

Session Initiation Protocol

**Funktionsweise, Einsatzszenarien, Vorteile und
Defizite**

von

**Dipl. Inform. Petra Borowka
Markus Schaub**

5 SIP: Erweiterungen und Mehrwertdienste

In den vorhergehenden Kapiteln wurden die Basisfunktionen von SIP dargestellt und Designmöglichkeiten für hochverfügbare Lösungen diskutiert. Damit allein wäre es kaum möglich klassische Telefonanlagen abzulösen.

Gegner führen gerne an, dass mit VoIP im Allgemeinen und SIP im Speziellen nicht alle Leistungsmerkmale abgedeckt werden könnten, die es bisher gab. Befürworter bestreiten das. Umgekehrt führen die Verfechter von VoIP an, dass durch die Verschmelzung der IT- und der TK-Welten ein erhebliches Potential zur Schaffung von Mehrwertapplikationen entstehen. Die Kritiker halten dagegen, dass das mit CTI schon lange möglich war.

Sicher ist, eine VoIP-TK-Lösung muss mindestens die Anforderungen erfüllen, die die bisherige PBX erbracht hat. Im Folgenden wird darum zunächst gezeigt, wie sich mit SIP klassische Leistungsmerkmale umsetzen lassen. In den weiteren Unterkapiteln werden die Standarderweiterungen, existierende Mehrwertdienste und typische TK-Anwendungen vorgestellt und wie sie sich mit SIP realisieren lassen. Dabei wird darauf verzichtet, die mittlerweile über 100 Standards rund um SIP detailliert vorzustellen. Auch wird nicht auf jede Erweiterung und Anwendung eingegangen, die es mittlerweile gibt. Stattdessen haben die Autoren eine Auswahl getroffen, von der sie meinen, dass sie für die Praxis besonders relevant sind. Das Augenmerk liegt dabei auf der grundsätzlichen Funktionsbeschreibung, der Anwendbarkeit und der Umsetzung in der Praxis. Zusätzlich werden die vorgestellten Funktionen und Mehrwertdienste wenn nötig bewertet und auf Stolpersteine im Praxiseinsatz hingewiesen.

Ein separates Unterkapitel ist den Schnittstellen gewidmet, da diese der Dreh- und Angelpunkt für das Zusammenwachsen von IT- und TK-Anwendungen ist.

5.1 Weiterführende Standards

Die Basis-Standards von SIP spezifizieren eine Reihe von Meldungen und Verfahren, die notwendig sind um eine Verbindung aufzubauen, zu beenden und ggf. die Parameter zu ändern. Sie definieren also, wie eine Verbindung gemanagt werden kann (Verfahren) und was dafür notwendig ist (Protokoll). Was sie jedoch nur in wenigen Ausnahmen abdecken ist das, was in der TK-Welt als Leistungsmerkmale bezeichnet wird.

Das heißt aber umgekehrt nicht, dass die klassischen Leistungsmerkmale mit den Mitteln von SIP nicht abgedeckt werden könnten. Zum Teil muss das SIP- oder das SDP-Protokoll dazu ergänzt werden. Jedoch ist es nicht notwendig, für jedes Leistungsmerkmal eigene Protokollparameter einzuführen.

Es ist gute Tradition bei der IETF Protokolle so generisch anzulegen, dass sie unterschiedlichste Aufgaben erfüllen können, ohne dafür modifiziert werden zu müssen. Es liegt dann an der Implementierung, was für Leistungsmerkmale

möglich sind. Wie mit Standardbordmitteln von SIP Leistungsmerkmale realisiert werden können, wird später detailliert am Beispiel der „Umleitung bei Besetzt“ vorgestellt.

Der Vorteil dieses Vorgehens ist es, dass Anwendungen sehr modular programmiert werden können: bereits geschriebene Subroutinen können immer wieder verwandt werden, so reicht es beispielsweise eine Routine für das SIP INVITE nur einmal zu programmieren. Diese kann dann sowohl für einen Gesprächsaufbau wie auch für eine Konferenzschaltung genutzt werden. Ein weiterer Vorteil ist, dass den Herstellern so viele Freiheiten gelassen werden, standardkonforme Leistungsmerkmale zu entwickeln, ohne dass dafür der oft langwierige Prozess der Standardisierung abgewartet werden muss.

Der Nachteil ist jedoch, dass viele Leistungsmerkmale zwar keine neuen SIP-Parameter oder -Pakete benötigen jedoch ein gemeinsames Vorgehen der Kommunikationspartner. Insbesondere bei Produkten unterschiedlicher Hersteller können sich sonst schnell Inkompatibilitäten ergeben. Ein Beispiel dafür wäre das Weiterleiten eines Telefonates: alle drei Clients müssen in der Abfolge und im Inhalt der SIP-Pakete übereinstimmen, wenn die Weiterleitung funktionieren soll.

Das Dilemma zwischen möglichst vielen Freiheiten bei der Implementierung und der Interoperabilität wird bei der IETF durch die BCPs gelöst. In ihnen werden erprobte Verfahren vorgestellt, ein Problem mittels bestehender Standards zu lösen.

Die Aufteilung zwischen Vorgehensweisen und Protokoll schlägt sich auch in der Teilung der Arbeiten in verschiedene SIP-Arbeitsgruppen nieder.

Im Folgenden werden die relevanten Arbeitsgruppen kurz beschrieben und anschließend praxisrelevante Standarderweiterungen vorgestellt.

5.1.1 Arbeitsgruppen

5.1.1.1 SIP-Arbeitsgruppe

Die SIP-Arbeitsgruppe ist für die Entwicklung der Kernelemente von SIP zuständig. D.h. sie standardisiert alle Dienste, Protokolle, Erweiterungen und Vorgehensweisen, die zwingend vorgeschrieben sein müssen, damit ein Signalingprotokoll interoperabel funktionieren kann.

Sie ist erklärter Weise nicht für die Entwicklung von Anwendungen auf der Basis dieser Signalisierungen zuständig. Tauchen jedoch in anderen Arbeitsgruppen Anforderungen an das SIP-Protokoll auf, so werden diese umgesetzt. Primär kommen diese Anforderungen von den Arbeitsgruppen SIPPING, SIMPLE und XCON.

Die vier Dogmen der Arbeitsgruppe lauten:

1. Dienste und Features werden wenn immer möglich Ende-zu-Ende realisiert.
2. Lösungen müssen generisch angelegt sein und dürfen sich nicht an speziellen Anwendungen ausrichten.
3. Jede Lösung muss so einfach wie möglich sein.
4. Nutzung von und Interoperabilität mit bestehenden Internetprotokollen hat oberste Priorität.

5.1.1.2 SIPPING-Arbeitsgruppe

Die SIPPING-Gruppe beschäftigt sich mit der praktischen Umsetzung des SIP-Standards, dazu hat sie sich fünf Ziele auf die Fahne geschrieben:

1. Evaluierung und Dokumentation von Anforderungen des Marktes
2. Dokumentation der SIP-Nutzung, um Anforderungen (Leistungsmerkmale) standardisiert zu implementieren
3. Beschreibung von Anforderungen notwendiger SIP-Erweiterungen
4. Zusammenführung der Gemeinsamkeiten verschiedener Task Groups
5. Beschreibung der Anforderungen und notwendigen Prozeduren für Entwicklung von Benutzerprofile

5.1.1.3 SIMPLE-Arbeitsgruppe

Die beiden Themen der dritten SIP-Arbeitsgruppe sind Instant Messaging und Präsenz. Wie die SIPPING-Gruppe spezifiziert sie selbst keine neuen Protokolle oder Protokollerweiterungen, bereitet sie im Bedarfsfall jedoch vor und übergibt sie an die SIP-WG. Die drei primären Ziele der Arbeitsgruppe sind:

1. Entwicklung eines Verfahrens, das SIP als Transportprotokoll für IM nutzt.
2. Weiterentwicklung von SIP, um Präsenz-Informationen auszutauschen.
3. Beschreibung einer Architektur, die es erlaubt, IM und Präsenz mittels Buddy-Listen zu implementieren.

5.1.1.4 P2PSIP-Arbeitsgruppe

Die jüngste Arbeitsgruppe zum Thema SIP hat sich im Februar 2007 konstituiert und beschäftigt sich mit SIP als Peer-to-Peer-Anwendung. Da es in einem P2P-Netzwerk keine „ausgezeichneten“ Geräte gibt, die a priori eine besondere Bedeutung besitzen, wie SIP-Server, muss jedes teilnehmende Gerät Dienste und Ressourcen zur Verfügung stellen, die für gewöhnlich von Servern erfüllt werden.

Erklärtes Ziel der Arbeitsgruppe ist es, eine möglichst einfache Technik für solche P2P Netze auf der Basis von SIP zu entwickeln, die zudem mit minimalen Konfigurationseinstellungen an den Clients auskommt.

Nähere Informationen zum Thema P2PSIP gibt es im Insider Artikel „P2P SIP – Das Ende von Skype?“ von Markus Schaub im Netzwerk Insider vom Mai 2006.

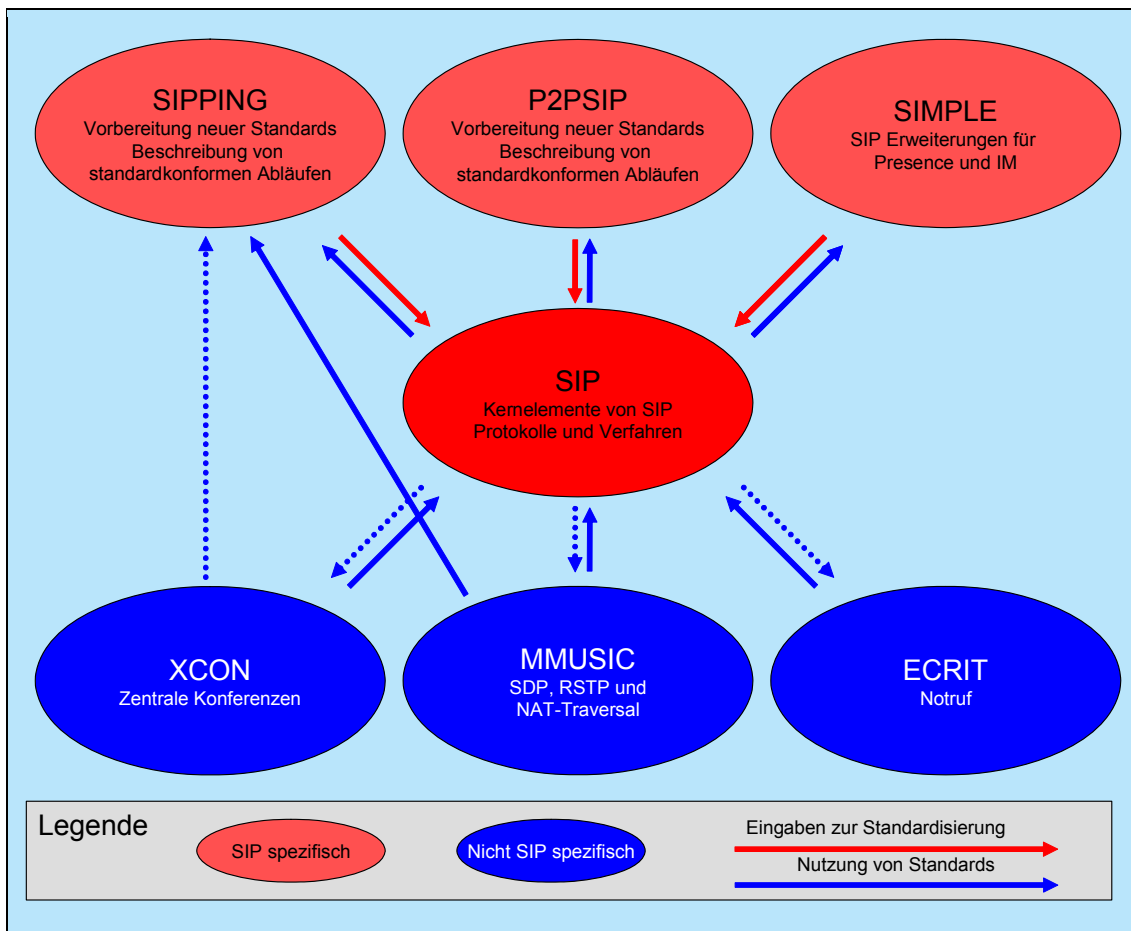


Abbildung 5.1: Verflechtung der Arbeitsgruppen

5.1.1.5 Nicht-SIP-spezifische VoIP-Arbeitsgruppen

Insgesamt beschäftigen sich aktuell 15 Arbeitsgruppen bei der IETF mit Themen Echtzeit-Anwendungen. Davon haben nur die vier bisher vorgestellten SIP als Schwerpunkt. Allerdings sind auch andere Gruppen für die SIP-Implementierung relevant. Die drei wichtigsten Gruppen sind:

1. XCON-Arbeitsgruppe

XCON steht für Centralized Conferencing. Konferenzen stellen zwar auch Anforderungen an die Signalisierung, jedoch liegen die besonderen Schwierigkeiten in anderen Bereichen, wie Rollendefinitionen (Besitzer, Moderator, Sprecher, etc.), Authentifizierung und Verschlüsselung. Die Umsetzung dieser Anforderungen ist nur zum Teil Aufgabe des Signalisierungsprotokolls.

Die XCON-Arbeitsgruppe beschäftigt sich ausschließlich mit einem zentralistischen Konferenzmodell, bei dem alle Teilnehmer mit einem Fokus und

einem Mixer verbunden sind, die die Konferenz steuern (Fokus) und die Datenströme mixen und verteilen (Mixer).

2. MMUSIC-Arbeitsgruppe

Schwerpunkt der Multiparty Multimedia Session Control (MMUSIC) Gruppe ist heute die Weiterentwicklung von SDP. Neben SDP werden von der Gruppe aktuell nur zwei weitere Themen vorangetrieben: RSTPv2 und NAT-Traversal.

3. ECRIT-Arbeitsgruppe

Die ECRIT-Gruppe beschäftigt sich mit dem Thema Notruf. Die Schwierigkeit von VoIP bzgl. Notrufen liegt darin, dass im Unterschied zur klassischen, kabelgebundenen Telefonie IP-Geräte nicht ohne weiteres lokalisiert werden können.

Ziel der Arbeitsgruppe ist es, so weit als möglich bestehende Techniken zu nutzen, um den Aufenthaltsort eines Anrufers bei einem Notfallanruf bestimmen zu können.

5.1.2 DTMF-Töne

DTMF steht für Dual Tone Multiple Frequency, im Deutschen auch MFV, Mehrfrequenzwahlverfahren, genannt. MFV löste mit dem Einzug digitaler Vermittlungsstellen das ältere Impulswahlverfahren, IWW, ab. Beim IWW wurden Nummern durch elektrische Impulse übertragen, die durch die Wählscheibe erzeugt werden konnten.

	1209 Hz	1336 Hz	1477 Hz	1633 Hz
697 Hz	1	2	3	A
770 Hz	4	5	6	B
852 Hz	7	8	9	C
941 Hz	*	0	#	D

Tabelle 5.1 DTMF: Zusammenhang Tastenbelegung¹ und Frequenzen

Für DTMF hingegen ist ein Nummernblock notwendig. Den Zeilen und Spalten werden feste Frequenzen zugeordnet und eine Ziffer wird durch die Überlagerung der Zeilen- und Spaltenfrequenz bei der Übertragung repräsentiert. Bei der Festlegung der Frequenzen hat man für die Zeilen niedrigere und für die Spalten höhere Frequenzen gewählt. Ferner hat man versucht, die Frequenzen so festzulegen, dass sie bei der Überlagerung eine Disharmonie erzeugen, um

¹ Die Tasten A-D finden in Deutschland keine Anwendung. Sie werden bspw. zur Priorisierung von Anrufen bei militärischen Anwendungen genutzt.

zu verhindern, dass zufällige Hintergrundgeräusche mit denselben Harmonien zu unvorhergesehenen Störungen führen.

Neben der Anwahl werden DTMF-Töne heute auch zur interaktiven Steuerung und Eingabe während eines Gespräches genutzt, um beispielsweise Anrufbeantworter abzufragen, Passwörter, Konto- oder Kundennummern einzugeben.

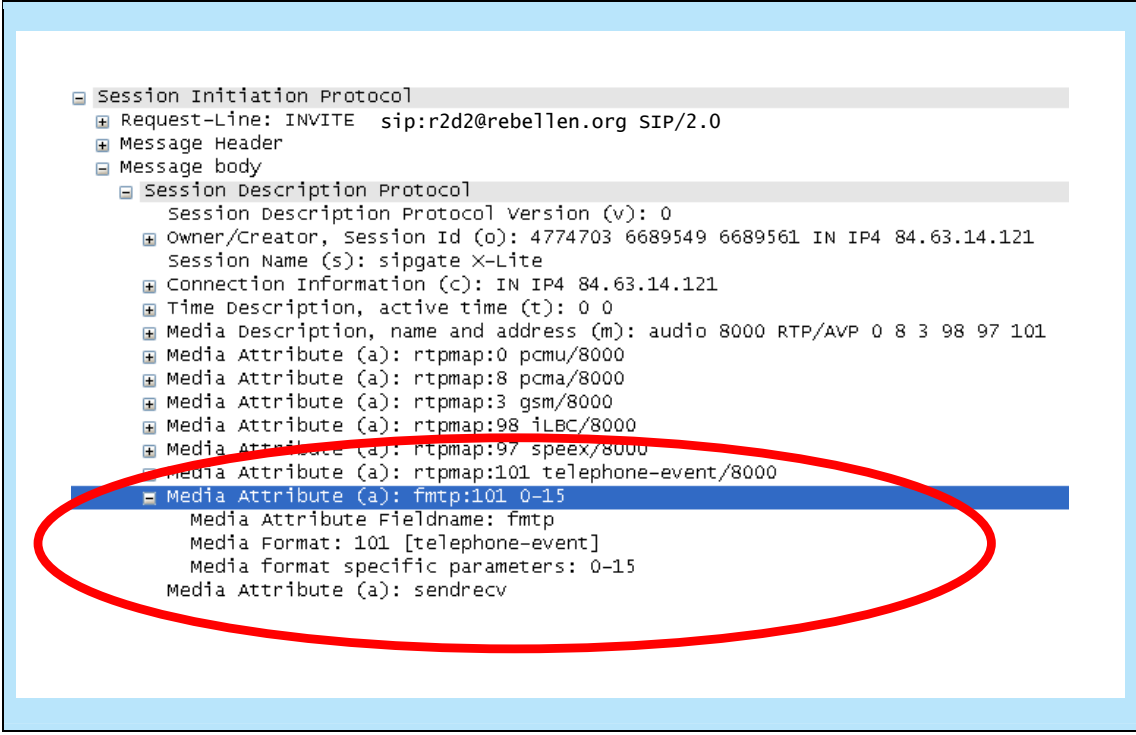
Für die Signalisierung von Telefonnummern und die Steuerung/Konfiguration von Telefonanlagen, wofür DTMF ursprünglich entwickelt wurde, gibt es bei SIP keine Anwendung mehr. Anders verhält es sich jedoch für die Eingabe von Nummern nach dem Verbindungsaufbau. In einer reinen SIP-Welt könnte grundsätzlich auf DTMF verzichtet werden und alle Befehle und Eingaben über entsprechende SIP-Nachrichten übertragen. Liegt zwischen Sender und Empfänger jedoch ein PSTN-Netz, so kann Stand heute nicht auf die Funktion verzichtet werden.

Für die Übertragung bieten sich zwei Varianten an:

1. DTMF wird als Ton kodiert und über RTP übertragen.
2. Die Information, dass und welche Taste gedrückt wurde, wird vom SIP-Client als SIP- oder RTP-Nachricht gesendet.

5.1.2.1 DTMF über RTP/G.711

Die erste Variante scheint den Status Quo zu erhalten: das Endgerät generiert bei Tastendruck einen DTMF-Ton, kodiert ihn gemäß dem genutzten Codec und sendet ihn anschließend als Voice-Daten über RTP.



```
Session Initiation Protocol
  Request-Line: INVITE sip:r2d2@rebellen.org SIP/2.0
  Message Header
  Message body
    Session Description Protocol
      Session Description Protocol Version (v): 0
      Owner/Creator, Session Id (o): 4774703 6689549 6689561 IN IP4 84.63.14.121
      Session Name (s): sipgate X-Lite
      Connection Information (c): IN IP4 84.63.14.121
      Time Description, active time (t): 0 0
      Media Description, name and address (m): audio 8000 RTP/AVP 0 8 3 98 97 101
      Media Attribute (a): rtpmap:0 pcmu/8000
      Media Attribute (a): rtpmap:8 pcma/8000
      Media Attribute (a): rtpmap:3 gsm/8000
      Media Attribute (a): rtpmap:98 iLBC/8000
      Media Attribute (a): rtpmap:97 speex/8000
      Media Attribute (a): rtpmap:101 telephone-event/8000
      Media Attribute (a): fmp:101 0-15
        Media Attribute Fieldname: fmp
        Media Format: 101 [telephone-event]
        Media format specific parameters: 0-15
        Media Attribute (a): sendrecv
```

Abbildung 5.2: Signalisierung der Inband-DTMF-Fähigkeit

So einfach das auch zunächst klingt ist es nicht. Da die Töne bewusst eine Dissonanz erzeugen, kann insbesondere eine komprimierender Sprachcodec dazu führen, dass der Empfänger die sich überlagernden Frequenzen nicht mehr voneinander trennen kann und somit auch nicht die Taste, die gedrückt wurde.

Um dieses Problem zu lösen, definiert RFC 4733 drei verschiedene Möglichkeiten, RTP zur DTMF-Übertragung zu nutzen:

1. Wechsel zu einem nicht-komprimierenden Codec wie G.711.
2. DTMF-Töne werden als „telephone-events“ über RTP gesendet (vgl. Abbildung 5.2).

Dazu werden beim Verbindungsaufbau über SDP die notwendigen Parameter ausgehandelt. Wird nun während des Gesprächs eine Taste gedrückt, so wird zwar derselbe Datenstrom genutzt, wie für das Gespräch selbst, jedoch zeigt der RTP-Paket Header ein „Event“ an und im Datenteil steht anstatt eines kodierten Tonsignals, welche Taste gedrückt wurde.

Ein Gateway zum PSTN muss diese Events im RTP unterstützen und als DTMF-Töne in das PSTN senden.

Vorteil dieses Verfahrens ist es, dass die gesendete Information minimal ist und somit kaum Bandbreite benötigt. Nachteil ist, dass alle möglichen Events bei allen Geräten gleich definiert sein müssen.

3. DTMF-Töne werden als RTP-Typ „tone“ übertragen.

Wie bei den „events“ handeln die Gesprächspartner über SDP aus, dass neben einem Auto-Profil ein weiteres „tone“-Profil von RTP genutzt werden kann. Bei diesem Verfahren wird nicht übermittelt, welche Taste gedrückt wurde, sondern aus welchen Frequenzen sich ein Ton zusammensetzt. Drückt man nun eine Taste während des Telefonates, so wird der DTMF-Tone als RTP-„tone“-Typ im bestehenden RTP Datenstrom gesendet.

Der Vorteil dieses Verfahrens ist, dass eine Vielzahl von „Tönen“ übertragen werden können, ohne dass sie im Vorfeld Sender und Empfänger bekannt sein müssen. Die Entwickler haben sich bei der Codierung der Frequenzen an den international vorkommenden Tönen orientiert, so dass dieses Verfahren auch genutzt werden kann, um ausländische Frei- und Besetzzeichen zu übermitteln.

Der Nachteil ist, dass mehr Daten übertragen werden, als beim „event“.

Für die Umsetzung von DTMF ist das zweite Verfahren hervorragend geeignet. Die dritte Variante ist hingegen offener für weitere Anwendungen.

5.1.2.2 DTMF über SIP

Anstatt RTP kann auch SIP genutzt werden, um einen Tastendruck am Telefon zu signalisieren. Dazu wird das in Kapitel 2.4 vorgestellte SIP-INFO-Paket aus