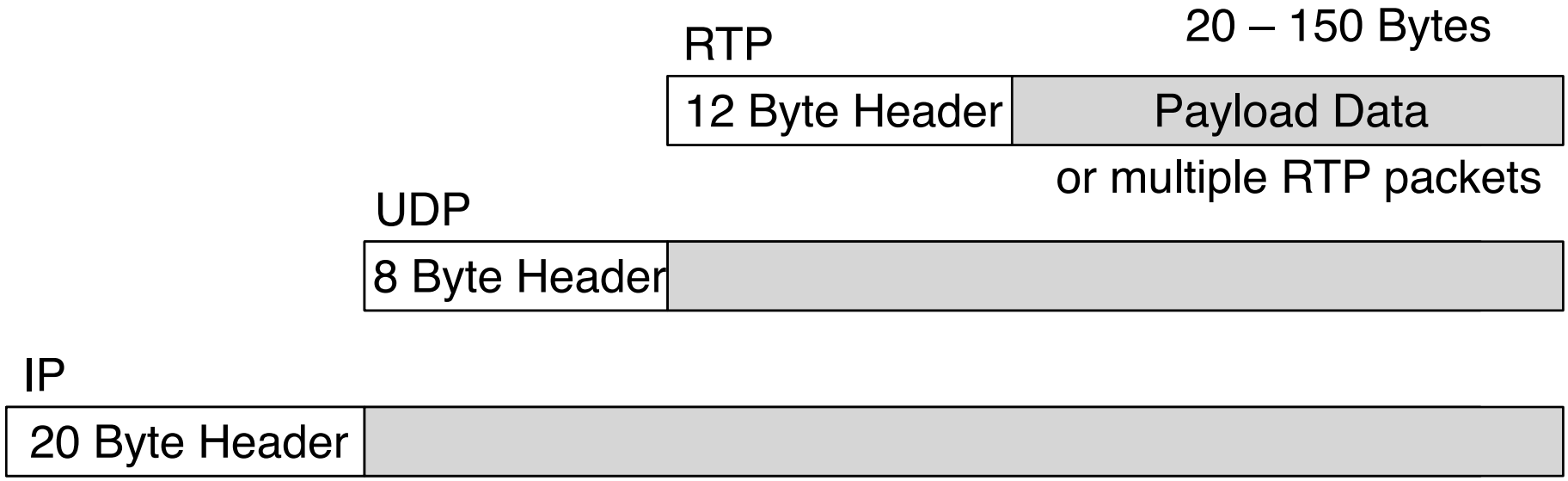
- Theoretically possible
  - Client requests next slice of audio/video information from server
  - Server responds with audio/video data
  - Popular in simple streaming applications
    - » e.g. MP3-Streaming with ShoutCast, Live365
- Main problem
  - HTTP usually run over TCP
  - Large overhead for ensuring correct transmission

- Consequence:
  - Specialized streaming protocols

Real-Time Transport Protocol RTP

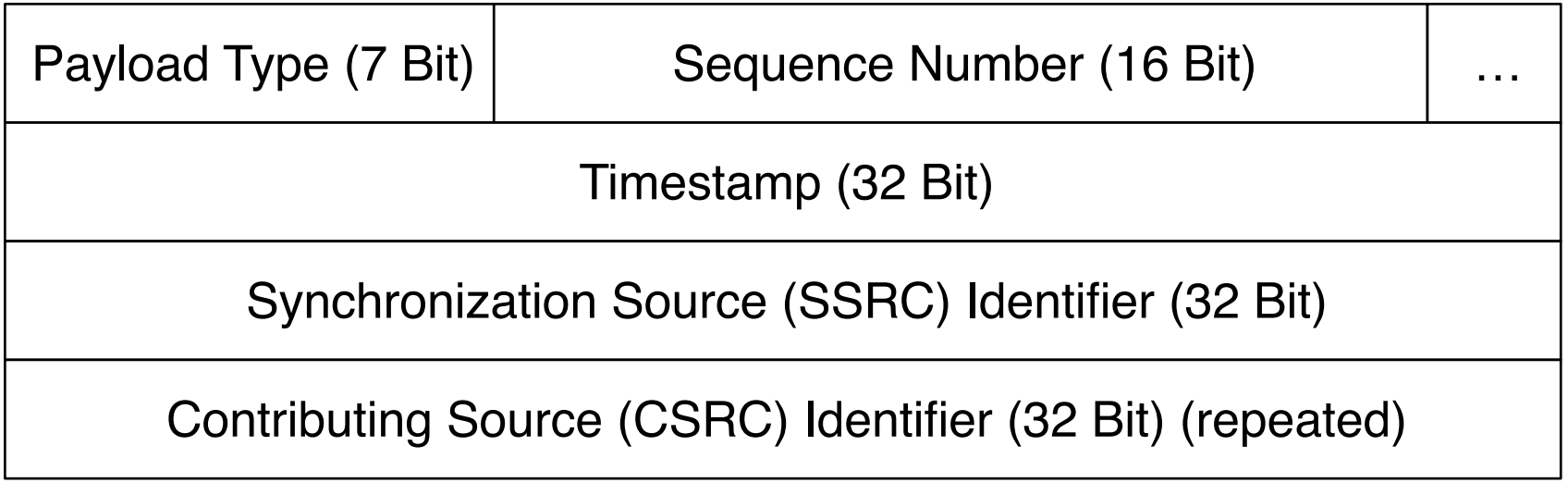
- Transport protocol specifically developed for streaming data
  - IETF (Internet Engineering Task Force) RFC (Request for Comments) 1889
- RTP packets contain
  - Sequence number
  - Time stamp
  - Identification of sender and destination
- RTP usually carried over UDP
- Very important:
  - **RTP does not at all change the way how IP packets are transferred in the network!**
  - To achieve “Quality of Service”, additional network technologies are required (see above)
- RTP used (for instance) by:
  - Apple QuickTime architecture
  - RealSystems streaming architecture

RTP Packets and Other Protocols



- IP Header:
  - Source address, destination address, length, time to live, ...
- UDP Header:
  - Port numbers (source and target processes), length, checksum

RTP Header Format

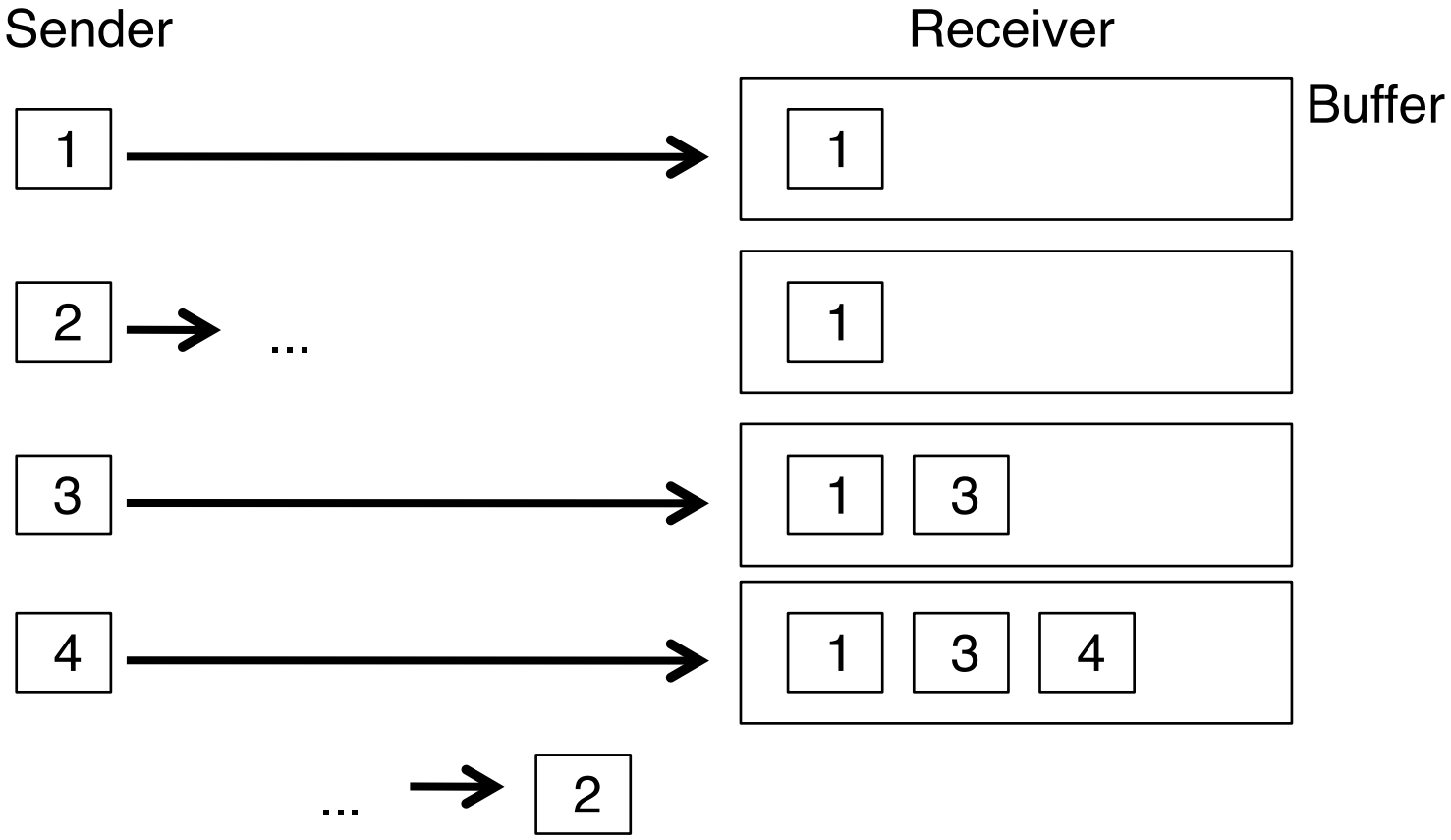


- Payload Type: Source coding (codec used)
- SSRC: Identification of sending data source, defined basis for sequence numbers and time stamps
- CSRC (optional): Identifications for original data sources which have been *mixed* together to form the synchronization source

Buffer Size Allocation

- Streaming quality relies on
  - Features of the network, and
  - *Adaptive* codecs on client/server side
  - E.g. intelligent use of buffering
- *Constant* delay does not require buffering
- Buffers are necessary to deal with *jitter (packet delay variation)*
  - Varying network delays
- Estimation of buffer sizes
  - Based on preliminary test transmissions
  - Adaptively during content transmission
- Problem:
  - Keep buffers small to achieve proper “real-time”
  - Avoid “buffer underrun” in case of exceptionally long delays

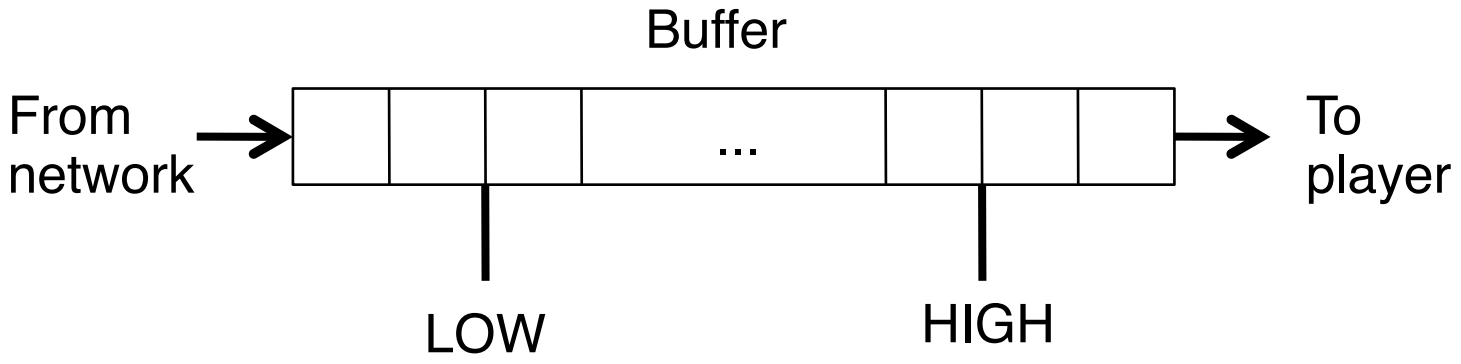
Jitter and Loss Compensation



- Options for application on receiver side:
  - Wait *(not* adequate), repeat last packet (1), interpolate (between 1 and 3)
  - Missing audio information is difficult, missing video can be compensated

Adaptive Transmission Rate Control

- Application-level mechanism
- Define “low” and “high” thresholds on buffer
- Communication between client and server
  - Decrease effective transmission rate when high threshold is reached
  - Increase effective transmission rate when low threshold is reached
- Changing the effective transmission rate
  - Change sending rate
  - Change content quality (frame rate, resolution etc.)



Real-Time Control Protocol RTCP

- RTCP controls the transmission (not the setup of connection)
- RTCP periodically sends monitoring information to all participants in a streaming session
- Main functions of RTCP:
  - Feedback on QoS of transmission
    - » Information for adaptive codecs, e.g. whether problem is local or global
  - Identification of sender by “canonical name”
    - » Helpful when synchronization source changes
    - » Supports lip synchronization between audio and video
  - Number of participants in a session
    - » Adaptation of sending rate of RTCP control information to number of participants, to avoid network overload
  - Transmission of additional information, e.g. names of session participants

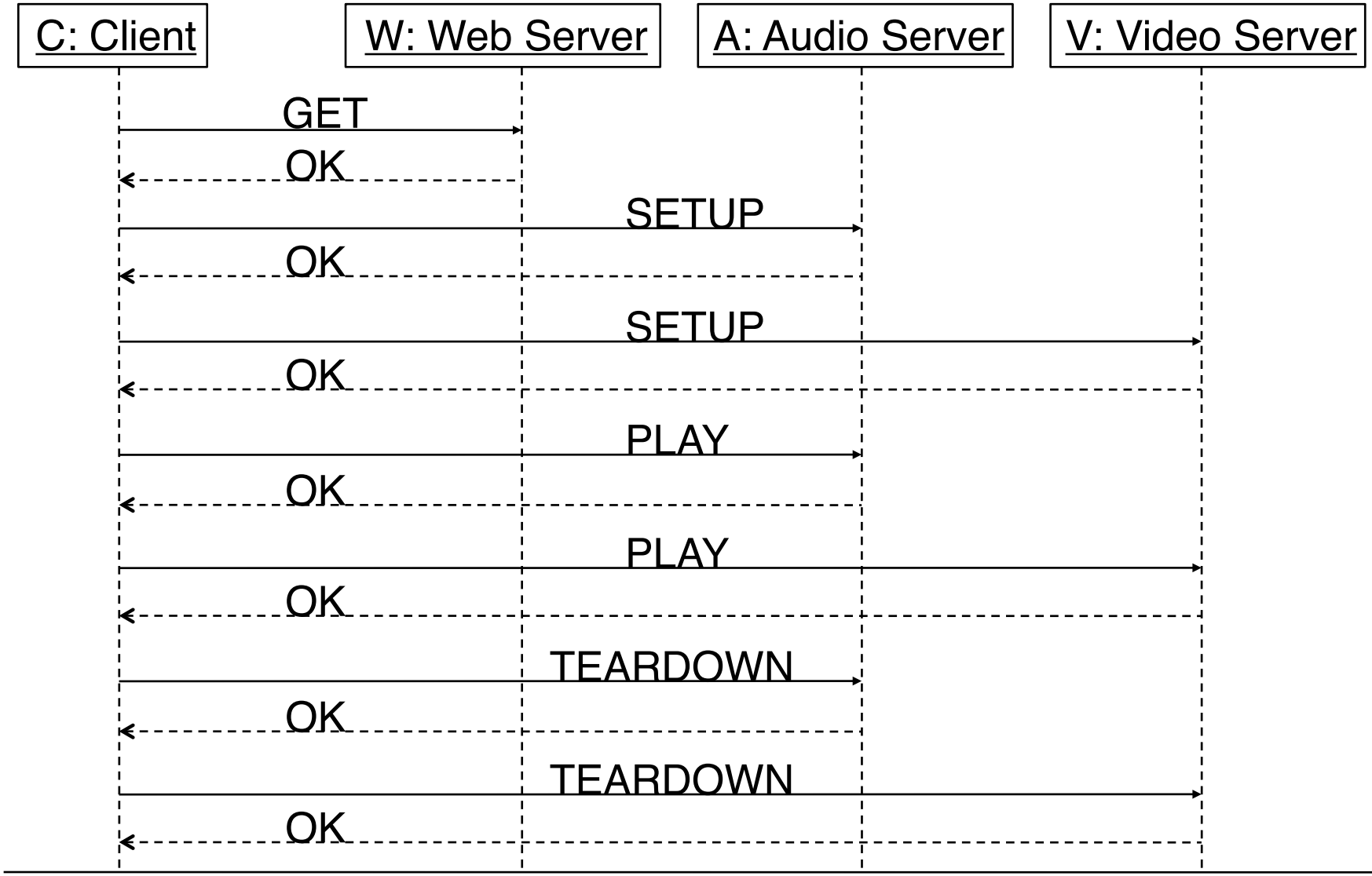
Real Time Streaming Protocol RTSP

- Client-server multimedia presentation protocol, designed specifically for streamed media
  - IETF (Internet Engineering Task Force) RFC (Request for Comments) 2326 (“MMUSIC” work group)
    - » February 1998, draft revision February 2004
  - “The Internet VCR remote control protocol” (www.rtsp.org)
  - Independent of the use of RTP for transport
  - Syntactically similar to HTTP 1.1 (carried over TCP or UDP)
- Main operations supported by RTSP:
  - Retrieval of media from media server
  - Invitation of a media server to a conference
- Key terminology
  - Aggregate control (e.g. for audio & video)
  - Server control (clients should be able to stop streaming from a server)
  - Transport & capability negotiation (e.g. disallowing a “seek” function)

Main Methods of RTSP

- SETUP:
  - Causes the server to allocate resources for a stream and create a RTSP session.
- PLAY:
  - Starts data transmission on a stream allocated via SETUP
  - Fast forward (scale ratio parameter)
- PAUSE:
  - Temporarily halts a stream without freeing server resources.
- REDIRECT:
  - Indicates that the session should be moved to new server / location
- PING:
  - Prevents the identified session from being timed out.
- TEARDOWN:
  - Frees resources associated with the stream. The RTSP session ceases to exist on the server.

Example Session with RTSP



Microsoft Media Server MMS

- Microsoft Technology for transmission of data packets and control messages from server to client
  - Based on TCP
- Protocol Rollover: Server tries the following protocols in sequence
  - MMSU: Microsoft Media Server Protocol/UDP
  - MMST: Microsoft Media Server Protocol/TCP
  - HTTP (may be successful in Firewall configurations)

- MSBD (Media Stream Broadcast Distribution Protocol):
  - Earlier solution for client-server connections
  - Currently limited in client numbers

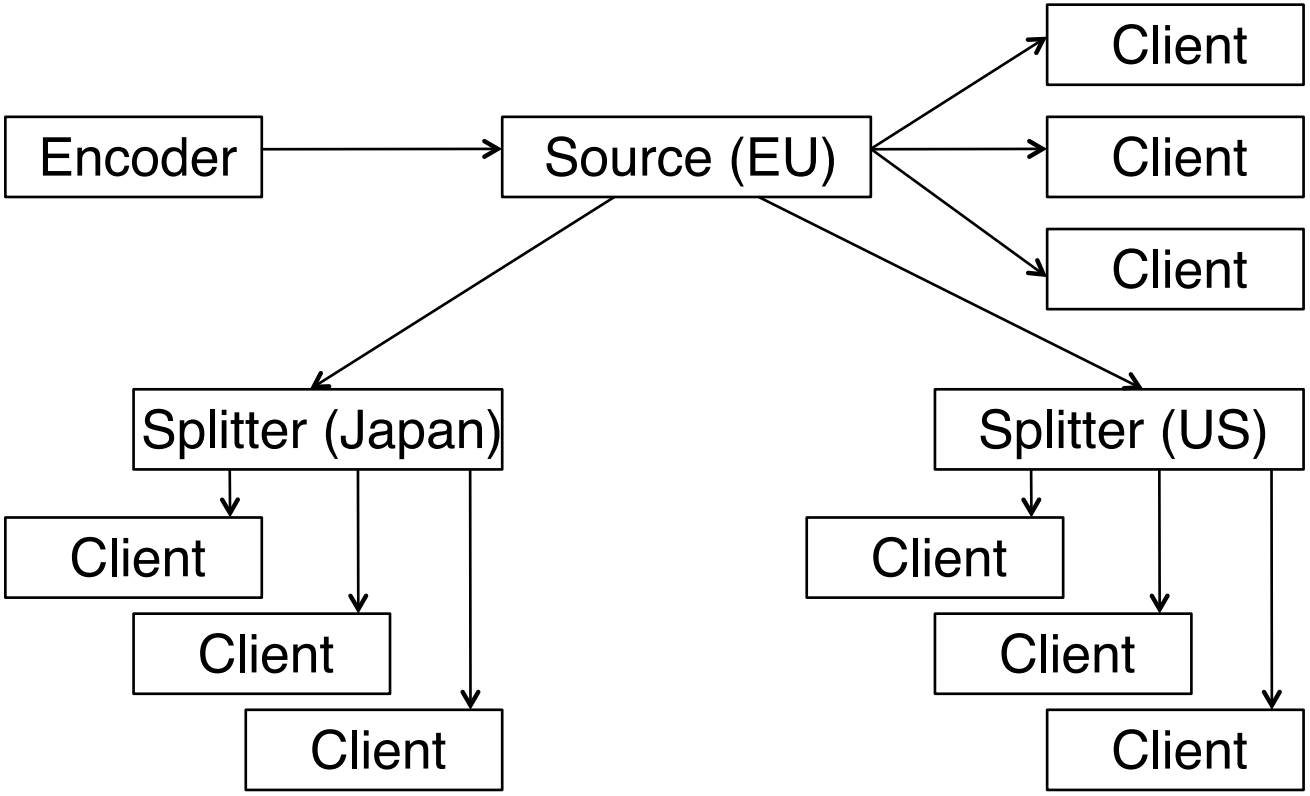
8 Streaming Architectures

Literature:

David Austerberry: The Technology of Video & Audio Streaming, Focal Press 2002
Gregory C. Demetriades: Streaming Media, Wiley 2003
Xueyan Tang et al.: Web Content Delivery, Springer 2005

Splitting

- Video servers are limited in capacity
- Assuming clients at spatially distant locations
  - Intermediate, forwarding server is useful: “splitter”



Content Delivery Networks (CDN)

- Serve content closer to to the user
  - “edge serving”
- Main components of CDN:
  - Smart routing
  - Edge caching of content
  - Proxy serving
  - Splitting of live webcasts

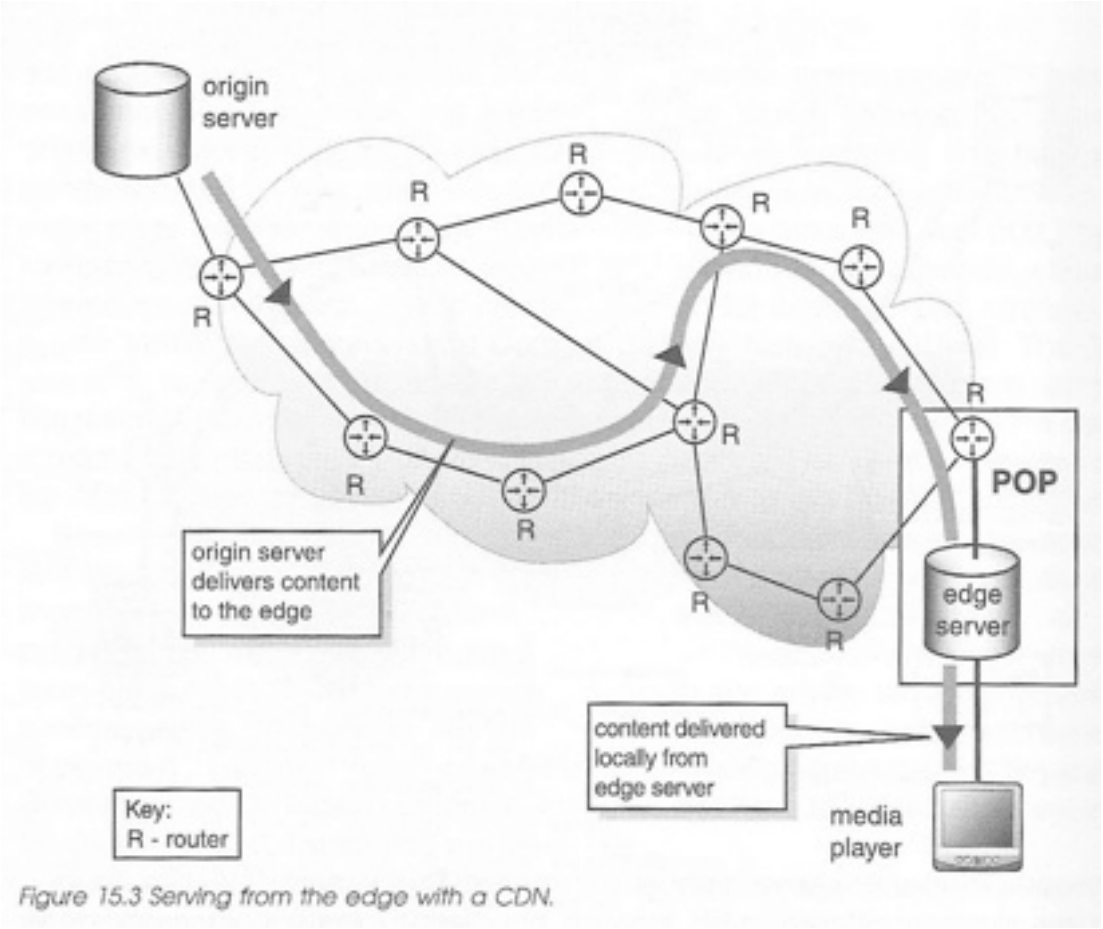


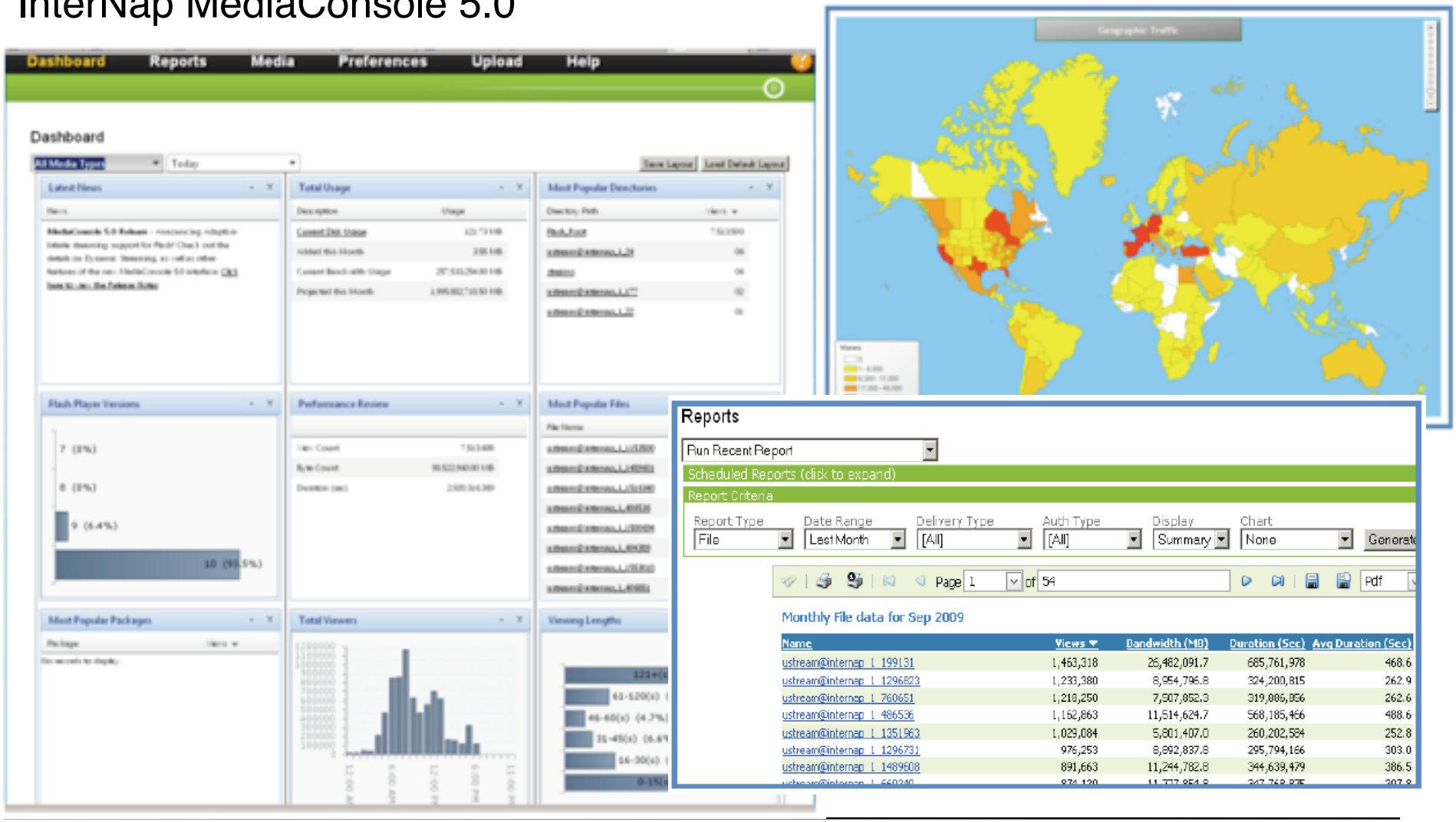
Figure from Austerberry 2002

Content Delivery Networks

- “Overlay server infrastructure”
  - Network of centrally operated Web and streaming servers
  - Geographically distributed, present in main ISP networks
  - Flexibly used for content from various sources
- Content delivery as a service
  - Content delivery service provider owns server infrastructure
  - Content owner pays for having content delivered to customers
- Examples:
  - Akamai (delivers 20% of all Web traffic according to akamai.com, 2009)
    - » Runs 50,000+ servers in 71 countries
  - InterNap CDN Services
    - » Traditionally specialized in media streaming (acquisition of VitalStream)
  - Amazon CloudFront
- Streaming of a/v media (e.g. movies) gains strategic importance

Example of Monitoring/Administration Interface

InterNap MediaConsole 5.0



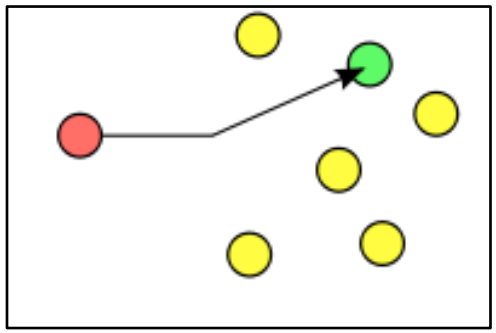
Key Problems in CDNs

- Replica placement:
  - Where to place copies of web sites or other content
  - Problem is in general NP-hard (Karlsson, Karamolis, 2004)
  - Replica placement algorithms (RPA) achieve a suboptimal solution within reasonable time frame
  - Global information is difficult or costly to get - RPA uses local information mostly
  - CDN providers typically try to observe global network performance to some extent
- Request routing:
  - Mechanism and policy of redirecting client requests to a suitable server containing the requested content
  - Redirection algorithm: Decides what node to direct a client request to
  - Redirection mechanism: Way of redirecting the request (client, network)

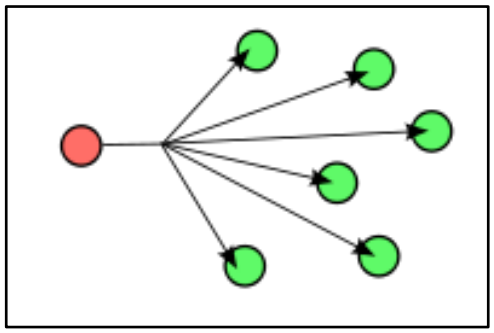
Streaming Media in CDN

- General idea: Local proxy caching. But: ...
  - Huge size (1 KB vs. 100 MB)
    - » To cache only portions of the original?
  - Intensive bandwidth usage
    - » Minimizing bandwidth consumption as primary consideration
  - High interactivity
    - » E.g. premature termination is frequent (Chen et al. 04: approx. 90 %)
  - However: Media content is rather static (compared to Web pages)
- Caching algorithms
  - Different for homogeneous and heterogeneous clients (in bandwidth/format)
  - Sliding interval caching: sequential access, mainly effective for similar requests in short time period
  - Prefix caching: Saves time to load remaining parts
  - Segment caching: Generalization of prefix caching to support fast forward
  - Rate-split caching: Lower layer from original server, higher layer from proxy
  - Co-operative proxy caching (e.g. Acharya/Smith: MiddleMan)

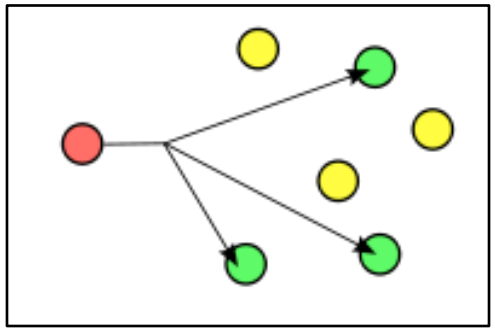
Unicast, Broadcast, Multicast, Anycast



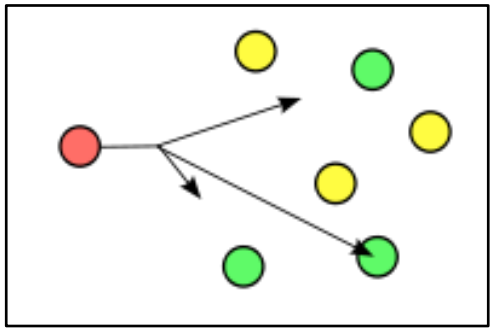
Unicast:
One specific receiver



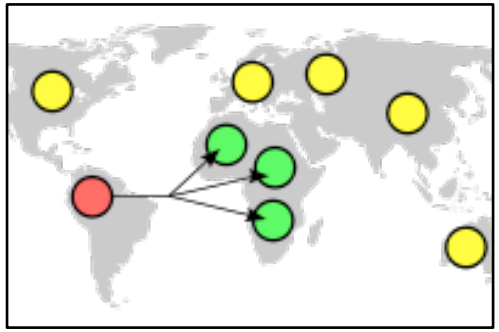
Broadcast:
Many receivers, all on the network



Multicast:
Many receivers, all of a specific group



Anycast:
One receiver, "nearest" of a specific group

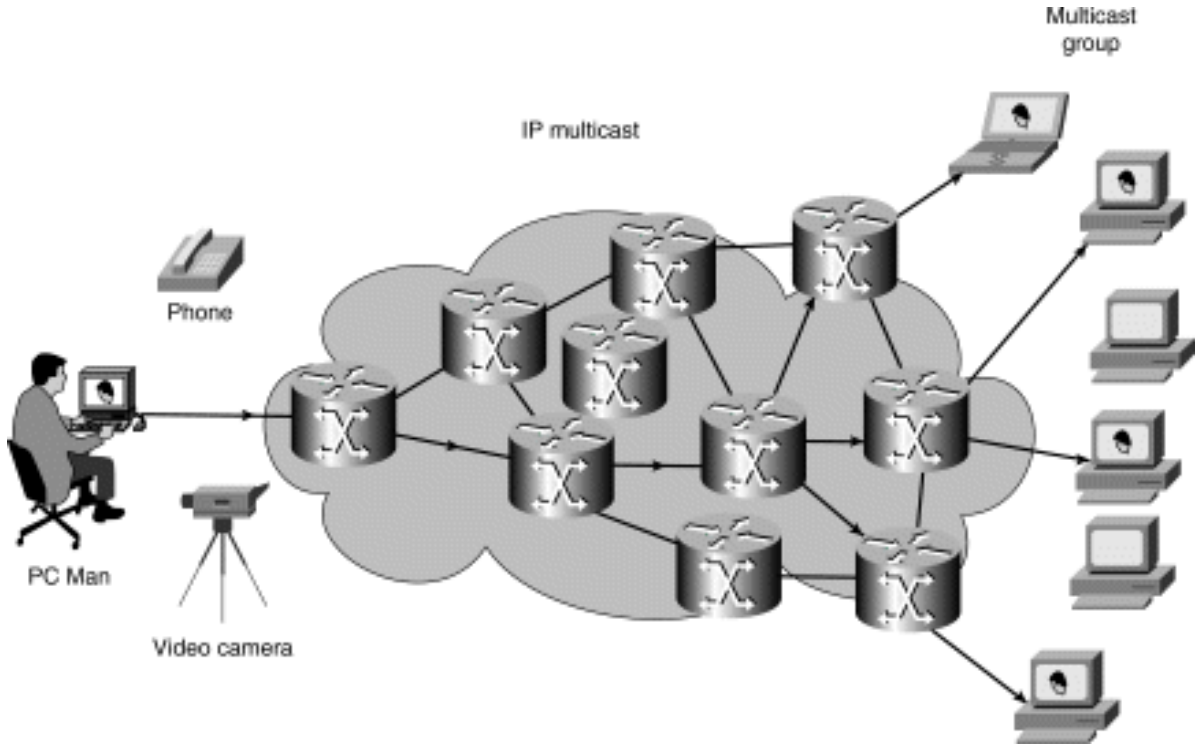


Geocast:
Many receivers, all of a geographic region

Pictures: Wikipedia

IP Multicast

- Multicast relatively easy to integrate in routers
- IP address class D (224.0.0.0 through to 239.255.255.255) reserved for multicast (multicast groups)
- Registration/deregistration with IGMP (Internet Group Management Protocol)



- Reliable multicast: e.g. “Mbone” overlay network
- Multicast still rarely used in today’s Internet

IP Version 6

- Next generation of the IP protocol
- 128 Bit address space
  - Intended to relieve shortage of IP v4 addresses
- Built-in support for multicast
  - Specific multicast addresses
- Uptake of IP version 6 is (strangely) slow

8 Streaming Architectures

8.1 High-Level Streaming Architecture

8.2 Real-Time Data Transport

8.3 Scalability and Multicast

8.4 Selected Commercial Streaming Architectures

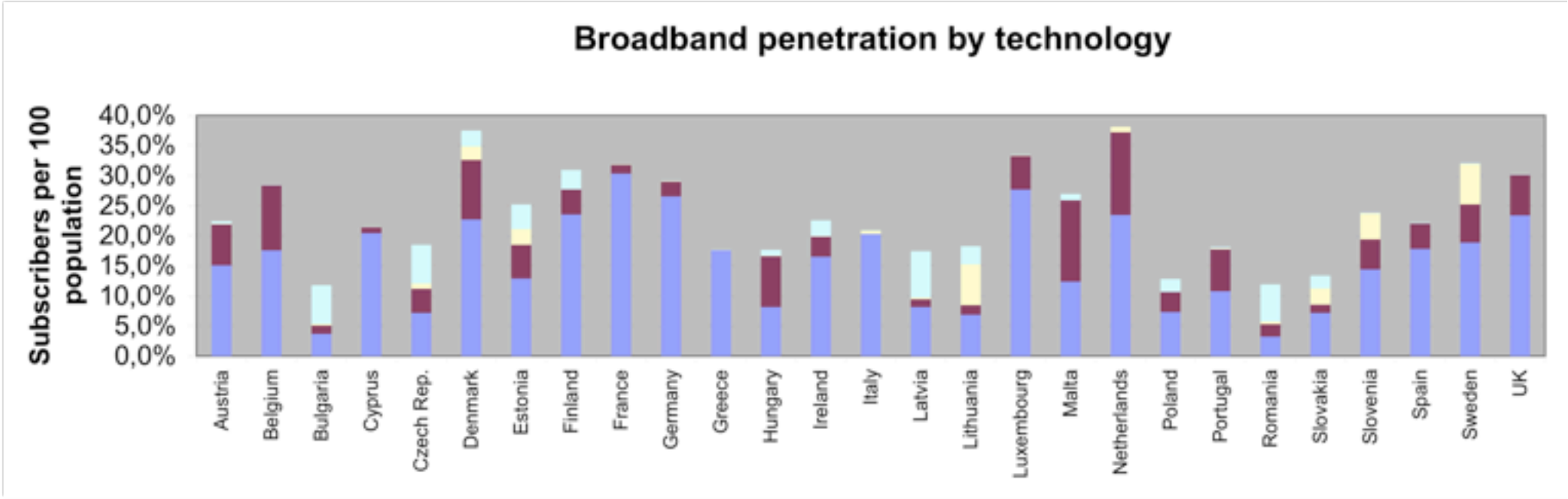
RealNetworks, QuickTime, Windows Media

Literature:

Tobias Künkel: Streaming Media – Technologien, Standards, Anwendungen, Addison-Wesley 2001

David Austerberry: The Technology of Video & Audio Streaming, Focal Press 2002

Broadband Internet in the EU



Source:
European Competitive Telecommunications Association (ECTA),
"Broadband Scorecard Q3 2009"

Actual Usage of Streaming

- 2004 (AccuStream iMedia research report):
  - Total number of video streams served/viewed rose 104% in 2003, following two years of consecutive growth
- June 2005 (comScore Media Metrix), USA:
  - 94 million people (56 percent of domestic Internet population), viewed a streaming video online
  - April/May/June 2005: the average consumer viewed 73 minutes of streaming video content per month.
- May 2010 (comScore Video Metrix), USA:
  - 183 million people view video online
  - YouTube videos: 14.6 billions watched (in May 2010)
- October 2009 (comScore Video Metrix), Germany:
  - 40.2 million people view video online (2009: 28.5 million)
  - 7.8 billion videos watched
- Clear leader in video site popularity: Google sites (YouTube)

comscore.com

Streaming Media Market 2009 – 2014

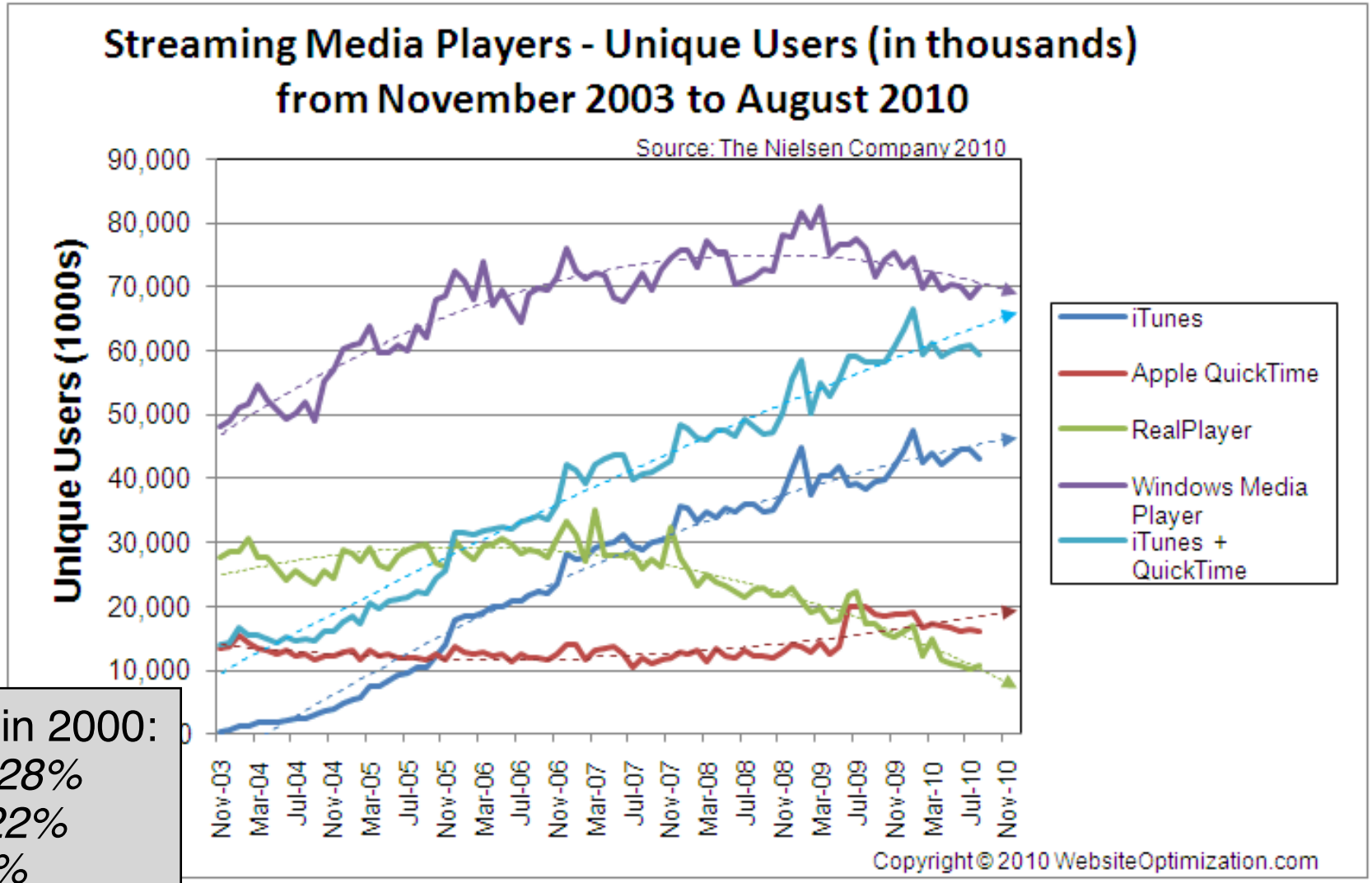
Insight Research, June 2009:

- Streaming media market will grow at a compound annual rate of 27% for the next six years
- Media over Internet market will generate a total revenue of $78 billion US$

Robert Rosenberg, Insight Research president:

"Over the past seven years as we've tracked the developments in streaming it has evolved from an esoteric niche to a mainstream market. What we predicted way back then is coming to fruition. The advertising revenue that long supported traditional TV is gravitating to this new medium, putting downward pressure on traditional TV distribution schemes,"

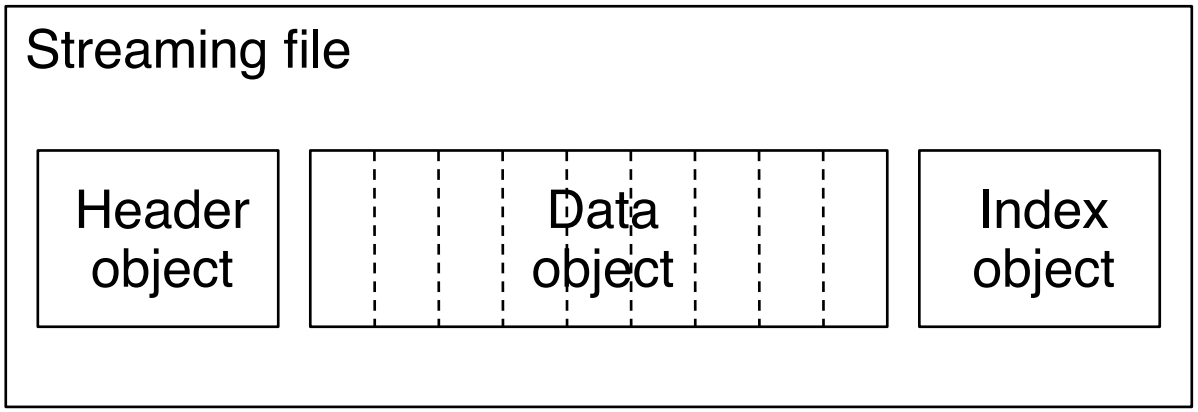
Market Shares of Streaming Players



Back in 2000:
Real 28%
Win 22%
QT 4%
(streaming media.com)

Prediction: iTunes+QT will pass Windows Media by 2011

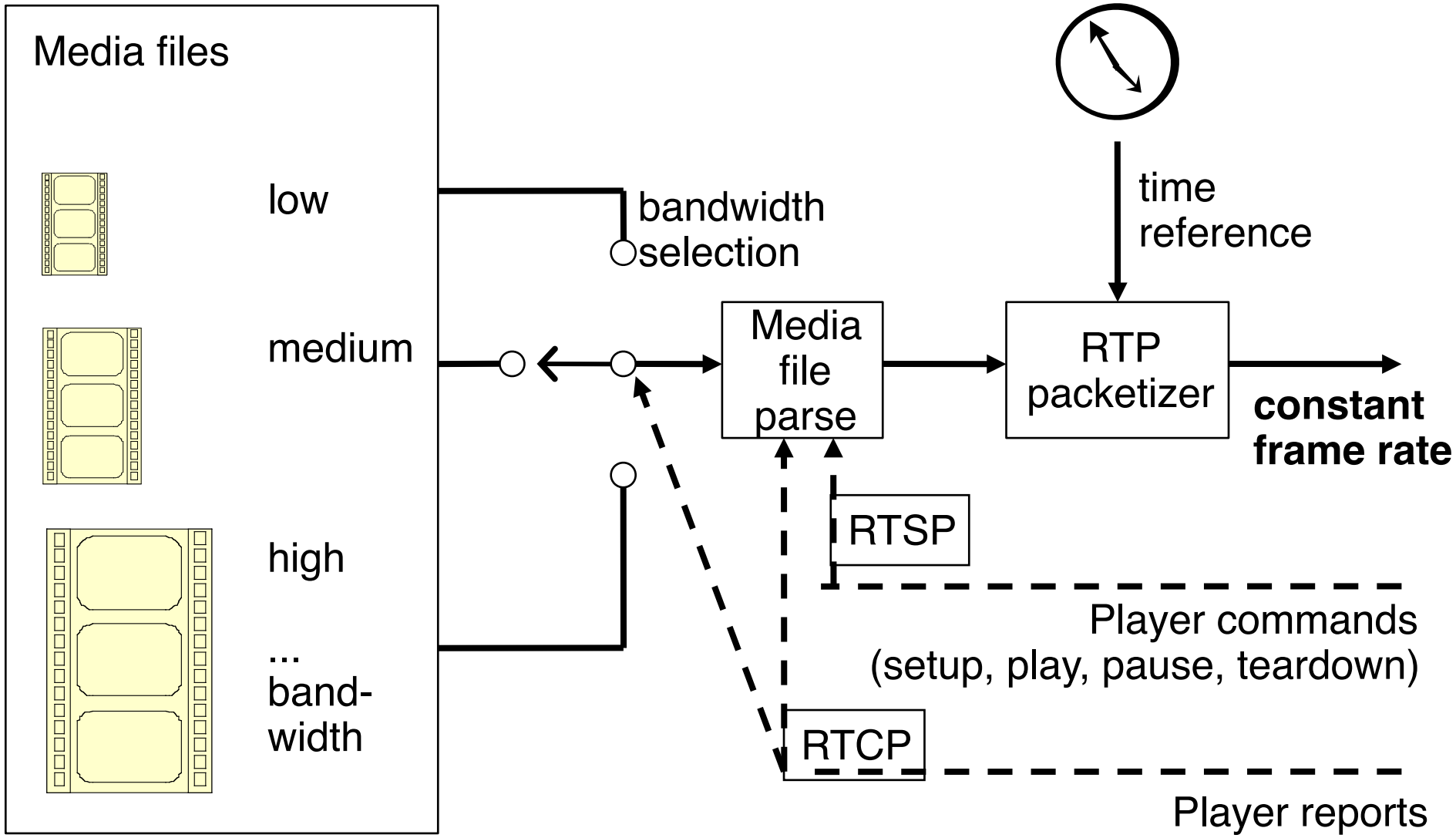
Streaming File Formats



(from ASF)

- Header, Data: As in other audio/video formats
  - Additional timing control information used to manage flow rate
- Index Object: Aid for client navigation
- Main streaming file formats:
  - Microsoft: Advanced Streaming Format (ASF), Windows Media Video (WMV), Windows Media Audio (WMA)
  - RealNetworks: RealMedia (RM), RealAudio (RA)
  - Apple: QuickTime Hinted Movie (MOV)

Adapting to Network Congestion



Realisations for Rate Adaptation

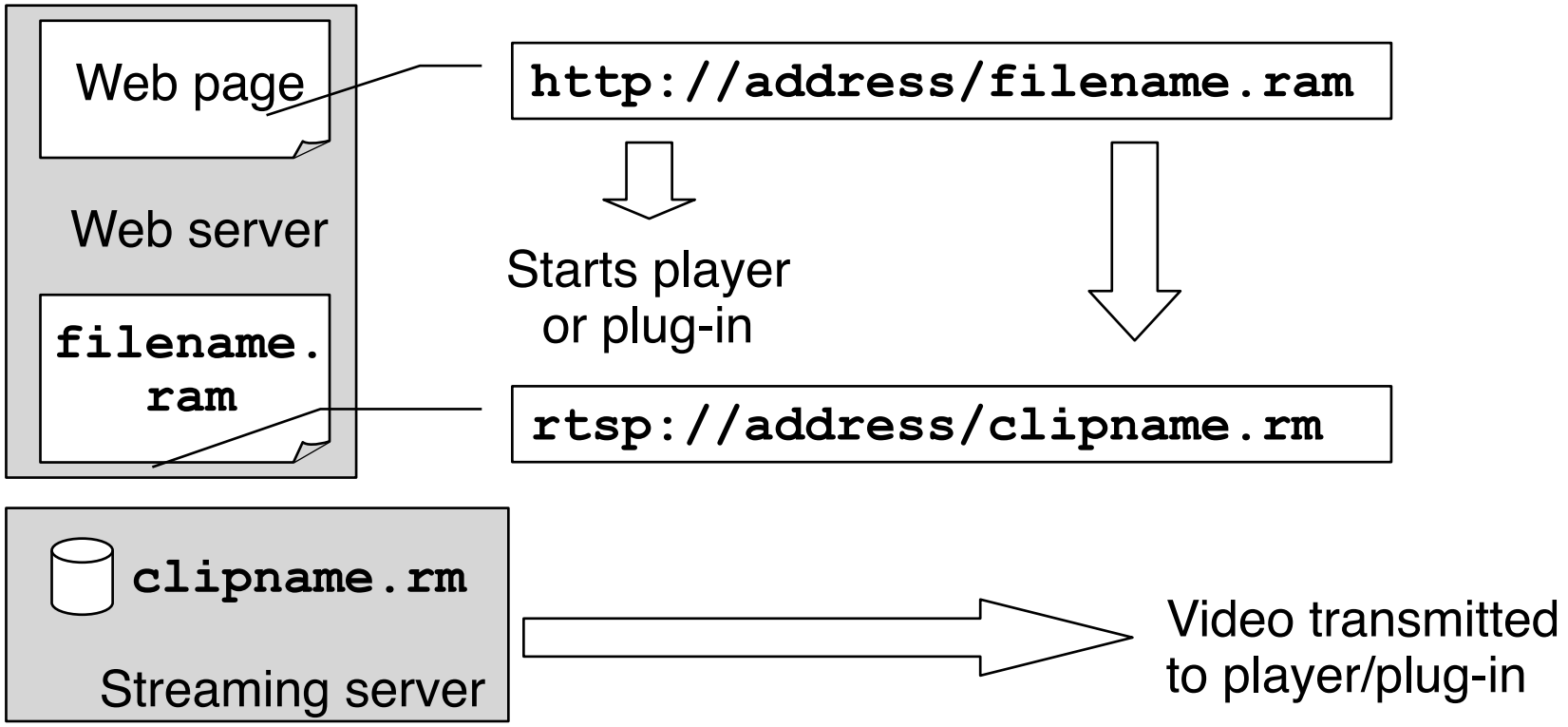
- Multiple bit rate files
  - RealNetworks “SureStream”, Windows Media “multiple bit rate”
  - Several bit rates in one file
  - Compatible only with streaming servers, not with Web servers
  - Adaptation by change of picture size not supported
- Alternate movies (QuickTime)
  - Player receives pointers to assemble the actual program
  - Usable for adapting bit rate and other parameters
  - Usable also for different language versions and other applications
- MPEG-4 Scalable Streams
  - Similar to “progressive” technique in picture compression
  - Basic low-resolution stream transmitted
  - Additional “helper” streams can add more detail and improve quality

Constant and Variable Bit Rate

- Constant Bit Rate (CBR)
  - Fixed bit rate (of decoded data stream) independent of content
- Variable Bit Rate (VBR)
  - Low bit rate for simple scenes (e.g. no movements)
  - High bit rate for complex scenes (e.g. quick changes)
  - Average bit rate can be optimized: Scenes “borrow” bandwidth from others
    - » Still an average bit rate limit is obeyed
    - » “Instantaneous bit rate” can be limited
  - Leads to increase of buffer lengths and buffering times
  - Optimal effect in combination with two-pass encoding
- Quality-Based Encoding
  - Maintains steady visual quality of output
  - No limit to average bit rate
    - » “Instantaneous bit rate” can be limited

Indirect Web Links

- Web link points to small file containing the actual media file link
  - Real: “Real Audio Metafile” (RAM)
  - Microsoft: “Stream redirector” (ASX)
- Metafile may contain list of files (playlist)

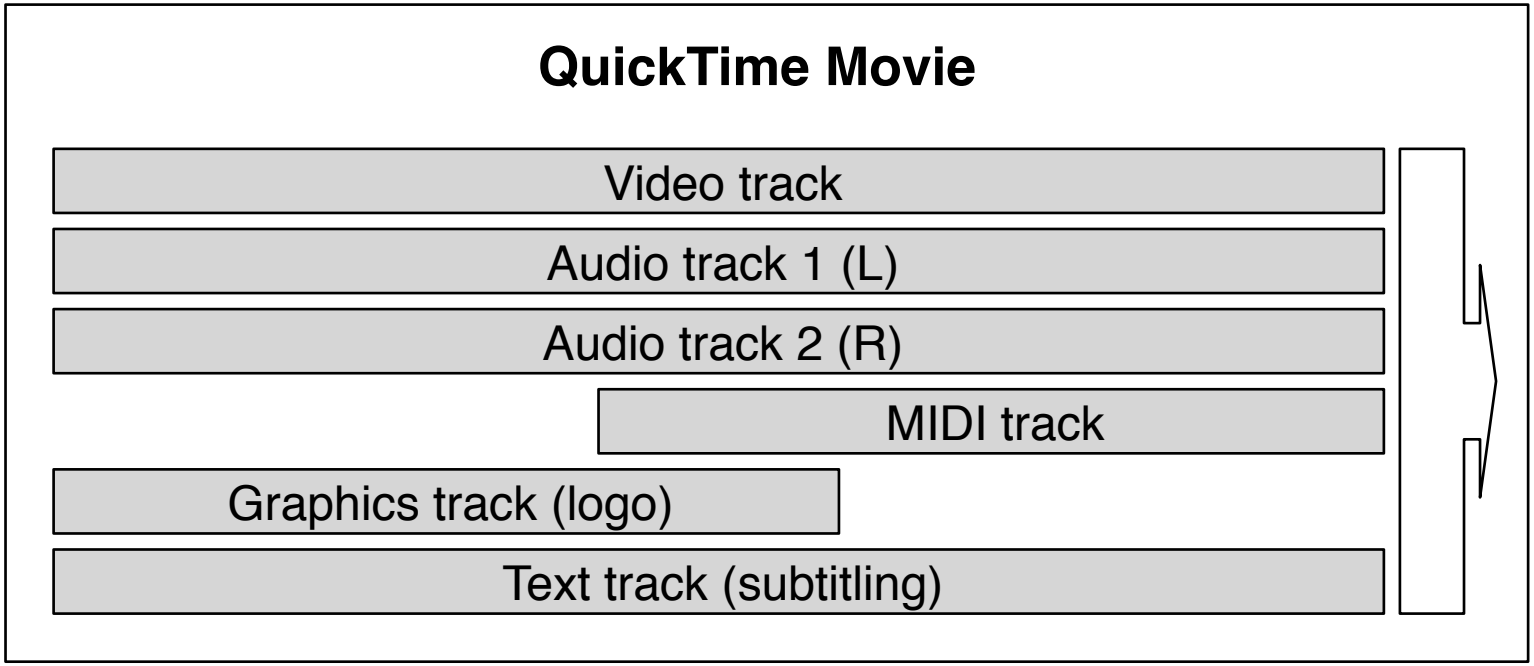


Apple QuickTime

- Platform-independent multimedia architecture
  - Supports MacOS and Windows
- History:
  - First version (1.0) for MacOS 6 in 1991
  - QuickTime 2.0 also for Windows 1994
  - Streaming support since version 4.0 (1999)
  - Current version 10.0 (or "X") for MacOS 10.6
- QuickTime consists of:
  - Framework, API, file format
- Applications using QuickTime:
  - QuickTime Player, iTunes, Logic, Final Cut, Premiere, Avid, ...
- Codec support:
  - Open plugin architecture (*QuickTime components*)
  - Huge selection of pre-installed components
- Digital Rights Management is intrinsic part of QuickTime

QuickTime Movie Files

- Modular and flexible architecture
  - Multimedia files organized in tracks
  - Example:



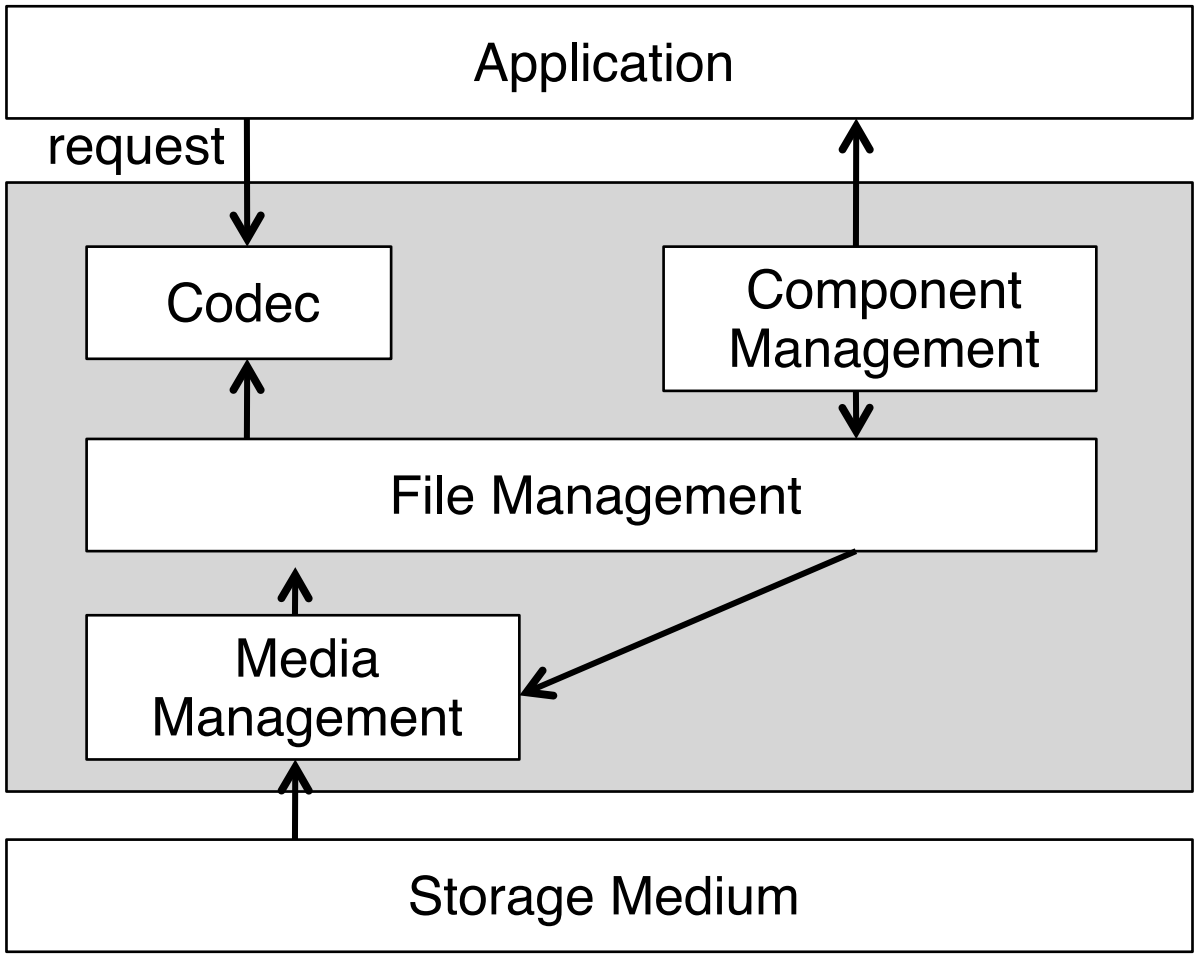
Types of QuickTime Tracks

- Movie track: Copyright info, annotations, ...
- Audio track(s)
- Text track: Titles, subtitles, credits, notes, ...
- Sprite track: Images with animatable, programmable behaviours
- Flash track: SWF animation
- QuickTime VR track: VR objects, panorama movies
- Video track: Digital video, 3D animation, ...
- Music track: MIDI
- Chapter track: Inserts addressable entry points
- 3D track: Contains QuickDraw 3D metafile objects
- Streaming track: References to streams from a server source
- Hint track: Additional information for streaming (see below)

Typical QuickTime Codecs

- Video
  - Animation: Specially for large colored areas
  - BMP: High-quality import/export
  - Cinepak: Exchange format for other platforms and old versions
  - Component video (3:2 compression, YUV colour model)
    - » High quality capturing with low CPU load
  - DV Stream (10:1 compression)
    - » Specially for capturing from DV devices (over FireWire)
  - H.261, H.263, H.264
  - Sorenson Video: Excellent compression/quality rate
  - ...
- Audio
  - IMA 4:1 (CD music quality with 80 kbps)
  - QDesign Music (comparable to MP3)
  - QualComm PureVoice

QuickTime Media Abstraction Layer

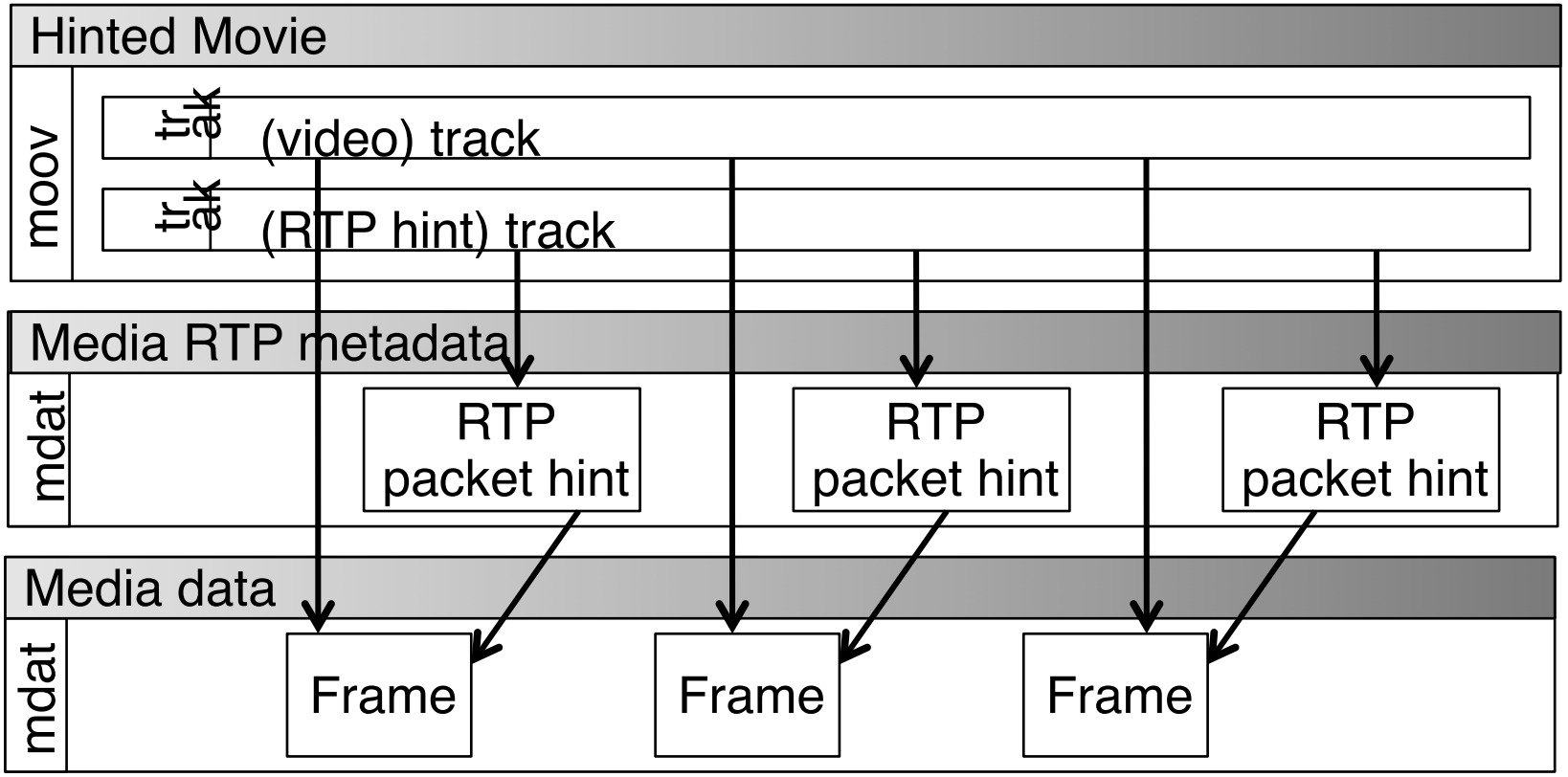


QuickTime Toolbox (Media Abstraction Layer)

Interactivity in QuickTime

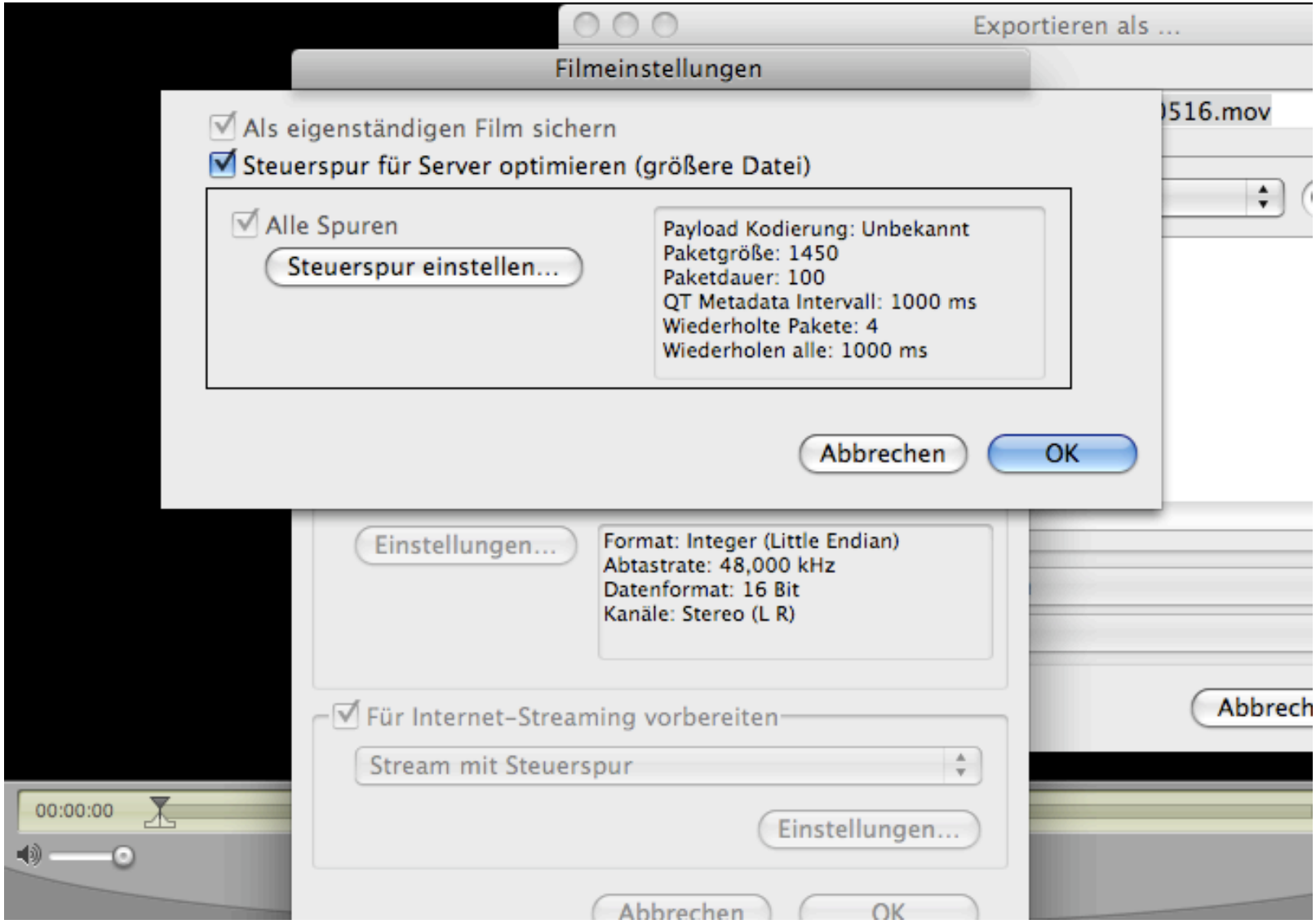
- Movies or video tracks as clickable links
  - “HREF tracks” allow dynamic change of target during runtime
- Sprites can react to system events (like mouse movement)
- Flash animations can be fully integrated
  - e.g. movie controllers developed in Flash
  - more efficient animations (vector-based)
- JavaScript support
- Full QuickTime API
  - Available also for Java (QuickTime for Java)

Hint Tracks in QuickTime and MPEG-4



- Hint track gives server software pointers to the RTP information to serve the relevant media chunks
- Concept from QuickTime, integrated in MPEG-4 (streaming)

Adding a Hint Track *(Steuerspur)*



History of Windows Media

- 1991-1992: Multimedia extensions of Windows 3.0
  - Playback of AVI
- 1992-1996: Netshow 1.0
  - First experiments with OnDemand Streaming
- 1997-1999: Windows Media Technologies in Windows 98
  - Completely reworked version of NetShow
- Since 2000:
  - Strong emphasis on compression/quality ratios comparable to competitors Real and Apple
- Since 2006:
  - Windows Media Services 9, Windows Media Player 11
- Latest version 2009:
  - Windows Media Services 2008, Windows Media Player 12 (for Windows 7)
- Digital Rights Management is intrinsic part of Windows Media

Windows Media and MPEG-4

- Windows Media MPEG-4 Video V3
  - MPEG-4 based Microsoft implementation
  - Probably used as a quick-start to catch up with competitors
- Windows media Video V7 and higher
  - No longer compatible with standard MPEG-4
- ISO MPEG-4 Video
  - Special, additional codec for MPEG-4 standard

Market Trends for the Future

- Generally, integrating platforms become more important than individual players/formats
  - QuickTime player can play WMA/WMV and Flash video using plugins
  - iTunes comprises QuickTime player (and more)
  - Windows media player (11 and up) as control device for digital media receiver in the home
- New competitors enter the market all the time
  - E.g. playing video through Adobe Flash (YouTube), HTML5
  - E.g. using game consoles as video streaming platform (Xbox 360)
- High definition content is coming up
  - Major players can already deal with HD content