Outline

1. Introduction and Motivation
2. Interactive Web Applications
3. Web Programming with Java
4. Communities, the Web, and Multimedia
5. Digital Rights Management
6. Cryptographic Techniques
7. Multimedia Content Description
8. Electronic Books and Magazines
9. Multimedia Content Production and Management
10. Streaming Architectures
11. Web Radio, Web TV and IPTV
12. Multimedia Conferencing
13. Signaling Protocols for Multimedia Communication
14. Visions and Outlook

Part I: Web Technologies for Interactive MM
Part II: Content-Oriented Base Technologies
Part III: Multimedia Distribution Services
Part IV: Conversational Multimedia Services
10 Streaming Architectures

10.1 High-Level Streaming Architecture
10.2 Real-Time Data Transport
10.3 Scalability and Multicast
10.4 Selected Commercial Streaming Architectures

Literature:
- David Austerberry: The Technology of Video & Audio Streaming, Focal Press 2002
- Gregory C. Demetriades: Streaming Media, Wiley 2003
A Classification of Multimedia Services

- According to ITU-T recommendation I.211 “B-ISDN Service Aspects”

Interactive Services
- Conversational Services
- Messaging Services

Distribution Services
- without user-individual presentation control
- with user-individual presentation control
Presentation Control in Distribution Services

- **Without** user-individual presentation control:
  - Continuous “live” stream of information from sender to receiver(s)
  - Replacement of other distribution media (e.g. radio) by digital networks
    » Real-time service (e.g. fixed start time for transmission)

- **With** user-individual presentation control:
  - Pause, resume, skip backward, skip forward
  - **Server** control:
    » Individual stream (or group of closely related streams)
  - **Client-local** control:
    » Outdated solution for interactivity: “Near Video-on-Demand” (NVOD)
      = Staggered broadcast of multiple transmissions of the same content
    » Time-shifted recording enhances interactivity (pause/resume)
  - Full transmission of video content on local storage before playback enables full interactivity
    » Modern solutions enabled by high-capacity local disks
      (e.g. AppleTV, Sky Anytime)
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

Bandwidth Economy

• Fully heterogeneous individual requests:
  – Required bandwidth = stream bandwidth x number requests

• Homogeneity of request helps saving bandwidth:
  – Same content for many clients, but different playback times:
    » Broadcast with local buffering
  – Same content at same time for many clients
    » Multicast (splitting streams)

• Pre-planning saves bandwidth
  – (Individual) transmission of pre-booked content in non-real time
    (“download and play”)
Streaming, Streaming Media

- Streaming media means real-time delivery of moving images, moving text and sound, from a server to client (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

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 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

A/V Capture → Encoding → Storage → Streaming → Distribution → Viewing → Serving
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
  – Pull model:
    » Request-response driven similar to Web server
  – Push model:
    » Server sends packets at regular intervals
• Interactivity
  – VCR-like control (PLAY, STOP, PAUSE, FFWD, REW)
  – Random access based on various criteria
  – Hyperlinks in A/V material (“hypervideo”)
• Transcoding
  – Encoding adapted to client capabilities
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)
10 Streaming Architectures

10.1 High-Level Streaming Architecture

10.2 Real-Time Data Transport

10.3 Scalability and Multicast

10.4 Selected Commercial 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
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
  – In particular for pull-model true streaming
  – 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?

• **Pull** model:
  – Client requests next slice of audio/video information from server
  – Server responds with audio/video data

• Very popular in Web-based multimedia services
  – e.g. YouTube video
  – e.g. MP3-Streaming with ShoutCast, Live365

• Advantage:
  – Ubiquitous protocol, no problems with firewalls etc.

• Main problems:
  – HTTP usually run over TCP, creates overhead
  – Progressive download rather than true streaming
  – Sometimes user has to select requested quality/bitrate

• Consequence:
  – Specialized streaming protocols: “Adaptive HTTP Streaming” (see later)
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

<table>
<thead>
<tr>
<th>Protocol</th>
<th>Header Size</th>
<th>Payload Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>RTP</td>
<td>12 Bytes</td>
<td>20 – 150 Bytes</td>
</tr>
<tr>
<td>UDP</td>
<td>8 Bytes</td>
<td>or multiple RTP packets</td>
</tr>
<tr>
<td>IP</td>
<td>20 Bytes</td>
<td></td>
</tr>
</tbody>
</table>
## RTP Header Format

<table>
<thead>
<tr>
<th>Payload Type (7 Bit)</th>
<th>Sequence Number (16 Bit)</th>
<th>…</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Timestamp (32 Bit)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Synchronization Source (SSRC) Identifier (32 Bit)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Contributing Source (CSRC) Identifier (32 Bit) (repeated)</td>
<td></td>
</tr>
</tbody>
</table>

- **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

C: Client | W: Web Server | A: Audio Server | V: Video Server

GET
OK
SETUP
OK
SETUP
OK
PLAY
OK
PLAY
OK
TEARDOWN
OK
TEARDOWN
OK
10 Streaming Architectures

10.1 High-Level Streaming Architecture
10.2 Real-Time Data Transport
10.3 Scalability and Multicast
10.4 Selected Commercial 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 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, 2012)
    » Runs 95,000+ 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
Example: Visualizing Akamai

Akamai handles 20% of the world’s total Web traffic, providing a unique view into what’s happening on the Web - what events are generating traffic, how much, from where, and why. Bookmark this page to get a feel for the world’s online behavior at any given moment - how much rich media is on the move, the sheer volume of data in play, the number and concentration of worldwide visitors, and average connection speeds worldwide.
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

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
10 Streaming Architectures

10.1 High-Level Streaming Architecture
10.2 Real-Time Data Transport
10.3 Scalability and Multicast

10.4 Selected Commercial Streaming Architectures

RealNetworks, QuickTime, Windows Media

Literature:


David Austerberry: The Technology of Video & Audio Streaming, Focal Press 2002
Usage of Video Streaming (1)

Online Video is Now a Huge Global Market

1.2 Billion people watched an average of 18 hours each of Online Video in October Worldwide

Marc Gosschalk, comscore.com, October 2011
Increasing Size of Video Files

Growth in length and in quality!

Source: WebSiteOptimization.com
Usage of Video Streaming (2)

Rapid Growth of Unique Viewers and Video Views Stopped in Q1 2010 for EU-3

EU-3 = Germany, France, UK

Marc Gosschalk, comscore.com, October 2011
Usage of Video Streaming (3)

Marc Gosschalk, comscore.com, October 2011
Usage of Video Streaming (4)

Marc Gosschalk, comscore.com, October 2011

The Last Six Months have seen the Birth of a New Era of Online Video

<table>
<thead>
<tr>
<th>Videos (Billions)</th>
<th>Unique Viewers (Millions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feb-2008</td>
<td>Aug-2008</td>
</tr>
<tr>
<td>Aug-2008</td>
<td>Feb-2009</td>
</tr>
<tr>
<td>Feb-2009</td>
<td>Aug-2009</td>
</tr>
<tr>
<td>Aug-2009</td>
<td>Feb-2010</td>
</tr>
<tr>
<td>Feb-2010</td>
<td>Aug-2010</td>
</tr>
<tr>
<td>Feb-2011</td>
<td>Aug-2011</td>
</tr>
</tbody>
</table>

- Minutes per Video: +4%
- Total Minutes: +33%
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 2012
Video Codec Usage

from WebSiteOptimization.com

RV = Real Video
WMV = Windows Media Video
Streaming File Formats

- 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)
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 10.1 for MacOS 10.7 (and 7.7.1 for Windows)

• 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:

```
QuickTime Movie

Video track
Audio track 1 (L)
Audio track 2 (R)
MIDI track
Graphics track (logo)
Text track (subtitling)
```

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)
QuickTime Media Abstraction Layer

Application

Codec

Component Management

File Management

Media Management

Storage Medium

QuickTime Toolbox (Media Abstraction Layer)
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)*
Adapting to Network Congestion

Media files

low

medium

high

... bandwidth

bandwidth selection

time reference

constant frame rate

Player reports

Player commands (setup, play, pause, teardown)

RTSP

RTCP

Media file parse

RTP packetizer

RTCP

RTSP

Player commands (setup, play, pause, teardown)
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
Trend: Adaptive HTTP Streaming

• Conventional streaming technologies (push-model, based on RTP, RTSP etc.):
  – Apple QuickTime Streaming
  – Microsoft Windows Media
  – Adobe Flash

• Emerging adaptive streaming technologies (pull-model, based on HTTP):
  – Apple HTTP Live Streaming
  – Microsoft Smooth Streaming
  – Adobe HTTP Dynamic Streaming

• Basic idea: Small video file fragments
  – Fragments exist at different bit rates
  – Index file for all available fragments
  – Client requests appropriate next fragment by GET request
3GP-DASH: Open Standard for HTTP Streaming

- Dynamic Adaptive Streaming over HTTP

**Figure 3 Solution overview – 3GP-DASH**

Source: Thomas Stockhammer, Qualcomm