

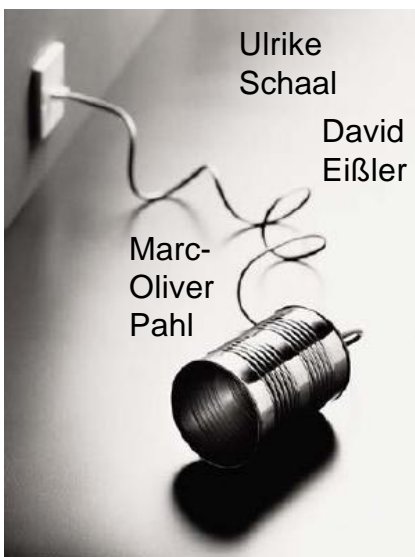


MultiMediaPraktikum

Universität Tübingen Wilhelm-Schickard-Institut für Informatik



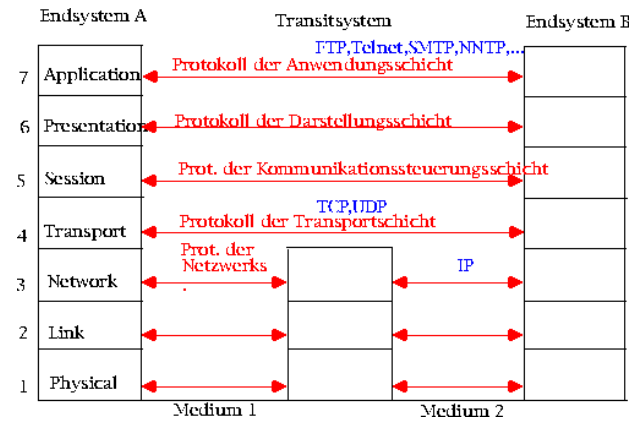
Wintersemester 2002/ 2003
Aufgabe 5 - Videokonferenz



Inhaltsverzeichnis

Theorie	4
UDP	4
TCP	5
UDP vs. TCP	6
RTP	6
h.32x	8
h.323H.32x	8
SIP	10
SIP/ h.323	11
Videokodierungsverfahren bei Videokonferenzen	12
Audiokodierungsverfahren bei Videokonferenzen	13
Praxis	15
i-visit	15
netmeeting	16

Theorie



Der Austausch von Daten über ein Netzwerk lässt sich durch das ISO/OSI-Schichtenmodell beschreiben.

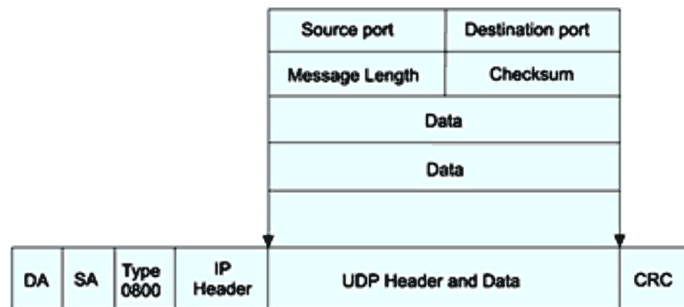
Mindestens einige der dargestellten Layer finden sich in jeder realen Implementierung/ Umgebung.

Zum Transport von Daten in einem Netzwerk sind Übertragungsprotokolle notwendig. Die untersten drei Layer (und die zugehörigen Protokolle) sind in Hardware realisiert, die Transportschicht ist meist das erste Softwareprotokoll. Von dieser Schicht übernehmen die Anwendungen dann ihre Nutzdaten.

UDP

User Datagram Protocol

UDP ist das User Datagram Protocol. Es bietet ein Minimum an Funktionalität zum Versenden von Daten. Das sieht man schon am Header:



Er beinhaltet nur Source-, Destination-Port, die Message-Length und eine Checksum, die sowohl über die Header-Daten, als auch über die Nutzdaten geht.

Es ist keine Möglichkeit vorgesehen, sicherzustellen, dass die Daten ankommen. Dies zu überprüfen obliegt der darüber liegenden Anwendung.

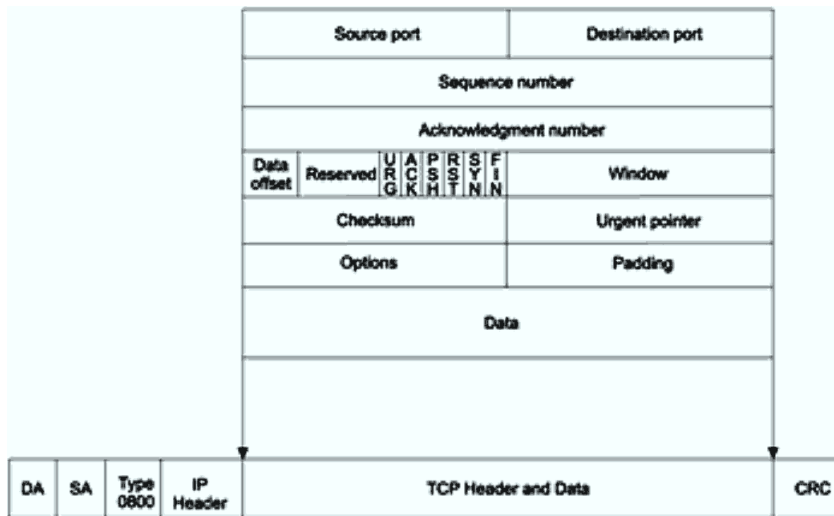
Klarer Vorteil von UDP ist die Geschwindigkeit. Mit weniger Overhead können die Daten nicht versendet werden.

UDP wird vor allem bei Multimedia-Anwendungen eingesetzt, wo es nicht so sehr darauf ankommt, dass jedes Paket ankommt, sondern wichtiger ist, dass der Durchsatz hoch ist.

Der TCP-Header sieht folgendermaßen aus:

TCP

Transmission Control Protocol



Wie man sieht übernimmt das TCP-Protokoll sehr viele Aufgaben, die bei UDP die Anwendung erledigen muss.

Es gibt eine „Sequence Number“, mit deren Hilfe sichergestellt wird, dass die Daten beim Empfänger (dessen über TCP liegenden Anwendung) in der richtigen Reihenfolge ankommen.

Desweiteren stellt TCP sicher, dass die Daten auch beim Empfänger ankommen, indem mithilfe der „Acknowledgement number“ der Empfang innerhalb einer festgelegten Zeit bestätigt werden muss. Bleibt die Bestätigung aus, wird das Paket erneut gesendet.

Der „DataOffset“ gibt die Länge des Headers an.

Die Bits danach dienen der Flusskontrolle:

- URG: Urgent Pointer [Headerfeld] ist gültig.
- ACK: Quittierungsnummer [Headerfeld] ist gültig.
- PSH: Die Daten sollen sofort gesendet werden (Push-Funktion).
- RST: Die Verbindung soll zurückgesetzt werden (Reset).
- SYN: Ein Verbindungsaufbau ist gewünscht, muss quittiert werden.
- FIN: Ein Verbindungsabbau ist gewünscht, muss quittiert werden.

„Window“ gibt an, wie groß das Empfangsfenster ist, das heißt, wie viele Pakete der Empfänger aufnehmen kann (wie viele Pakete dürfen gesendet werden, ohne das eine Bestätigung eingegangen ist). Diese Größe wird dynamisch angepasst.

Die Checksum geht wieder über Daten und Header.

Der „Urgent pointer“ ist eine Möglichkeit, wichtige Daten (z.B. Interrupts) sofort an den Empfänger zu versenden (ohne auf Platz in einem Paket zu warten). Die dringenden Daten werden an die Nutzdaten angehängt. Der Empfänger sollte in der Lage sein, solche Daten sofort zu verarbeiten.

In „Options“ können weitere Eigenschaften festgelegt werden, wie z.B. MSS (Maximum Segment Size), WSopt (Window Scale), ...

„Padding“ gibt an, ob die Nutzdaten auf die volle Blockgröße angefüllt werden.

UDP vs. TCP

Beide Schicht-4-Protokolle unterscheiden sich also grundlegend schon in ihrer Konzeption:

UDP arbeitet verbindungslos und stellt den Transfer nicht sicher. Klarer Vorteil ist die Geschwindigkeit.

TCP arbeitet verbindungsorientiert und stellt sicher, dass beim Empfänger auch alle Daten ankommen, die vom Versender geschickt wurden. Der Preis der Sicherheit ist ein deutlich höherer Overhead bei den Daten und ein größerer Aufwand in der Protokollschicht.

Die vermeintlichen Vorteile von TCP gegenüber UDP sind für Multimediaanwendungen wie Streaming oder Videokonferenz Nachteile.

Hier kommt es nicht darauf an, dass jedes Paket ankommt, sondern dass die Pakete rechtzeitig ankommen.

Für Realtime-Dienste wird daher in der Regel nicht TCP verwendet. Stattdessen kommen Protokolle wie das im Folgenden beschriebene RTP zum Einsatz, die auf „leichteren“ Protokollen wie UDP aufbauen und die spezifischen Eigenschaften von Multimediadaten berücksichtigen.

http://ethernet.industrial-networking.com/articles/i04_georgethomas.asp

<http://goethe.ira.uka.de/seminare/internet/tcp+udp/#ToC1>

<http://www.substream.org/streaming-netzwerk3.html>

RTP

Realtime Transport Protocol

Das Realtime Transport Protocol ist ein durch den RFC 1889 als Internet-Standard festgelegtes Protokoll zum Transport von Echtzeitdaten über das Netz. Es stellt eine Ende-zu-Ende-Verbindung (verbindungsorientiert) zwischen zwei oder mehr Rechnern her.

Zum eigentlichen Transport der Daten verwendet RTP zumeist UDP, kann aber auch auf TCP, ATM oder ähnliche Protokolle aufsetzen. RTP erweitert das jeweilige Protokoll um die echtzeitspezifischen Erfordernisse.

Es kann im *unicast*- oder im *multicast*- Modus betrieben werden.

Unicast bedeutet, dass der Sender selbst sich darum kümmert, dass die Daten an alle Empfänger versendet werden.

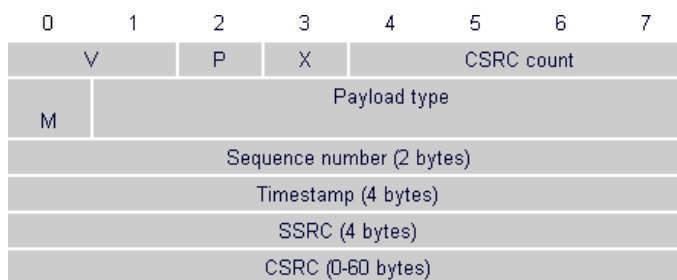
Bei Multicast werden die Daten nur einmal gesendet und das Netzwerk kümmert sich um die Verteilung, was auch eine Verringerung des Traffic mit sich bringen kann, da die Daten erst beim letzten gemeinsamen Router an die einzelnen Teilnehmer kopiert werden.

Die Verbindungsqualität wird mit sogenannten RTCP-Paketen kontrolliert, die in bestimmten Abständen verschickt werden.

Innerhalb des der Nutzdaten (Payload) befindet sich zumeist ein weiterer anwendungsspezifischer Header, der anwendungsspezifische Informationen zu den Daten (z.B. Bildgröße o.ä.) bereitstellt.

Das RTP selbst definiert also nur ein Minimum an Parametern.

Der RTP-Header:



<i>V</i>	RTP-Version.
<i>P</i>	Padding an? Wenn ja werden die Nutzdaten (Payload) auf die volle Paketgröße aufgefüllt.
<i>X</i>	Header-Extension? Wenn ja folgt die dem Header.
<i>CSRC</i>	Anzahl der ContributingSource-Identifizier, die dem Header folgen.
<i>M</i>	Marker, der der Anwendung z.B. das Ende einer Einheit anzeigen kann.
<i>Payload type</i>	ID für die Art der Nutzdaten (z.B. 34 = H.263).
<i>Sequence number</i>	Die Pakete werden durchnummeriert, damit der Empfänger die Originalreihenfolge rekonstruieren kann.
<i>Timestamp</i>	gibt die Zeit an zu der das erste Byte galt. Die Zeit kann z.B. relativ zum Start des Streams sein.
<i>SSRC</i>	Identifiziert die Synchronisationsquelle (es kann mehrere geben).
<i>CSRC</i>	Identifiziert die ContributingSource der einzelnen Pakete im Stream.

http://www.protocols.com/voip/rtp_header.htm

Bei Videokonferenzen bietet es sich an, bei seiner Anwendung ein RTP-konformes (darauf aufbauendes) Protokoll zu implementieren.

RTP kümmert sich dabei darum, dass die Frames in der richtigen Reihenfolge abgespielt werden (Sequence number) und zur richtigen Zeit (Timestamp). Weitere Parameter müssen dann im Payload spezifiziert werden.

Da die Payload-Typen auch standardisiert sind, kann man so Kompatibilität sicherstellen.

h.32x

Die h.32x sind eine Familie von Standards zur Videotelefonie über verschiedene Netzwerke. Diese wurden vom International Multimedia Teleconferencing Consortium (IMTC), dem 125 Firmen aus 20 Ländern (1998) angehören, innerhalb der International Telecommunications Union (ITU) verabschiedet.

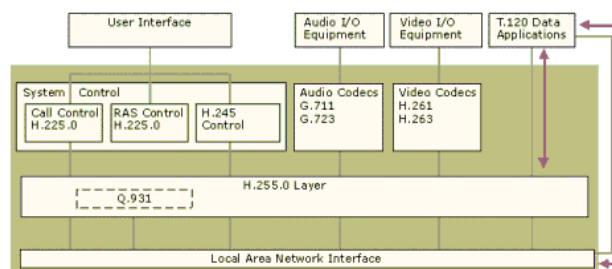
Innerhalb des Standards gibt es folgende Familienmitglieder:

- H.320 Sprache/Video über ISDN
- H.321 Sprache/Video über ATM
- H.322 Sprache/Video über LANs mit QoS
- H.323 Sprache/Video über Packet Switched Networks
- H.324 Sprache/Video über analoges Telefon oder Mobilfunk

h.323H.32x

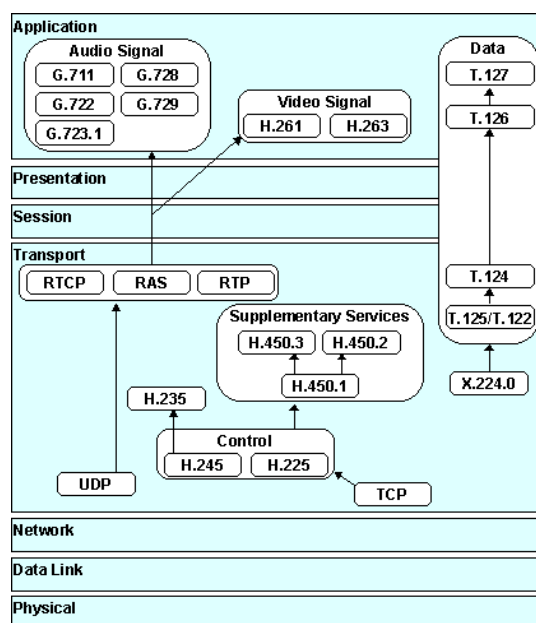
h.323 ist dabei der Standard für Netzwerke mit nicht garantierter QoS (Quality of Service). Dazu zählen das Internet, aber auch die meisten LANs oder WANs in Firmen. Dem h.323-Protokoll liegt RTP zugrunde.

Eine h.323-Implementierung sieht z.B. so aus:



Quelle: Microsoft (z.B. Netmeeting)

Folgende Unterprotokolle (hier im ISO/ OSI-Modell dargestellt) sind in h.323 zusammengefasst:



Quelle: <http://www.protocols.com/pbook/h323.htm>

ITU-Standard h.323 Unterprotokolle (nur die oben genannten gehören zum Standard; der Rest steht nur als Erklärung zur Grafik links da):

Audio: G.711, G.722, G.723.1, G.728, G.729

- G.711** Pulse Code Modulation of voice frequencies (PCM), where 3.1 kHz analogue audio is encoded into a 48, 56 or 64 kbps stream. Used when no other standard is equally supported.
- G.722** 7 kHz audio encoded into a 48, 56 or 64 kbps stream. Provides high quality, but takes bandwidth.
- G.722.2** Coding of speech at around 16 kbps using Adaptive Multi-Rate Wideband, **AMR-WB**. Five mandatory modes, 6.60, 8.85, 12.65, 15.85 and 23.85 kbps.
- G.723.1** Dual rate speech codec for telecommunications at 5.3 kbps & 6.4 kbps.
- G.728** Low Delay Code Excited Linear Prediction (LD-CELP) where 3.4 kHz analogue audio is encoded into a 16 kbps stream. This standard provides good quality results at low bitrates.
- G.729 A/B** Speech codec that provides near toll quality audio encoded into an 8 kbps stream using the AS-CELP method. Annex A is a reduced complexity codec and Annex B supports silence suppression and comfort-noise generation.

Video: H.261, H.263

- H.261** video codec for audiovisual services at p x 64Kbps.
- H.263** video codec for narrow telecommunications channels at < 64 Kbps.

Daten: T.120

- T.120** defines protocols and services for Multipoint Data Conferencing.
- T.121** Generic Application Template (GAT). Defines a template as a guide for developers in managing T.120 resources.
- T.122** defines Multipoint Communication Services (MCS) available to developers.
- T.123** defines Network Specific Data protocol for MultiMedia conferencing.
- T.124** defines Generic Conference Control (GCC), mandatory for 'group' conferences.
- T.125** defines MCS data transmission protocol.
- T.126** Multipoint Still Image and Annotation protocol. Defines the protocol used to provide interoperability with graphics data in applications such as whiteboarding, annotated image exchange, screen sharing and remote apps control.
- T.127** Multipoint Binary File Transfer protocol. Defines the protocol used to support binary file transfer within a conference.
- T.128** defines Multipoint Application Sharing protocol (also known as T.SHARE)
- T.134** defines Multimedia Application Text Conversation protocol (also known as T.CHAT).
- T.135** User-to-reservation system transactions within T.120 conferencing.
- T.136** Remote device control application protocol.
- T.137** Virtual meeting room management - services & protocol.
- T.140** Protocol for multimedia application text conversation.

Control: H.225, H.245

- H.225** defines the multiplexing transmission formats for media stream packetisation & synchronisation on a non-guaranteed QoS LAN.
- H.235** Security and encryption for H-Series multimedia terminals.
- H.245** defines the control procedures and protocol for H.323 & H.324 multimedia communications.

<http://www.teamsolutions.co.uk/tsstds.html>

SIP

Wie man schon an den beinhalteten Protokollen sieht, ist eine h.323-Implementierung sehr komplex. Es müssen zwar nicht alle, aber dennoch eine Reihe der aufgelisteten Protokolle implementiert werden.

Da es so viele mögliche Protokolle gibt, ist der unnötige Overhead, wenn man nur ganz bestimmte Dienste nutzen will (z.B. nur VoiceOverIP) recht groß.

Als Adaption des h.320-Standards (ISDN) auf paketorientierte Netzwerke ist h.323 vermutlich auch nicht optimal an deren Gegebenheiten angepasst (z.B. n x 64K Video bei h.261).

Für die Verbreitung von Telefonie über das Internet ist auch das relativ aufwendige Verfahren zum Herstellen einer Verbindung und die damit einhergehende Wartezeit nicht akzeptabel.

Deshalb wurde der SIP-Standard (Session Initiation Protocol; 1999 RFC2543) entwickelt, der den Verbindungsaufbau vereinfacht und es auch sehr leicht ermöglicht, dass ein Teilnehmer seine „Telefonnummer“ (z.B. e-Mailadresse) mitnimmt.

SIP bietet dabei ähnlich http einen Standard zum Verbindungsaufbau und greift auf schon vorhandene Infrastruktur des Internet (URI, DNS, MIME) zurück.

Der Standard wurde absichtlich sehr einfach gehalten, damit er auch problemlos und billig in Endgeräte eingebaut werden kann, was dann z.B. in einer der nächsten Mobilfunkgenerationen (3gpp) auch geschehen soll, womit der Durchbruch dann geschafft wäre.

<http://www.dynamicsoft.com/sip/index.php>
<http://www.teamsolutions.co.uk/tsstds.html>

Die konzeptionellen Grundlagen von SIP sind:

- vorrangig entworfen, um Verbindungsaufbau zu ermöglichen
- geringe Anforderungen an Infrastruktur (Verwaltung der Sitzungsinformationen wird konsequent ins Endgerät verlagert)
- kein Support für Multimedia- & Datenanwendungen
- Möglichkeiten zur Konferenzkontrolle fehlen
- Jeder Nutzer hat einen Protokoll-Client (initiiert Anrufe) und einen Protokoll-Server (beantwortet Anrufe).
- Der Netzwerkservers, welcher aus Proxy- und Redirect-Server besteht, ist für die Weiter- und Umleitung zuständig.

wesentliche Kommandos:

- „Invite“ initiiert einen Anruf
- „Options“ kann Zusatzinformationen über den Anwender enthalten
- „Ack“ bestätigt den zuverlässigen Nachrichtenaustausch
- „Register“ übermittelt Standortinformationen an einen SIP-Server
- „Cancel“ versucht einen bereits gesendeten Request zu löschen
- „Bye“ beendet die Konferenz

Vorteile:

- SIP kann einen eingehenden Anruf auf mehrere Endgeräte aufteilen (sogenanntes „Forking“)
- schneller Rufaufbau durch gleichzeitiges Übertragen der Anforderung einer Session und Eigenschaften des Clientes
- Adressierungsschema sehr flexibel: nutzt Informationen aus existierenden DNS-Systemen
- Adresse ist e-mail ähnlich: user@company.com
- Unterstützung durch Windows XP und UMTS

Nachteile:

- Sicherheit des Protokolls noch verbesserungswürdig
- Forking funktioniert zwischen verschiedenen Herstellern noch unzureichend
- Integration in eine Nebenstellenanlage ist noch schwierig
- Voice-Mail ist potentieller Stolperstein
- NAT und niedrige Bandbreite bereiten Probleme

<http://vcc.urz.tu-dresden.de/Grundlagen/sip.html>

Da SIP extra für „erweiterte Telefondienste“ geschaffen wurde und viel weniger spezifiziert, hinkt der Vergleich mit h.323 etwas, zeigt aber die Vorteile von SIP auf diesem Gebiet auf.

SIP/ h.323

Die Hauptpunkte sind:

	SIP	h.323
Encoding	textual (textual coding leads to less compression but remains parsable for Switches etc.)	binary
Call Setup delay	1,5 RTT (RoundTripTime)	<= 7 RTT
Complexity	low (h.323 specifies much more...)	high
Extensibility	content not specified => open	only limited possibilities
Architecture	modular	monolithic
Inter-Domain routing	especially supported	complicated
Standardisation	low; only protocol (37 headers)	nearly everything

Videokodierungsverfahren bei Videokonferenzen

Bei Videokonferenzen werden heute z.B. folgende Videokompressionsverfahren eingesetzt:

h.261 (muss von jeder h.323-Implementierung unterstützt werden)

Eingangsdaten werden im Common Intermediate Format (CIF - 288 Zeilen x 352 Spalten) oder aCIF (quarter CIF - halbe Auflösung in vertikaler und horizontaler Richtung) an den Kodierer übergeben (um Inkompatibilitäten zwischen verschiedenen Formaten wie NTSC und PAL zu vermeiden).

Die Unterstützung von qCIF (es werden auf 4 Y-Helligkeitsblöcke ein Cb- und ein Cr-Farbblock übertragen, die Datei ist hierarchisch in Blöcke, Makroblöcke und Blockgruppen eingeteilt) ist für alle Encoder und Decoder von H.261 verpflichtend, CIF ist optional.

- für Netzwerke mit Bandbreiten von 64 kBit/ s bis 2 MBit/ s

Vorteile:

- bei jedem h.323 konformen Gerät dabei (weit verbreitet)

Nachteile:

- nur zwei Auflösungen (CIF, QCIF)
- out of date (Niedrige Auflösung, zu schlechte Qualität)

h.263 (kann von einer h.323-Implementierung unterstützt werden)

Eine abwärtskompatible Weiterentwicklung von H.261, die zum einen die Qualität deutlich verbessert und zum zweiten durch eine Huffman-Code-Tabelle eine Optimierung für schmalbandige Verbindungen (<64kBit) mit sich bringt.

- für Netzwerke mit Bandbreiten von 20 kBit/ s bis 2 MBit/ s

Vorteile:

- kann h.261 abspielen
- deutlich bessere Bildqualität

Nachteile:

- nur wenige Auflösungen werden unterstützt

mpeg4

Weiterentwicklung von h.263, MPEG-1 und MPEG-2.

- 5 kBit/ s bis > 1 GBit/ s (konstant und variabel)

Vorteile:

- hohe Qualität möglich (Skalierbarkeit!)
- sehr variabel einsetzbar
- interaktive Anwendungen möglich
- inhaltsbasierte Kodierung (dadurch z.B. Empfang von nur bestimmten Teilen möglich)

Nachteile:

- hohe Rechenleistung erforderlich (>> h.263)
- Lizenzgebühr
- Verschiedene Implementierungen möglich (=> inkompatibel)

www.itsc.org.sg/synthesis/2001/itsc-synthesis2001-csp-video-over-internet.pdf

Da h.323 weit verbreitet ist (Netmeeting) folgt nochmal eine Auflistung von dessen Audio-Protokollen und deren Eigenschaften:

Audiokodierungsverfahren bei Videokonferenzen

Name	description	bit rate (kb/s)	sampling rate (kHz)	frame size (ms)
G.711 mu-law (US, Japan) and A-law (Europe) companding	Pulse code modulation (PCM)	64	8	sample
G.721 Now described in G.726; obsolete.	Adaptive differential pulse code modulation (ADPCM)	32	8	sample
G.722 Subband-coder that divides 16 kHz band into two subbands, each coded using ADPCM	7 kHz audio-coding within 64 kbit/s	64	16	sample
G.722.1	Coding at 24 and 32 kbit/s for hands-free operation in systems with low frame loss	24/32	16	20
G.723 Superseded by G.726; obsolete. This is a completely different codec than G.723.1.	Extensions of Recommendation G.721 adaptive differential pulse code modulation to 24 and 40 kbit/s for digital circuit multiplication equipment application	24/40	8	sample
G.723.1 Part of H.324 video conferencing. DSP Group . It encodes speech or other audio signals in frames using linear predictive analysis-by-synthesis coding. The excitation signal for the high rate coder is Multipulse Maximum Likelihood Quantization (MP-MLQ) and for the low rate coder is Algebraic-Code-Excited Linear-Prediction (ACELP).	Dual rate speech coder for multimedia communications transmitting at 5.3 and 6.3 kbit/s	5.6/6.3	8	30
G.726 ADPCM; replaces G.721 and G.723	40, 32, 24, 16 kbit/s adaptive differential pulse code modulation (ADPCM)	16/24/32/40	8	sample
G.727 ADPCM. Related to G.726.	5-, 4-, 3- and 2-bit/sample embedded adaptive differential pulse code modulation (ADPCM)	var.	?	sample
G.728 CELP. Annex J offers variable-bit rate operation for DCME.	Coding of speech at 16 kbit/s using low-delay code excited linear prediction	16	8	
G.729 Low delay (15 ms)	Coding of speech at 8 kbit/s using conjugate-structure algebraic-code-excited linear-prediction (CS-ACELP)	8	8	10

Daten übernommen von <http://www.cs.columbia.edu/~hgs/audio/codecs.html>

Alle aufgeführten Formate sind speziell für die Übertragung von Sprache gemacht und bieten, wie man an der Tabelle sieht, unterschiedliche Kodierungsverfahren und Samplingraten, was sich natürlich auf Qualität und Datenrate, aber auch auf die Ende-zu-Ende Verzögerung auswirkt.

G.711 ist der Codec, der für ISDN-Telefonie eingesetzt wird:

Datenrate: fest bei 64kbps

- + sehr gute Qualität für Sprache
- hohe und konstante Datenrate notwendig

G.722 ist eine Verbesserung des G.711 und verwendet anstelle der pulse code modulation (PCM) adaptive differential pulse code modulation (ADPCM), kodiert also nur die Differenzen und benutzt einen Prädiktor:

Datenrate: fest bei 64kbps

- + 16kHz Samplingrate, also doppeltso viel, wie G.711
- hohe und konstante Datenrate notwendig
- etwas aufwendiger zu de- und encodieren

G.723.1 ist der Audio-Codec, der in Netmeeting standardmäßig Anwendung findet. Er verwendet lineare prädiktive analysis-by-synthesis Kodierung.

Datenrate: 5.3 und 6.3 kbps

+ erkennt Sprachpausen und sendet dann nichts bzw. weniger

G.729 verwendet „Conjugate Structure Algebraic-Code-Excited Linear-Prediction (CS-ACELP)“. Der Codec ist besonders auf kurze Latenzzeiten optimiert.

Datenrate: 8 kbps

+ Latenzzeit < 15ms

MPEG-4-Audio ist ein sehr breites Format und vereint viele Vorteile von bisherigen Kompressionsformaten, wie mp3 oder G.???

Es ist z.B. auch synthetisches Audio – Spracherkennung -> übertrage „hallo“ -> synthetisiere das Wort wieder in Sprache – möglich.

In mpeg4-v.1 gibt sind folgende Profile vorgesehen:

- **The Speech Profile** provides HVXC, which is a very-low bit-rate parametric speech coder, a CELP narrowband/wideband speech coder, and a Text-To-Speech interface.
- **The Synthesis Profile** provides score driven synthesis using SAOL and wavetables and a Text-to-Speech Interface to generate sound and speech at very low bitrates.
- **The Scalable Profile**, a superset of the Speech Profile, is suitable for scalable coding of speech and music for networks, such as Internet and Narrow band Audio Digital Broadcasting (NADIB). The bitrates range from 6 kbit/s and 24 kbit/s, with bandwidths between 3.5 and 9 kHz.
- **The Main Profile** is a rich superset of all the other Profiles, containing tools for natural and synthetic Audio.

mit Version 2 kommen hinzu:

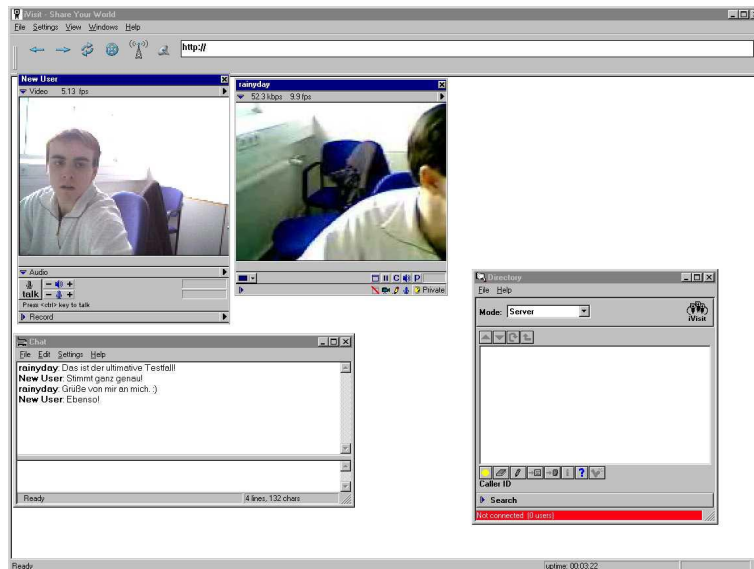
- **The High Quality Audio Profile** contains the CELP speech coder and the Low Complexity AAC coder including Long Term Prediction. Scalable coding can be performed by the AAC Scalable object type. Optionally, the new error resilient (ER) bitstream syntax may be used.
- **The Low Delay Audio Profile** contains the HVXC and CELP speech coders (optionally using the ER bitstream syntax), the low-delay AAC coder and the Text-to-Speech interface TTSI.
- **The Natural Audio Profile** contains all natural audio coding tools available in MPEG-4, but not the synthetic ones.
- **The Mobile Audio Internetworking Profile (MAUI)** contains the low-delay and scalable AAC object types including TwinVQ and BSAC. This profile is intended to extend communication applications using non-MPEG speech coding algorithms with high quality audio coding capabilities.

<http://mpeg.telecomitalia.com/standards/mpeg-4/mpeg-4.htm#3.4>

Datenrate: 6 kbps bis mehrere GBps

- + Skalierbarkeit (Dabei ist es sogar möglich, dass das Netz, bestimmte Teile des Datenstromes droppt; sonst natürlich auch VBR u.ä.)
- + beliebige Qualität
- + ...
- wesentlich höherer Aufwand beim de- und encodieren

theLOOK:



Praxis

i-visit

theFEEL:

- Video und Tonqualität ausreichend aber nicht besonders gut (Ton sehr dumpf, Bild schlecht aufgelöst [in Stufen von 1 bis 10 beim Sender einstellbar], h.323 halt)
- Übertragung funktioniert gut
- Bedienung nicht besonders übersichtlich (z.B. Chat weg, wie bekomme ich den wieder?)
Wenn man das Programm mal benutzt hat geht's wohl
- keine Zusatzfeatures (Whiteboard), nur Bild, Ton und Chat

aUFzeichnung:

Man kann selbst bestimmen, ob man sein Videobild von einem Gesprächsteilnehmer aufzeichnen lassen will. Dazu muss man seine Bereitschaft erklären „Others may record me“ im Menu Record und dann der Aufzeichnung noch per Dialog im konkreten Fall zustimmen.

Aufgezeichnet wird alles, also Bild, Ton und Chat. Abgespielt wird es dann natürlich auch und zwar kommt der Chat auch zu rechten Zeit ins Chatfenster, falls man das geöffnet hat, denn von sich aus öffnet er es nicht.

auDioCoDecs:

iVision bietet Lucent SX8300 (8 kbps) und ADPCM (32 kbps).

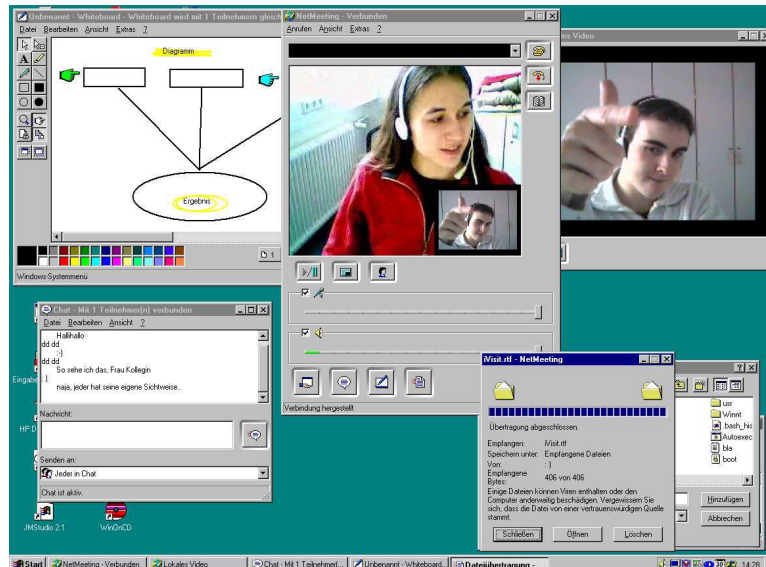
Zu Anfang des Tests war der Lucent Codec deutlich schlechter, sowohl bei der Qualität (ständig übersteuert, was sich als plop übertrug) als auch mit viel mehr Latenzzeit (subjektiv fast doppelt).

Das lag aber wohl eher daran, dass das Mikrofon schlecht eingestellt war, denn später waren beide ungefähr gleich. Der ADPCM hat bei Sprache nur ein bisschen weniger Rauschen, aber dafür ja auch die vierfache Datenrate.

Die Latenzzeit des Tons hing bei unserem Test sehr stark von der eingestellten Videoqualität ab. Bei Stufe 10 (bestes Video) hing der Ton merklich stärker hinterher.

netmeeting

theLOOK:



theFEEL:

- Übertragungsgeschwindigkeit der Bilddaten in NetMeeting wesentlich besser
- Übertragungsgeschwindigkeit bei Audiodaten dafür verzögerter
- Bildauflösung: beide gleich schlecht (h.323)

teXtChat:

Öffnet ein Teilnehmer einen Textchat, so öffnet sich der Chat auch beim anderen.

Wird er dann wieder geschlossen, so öffnet er sich beim zweiten Mal nicht mehr automatisch auch beim Gegenüber.

Wird das Fenster erneut geöffnet, so ist der Text von vorher nur dann erhalten, wenn er gespeichert wurde.

whiteBoard:

Ist die Ansicht synchronisiert, sehen alle Konferenzteilnehmer immer die gleiche Seite.

Ist die Ansicht nicht synchron wirken sich die Änderungen zwar auch aus, aber es kann sein, dass ich gerade auf Seite 1 bin, während der Rest Seite 5 editiert.

Datei versenden:

Im „Datei versenden“-Fenster kann jeder Teilnehmer den Ordner für die gemeinsamen Dateien angeben.

Versendet ein Teilnehmer eine Datei, so landet sie in diesem Ordner.

Der Empfang kann nicht verhindert werden. Nach Erhalt kommt aber ein Dialog, in dem man die Datei sofort löschen kann (Sinn?).

programmFreigabe:

Gibt man ein beliebiges Programm (in unserem Fall WordPad) frei, so erscheint auf den anderen Rechnern ein Fenster, in dem das Programm läuft. Konkret wird die Oberfläche des Programms gecaptured und übertragen.

Dabei haben wir festgestellt, dass der Algorithmus zum Übertragen des Fensters nicht so toll ist, weil er wenn man das Fenster nach dem Verkleinern wieder groß macht sehr lange braucht, bis man wieder den gesamten (neuen) Text sieht.

In der Menüleiste des RemoteFensters kann die Steuerung für WordPad angefordert werden.

Das führt auf dem Hostrechner dazu, dass die Maussteuerung komplett weg. Für alle Programme... das ist natürlich doof, aber auch logisch, weil der Hostcomputer ja nur ferngesteuert wird und der eben nur eine Maus hat...

Ganz verloren ist man als Host aber nicht, mit Escape kann man wieder Herr über den Mauszeiger werden...