

Seminarvortrag
„Voice over IP“

Sylvia Kreuzer
Bernd Maurer

2. Juli 2004



Inhaltsverzeichnis

1	Einführung	2
1.1	Geschichtlicher Überblick	2
1.2	Was ist VoIP?	2
1.3	Wie funktioniert es?	2
1.4	Wo liegt das Problem?	2
2	Die Funktionsweise von VoIP	3
2.1	A-D-Wandlung und Kompression	3
2.1.1	Silence Compression	4
2.1.2	G.711 Codec mit Samples von 60 ms je Paket	5
2.1.3	G.711 mit Samples von 30 ms je Paket	6
2.2	RTP, UDP, IP	7
2.3	H.323	7
2.3.1	RAS	9
2.3.2	Die Kommunikation im Detail	9
3	Qualität	11
3.1	Qualitätsanforderungen bzgl. Paketverlust/Delay gemäß ITU G.114	11
3.1.1	Paketverlust	11
3.1.2	Delay	12
3.2	Wie misst man die Sprachqualität?	12
4	Systemvoraussetzungen für eine PC-Implementierung	13
4.1	Hardware	13
4.2	Software	13
5	Sicherheit	13
5.1	Abhörbarkeit	13
5.2	Verschlüsselung	14
6	Realisierung am Beispiel der Universität Ulm	14
6.1	Das Telefonnetz der Universität Ulm	14
6.2	Das IP-Telefonnetz der Universität Ulm	15
7	Möglichkeiten für Privatkunden	16
7.1	Angebot der T-Com	16
7.2	Angebot von Freenet	17
8	VoIP – Ein kritischer Blick	17
9	Ausblick	18

Zusammenfassung

Thema des Seminars ist die Verwendung des IP-Protokolls zum Telefonieren. Der wesentliche Unterschied zur klassischen Telefonie besteht darin, dass nicht für jedes Telefonat eine Leitung vorhanden ist (Leitungsvermittlung), sondern das Gespräch aufgeteilt in einzelne IP-Pakete über Datenleitungen übertragen wird. Dadurch entstehen neue Probleme wie Paketverlust, unterschiedliche Paketlaufzeit, möglichst geringer Bandbreitenbedarf etc. Wirtschaftlich wird VoIP dadurch interessant, dass es genügt, ein Datennetzwerk zu warten; das bisherige Telefonnetz kann entfallen.

1 Einführung

Die folgenden Informationen sind entnommen aus [1].

1.1 Geschichtlicher Überblick

Vor mehr als dreißig Jahren gab es kein Internet; es gab das Telefon (das erste konstruierte der deutsche Physiker Johann Philipp Reis 1861), und um Daten auszutauschen musste sehr viel Geld ausgegeben werden, denn es war erforderlich, dass sie von einem Menschen auf einem Datenträger von *A* nach *B* gebracht wurden. Doch vor einigen Jahren begannen massive Veränderungen: Der PC wurde ein Massengut, neue Kommunikationsmittel wie Handys eroberten sich Märkte, die es zuvor gar nicht gab, und schließlich entstand das Massenmedium Internet mit Killerapplikationen wie e-Mail, Chatting und dem Surfen.

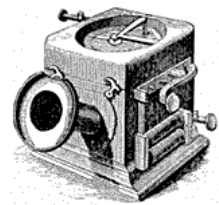


Abbildung 1:
Eines der ersten
Telefone

Als Ergebnis sind wir heute Zeugen einer echten Revolution in der Kommunikation. PCs werden von nahezu jedem privat wie auch geschäftlich genutzt, um mit anderen Menschen in Kontakt zu bleiben oder gar neue Kontakte aufzubauen. Im Moment noch beschränken sich dabei die meisten auf triviale, da nicht-zeitkritische Daten: Texte, Bilder, Musik,... Doch immer mehr Menschen wollen auch über das Internet telefonieren.

1.2 Was ist VoIP?

VoIP ist die Abkürzung für *Voice over Internet Protocol*, und wie dieser Name sagt, ist das Ziel, die menschliche Stimme beim Telefonieren mit Hilfe eines Datennetzes (häufig das Internet), aufgeteilt in IP-Pakete, zum Gesprächspartner zu bringen.

1.3 Wie funktioniert es?

Da die Stimme in Form von (analogen) Schwingungen aufgenommen wird, muss sie zuerst digitalisiert werden – dies ist ein aus der Vergangenheit bekannter Prozess, für den es mehrere Möglichkeiten gibt. Diese digitalisierte Sprache kann dann auf die Reise gehen.

1.4 Wo liegt das Problem?

VoIP war bis vor einigen Monaten noch nicht weit verbreitet; warum?

In der Entstehung des Internets ging es darum, *dass* die Daten ankommen, nicht *wann*. Beim Telefonieren ist aber eine zwingende Voraussetzung, dass alles in *Echtzeit* ankommt;

dazu braucht man hohe Leitungskapazitäten. Dieser entscheidende Unterschied in den Anforderungen und Zielsetzungen ist der Hauptgrund, dass erst jetzt mit der Einführung von VoIP begonnen wird. Zum genaueren Verständnis müssen wir etwas weiter ausholen:

2 Die Funktionsweise von VoIP

Die folgenden Informationen sind weiterhin entnommen aus [1], hinzu kommen Gespräche mit [7].

Zum Aufbau einer VoIP-Verbindung benötigt man:

1. Einen A/D-Wandler, um die Stimme zu digitalisieren
2. Eventuell eine Kompression, um die Datenmenge zu reduzieren
3. Nun werden die einzelnen „Sprachschnipsel“ mit Hilfe eines Echtzeitprotokolls in Datenpakete gepackt (meist *RTP over UDP over IP*)
4. Um den Gesprächspartner erreichen zu können bedarf es eines Protokolls (*ITU-T H.323*)
5. Bei diesem wird nun die Sprache wieder extrahiert, dekomprimiert und letztendlich nach Analogwandlung ausgegeben

All dies sollte natürlich möglichst in Echtzeit geschehen und es sollten außerdem keine Zeitabschnitte verloren gehen.

2.1 A-D-Wandlung und Kompression

Zur Digitalisierung der analogen Audiosignale werden in definierten Zeitabständen die notwendigen Daten gemessen und als Zahlenwerte gespeichert. Die Frequenz, mit der dies geschieht, ist die *Sampling-Rate*, die einzelnen Datensätze nennt man *Samples*. Schmalbandcodecs fassen häufig mehrere Samples zu einem Frame zusammen. Die Komprimierungsalgorithmen der Codecs benötigen häufig zusätzlich Informationen aus vorangegangenen und nachfolgenden Samples, weshalb diese solange zwischengespeichert werden müssen, was eine Verzögerung bewirkt.

Weit verbreitete Codecs sind:

G.711: PCM- (Pulse Code Modulation) basiert benötigt dieser 1965 von der ITU akzeptierte Codec 64 kbit/s, also genau die Bandbreite, die eine ISDN-Verbindung bereit stellt. Er wird bei ISDN-Telefonaten benutzt und zeichnet sich durch eine sehr geringe Prozessorleistung aus.

G.726: Dieser Codec aus dem Jahre 1990 benötigt durch die Benutzung von ADPCM nur 32 kbit/s; erreicht wird dies, indem jeweils nur die Differenzinformation zwischen zwei aufeinander folgenden Sprachpaketen übertragen wird. Die Anforderungen an den Prozessor sind allerdings deutlich höher.

weitere: Mit weiterer Verbesserung der Kompression (auf Kosten noch höherer Rechenleistung) erreicht man Datenraten von mindestens 2,5 kbit/s

Allerdings kommt zu den reinen Sprachdaten noch ein Overhead von 40 Byte durch die drei Header von RTP, UDP und IP hinzu. Dadurch wird die Datenmenge je nach Paketgröße mehr oder weniger stark aufgebläht. Jedes dieser Pakete enthält nur einen mehrere Samples umfassenden bestimmten Zeitabschnitt des Gesprächs – meist 20 ms, da hier ein gutes Verhältnis zwischen Sprachdaten und Overhead vorliegt (siehe auch die beiden folgenden Abschnitte) und außerdem der Verlust einzelner Pakete noch nicht zu große *Löcher* in das Telefonat reißt. Kommt ein Paket nicht oder zu spät an (eine sinnvolle Wartezeit sind 60 ms, mehr als 150 ms Verzögerung empfindet man als störend [9]), so muss der Empfänger das Paket als verloren ansehen. Er kann dann entweder das letzte Paket wiederholen oder eine kurze Lücke lassen. Damit ist ein Paketverlust von bis zu 5% noch nahezu unmerklich auszugleichen.

Codec		Bandbreite <i>kbit/s</i>	Framedauer <i>ms</i>	IP Bandbreite <i>kbit/s</i>
G.711	PCM	64	0.125	80
G.723.1	ACELP	5.6	30	16.27
G.723.1	ACELP	6.4	30	17.07
G.726	ADPCM	32	0.125	48
G.728	LD-CELP	16	0.625	32
G.729(A)	CS-ACELP	8	10	24

Tabelle 1: Datenraten (Entnommen aus [8])

Achtung: In der Tabelle sind zwei Algorithmen aufgeführt, die nicht einen Overhead von 16 kbit/s erzeugen: G.723.1 ACELP mit 5.6 kbit/s und 6.4 kbit/s. Da bei ihnen ein Frame 30 ms statt 20 ms umfasst, werden nur 33 Frames pro Sekunde gesendet.

2.1.1 Silence Compression

Die meiste Zeit während eines Telefonats besteht aus Pausen. Es wäre unnötig, in diesen Zeitabschnitten mit der vollen Datenrate zu arbeiten. Daher enthalten Codecs wie der G.723.1 oder der G.729 eine Silence Compression. Sie besteht im Wesentlichen aus den folgenden drei Komponenten:

VAD

Die Aufgabe von VAD (voice activity detector) ist es, festzustellen, wann ein Gesprächsteilnehmer spricht und wann er still ist. Hierzu muss der Algorithmus schnell reagieren, um zu verhindern, dass nach einer solchen Ruhe die erste Silbe verloren geht. Für die sichere Unterscheidung zwischen Gespräch und Stille benötigt der Codec einen Zwischenspeicher, der einen zusätzlichen Delay verursacht.

DTX

DTX (discontinuous transmission) ermöglicht es einem Codec theoretisch, wenn VAD Stille erkannt hat, die Verbindung zu unterbrechen. Da eine solche Unterbrechung aber absolute Stille beim Gesprächspartner bedeuten würde, wird die Verbindung nicht wirklich komplett unterbrochen, sondern es wird ein kleiner Satz an Daten übertragen, der die Erzeugung von Hintergrundgeräuschen beim Empfänger ermöglicht.

CNG

Der comfort noise generator (CNG) setzt genau an dieser Stelle an: er ist in der Lage, selbstständig Hintergrundgeräusche zu erzeugen; dazu benutzt er die bei der vorherigen Gesprächsphase vorhandenen Hintergrundgeräusche (alles aus [17] entnommen).

Wir wollen nun die Auswirkungen verschiedener Parameter auf die Gesprächsqualität untersuchen:

2.1.2 G.711 Codec mit Samples von 60 ms je Paket

G.711 ist einer der ältesten digitalen Audio-Codecs. 1965 wurde er von der ITU (International Telecommunications Union) offiziell als Protokoll angenommen und ist daher noch weit verbreitet. Der Codec G.711 nimmt keine Kompression an den Samples vor und hat somit eine zu ISDN identische Netto-Datenrate von 64 kbit/s, weshalb ISDN-Telefonate darüber abgewickelt werden; brutto benötigt er wegen den zusätzlichen Headern 80 kbit/s (vgl. Tabelle 1)[17].

Paketverlust

Bei jedem Paketverlust gehen 60 ms an Sprachdaten verloren. Dies ist ein so langer Zeitraum, dass jeder einzelne Paketverlust, der nicht in eine Sprachpause fällt, bemerkt wird und zu einer deutlichen Qualitätseinbuße führt. Die Sprachqualität kann noch als gut bezeichnet werden, wenn höchstens ein bis zwei Prozent der Pakete verloren gehen. Liegt der Paketverlust über fünf Prozent, ist die Qualität so schlecht, dass eine Kommunikation unter normalen Umständen nicht mehr sinnvoll ist. In Abbildung 2 kann man den Abfall der Sprachqualität gut erkennen. Durch Paketverluste können einzelne Wortteile bis hin zu ganzen Worten verloren gehen, bei entsprechend hohen Verlusten ist eine Wort- oder sogar Satzverständlichkeit nicht mehr gewährleistet. Der Hörende bekommt dies durch abgehackte Sprache zu spüren (zur Messung der Sprachqualität siehe 3.2).

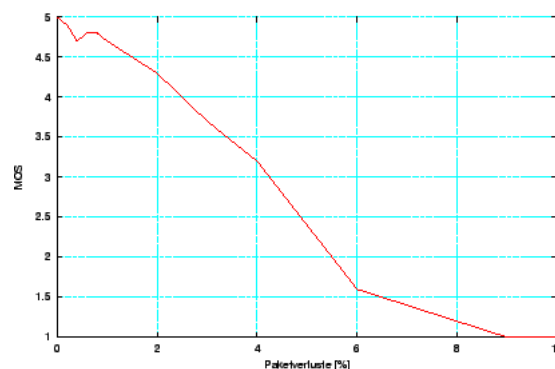


Abbildung 2: Paketverlust bei G.711-Codec mit 60 ms Samples/Paket

Jitter

Jitter ist der Zeitunterschied zwischen den Paketankunftszeiten der unterschiedlichen Pakete. Dieser kann verursacht werden durch Netzüberlastung oder dadurch, dass jedes Paket einen anderen Weg geht. Mit Hilfe des Jitterpuffers versucht man den Zeitunterschied, der

zwischen den einzelnen Paketen sehr variieren kann, zu glätten:

Ein Jitterpuffer ist ein Speicherbereich, in dem die Samples gesammelt, gelagert und von dem aus sie in gleichbleibenden Abständen zum Sprachprozessor geschickt werden. Der Jitterpuffer, der am Empfangsende der Sprachverbindung installiert ist, versucht die Sprache um die mittlere Laufzeit verzögert wiederzugeben, so dass der Hörende eine klare Verbindung mit geringem Klirrfaktor empfindet [13].

Durch den Jitterpuffer wird die Auswirkung von geringem Jitter auf die Sprachqualität gemildert oder sogar ganz vermieden. Erst wenn der Jitter in den Bereich der Puffergröße kommt, ist er bemerkbar. Dann ist die Verschlechterung der Qualität sehr stark und ändert sich sehr sprunghaft. Setzt man die Größe des Jitterpuffers auf die Größe eines Paketes (also in diesem Fall 60 ms), werden durch ihn Jitter bis knapp 60 ms nicht hörbar. Die Grenze bei den Siemens-Telefonen liegt bei 59 ms. Wenn der Jitter diesen Wert erreicht oder sogar überschreitet fällt die Sprachqualität auf ein inakzeptables Maß ab (siehe Abbildung 3). Die Auswirkung des Jitters auf die Sprache läßt sich am ehesten mit abgehacktem, kurzzeitigem Echo oder Stottern beschreiben. Wenn die Schwankung sehr groß wird (ab ca. 150 ms) und gleichzeitig ein sehr hoher Anteil der Pakete diese Schwankung erfährt (ab etwa 20% aller Pakete), entsteht ein sehr deutlicher Echoeffekt.

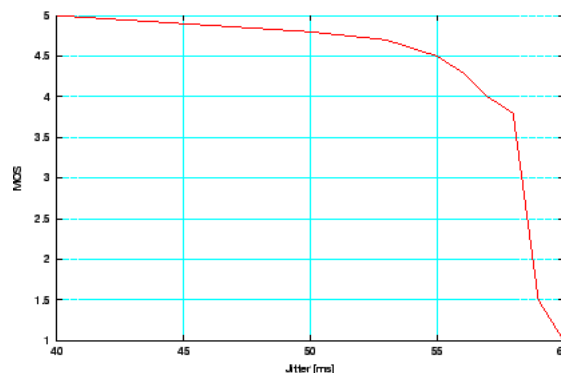


Abbildung 3: Jitter bei G.711-Codec mit 60 ms Samples/Paket

2.1.3 G.711 mit Samples von 30 ms je Paket

Halbierte Samples

Was passiert, wenn man die Samples halbiert? Wenn hier ein Paket verloren geht, haben wir „nur“ einen Verlust von 30 ms an Samples. Bei gleichem Prozentsatz an Paketverlusten verbessert sich die Qualität der Übertragung subjektiv kaum, denn das Fehlen von Wortteilen fällt ebenso deutlich auf.

Die Grenze zwischen guter und mittelmäßiger Qualität liegt etwa bei ein bis zwei Prozent, zwischen mittelmäßiger und dürrtiger Qualität bei vier bis fünf Prozent (siehe Abbildung 4).

Bei einer Samplegröße von 30 ms liegt die Größe des Jitterpuffers auch bei 30 ms. Durch die halbierten Samples ist der Jitterpuffer in der Lage, nur eine halb so große Schwankung bei der Paketankunft auszugleichen. Die Grenze, ab der man die ersten vereinzelt Schwankungen hört, liegt bei 26 bis 27 ms, ab 28 bis 29 ms ist die Qualitätseinbuße durch den Jitter für eine sinnvolle Kommunikation zu hoch (siehe Abb. 5). Der Charakter ist derselbe wie bei dem G.711 Codec mit 480 Samples (60 ms) je Paket (Quelle: [12]).

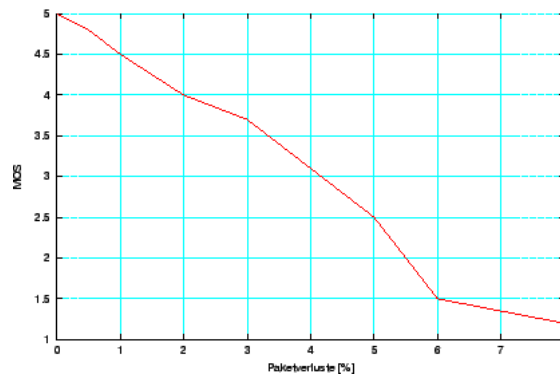


Abbildung 4: Paketverlust bei G.711-Codec mit 30 ms Samples/Paket

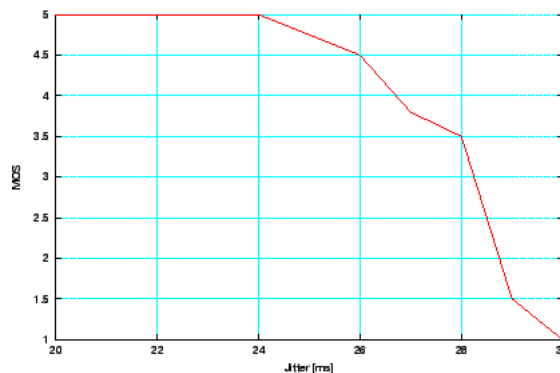


Abbildung 5: Jitter bei G.711-Codec mit 30 ms Samples/Paket

2.2 RTP, UDP, IP

Die komprimierten Daten müssen nun verschickt werden. Dazu benutzt man drei Protokolle:

1. Real-Time-Protocol (RTP)
2. User-Datagram-Protocol (UDP)
3. Internet Protocol (IP)

Die Sprachdaten werden zunächst mit einem Header gemäß eines Real-Time-Protocols versehen, damit der Empfänger weiß, an welche Stelle das Gesprächsstück gehört und ob er es gegebenenfalls verwerfen muss, da es zu spät kam. Dieses Datenpaket wird nun mit einem UDP-Header versehen und über das IP-Protokoll versendet.

2.3 H.323

Für VoIP gibt es ein wichtiges Protokoll: H.323. Es erlaubt vielen verschiedenen „Elementen“, miteinander zu kommunizieren:

Terminals: Telefone, Videotelefone, Voicemail Systems, Soft Phones (z.B. Netmeeting)

Clients die VoIP-Verbindungen initialisieren

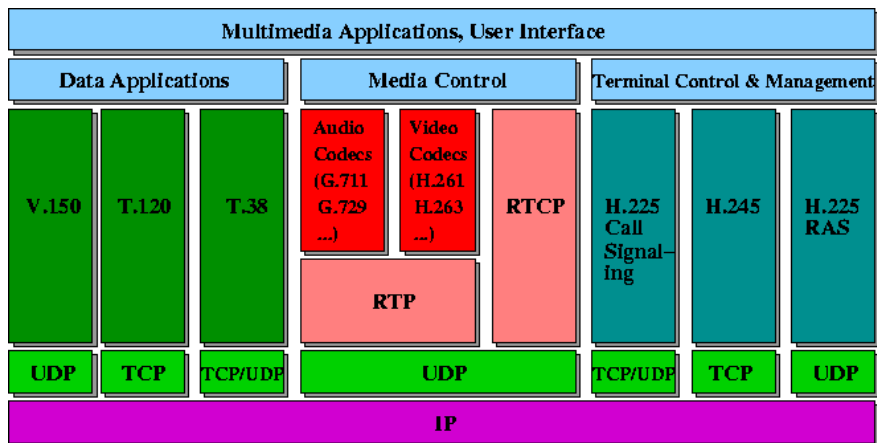


Abbildung 6: H.323-Stack

Gatekeepers: (optional): übersetzt Namen in IP-Adressen *et vice versa*, kontrolliert Zugangsberechtigungen und erlaubt, die Bandbreitennutzung zu beeinflussen; außerdem kann er „Punkt-zu-Punkt“-Verbindungen (d.h. direkt von Client zu Client) oder Gespräche nur „über sich“ erlauben, um Funktionen wie *follow-me*, *find-me*, *forward on busy* etc. zu ermöglichen

Gateways für den Übergang zum Telefonnetz (*PSTN*) (Public Service Telephone Network), zu anderen H.323-Netzen oder zu H.320-Systemen

Multipoint Control Units (MCUs), um Konferenzschaltungen zu ermöglichen; Dabei gibt es einen *Multipoint Controller* (MC), der die Steuerung des Gespräches übernimmt

Proxie Server

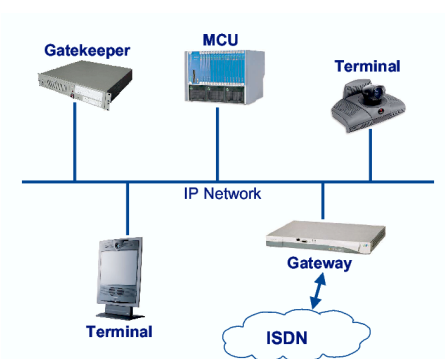


Abbildung 7: H.323-Komponenten

H.323 regelt des weiteren Video- und Datentransfers. Das Protokoll wurde von der *International Telecommunications Union - Telecommunications Sector* (ITU-T) [14] entwickelt und verabschiedet.

2.3.1 RAS

RAS steht für *Registration, Admission* und *Status*. Es dient in der Kommunikation zwischen Gatekeeper und Endpunkt dazu

- dass der Gatekeeper einen Endpunkt „managen“ kann
- dass der Endpunkt beim Gatekeeper um Erlaubnis für ein Gespräch bitten kann
- dass der Gatekeeper die Adressauflösung für den Endpunkt vornehmen kann

RAS muss benutzt werden, wenn es im Netz einen Gatekeeper gibt.

2.3.2 Die Kommunikation im Detail

Wird ein Endpunkt an das Netz angeschlossen, so sendet er zunächst eine *Gatekeeper Request*, um den zuständigen Gatekeeper zu finden. Dabei ist es auch möglich, dass mehrere Gatekeeper antworten. Ist der Gatekeeper nicht bereit, den Endpunkt zu akzeptieren, so antwortet er mit einer *Gatekeeper Reject*, die auch den Grund für die Absage angibt. Ist er dagegen einverstanden, so sendet er eine *Gatekeeper Confirm* – dabei erhält der Endpunkt für die spätere Kommunikation wichtige Daten. Der Endpunkt sendet noch eine *Gatekeeper Registration* und je nachdem, ob er darauf eine positive oder negative Antwort erhält, ist er anschließend in der Lage, Services des Gatekeepers zu erhalten oder nicht; er ist aber auf jeden Fall grundsätzlich im Netz erreichbar. Desweiteren wird nun ausgehandelt, welche Nummern bzw. Namen der Endpunkt erreichen darf und wie lange er im Netz akzeptiert wird (*time to live*).

Möchte der Endpunkt nun ein Gespräch aufbauen, so sendet er eine *Admission Request (ARQ)* an den Gatekeeper, der sie akzeptieren (*ACF*) oder ablehnen (*ARJ*) kann. Im ersten Fall wird ein (eindeutiger) *call reference value (CRV)* für das Gespräch zwischen Endpunkt und Gatekeeper ausgehandelt; anschließend wird eine *CallID* erzeugt. Diese gilt nur für dieses Gespräch.

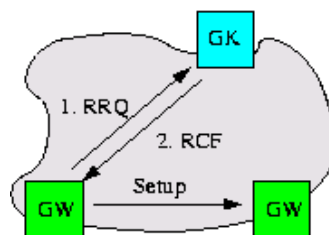


Abbildung 8: Admission Request

Um den Gesprächspartner zu finden, schickt der eigene Gatekeeper oder der Endpunkt selbst an den Gatekeeper der Gegenseite eine *Location Request (LRQ)*, um ihn zu bitten, die Alias-Adresse aufzulösen, z.B. um aus der Telefonnummer eine IP-Adresse zu machen.

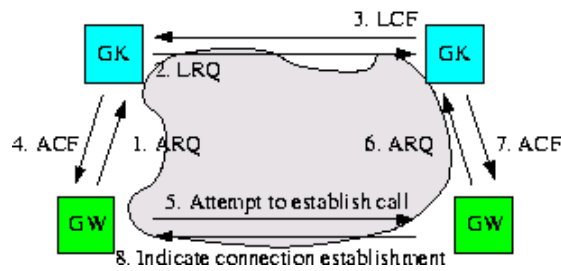


Abbildung 9: Location Request

Auf diesen ersten Teil des Gesprächsaufbaus kann nun das Gesuch des Endpunktes folgen, mehr (oder auch weniger) Bandbreite als beim letzten Gespräch zu bekommen (*Bandwidth Request (BRQ)*).

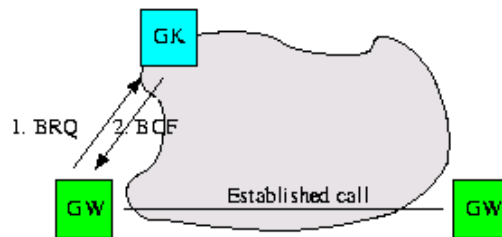


Abbildung 10: Bandwidth Request

Jetzt kann telefoniert werden!

Soll das Gespräch beendet werden, geschieht folgendes: der Endpunkt sendet eine *Disengage Request (DRQ)* an den Gatekeeper. Auch der Gatekeeper kann eine solche Aufforderung an den Endpunkt senden, um den Abbruch eines Gesprächs zu erzwingen.

Darüber hinaus gibt es *Information Requests (IRQ)*, um weitere Daten zwischen Endpunkt und Gatekeeper auszutauschen, *Request In Progress (RIP)*-Messages um anzuzeigen, dass die Bearbeitungszeit momentan höher als gewöhnlich liegt, *Resource Availability (RAI)*-Messages, um vor drohender Ressourcenknappheit zu warnen oder sie anzuzeigen und einige andere.

Der Timeout auf die obengenannten Messages liegt zwischen drei und fünf Sekunden. Um den oben beschriebenen, doch recht aufwändigen Gesprächsaufbau zu beschleunigen, gibt es die Möglichkeit des *Fast Connect*, durch den die Prozedur unter Umständen auf zwei Messages verkürzt werden kann.

Wie in Abschnitt 6.2 noch an Beispielen erläutert wird, gibt es proprietäre Erweiterungen von H.323.

Für die Identifikation eines einzelnen Endpunkts gibt es in H.323 die folgenden Möglichkeiten:

- dialedDigits
- h323-ID
- url-ID
- transportID
- email-ID
- partyNumber
- mobileUIM

Davon wird dialedDigits am meisten eingesetzt, um Gespräche aufzubauen. h323-IDs sind dagegen nur zur Kommunikation zwischen einem Endpunkt und seinem Gatekeeper anzuwenden. url-ID und email-ID sind gerade erst im Kommen (Quelle: [15]).

Da das Protokoll insbesondere von Firmen wie Microsoft, Siemens und Cisco benutzt wird, fanden sich OpenSource-Anhänger, die nun das Protokoll *Open H.323* entwickeln, das garantieren soll, dass jedermann kostengünstigen Zugang zu den auf H.323 basierenden Produkten bekommen kann. Für nähere Informationen hierzu siehe [2].

3 Qualität

TCP/IP kennt keine Echtzeit. Das einzige Mittel, um die Übermittlungsdauer der einzelnen Pakete zu reduzieren, besteht darin, im TOS-Feld der IP-Pakete die Dringlichkeit auf den Minimalwert zu setzen (das bedeutet höchste Dringlichkeit). Allerdings müssen alle Knotenpunkte, die die Pakete passieren, diese Information auch nutzen und dementsprechend die ankommenden Pakete weiter vermitteln. Da man keinen Einfluss darauf hat, welchen Weg die Pakete nehmen, ist dies ein großes Problem.

Wie bereits in Kapitel 2.1 beschrieben, gibt es verschiedene Möglichkeiten, auf den Verlust von Paketen zu reagieren. Wichtig für die Qualität ist aber auch die Laufzeit der einzelnen Pakete: Kommen die Pakete sehr unterschiedlich an, so ist dies ein Problem, vor allem wenn die Laufzeitunterschiede um die Schwelle der Jitterpuffergröße herum liegen (siehe Kapitel 2.1.2). Unterscheidet sich die Laufzeit einzelner Pakete dagegen nur gering, so können die Pakete beim Empfänger wieder gut in die Originalreihenfolge gebracht werden und der Benutzer merkt nichts – es sei denn, die durchschnittliche Laufzeit der Pakete ist höher als die oben erwähnten 150 ms (siehe Kapitel 2.1.2).

3.1 Qualitätsanforderungen bzgl. Paketverlust/Delay gemäß ITU G.114

3.1.1 Paketverlust

Nach der Empfehlung G.114 der ITU wird die Qualität der Sprache nicht von der Menge des Paketverlustes beeinflusst, solange dieser unter 5% bleibt; signifikant ist vielmehr der Delay der einzelnen Pakete.

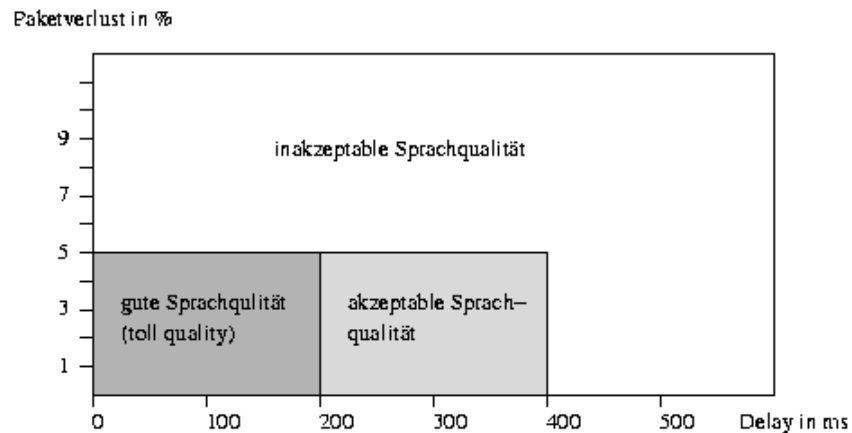


Abbildung 11: Qualitätsanforderungen

3.1.2 Delay

Aus Abbildung 11 wird ersichtlich, dass als höchstmöglicher Delay-Wert 200 ms angegeben werden, um noch eine gute Sprachqualität zu erreichen. Diese 200 ms wurden als einfacher Wert gemessen, d.h. es sind 200 ms von Gesprächspartner A bis zu Gesprächspartner B erlaubt, um noch eine gute Sprachqualität zu erreichen. Liegt die Verzögerung jedoch zwischen 200 und 400 ms, so ist die Sprachqualität nur noch akzeptabel. Kommt es zu einem noch größeren Delay, so ist der Bereich der inakzeptablen Sprachqualität erreicht [17].

3.2 Wie misst man die Sprachqualität?

Um die Sprachqualität in Werten auszudrücken, wurde die Einheit *MOS* (mean opinion score) eingeführt. Um den MOS-Wert zu erreichen, werden verschiedenen Testpersonen zuvor aufgenommene Hörproben von Sprache, Musik und Sprache mit Hintergrundgeräuschen vorgespielt, worauf die Testpersonen das Gehörte mit Noten von 1 (= schlecht) bis 5 (= gut) bewerten müssen. Anschließend werden statistische Mittel darüber gebildet und somit der MOS-Wert ermittelt (siehe Tabelle 2).

Gesprächstyp	MOS-Wert
ISDN	4.2
IP	4.0
PSTN (toll-quality)	3.8
GSM (Mobilfunk)	3.0
military-quality	2.5

Tabelle 2: Sprachqualität

4 Systemvoraussetzungen für eine PC-Implementierung

Die folgenden Angaben sind [1] entnommen.

4.1 Hardware

Um eine Vorstellung zu geben, was man der Hardware abverlangt, folgende Mindestvoraussetzungen:

- PC 386 oder höher, 32 MB RAM
- Duplex-fähige Soundkarte
- Netzwerkverbindung mit mind. 64 kbit/s (ISDN)

Die Anforderungen sind also nahezu minimal.

4.2 Software

Es gibt proprietäre wie auch freie Software, jeweils für Windows (z.B. Netmeeting) wie auch für UNIX / Linux (z.B. GnomeMeeting).

5 Sicherheit

„FBI dringt auf bessere Lauscmöglichkeiten für Internettelefonie“^[3]

„[...] Das FBI moniert schon seit längerem, dass die Überwachungsmöglichkeiten für „VoIP“ nicht ausreichen. Nach dem CALEA-Gesetz (Communications Assistance for Law Enforcement Act) vom Jahr 1994 müssen zwar Telekom-Firmen, die die digitalen Netzwerke einrichten, die Möglichkeiten zur Überwachung der Telefongespräche einbauen, für VoIP gibt es in den USA entsprechende Abhörstandards noch nicht. Nach CALEA sind Informationsdienste grundsätzlich von den Vorschriften ausgenommen. Das FBI will das CALEA-Gesetz auch auf das Internet erweitern und hat auch schon längst genauere technische Vorstellungen. Auch in der EU ist die Einführung von Standards für das Abhören von VoIP geplant. VoIP-Gespräche können nur gezielt direkt über den Zugang zu den Servern des jeweiligen Internetproviders abgehört werden. Das juristische Problem ist, wie sicher gestellt werden kann, dass die Sicherheitsbehörden nur die richterlich genehmigte Überwachung von bestimmten VoIP-Teilnehmern durchführen, aber nicht damit auch einen unkontrollierbaren Zugang zu Emails oder der gesamten Internetnutzung aller Kunden erhalten. [...]“

5.1 Abhörbarkeit

Bisher sind alle Telefonate, die über das öffentliche Telefonnetz gehen, theoretisch abhörbar; zwar auf Grund des großen Aufwands nicht von Privatpersonen, aber man kann davon ausgehen, dass Geheimdienste seit Jahrzehnten jede Kommunikation per Telefon oder Fax abhören. Dazu braucht man nur Zugang zu der Leitung – indem man die eigene Leitung damit verbindet oder die elektromagnetische Strahlung abfängt und daraus das Signal ermittelt.

Bei VoIP werden die Gesprächsdaten über das Internet übertragen; auch hier ist es möglich, das Gespräch – wie auch jedwede Kommunikation per eMail etc. – abzuhören. Weitere Überwachungsmöglichkeiten bieten sich den Providern, über deren Leitungen die Daten übertragen werden. Unternehmen haben dagegen das Problem, dass das Überwachen durch Anzapfen fremder Leitungen illegal ist und sich daher, wenn es publik wird, stark rufschädigend auswirken kann.

5.2 Verschlüsselung

Eine Verschlüsselung der Telefonate wäre technisch zwar machbar (wenn man in die Telefone Prozessoren mit höherer Leistung einbaute), aber es fehlt an der Möglichkeit, die Authentizität des Gesprächspartners zu überprüfen. Wie bei *PGP* Keyserver-basiert mit Schlüsseln zu arbeiten ist bei Millionen von Telefonen nicht machbar. Des weiteren besteht kein Interesse des Staates, eine Zentralstelle für Zertifikate zu schaffen. Allerdings bleibt größeren Unternehmen die Möglichkeit, mit Partnern, mit denen sie häufig in Kontakt treten, auf dem Gateway basierend eine Verschlüsselung aufzubauen; dies ist sinnvoller, als jedem Telefon die Verschlüsselung einzeln zu überlassen, da man Rechenleistung einspart und die für die Verschlüsselung notwendigen Daten nur an einer Stelle vorgehalten werden müssen. Für Privatkunden wird sich dagegen aller Wahrscheinlichkeit nach *in puncto* Abhörbarkeit und Schutz der Privatsphäre nichts ändern.

Selbst bei Benutzung einer Verschlüsselung (durch das Telefon oder den Gateway) wäre es möglich, eine Verkehrsstatistik zu erstellen; bereits die Information, wer wann mit wem telefoniert, kann wichtige Indizien bringen.

Eine größere Gefahr als das Abhören der Verbindung „unterwegs“ besteht allerdings darin, dass es jemandem gelingt, in das dahinter liegende Netzwerk (bei Unternehmen) einzudringen.

Ein weiteres Sicherheitsrisiko besteht darin, dass der Käufer eines Telefons nicht weiß, ob es dem Hersteller die Möglichkeit bietet, einzudringen und Daten zu klauen. (=> **backdoor-Problem**). Bei Cisco, momentan Marktführer bei VoIP-Telefonen, waren solche Funktionen als Supportfunktionen schon des öfteren implementiert worden. Ist eine solche Möglichkeit bei allen Geräten eines Typs vorhanden, so ist es nur eine Frage der Zeit, bis diese Möglichkeit, in das Telefon bzw. die Telefonanlage einzudringen, beispielsweise über entlassene Mitarbeiter, bekannt wird.

6 Realisierung am Beispiel der Universität Ulm

6.1 Das Telefonnetz der Universität Ulm

Das Telefonnetz der Universität Ulm hat ungefähr 9200 Telefonanschlüsse. Dazu gehören nicht nur der Campus, sondern auch die Kliniken am Michelsberg, Safranberg, Eselsberg etc. Diese Außenstellen sind über Glasfaser mit dem „Unizentralnetz“ verbunden. Diese Telefone brauchen keine Stromversorgung [7].

Wird nun ein Telefonat von Intern nach Intern/Extern geführt, so geht das Gespräch zunächst von dem Apparat, an dem man sitzt, zum Verteiler des entsprechenden Niveaus. Von dort wird es in den Keller weitergeleitet. Dort wird entschieden, ob das Gespräch nun

das Intranet verlässt oder nicht [7]. Hierbei spricht man von einer *Leitungsvermittlung* [10].

Da die Telefonanlage der Universität Ulm nun 15 Jahre alt ist, kommt es zur Neuausschreibung. Dabei will die Universität Ulm versuchen, in das neue Telefonsystem eine VoIP-Telefonanlage zu integrieren, die 20 Prozent der Anschlüsse einnimmt. (Die Telefonanlage der Universität Stuttgart besteht derzeit aus 60 Prozent VoIP-Telefonen und 40 Prozent herkömmlichen Telefonen)[7].

6.2 Das IP-Telefonnetz der Universität Ulm

Derzeit hat die Universität Ulm in ihrem Pilotprojekt eine VoIP-Telefonanlage mit ca. 30 VoIP-Telefonen von Siemens. Bei diesen Apparaten wird die Sprache in 10-ms-Samples aufgeteilt. Die Größe des Jitterpuffers liegt somit bei 10 ms (maximal bei 20 ms, also dem 2-fachen der Sample-Größe). Diese VoIP-Telefone sind weniger fehlertolerant, so dass es auch mal sein kann, dass das Gespräch abgebrochen wird [7].

Bei der Internet-Telefonie werden die Sprach-Pakete von einem intelligenten Netz an ihr Ziel weitergeleitet: Die Telefone der VoIP-Telefonanlage der Universität Ulm wandeln die Sprache in digitale Sprachpakete um. Diese werden dann zum nächsten *Switch* und *Router*, den es in jedem Festpunkt gibt, geschickt. Von hier nehmen die Pakete den direkten Weg über das Internet zum Gesprächspartner. Deshalb spricht man hier auch von *Paketvermittlung* [10].

Daher können bei der IP-Telefonie die physikalischen Leitungen von allen Teilnehmern simultan zum Transport von Sprach- und Datenpaketen genutzt werden [10]. Somit sind Gespräche, die über das Internet geführt werden, auch sehr leicht „abhörbar“ (siehe Kapitel 5.1). Leider ist derzeit noch von keinem Hersteller eine Verschlüsselung für Gesprächspakete implementiert worden (siehe Kapitel 5.2). Diese Verschlüsselung wäre wohl mit der Rechenleistung eines 486er gut realisierbar [7].

So problemlos, wie es sich anhört, ist es jedoch nicht:

Die Switches müssen über eine unterbrechungsfreie Stromversorgung verfügen. Zum einen muss von Gesetzeswegen her bei Stromausfall noch mindestens zwei Stunden lang telefoniert werden können, und zum anderen muss jedes Telefon, das angerufen werden will, kontinuierlich am „Netz“ hängen, was eben diese Stromversorgung verlangt. Dies bedeutet höhere Kosten, da an den einzelnen Plätzen, an denen mit VoIP-Telefonen telefoniert wird, eine solche Stromversorgung garantiert werden muss [7]. VoIP-Telefone von Siemens oder Cisco haben zum Beispiel einen Stromverbrauch von 4,5 Watt, sowohl im Ruhezustand als auch während des Telefonats [11].

Die Switches und damit die gesamte Telefonanlage bzw. Teile von ihr können ausfallen, wenn ein solcher Switch zum Beispiel zu heiß wird.

Die Flexibilität der IP-Protokoll-basierten Datennetzwerke zeigt sich im weltweiten Erfolg des Internets mit vielen 100 Millionen Teilnehmern. Die miteinander vernetzten Endgeräte sind völlig heterogen; sie haben aber eine Gemeinsamkeit: die Nutzung des IP-Protokolls zum Austausch der Daten über das Netz [10].

Für das Telefonieren wurde das Protokoll H.323 entwickelt. Naheliegender wäre dann, dass man Telefone verwendet die „direkt“ H.323 unterstützen. Diese können jedoch bei hohem

Datenfluss in der gesamten Netzleitung Schwachstellen aufweisen. Es kommt zu Störungen wie Echos oder Aussetzern. Es können aber auch „nicht reine“ H.323-Telefone verwendet werden, bei denen ein hoher Datenfluss im Netz keine Auswirkungen auf die Sprachqualität hat. Allerdings müssen dann im gesamten Intranet Systemtelefone vom selben Hersteller verwendet werden.

Diese Telefone besitzen proprietäre Leistungen wie zum Beispiel:

- Der Name und die Abteilung des Anrufers werden dem Angerufenen übermittelt; die Anrufnummer ist jedoch nicht, wie beim Handy, mit Namen im Telefon gespeichert; es wird statt dessen der Name übermittelt.
- Am Telefon sind Leuchtdioden für „Sekretärinnenfunktionen“; das Telefon kann also signalisieren, dass Anrufe nicht angenommen wurden, desweiteren besteht die Möglichkeit des Aufschaltens, u.v.m.

Die Inkompatibilitäten zwischen dem in der Universität Ulm verwendeten „nicht reinem“ H.323 und dem „Standard“ H.323 werden, wenn man ein Gespräch nach Extern führt, vom Gateway ausgeglichen.

Die Umstellung der Universität Ulm auf eine reine VoIP-Telefonanlage wäre ein sehr großes Unterfangen, da es tausende von Telefonanschlüssen gibt. Deshalb ist es auch derzeit noch ein Pilotprojekt. Außerdem soll das Zwei-Leitungsnetz erhalten bleiben. Herr Borchert und Herr Klenk haben zum Beispiel ein VoIP-Telefon in Benutzung; andere telefonieren noch auf die herkömmlich Art mit Analog- bzw. Digitaltelefonen [7].

7 Möglichkeiten für Privatkunden

Nachdem im vorangegangenen Abschnitt die Möglichkeiten für den Einsatz von VoIP im großen Stil in Unternehmen oder Verwaltungen im Vordergrund stand und die Telekommunikationsunternehmen das Internet zur Übertragung von Fern- und Interkontinentalgesprächen schon länger nutzen, da sie dadurch ihre Kosten reduzieren können, geht es nun um die Möglichkeiten für den einzelnen Bürger und Telefonkunden, aus VoIP seinen Nutzen zu ziehen.

7.1 Angebot der T-Com

Beim Stichwort Telefon denken die meisten Deutschen auch Jahre nach Beginn der Deregulierung des deutschen Telefonmarktes immer noch zuerst an die Deutsche Telekom. Allerdings sah man sich dort außer Stande, Informationsmaterial für Privatkunden zu verschicken – die schriftliche Antwort der Festnetzsparte T-Com des Konzerns bestand ausschließlich im Hinweis auf das eigene Internetangebot. Doch dort war kein Angebot für Privatkunden zu finden! Im Internet kursieren dafür bereits plausibel klingende Erklärungen: Im Moment verdient die Telekom an den Privatkunden recht gut, da die Vermittlungsstellen allesamt abbezahlt sind und die Tarife durch das quasi-Monopol auf der letzten Meile viel höher als der tatsächliche Aufwand sind. Es gibt also keinen Grund, an diesem sehr zufrieden stellenden *status quo* etwas zu ändern. Daher wartet der ehemalige Staatskonzern ab, was von Seiten der Wettbewerber geschieht. Schön für die Aktionäre der Deutschen Telekom, schlecht für die Privatkunden in Deutschland.

7.2 Angebot von Freenet

Freenet startete vor einigen Monaten das Angebot *freenet iPhone*, mit dem es möglich sein soll, ab 0 cent pro Minute zu telefonieren. Dabei wird VoIP benutzt. Allerdings ist iPhone nicht sofort voll ausgebaut, sondern läuft momentan erst in Stufe zwei von vier.

iPhone benutzt einen freenet-Internetzugang, am Besten natürlich DSL, aber auch ISDN soll reichen. Besitzt man einen PC und möchte mit diesem telefonieren, so ist dies einfach über die freenet-Software möglich; oder man benutzt ein beliebiges eigenes Softwaretelefon. Stattdessen kann auch ein handelsübliches IP-Telefon benutzt werden. Erreichen kann man damit jeden Telefonanschluss, angerufen werden kann man aber nur von anderen iPhone-Anschlüssen aus. Seit der CeBIT funktioniert auch das Telefonieren vom PDA aus, sofern auf diesem Windows CE oder Pocket PC 2003 läuft. Seit Mai gibt es den *freenet DSL Router*, an den man vorhandene Endgeräte wie DECT-Telefone anschließen kann. In Phase 4, deren Starttermin noch nicht feststeht, bekommt dann jeder iPhone-Nutzer eine eigene Telefonnummer (Schema: 032...). Erreichbar sind alle in- wie ausländischen Rufnummern im Fest- wie Mobilfunknetz. Auch kann der Nutzer von überall her angerufen werden. Über die dann gültigen Preise sowie den genauen Einführungsstermin war freenet leider nicht bereit, uns nähere Auskünfte zu geben.

8 VoIP – Ein kritischer Blick

Ist also VoIP die Lösung aller Probleme, gar solcher, von deren Existenz wir noch gar nichts wussten? Oder ist es vielmehr ein *Hype* wie die Kursentwicklung an den Aktienmärkten Ende des letzten Jahrzehnts des 20. Jahrhunderts?

Die Wahrheit liegt sicher wie so oft irgendwo dazwischen; zwar steht mittlerweile in etwa 80% aller deutschen Netzprojekte das Thema VoIP mit auf der Tagesordnung, aber dennoch wird das traditionelle Telefonnetz noch auf Jahre hinaus benutzt werden. Ein Grund hierfür besteht in der im Vergleich zu ISDN noch zu geringen Funktionalität und unter Umständen auch Qualität. Ferner funktioniert das Roaming bei Ferngesprächen zwischen den verschiedenen Carriern noch nicht reibungslos. Und auch ein deutlicher Kostenvorteil ist im Moment noch nicht zu beobachten, zumal eventuell auch auf bisher verfügbare Leistungsmerkmale verzichtet werden muss – hierdurch können Umsatzausfälle entstehen, die zwar nur schwer abzuschätzen sind, die aber eigentlich mit berücksichtigt werden müssen, wenn man einen fairen Vergleich von PSTN und VoIP durchführen möchte.

Bei *Level 3*, einem international agierenden Carrier, der Backbones für andere Telefongesellschaften bereit stellt, rechnet man in Europa erst in zehn Jahren mit einem Durchbruch von VoIP in den Weitverkehrsnetzen – in den USA dagegen schon in spätestens fünf Jahren, da dort bereits homogene Netzstrukturen existieren, im Gegensatz zur EU, wo noch in jedem Land ein Carrier eine Quasi-Monopolstellung inne hat.

Ein weiterer wichtiger Punkt ist die finanzielle Lage der Telekommunikationsunternehmen in Europa: in vielen Staaten haben sie sich im Zuge der UMTS-Lizenz-Versteigerungen stark verschuldet; ein Plus für die bereits abgeschriebenen, kostengünstigen Vermittlungsstellen.

Zu beachten bei der Migration von Unternehmensnetzwerken ist außerdem die Verfügbarkeit des firmeninternen Netzes. Wer da auf ältere Windowsversionen setzt, hat bei VoIP

schon so gut wie verloren: eine Verfügbarkeit von 99,9999% wie bei ISDN-Telefonanlagen ist mit diesen Betriebssystemen nicht zu erreichen. Doch ein Netz-Upgrade kostet Geld – ein weiteres Minus für VoIP. Ein weiteres Problem sind die immer wieder auftretenden Viren-epidemien mit ihrer großen Netzlast, auch das kann Auswirkungen auf die Verfügbarkeit des Kommunikations- und Arbeitsmittels Telefon haben [16].

9 Ausblick

Was bringt uns die Zukunft, wie werden wir im Jahre 2020 telefonieren?

„Die Deutsche Telekom plant, ihr Telefonnetz in den nächsten Jahren vollständig auf Internet-Technologie umzustellen. Nach Aussagen des Bereichsvorstandes Roland Kittel der T-Com, der Festnetzsparte der Telekom, soll die Umstellung deutlich vor dem Jahr 2020 abgeschlossen sein. Marktforscher erwarten, dass bis zu diesem Zeitpunkt jeglicher Telefonverkehr in Europa über das Internet abgewickelt wird.“ [4]

Aus Kostengründen kann man davon ausgehen, dass nicht auf Dauer ein Telefonnetz neben dem Internet existieren wird, denn das Internet kann die Übertragung von Telefonverbindungen leicht mit übernehmen. Die Kunden werden davon genauso wenig etwas merken, wie sie bisher bemerkten, dass die Telekommunikationsunternehmen bereits seit längerem Ferngespräche über das Internet schalten, um Kosten zu sparen, denn sie werden weiterhin eine Telefonnummer wählen, und auch das Telefon als solches wird sich mit seinem Design nicht *verraten*. Nur die Älteren werden noch wissen, dass man früher pro Zeiteinheit bezahlen musste und dass damals eine Minute so viel kostete wie dann eine Stunde oder noch länger... (Quellen: [5] und [6]).

Tabellenverzeichnis

1	Datenraten (Entnommen aus [8])	4
2	Sprachqualität	12

Abbildungsverzeichnis

1	Eines der ersten Telefone	2
2	Paketverlust bei G.711-Codec mit 60 ms Samples/Paket	5
3	Jitter bei G.711-Codec mit 60 ms Samples/Paket	6
4	Paketverlust bei G.711-Codec mit 30 ms Samples/Paket	7
5	Jitter bei G.711-Codec mit 30 ms Samples/Paket	7
6	H.323-Stack	8
7	H.323-Komponenten	8
8	Admission Request	9
9	Location Request	10
10	Bandwidth Request	10
11	Qualitätsanforderungen	12

Literatur

- [1] Roberto Arcomano: *VoIP Howto*, fatamorgana.com
- [2] openh323.org
- [3] Florian Schötzer in iX, 06.02.2004
- [4] optel Media Services GmbH auf www.tariftip.de über VoIP
- [5] freenet.de AG
- [6] onlinekosten.de GmbH, 21.02.2004
- [7] Markus Klenk, Mitarbeiter im KIZ der Universität Ulm
- [8] Westbay Engineers Limited, www.erlang.com
- [9] Mario Gongolsky auf www.firstsurf.de, 11. Juni 2001
- [10] www.rus.uni-stuttgart.de
- [11] www.belwue.de
- [12] www.rvs.uni-hannover.de
- [13] searchnetworking.techtarget.com
- [14] www.itu.int
- [15] Paul E. Jones: *H.323 Protocol Overview*, February 2003, siehe [Packetizer](#)
- [16] Computerwoche vom 02.06.2004 www.computerwoche.de
- [17] Markus Klenk: Voice-over-IP-Lösungen im Weitverkehrsnetz unter Berücksichtigung von Quality-of-Service-Aspekten (Diplomarbeit)