IT6 - Lerneinheit VOIP

Aufgabe:

Machen Sie sich mit dem Kapitel 1
Ausgangssituation - der Internettelefonie vertraut!
(Stellen Sie sich vor Sie sind in Ihrer Firma mit der Umstellung der TK-Anlage auf VoIP beauftragt worden!)

Zeit: 60 min

Bearbeiten Sie dabei folgende Aufgaben in Form von Stichpunkten oder einer Kurzbeschreibung:

- ☐ 1. Welche **Planungskriterien** gilt es zu beachten?
- ☐ 2. Welche **Sicherheitsaspekte** sollten beim Einsatz von VoIP betrachtet werden?
- ☐ 3. Was versteht man unter **QoS** und welche **Mess- und Bewertungsgrößen** werden vorgestellt?
- ☐ 4. Welche der gebräuchlichsten **Protokolle** für die **Signalisierung** werden vorgestellt!
- ☐ 5. Beschreiben Sie das **Protokoll** für die **Datenübertragung**!
- ☐ 6. Wie funktioniert das **Telefon Number Mapping?**

1. Ausgangssituation

Bereits in den 1990er Jahren hat die VoIP schon einmal das Interesse der Öffentlichkeit geweckt, als sie billiges Telefonieren über das Internet versprach. Das Internet bot damals jedoch noch nicht die für die Übertragung von Telefongesprächen in Echtzeit erforderlichen Voraussetzungen, was dazu führte, dass Gesprächsverzögerungen von mehreren Sekunden, Sprachaussetzer und schlechte Tonqualität die Regel waren.
Inzwischen ist die Technologie zur Marktreife weiterentwickelt worden.

1.1 Grundlagen

Die IP-Telefonie ist eine Technologie, die es ermöglicht, den Telefondienst auf der IP-Infrastruktur zu realisieren, sodass diese die herkömmliche Telefontechnologie samt ISDN, Netz und allen Komponenten ersetzen kann. Zielsetzung dabei ist eine Reduzierung der Kosten durch ein einheitlich aufgebautes und zu betreibendes Netz (Konvergenz der Netze - „Everything over IP"), wodurch sich für die Betreiber Kostenvorteile ergeben, die an den Endverbraucher weitergegeben werden können. Die Kostenvorteile sind u. a. dadurch begründet, dass VoIP wesentlich ressourcenschonender mit dem zur Verfügung stehenden Übertragungsmedium umgeht. So lassen sich über eine IP-gesteuerte Breitband-Verbindung mehr Sprachverbindungen realisieren als bei der klassischen Nutzung einer Telefonleitung.

Aufgrund der hohen Einsatzdauer klassischer Telefoniesysteme und der notwendigen Neuinvestitionen für IP-Telefonie wird der Wechsel bei bestehenden Anbietern oft als lang andauernder, gleitender Übergang realisiert. Währenddessen existieren beide Technologien parallel (sanfte Migration). Daraus ergibt sich ein deutlicher Bedarf an Lösungen zur Verbindung beider Telefoniesysteme (z. B. über VoIP-Gateways) und die Notwendigkeit zur gezielten Planung des Systemwechsels unter Berücksichtigung der jeweiligen Möglichkeiten für Kosten- und Leistungsoptimierung.
Neben Sprache lassen sich per IP-Verbindung zugleich weitere multimediale Daten austauschen (Video over IP, verteiltes kooperatives Arbeiten usw.). Allerdings bringt die Integration von Sprache und Daten in einem gemeinsam genutzten IP-Netz einige neue Probleme mit sich, da jetzt Dienste mit völlig unterschiedlichem Verkehrsverhalten über eine Leitung übertragen werden. Die Integration von zeitkritischen Diensten wie Sprache, Streaming Audio und Video ist zum Teil mit neuartigen Anforderungen an die Infrastruktur verbunden.
Neue Anbieter drängen dennoch zunehmend mit neuer Technologie (IP-Telefonie statt herkömmlichem Telefon) auf den Markt.

Voice over IP im OSI-Schichtenmodel

Schicht	Protokoll
7. Anwendung	Softphone
6. Präsentation	G.729 / G.723 / G.711
5. Session	H.323 / SIP
4. Transport	RTP / UDP / TCP / RSVP[2]
3. Netzwerk	IP
2. Verbindung	ATM / Ethernet
1. Bitübertragung	DSL / Ethernet

Bild 1: VoIP-Protokolle im OSI-Schichtenmodell

Bei VoIP muss zwischen den Datenpaketen zum Verbindungsauf- und -abbau (Signalisierung) und den eigentlichen Sprachpaketen (Datenstrom) unterscheiden werden.

Die ***Signalisierungsdaten*** (sie steuern die Verbindung) müssen möglichst sicher übertragen werden. Sie können länger übertragen werden und einen größeren Protokoll-Overhead haben. Entscheidend ist, dass die Verbindung zustande kommt (TCP).

Die **Sprachpakete** dagegen müssen möglichst schnell und verzögerungsfrei übermittelt werden. Sollte ein Datenpaket verloren gehen, so kann dies toleriert werden. Daher kann man sich hier eine unsichere Übertragung leisten (UDP).

In der Praxis werden die Sprachpakete zuerst in RTP-Pakete und dann in UDP-Pakete verpackt und zur Adressierung zusätzlich mit einem IP-Header versehen.

1.2 Planung des VoIP-Einsatzes

Eine wesentliche Voraussetzung für den sicheren Einsatz von VoIP ist eine gründliche Planung im Vorfeld.

Es sollten beispielsweise folgende Fragestellungen behandelt werden:

- Soll vollständig oder teilweise auf VoIP umgestiegen werden?
- Gibt es besondere Anforderungen an die Verfügbarkeit von VoIP oder an die Vertraulichkeit und Integrität der Telefonate bzw. der Signalisierungsinformationen?
- Welche Signalisierungs- und Medientransportprotokolle sollen eingesetzt werden?
- Wie vielen Benutzern soll die Kommunikation über VoIP ermöglicht werden?
- Wie soll die Anbindung ans öffentliche Telefonnetz erfolgen? Sollen VoIP-basierte Kommunikationsverbindungen direkt aus dem öffentlichen Datennetz gestattet werden?
- Kann die Sicherheit des vorhandenen LANs durch VoIP beeinträchtigt werden? Ist das vorhandene LAN für die Nutzung von VoIP ausreichend dimensioniert? Müssen Änderungen an der Netzarchitektur vorgenommen werden?

[2] Das Resource Reservation Protocol (RSVP) ist ein Signalisierungsprotokoll im Internet Protocol-Stack. Es erlaubt Empfängern außerhalb einer Multicast-Gruppe, deren Dienstanforderungen festzulegen. Damit können für bestimmte Anwendungen, z. B. für die Übertragung von Video-Streams, bestimmte Übertragungsraten für einzelne Verbindungen reserviert werden.

1.3 Sicherheitsrisiken von VoIP

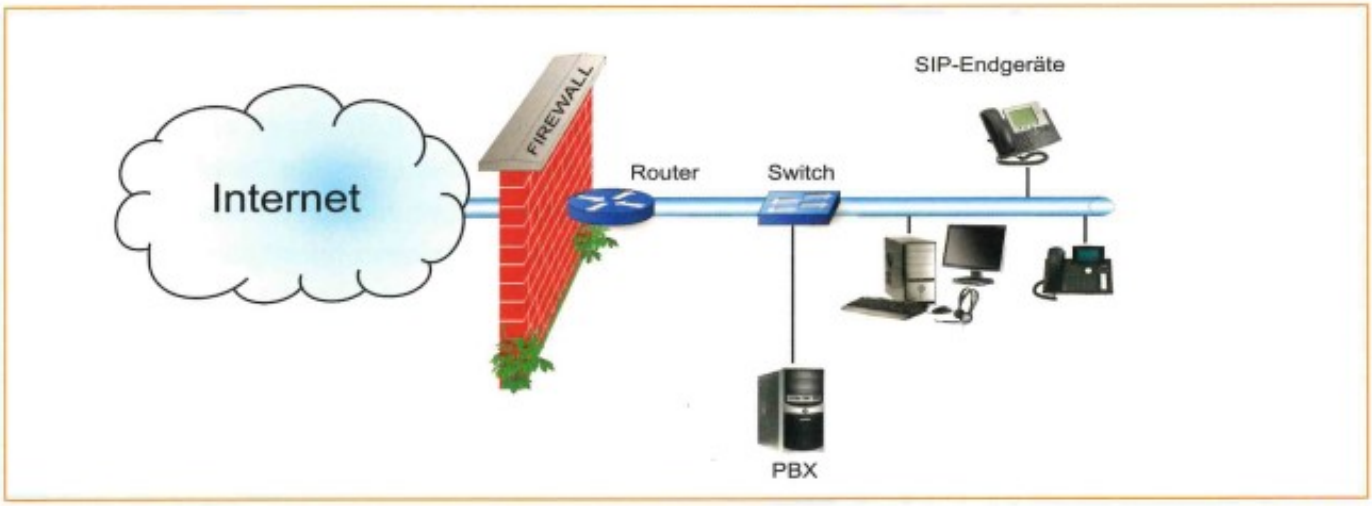

Bild 2: Installieren Sie den PBX-Server im selben Netz wie die IP-Endgeräte

Im Folgenden sollen einige Sicherheitsaspekte von VoIP betrachtet werden.

1.3.1 Anmeldung

Die Anmeldung des SIP-Clients am Registrar Server (oft identisch mit dem SIP-Proxy) des Providers erfolgt normalerweise unverschlüsselt und unsigniert (siehe Bild 8: Übersicht eines SIP-Kommunikationsaufbaus). Wer den Netzverkehr mitlesen kann erhält so die Anmeldeinformationen und kann diese auch selbst nutzen. Ein Angreifer kann dem Server also vortäuschen, selbst der Benutzer zu sein und dann auf Kosten des echten Kunden telefonieren und seine Anrufe entgegennehmen.

Die Lösung dieses und vieler verwandter Probleme wäre der Einsatz des SIPS (SIP over TLS).

1.3.2 Gesprächsübertragung

Weil der Datenverkehr während des Telefonats per RTP ebenfalls unverschlüsselt und unsigniert erfolgt, kann jeder, der den Verkehr mitlesen kann, das Gespräch aufzeichnen und wiedergeben oder manipulieren. Insbesondere das Aufzeichnen und Wiedergeben lässt sich mit freier Software und auch unter Windows sehr einfach erledigen. Hängt der Angreifer am gleichen Switch wie das Opfer, kann z. B. Cain & Abel[3] (siehe auch unter „5. Anlage zum Datenrecht") benutzt werden, um einen Man-In-The-Middle-Angriff auszuführen. Damit lässt sich dann der Verkehr mitschneiden und die RTP-Pakete per Filter isolieren. Cain & Abel bietet auch Funktionen an, um den RTP-Datenstrom zu analysieren, dabei zusammenzusetzen und als Audiodatei zu exportieren.

Auch hier existiert es eine Weiterentwicklung des Protokolls, nämlich SRTP, die diese Probleme verhindern könnte. Im Markt hat sich SRTP wie auch SIPS aber noch nicht voll durchgesetzt.

1.3.3 Trennung des Daten- und VoIP-Netzes

IP-Telefonie ermöglicht das Telefonieren über existierende IP-Datennetze. Jedoch können zur Erhöhung von Skalierbarkeit, Dienstqualität (QoS), Administrierbarkeit und Sicherheit die Datennetze von den Sprachnetzen auch logisch getrennt werden. Es muss überprüft werden, ob eine Trennung von Daten- und VoIP-Netz erforderlich ist. Eine Trennung ist sinnvoll, wenn Daten- und VoIP-Netz einen unterschiedlichen Schutzbedarf haben.

[3] **Rechtliches**
Da Cain & Abel Sicherheitsvorkehrungen umgeht, muss es nach Inkrafttreten des so genannten Hackerparagrafen (§202c StGB) in Deutschland als Computerprogramm zum Ausspähen von Daten aufgefasst werden. Somit kann die illegale Benutzung der Software unter Strafe gestellt werden.

1.3.4 Trennung der Netze über VLANs

Lokale Netze können physikalisch durch aktive Netzkomponenten oder logisch durch eine entsprechende VLAN-Konfiguration segmentiert werden. Eine logische Trennung kann mit VLAN-Technologie auf OSI-Schicht 2 mit VLAN-fähigen Switches aufgebaut werden. VLANs alleine bieten jedoch keinen Schutz vor Angreifern, die sich mit ihrem IT-System (PC, Laptop) physikalisch an ein VLAN anschließen. Da die Netzdose, also der VLAN-Port des Telefons jedem unmittelbar zugänglich ist, könnte ein Angreifer direkt die Telefone im VLAN angreifen, indem er z. B. anstatt eines Telefons seinen PC mit dem VLAN verbindet.
Aus diesem Grunde sollten weitere, über die logische Netztrennung hinausgehende Maßnahmen getroffen werden, um derartigen Angriffe zu begegnen.

1.3.5 Physikalische Trennung der Netze

Bei erhöhten Sicherheitsanforderungen kann eine komplette physikalische Trennung des Sprachnetzes vom Datennetz sinnvoll sein. Die physikalische Trennung von Daten- und Sprachnetzen verringert deutlich die Angriffsmöglichkeiten. Außerdem kann bei dem Ausfall eines Netzes, beispielsweise durch den Ausfall der aktiven Netzkomponenten oder einem Kabelbruch, weiterhin über das verbleibende Netz kommuniziert werden. Durch die Trennung hat die Auslastung des Datennetzes keinen Einfluss auf die Auslastung des Sprachnetzes.

1.3.6 Verschlüsselung von VoIP

Gelingt es einem Angreifer, sich Zugang zu einem internen Netz zu verschaffen, kann er die gesamte Netzkommunikation im LAN protokollieren. Falls die VoIP-Nutzlast nicht verschlüsselt ist, kann der Angreifer sämtliche Informationen mitlesen. Daher sollte überlegt werden, ob eine Verschlüsselung der VoIP-Nutzdaten möglich ist. Eine Verschlüsselung muss jedoch von alle beteiligten TK-Systeme unterstützen werden.
Bei dieser Überlegung ist es zweckmäßig, zwischen interner und externer Kommunikation zu unterscheiden.

Für VoIP-Telefonate innerhalb eines LANs kann überlegt werden, ob auf eine Verschlüsselung verzichtet wird, wenn sichergestellt werden kann, dass auf Informationen nicht über einen unsicheren Netzbereich (wie einem WLAN) durch einen Außentäter zugegriffen werden kann. Um die internen Gespräche vor dem Zugriff durch Innentätern zu schützen, kann der Einsatz einer Verschlüsselung dennoch sinnvoll sein. Hierfür ist der Betrieb der VoIP-Endgeräte in einer VPN-End-to-End–Struktur oder die Nutzung eines verschlüsselten Transportprotokolls (SRTP) denkbar.
Verlassen Pakete mit VoIP-Inhalten das gesicherte LAN, müssen sie mit entsprechenden Verfahren geschützt werden. Für den Schutz der VoIP-Kommunikation ist eines oder mehrere der folgenden Verfahren auszuwählen:

- Nutzung verschlüsselnder Transportprotokolle, wie SRTP (Secure Realtime Transport Protocol).
- Verschlüsselung der Signalisierungsprotokolle, wie SIPS (SIP over TLS).
- Virtual Private Networks (VPNs):
  Durch den Einsatz von VPN-Gateways können Informationen verschlüsselt zwischen entfernten LANs übertragen werden. Einzelne Geräte können als VPN-Endpunkte betrieben werden. Ohne eine direkte Unterstützung von verschlüsselnden Signalisierungs- und Transportprotokollen kann auf dieser Weise eine protokollunabhängige Verschlüsselung eingesetzt werden.
- Verschlüsselung des Funknetzes:
  Sind die VoIP-Gesprächsteilnehmer über ein WLAN miteinander verbunden, muss ein qualifizierter Schutz für das WLAN, wie WPA2, genutzt werden. Da sich diese Verschlüsselung auf das Funknetz beschränkt, ist zu beachten, dass die Informationen im restlichen LAN ungeschützt übertragen werden.
- Soll ein Gespräch zu einem Telefonteilnehmer über ein öffentliches Telefonnetz aufgebaut werden, kann die Verbindung zwischen dem VoIP-

Endgerät und dem Gateway, der zwischen dem IP-Netz und dem öffentlichen leitungsvermittelnden Netz eingesetzt wird, gegebenenfalls mit VPNs oder verschlüsselnden Signalisierung- und Transportprotokollen geschützt werden. Da nur sehr wenige Telefone für leitungsvermittelnde Netze Schutzmechanismen bereitstellen und deren Einsatz vom jeweiligen Empfänger abhängig ist, ist eine Verschlüsselung zwischen VoIP-Gateway und dem Gesprächspartner meist nicht realistisch.

Ist eine verschlüsselte Kommunikation, beispielsweise zu externen Gesprächspartnern, nicht möglich, müssen die Benutzer hierüber informiert und sensibilisiert werden. Vertrauliche Gespräche sollten bei einer fehlenden Verschlüsselung nicht über das Telefon geführt werden.
Bei der Beschaffung von VoIP-Komponenten sollte darauf geachtet werden, dass diese verschlüsselnde Signalisierungs- und Transportprotokolle wie z. B. TLS und SRTP unterstützen.

1.4 Quality of Service (QoS)

Die Anforderungen an das Netz für Datenübertragung und IP-Telefonie unterscheiden sich erheblich. Voice over IP braucht immer einen verzögerungsfreien und kontinuierlichen Datenstrom.
Neben der erforderlichen Übertragungskapazität (ca. 100-120 kbit/s für ein Gespräch kodiert mit G.711) haben insbesondere Qualitätsmerkmale wie Verzögerung, Schwankungen in der Übertragung (Jitter) und Paketverlustrate erheblichen Einfluss auf die resultierende Sprachqualität. Durch Priorisierung und geeignete Netzplanung ist es möglich, eine mit der herkömmlichen Telefonie vergleichbare Sprachqualität und Zuverlässigkeit unabhängig von der Verkehrslast zu erreichen.

Qualitätsmerkmale

Um eine qualitativ hochwertige Kommunikation über Voice-over-IP führen zu können, müssen die für den Sprachtransport verwendeten Datenpakete so beim Empfänger ankommen, dass sie zu einem getreuen Abbild des ursprünglichen, zeitlich zusammenhängenden Datenstrom zusammengesetzt werden können.
Übertragungsfehler, Verzögerungen und Laufzeitunterschiede lassen sich nur durch ausreichende Bandbreite und/oder Protokollzusätze vermeiden. Diese Maßnahmen werden unter dem Begriff Quality-of-Service (QoS) zusammengefasst.

Mean Opinion Score (MOS)

Eine verbreitete subjektive Messgröße zur Bewertung der Sprach-Codec-Qualität ist der mittlere Meinungswert (Mean Opinion Score – MOS). MOS-Tests werden jeweils bei einer Gruppe von Zuhörern durchgeführt, die Sprach- und Klangqualität auf Grund ihrer subjektiven Empfindungen mit Noten von 1 (schlecht) bis exzellent (5) bewerten müssen. Daneben finden auch rechnergestützte Analyseverfahren Einsatz.

Wert	Qualität	Grad der Beeinflussung
5	exzellent	unmerklich; es ist keine Anstrengung nötig, um die Sprache zu verstehen
4	gut	gerade bemerkbar, nicht störend; durch aufmerksames Hören kann die Sprache ohne Anstrengung wahrgenommen werden
3	ordentlich	bemerkbar, leicht störend; die Sprache kann mit leichter Anstrengung wahrgenommen werden
2	mäßig	störend; es bedarf großer Konzentration und Anstrengung, um die übermittelte Sprache zu verstehen
1	mangelhaft	sehr störend und unangenehm; trotz großer Anstrengung kann man sich nicht verständigen

Bild 3: Subjektive Bewertung der Sprachqualität – Mean Opinion Score

Die im nachfolgenden aufgeführten Faktoren bestimmen die Qualität des Systems.

Durchsatz

Der erforderliche Datendurchsatz hängt in erster Linie von der verwendeten Codierung ab. Ein unkomprimiertes Gespräch hat typischer Weise eine Datenrate von 64 kbit/s (ISDN). Abhängig vom verwendeten Kompressionsverfahren beträgt die für die reine IP-Telefonie benötigte Bandbreite rund 100 kbit/s (64 kbit/s netto zuzüglich der Overheads der verschiedenen Kommunikations-Protokolle) in beide Richtungen.
Die Wahl des Codecs (Codec (***co***der und ***de***coder) bezeichnet ein Verfahren bzw. Programm, das Daten oder Signale digital kodiert und dekodiert) ist immer ein Kompromiss zwischen

- der Sprachqualität,
- dem Bandbreitenbedarf,
- dem Codierungsdelay,
- der benötigten Rechenleistung.

Die folgende Tabelle zeigt den MOS für verschiedene ITU-standardisierte Telefonieverfahren. Es lässt sich ersehen, dass ausgefeilte LPC-Verfahren den Bandbreitenbedarf deutlich reduzieren, ohne einen subjektive Qualitätsverlust zu verursachen. So erreicht ein algebraisches CELP-Verfahren (Conjugate Structure ACELP) nahezu den gleichen Akzeptanzwert beim Benutzer wie PCM. Dabei reduziert es jedoch die Transferrate auf ein Achtel.

	G.711	G.726	G.728	G.729	G.723.1	G.723.1
Bit Rate [kbit/s]	64	16	16	8	5.3	6.3
Kodierung	PCM	ADPCM	LD-CELP	CS-ACELP	CELP	MP-MLQ
Algorithmic Delay [ms]	0.125	0.125	0.625	10	30	30
Mean Opinion Score	4.1	3.85	3.61	3.92	3.65	3.9

Bild 4: ausgewählte Codecverfahren im Zusammenhang mit Bandbreitenbedarf und MOS-Wert

Laufzeit (Latenz)

Die für den Transport von Daten benötigt Zeit wird als Laufzeit bzw. Latenz (engl. Delay oder latency) bezeichnet und ist bei herkömmlicher Telefonie im Wesentlichen die Summe der Signallaufzeiten auf den Übertragungskanälen. Bei Telefonie über IP-Netze kommen weitere Verzögerungen durch die Paketierung und Zwischenspeicherung sowie gegebenenfalls Kompression und Dekompression der Daten hinzu. Bei der Telefonie stellen gemäß ITU-T Empfehlung G.114 bis 400 Millisekunden Einweglaufzeit (Mund zu Ohr) die Grenze dar, bis zu der die Qualität von Kommunikation in Echtzeit noch als akzeptabel gilt. Ab ungefähr 125 Millisekunden kann die Laufzeit vom Menschen jedoch schon als störend wahrgenommen werden.

Verzögerung (Delay)

One Way Delay	Beschreibung
0 bis 150 ms	akzeptabel für die meisten Anwendungen
150 bis 400 ms	akzeptabel, wenn die Administration sich des Einflusses der Übertragungszeit auf die Qualität der Übertragung der Anwendungsdaten bewusst ist
über 400 ms	nicht akzeptabel für die meisten Netzwerkplanungen; unter Umständen kann aber auch dieses Limit überschritten werden

Bild 5: Verzögerungszeiten und ihre Auswirkung für die Übertragungsqualität

Ursache	Laufzeit
A-D-Wandlung	20 ms
Paketerstellung	30 ms
sonstige Servicezeiten	10 ms
Routing über 800 Kilometer	50 ms
Jitter Buffering	30 ms
D-A-Wandlung	20 ms
Laufzeit gesamt	160 ms

Bild 6: Faktoren, die die Übertragungszeit bestimmen

Laufzeit mit Ping messen

Um Verzögerungen auf einer Übertragungsstrecke zu messen, bietet sich der Ping als grobe Abschätzung an.

Soll die Messung einigermaßen realistisch gestaltet werden, muss die Paketgröße von Ping eingestellt werden. Geht man von der Kodierung mit G.711 und 20 ms Sprachdaten pro Paket aus, dann entspricht das 160 Byte (64 kBit/s x 0,02 s). Dazu müssen noch 40 Byte für den IP/UDP/RTP-Header-Anteil hinzugerechnet werden. Der Ping sollte somit 200 Byte pro Paket verschicken.

Unter Windows lautet das Ping-Kommando demnach

ping –l 200 –t {Hostname}.

Durch das Attribut –t wird der Ping solange wiederholt, bis die Tastenkombination Strg + C gedrückt wird. Unter Linux würde das Ping-Kommando

ping –s 200 {Hostname} lauten.

Dabei ist zu beachten, dass der Ping die Gesamtverzögerung von Hinweg und Rückweg (Round-Trip-Time, RTT) misst. Sprachdaten dagegen werden nur in eine Richtung übertragen und enden beim Empfänger. Der Empfang des Paketes wird auf Transportebene nicht bestätigt. Deshalb muss der Wert, den Ping liefert, halbiert werden.

Beachte:
Diese Messung dient nur als eine erste grobe Abschätzung; eine Messung, die zu aussagekräftigen und korrekten Werten führt, wird dadurch nicht ersetzt.

Laufzeitschwankungen (Jitter)

Als Jitter bezeichnet man die zeitliche Schwankung zwischen dem Empfang von zwei Datenpaketen. Um diese zu kompensieren, werden „Pufferspeicher" (Jitterbuffer, in dem die Daten zunächst für 30-50 Millisekunden gesammelt werden) eingesetzt, die eine zusätzliche, absichtliche Verzögerung der empfangenen Daten bewirken, um anschließend die Daten isochron auszugeben. Pakete, die noch später ankommen, können nicht mehr in den Ausgabedatenstrom eingearbeitet werden. Die Größe des Pufferspeichers (in Millisekunden) addiert sich zur Laufzeit. Sie erlaubt somit die Wahl zwischen mehr Verzögerung oder mehr Paketverlustrate.

Paketverlust

Von Paketverlust spricht man, wenn gesendete Datenpakete den Empfänger nicht oder nicht in der richtigen Reihenfolge erreichen und deshalb verworfen werden. Bei Echtzeitanwendungen spricht man auch von Paketverlusten, wenn das Paket zwar den Empfänger erreicht, aber zu spät eintrifft, um noch in den Ausgabestrom eingefügt werden zu können. Für Telefonie wird nach ITU-T G.114 eine Paketverlustrate (packet loss rate) bis maximal 5 % noch als akzeptabel eingestuft.

Verfügbarkeit

Die Verfügbarkeit des Gesamtsystems ergibt sich aus den Einzelverfügbarkeiten der beteiligten Komponenten und deren Zusammenschaltung (kaskadiert – in Reihe, oder redundant – parallel). Somit hängt die Verfügbarkeit eines IP-Telefonie Systems in erster Linie vom Netzdesign ab. Eine US-amerikanische Studie vom Juni 2005 untersuchte die Verfügbarkeit von IP-Telefonie in den USA. Im Durchschnitt wurden knapp 97 % erreicht. Das entspricht einem Ausfall an insgesamt 11 kompletten Tagen im Jahr. Zudem gibt es bei vielen DSL-Providern eine so genannte 24 Stunden-Zwangstrennung, die dazu führt, dass bei ständig benutzter Leitung eine Trennung stattfindet. Die daraufhin nötige Neueinwahl kann unter Umständen mehrere Minuten dauern.

1.5 Signalisierungs- und Übertragungsprotokolle

Da die meisten Anwender im IP-Netz ihre IP-Adresse dynamisch beziehen, stellt der Verbindungsaufbau im VoIP-Netz eines der größten Probleme dar. Hier ist ein zentraler Server erforderlich, auf dem die URIs (SIP), Aliases (H.323) bzw. Benutzernamen (Skype) aufgelöst und auf die entsprechende IP-Adresse gemappt werden.

Die Aufgaben eines Signalisierungsprotokolls lassen sich wie folgt zusammenfassen:

- Aufbau der Sitzung (Session)
  - Akzeptieren / Ablehnen des Anrufs
  - Umleitung des Anrufs (zu anderen Personen, zur Voice Mail, zur Web Page, ...)
- Aushandeln der Sitzungsmerkmale
  - Festlegung der Kompressionsalgorithmen (für Audio und Video)
  - Festlegung der Ports
- Hinzufügen neuer Teilnehmer, Verlassen der Konferenz

1.5.1 Signalisierungsprotokolle – H323 vs. SIP

Beim Einsatz von VoIP werden die Steuerinformationen und die Sprachdaten getrennt voneinander, mittels unterschiedlicher Übertragungsprotokolle transportiert. Steuerinformationen, z. B. "besetzt", werden über Signalisierungsprotokolle, wie H.323 oder SIP (Session Initiation Protocol), übermittelt. Für die Übertragung der Sprachdaten ist ein Medientransportprotokoll, wie RTP (Real-Time Transport Protocol), zuständig.

Da die verschiedenen Signalisierungsprotokolle untereinander nicht kompatibel sind, spielt die Auswahl für den Aufbau eines VoIP-Netzes eine wichtige Rolle. VoIP-Komponenten, die kein gemeinsames Protokoll unterstützen, können ohne ein Gateway nicht miteinander kommunizieren.

Die gebräuchlichsten Signalisierungsprotokolle sind H.323 und SIP.

1.5.1.1 H.323

Die Protokollgruppe um H.323 beschreibt die Übertragung von Echtzeitinformationen (Video, Audio, Daten) in paketorientierten Transportnetzen. H.323 wurde ursprünglich als Umsetzung des ISDN D-Kanal Protokolls Q.931 auf ein IP-basiertes Netz entwickelt. Innerhalb von dieser Protokollgruppe sind die Protokolle H.225.0, H.245 und H.450 und H.235 definiert. H.323 beschreibt den Rahmen der Signalisierungsprotokolle, H.225.0 die eigentliche Signalisierung, H.245 die Kontrolle der Übertragung der Sprachinformationen und H.450 die eigentliche Telefonie-Funktion. Die optionale Unterstützung von H.235 bietet Schutz der Integrität und Vertraulichkeit der Signalisierung. Audio- und Videodaten werden per UDP, Faxdaten per UDP oder TCP übertragen. Vor der Übertragung dieser Echtzeitdaten werden logische RTP- und RTCP-Kanäle zwischen den Endpunkten (VoIP-Endgeräten) aufgebaut.

An einer H.323-Kommunikation können folgende Komponenten beteiligt sein:

- Terminals (VoIP-Endgeräten) stellen die Endpunkte einer Kommunikation beim Benutzer dar. Eine direkte Verbindung zwischen den Endgeräten ist nur bei bekannter IP-Adresse möglich.
- Gatekeeper werden zur Verwaltung eingesetzt. Da die direkte Verbindungsaufnahme zwischen Terminals nur bei bekannten IP-Adressen möglich ist, agiert ein Gatekeeper als zentrale Steuerkomponente in H.323-Netzen.
- Die Multipoint Control Unit (MCU) ermöglicht Gespräche zwischen mehr als zwei Anwendern (Konferenzen).
- Gateways realisieren die Übergänge in andere Netze und nehmen dabei die Anpassung der Nutzdaten und der Signalisierungsinformation vor. Beispielsweise vermitteln Gateways zwischen IP- und leitungsvermittelnden Telefonnetzen.

Der größte Nachteil von H.323 ist die Komplexität des Protokolls. Die Vielzahl der verschiedenen Protokolle lässt H.323 sehr unübersichtlich und aufwendig wirken. Diese Komplexität erschwert die Fehlersuche und kann zu Mehrkosten führen. Erschwerend kommt hinzu, dass das im Folgenden vorstellte SIP von vielen Herstellern bei neueren Produkten priorisiert wird.

1.5.1.2 Session Initiation Protocol (SIP)

Das Session Initiation Protocol ist ein textbasierendes Sitzungssignalisierungsprotokoll der IETF (Internet Engineering Task Force) zum Aufbau, zur Steuerung und zum Abbau einer Kommunikationssitzung zwischen zwei und mehr Teilnehmern. Dabei ist es nicht auf Internet-Telefonie beschränkt, sondern Sessions können beliebige Multimediaströme, Konferenzen, Computerspiele usw. verwalten. In der IP-Telefonie ist das SIP ein häufig angewandtes Protokoll. Es wird in RFC 3261 beschrieben.

Das Adressierungsschema von SIP ähnelt dem einer e-Mail-Adresse (sip:benutzername@provider-name.org). Die Lokalisierung erfolg über DNS (Domain Name System). SIP unterstützt Punkt-zu-Punkt- und Punkt-zu-Mehrpunkt-IP-Verbindungen. Durch das einfache Klartextdesign der SIP-Pakete und der geringen Komplexität erfährt SIP eine immer größere Verbreitung. Die an einer Kommunikation über SIP beteiligten VoIP-Komponenten werden im Punkt 1.5.1.2.1. SIP-Architektur und Komponenten beschrieben.

Da über eine SIP-Adresse die aktuelle IP-Adresse eines Teilnehmers ermittelt werden kann, bietet sich auch die Möglichkeit, dass man in Zukunft über eine Adresse erreichbar sein wird, die dann sowohl für E-Mail als auch Telefonie verwendet werden kann. (Hier gibt es z. Zt. Jedoch noch einige Probleme, die sich aus der Technik von VoIP ergeben, denn durch die Rufnummern bei VoIP kann grundsätzlich nicht wie beim Festnetz geschlossen werden, woher der Anruf kommt, was beim Wählen von Notrufnummern zum Beispiel entscheidend sein kann.)

SIP-Architektur und Komponenten

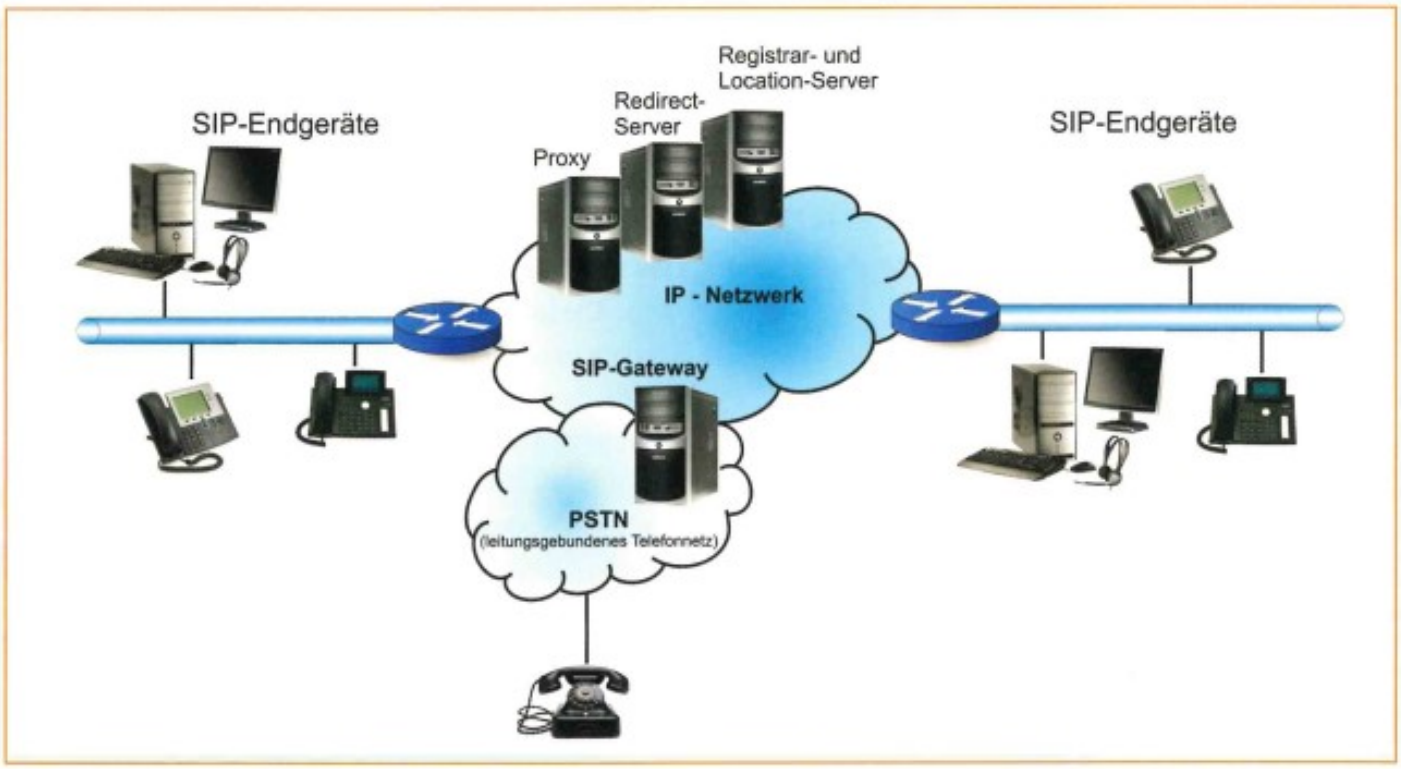

Bild 7: SIP-Komponenten

SIP-Endgeräte: alle SIP-fähigen Endgeräte (d.h. SIP-Telefone, Softwaretelefon, PDAs, Gateways etc.); Es gibt zwei Arten von SIP- Endgeräten (User Agent – UA):

- User Agent Clients (UAC), initiiert die Session
- User Agent Server (UCS), beantwortet die Anfrage von UACs.

Proxy Server: übernimmt das Routing der SIP-Nachrichten zum Aufbau einer SIP-Verbindung. Verbindungsgesuche eines Anrufers können über mehrere Proxies geleitet werden.

Redirect Server: eingehenden Anfragen zum gewünschten Empfänger werden aus der location database ausgelesen. Die gefundenen Aufenthaltsorte (es ist bei SIP möglich, dass man sich zur gleichen Zeit mit verschiedenen Endgeräten von verschieden Orten aus mit dem gleichen Benutzer anmelden kann) werden dann dem Sender der Anfrage mit einer entsprechenden Nachricht zurückgesandt.

Registrar Server: empfängt in bestimmten Zeitintervallen Register Requests von Teilnehmern. Die Requests beinhalten aktuelle Location des Teilnehmers (aktuelle IP-Adresse, Port, Benutzername, Domain). Die Informationen des Requests werden zum Location Server weitergeleitet und dort in einer sog. ,'location database" hinterlegt. Auf diese kann dann der Proxy-Server des Angerufenen zugreifen, um den Aufenthaltsort herauszufinden.
Bei einem Registrar handelt es sich in der Regel um eine logische Einheit – wegen der engen Verzahnung von Registrar und Proxy sind diese meist in einem Gerät (Soft- oder Hardware) zusammengefasst.

Location Server: teilt anfragenden Teilnehmern die aktuelle Locations der gewünschten Teilnehmer mit.

SIP-Gateway: stellt Schnittstellen zwischen VoIP und konventionellen Telekommunikationsnetzen zur Verfügung.

■ Darstellung einer SIP-Kommunikation

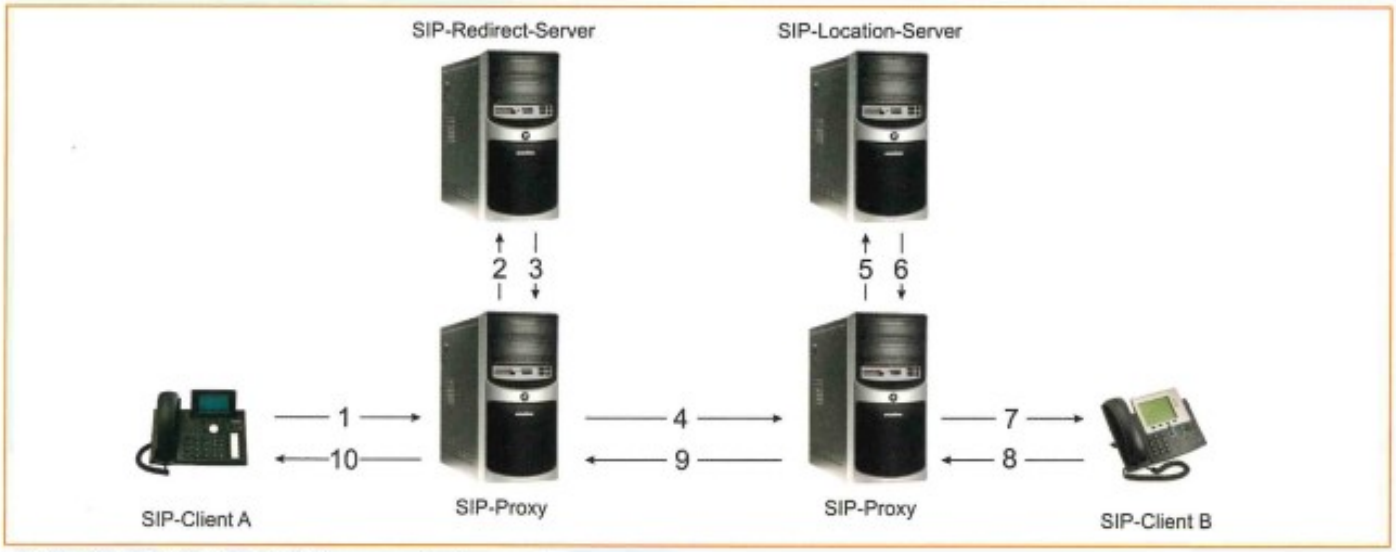

Bild 8: Übersicht eines SIP-Kommunikationsaufbaus

1. Der SIP-Client A sendet eine Anfrage (INVITE) an den SIP-Proxy.
2. Der SIP-Proxy richtet diese Anfrage an den für die Domain der Zieladresse zuständigen Server.
3. Da der SIP-Client B z. Zt. außerhalb seiner Heim-Domain erreichbar ist, erhält der SIP-Proxy eine Antwort mit der neuen URI.
4. Der Proxy richtet nun die INVITE–Nachricht an den zuständigen Ziel-Proxy.
5. Der Ziel-Proxy fragt den Location-Server nach der Lokation (aktuelle IP-Adresse, Port, Benutzername, Domain) des SIP-Client B.
6. Der Location-Server übermittelt die benötigten Informationen an den SIP-Proxy.
7. Der SIP-Proxy sendet die Nachricht an den SIP-Client B.
8. Der SIP-Client antwortet an den SIP-Proxy.
9. Der SIP-Proxy leitet die Antwort an den ursprünglichen SIP-Proxy.
10. Über den ursprünglichen SIP–Proxy gelangt die Antwort an den SIP-Client A.

SIP-Nachrichten

SIP-Requests (Anfragen)

SIP definiert Nachrichten, die für die Kommunikation zwischen Client und SIP-Server verwendet werden. Diese Nachrichten sind:

INVITE:	um einen Benutzer zu einem Anruf einzuladen.
BYE:	um eine Verbindung zwischen zwei Endpunkten zu beenden.
ACK:	für den zuverlässigen Austausch von INVITE-Nachrichten.
OPTIONS:	um Informationen über die Parameter eines Anrufs zu erfahren.
REGISTER:	gibt die Informationen über die Adresse des Benutzer an den SIP Registrierungsserver.
CANCEL:	um die Suche nach der Adresse eines Benutzers abzubrechen.
INFO:	übermitteln zusätzlicher Informationen während einer bestehenden RTP-Session.

SIP-Responses (Rückmeldungen)

Anhand der Status-Codes wird zwischen vorläufigen Antworten (im Bereich von 100 -199) und finalen Antworten (ab 200) unterschieden.
Die Status-Codes können in folgende Klassen eingeteilt werden:

Informational (100 – 199):
Dieser Bereich wird benutzt, um den Status während des Sitzungsaufbaus (100 Trying) zu übermitteln. Beispielsweise wird durch 180 Ringing signalisiert, dass es auf der Gegenseite „klingelt".

Success (200 – 299):
Dieser Bereich zeigt die erfolgreich verlaufene Bearbeitung einer Anfrage an. Ein Beispiel hierfür ist 200 OK.

Redirect (300 – 399):
Mit einer Redirect-Antwort wird eine Weiterleitung eines UACs erreicht. In dieser Antwort kann sich eine Liste von alternativen Adressen befinden.
Beispiele sind 300 Multiple Choices und 301 Moved Permanently.

Client Error (400 – 499):
Diese Klasse findet Verwendung, wenn die Anfrage vom Server nicht bearbeitet werden kann, z. B. 400 Bad Request 401 Unauthorized oder 404 Not Found.

Server Error (500 – 599):
Der Server kann eine gültige Nachricht nicht bearbeiten, beispielsweise 500 Internal Server Error oder 505 Version Not Supported.

Global Error (600 – 699):
Die Anfrage kann generell bei keinem Server erfüllt werden, beispielsweise 600 Busy Everywhere oder 606 Not Acceptable.

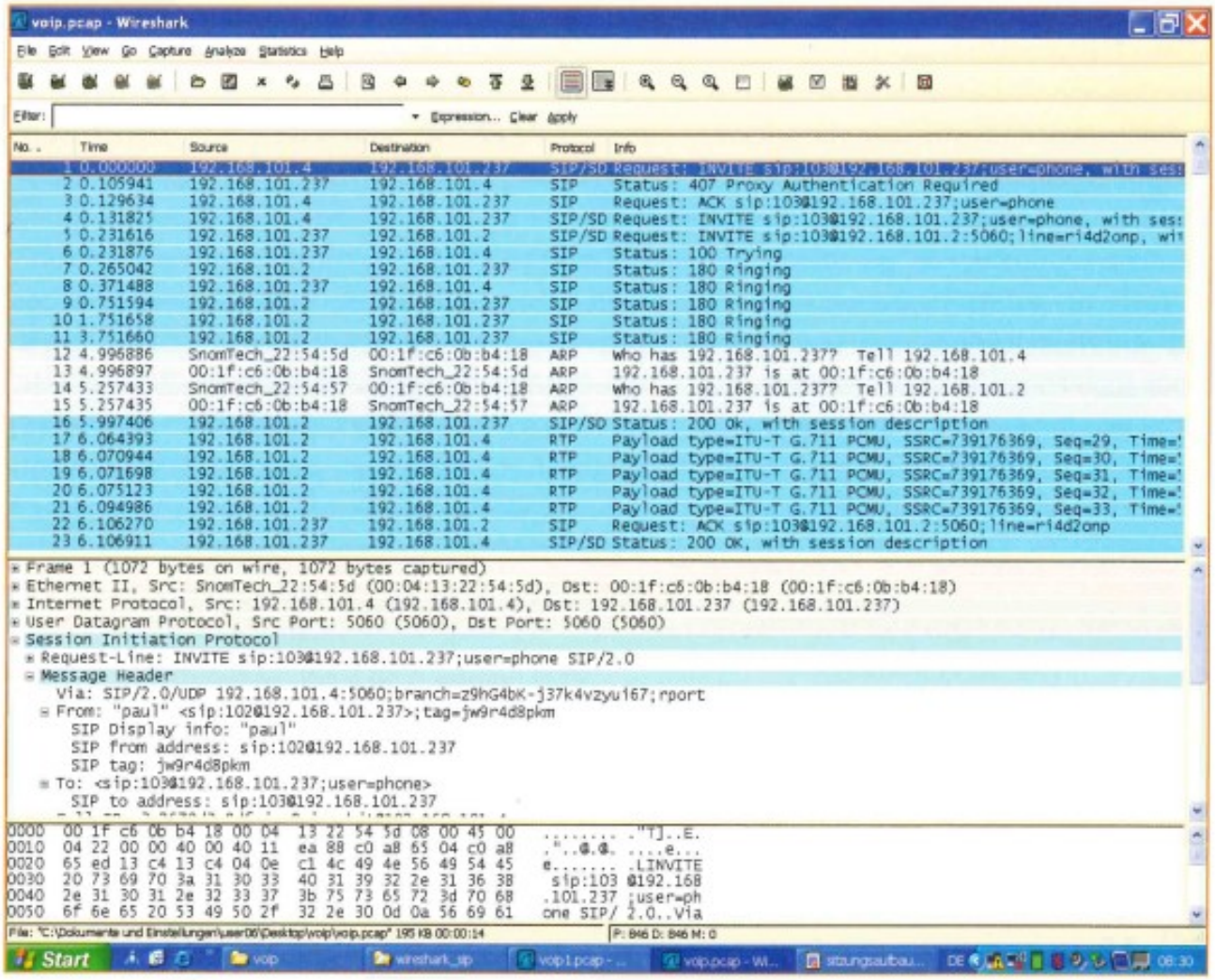

Bild 9: SIP-Kommunikation (Aufzeichnung mit Wireshark)

■ Struktur von SIP Adressen

SIP basiert auf Textnachrichten und verwendet URLs.

Die Adressen werden ähnlich wie die E-mail-Adressen aufgebaut (sind zu diesen aber nicht kompatibel), z. B.:

user@domain

user@ip-address

phone-number@domain

Allgemein: **userinfo@hostport[;url-parameter]**

(transportparameter | .....)

transport = udp | tcp

Beispiele: fritz@voip.de

fritz@voip.de; transport=udp

für die SIP-Nachrichtenübermittlung wird UDP verwendet

fritz@192.168.100.12

301@voip.de

Die SIP-Adresse identifiziert einen Teilnehmer in einer Internet-Domain.

1.5.2 Real-Time Transport Protocol (RTP)

Das Real-Time Transport Protocol (RTP) ist ein Protokoll zur kontinuierlichen Übertragung von audiovisuellen Daten (Streams) über IP-basierte Netzwerke. Das Protokoll wurde erstmals 1996 im RFC 1889 standardisiert. 2003 wurde ein überarbeiteter RFC veröffentlicht. Auf RTP basieren Produkte wie LiveMedia von Netscape oder NetMeeting von Microsoft, aber auch komplette Anwendungsprotokolle wie H.323 und SIP. Es wird in der IP-Telefonie zur Übertragung echtzeitsensitiver Datenströme verwendet.

Mit RTP wird der Medienstrom in Datenpakete aufgeteilt. Im Header der Datenpakete sind Informationen über Codec, Sequenznummer, Zeitstempel, Synchronisation und evt. Verschlüsselung (SRTP) enthalten.
RTP, das unabhängig von den darunterliegenden Protokollschichten ist, nutzt zum Transport der auftretenden Datenpakete das verbindungslose UDP. Es gewährleistet einen durchgängigen Transport von Daten in Echtzeit. Hier steht die schnelle und effiziente Übertragung der Daten im Vordergrund. (Gehen UDP-Daten im Netz verloren, werden sie nicht - wie beim verbindungsorientierten TCP - erneut angefordert.)

RTP (Real-Time-Transport Protocol)					
Familie:	Netzwerkprotokoll				
Einsatzgebiet:	Transport von Medien-Streams				
Port:	beliebiger freier, gerader Port größer 1024				
RTP im TCP/IP-Protokollstapel					
Anwendung	RTP				
Transport	UDP				
Internet	IP (IPv4, IPv6)				
Netzzugang	Ethernet	Token Bus	Token Ring	FDDI	...
Standard:	RFC 3550 (RTP: A Transport Protocol for Real-Time Applications, 2003)				

Bild 10: Real-Time-Transport-Protocol im OSI-Modell

1.6 Telefon Number Mapping (ENUM)

1.6.1 Was ist ENUM?

ENUM verknüpfen Telefonnummern mit dem Internet. Dabei ermöglicht ENUM unter einer Telefonnummer verschiedene Dienste wie Webseiten, e-Mail und anderen denkbaren Services verfügbar zu machen sowie Prioritäten für die einzelnen Dienste zu setzen.
Das Ziel von ENUM ist, verschiedene Adressen, Nummern und URLs unter einer einzigen Nummer zu vereinen. So können unter einer einzigen ENUM-Nummer das private Telefon zu Hause, das Telefon in der Firma, die Faxnummer, die Mobilfunknummern, geschäftliche und private e-Mail-Adressen, Videokonferenzadressen, die eigene Website und anderen Kommunikationsadressen zusammengefasst werden. Je nach gerade aktuell verwendeter Applikation (z. B. Telefon, e-Mail-Programm usw.) sucht sich diese unter der angegebenen ENUM-Nummer die eigentliche Zieladresse.

1.6.2 Wie funktioniert ENUM?

ENUM benutzt das DNS (Domain Name System), um die Telefonnummern auf die hinterlegten Webadressen bzw. URLs zu verknüpfen und aufzulösen.
Das Protokoll zu ENUM ist im Standard "RFC2916" (E.164 Number and DNS) definiert.
Um die Telefonnummern in das DNS-System zu überführen, wird eine Telefonnummer in eine eindeutige Domain verwandelt. Dazu wurde eine neue Domain namens "e164.arpa" eingeführt und definiert. Unter dieser Domain werden nun einfach Subdomains für jede Ziffer der bisherigen Telefonnummer eingeführt. Da das DNS-System aber rückwärts arbeitet (die Top-Level-Domain ist immer die ganz rechte, die Second-Level-Domain ist die zweite von rechts und die Subdomains sind dann die Weiteren nach links), wird die Telefonnummer auch rückwärts überführt.

Im Beispiel wird die Telefonnummer (+49 3328 3507-10) in das DNS-System überführt.

Zunächst werden alle Leerzeichen und sonstigen optischen Strukturierungshilfen wie Bindestriche usw. entfernt: 493328350710
Die Nummer wird umgedreht: aus 493328350710 wird damit 017053823394
Zwischen den Ziffern werden Punkte eingefügt: 0.1.7.0.5.3.8.2.3.3.9.4.
Die ENUM-Domain "e164.arpa" wird angehängt (die Top-Level-Domain "arpa", gefolgt von der Second-Level-Domain "e164") Eine Telefonnummer wird dann (einschließlich Landeskennzahl) in umgekehrter Reihenfolge geschrieben, wobei jede Ziffer durch Punkt getrennt wird und damit eine eigene Sub-Domain bildet. Da Telefonnummern weltweit eindeutig sind, gibt es auf diese Art und Weise auch zu jeder Telefonnummer eine eindeutige ENUM-Domain im DNS:

0.1.7.0.5.3.8.2.3.3.9.4.e164.arpa.

Aus der bisherigen Telefonnummer +49 3328 3507-10 als Ausgangsbasis ist im Ergebnis einen vollständigen Domainnamen (sog. "Fully Qualified Domain Name" = FQDN) 0.1.7.0.5.3.8.2.3.3.9.4.e164.arpa gebildet worden. Dieser vollständige Domainname kann ganz normal im Nameservicebetrieb verwendet werden.

Laut dem im RFC2915 definiertem ENUM-Protokoll werden im Nameserver spezielle Einträge verwendet um auf die einzelnen Kommunikationsadressen (Telefon, e-Mail, ...) zu verweisen. Verwendet werden dabei die sog. "Naming Authority Pointer Records" (NAPTR). Für jede Domain können mehrere dieser NAPTR-Einträge eingesetzt werden, eben für jede Kommunikationsadresse genau eine. Außerdem ist es möglich, Prioritäten (Preferenzen) zu definieren.
Wenn ein Anruf getätigt wird, wird über ENUM also zukünftig zunächst über das DNS-System die Nummer bis zum NAPTR-Eintrag aufgelöst, welcher dann auf eine Nummer eines Telefons, eines Handys, eines Fax-Anschlusses usw. zeigt bzw. diese als Ergebnis liefert (welche dann entsprechend angewählt werden kann).