DISSERTATION

Connection Control Signaling in next generation satellite network using Voice over IP Protocols

ausgeführt zum Zwecke der Erlangung des akademischen Grades eines Doktors der technischen Wissenschaften unter der Leitung von

O.Univ.Prof. Dipl.Ing. Dr.techn. Richard EIER
Institut für Computertechnik (E384)

ICT

eingereicht an der Technischen Universität Wien
Fakultät für Technische Naturwissenschaften und Informatik
von

Dipl.-Ing. Edith DUSCH
Matr. Nr.: 9525813
Leibnizgasse 55/16
A-1100 WIEN

Wien, im Mai 2004

DI Edith Dusch

Danksagung

An erster Stelle möchte ich meinem Betreuer Herrn o. Prof. Dr. Richard Eier für die Möglichkeit danken, neben meiner beruflichen Tätigkeit diese Dissertation fertig-zustellen. Ebenfalls für die genaue Durchsicht meiner Ausarbeitungen und für viele hilfreiche Anmerkungen, welche die Lesbarkeit der Arbeit deutlich verbessert haben.

Auch meinem zweiten Betreuer Herrn Prof. Dr. Kurt P. Judmann, möchte ich danken, ich habe aus seinem Interesse am Thema und aus den lebhaften Diskussionen viele wertvolle Anregungen erhalten.

Mein besonderer Dank gilt meinem Ehemann Alexander, der trotz vieler Stunden, in denen er mich entbehren musste, immer hinter mir gestanden hat. Ihm möchte ich diese Dissertation widmen.

Kurzfassung der Dissertation

Der Zweck dieser Arbeit ist die Untersuchung der Eignung von Voice over IP (VoIP) Protokollen und Technologien für die Signalisierung einer Verbindung in einem Satelliten- Netzwerk, ebenfalls die nächste Generation von Satelliten-Netzen genannt, die multimedialen Satelliten-Netze.

Für diese Untersuchung wurde ein konkretes Satelliten-Netz ausgewählt: das EuroSkyWay-Programm, welches von der Europäischen Weltraumbehörde ESA (European Space Agency) entwickelt wurde. Die potentiellen VoIP-Protokolle, ausgewählt für die Untersuchung, sind H.323, SIP und Megaco (H.248).

In der ersten Generation von multimedialen Satelliten-Netzen wurde das B-ISDN-Protokoll für die Signalisierung einer Verbindung verwendet. Der Grund für diese Entwicklung war die Tatsache, dass B-ISDN weit entwickelt, voll etabliert, gut getestet und sich in der terrestrischen Kommunikation durchgesetzt hat. Das B-ISDN Protokoll bietet eine umfangreiche Funktionalität und hohe Komplexität, welche sich in der terrestrischen Kommunikation voll bewährt hat, aber nicht notwendigerweise in Satelliten-Netzen benötigt wird. In einem Satelliten-Netz wäre für die Signalisierung einer Verbindung ein viel einfacheres Protokoll ausreichend. Derzeit, während das Internet dabei ist sich voll und ganz zu entwickeln und durchzusetzten, sind neue Protokolle entstanden, welche in der Audiokommunikation über das Internet verwendet werden und somit als Kandidaten für die Satelliten-Netzwerke eingesetzt werden könnten.

Bei der Untersuchung wurde eine komparative Methode eingesetzt um die einzelnen Kandidaten (H.323, SIP und H.248) zu vergleichen und zu analysieren. Dabei wurde untersucht und verglichen welches der drei Protokollen sich am besten für die Signalisierung in einem Satelliten-Netzwerk eignet. Für den Vergleich wurden Kriterien erarbeitet, welche für ein Satelliten-Netz von Bedeutung sind. Die Kriterien sind gegliedert in Anwendbarkeit, Komplexität, Performance, Erweiterbarkeit, Ressourcen und Bandbreiten-Nutzung und weiters unterteilt in Unterkriterien, durch die eine Evaluierung für Satelliten-Zwecke erst möglich ist.

Das Ergebnis der Analyse brachte das am besten geeignetste Protokoll hervor: das SIP Protokoll. Allerdings bevor noch SIP in einem Satelliten-Netzwerk für die Verbindungssignalisierung eingesetzt werden kann, müssen einige Anpassungen vorgenommen werden, um ein maßgeschneidertes Protokoll zu bekommen. Zwei Arten von Anpassungen sind notwendig: einerseits eine Erweiterung des Protokolls und andererseits eine Vereinfachung (Reduzierung) des Protokolls. Die Erweiterung ist relevant und notwendig um die satelliten-spezifischen Anforderungen in das SIP-Protokoll zu integrieren. Die Vereinfachung erlaubt, die für Satelliten-Netze, nicht verwendete Funktionalität des SIP-Protokolls zu entfernen. Diese notwendigen Anpassungen für die Signalisierung einer Verbindung im Satelliten-Netzwerk sind ebenfalls ein wichtiger Aspekt dieser Arbeit.

Abstract

The purpose of this thesis is to introduce the next generation satellite network also called multimedia satellite network and to investigate the applicability of Voice over IP (VOIP) protocols and technologies for connection control signaling in a satellite environment.

For this analysis a concrete satellite system is chosen, the EuroSkyWay Program developed by ESA (European Space Agency). The potential VoIP protocols chosen for the investigation are H.323, SIP and Megaco (H.248).

In the first generation of multimedia satellite networks the B-ISDN protocol was used for call control signaling. The reason for this choice was that B-ISDN was well developed, established and proven in terrestrial voice communication. The powerful B-ISDN protocol supports the high complexity and functionality which is needed in the terrestrial network. However, for satellite purposes a simpler protocol for connection control signaling would be sufficient. Also, the Internet revolution has resulted in a number of new protocols which are used in voice communication over the Internet and these are therefore potential candidates for the satellite network for connection control signaling.

This analysis performs a comparative analysis between the three candidate VoIP call control protocols (SIP, H.323 and Megaco). The analysis examines and compares the suitability of the VoIP protocols and outlines the most suitable protocol for connection control signaling in a multimedia satellite network environment. The following criteria are defined relating to the requirements of a satellite network: Applicability, Complexity, Performance, Extensibility, and Resource and Bandwidth Efficiency. Each of them contains well-defined sub-criteria, which allow the evaluation of the protocols for satellite purposes.

This analysis finds that the most suitable protocol is SIP, but before the usage of SIP for satellite networks is ensured, some modifications and adaptations are necessary. Two aspects are quite simple: extensions and simplification. Extension is foreseeable: that some satellite-specific extensions to the protocol standard will be necessary e.g. due to the fact that the chosen protocol does not support the whole functionality required for connection control in a multimedia satellite network. Simplification allows the usage of a subset of the standard functionality needed. The required adaptation for satellite purposes of the chosen protocol is also a relevant aspect of this thesis.

Table of Contents

1. Introduction

1.1. Purpose of this thesis

The purpose of this thesis is to introduce the next generation satellite network also called multimedia satellite network and to investigate the applicability of Voice over IP protocols and technologies for connection control signaling in a satellite environment. For this analysis a concrete satellite system is chosen, called EuroSkyWay Program developed by ESA (European Space Agency). In this thesis the term Multimedia Satellite Network is used for such satellite systems.

This thesis is originated from a research work in company "Siemens AG Österreich" on behalf on the European Space Agency. The research work was handled as a project with four project members divided in following four fields of responsibility: the scientific part of the project, project management, quality assurance and the implementation of the prototype. The scientific part of the project was my task, which included investigations, analysis and the responsibility of the project documentation. The research work outlined a new possible standard for multimedia satellite network purposes, additionally a patent [1] has been proposed with Austrian Patent Office. Although this thesis only introduces and outlines the research work, the investigations, the analysis and the results. The possible standard and the patent are for this thesis not a relevant aspect and therefore they are not included in this work.
The content of this thesis emerge through the research work and through the investigation. One exception is the description of the VoIP protocols in Annex B, Annex C and Annex D, which are originated from the respective standard.

This thesis is organized as follows.

- Section **1.2 "The next generation satellite system"** introduces shortly the next generation satellite network. Additionally detailed information about satellite system is given in **Annex A. Annex A: "Multimedia Satellite Network"** gives a short overview of the EuroSkyWay program.
- Section **1.3 "Overview of VoIP Protocols", Annex B, Annex C** and **Annex D** present the protocols used for voice communication over internet: H.323, SIP and H.248/Megaco, which are the potential candidate for connection control signaling a satellite network. These descriptions of given protocols is originated from the particular protocol standard.
- Section **1.4 "Requirements of the Multimedia Satellite Network"** points out the requirements of a multimedia satellite network. These requirements determine the criteria's for the comparative analysis.
- Section **2 "Analysis of VoIP Protocols for Satellite Network"** analyses upon following defined criteria: Applicability, Complexity, Performance, Extensibility, Resource and Bandwidth Efficient the suitability of the VoIP protocols for multimedia satellite network. The conclusion of this section presents the must suitable protocol for connection control signaling in a satellite environment.

- Section 3 **"Adaptation of the VoIP Protocol for Satellite Network"** discusses which modification and adaptation of the chosen protocol is needed to fulfill the whole requirements of a satellite network.
- Section 4 **"Summary and Conclusion"** finishes this study according a short summary and conclusion.
- At the end section **"References"** holds the list of references, section **"List of Acronyms"** lists the acronyms, section **"List of Figures"** lists the figures and section **"List of Tables"** lists the tables used in this thesis.

1.2. The next generation satellite network

The first usage of the satellite systems began in 1957 through the initiation of the U.S. and the Soviet Union, primarily for navigation and espionage purposes. Afterwards the usage of the conventional satellite systems was to provide television access through satellite dishes. Today satellites are an integrated part of the global communications system and serve one or more functions, such as communication, navigation, weather and environmental purposes.

The next generation of the satellite network, also called multimedia satellite network, should provide different services for private, public and business markets. The key services for broadband satellite systems are:

- Internet/Intranet connectivity for both private consumers and all categories of business users
- Interactive TV
- Point to point connections (for example: between two branch offices)
- Shared connectivity, short term links on demand for transaction based data (for example: to transmit video conferences or to transmit the Olympic games)
- Telecommunications transport (carrying telecommunications on a wholesale basis)
- Fixed telephony (basic services in underdeveloped regions)

These are only some of the possible application for the satellite systems, which of them will be used is a question of the technical, social and financial development. Industrial analyses believe that only a few of the services will survive and be established, but this may not be the case. Wherever terrestrial broadband access is available, satellite solutions will probably not be used, but everywhere else where terrestrial broadband is not available, the satellite solution will be chosen. The majority of the world, except the large concentration of population in the developed world, will not have a terrestrial broadband access available in the near future.

One example for the next generation satellite system, also called multimedia satellite network, is the EuroSkyWay Program developed by ESA (European Space Agency). This program is chosen as a multimedia satellite network respectively next generation satellite system for this thesis. For more detailed information about this program see Annex A.

1.3. Overview of VoIP Protocols

In the first generation of multimedia satellite network B-ISDN protocol was used for call control signaling. The reason for this chose was that B-ISDN was well developed, established and proven in the terrestrial voice communication. The powerful B-ISDN protocol support high complexity and functionality, which is needed in the terrestrial network, but not adequate necessary in a satellite environment. For example in the terrestrial network the topology contains a complex hierarchical structure, during the satellite network involve a simple structure, where only one central control element controls all terminals. In satellite environment the routing functionality is not needed, whereas the terrestrial network requires the routing functionality. Therefore for satellite purposes a most simple protocol for call control signaling would be sufficient.

In meantime, the Internet revolution has outlined a number of new protocols, which are used in the voice communication over Internet and therefore which are potential candidate for the satellite network for call control signaling. This subsection introduces shortly these potential "Voice over Internet Protocol" (VoIP) protocols. The protocols are H.323, SIP and Megaco. Annex B, Annex C and Annex D outline the main functionality and feature of these protocols.

1.3.1. H.323

The H.323 standard is issued by the Telecommunication Standardization Sector of ITU-T Study Group 16, named "H.323, Packet-based multimedia communications systems", ITU-T Recommendation H.323 [44]. The H.323 standard is part of the H.32x family of recommendations specified by ITU-T. H.323 is an umbrella recommendation, which sets standards for multimedia communications over Local Area Networks (LANs). H.323 is a standard which specifies components, protocols and procedures to provide multimedia communication services such as real-time audio, video, and data communications over packet networks, including Internet protocol (IP)-based networks.

The packet-based network over which H.323 entities communicate may be a point-to-point connection, a single network segment, or a network having multiple segments with complex topologies. H.323 entities may be used in point-to-point, multipoint, or broadcast configurations. H.323 entities may be integrated into personal computers or implemented in stand-alone devices such as video telephones. H.323 can be applied in a variety of mechanisms - audio only (IP telephony); audio and video (video telephony); audio and data; and audio, video and data.

For more details about H.323 see Annex B.

1.3.2. SIP

The Session Initiation Protocol (SIP) [40] is under development within the Internet Engineering Task Force (IETF) Multiparty Multimedia Session Control (MMUSIC) working group and published as RFC 3261: The Session Initiation Protocol (SIP), which obsoletes RFC 2543 in June 2002. They have based SIP on some other protocols coming out of the IETF, in particular the Simple Mail Transfer Protocol (SMTP) [36] and Hyper Text Transfer Protocol (HTTP) [25]. Like these SIP is a textual protocol based on the client-server model. SIP reuses much of the syntax and semantics of HTTP, including its response code architecture[i].

SIP is an application-layer control (signaling) protocol designed for signaling and session management in a packet telephony network. Session management functionality provides the capability of creating, modifying and terminating sessions between two or more participants. The purpose of signaling is to carry call information across network boundaries.

SIP sessions include Internet multimedia conferences, Internet telephone calls, distance learning, multimedia distribution and similar applications. Members in a session can communicate via multicast or via a mesh of unicast relations, or a combination of these. SIP invitations are used to create sessions. They carry session descriptions which allow participants to agree on a set of compatible media types. SIP can invite both persons and "robots", such as a media storage service. Media and participants can be added to an existing session. SIP supports user mobility by proxying and redirecting requests to the user's current location. Users can register their current location. SIP is not tied to any particular conference control protocol. SIP is designed to be independent of the lower-layer transport protocol and can be extended with additional capabilities.

SIP foresees the separation of call establishment and call description. For the latter one, the Session Description Protocol (SDP) (RFC 2327) [31] is used. It is intended for describing multimedia sessions for the purposes of session announcement, session invitation, and other forms of multimedia session initiation.

For more details about SIP see Annex C.

1.3.3. H.248

MEGACO [39] was standardized by the IETF Megaco Working Group as RFC 3015 in close cooperation with the ITU-T Study Group 16. Additionally it was published as ITU-T Recommendation H.248 [51].

The Media Gateway Control Protocol (MGCP) [35] specified by the IETF in RFC 2705 is a predecessor of MEGACO, which has some market share, but will not be enhanced further.

Megaco defines the protocol used between elements of a physically decomposed multimedia gateway. Megaco is an acronym for MEdia GAteway COntrol Protocol. In this document both terms Megaco and H.248 are used equally. The H.248 Gateway Control Protocol is the emerging standard, which is intended for the decomposition of Voice over IP (VoIP) gateways into Signaling Gateway (SG), Media Gateway Controller (MGC) and Media Gateway (MG). It specifies the interface between the MGC and the MG. Megaco does not define how multiple gateways or gateway controllers communicate with each other. That is defined under the umbrella of the H.323 protocol. But in more general H.248 may be used to allow for scalable and distributed architectures. So it allows separating the session signaling

---

[i] The expression "response code architecture" is used to describe the set of messages, which are responses in the particular protocol. I case of SIP this expression is related to the response codes 1xx to 6xx.

functionality independently from the session media and resource handling functionality depending upon the customer requirements. Megaco ensures that the VoIP traffic travels between IP networks and traditional circuit switched networks. H.248 is an asymmetrical protocol, which is used between an intelligent master (MGC) and a dumb slave (MG).
For more details about H.248 see Annex D.

1.4. Requirements of the Multimedia Satellite Network

Alter the next generation satellite network has been shortly introduced in subsection 1.2 and also the candidate Voce over IP protocols has been named in subsection 1.3, the next step or actually the next question, before the investigation in subsection 2 can be started, is: what are the requirement of the multimedia satellite network, which has to be fulfilled by the candidate VoIP protocols? The answer of this question will influence and determinate the criteria's for the analysis in subsection 2.

A multimedia satellite network has two main requirements: first the connection control and second the resource management control. Both requirements are handled by the network element Network Control Center (NCC) and must work in close cooperation with each other.

The connection control has to support following services: establishment of different kind of connections, release the connection and change the connection participants. The connection can be a permanent or on-demand connection, a point-to-point or point-to-multipoint connection, a uni-directional or bi-directional connection, a real-time or non real-time connection.

The resource management control has to support a dynamic resource allocation, utilization and re-negotiation procedures. In a multimedia satellite network an important task is the QoS support. This means that such system has to provide a guaranteed throughput level or to guarantee that end-to-end latency will not exceed a specified level. This implies that the user of such systems must inform the resource management about the desire resource allocation and utilization through the QoS parameters. This includes the support of different QoS levels and bandwidth management procedure for resource distribution, which is assured by the resource management control. The resource management uses a connection admission control (CAC) algorithm to avoid overload network situation.

The significant advantage of the multimedia satellite network over the packet based network such as IP is the connection oriented payload traffic, which make possible the support of guaranteed QoS.

This analysis should outline a VoIP protocol, which can optimum fulfill the requirements given by the multimedia satellite network.

2. Analysis of VoIP Protocols for Satellite Network

2.1. Criteria's for the Analysis and Selection

The purpose of this section is to perform a comparative analysis between the three introduced VoIP call control protocols (SIP, H.323 and Megaco) described in subsection 1.3, Annex B, Annex C and Annex D. The analysis examines and compares through the following criteria the suitability of the VoIP protocols for connection control signaling in a multimedia satellite network environment. Additionally the analysis should outline the most suitable protocol. The following criteria's are defined due to the requirement of a satellite network: Applicability, Complexity, Performance, Extensibility, Resource and Bandwidth Efficient. Each of them contains well-defined sub-criteria's, which allow the valuation of the protocols for satellite purposes. The criteria's with included the sub-criteria's are not part of any standards but rather an own defined criteria's suitable for the required analysis.

These criteria's are chosen due to the described requirement of a satellite network in subsection 1.4. Whether these criteria's are also suitable for an analysis of a terrestrial network is not investigated.

The defined criteria's investigate the candidate protocols for connection control signaling in a satellite environment upon the described requirements. The usage of the connection and the payload type of the connection is out of scope of this thesis.

2.2. Applicability

The criteria applicability is defined via a set of sub-criteria which are listed here:

1. **Connectivity/Symmetry**: the supported connectivity can be ptp (point to point), ptmp (point to multipoint); from symmetry point of view uni- or bi-directional calls between two or more participants/terminals are possible.
2. **Communication Model**: the communication model of the protocol can be a peer-to-peer or a master/slave.
3. **Basic Services**: the basic service of the protocol is the establishment and release of ptp and ptmp connection; for ptmp connections an additional basic capability is to add a user to or drop a user from the connection.
4. **Additional Services**: additional services supported by the protocol include registration and authentication of the participant/terminal. The registration process serves as a localization of the participant/terminal, and the authentication process verifies, if the terminal can have legal access to the network and to the services it is asking for.
5. **QoS Parameter Transport**: flexibility of protocol for supporting the transfer (and/or negotiation/renegotiation) of QoS parameters, which are:

- Service Category: type of contract (business, residential),connection typology (tributary traffic, signaling) and different service classes
- connection configuration / symmetry (ptp uni, ptp bi, ptmp uni, ptmp bi)
- QoS parameters (maxCTD - cell transfer delay, peak-to-peak CDV - cell delay variation, CLR - cell loss ratio)
- Traffic parameters (PDR - peak data rate, MDR - mean data rate, MBS - max burst size, UF - utilization factor)

6. **Core Supplementary Services**: support for core supplementary services. (These are call forwarding, call hold, call transfer, conferencing, call back, call pick up and park, call offering)
7. **Networks Elements**: the possibility to make a functional mapping[ii] of the supported protocol components on the network elements needed in a multimedia satellite network architecture. The following key network elements are required: multimedia satellite endpoints (such as User-, Gateway- or, Provider terminals), a Network Operation Center (NOC) including Call Control functionality with CAC (Connection Admission Control), Localization Management (LM) and Authentication Management (AM).
8. **Adaptability**: the ease with which the candidate protocol satisfies differing system constraints and user needs, or more specifically how the functionality existing in candidate protocols (i.e. messages and procedures) can be adapted, simplified or extended in order to comply with the needs of future MM satellite systems.

2.2.1. Connectivity/Symmetry

2.2.1.1. H.323

H.323 supports the following four types of connectivity:

- *point-to-point connectivity (correspond to ptp)*: point-to-point connectivity for bi-directional calls between two H.323 endpoints (terminals) is supported. Ptp unidirectional connectivity is not supported.
- *multipoint connectivity (correspond to ptmp)*: bi-directional conference calls between three or more H.323 endpoints (terminals) are supported. Ptmp unidirectional connectivity is not supported.
- *broadcast connectivity*: unidirectional broadcast connectivity where there is one transmitter of media streams and many receivers is supported. No bi-directional transmission of control or media streams in broadcast case is possible.
- *broadcast panel connectivity*: broadcast panel connectivity is a combination of a multipoint connectivity and broadcast connectivity. In this conference, several terminals are engaged in a multipoint conference, while many other terminals are only receiving the media streams. There is bi-directional transmission between the terminals in the multipoint portion of the conference and no bi-directional, unidirectional transmission between them and the listening terminals.

[ii] The expression functional mapping is used to describe the mapping of the functionality of the single component. The mapping is more and more successful if the functionality of the components is similar or equal.

2.2.1.2. SIP

Members in a SIP connection/session can communicate:

- *via unicast (correspond to ptp)* relation: two party are established to one session.
- *via a mesh of unicast (correspond to ptp)* relations: two party are established to one session and one of the participants invite a third party to the session. In this case every participant has a signaling session with each other. This type of session is called fully meshed unicast conference.
- *via multicast (correspond to ptmp)* relation: if the fully meshed conference will transform to a multicast session by redirecting all the single sessions to a conference bridge.

Both, ptp and ptmp sessions, are bi-directional. Uni directional and broadcast sessions are not supported by SIP.

In a group invitation according to the way how a connection is established and how the parties are invited, the following two scenarios can be discriminated:

- reach-first variant: is unicast whereby each call reaches a single user individual and therefore reaches the first available individual from a group.
- reach-all variant: instead of calling sequential each callee party, a multicast invitation will be sent

2.2.1.3. Megaco

Megaco supports:

- *a bi-directional unicast communication (correspond to ptp)* between a Media Gateway Controller (MGC) and Media Gateway (MG).
- *a bi-directional multipoint communication (correspond to ptmp)* between three or more parties is possible, if the MG supports multipoint conferences.

Unidirectional and broadcast communication between MGC and MG is not foreseen.

2.2.2. Communication Model

2.2.2.1. H.323

H.323 uses a peer-to-peer communication model, but sometimes is it necessary to use master/slave model. The H.245 master-slave determination procedures are used to resolve conflicts between two endpoints which can both be the Multipoint Controller (MC) for a conference, or between two endpoints which are attempting to open a bidirectional channel. In this procedure, two endpoints exchange random numbers in the H.245 masterSlaveDetermination message, to determine the master and slave endpoints. H.323 endpoints shall be capable of operating in both master and slave modes.

2.2.2.2. SIP

Peer-to-peer is a communications model in which each party has the same capabilities and either party can initiate a communication session. SIP uses a peer -to-peer model, because either of two devices involved can initiate a communication session. SIP does not use the master/slave communication model.

2.2.2.3. Megaco

Megaco uses the master/slave communication mode, whereby the MGC (Media Gateway Controller) plays the role of the intelligent master and the MG (Media Gateway) plays the role of the non-intelligent slave. Megaco does not use the peer-to-peer communication model.

2.2.3. Basic Services

2.2.3.1. H.323

The basic service such as the connection establishment and connection release are supported by the H.323 protocol through Q.931 messages (Setup, Call Proceeding, Alerting, Connect, Release Complete).

The possibility of adding a user to the connection is given through the Multipoint Controller Unit (MCU), which is an endpoint on the LAN, enabling three or more terminals and Gateways to participate in a multipoint conference: after a point-to-point connection between two endpoints is established, a third party can be added.

The drop party functionality is not supported. Any participant of a multipoint conference, also the originator, can leave the conference without terminating it. There is no possibility for the originator of the multipoint conference or for other party of the conference to drop an unwanted participant.

A conference can be dropped, but not within a termination of a call. Terminating a call does not terminate a conference; a conference may be explicitly terminated using an H.245 message (dropConference). In this case, the endpoints shall wait for the Multipoint Controller (MC) to terminate the single calls. If a multipoint call is reduced to a point to point call between the last two users, the termination of the call will drop the conference.

2.2.3.2. SIP

The basic services such as establishment / release of a connection are supported by SIP (through the SIP request INVITE and BYE).

Adding a party to an already existing session is also supported through the invitation of a third party to the session or through changing a session to a multipoint conference.

Dropping a party from an existing session is not supported. At any time it is possible for a party, including the originator, to leave a multipoint connection without terminating it, but to drop an unwanted party is not possible.

2.2.3.3. Megaco

The basic services such as establishment / release of a connection and add / drop party service in the sense of a telephone connection are not supported by Megaco, because Megaco is not a real connection control protocol. Megao describes the communication between a MGC and a MG and the translation of media streams from one type of network to the format required in another. Commands like *Add*, *Subtract* and *Modify* are sent from the MGC to MG to add, drop, move and control a media stream, but not for connection control between MGC and MG.

2.2.4. Additionally Services

2.2.4.1. H.323

Additional services such as registration, deregistration and authentication are supported by H.323. For this purpose a H.225 RAS (Registration Admission and Status) unreliable channel will be used to convey the information between H.323 endpoints (terminals and gateways) and gatekeepers.
The registration procedure is handled through the Registration Request (RRQ), which is sent from an endpoint to the Gatekeeper. The Gatekeeper responds with either a Registration Confirmation (RCF) or a Registration Reject (RRJ). The deregistration procedure is handled in same way through following messages: URQ (Unregistration Request), UCF(Unregistration Confirm), URJ (Unregistration Reject). The authentication of the H.323 endpoint is ensured by an encrypted password in the registration procedure.
Additionally each H.225 RAS message containing a parameter integrityCheckValue, which provides improved message integrity/message authentication of the RAS messages. The cryptographically based integrity check value is computed by the sender applying a negotiated integrity algorithm and the secret key upon the entire message.

2.2.4.2. SIP

Additional services such as registration, de-registration and authentication are supported by SIP.
Registration is a mandatory procedure in which each SIP user can dynamically be registered through sending a SIP REGISTER request from the chosen location. The REGISTER request allows a SIP user to inform the registrar (registration server) about the address(es) where it can be reached. SIP Client can be registered from different locations.
Authentication is an optional procedure and supported through the message header fields "WWW-Authentication" and "Authorization" in a REGISTER or in another message to prevent an active attacker from modifying and replaying SIP requests and responses. The same cryptographic measures that are used to ensure the authenticity of the SIP message also serve to authenticate the originator of the message.

2.2.4.3. Megaco

Additional services such as registration, deregistration and authentication in the sense of telephony services are not supported by Megaco.
Registration of the MG in Megaco protocol is used in following sense: the MG announces, through the ServiceChange message, its availability to a MGC (=registration) and notifies the MGC of an impending or completed restart of the MG.
For authentication Megaco has no specific mechanism, but uses the following standard: RFC 2402, "IP Authentication Header" [34] to prevent unauthorized entities from using the protocol messages between MGC and MG. This is a part of the security consideration in the IP environment.

2.2.5. QoS Parameters Transport

2.2.5.1. H.323

Resource parameters as required in a satellite network for resource negotiation, such as the traffic parameters (PDR - Peak Data Rate, MDR - Mean Data Rate, MBS - Max Burst Size, UF - Utilization Factor) and such as the QoS parameters (maxCTD – Max Cell Transfer Delay, CDV - Cell Delay Variation, CLR - Cell Loss Ratio), are not defined and available in a H.323 telephony environment. However, audio, video, data and security capabilities are supported through the capability exchange process.

The H.245 control signaling over the H.245 control channel carries H.245 control messages for capabilities exchange. Capabilities exchange is a process through which the communicating terminals exchange messages to provide their transmit and receive capabilities to the peer endpoint. The capabilityTable in the TerminalCapabilitySet message contains the capability entries. The capability entries describes the video, audio, data, conference and security capabilities, which the terminal is able to receive, transmit and receive/transmit.

The possibility to transport undefined parameters through the H.323 messages is only achievable if placeholders are available in the form of nonStandardData. The NonStandardData parameter is defined in some of H.323 messages, such as Admission Request, Bandwidth Request, to carry information not defined in the H.323 Recommendation, for example, proprietary data. This solution has its limitations, because only one QoS parameter can be transported through one specific H.323 message.

In a H.323 telephone environment there exist some similar parameters to QoS parameters desirable in a MM satellite network. The subsection 2.5.3.1, Modularity of Protocol, H.323 describes this parameters and the possibility of their usage.

2.2.5.2. SIP

The basic SIP standard does not provide any kind of network resource reservation or capabilities negotiation. For network resource reservation the Resource Reservation Protocol (RSVP) described in RFC 2205 [28] is used. For describing multimedia sessions the Session Description Protocol (SDP) described in RFC 2327 [31] is used.

How to support QoS and resource reservation is not described in the SIP standard [40] but there are many papers and internet Drafts, which give a survey and introduction of methods how to support QoS and resource reservation. Following two papers are useful: Integration of Resource Management and SIP [42], SIP Extension for QoS support in Diffserv Networks [13] and Extending SIP for QoS support [8] .

The possibility of transport for QoS parameters is however available without changing the SIP standard, through the SIP message body. The message body is mostly used to carry the SDP parameters, which serves as a description of the session. The body can be also used to carry additionally the needed OoS parameters. For this purpose the definition of the QoS Parameters for SDP, referred to as SDP Plus, is required.

2.2.5.3. Megaco

No procedures for QoS parameter negotiation are foreseen in the Megaco protocol specification. The MGC reserves and requests MG resources for media decoding and encoding for a given Stream(s) and Termination to which they apply. This resource reservation/request is done with local or remote descriptors as described in the MEGACO Standard [39] or [51].

The transport of QoS parameters in Megaco can be done through the definition of additional Packages. Packages are the primary mechanism for extension and they define additional properties, events, signals and statistics that may occur on terminations. Packages defined by IETF will appear in separate RFCs. Packages defined by ITU-T may appear in the relevant Recommendations (e.g. as annexes).

2.2.6. Supplementary Services

2.2.6.1. H.323

The core supplementary services such as Call Transfer (H.450.2) [56], Call Diversion (H.450.3) [57], Call Hold (H.450.4) [58], Call Park and Pick Up (H.450.5) [59], Call Waiting (H.450.6) [60], Message waiting Indication (H.450.7) [61], Name Identification (H.450.8) [62], Call Completion (H.450.9) [63], Call Offering (H.450.10) [64], Call Intrusion (H.450.11) [65] and Common Information Additional Network Feature (H.450.12) [66] are supported by H.323 as described in the respective ITU-T standard.

2.2.6.2. SIP

The core supplementary services such as group invitation, call forwarding and conferencing are supported by SIP as described in subsection C.3.1.2.

Other cure supplementary services such as call hold, call transfer, call back, call pick up and park are not defined in the SIP Standard, but there exits several Internet Draft which gives a proposal how to implement such services.

2.2.6.3. Megaco

The Megaco Protocol specification does not contain specific supplementary services.

2.2.7. Network Elements

2.2.7.1. H.323

The H.323 protocol architecture contains the following network elements: H.323 endpoints (clients and gateways), MCU and H.323 Gatekeeper. These may be mapped to the following satellite network elements:

- H.323 endpoints such as H.323 software client, H.323 phone or a H.323 gateway to another telephone network may correspond to satellite terminals. Both, H.323 endpoints and satellite terminals, are considered from the call control point of view as one party.
- The H.323 Gatekeeper and the MCU may adopt functionality of the Network Operation Center (NOC) with respect to localization and authentication management.

2.2.7.2. SIP

SIP contains the network elements client and server, which may be mapped to the corresponding satellite network elements:

- SIP clients, such as Softclients, IP Phones or Gateways may adopt the functionality of satellite terminals.
- A SIP Proxy Server collocated with Registrar may adopt the functionality of a satellite network element called Network Operation Center (NOC) with localization and authentication management functions.
- The Location Server is a database and may support the same functionality as the System Data Base (Network Center Data Base).
- A SIP Redirect Server is not needed in a satellite network, because the call control and call admission control is always supported by the NOC and not by the terminal itself.

2.2.7.3. Megaco

The MGC and MG have no corresponding network elements in a satellite network, because there is a master/slave relationship between the MGC and one or more MGs, while satellite terminals and the NOC communicate in a peer-to-peer relationship.

2.2.8. Adaptability

2.2.8.1. H.323

The H.323 protocol specification is an umbrella specification covering several protocol standards (e.g. H.225, H.245 and Q.931) and defining a large number of messages and procedures for provision of real-time multimedia services over packet networks (including IP-based networks). In MM satellite network only a subset of the whole complex functionality needs to be supported. Useful simplifications refer to the reduction of the message set and the adaptation of content of messages where needed:

- Some H.323 procedures can be omitted to reduce the protocol complexity. Some examples are:
  - Given that an centralized gatekeeper functionality has to be located in the Network Control Center (NCC) of a MM satellite network, e.g. gatekeeper discovery RAS procedures can be omitted.
  - Furthermore H.225 call signaling messages shall always be exchanged after being routed through the gatekeeper. The direct call routing case can be omitted.
- Extending the H.3232 functionality for QoS support, through extension of QoS parameters transfer (the Admission Request (ARQ) message specifies *only* the requested Call Bandwidth which is an upper limit on the aggregate bit rate for all transmitted and received audio and video channels)
- Simplifications are probably possible in H.245 control signaling flows. Due to the existence of a centralized call control in the MM satellite network the messages and procedures for master & slave determination, round trip delay and maintenance loop can probably be omitted. Thus the number of H.245 messages which have to be exchanged can be significantly reduced.

However, due to the complex interactions between the different protocol standards encompassing the H.323 recommendation, the adaptability is low and requires detailed analysis in each case.

2.2.8.2. SIP

SIP is an ASCII-based, application-layer protocol for creating, modifying, managing and terminating sessions between one or more participants. For the provision of QoS and for resource reservation existing QoS protocols (such as RSVP) may be used. Thus adopting SIP as the call control protocol in MM satellite requires modifications in the following respects:

- Reduction of the number of messages (e.g. response messages)
- Reduction of message size (encoded in text format) e.g. through omitting non-mandatory headers
- Extending the SIP functionality for QoS support, perhaps through extension of SDP parameters.

2.2.8.3. Megaco

Megaco uses a master and slave approach for communication between intelligent network servers (masters) and non-intelligent devices (slaves). Therefore the Megaco commands and procedures do not support essential functionality required for MM satellite networks. The adaptation of the protocol would mean to extend the protocol functions and components:

- Changing the master/slave relationship to peer-to-peer communication form
- Introducing the call connection control instead of only control functionality
- Support of additionally services such as registration, admission and supplementary services.
- Definition of Packages for transfer of QoS parameters.

2.2.9. Conclusion on Applicability

From Table 1 below it can be seen that Megaco does not support most of the criteria comprising applicability as call control protocol for MM satellite networks. Megaco uses a master/slave communication model, which is not applicable for MM satellite environment. Further Megaco does not supports the basic, additionally and supplementary services, which are the main characteristic of a call control environment.

SIP and H.323 are both suited in principle. Both support the main functionality and the network architecture also corresponds to the MM satellite environment. However, H.323 is the more complex protocol and therefore will require more and substantial changes. One advantage of SIP compared with H.323 is the possibility of extension of the SIP message body through transferring the parameters for traffic and QoS purposes in a MM satellite network environment.

Table 1: Summary on Applicability Criteria's

	H.323	SIP	Megaco
Connectivity/Symmetry	- ptp bi-directional - ptmp bi-directional - broadcast uni-directional - broadcast multicast combination	- ptp bi-directional - ptmp bi-directional	- ptp bi-directional - ptmp bi-directional
Communication Model	peer-to-peer and also master/slave	peer-to-peer	master/slave
Basic Services	establish, release, add,	establish, release, add,	not supported
Additional Services	registration, deregistration, authentication	registration, deregistration, authentication	not supported
QoS Parameter Transport	partly supported only through placeholder	supported through SDP	supported through packagers
Supplementary Services	supported	partly supported	not supported
Network Elements	correspond	correspond	not correspond
Adaptability	low	high	low

2.3. Complexity

The criteria complexity is defined via a set of sub-criteria which are listed here:

1. **Message Encoding/Decoding**: message encoding in binary or text format, complexity of encoding, generation and parsing process as well as decoding and parsing process (separation between processes)
2. **Number of Messages for Basic Call**: the number of messages exchanged for basic call establishment and release
3. **Function Specific Message Set**: the message set shows the number of messages used by within the protocol in order to provide a specific protocol functionality (e.g. registration, admission, call initiation, capability exchange, status, release, etc.)
4. **Supported Message Set**: the message set shows the number of the messages supported by the protocol to cover the whole protocol functionality, without relation to a specific functionality
5. **Number of States for Basic Call**: the number of states required for covering selected protocol functionality for basic call establishment and release
6. **Supported State Set**: the state set shows the number of the states supported by the protocol to cover the hole protocol functionality, without dependences of a selected functionality
7. **Protocol Dependencies and Interactions**: the number of required protocols to support the main functionality (main functionality is defined in this chapter as basic call signaling for call establishment and release), interworking and usability with other protocols
8. **Degree of Protocol Optimization**: this criterion provides a quantitative estimation for the extent of complexity reduction in candidate protocols e.g. by reducing the number of protocol messages, parameters, or functionality. The percent value is based on the information given in subsection 2.2.8 "Applicability" under criterion Adaptability.

2.3.1. Message Encoding/Decoding

2.3.1.1. H.323

H.323 uses a binary representation for its messages, based on Abstract Syntax Notation One (ASN.1) and the Packed Encoding Rules (PER). ASN.1 generally requires special code-generators to parse. A text-processing tool is a necessity to make it readable and make it debugable for humans. One advantage of binary representation is that binary codecs are significantly faster than the fastest text codecs and the message size is lower through the binary representation.

2.3.1.2. SIP

SIP is text-based, using ISO 10646 in UTF-8 encoding throughout. This textual encoding simplifies debugging, allows manual entry and analysis of messages, and is suitable for humans to read. Coding with text based will make easier its implementation in programming languages such as JAVA, Tcl, and Perl, and simplifies the debugging process. As a consequence, the messages are large and less suitable for networks where bandwidth, delay, and/or processing are a concern. However this disadvantage can be overcome by efficient compression.

2.3.1.3. Megaco

The Megaco protocol may be encoded both in binary form (in Abstract Syntax Notation One - ASN.1) as well as in text form (Augmented BNF for Syntax Specifications - ABNF).

2.3.2. Number of Messages for Basic Call

2.3.2.1. H.323

Basic Call Establishment:

When establishing a basic call using H.323, the number of messages depends on the type of call establishment method used. There are three possibilities: the common way (1), the "Fast Connect" procedure (2) or the "Fast Connect" procedure combined with tunneling (3). In all procedures the Gatekeeper routed case is used, the H.225 messages are routed by the Gatekeeper and the H.245 messages are only exchanged between the two Terminals.

(1) The common way of call establishment is the standard procedure, which is supported by all H.323 entities defined through the H.323 Protocol Version 1 (1996). The common way requires:
   - two RAS messages (ARQ – Admission Request, ACF – Admission Confirm) per user,
   - four H.245 messages (TerminalCapabilitySet, TerminalCapabilitySetAck, OpenLogicalChannel, OpenLogicalChannelAck) per user and
   - four Q.931 messages (Setup, Call Proceeding, Alerting, Connect) from one user to the Gatekeeper and from the Gatekeeper to he other user, therefore eight Q.931 messages.

   So altogether (2*2 + 4*2 + 8 =) 20 messages have to be sent for establishment of a ptp connection. Figure 1 below shows the corresponding message flow.

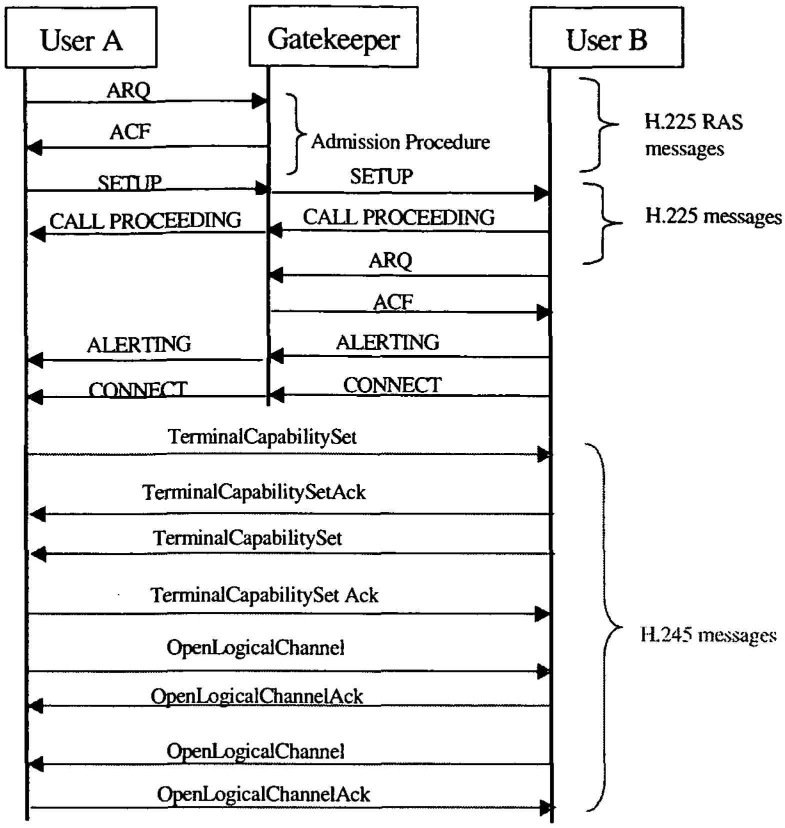


Figure 1: H.323 Basic Call Establishment

(2) The H.323 Protocol Version 2 (1998) defines additionally to the common way the "Fast Connect" variants. If a client supports the protocol version 2.0 functionality, then the Fast Connect procedure can be used, otherwise the common way will be used. In the "Fast Connect" procedure the H.245 messages OpenLogicalChannel and OpenLogicalChannelAck per user are spared by putting this information into the "Setup" message. This results in a total number of 16 messages sent (4 messages less than in the case of common way establishment).

(3) In the H.323 protocol version 4.0 (2000) is the third variant of call establishment the "Fast Connect procedure combined with tunneling" is defined. In this variant all H.245 messages (TerminalCapabilitySet, TerminalCapabilitySetAck, OpenLogicalChannel, OpenLogicalChannelAck) are spared. Instead of sending H.245 messages, the information is put to the "Setup" message before and a "Facility" message after a connection is established. This results minus 8 messages and plus one new messages (Facility) from one user to the gatekeeper and from the gatekeeper to the other user (two time the "Facility" msg.), among a total of 20-8+2 = 14 messages.

A quantitative comparison for the call establishment is shown in Table 2 below. This table realized that the three different kind of call establishing methods use a different number of messages.

Table 2: H.323 Basic Call Establishment

	RAS	H.245	Q.931	**Total number of messages**
(1) Common Way	4	8	8	**20**
(2) Fast Connection	4	4	8	**16**
(3) Fast Connection with Tunneling	4	0	10	**14**

Basic Call Release:

For basic call release two H.245 messages (CloseLogicalChannel, EndSessionCommand) per user, two RAS messages (DRQ – Disengage Request, DCF – Disengage Confirm) per user and one Q.931 messages (Release Complete) are needed. This amounts to 2*2 + 2*2 + 1 = 9 messages. Figure 2 below shows the corresponding message flow.

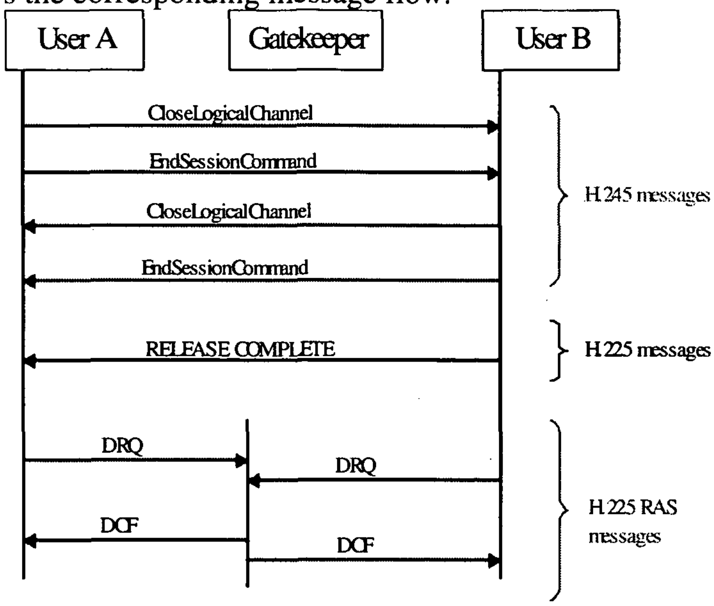


Figure 2: H.323 Basic Call Release

Table 3 summarizes the call release case.

Table 3: H.323 Basic Call Release

	RAS	H.245	Q.931	**Total number of messages**
Call Release	4	4	1	**9**

2.3.2.2. SIP

In the SIP basic call considering the case, where both users are in the same SIP domain and uses the same proxy server, is the mostly useful case, because in a MM satellite environment, only one NOC and one satellite is involved and therefore the two communication satellite terminals will always contact the same operation center for connection control. The next Figure 3 illustrate this case. The number of messages is compound of:

- the messages, which are exchanged between the first user and the proxy server (request: INVITE and response: 100 Trying, 180 Ringing, 200 OK). The "100 Trying" provisional response could be spared out if the proxy server can send a final response within of 200 ms. The proxy server and the location server are co-located on one device and has a internal communication interface.

- the messages, which are exchanged between the proxy server and the second user: (request: INVITE and response: 180 Ringing, 200 OK)
- messages, which are exchanged directly between the communicating user: (request: BYE and response: 200 OK)

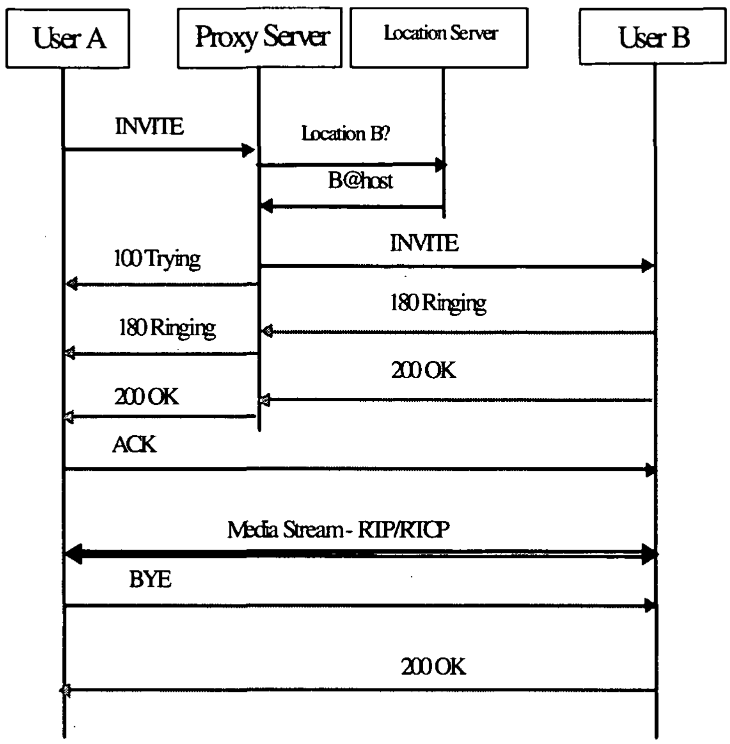


Figure 3: SIP Basic Call

Table 4 below provides a summary of request and response messages exchanged during SIP basic call establishment/release.

Table 4: SIP Basic Call Establishment and Release

	Request	Response	**Total number of messages**
Call Establishment	4 (INVITE, ACK)	5 (100, 180, 200)	**9**
Call Release	1 (BYE)	1 (200)	**2**

2.3.2.3. Megaco

The basic services such as establishment and release of a connection are not supported by Megaco. Commands for adding and dropping media streams (*Add*, *Subtract* and *Modify)* are exchanged between the MGC to MG to add, drop, move and control a media stream, For call signaling any other protocol can be used depending on the network type (e.g. DSS1, H.323, SIP etc.). To add a media stream the following 8 messages are necessary: ADD, Add Reply (to both MG), NOTIFY

Setup, Notify Reply, NOTIFY Connect, Notify Reply. To drop a media stream the following 8 are messages necessary: DROP, Drop Reply (to both MG), NOTIFY Disconnect, Notify Reply (from both MG). The following Figure 4 illustrates this message flow.

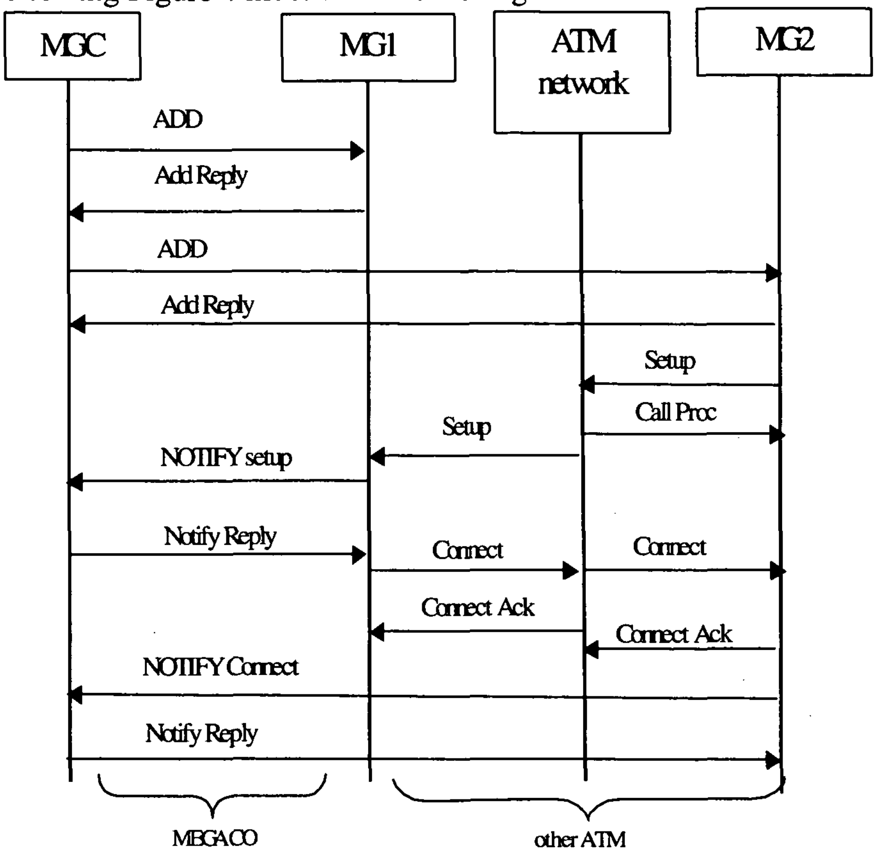


Figure 4: MEGACO Basic Call Establishment

Table 5: Basic Call Establishment and Release (not supported[iii])

	Exchanged Messages between MG and MGC
Add media stream	8
Drop media steram	8

2.3.3. Functions Specific Message Set

2.3.3.1. H.323

Table 6 provides an overview on the number of H.323 signaling messages needed for a specific functionality.

[iii] Megaco is not a protocol for call related signaling, therefore the number of messages is not directly comparable. Only the messages during adding or dropping media streams are considered.

Table 6: Function Specific Message Set of H.323

Function	**Messages**	**Protocol**	**Number of Messages**
Registration	RRQ, RCF, RRJ, URQ, UCF, URJ	H.225 RAS	6
Admission	ARQ, ACF, ARJ	H.225 RAS	3
Call Initiation	Setup	Q.931	1
	OpenLogicalChannel, OpenLogicalChannelAcknowledge, OpenLogicalChannelReject, OpenLogicalChannelConfirm,	H.245	4
Status	Alerting, Call Proceeding, Connect	Q.931	3
Call Release Initiation	Release Complete	Q.931	1
	CloseLogicalChannel, EndSessionCommand	H.245	2
Capability Exchange	TerminalCapabilitySet, TerminalCapabilitySetAck, TerminalCapabilitySetReject, TerminalCapabilitySetRelease	H.245	4

2.3.3.2. SIP

Table 7 provides an overview on the number SIP signaling messages needed for a specific functionality.

Table 7: Function Specific Messages Set of SIP

Function	**Messages**	**Number of Messages**
Registration	REGISTER, 100 Trying, 200 OK	3
Admission	INVITE	1
Call Initiation	INVITE	1
Status	46 response messages for example: 100 Trying, 180 Ringing, 200 OK 1 request message: ACK	46 1
Call Release Initiation	BYE	1
Capability Exchange	(through the SDP Protocol in) INVITE and OPTION message	2

2.3.3.3. Megaco

Table 8 provides an overview on the number of MEGACO messages needed for a specific functionality.

Table 8: Function Specific Message Set of MEGACO

Function	Messages	Number of Messages
Registration	ServiceChanged	1
Admission	*** not supported ***	0
Call Initiation	Add	1
Status	Notify any 4xx and 5xx response code (= 31 and 21)	1 52
Call Release Initiation	Subtract, Move	2
Capability Exchange	AuditCapabilities, AuditValue, Modify	3

2.3.4. Supported Message Set

2.3.4.1. H.323

The H.323 message set can be divided into H.225 RAS messages, H.225 call signaling (the Q.931/H.450 messages) and the H.245 call control messages.
The set of H.225 RAS messages contains 27 messages: GRQ, GCF, GRJ, RRQ, RCF, RRJ, URQ, UCF, URJ, ARQ, ACF, ARJ, BRQ, BCF, BRJ, DRQ, DCF, DRJ, LRQ, LCF, LRJ, IRQ, IRR, IACK, RIP, RAI, RAC. (the specific messages are described in H.225 specification [49] and an overview in subsection B.3.2.1).
The set of H.245 call control messages contains 43 messages. For more detail and description of the specific messages see H.245 specification [50] and an overview in subsection B.3.2.3.
The set of H.225 call signaling (the Q.931/H.450 messages) contains following 8 messages: Alerting, Call Proceeding, Connect, Setup, Status, Status Inquiry, Facility and Release Complete. For more detail and description of the specific messages see H.225 [49] and Q.931 [68] specifications, and an overview in subsection B.3.2.2.
A summary of the supported message set is shown in Table 9 below.

Table 9: H.323 Supported Message Set

	RAS	H.245	Q.931/H.450	**Total number of messages**
Supported Message Set	27	43	8	**78**

2.3.4.2. SIP

The group of request messages contains the following 6 messages: INVITE, ACK, BYE, CANCEL, OPTION und REGISTER. For more detail and description of the specific messages see SIP [40] specification and an overview in subsection C.3.2.1.
The group of response messages contains 46 response messages, which can be further subdivided into the following 6 groups: 1xx: Informational (5 messages), 2xx: Success (1 message), 3xx: Redirection (5 messages), 4xx: Client Error (25 messages), 5xx: Server Error (6 messages) and

6xx: Global Failure (4 messages). For more detail and description of the specific messages see SIP [40] specification and an overview in subsection C.3.2.2.
An overview on the supported message set is shown in Table 10 below.

Table 10: SIP Supported Message Set

	Request	Response	**Total number of messages**
Supported Message Set	6	46	**52**

2.3.4.3. Megaco

In the Megaco/H.248 standard the messages are called commands. The commands and error codes are used for the communication between Media Gateway Controller and Media Gateway for controlling the media streams.
The Megaco command set contains the following 9 commands: Add, Subtract, Modify, Move, Audit Value, AuditCapabilities, Notify and ServiceChange. For more detail and description of the specific messages see H.248 [51] specification and an overview in subsection D.3.2.1.
The Megaco error code set contains 52 error codes, which are listed in the subsection D.3.2.3.
An overview on the supported message set is shown in Table 11 below.

Table 11: MEGACO Supported Message Set

	Commands	Error Codes	**Total number of messages**
Supported Message Set	9	52	**61**

2.3.5. Number of States for Basic Call

2.3.5.1. H.323

The number of the states required for basic call establishment is counted on the incoming and outgoing side. From the H.323 client (user) point of view the incoming side represents the calling party and the outgoing side represents the called party. The number of states for call establishment on the incoming side is fife, namely "null" state, "call present", "call received", "connect request" and "active". There are six states on the outgoing side: "null state", "call initiated", "proceeding indication", "outgoing call proceeding", "call delivered" and "active".
The number of the states required for basic call release on both sides (calling or called) is three states. But the states for call release depend on which user (calling or called) hangs up. If the calling party (user A) hangs up, then following three states are possible: "disconnect request", "release request" and "null" state. The called user (user B) has following three states: "disconnect indication", "release request" and "null" state.
The Table 12 below provides a summary of the number of states of a H.323 basic call.

Table 12: Number of states for a H.323 call (Q.931 states only)

	Call Establishment	Call Release
Calling Party	5	3
Called Party	6	3

2.3.5.2. SIP

The number of states during call establishment on the User Agent Client (UAC) or on the User Agent Server (UAS) side is four. From the SIP client (user) point of view the UAC represents the calling party and the UAS represents the called party. The UAC, who send a request for a SIP invitation has following four states for call establishment: “Calling”, “Proceeding”, “Completed” and “Terminated”. The UAC uses the INVITE-Client-Transaction for this purpose. The UAS, who answers a SIP invitation through a response, has following four states: “Proceeding”, “Completed” “Confirmed” and “Terminated”. The UAS uses the INVITE-Server-Transaction for this purpose.
The number of states during call release on the UAC or on the UAS side is four. From the SIP client (user) point of view the UAC represents the party, who disconnect the call through hanging up and the UAS represents the disconnected party. The UAC uses the Non-INVITE-Client-Transaction with following four states: “Trying”, “Proceeding”, “Completed” and “Terminated”. The UAS uses the Non-INVITE-Server-Transaction with following four states: “Trying”, “Proceeding”, “Completed” and “Terminated”.
The Table 13 below provides a summary of the number of states of a SIP basic call.

Table 13: Number of states for a SIP call

	Call Establishment	Call Release
Calling Party	4	4
Called Party	4	4

2.3.5.3. Megaco

Because Megaco is not a signaling call control protocol for telephony applications, there are no states in sense of a call connection control. The MGC uses information from the signaling protocol (e.g.: SS7) about the current call state to control the media channels in a MG.
The connection model for the Megaco protocol describes the logical entities within the MG that can be controlled by the MGC. The key terms used in the connection model are Terminations and Contexts. A Termination can be in one of the following states: "test", "out of service", or "in service". The "test" state indicates that the Termination is being tested. The state "out of service" indicates that the Termination cannot be used for traffic. The state "in service" indicates that a Termination can be used or is being used for normal traffic, "in service" is the default state.

2.3.6. Supported State Set

2.3.6.1. H.323

The state set of H.323 supports 11 states at the user side of the interface (on the client side) and 12 states at the network side of the interface (on the server side), as described in the Q.931 Recommendation.

The 11 states at the user side of the interface are: null state, call initiated, outgoing call proceeding, call present, call received, connect request, incoming call proceeding, active, disconnect request, disconnect indication, release request.

The 12 states at the network side of the interface are: null state, call initiated, outgoing call proceeding, call present, call received, connect request, incoming call proceeding, active, disconnect request, disconnect indication, release request, call abort.

The Table 14 below provides a summary of the number of H.323 state set.

Table 14: Number of H.323 state set

	Number of states
Client Side	11
Server Side	12

2.3.6.2. SIP

SIP is based on a request/response transaction model. A transaction is a request sent by a client transaction to a server transaction, along with all responses to that request sent from the server transaction back to the client. The client (UAC) side is the calling party and the server side (UAS) is the called party, therefore exists a Client Transaction and a Server Transaction.

Depending on the kind of the request SIP differentiates between: INVITE Transaction (for INVITE request) and Non INVITE Transaction (for BYE, OPTIONS, REGISTER, CANCEL and ACK requests).

The set of states in SIP resulted from the client and the server transaction and also from an INVITE or non-INVITE transaction, are following:

- INVITE-Client-Transaction with states: “Calling”, “Proceeding”, “Completed” and “Terminated”.
- INVITE-Server-Transaction with states: “Proceeding”, “Completed” “Confirmed” and “Terminated”.
- Non-INVITE-Client-Transaction with states: “Trying”, “Proceeding”, “Completed” and “Terminated”.
- Non-INVITE-Server-Transaction with states: “Trying”, “Proceeding”, “Completed” and “Terminated”.

All SIP entities have to support this state set. The client side has to support the INVITE-Client-Transaction and the Non-INVITE-Client-Transaction each with four states, this amount to 8 states.

The server side has to support the INVITE-Server-Transaction and the Non-INVITE-Server-Transaction each with four states, this amount to 8 states. The Table 15 below provides a summary of the number of SIP state set.

Table 15: Number of SIP state set

	Number of states
Client Side	8
Server Side	8

2.3.6.3. Megaco

See description in subsection 2.3.5.3 (Number of States for Basic Call)

2.3.7. Protocol Dependencies and Interactions

2.3.7.1. H.323

The umbrella H.323 recommendation requires <u>three</u> protocols for maintaining the basic call functionality: *H.225* for RAS and H.225 for call signaling (= *Q.931*), *H.245* for call control. Additionally the following protocols defined in H.323 umbrella recommendation are needed for successful communication: RTP / RTCP, audio codecs (G.711, G.723.1, G.728, etc.), video codecs (H.261, H.263) and data codec (T.120).
In addition to the protocols inside of H.323's umbrella recommendation the following protocols are significant: H.230 for frame-synchronous and signal indication, H.235 for security and encryption, H.332 for conference and H. 450 for supplementary services.

2.3.7.2. SIP

The SIP protocol contains the following <u>two</u> protocols for the main functionality (basic call signaling): SIP and SDP (Session Description Protocol). Additionally the RTP / RTSP protocols are needed for successful communication (payload transport).
SIP is incorporating protocols such as the Real-time Streaming Protocol (RTSP) described in RFC 2326 [30], the Real-time Transport Protocol (RTP), described in RFC 1889 [26], the Resource Reservation Protocol (RSVP) described in RFC 2205 [28], the Session Announcement Protocol (SAP) described in RFC 2974 [38] and the Session Description Protocol (SDP), described in RFC 2327 [31].

2.3.7.3. Megaco

Megaco was standardized by the ITEF as RFC3015 [39] and also was published as ITU-T Recommendation H.248 [51]. There are no dependencies on other protocols and Megaco does not use other protocols to cover its own functionality. The Media Gateway Control Protocol (MGCP) described in RFC 2705 [35] is only a predecessor of Megaco, which has some market share, but will not be enhanced further.

2.3.8. Degree of Protocol Optimization

2.3.8.1. H.323

In accordance with the considerations in subsection 2.2.8.1 the following H.245 messages can probably be omitted in the present context (subject to confirmation through a detailed analysis):

- Master and Slave Detection Messages: Determination, Determination Acknowledge, Determination Reject, Determination Release
- Round Trip Delay Messages: Round Trip Delay Request, Round Trip Delay Response
- Maintenance Loop Messages: Maintenance Loop Reject, Maintenance Loop Command Off
- GRQ (Gatekeeper Request), GCF (Gatekeeper Confirm), GRJ (Gatekeeper Reject)

This reduces the number of H.323 message by 15 percent. This reduction is based on examples due to estimate, an accurate analysis should be done in the next chapter.

2.3.8.2. SIP

In accordance with the considerations in subsection 2.2.8.2 a number of SIP messages can probably be omitted when applying SIP as call control protocol in a MM satellite network. This reduces the number of SIP messages by 20 percent.

Examples for response messages, which may be omitted are:

- 1xx: Informational: "181" - Call Is Being Forwarded
- 3xx: Redirection: "300" - Multiple Choices, "301" - Moved Permanently, "302" - Moved Temporarily, "305" - Use Proxy, "380" - Alternative Service
- 4xx: Client Error: "482" - Loop Detected, "483" - Too Many Hops, "407" - Proxy Authentication Required

Example for header parameters, which can be omitted are: Call-Info, Contact, Date, Expires, Max-Forwards, Min-Expires, Record-Route, Route, Subject, Via.

2.3.8.3. Megaco

A reduction of the Megaco protocol is not considered, because the Megaco protocol does not support the whole functionality which is required in a MM satellite network.

2.3.9. Conclusion on Complexity

Considering Table 16 below it can be seen that H.323 has the highest complexity considering all criteria. This is mainly due to the fact that the H.323 umbrella recommendation covers much more functionality than is required for connection related signaling. SIP has a considerably lower complexity in comparison to H.323. The low complexity of Megaco can be explained by the fact that Megaco is not a protocol intended for connection related signaling and does support not the same functionality as the other two protocols.

Table 16: Summary on Complexity Criteria

	H.323	SIP	Megaco
Message Encoding/ Decoding	binary	text (optional compression)	binary or text
Number of Messages for Basic Call	20 or 16 or 14 for call establishment 9 for call release	9 for call establishment 2 for call release	(Note 1) 8 for add media stream 8 for drop media stream
Function Specified Message Set	Registration = 6	Registration = 3	Registration = 1
Supported Message Set	8 - Q.931 27 - H.225 43 - H.245 => 78 total	6 request 46 response => 52 total	9 messages 52 response => 61 total
Number of States for Basic Call	5 / 6 for call establishment 3 for call release	4 for call establishment 4 for call release	(Note 2) 3 state of a Termination
Supported State Set	11 client side 12 server side	8 client side 8 server side	(Note 2) 3 state of a Termination
Protocol dependencies	3 protocols	2 protocols	1 protocol
Degree of Protocol Optimization	15 % (estimate)	20 % (estimate)	not necessary

Note 1 ... Megaco is not a protocol for call related signaling, therefore the number of messages is not directly comparable. Only the messages during adding or dropping media streams are considered..

Note 2 ... Megaco is not a protocol for call related signaling , therefore the number of states is not directly comparable. Only the logical entity: Termination of a MG controlled by the MGC can be in one of the three states.

2.4. Performance

The criteria performance is defined via a set of sub-criteria which are listed here:

1. **Basic Call Signaling Characteristics**: the type, number, and uncompressed sizes of candidate protocol messages required in basic call establishment and release procedures
2. **Signaling Channels**: the number of signaling channels required by the protocol.
3. **Signaling Load** (Note 1): the signaling load at the RASC (Random Access Channel) and DSC (Dedicated Signaling Channel) channels are used by satellite terminals to signal call establishment and release to the NOC. Especially the RASC dimensioning is critical, since slotted ALOHA [iv] access is used and therefore the probability of collisions of signaling messages will increase with signaling load. The signaling load on these channels depends on the following parameters:
   - the number and size of signaling messages transmitted to the NOC e.g. for call establishment and release.
   - the number of busy hour call attempts (BHCAs) (depends e.g. on the total number of terminals within the MM satellite network, the service mix and terminal characteristics etc.)

Note 1: The potential RASC and DSC signaling load for the ESW L3 (EuroSkyWay Layer 3) protocol as well as for the unmodified candidate protocols will be calculated for three different scenarios ("consumer-oriented", "mixed" and "business-oriented") in a typical Satellite Multimedia network. The consumer-oriented scenario contains twice as much small terminals than the business-oriented scenario. Table 17 summarizes scenario characteristics which represent the output from a previous project [9].

[iv] ALOHA, also called the Aloha Method, refers to a simple communications scheme in which each source (transmitter) in a network sends data whenever there is a frame to send. If the frame successfully reaches the destination (receiver), the next frame is sent. If the frame fails to be received at the destination, it is sent again. This protocol was originally developed at the University of Hawaii for use with satellite communication systems in the Pacific. In a wireless broadcast system or a half-duplex two-way link, Aloha works perfectly. But as network become more complex, for example in an Ethernet system, trouble occurs because data frames collide. To minimize the number of collision a scheme so called slotted Aloha was developed. This system employs signals called beacons that are sent at precise intervals and tell each source when the channel is clear to send a frame. (A further improvement can be realized by protocol called Carrier Sense Multiple Access with Collision Detection (CSMA))

Table 17: Scenario characteristics

Scenario	Scenario Type	Total Number of Terminals	Call Acceptance Rate [1/s]	Call Blocking Rate [1/s]
Scenario A	Consumer-Oriented	19401	40	10
Scenario B	Mixed	10884	32	5
Scenario C	Business-Oriented	8342	27	5

The signaling load will be calculated according to the following formula:

$$L = CAR * SC + CBR * UC$$

whereby:
L ... Signaling Load
CAR ... Call Acceptance Rate [l/s]
CBR ... Call Blocking Rate [l/s]
SC ... Number of transmitted bytes for successful Basic Call
UC ... Number of transmitted bytes for unsuccessful Basic Call

2.4.1. Basic Call Signaling Characteristic

2.4.1.1. H.323

The size of H.323 messages varies depending on the kind of message. For basic call establishment and release three types of messages are used:

- H.225 call signaling (Q.931 messages): the message size varies between 80 and 300 bytes.
- RAS messages: the message size varies between 35 and 200 bytes.
- H.245 messages: the message size varies between 20 and 360 bytes.

Table 18 below provides a summary of the message characteristics of the various H.323 messages used for call establishment and call release. The number of messages is calculated based on the assumption that gatekeeper routing is used.

Table 18: Signaling Characteristics for H.323 Basic Call Procedures using the common way

Basic Call Procedure	**Message Type**	**Message Size (in Bytes)**	**Number of Messages transmitted**	**Number of Bytes transmitted**
Call Establishment => successful case	ARQ (H.225)	232	2	464
	ACF (H.225)	26	2	52
	Setup (Q.931)	306	2	606
	Call Proceeding (Q.931)	80	2	160
	Alerting (Q.931)	84	2	168
	Connect (Q.931)	120	2	420
	TerminalCapabilitySet (H.245)	223	2	446
	TerminalCapabilitySetAck (H.245)	3	2	6
	OpenLogicalChannel (H.245)	20	2	40
	OpenLogicalChannelAck (H.245)	23	2	46
	Total Number of Bytes transmitted for Call Establishment			*2408*
Call Release => successful case	Release Complete (Q.931)	50	2	100
	CloseLogicalChannel (H.245)	20	2	40
	EndSessionCommand (H.245)	2	2	4
	DRQ (H.225)	155	2	310
	DCF (H.225)	3	2	6
	Total Number of Bytes transmitted for Call Release			*460*
Total Number of Bytes transmitted for successful Basic Call				**2868**
Call Establishment => unsuccessful case	ARQ (H.225)	232	1	232
	ACF (H.225)	26	1	26
	Setup (Q.931)	306	1	306
	Release Complete (Q.931)	50	1	50
	DRC (H.225)	155	1	155
	DCF (H.225)	3	1	3
	Total Number of Bytes transmitted for unsuccessful Basic Call			**772**

Table 19: Message Size for Basic Call using the Fast Connect Procedure

Basic Call Procedure	Message Type	Message Size (in Bytes)	Number of Messages transmitted	Number of Bytes transmitted
Call Establishment => successful case	ARQ (H.255)	232	2	464
	ACF (H.255)	26	2	52
	Setup (Q.931)	734	2	1468
	Call Proceeding (Q.931)	80	2	160
	Alerting (Q.931)	84	2	168
	Connect (Q.931)	120	2	240
	TerminalCapabilitySet (H.245)	223	2	446
	TerminalCapabilitySetAck (H.245)	3	2	6
	Total Number of Bytes transmitted for Call Establishment			*3004*
Call Release => successful case	Release Complete	50	2	100
	CloseLogicalChannel (H.245)	20	2	40
	EndSessionCommand (H.245)	2	2	4
	DRQ (H.255)	155	2	310
	DCF (H.255)	3	2	6
	Total Number of Bytes transmitted for Call Release			*460*
Total Number of Bytes transmitted for successful Basic Call				**3464**
Call Establishment => unsuccessful case	ARQ (H.255)	232	1	232
	ACF (H.255)	26	1	26
	Setup (Q.931)	734	1	734
	Release Complete (Q.931)	50	1	50
	DRC (H.255)	155	1	155
	DCF (H.255)	3	1	3
	Total Number of Bytes transmitted for unsuccessful Basic Call			**1200**

The total message size described in Table 18 and Table 19 contains H.225, H.245 and Q.931 messages and is the total number of bytes which are transmitted, without the consideration between which parties they are transmitted. If only two terminals were involved without Gatekeeper, then only the Q.931 and H.245 messages were required. In a MM satellite network call establishment and release between two terminals without a NCC is not realizable, therefore only the Gatekeeper routed variants can be considered. In the Gatekeeper-routed case the H.225 and Q.931 messages are routed by the Gatekeeper in call establishment and call release procedure. The H.245 control signaling connection exists only between the two terminals.

For the third criterion of performance (for "signaling load") the total number of bytes transmitted for successful and unsuccessful basic call routed by the Gatekeeper is required. The following Table 20 shows this calculation, which originates from the calculation of the H.225 and Q.931

messages, without consideration of the H.245 messages. The calculation is only for the criterion "Signaling Overhead" and contains only the incoming messages from the two terminals to the NOC and not the outgoing messages from NOC to terminals.

Table 20: Total number of bytes transmitted in Gatekeeper routed case for signaling load calculation

	Successful Basic Call		**Unsuccessful Basic Call**	
Common Way	1182 bytes	ARQ, Setup, Call Proceeding, ARQ, Alerting, Connect, Release Complete, DRQ, DRQ	693 bytes	ARQ, Setup, DRQ
Fast Connect	1610 bytes		1121 bytes	

2.4.1.2. SIP

Typical SIP message sizes are in the range from a few hundred bytes (400 - 500 bytes) to as much as 2000 bytes. Each SIP message consists of a request or response line, a message header and a message body. The request or response line has a size between 15 und 70 bytes. The message header has a size between 200 and 300 bytes. The message body length in bytes is provided by the Content-Length header field and is between 0 and 150 bytes. If no body is present in a message, then the Content-Length header field value must be set to zero. This size of the message covering only mandatory parameters is between 250 and 500 bytes. Inclusion of optional parameters may increase the messages size to as much as 2000 bytes.

Table 21 below provides a summary of the message characteristics of the various SIP messages used for call establishment and call release. The SIP messages in the Table 21 are taken from the SIP standard [40]. In this call scenario one Proxy server and two terminals are involved, therefore the number of messages transmitted is 2 (terminal to Proxy and Proxy to the other terminal). The number is 1 in the case of an exchange between Proxy and terminal once in only one direction and in the case of the ACK messages the exchange occurs between two terminals.

Table 21: Signaling Characteristics for SIP Basic Call Procedures

Basic Call Procedure	**Message Type**	**Message Size (in Bytes)**	**Number of messages transmitted**	**Number of Bytes transmitted**
Call Establishment => successful case	INVITE	450	2	900
	100 Trying	197	1	197
	180 Ringing	257	2	514
	200 OK	414	2	828
	ACK	233	1	233
	Total Number of Bytes transmitted for call establishment			*2672*
Call Release => successful case	BYE	233	2	466
	200 OK	194	2	388
	Total Number of Bytes transmitted for call release			*854*
Total Number of Bytes transmitted for successful basic call				**3526**
Call Establishment => unsuccessful case	INVITE	450	1	450
	404 Not found	208	1	208
	ACK	233	1	233
Total Number of Bytes transmitted for unsuccessful Basic Call				**891**

For the third performance criteria "Signaling Load" the number of byte transmitted from the terminals to the Proxy are counted; these are only the incoming messages to the Proxy, shown in the Table 22 below. The outgoing messages from the Proxy are not calculated for the Signaling Overhead.

Table 22: Total number of bytes transmitted for signaling load calculation

Successful Basic Call		**Unsuccessful Basic Call**	
1548 bytes	INVITE, 180 Ringing, 200 OK, Bye, 200 OK	683 bytes	INVITE, ACK

2.4.1.3. Megaco

The Megaco/H.248 standard defines both a plain text encoding and a binary encoding (ASN.1 BER). The following possible encoding /decoding schemes can be used:

- pretty printed text: in the text encoding, the protocol stack implementers have the choice of using a mix of short and long keywords. It is also possible to add white spaces to improve readability. The pretty text encoding utilizes long keywords and an indentation style like the text examples in the Megaco/H.248 specification.
- the compact text encoding uses the shortest possible keywords and no optional white spaces.
- ASN.1 BER, native configuration.
- ASN.1 PER, not standardized as a valid Megaco/H.248 encoding, but included for the matter of completeness as its encoding is an extremely compact, native configuration.

In the measurements there are no significant differences in message sizes between ASN.1 BER and the compact text format. The Table 23 shows the average message size using the specified encoding/decoding schemes. Average message sizes are between 86 and 252 bytes. Considering the minimum and maximum message size, then the size of messages can vary between 30 and 600 bytes depending on the message contents and of the encoder/decoder.

Table 23: Signaling Characteristics for MEGACO Basic Call Procedure

Encoder / Decoder	**Average Message Size**	**Number of Messages transmitted** successful case	**Total Number of Bytes transmitted for Basic Operation** successful case	**Number of Messages transmitted** unsuccessful case	**Total Number of Bytes transmitted for Basic Operation** unsuccessful case
pretty printed text	252 bytes	8	**2016**	2	**504**
compact text encoding	140 bytes		**1120**		**280**
ANS.1 BER	153 bytes		**1224**		**306**
ANS.1 PER	86 bytes		**688**		**172**

The successful operation contains 8 Megaco messages, described in the subsection D.3.2. The unsuccessful case contains 6 messages left, amount following 2 messages: ADD and AddReply (example: 430 – Unknown TerminationID).

2.4.2. Signaling Channels

2.4.2.1. H.323

Each H.323 call consists of the collection of reliable and unreliable channels between the endpoints. A reliable channel is a transport connection used for reliable transmission of an information stream from its source to one or more destinations. An unreliable channel is a logical communication path used for unreliable transmission of an information stream from its source to one or more destinations. H.323 uses reliable and unreliable channels in the following way:

- Reliable Channel: (e.g. TCP) end-to-end service is mandatory for the H.245 Control Channel, the Data Channels and the Call Signaling Channel.
- Unreliable Channel: (e.g. UDP) end-to-end service is mandatory for the Audio Channels, the Video Channels and the RAS Channel.

Usage of reliable channels:

- *H.245 Control Channel*: used to carry the H.245 control information messages [ITU-T H.245] between two H.323 endpoints
- *Data Channel*: data applications and associated user interfaces which use T.120 or other data services over the data channel [ITU-T T.120]
- *Call Signaling Channel*: used to convey the call setup and teardown messages [ITU-T H.225.0] between two H.323 entities.

Usage of unreliable channels:

- *Audio Channel*: one or more audio channels can be used to allow one or more language (codes) to be conveyed or for using multipoint conferencing.
- *Video Channel*: unreliable channel used to carry video streams, video codec is optional and more than one video channel may be transmitted and/or received, as negotiated.
- *RAS Channel*: unreliable channel used to convey the registration, admissions, bandwidth change, and status messages [ITU-T H.225.0] between two H.323 entities

All the above mentioned H.323 channels are logical channels and there is no relationship between a logical channel and a physical channel. A logical channel is used to carry the information streams between two H.323 endpoints. These channels are established / released following the H.245 OpenLogicalChannel / CloseLogicalChannel procedures. The number of required physical channels is one. The 6 logical channels can be divided into three reliable and three unreliable channels, some of them are for signaling and some of them for payload, as shown in the Table 24 below.

Table 24: H.323 logical channel usage

	Reliable	**Unreliable**
Signaling	H.245 Control Channel, Call Signaling Channel	H.225 RAS Channel
Payload	Data Channel	Audio Channel, Video Channel

2.4.2.2. SIP

One physical channel is required for sending and receiving the data. The transport layer uses the physical channel to deliver the data from source to destination and reverse. On the transport layer SIP allows the use of a reliable transport protocol such as TCP, SCTP, TLS (TLS channel is a secure encrypted transport) or an unreliable transport protocol such as UDP. Using these transport protocols one or more logical channels can be connected.

One logical cannel will be used for signaling, and this can be a non-secure or a secure TLS (transport layer security) channel. This secure channel is also called authenticated channel, because before using this secure channel a previous authentication of user is required. A second logical channel is the communication channel, which is used for the actual exchange of payload between the sender and receiver, using a separate protocol. A typical example is RTP/RTSP. The communication channel can be a voice channel, a video channel or a data channel.

Summary: One physical channel und at least two logical channels are required. One logical channel is used for signaling and one or more logical channels for payload such as audio, video or data.

2.4.2.3. Megaco

The Megaco protocol describes the communication between the MGC and MG and needs one physical and one logical channel for the communication (signaling) between them.

2.4.3. Signaling Load

2.4.3.1. H.323

For each of the MM Satellite Scenarios described in Table 17 above the signaling load for the H.323 protocol taking into account the various call establishment methods can be calculated. The calculation is based on the assumption that the unchanged H.323 message formats are used (without specific enhancements or optimizations required for call signaling in a MM satellite system). Table 25 below shows a comparison of the results for the three Scenarios.

Table 25: Comparison of H.323 signaling load

Scenarios	**Call Acceptance Rate [1/s]**	**Call Blocking Rate [1/s]**	**Call Establish Method**	**Transmitted bytes for basic call**		**Signaling Load on RASC and DSC** In bytes/s
				successful case	unsuccessful case	
Scenario A	40	10	Common Way	1182	692	**54200**
			Fast Connect	1610	1121	**75610**
			Fast Connect with Tunneling	not available	not available	-
Scenario B	32	5	Common Way	1182	693	**41289**
			Fast Connect	1610	1121	**57125**
			Fast Connect with Tunneling	not available	not available	-
Scenario C	27	5	Common Way	1182	693	**35379**
			Fast Connect	1610	1121	**49075**
			Fast Connect with Tunneling	not available	not available	-

The number of transmitted bytes for the "Fast Connect with Tunneling" Procedure is not available yet, because the product supporting this procedure are in the development and no test data are available yet. This procedure is only described in the protocol version 4 which is in status "pre-published" at this time.

2.4.3.2. SIP

The Table 26 below shows the signaling overhead of the SIP protocol for call establishment. The calculation is based on the assumption that the unchanged SIP message formats are used (without specific enhancements, compression or optimizations required for call signaling in a MM satellite system).

Table 26: Comparison of SIP signaling load (without compression and optimization)

Scenarios	Call Acceptance Rate [1/s]	Call Blocking Rate [1/s]	Transmitted bytes for basic call		Signaling Load on RASC and DSC [bytes/s]
			successful case	unsuccessful case	
Scenario A	40	10	1548	683	**68750**
Scenario B	32	5			**52951**
Scenario C	27	5			**45211**

2.4.3.3. Megaco

The calculation of Signaling Overhead in case of Megaco is not useful, because the Megaco architecture cannot be mapped to the MM satellite network architecture. Figure 5 shows the message flow and architecture differences. The signaling overhead counts the incoming request from the terminal to the NOC. In case of Megaco the incoming messages are only response messages and not request messages.

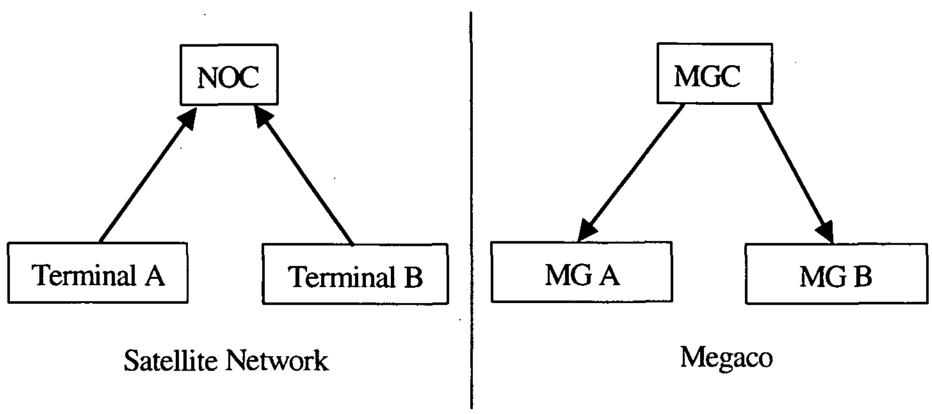


Figure 5: Communication flow in MM satellite architecture and in Megaco architecture

2.4.4. Conclusion on Performance

Table 27 summarizes the conclusions on protocol performance, defined in terms of message size and signaling load. As can be seen, the H.323 family of protocols is the most resource intensive signaling framework of the four protocols considered. This is mainly due to its complexity, as outlined in previous subsections. The baseline version of SIP involves a similar signaling characteristics and load as H.323.

Table 27: Summary on the performance criteria

	H.323	**SIP**	**Megaco**
Basic Call Signaling Characteristics **(in transmitted bytes)**	successful call establishment: 2408/3004 successful call release: 460/460 unsuccessful call establishment: 772/1200 (Note 1)	successful call establishment: 2672 successful call release: 854 unsuccessful call establishment: 891	not applicable
Signaling Channels	1 physical 6 logical channels: (3 reliable + 3 unreliable): - 3 for signaling and - 3 for payload	1 physical 4 logical: - 1 for signaling and - 3 for payload	1 physical 1 logical for signaling
Signaling load **(in Bytes/s)**	Common Way: Scenario A: 41289 Scenario B: 41289 Scenario C: 35379 Fast Connect: Scenario A: 75610 Scenario B: 52125 Scenario C: 49075	Scenario A: 68750 Scenario B: 52951 Scenario C: 45211	not applicable

Note 1: The Basic Call signaling characteristics are calculated for H.323 using the Common Way and Fast Connect Method.

2.5. Extensibility

The criteria extensibility is defined via a set of sub-criteria which are listed here:

1. **Version Compatibility**: existence of protocol built-in version identifiers that can be used to control extensibility mechanisms based on version.
2. **Feature evolution**: how can evolving features can be integrated into the protocol or how does the protocol provide means to incorporate vendor specific (proprietary) features/services.
3. **Modularity of protocol**: easiness of integrating advanced services/features (such as QoS support, capability exchange, service discovery, conference control etc.). For MM satellite network a most important feature is the support of guaranteed QoS for each connection, therefore a candidate protocol have to transport the QoS parameters for the negotiation of this feature. This sub criteria will also analyze the extensibility of the protocol for QoS support.

2.5.1. Version Compatibility

2.5.1.1. H.323

H.323 requires full backward compatibility. This ensures continuous support of existing features. H.323 has version identifiers that can be used to control extensibility mechanisms based on version.

2.5.1.2. SIP

SIP does not have explicit requirements for compatibility among versions. Unknown/unsupported headers are ignored by default. This reduces code size and protocol complexity. Also, this provides flexibility in terms of developing/evolving features and makes encoding/decoding clean and concise. An adverse effect of this may be that features supported by the older versions may not be supported by newer version. Clients can indicate what feature the server must understand in the Require header. If the server does not support the named feature the request will send an error response with a list of supported features.

Features can be registered with the Internet Assigned Number Authority (IANA) by any creator of a feature. SDP is also easy to extend with new attributes.

2.5.1.3. Megaco

Meantime Megaco protocol exists only in version 1.0.

2.5.2. Feature Evolution

2.5.2.1. H.323

H.323 provides extensibility mechanism only if placeholders were defined. H.323 defines a NonStandardParameter parameter field in various locations in ASN.1 to incorporate vendor specific (proprietary) features. These parameters contain a vendor code, followed by an opaque value that has meaning only for that vendor. This allows different vendors to develop their own extensions. However it has its limitations. First, extensions are limited only to those places where a non-standard parameter has been added. If a vendor wishes to add a new value to an existing parameter, and there is no placeholder for a non-standard element, it cannot be added. Secondly, H.323 has no mechanism for allowing terminals to exchange information about the extensions supported by the various terminals. If changes are made to existing capabilities or control message parameters other than NonStandardParameter, a new version of corresponding specifications may need to be issued.

2.5.2.2. SIP

Using SIP, a new feature may be evolved by extending or defining new SIP header information. Current SIP RFC defines default headers and some extensions. New extensions can be added as a part of separate RFCs or defined as Internet Draft.

2.5.2.3. Megaco

To extend Megaco, Megaco provides the usage of new/additionally packages. Packages are the primary mechanism for extension and they define additional properties, events, signals and statistics that may occur on terminations. Packages defined by IETF will appear in separate RFCs. Packages defined by ITU-T may appear in the relevant Recommendations

2.5.3. Modularity of Protocol

2.5.3.1. H.323

H.323 is an umbrella recommendation comprising vertically integrated sub-protocol suites of H.225, H.245, Q.931, H.450 etc. Hence, from the feature/service perspective, there is no clean separation between these sub-protocols. This results in interactions between its sub-protocols. Also, most of the services are in-built and intertwined between more than one sub-protocol.

If H.323 should be extended in order to serve as a connection control protocol for MM networks, at least support of the following features is required: Service Category, Connection Configuration, QoS Parameters and Traffic Parameters. For more details for the QoS parameters see description in subsection 2.2 (applicability criteria “QoS Parameter Transport”)

1. Service Category (The type of services: Consumer-Oriented or Business-Oriented ...) is not supported by the protocol standard.
2. Connection Configuration is supported in the messages ARQ and Setup before the call and BRQ during the call through the parameter “call type”, which can be set to pointToPoint, oneToN, nToOne, nToN. Using this value, called party's gatekeeper can attempt to determine 'real' bandwidth usage. The default value is pointToPoint for all calls; it should be recognized that the call type may change dynamically during the call.
3. QoS Parameters (maxCTD, CDV and CLR) are not supported through the H.323 protocol standard. There is only the possibility to indicate resource reservation through the endpoint, but H.323 does not support the resource reservation mechanism. This indication is supported through the parameter “transportQOS” in an ARQ. For more details see subsection 2.5.3.3.
4. Traffic Parameters are such as MDR, MBS, UF are not supported, only the PDR is supported in the ARQ message through the “bandwidth” parameter. A terminal can use this parameter to

request the max bandwidth required for the call. The Gatekeeper can agree or change the requested bandwidth in the ACF message.

Due to the complex extensions mechanism as described above, the extensibility of H.323 protocol standard with respect to QoS, and traffic parameters may prove to be difficult.

2.5.3.2. SIP

In general, SIP defines methods, default request/response headers and status codes. Service definition is included in the header itself. The extensions to existing headers may require changes in RFCs where the original header is defined. New headers can be easily added by having definition in separate RFC, Internet Draft or other standards process.

As far as QoS support as a new SIP feature is concerned, a draft exists, which discusses how network QoS can be made a precondition to establishment of sessions initiated by SIP (IETF, Internet Draft, draft-ietf-sip-manyfolks-resource-07.txt, G. Camarillo, W. Marshall, Jonathan Rosenberg "Integration of Resource Management and SIP") [14] and RFC3312 [42].) A precondition is a set of constraints about the session which are introduced in the offer. These preconditions simply require a participant to use existing resource reservation mechanisms (e.g. using RSVP) before beginning the session. The recipient of the offer generates an answer, but does not alert the user or otherwise proceed with session establishment. That only occurs when the preconditions are met.

Some other methods for introducing QoS into SIP are available and published as white paper or publication. One of them is the Q-SIP method described in [8] (D. Papalilo, S. Salsano, L.Veltri, "Extending SIP for QoS support", Università di Roma "La Sapienza" (Italy)]. The main idea behind this solution is to eliminate the RSVP signaling from the terminals, and to use SIP as unique call setup protocol for QoS calls. The additional advantage is that all the QoS related functions can be moved to the SIP proxy server that will control both call setup and resource reservation, thus relieving the terminals from unneeded complexity.

A possible solution for a MM satellite network could be an extension of the SDP protocol for negotiation of the QoS parameters. This solution can be similar as proposed in the description in [42] and [14]. In this case after the INVITE message is sent, the response will be a "183 Proceeding" and the signaling will stop until the requested resources are reserved and resume after the reservation. The body of each INVITE message, described through the SDP protocol, will contain the required QoS parameters.

2.5.3.3. Megaco

The Megaco protocol hardly supports the required basic functionality described in the subsection Applicability. Extensibility would require to change the main character, behavior and functionality of the protocol. Such an extension would require a high effort and thus can not be recommended.

2.5.4. Conclusion on Extensibility

Extension of the existing protocol functionality is needed for all candidate protocols (at least for supporting QoS parameters). SIP and SDP are very easy to extend, because the unknown header will be ignored and the main functionality is not distributed. H.323 is very hard to extend, because the extension is only possible if placeholders were defined. Megaco has the highest need for extension to support the basic functionality for connection related signaling. Table 28 summarizes the conclusions on protocol extensibility.

Table 28: Summary on the extensibility criteria

	H.323	SIP	Megaco
Version Compatibility	full backward compatibility	no backward compatibility (unknown header ignored)	only version 1.0 available
Feature Evolution	extensibility only if placeholder available	extension through defining a new SIP header	extension through packages
Modularity of Protocol	complex mechanisms to extend required for Service Category, QoS Parameters and Traffic Parameters	easiness to extend by separate RFC required for Service Category, QoS Parameters and Traffic Parameters	high effort to extend required for the hole funtionality

2.6. Resource and Bandwidth Efficiency

The criteria resource and bandwidth efficiency is defined via a set of sub-criteria which are listed here:

1. **Bandwidth Management**: the possibility and kind of bandwidth management before and during the call
2. **Compression Gain**: the compressed size of the messages vs. uncompressed size
3. **Resource Allocation / Reservation**: the possibility and kind of resource reservation/allocation mechanism used

2.6.1. Bandwidth Management

2.6.1.1. H.323

The H.323 Gatekeeper provides bandwidth control as mandatory services and bandwidth management as optional services. The bandwidth control is done through the RAS channel using H.225 messages to convey bandwidth information through the following RAS messages: BRQ (Bandwidth Request), BCF (Bandwidth Confirm) and BRJ (Bandwidth Reject). Bandwidth management is the control of the number of H.323 terminals permitted simultaneous access to the network. Through the use of H.225.0 signaling, the Gatekeeper may reject calls from a terminal due to bandwidth limitations. This may occur if the Gatekeeper determines that there is not sufficient bandwidth available on the network to support the call.

Both functions operate before a call and also during an active call when a terminal requests additional bandwidth. Before a call the ARQ (Admission Request) message will be used from the terminal to request the upper limit of required bandwidth for a call from the Gatekeeper. The Gatekeeper can reduce the requested call bandwidth in the ACF (Admission Confirm) message. An endpoint or the Gatekeeper can modify the call bandwidth during a call using the following messages: BRQ, BRJ and BCF.

An endpoint shall assure that the aggregate for all transmitted and received audio and video channels, excluding any RTP headers, RTP payload headers, network headers, and other overhead, is within this bandwidth. Data and control channels are not included in this limit. After the call the Gatekeeper needs to know about the release of bandwidth, which is done with the DRQ (Disengage Request), DCF (Disengage Confirm) and DRJ (Disengage Reject) messages.

2.6.1.2. SIP

SIP does not support bandwidth management in the sense of controlling the available bandwidth on the network. SIP only supports the exchange of bandwidth information for session or media between two endpoint using the SDP protocol. In the SDP protocol the parameter b (bandwidth information) is defined, which is used as follows: b = <modifier> : <bandwidth-value>. The parameter b specifies the proposed bandwidth to be used by the session or media, and is optional. The <modifier> is a single alphanumeric word giving the meaning of the bandwidth figure. The <bandwidth-value> is in kilobits per second.

2.6.1.3. Megaco

The Megaco Protocol specification does not contain any specification about bandwidth management.

2.6.2. Compression Gain

2.6.2.1. H.323

H.323 messages are coded in binary form; this ensures a reduced size of the signaling messages and saves bandwidth. Further compression is not foreseen for H.323 messages.

2.6.2.2. SIP

SIP, along with many other IP based protocols used for multimedia communications, are UDP-based, textual protocols engineered for environments where bandwidth is not a major concern. As a result, these protocols have not been optimized for message size. With the planned usage of these protocols in wireless handsets as part of 2.5G and 3G wireless, the large size of these messages is problematic, primarily for latency reasons. As a result, for usage in these environments, it has been proposed to efficiently compress SIP messages. Currently, however, there is no agreed standard for compression of SIP messages, the SIP standard [40] describes only the protocol functionality and the message presentation in text format. There are however some Internet Drafts available which describe how to compress and how to negotiate the compression parameters. Some of the Drafts describe how to negotiate the compression parameters and some others introduce the compression method. These are

- IETF, Internet Draft, draft-camarillo-sip-sigcomp-00.txt, G. Camarillo, “SigComp discovery for SIP” [15],
- IETF, Internet Draft, draft-ietf-rohc-sigcomp-algorithm-00.tx, Richard Price, Jonathan Rosenberg, Abigail Surtees, Mark A West, Lawrence Conroy, “Universal Decompression Algorithm” [16],
- IETF, Internet Draft, draft-ietf-rohc-signaling-req-assump-06.txt, Hans Hannu, “Signaling Compression Requirements & Assumptions” [17],
- IETF, Internet Draft, draft-ietf-sip-compression-01.txt, G. Camarillo, “Compressing the Session Initiation Protocol” [18],
- IETF, Internet Draft, draft-rosenberg-rohc-sip-udpcomp-00.txt, Rosenberg, “Compression of SIP” [19]
- and Hans Hannu, „Signaling compression, Application signaling over cellular links” (based on <draft- hannu- rohc- signaling- cellular- 01. txt>) [10],

which are also useful and necessary in a MM satellite network to improve bandwidth management. The importance of the compression is increased and the compression of SIP is also considered from other technologies, for example in the 3G Wireless Network, described in Jouni Korhonen “SIP Signaling Compression for 3G Wireless Network” [11] and in Dean Willis, “SIP and 3G Wireless” [12].

2.6.2.3. Megaco

Megaco messages can be coded in text (pretty printed text or compact text encoding) or in binary (ASN.1 BER or ASN.1 PER). The binary encoding using ANS.1 PER result in the smallest average message size of 86 bytes and this ensures a reduced bandwidth. For this reason no compression is foreseen.

2.6.3. Resource Allocation/Reservation

2.6.3.1. H.323

H.323 supports resource reservation through RSVP. RSVP is the transport level signaling protocol for reserving resources in unreliable IP-based networks. Using RSVP, H.323 endpoints can reserve resources for a given real-time traffic stream based on its QoS requirements. When an endpoint requests admission with a Gatekeeper, it should indicate in the ARQ (Admission Request) message whether or not it is capable of reserving resources. The Gatekeeper should then decide, based on the information it receives from the endpoint and on information it has about the state of the network, either:

- to permit the endpoint to apply its own reservation mechanism for its H.323 session; or
- to perform resource reservation on behalf of the endpoint; or
- that no resource reservation is needed at all. Best-effort is sufficient.

This decision is conveyed to the endpoint in the ACF (Admission Confirm) message, in the "transportQoS" element. The endpoint shall accept the Gatekeeper's decision in order to place a call.

2.6.3.2. SIP

For resource reservation the SIP protocol relies on other protocols such as the Resource Reservation Protocol (RSVP). For more information how to support resource reservation, which is not described in the SIP RFC, there exist many papers and internet Drafts, which give a survey and introduction of methods. The following two papers are useful: Integration of Resource Management and SIP [42] and Extending SIP for QoS support [8].

2.6.3.3. Megaco

Procedures for resource reservation on the network are not supported by Megaco. The resource reservation is sent from the MGC to the MC for reserving the MG resources for media decoding and encoding for the given Stream(s) and Termination.

2.6.4. Conclusion on Resource and Bandwidth Efficiency

Considering Table 29 below it can be seen that SIP and H.323 both support the required criteria in a similar way, while Megaco does not.

Table 29: Summary on the resource and bandwidth efficiency criteria

	H.323	SIP	Megaco
Bandwidth Management	supported through H.225 (BRQ, BCF, BRJ)	not supported (optional bandwidth information through SDP)	not supported
Compression Gain	binary coding => no compression foreseen	compression => Internet Drafts available in the context of SIP for wireless devices	binary coding => no compression foreseen
Resource Reservation	H.225 signaling + RSVP	RSVP	not supported

2.7. Final Conclusion and Decision

The analysis presented above shows for each of the candidate protocols, though to a different degree, the necessity for both:

- **simplification** (in terms of reduction of unnecessary functionality, features, message set and contents) and
- **extension** of protocol functionality in order to incorporate non-existing features.

These two steps shall guarantee that the outcome is a protocol solution which is optimized for a MM satellite network.

H.323 is a complex set of protocols (H.225, H.245 and Q.931) with a considerable number of messages, states and interdependencies. H.323 provides most of the required functionality except the support of negotiation of all QoS parameters. It uses binary representation for its messages, which ensures a small message size, fast transfer of messages, lowest overhead and a good performance. A reduction of protocol complexity will require much effort due to the interactions between the different protocols. Extension of the protocol to support all OoS parameters may be a difficult challenge, because the extension is not foreseen and only possible if placeholders were defined. This is the biggest problem and obstacle to use H.323 for a MM satellite network.

Table 30: Conclusions on H.323

H.323 Advantages:	H.323 Disadvantages
+ similar to ESW L3 (ISDN origin)	- does not support all QoS parameter
+ support of required functionality	- high complexity, high message number
+ binary representation, small message size	- difficult extensibility
+ performance	- difficult protocol reduction

The advantage of SIP is the flexibility, simplicity and extensibility of the protocol standard. SIP supports the required functionality, except for the negotiation of QoS parameters. Incorporation of this feature requires an extension of the SIP or SDP protocol, which is possible in a straightforward way without changing the protocol standard described in the RFC 3261. SIP handles extensions in such a way that unknown headers within a message will be ignored and the call handling procedures are not affected. SIP is a simple, lightweight protocol, it uses only 6 request messages and during a call only 4 protocol states. The flexibility of message definition allows and facilitates the reduction of the protocol by omitting unnecessary message headers when applying it as a call control protocol for a MM satellite network. Since SIP messages are presented in textual form, a reduction of message sizes is called for to improve the performance and avoid wasting signaling resources, for this reason a compression mechanism shall be foreseen.

Table 31: Conclusions on SIP

SIP Advantages:	SIP Disadvantages
+ support of required functionality + low complexity, lightweight protocol + simple extensibility, including for QoS support + simple protocol reduction	- no explicit support for all QoS parameters - large message sizes, compression required

The Megaco protocol does not fulfill the basic requirements on the applicability of the protocol. The main functionality of the protocol is to control the MG from the MGC in a master and slave relationship. Also with respect to the other criteria it is not comparable with the other two candidates and does not support the basic functionality required for a call control protocol in a MM satellite network. Thus the effort would be too high in order to bring it to the same level. This is a sufficient reason to eliminate Megaco as a candidate protocol.

In conclusion, after consideration of all relevant criteria, SIP is considered the most attractive candidate protocol for call control in a MM satellite network. Certain adaptations and optimizations are however necessary for a satellite network environment. These will be specified in detail in section 3 of this study.

3. Adaptation of the VoIP Protocol for Satellite Network

3.1. Overview

The analysis of this study outlined the most suitable protocol: the SIP protocol. Before the usage of SIP for satellite network is ensured, some modifications and adaptations are necessary. The general intension is to adhere as far as possible to the SIP standard as described in [40]. Any deviations will have to be motivated and justified. The purpose of the current section is to provide reasons and justifications for deviations from the SIP standard. Two aspects are quite simple:

- **Extensions**: It is foreseeable, that some satellite specific extensions to standard SIP will be necessary e.g. due to the longer transmission delay of signaling messages over a satellite link, due to the fact that SIP does not support additional functionality required for call control in a multimedia satellite network (such as transport of certain QoS and traffic parameters) etc.
- **Simplification/Reduction**: In the call control prototype only a subset of the standard SIP functionality will be used. However, if in a future version of the satellite based SIP functionality is required, which is already supported by standard SIP, this has not to be specified in addition but can be simply modified and activated.

This section specifies the satellite specific modifications and extensions necessary to the standard SIP in more detail. The new developed protocol, which arise from the SIP standard and modified for MM satellite network purposes, is called "Satellite-SIP" in this document.

The "Satellite-SIP" protocol has to be optimized for connection control signaling in a satellite environment, the payload type of the connection is out of scope of this thesis.

3.2. Lower Layer Adaptation

3.2.1. Requirements on Lower Layer

SIP was designed originally as a VoIP signaling protocol, and the usual way to use SIP is within IP networks according to the specifications in the standard. The SIP standard specifies the following requirement for Layer 4: "all SIP elements must implement UDP and TCP. SIP elements may implement other protocols", such as SCTP. The reason of this specification is the limited message size supported by these underlying protocols. The SIP message size is limited only by the underlying transport protocol. Some datagram-oriented protocols such as UDP, which are commonly used by SIP applications, perform best with small message sizes. Because of this, the SIP specification limits the message size only when using such transport protocols. SIP applications using stream-oriented protocols, such as TCP, have no such limit. For UDP, this size of the maximum transmission unit (MTU) is limited to 65,535 bytes, including IP and UDP headers. The default Ethernet MTU amounts to 1500 bytes. If the SIP message size is larger than the MTU, a congestion controlled transport protocol, such as TCP, has to be used.

If, in some application it is not desirable to use SIP on top of TCP or UDP, then SIP can easily be adapted for another underlying layer, because the protocol specification makes minimal assumptions about the underlying transport and network layer protocols:

- The lower layer can provide either a packet or a byte stream service, and it can directly use any datagram or stream protocol with the only restriction that a whole SIP request or response has to be either delivered completely or not at all.
- The underlying protocol can be a reliable or unreliable protocol; no reliable protocol is needed, because SIP supports its own reliable mechanism through the state machine.
- The necessary naming conventions are described in the SIP standard ([RFC 3261]). Network addresses within SIP are not restricted to Internet addresses, but could be E.164 (PSTN) addresses, OSI addresses or private numbering plans.
- SIP provides two essential pieces of information to the underlying layer for the successful delivery: the data port of the other peer, and the address of the other peer. The port is necessary for successful delivery the data to the correct application, and the address is necessary for the successful delivery of data to the correct destination.

3.2.2. Lower Layer Adaptation-Options

In the VoIP environment the underlying layer for SIP is the layer 4 of OSI Model, the transport layer. The transport layer determines through the port number which application is requested. Under the layer 4 lies the network layer, which is responsible for the transfer of the message from one peer to another peer, determined through the IP Address. In the IP - Telephony environment Layer 4 can be UDP, TCP or SCTP, and IP used in the network layer. The following trade-off analyzes the possibilities of the underlying protocol stack and it offers suitable variants for a MM satellite network environment.

These four adaptations are considered:

- Option 1: Standard IETF protocol stack,
- Option 2: Adaptation Layer between SIP and Sat-L2
- Option 3: Sat-L2 adaptation to support SIP, and
- Option 4: SIP adaptation to support Sat-L2

They will be introduced and discussed in the following subsections.

3.2.2.1. Option 1: Standard IETF Protocol Stack

The most straightforward solution is to employ the usual protocol stack, i.e. to use TCP, UDP or SCTP in Layer 4, and IP in Layer 3. This option is depicted in Figure 6. It is noted that this solution also requires a (small) adaptation layer between IP and Sat-L2 to provide compatibility between the layers.

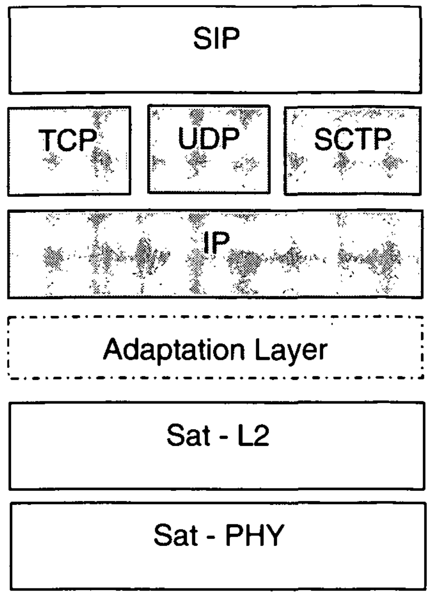


Figure 6: Option 1 - Standard IETF Solution

The advantages of this solution are that only satellite-network specific adaptations and additions are required in SIP, and for additional services the standard TCP/UDP/IP environment is available.
The main disadvantage concerns the overhead incurred by the use of the standard transport and network layer Internet protocols, where only a small subset of functions is in fact required in the present satellite network environment. Furthermore, an adaptation layer between IP and Sat-L2 is required in addition to the standard Internet layers (unless Sat-L2 already supports IP).
This solution also requires a decision as to which one of the three protocols on layer 4 should be used. Table 32 shows a comparison of the three protocols. Taking into account also the considerations in earlier sections, UDP is considered the preferred choice in this scenario. Its main advantage is its simplicity and the fact that the more complex features of TCP and SCTP, such as their connection orientation and reliability, are not needed in the present environment.

Table 32: Comparison of the Layer 4 protocols

	TCP	UDP	SCTP
Connection	connection oriented	connectionless	connection oriented
Reliability	reliable	unreliable	reliable
Complexity	High	Low	High
Data Presentation	Byte stream oriented	datagrams	datagrams
Order of Data	bytes stream ordered, but wait till all data are arrived, blocking other data	datagrams unordered, don't block other data	datagrams ordered, don't block other data

3.2.2.2. Option 2: Adaptation Layer between SIP and Sat-L2

The solution described in this section uses an adaptation layer between SIP and the satellite layer 2. The adaptation layer is an interface between SIP and the existing satellite layer 2, as shown in Figure 7 below.

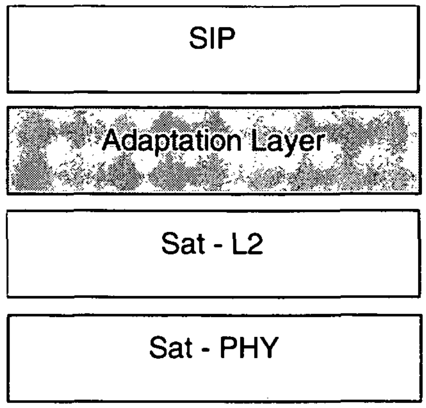


Figure 7: Option 2 - Adaptation Layer Solution

The adaptation layer takes into consideration the requirement from both layer (SIP and Sat-L2) and supports the interface between them. A key benefit is that no modification of Sat-L2 is necessary to support SIP compatibility, and that changes of the Sat-L2 have no effect on SIP. Furthermore, no modification of SIP (beyond satellite-network specific adaptations and modifications) is necessary to support Sat-L2 compatibility and changes of SIP have no effect to Sat-L2. This solution is open for new implementation of SIP or Sat-L2. In the adaptation layer support can also be provided for the compression of SIP messages in binary code for efficiency and performance improvement. A comparison between Figure 6 and Figure 7 shows that the role of the adaptation layer is essentially to comprise those functions from the Layer 3 and Layer 4 Internet protocols that are actually needed for the present environment. In this way, a lightweight and optimized protocol stack solution can be provided.

A disadvantage of this solution is that, for each modification of SIP or Sat-L2, the adaptation layer must be updated accordingly.

3.2.2.3. Option 3: Sat-L2 Adaptation to Support SIP

The solution described in this section proposes to modify the satellite layer 2 protocol such that it is able to fulfill the service requests it receives from SIP. This option, which is shown in Figure 8, requires the implementation of functionalities usually covered by Layer 4 protocols into the Sat-L2 protocol.

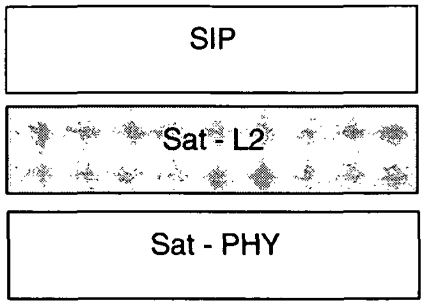


Figure 8: Option 3 - Satellite Layer 2 adaptation

The advantage of this solution is that no additional layer is required, thus yielding a protocol stack where SIP assumes the role of a traditional OSI Layer 3 protocol.

A severe drawback is however that the Sat-L2 protocol must be modified substantially to cover functionalities not usually provided by a Layer 2 protocol. This option also severely limits the upgradability of the protocol, and does not lead to a modular implementation.

3.2.2.4. Option 4: SIP Adaptation to Support Sat-L2

The solution described in this section employs an existing satellite layer 2 protocol and proposes to adapt SIP such that it complies with the requirements of Sat-L2. This scenario, shown in Figure 9 below requires substantial modifications of SIP well beyond those required from the general satellite network environment alone.

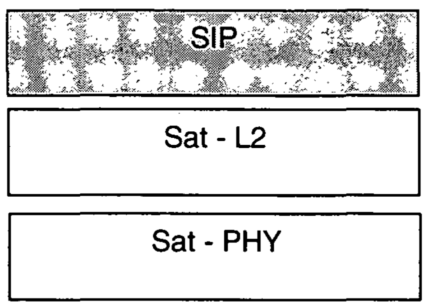


Figure 9: Option 4 - SIP adaptation for Sat-L2

The advantage of this solution is that no additional layer is required, thus yielding a protocol stack where SIP assumes the role of a traditional OSI Layer 3 protocol.

Analogous to Option 3, a severe drawback is however that SIP must be modified substantially to cover functionalities well beyond the originally indented ones. This option also severely limits the upgradability of the protocol (e.g. if Sat-L2 is changed), and does not lead to a modular implementation.

3.2.3. Conclusions on Protocol Stack Adaptation

In conclusion, the above trade-off analysis for the four protocol stack options reveals a clear preferred solution: Option 2. This solution introduces a lightweight adaptation layer between SIP and Sat-L2, with these capabilities:

- It provides to the satellite-optimized SIP layer an optimized minimal set of services usually offered by a transport layer protocol such as UDP, and
- It complies with the requirements of a given satellite-optimized Sat-L2 protocol.

As seen from the point of view of the satellite-optimized SIP layer, therefore, in terms of the services actually requested by SIP, the adaptation layer behaves exactly like a usual Internet transport layer. This provides an additional advantage during development and testing: Satellite-optimized SIP and the adaptation layer can be specified and implemented largely in parallel, with very little interactions necessary. In particular, for the development and testing of the satellite-optimized SIP protocol, a standard Layer 4 protocol such as UDP may be used.

3.3. Optimization of the Message Set

The optimization of the SIP message set starts from the standard SIP message set as described in the SIP standard RFC 3261 [40]. Optimization can occur in two ways:

- Through removal of messages (requests or responses) or message elements from the standard SIP which does not need to be supported in Satellite-SIP (since the functionality is not required). = Simplification/Reduction
- Through extension of the SIP standard in order to support additional functionality required for Satellite-SIP. = Extension

The first kind of the optimization is described in subsection 3.3.1 and the second kind is described in the subsection 3.3.2.

3.3.1. Message Set Reduction

One possibility for optimization of the Standard SIP Protocol is to optimize the available SIP message set (comprising SIP requests and responses) as described in RFC 3261 [40]. Other Requests for Comment, such as RFC 2976 [33] and RFC 3265 [41], will not be considered for Satellite-SIP. Thus, requests such as SUBSCRIBE, NOTIFY and INFO will not be taken into account, neither the responses “202 Accepted” and “489 Bad Event”.
This optimization through reduction is performed through removal of messages (requests or responses) or message elements from the standard SIP, which does not need to be supported in Satellite-SIP (since the functionality is not required).

3.3.1.1. Optimization though reduction of SIP Requests

The SIP Standard [RFC 3261] defines 6 requests, exactly called methods: This subsection analyzes the possibility of the optimization of this set of requests with respect to their applicability for Satellite-SIP:

- REGISTER Request: the REGISTER message will be part of the message set of the Satellite-SIP protocol in order to cover this functionality with the same protocol.

- INVITE Request: the INVITE method indicates that the user or service is being invited to participate in a session. It will be used in the Satellite-SIP protocol to initiate a session or to add one or more parties to an already existing session. The message will contain description of the session e.g. in terms of traffic and QoS parameters initiated by the calling terminal. The INVITE message will be sent to the Network Control Centre (NCC), where the Call Admission Control (CAC) is executed and it is decided whether to admit or reject the connection.
- ACK Request: In the SIP standard the ACK request confirms that a client has received a final response to an INVITE request. In the Satellite-SIP protocol the ACK request will be used in same way. In case of ptm connections, several ACK requests are possible in response to one INVITE message (one for each added party).
- CANCEL Request: according to the SIP standard a CANCEL request cancels a pending request, but does not affect a completed request, therefore a CANCEL request can only be sent before a final request (ACK) is sent. After the final request has completed the transaction, the CANCEL request is not usable. The CANCEL request will be used in the Satellite-SIP protocol in order to cancel an on-going session establishment, add party procedure.
- BYE Request: according to the Standard SIP the client user uses BYE to indicate to the server that it wishes to release the session. In the Satellite-SIP protocol this request will be used to release a session, to indicate that a party is leaving a session or to drop a party from a session.
- OPTIONS: the SIP request OPTIONS allows a UA to query another UA or a proxy server as to its capabilities. This allows a client to discover information about the supported methods, content types, extensions, codecs, etc. without "ringing" the other party. In the Satellite-SIP based call control prototype the OPTIONS request can in principle be used for several purposes:
  - as an admission request, for negotiation of QoS parameters, before sending of an INVITE message.(which is not feasible due to the long transmission delay),
  - for re-negotiation of QoS parameters during session establishment,
  - for re-negotiation of QoS parameters after the session was already established (e.g. in order to request more capacity).

The result of standard SIP request optimization outlined that all possible requests messages (REGISTER, INVITE, ACK, CANCEL, BYE and OPTIONS) will be supported in the Satellite-SIP protocol version. The use of all requests has to be adapted due to the needs of the satellite network.

3.3.1.2. Optimization through reduction of SIP Responses

A SIP response as defined in [40] is a SIP message sent from a server (UAS) to client (UAC) to answer to the received request. Each SIP response contains a Status-Code. The first digit of the Status-Code defines the class of response. The last two digits do not have any categorization role. For this reason, any response with a status code between 100 and 199 is referred to as a "1xx response", any response with a status code between 200 and 299 as a "2xx response", and so on.

SIP/2.0 in the standard [40] allows 6 values for the first digit and consequently this results 6 response classes:

- 1xx: Provisional - request received, continuing to process the request;
- 2xx: Success - the action was successfully received, understood, and accepted;
- 3xx: Redirection - further action needs to be taken in order to complete the request;
- 4xx: Client Error - the request contains bad syntax or cannot be fulfilled at this server;
- 5xx: Server Error - the server failed to fulfill an apparently valid request;
- 6xx: Global Failure - the request cannot be fulfilled at any server.

Response classes are divided into final and provisional/informational responses. A final response terminates a SIP transaction, as opposed to a provisional response that does not. Final responses are 2xx, 3xx, 4xx, 5xx and 6xx responses. The provisional response is used by a server to indicate a progress, but that does not terminate a SIP transaction, 1xx responses are provisional. This subsection analyzes the possibility for optimization of the response classes with respect to their envisaged use in a Satellite-SIP protocol version.

3.3.1.2.1. Provisional 1xx

Provisional responses indicate that the server contacted is performing some further action and does not yet have a definitive response. A server sends a 1xx response, if it expects that it will take more than 200 ms to obtain a final response (which is actually the case in a satellite network with a geo-stationary satellite). The following 5 provisional responses are defined in the SIP standard: 100 Trying, 180 Ringing, 181 Call Is Being Forwarded, 182 Queued, and 183 Session Progress.

For the Satellite-SIP version only the response 100 Trying is needed to indicate that a request is received and some unspecified action is being taken. Since the transmission delay between a terminal and the NCC in a MM satellite network with a GEO satellite is at least 250 ms, this response has to be sent from the NCC to the calling terminal when establishing a session. Furthermore it has to be sent from the called terminal to the NCC after receiving the INVITE message.

Other 1xx responses are not needed:

- 180 Ringing - indicates that the other party is being alert. This response will not be used in the Satellite-SIP protocol, because indication of ringing to the calling terminal is not necessary. Furthermore omitting this message can reduce the signaling traffic between NCC and the satellite terminals.
- 181 Call Is Being Forwarded - indicates that the call is being forwarded to a different destination. This response is not required and will not be used in the Satellite-SIP protocol, because multiple choices of destinations of the same terminal are not foreseen.
- 182 Queued - is used to indicate that the called party is temporary not available, but the server decides to queue the call rather than reject it. This response will not be used in the Satellite-SIP

protocol, because if a terminal is not available, the session has to be rejected immediately in order to avoid unnecessary blocking of traffic resources.

- 183 Session Progress - is used to convey information about the progress of the session that is not otherwise classified. This response is not required and will not be used in the Satellite-SIP protocol, since progress of session establishment is already indicated by using 100 Trying response.

3.3.1.2.2. Successful 2xx

This class of responses contains only one response message: 200 OK, which indicates that the request has succeeded. This response is required in the Satellite-SIP protocol. It will be sent from NCC to the calling terminal and from called terminal to NCC.

3.3.1.2.3. Redirection 3xx

3xx responses give information about the user's new location, or about alternative services that might be able to satisfy the session. The following 5 responses are defined in the SIP standard: 300 Multiple Choices, 301 Moved Permanently, 302 Moved Temporary, 305 Use Proxy, 380 Alternative Service.

For the Satellite-SIP version following responses will be supported:

- 301 Moved Permanently - the called user cannot be found at the address in the request. The response delivers a new address of the called user. The calling terminal can retry at the new address. This response will be used in the Satellite-SIP protocol, whereby the Satellite-SIP protocol only handles terminal addresses, not user the addresses behind a terminal. The address of a terminal can be changed, if it registers again at the NCC, therefore the NCC can provide the new address.
- 380 Alternative Service will be supported to indicate that the requested service was not available, but alternative services are available.

Other 3xx responses are not needed:

- 300 Multiple Choices - the called user is reachable in several locations, the calling terminal can select a preferred communications end point and redirect its request to that location. Since different locations of a user behind a satellite terminal will not be communicated to the calling user, this response will not be used in the Satellite-SIP protocol.
- 302 Moved Temporary - the called user is not reachable at the requested address, the new address returned in the response can be retrieved by the calling terminal. This response will not be used in Satellite-SIP protocol, because a satellite terminal cannot change its address or location temporarily.
- 305 Use Proxy - the calling user has sent a request to a registrar server or to a redirect server, which answers to send the request to the proxy returned in response. The proxy functionality is covered by the NCC in the satellite network. For the first protocol version, where only one

NCC is available, no other entities are able to receive requests; therefore this response will not be used.

3.3.1.2.4. Client Error (Request Failure) 4xx

4xx responses are definite failure responses from a particular server. The client should not retry the same request without modification (for example, adding appropriate authorization). In the SIP standard 28 responses are defined.

For the Satellite-SIP protocol version following 4xx responses will be supported:

- 400 Bad Request - the request could not be understood due to malformed syntax. This response will be used in the Satellite-SIP protocol by NCC or by a terminal, if the received request contains syntax failures or/and cannot be understood. This response will be sent from NCC to a terminal or from a terminal to NCC.
- 401 Unauthorized - the request requires user authentication, is send from redirect or registrar server to the SIP user to challenge the identity of the user. This response will be supported by the Satellite-SIP protocol in the registration procedure of the satellite terminal.
- 403 Forbidden - the server understood the request, but is refusing to fulfill it. Authorization will not help, and the request should not be repeated. The NCC will use this response to decline the request of a satellite terminal (e.g. if the terminal of a specific type has not been authorized to request such a service).
- 404 Not Found - the server has definitive information that the user does not exist at the domain specified in the request. This response will be used in the Satellite-SIP protocol, if the NCC could not resolve the terminal address, either because the address of the terminal does not exist or the address is not complete or the address is false. This response will be sent from NCC to the calling terminal.
- 408 Request Timeout - the server could not produce a response within a suitable amount of time. The client may repeat the request without modifications at any later time. This response will be used in Satellite-SIP by NCC or by called satellite terminal to indicate that the request timed out due to NCC internal reasons (e.g. overload conditions). This response will be sent from NCC to the calling terminal or from called terminal to NCC.
- 416 Unsupported URI Scheme - the server cannot process the request because the scheme of the URI in the Request-URI is unknown to the server. This response will be used in the Satellite-SIP protocol, by NCC to indicate, if the requested address scheme is not supported, or is too large or is not completed, etc. This response will be sent from NCC to the calling terminal.
- 481 Call/Transaction Does Not Exist - this status indicates that the UAS received a request that does not match any existing dialog or transaction. This response will be used in the Satellite-SIP by NCC or by a terminal, if the received request is not part of an existing session. This response indicates an unrecognized session. This response will be sent from NCC to the terminal or from terminal to NCC.

- 486 Busy Here - the calling user's end system was contacted successfully, but the calling user is currently not willing or able to take additional calls at this end system. The response will be used in the Satellite-SIP to indicate, if the satellite terminal is busy or cannot take an additional call. This response will be sent from NCC to the terminal or from terminal to NCC.
- 487 Request Terminated - the request was terminated by CANCEL request. This response will be used in Satellite-SIP, if the calling terminal, through the CANCEL request, terminates the existing INVITE request, before the call is connected, then the NCC will send this response to the called terminal to inform the called terminal about the terminated session initiation. The calling terminal will also get this response as acknowledgment for successful cancellation of the session initiation. This response will be sent from NCC to the terminal.
- 488 Not Acceptable Here - the response has the same meaning as 606 (Not Acceptable), but only applies to the specific resource addressed by the Request-URI and the request may succeed elsewhere. The user's agent was contacted successfully but some aspects of the session description such as the requested media, bandwidth, or addressing style were not acceptable. This response will be used in Satellite-SIP protocol to indicate that some QoS parameters cannot be provided. This response will be sent from NCC to the terminal.

Other 4xx responses are not needed:

- 402 Payment Requirement - reserved for future use. This response will not be used in the first version of the Satellite-SIP protocol.
- 405 Method Not Allowed - the method (request) specified in the Request-Line is understood, but not allowed for the address identified by the Request-URI. This response will not be used in the Satellite-SIP protocol. If the method, for example, is not implemented the response "501 Not Implemented" will be used. If a service must not be requested by a terminal, the response "403 Forbidden" will be used.
- 406 Not Acceptable - the resource identified by the request is only capable of generating response entities that have content characteristics not acceptable according to the Accept header field sent in the request. The response is not needed and will not be used in the Satellite-SIP protocol, because the response "488 Not Acceptable Here" will be used, if the QoS description is not acceptable.
- 407 Proxy Authentication Required - this code is similar to 401 (Unauthorized), but indicates that the client must first authenticate itself with the proxy. This response is used by the proxy server as response to an INVITE request to challenge the identity of the user. This response will not be included in the Satellite-SIP protocol specification, because only the registration procedure will be authenticated, not the session setup procedure. Using the authentication during the session initiation would increase the signaling traffic and therefore Satellite-SIP will only support the authentication during the registration procedure of a satellite terminal.
- 410 Gone - the requested resource is no longer available at the server and no forwarding address is known for the moment. This response will not be used in the Satellite-SIP, as

temporary availability of satellite terminals is not foreseen. If a terminal is unavailable (for what reason ever) the response "404 Not Found" will be used.

- 413 Request Entity Too Large - the server is refusing to process a request because the request entity-body is larger than the server is willing or able to process. This response is not needed, because the messages sizes in Satellite-SIP will be optimized.
- 414 Request-URI Too Long - the server is refusing to service the request because the Request-URI is longer than the server is willing to interpret. This response is not needed in Satellite-SIP, because if the Request-URI does not match (incomplete/to large/not supported URI/etc) the response "416 Unsupported URI Scheme" will be used.
- 415 Unsupported Media Type - the server is refusing to service the request because the message body of the request is in a format not supported by the server for the requested method. This response will not be used in the Satellite-SIP version, because the media type, which will be transported after the session establishment as payload data, is not part of the Satellite-SIP protocol.
- 420 Bad Extension - the server did not understand the protocol extension specified in a Proxy-Require or Require header field. This response will not be used because the Satellite-SIP protocol version will not support any additional extensions, only the defined protocol specifications.
- 421 Extension Required - the UAS needs a particular extension to process the request, but this extension is not listed in a Supported header field in the request. This response will not be used, because the Satellite-SIP protocol version will not support any additional extensions.
- 423 Interval Too Brief - the server is rejecting the request because the expiration time of the resource refreshed by the request is too short. The registrar in a registration procedure uses this response. This response will not be used in the Satellite-SIP protocol during the registration procedure of a satellite terminal, because the satellite terminal has no re-registration interval. The satellite terminal will register itself by joining to satellite network and un-register itself by leaving the satellite network; the re-registration procedure of SIP will not be used.
- 480 Temporarily Unavailable - the calling terminal's end system was contacted successfully but the calling terminal is currently unavailable. At NCC level only the availability of the terminal not the users behind a terminal can be checked. If the called terminal does not respond to an offered session request since he is busy or not available, response 486 (Busy here) will be used.
- 482 Loop Detected - the server has detected a loop. This response will not be used in the Satellite-SIP protocol, because due to satellite network topology, where each request must first pass through NCC this response is not applicable.
- 483 Too Many Hops - the server received a request that contains a Max-Forwards header field with the value zero. This response is not needed in Satellite-SIP, because each request makes only a single hop from terminal to NCC and will not be transported from one NCC to another.
- 484 Address Incomplete - the server received a request with a Request-URI that was incomplete. This will not be supported in the Satellite-SIP, because the response "416

Unsupported URI Scheme" will be used, if the address (Requested-URI) of the satellite terminal is not supported or not completed.

- 485 Ambiguous - the Request-URI was ambiguous. The address of the satellite terminal will always be unique and never ambiguous, therefore this response will not be used in the Satellite-SIP protocol.
- 491 Request Pending - the request was received by a UAS that had a pending request within the same dialog. (Both parties send each other a re-INVITE simultaneously during the connection) This response is not needed and will not be used, because a request in Satellite-SIP will always be sent to NCC and not direct to the other satellite terminal.
- 493 Undecipherable - the request was received by a UAS that contained an encrypted MIME body for which the recipient does not possess or will not provide an appropriate decryption key. This response will be not used in Satellite-SIP, since encryption/decryption is out of scope of the protocol.

3.3.1.2.5. Server Error (Server Failure) 5xx

5xx responses are failure responses indicating server errors. The following 7 responses are defined: 500 Server Internal Error, 501 Not Implemented, 502 Bad Gateway, 503 Service Unavailable, 504 Server Time-out, 505 Version Not Supported, 513 Message Too Large.

For the Satellite-SIP protocol version following 5xx responses will be supported:

- 500 Server Internal Error - the server encountered an unexpected condition that prevented it from fulfilling the request. This response is needed and will be used in Satellite-SIP to indicate an internal failure in NCC. The response will be sent from NCC to the terminal.
- 501 Not Implemented - the server does not support the functionality required to fulfill the request. This is the appropriate response when a UAS does not recognize the request method and is not capable of supporting it for any user. In Satellite-SIP this response will be sent from the NCC to the terminal in response to an unsupported method/request.
- 505 Version Not Supported - the server does not support, or refuses to support, the SIP protocol version that was used in the request. The NCC will send this response to a calling satellite terminal to indicate that the requested SIP protocol version is not supported.

Other 5xx responses are not needed:

- 502 Bad Gateway - the server, while acting as a gateway or proxy, received an invalid response from the downstream server it accessed in attempting to fulfill the request. This response is not needed and will not be supported in the Satellite-SIP protocol version, because the Satellite-SIP protocol will not be used for a server-to-server communication.
- 503 Service Unavailable - the server is temporarily unable to process the request due to a temporary overloading or maintenance of the server. This response will not be used in the satellite network. If the NCC is unable to process the request due to overload conditions, the response "408 Request Timeout" will be used.

- 504 Server Time-out - the server did not receive a timely response from an external server it accessed in attempting to process the request. This response will not be used in the Satellite-SIP protocol, because the response "408 Request Timeout" available will be used in the Satellite-SIP version for rejection due to timeout.
- 513 Message Too Large - the server was unable to process the request since the message length exceeded its capabilities. This response will not be used in the satellite network, because the Satellite-SIP protocol messages will be optimized and compressed and message length will be checked at sending side.

3.3.1.2.6. Global Failures 6xx

6xx responses indicate that a server has definitive information about a particular user, not just the particular instance indicated in the Request-URI. The following 4 responses are defined: 600 Busy Everywhere, 603 Decline, 604 Does Not Exist Anywhere, 606 Not Acceptable.

For the Satellite-SIP protocol version only one 6xx response is foreseen:

- 603 Decline - the calling user's machine was successfully contacted but the user explicitly does not wish to or cannot participate. In case of the satellite network the NCC does not have any information about the satellite user and will therefore not use this response. However, if the user behind the terminal does not accept the call, this response can be sent from the called terminal to NCC (optionally).

Other 6xx responses are not needed:

- 600 Busy Everywhere - the calling user's end system was contacted successfully but the calling user is busy and does not wish to take the call at this time. This response will not be used, because the NCC does not have any information whether the satellite user wishes or does not want to accept the call. If the terminal or the user behind the terminal is busy or does not accept the call, the response "486 Busy Here" will be used in the Satellite-SIP protocol.
- 604 Does Not Exist Anywhere - the server has authoritative information that the user indicated in the Request-URI does not exist anywhere. In the satellite network this response will not be used. If the satellite terminal or the NCC have information that the satellite user does not exist, the response "404 Not Found" will be used, to indicate that called number does not exist.
- 606 Not Acceptable - the user's agent was contacted successfully but some aspects of the session description such as the requested media, bandwidth, or addressing style were not acceptable. This response is not needed in the Satellite-SIP protocol, because the response "488 Not Acceptable Here" will be used in Satellite-SIP, if the session description is not acceptable.

3.3.1.2.7. Result of Response Optimization

The Table 33 below shows the results of the optimization of the standard SIP response messages.

Table 33: Results of Standard SIP Response Optimization

Response Classes	Not required response messages for Satellite-SIP protocol	Required response messages for Satellite-SIP protocol
1xx	180 Ringing, 181 Call Is Being Forwarded, 182 Queued, 183 Session Progress	100 Trying
2xx		200 OK
3xx	300 Multiple Choices, 302 Moved Temporarily, 305 Use Proxy	301 Moved Permanently, 380 Alternative Services
4xx	402 Payment Requirement, 405 Method Not Allowed, 406 Not Acceptable, 407 Proxy Authentication Required, 410 Gone, 413 Request Entity Too Large, 414 Request-URI Too Long, 415 Unsupported Media Type, 420 Bad Extension, 421 Extension Required, 423 Interval Too Brief, 480 Temporarily Unavailable, 482 Loop Detected, 483 Too Many Hops, 484 Address Incomplete, 485 Ambiguous, 491 Request Pending, 493 Undecipherable	400 Bad Request, 401 Unauthorized, 403 Forbidden, 404 Not Found, 408 Request Timeout, 416 Unsupported URI Scheme, 481 Call/Transaction Does Not Exist, 486 Busy Here, 487 Request Terminated, 488 Not Acceptable Here
5xx	502 Bad Gateway, 503 Service Unavailable, 504 Server Time-out, 513 Message Too Large	500 Server Internal Error, 501 Not Implemented, 505 Version Not Supported
6xx	600 Busy Everywhere, 604 Does Not Exist Anywhere, 606 Not Acceptable	603 Decline
Number of Responses	32	18

32 responses can be omitted and only 18 response messages are part of the first call control protocol version of the Satellite-SIP protocol specification.

3.3.2. Message Set Extensions

This subsection analyzes the extension of the SIP standard message set by new messages and functionality, which are required for Satellite-SIP. The SIP standard does not support messages and consequently the functionality of following two services, which are required in the satellite network:

1. Support of ptm connection with add party and drop party functionality.

2. Support of transport the QoS parameters for requesting, modifying, updating and clearing the satellite resources for the connection.

This subsection investigates the possible solutions for this mentioned functionality.

3.3.2.1. Ptm Support

The first important functionality required in a call control protocol for satellite network is the capability to establish and release a ptp and a ptm sessions and to add and drop parties from existing ptm sessions. Ptp functionality is well defined in the SIP standard. Ptm sessions in the sense of the SIP standard are established in a step-wise manner as several ptp sessions via a conference bridge. The SIP standard does not define separate messages for adding and dropping parties from such a point-to-multipoint connection. For creating multipoint connection by adding a party to the connection an INVITE message is used, which is sent through the proxy to the desired user. Functionality such as dropping a party from the multipoint connection is not supported through the SIP standard, each user (calling or called user) can leave the multipoint connection though sending the BYE message, without terminating the connection, but the originator/calling user of the multipoint connection has no possibly to drop undesired user.

In the SIP standard the connection characteristics (i.e. whether the connection is uni- or bi-directional, point-to-point or point-to-multipoint) are handled through the SDP protocol in the message body:

- For a point-to-point connection, the SDP protocol parameter "a" specifies whether a connection is uni- or bi-directional: if "a=sendonly" then only a uni-directional connection is required, if "a=sendrecv", then a bi-directional connection is required.
- For a point-to-multipoint connection a conference bridge or a multicast address is needed. This point, where the RTP streams are handled and join together, is defined in the SDP message body. SIP signaling messages are directly exchanged between the users, but the payload connection is between the user and this defined point. The calling user sends an INVITE with SDP description of the multicast address to the called user; the called used answers with 180 and 200 to the calling user. The payload connection will be switched from the called user to the multicast address and not to the calling user. In this way the signaling handling is separated from the payload handling.

For the Satellite-SIP protocol the following two options are considered for signaling the connection characteristics (ptp, ptm, uni, bi-directional):

1. <u>Using SIP with SDP:</u> this solution allows the usage of SIP for call control signaling in the message header and SDP for describing the connection characteristics in the message body. The standard SIP procedure for establishment of a ptm connection is not suitable in a MM satellite network, because the SIP server (proxy server) does not directly handle and join together the traffic streams. In the satellite network, NCC handles both call control signaling and payload switching commands. The signaling traffic makes a double hop and passes through the NCC, whereas the traffic only makes a single hop and is through-connected at the satellite. Therefore

an extension to the SIP and SDP standard is needed to provide multipoint connections in a suitable way. Additionally an extension of the SIP protocol is required in order to support a drop party functionality. The SDP protocol also needs an extension to provide the transport of satellite specific QoS parameters.

2. Using SIP without SDP: in this option, SIP is used for signaling and the body of the message contains satellite specific information, such as the connection profile with service category, connection configuration/symmetry etc. This would be a more suitable solution, because only adaptation of the SIP protocol is necessary, no changes and adaptation of SDP is needed. For adding a party to an existing connection an INVITE message with the same "Call-ID", can be used to identify to which connection the new satellite terminal/user should be added. If the called terminal wants to leave the connection, it will use the BYE message. If the calling terminal wants someone to leave the connection, an extension of the BYE message header parameter is needed. Table 34 summarizes the functionality required for Satellite-SIP. In some lines, where the word "extension" is present, also a possible solution is included. In this solution the message body is used to provide QoS satellite specified parameters.

Table 34: Required Functionality of Satellite-SIP

Functionality	**Satellite-SIP requests**	**Satellite-SIP responses**
ptp establish	INVITE	200 (OK), ...
ptp release	BYE	200 (OK), ...
ptm establish	extension => INVITE with more then one receiver (more than one terminals for the "To" header field or more than one "To" header fields)	Extension
ptm add one party	extension => INVITE, "To" header contains the new terminal, using the same Call-ID to identifies the session (Note 1)	200 (OK), ...
ptm drop one party	extension => BYE with "To" header contains the terminal, which should be dropped, Call-ID identifies the session (Note 1)	200 (OK), ...
ptm leave	BYE	200 (OK), ...
ptm release	extension => BYE with one or more receiver in the "To" header field or more than one "To" header fileds, depending on the number of communicating terminals. (Note 2)	Extension

Note 1: add and drop party functionality is only for the originator of the session allowed, that is only the calling terminal can add or drop parties.

Note 2: ptm release functionality is only for the originator of the session allowed, that is only the calling terminal can release the ptm session. ptm leave functionality is for each terminal applicable to leave the session. However if the originator leaves the session, the ptm session will be released.

The second discussed variant is most suitable one, because only the extension of the SIP Standard is necessary and no usage of DSP protocol is needed. Thus the message body of the Satellite-SIP messages is free for other purposes (e.g. for transfer of QoS parameters, including the connection configuration/symmetry parameters such as ptp, ptm, uni or bi-directional).

3.3.2.2. QoS support

In the SIP standard there are no messages for requesting, modifying, updating and clearing satellite specified resources before, during and after the session. The problem of resource negotiation can be solved by adaptation the INVITE and the OPTIONS messages in the following way:

- The INVITE message, which is sent to the NCC during session initiation, will contain the desired traffic and QoS parameters in the Satellite-SIP message body. Based on the information contained in this message a (statistical) check for the availability of resources and the required QoS will be made. If the NCC cannot provide the desired resources with the requested QoS, the NCC will send a 488 (Not Acceptable Here) and reject the request. The 488 response body may contain the information about the available resources and QoS, which the NCC is prepared to provide. If the NCC can provide the resources with the requested QoS, the session establishment procedure will be continued and an INVITE request will be sent to the called satellite terminal.
- If resource re-negotiation is necessary for an already established connection the OPTIONS request can be used. The OPTIONS request will contain the same "Call-ID" header parameter to identify and assign the requests to the same connection and the message body will also contains the QoS parameters. The response for the OPTIONS request can be 200 (OK) as acknowledgment for resource changes or 488 (Not Acceptable) as rejection.
- Upon reception of a BYE request, the NCC initiates release of the resources assigned for that session. An additionally message for resource clearing is not necessary and will only increase the signaling overhead.

The same messages for resource negotiation will be used for all types of session, independent of the connection symmetry. More about the QoS parameters see next subsection 3.4 and 3.4.2.

3.4. Optimization of the Parameter Set

A SIP message is composed of (1) a message (request or response) line, (2) message header and (3) the message body. In this section, considerations will be made about optimizing the (2) message header, which contains the SIP message parameters. The (1) message (request or response) line optimization has already been described in subsection 3.3. The standard SIP (3) message body contains the session description through the SDP protocol (the media description). In a satellite network media parameters for a connection are not required, but the traffic and the QoS parameters are necessary to be negotiated. Therefore the next subsection 3.4.2 will optimize the (3) message body and in this way analyze the container for traffic and QoS parameters.

The SIP message parameters, actually called the header fields, consist of a field name followed by a colon (":") and the field value. In the SIP standard [40] 44 header parameters are specified, which will be optimized for the Satellite-SIP protocol in subsection 3.4.1. The optimization will outline for the Satellite-SIP protocol the header fields required from the SIP standard and also introduce the header fields from the SIP standard, which are not needed for satellite purposes and therefore they can be removed = Simplification/Reduction. After this optimization, subsection 3.4.2 will analyze whether additional header parameters are required, which are currently not supported by the SIP standard = Extension.

3.4.1. Parameter Set Reduction

The parameter set reduction is performed through removal of header fields of messages from the standard SIP which does not need to be supported in Satellite-SIP (since the functionality is not required).

The following parameters are required for the Satellite-SIP Protocol:

- *Accept*: this header field indicates the format, which is used in the message body. If no accept header field is present, the default format of the message body is application/sdp. In Satellite-SIP, the accept header parameter will be used in messages: INVITE, OPTIONS, 200, 488 to indicate, that the content of the message body is not default application/sdp, but satellite specific (for transport of QoS parameters). If the SDP protocol will not be used in Satellite-SIP, then the present of this header in each message will indicate that the message body is not empty and the value of this header will indicate the used format of the message body. If this header is not present in a message, then the message body of this message is empty.
- *Allow:* this header field lists the set of methods supported by the UA generating the message. Example: Allow: INVITE, ACK, REGISTER, OPTIONS, CANCEL, BYE. This header field will be used in Satellite-SIP in the response 501 (Nor Implemented) to indicate the supported methods.
- *Authentication-Info:* this header field provides information for mutual authentication with HTTP Digest. This header parameter will be included in the Satellite-SIP protocol specification and is only mandatory included in a 200 (OK) response to a REGISTER request that was successful, using digest based on the Authorization header field.
- *Authorization*: this header field contains authentication credentials of a UA. This header parameter will be included in the Satellite-SIP protocol specification. The Authorization header field contains authentication credentials of a satellite terminal. This header is only mandatory and only used in the REGISTER request sent from a satellite terminal to the NCC for authentication challenge.
- *Call-ID:* this header field uniquely identifies a particular invitation. In the Satellite-SIP protocol version this header will uniquely identify the session and will be mandatory in all messages.
- *Cseq:* this header field in a request contains a single decimal sequence number and the request method. The Cseq header field serves to order transactions within a dialog, to provide a means to uniquely identify transactions, and to differentiate between new requests and request retransmissions. This header field will be used in Satellite-SIP to identify another transaction within a session. For example: first INVITE with "Cseq: 1 INVITE" and then the second INVITE with "Cseq: 2 INVITE" identifies a second transaction to the same session and not a

retransmission of a request. The Call-ID of both INVITE messages is identical, which allows relating them to the same session. This is a mandatory field for all Satellite-SIP messages.

- *Contact:* this header field is used with the registration procedure and indicates the URI, where the user wants to be reachable. This header parameter will be supported by the Satellite-SIP protocol, to indicate the terminal address in the response 301 Moved Permanently.
- *Error-Info:* this header field provides a pointer to additional information about the error status response. This header will be used in the Satellite-SIP protocol version to specify in a response message the cause of the rejection of the certain request. This so called "reject cause" is satellite specific and not a SIP standard value. This optional header field will not be used in each response, only if it is required.
- *Expires:* this header field gives the relative time after which the message (or content) expires and is used in the registration procedure to indicate the lifetime of the registration. This header field is not required for the registration procedure of a satellite terminal, but for the de-registration procedure. If a terminal send a REGISTER request with Expire = 0, then this request correspond to the de-registration procedure.
- *From:* this header field indicates the initiator of the request. In the Satellite-SIP protocol version the From header field will contains the calling satellite terminal identifier. This is a mandatory heard field for all messages.
- *Retry-After:* this header field can be used with a 500 or 503 response to indicate how long the service is expected to be unavailable to the requesting client and with a 404, 413, 480, 486, 600, or 603 response to indicate when the called party anticipates being available again. The Retry-After header field can be optionally used in Satellite-SIP protocol with a 500 (Server Internal Error) response to indicate how long the service is expected to be unavailable to the requesting UAC and with a 486 (Busy Here) response to indicate when the invited terminal anticipates being available again.
- *To:* this header field specifies the logical recipient of the request. In the satellite environment, this header field will contain the satellite terminal address, which identifies the satellite terminal. This is a mandatory heard field for all messages. A Satellite-SIP message may contain more than one To header fields or more than one terminals listed in one To header fields for ptm sessions.
- *WWW-Authenticate:* this header field value contains an authentication challenge. This header will be part of the Satellite-SIP protocol specification. The WWW-Authenticate header field value contains an authentication challenge from the NCC. This header is only used and mandatory in the 401 (Unauthorized) response sent from the NCC to satellite terminal.

The following parameters are not needed in the Satellite-SIP Protocol:

- *Accept-Encoding*: this header field is similar to Accept, but restricts the coding of the content, that are acceptable in the response. This parameter will not be used in Satellite-SIP, because the message body will only transport SDP for session description or satellite specified for QoS. No transport of other data is foreseen, and therefore the acceptable description of coding of data is not needed.

- *Accept-Language*: this header field is used in requests to indicate the preferred languages for reason phrases, session descriptions, or status responses carried as message bodies in the response. This header will not be used in Satellite-SIP, because the default language will be English as defined in the standard, no support of other language will be provided.
- *Alert-Info:* when present in an INVITE request, the Alert-Info header field specifies an alternative ring tone to the UAS. When present in a 180 (Ringing) response, the Alert-Info header field specifies an alternative ringback tone to the UAC. Using this parameter to defined the ringback tone is not required in the satellite network, therefore will not be used.
- *Call-Info:* this header field provides additional information about the calling or called user, depending on whether it is found in a request or response. (For example: web page of the calling or called user). This header will not be used in the Satellite-SIP, because Satellite-SIP will be primarily used for call control purposes between NCC and terminals.
- *Content-Disposition:* this optional header field describes how the message body or, for multipart messages, a message body part is to be interpreted. For example the value "session" indicates that the body part describes a session for a call, the value "render" indicates that the body part should be displayed to the user. This header field will not be used, because the Satellite-SIP protocol will only be used as a call control protocol, not for other purposes.
- *Content-Encoding*: this header field indicates what additional content codings have been applied to the entity-body, and thus what decoding mechanisms must be applied in order to obtain the media-type. This header parameter will not be used in Satellite-SIP, because Satellite-SIP will not transport other data, especially default SDP or satellite specific data in the message body.
- *Content-Language:* this header filed indicates the used language for the session description in the message body. This header is not required since only English language will be supported.
- *Content-Length*: this header field indicates the size of the message-body, in decimal number of octets. This header filed is mandatory used in SIP standard, but in satellite network is it not primary needed and one header field more in a Satellite-SIP message, would only increase the message size, therefore will not be used in Satellite-SIP.
- *Content-Type:* this header field indicates the media type of the message-body sent to the recipient, if the message body is not empty. If the message body is empty, then this header field is not present in the message. This header field will not be used in Satellite-SIP, because the header field Accept will indicate the content type of the message body, if the message body is not empty.
- *Date:* this header field contains the date and time and reflects the time when the request or response is first sent. This header parameter is not required and will not be used in the Satellite-SIP version.
- *In-Reply-To:* this header field enumerates the Call-IDs that this call references or returns, this allows call distribution systems to route return calls to the originator of the first call. This

header filed will not be used in Satellite-SIP, because call routing is not a required functionality.

- *Max-Forwards:* this mandatory header field must be used with any SIP method to limit the number of proxies or gateways that can forward the request to the next downstream server. This header will not be used in Satellite-SIP, because a SIP message will only be sent from satellite terminal to NCC or in reverse direction, therefore the number of hops is limited to one.
- *Min-Expires:* this header field conveys the minimum refresh interval supported for soft-state elements managed by that server, mandatory used by registrar for registration expire interval. This header field will not be used in Satellite-SIP during registration of a terminal in the NCC, because the satellite terminal registration does not contain a registration interval.
- *MIME-Version:* this header field indicates the used MIME version. The format of the message body will not be defined through MIME, because this would increase the message header and the message body will only be a container for defined satellite specific parameters, therefore this header field is not needed in Satellite-SIP.
- *Organization:* this header field conveys the name of the organization to which the SIP element issuing the request or response belongs. This header field is not required for call control, therefore it will not be used in Satellite-SIP.
- *Priority:* this header field indicates the urgency of the request as perceived by the client, value can be for example "emergency" and is address to the receiving human. This header field will not be used in Satellite-SIP, because the receiver of the connection is a satellite terminal, never a human.
- *Proxy-Authenticate:* this header field value contains an authentication challenge in a session setup procedure with the proxy server, and has a usage in responses 407 and 401. This header field will not be used in Satellite-SIP protocol, because the authentication will only be a part of a registration procedure, not of the call setup procedure.
- *Proxy-Authorization:* this header field allows the client to identify itself (or its user) to a proxy that requires authentication. Authentication will be only part of the registration procedure in the Satellite-SIP protocol, therefore this header field, which is required by proxy for the session initiation functionality, will not be supported by Satellite-SIP.
- *Proxy-Require:* this header field is used to indicate proxy-sensitive features that must be supported by the proxy. This header field will not be supported by Satellite-SIP, because no additionally features are required by the NCC or terminal.
- *Record-Route:* this header field is inserted by proxies in a request to force future requests in the dialog to be routed through the proxy. In Satellite-SIP all requests and response will be routed from one satellite terminal to the NCC, therefore this heard field is not needed and will not be used in Satellite-SIP.
- *Reply-To:* this header field contains a logical return URI that may be different from the From header field. This header field will not be used in Satellite-SIP, because the satellite terminal address always will be identified in the From header field.

- *Require:* this header field is used by UACs to tell UASs about options that the UAC expects the UAS to support in order to process the request. The order to process the request will be unique defined in the Satellite-SIP protocol; the signaling of the order to process is therefore not needed. This header will not be used.
- *Route:* this header field is used to force routing for a request through the listed set of proxies. In Satellite-SIP protocol version this header field will not be used, because routing of requests is not required.
- *Server:* this header field contains information about the software used by the UAS to handle the request. This information, what software is used by the requesting part is not needed in the satellite network, therefore this header field will not be used.
- *Subject:* this header field provides a summary or indicates the nature of the call, example: Subject: come for a coffee today. Indication of the subject of a session to call control is not foreseen. This header will not be used in Satellite-SIP.
- *Supported:* this header field enumerates all the extensions supported by the UAC or UAS. This header field will not be used in Satellite-SIP, because the Satellite-SIP will not support any SIP extensions.
- *Timestamp:* this header field describes when the UAC sent the request to the UAS. This header field is not required in the satellite environment and will not be used in Satellite-SIP.
- *Unsupported:* this header field lists the features not supported by the UAS, used header in 420 (Bad Extension) response. This header field will not be used in Satellite-SIP, because the response 420 will not be used in Satellite-SIP.
- *User-Agent:* this header field contains information about the UAC originating the request, example: User-Agent: Softphone Beta1.5. This additional information is not required in the NCC or terminal. This header will not be used.
- *Via:* this header field indicates the path taken by the request so far and indicates the path that should be followed in routing responses and is used by proxies to detect loops. Indication of the path, which is taken by the request, is not needed in the satellite network; because only one possible path is available, not many paths such in IP network are possible.
- *Warning:* this header field is used to carry additional information about the status of a response. This header will not be used in Satellite-SIP, because the status code of the response message is sufficient to indicate the response reason.

Table 35 shows result of the optimization of SIP headers and their usage in the Satellite-SIP messages.

Table 35: Header Fields for Satellite-SIP

Name of the Header	Description of Satellite-SIP Header	Usage of the Header in Satellite-SIP	
		Header for Call Control	Header for Registration/Authentication
Accept	indicates the format of message body	Mandatory in all messages, where the message body is not empty	
Allow	lists the set of supported methods	Mandatory in 501 (Not Implemented) response	
Authentication-Info	provides information for authentication		optional in 200 (OK) response
Authorization	contains authentication credentials		Mandatory REGISTER request
Call-ID	uniquely identifies the session	Mandatory in all messages	
Contact	indicates the terminal address	Mandatory in Moved Permanently Response	
Cseq	identifies another transaction within a session (distinction between a new transaction and a retransmission possible)	Mandatory in all messages	
Error-Info	identifies the satellite specific "reject reason", (additional information to the response)	optional in all responses; used only, if some satellite specified "reject cause" available	
Expires	indicates the de-registration		mandatory REGISTER request
From	identifies the calling satellite terminal	mandatory in all messages	
To	identifies the called satellite terminal	mandatory in all messages	
WWW-Authenticate	contains an authentication challenge		mandatory in 401 response

3.4.2. Parameter Set Extension

This subsection describes the required satellite specific header parameters, which are not covered by the SIP standard. The following Connection Profile parameters for QoS are not supported through the SIP standard:

- Service Category: it specifies the service class to which the connection belongs (real-time or non real-time), the type of contract (business or residential) and the connection typology (traffic or signaling).
- Connection Configuration/Symmetry: specifies whether the connection is ptp, ptm, uni or bi-directional.
- Maximum Cell Transfer Delay: specifies the maximum transfer delay tolerated (in frames).

- Cell Loss Ration: specifies the ratio between the cells received with errors and the total transmitted ones.
- Peak Data Rate: specifies the maximum required transmission bit rate expressed in terms of 16 kbit/s multiples.
- Utilization Factor: specifies the ratio between the average cell rate and the peak cell rate.
- Maximum Burst Size: specifies the maximum size of a traffic burst (in cells)
- Peak-to-peak Cell Delay Variation: it defines the maximum on-board queuing latency expressed in terms of couple of frames.

To support these parameters two options exist:

1. Introducing new SIP header parameters and using them in the SIP message header part. This solution requires the extension of standard SIP. The parameter format for the new header parameters has to be defined and parameter names and values have to be introduced.
2. Using the message body as container for satellite specific data. When applying this solution two further options are possible: the format of the satellite specific parameters can be DSP like textual format or satellite specific own defined parameters. Both options allow a compact format, which would be smaller than the usage of additional SIP header parameters. This solution does not require the extension of standard SIP and the SIP textual part of the message would not be increased.

The second variants using the Satellite-SIP message body as container with satellite specific own defined parameters is efficient, because no extension of the SIP header fields is necessary and also no extension of SDP parameters is needed. The Satellite-SIP message body would contain these defined QoS parameters, which can be coded textual numeric value or also binary encoding, independent from the SIP and DSP message format. The description, in which messages the Traffic and QoS Parameters will be transport and how they will be exchanged, is already described in previously subsection 3.3.2.2.

3.5. Optimization through Message Size Reduction

The Session Initiation Protocol (SIP), along with many other IP protocols, is an UDP-based, textual protocol engineered for environments where bandwidth is not a serious limitation. As result, the SIP message size has not been optimized through the SIP standard. Typical SIP messages are from a few hundred bytes (500 bytes for a SIP INVITE message) to as much as 2000 bytes. In contrast, in a satellite environment the signaling traffic is a bandwidth overhead, which has to be minimized. Consequently for the Satellite-SIP implementation the message size reduction is a major concern. This can be done in two different ways: either message compression or using a binary encoding for Satellite-SIP messages.

3.5.1. Message Compression

Message compression is out of scope of the SIP standard. There are some Internet Drafts, listed in subsection 2.6.2.2, which introduces compression methods for SIP messages. There are basically two points to be considered for compression of SIP messages:

1. Which compression algorithms and method should be used, which tool should it be used for the compression.
2. How to negotiate the compression parameters / how to negotiation of the compression algorithm used between the sending and the receiving side.

3.5.2. Binary Encoding

The second possible solution is the binary encoding of the SIP messages. Due to the fact that in MM satellite environment most of the message parameters are numerical elements, the binary representation would be most suitable instead of the textual representation. The SIP Standard uses textual information elements (e.g. email addresses to identify the invited user). In opposite to, the Satellite SIP uses numerical values for numbering plans and identifies, therefore the numeric values would be most sufficient.

3.5.3. Conclusion

The message size reduction must be a part of the Satellite SIP implementation. There are two possibilities on which layer the message size reduction should be implemented: first the Satellite SIP layer and the second variant is the layer below Satellite SIP, the Adaptation layer, as described in subsection 3.2.2.2. The second variant is the famous one, because in this solution the message size reduction procedure is independent of the Satellite SIP protocol implementation.

The selection of the strategy – compression / compression algorithm / binary encoding – is done most efficiently for a given satellite system. Once the messages and message contents are specified in detail, the performance (in terms of size reduction and computing effort) of a message reduction strategy can be measured.

3.6. Support of Satellite Specific Services

This subsection describes the satellite specific services, which have to be supported by the Satellite-SIP protocol version and protocol implementation. At the same time this subsection analyses the required adaptation of the SIP Standard to support satellite specific services. In general is to considered that the most of the point-to-point services are adopted with only a few extensions from the SIP standard, during the point-to-multipoint serves represent a pure extension to the standard.

3.6.1. Registration/Deregistration Services

Each satellite terminal is responsible for its own registration and deregistration by the NCC (Network Control Center). The registration procedure must be performed, before a terminal can become active in the satellite network. If the registration has been completed successfully by the NCC, the terminal is marked as active.

The registration and deregistration procedure in Satellite-SIP is preformed through a REGISTER request as described in the SIP Standard. For both services authentication is required, which is perform by MD5 algorithms, as well applied in the SIP standard. The following useful deviations from the SIP Standard are required for satellite environment:

- When a terminal sends a REGISTER request, it will not suggest an expiration interval (as in the SIP standard) indicating how long the terminal would like the registration to be valid. The registration will be valid until the terminal initializes the deregistration procedure. Therefore the Expire header field will not be used to indicate the lifetime of the registration, but in the deregistration procedure the terminal sends a REGISTER request with Expire=0 to indicate the deregistration procedure.
- An update of the registration is not foreseen.
- A registration on behalf of a particular address (TerminalIDs) cannot be performed by an authorized third party. The satellite terminal is responsible for its own registration and deregistration

3.6.2. Point to Point Services

3.6.2.1. Point to Point Connection Initiation

A ptp session is defined as a session between two terminals. The ptp connection initiation allows two terminals to initiate/establish the session and to begin the communication. For the point-to-point call/session establishment the calling terminal uses an INVITE request. The request line contains the address of the NCC, where the request is sent. The request header includes the required header fields, as well used in the SIP Standard. Deviation and extension of the standard consists of using the SIP message body for transport of the QoS parameters. Further the message body gives information about the connection symmetry. In this case, the symmetry is ptp and either uni- or bi-directional. Once a ptp session is initiated, adding a new terminal is always possible. If one or more new parties are added to a ptp session, the session type switches to ptm.

3.6.2.1.1. The different kind of the INVITE request

In Satellite-SIP different kind of naming convention and usage of the INVITE request exists. In this case, establishing a connection, the INVITE request is called init-INVITE. Other possibilities are add-party-INVITE and re-INVITE. Due to the reason that the INVITE request is used for different kinds of services; therefore a different name of the request is used for each service in this document, as described in the next Table 36.

Table 36: Naming Convention and Usage of the INVITE request

Term	Usage
INVITE	The term "INVITE" is a common usage of the INVITE request; no specific usage or service is being addressed. According to the previous statement the term INVITE refers to all kind of INVITE messages. (For example: "The INVITE request does not carry authentication credentials.", means that this statement is valid for all services - init-INVITE, re-INVITE and add-party-INVITE)
init-INVITE	The term "init-INVITE" refers to the INVITE message sent for session initiation.
add-party-INVITE	The term "add-party-INVITE" refers the INVITE message used for adding one or more terminals to an already existing session, with or without changing the QoS parameters of the session.
re-INVITE	The term "re-INVITE" refers to the INVITE message used for changing the QoS parameters.

In the Satellite-SIP the INVITE message header contains the signaling data whereas the message body contains the QoS parameters for the payload connection. The message body provides also information about the connection symmetry, which can be point-to-point or point-to-multipoint, uni- or bi-directional. The request body of an INVITE request is not allowed to be empty, with one exception when adding one or more party/parties to an existing connection whereby the QoS parameters are not going to change. The INVITE request does not contain authentication credentials.

The INVITE request is sent from the calling satellite terminal to the NCC and from NCC to the called satellite terminal. If the NCC receives an INVITE request, it first examines - using the CAC (Call Admission Control) - whether the session with the requested QoS can be admitted. If this is the case, the RM (Resource Management) is configured and the NCC sends the INVITE to the called terminal. If the session cannot be accepted with the requested QoS, the NCC returns a final response to the calling terminal, without notification to the called terminal. In the acceptance case the called terminal is invited to session by an INVITE request and, at the same time, informed about the QoS characteristics of the session (as requested by the calling terminal). This procedure is only valid for the session initiation and during the session for renegotiation of the QoS parameters, consequently for the init-INVITE and re-INVITE requests. The add-party-procedure using the add-party-INVITE is identical but without changing the QoS parameters.

3.6.2.1.2. The init-INVITE

The init-INVITE request is used for a ptp or ptm session initiation. If no session is available, sending an init-INVITE request initiates a new one (either ptp or ptm). The init-INVITE always requests two services: first the assignment of QoS parameters and second the invitation of a terminals. Only after successful assignment of the QoS parameters, the session initiation can proceed.

3.6.2.2. *Point to Point Connection Parameter Re-Negotiation*

In a ptp connection the calling terminal can change the existing QoS parameters using the re-INVITE, which is described in the next subsections 3.6.2.2.1. The request body contains the new desired QoS parameters. This re-INVITE request allows changing all QoS parameters. This includes the bandwidth parameters and also the connection symmetry, i.e. a switch from a uni- to bi-directional session or reverse. The re-INVITE procedure is also known by the SIP standard, however it is used for different purposes. In the Satellite-SIP the re-INVITE procedure is adapted for satellite specific requirements.

3.6.2.2.1. *The re-INVITE*

The initiator of a session may use a re-INVITE request during the session to change the QoS parameters. For this purpose, the calling terminal sends a re-INVITE containing the new QoS parameters (in the message body) to the NCC. This re-INVITE references the existing session through the equal To, From and Call-ID header fields, so that the NCC knows, that it is a modification of the QoS parameters of the existing session (instead of a request to establish a new session). The CSeq number is incremented by one, which indicates a new request within a session and not a retransmission of the initial request to establish the session.

The NCC verifies this requested QoS parameters and – if the NCC accepts the changes – informs all called terminal(s) through the same re-INVITE request about the changes. Each called terminal sends a 200 (OK) to indicate the reception of the request, however the called terminal(s) has/have no possibility to refuse the changes. If the terminal does not accept the changes, it can drop itself from the connection with a BYE request afterwards. After timeout or reception of all notifications from the called terminals the NCC responds to the calling terminal with a single 200 (OK) containing all identifications of terminals, for which a 200 (OK) has been received. The calling terminal responds with an ACK to the 200 (OK). If the NCC does not accept the change or cannot assign the requested QoS parameters, the NCC sends an error response 488 (Not Acceptable Here), which is acknowledged with an ACK from the calling terminal. If the re-INVITE fails, the existing session continues with the previously negotiated QoS parameters.

3.6.2.3. *Point to Point Connection Add Party*

The add-party-procedure can only be performed by the calling terminal. Consequently the ptp session will automatically changed from ptp to a ptm session. The add-party-INVITE with an empty message body is used for this purpose during the session. The changes of the QoS parameters is not foreseen during this procedure. The add-party-INVITE request header contains the new terminal or terminals through the To header field. The description of the add-party-INVITE request is provided in the next subsection 3.6.2.3.1. The add-party-functionality is a pure satellite specific service, therefore an extension of the SIP standard.

3.6.2.3.1. The add-party-INVITE

The add-party-INVITE request is sent before a session is terminated and allows adding one or more satellite terminals to the existing session. In the add party procedure, the Call-ID and the From header fields of the add-party-INVITE are equal to those of the init-INVITE request. The CSeq contains an incremented sequence number (with respect to the previous transaction for that session) and the To header field contains the address (TerminalID(s)) of the new terminal(s).

The request body in the add-party-INVITE is always empty, because the add-party-INVITE requests only one service, to add new terminals to the session. The second service to change the QoS

parameters is independent of add-party-procedure and is done through a re-INVITE.

The unsuccessful add-party-INVTE can have two reasons: the NCC does not accept the request or the called terminal refuses the invitation.

3.6.2.4. Point to Point Connection Termination

The ptp connection termination is used in the same way as described in the SIP standard. The BYE request terminates a ptp session. It may also be used to terminate an attempted session. The BYE request can be sent from the called or from the calling terminal. The request line of the BYE request contains the NCC or the terminal, for which the termination is meant and where the request is sent. The request header includes the required header fields and the message body is empty. The answer is a 200 (OK) and therefore the session is terminated. After the reception of the final response to the BYE request, the CAC is informed about the no longer needed resources and the RM is configured accordingly.

The ptp connection termination can also be unsuccessful, the following cases has to be considered:

The calling or the called terminal of a ptp session sends a BYE request to terminate a session or an attempted session. If the final response is 4xx or 5xx, the session was not successfully terminated. The reason depends on the response code. In an unsuccessful case the unsuccessful final response can be sent from the other terminal or from the NCC.

The satellite terminal which wants to terminate the session must try again by sending the BYE request, if the final response to the first BYE request is 400, 408, 416, 500 or 505. It must consider the session is terminated if the final response is 481, 487 or no response is received.

If no final response for the BYE request is sent or if the final response is 481 or 487, the NCC informs the CAC about not needed resources and configures the RM accordingly.

3.6.3. Point to Multipoint Services

3.6.3.1. Point to Multipoint Connection Initiation

A ptm session is defined as a session between more than two terminals with one calling and several called terminals. The ptm session initiation is very similar to a ptp case however an extension to the SIP Standard, where only the ptp case is discussed. The main difference is, that not only one terminal is invited (as in ptp), but several terminals are invited. Therefore the INVITE request contains one To header field with several invited terminals listed in this field. This means that also in the ptm case only one INVITE message is sent to the NCC. The NCC receives this ptm INVITE request and is responsible for invitation of the called terminals. The NCC operates as a "distributor": It creates an INVITE message for each invited terminal where only this terminal is included in the To header; then the NCC distributes these INIVTE requests to the called terminals. Each called terminal can optionally answer with a provisional response and must answer with the final response. The NCC waits until all terminals have answered with a final response or a timeout occurs. After the NCC has collected all final responses, it composes a final response, which it sends to the calling terminal. This final response can either be:

- a 200 OK with all called terminals in the To header (in the case that all called terminals answered with a 200 OK),
- a 200 OK with some of the called terminals in the To header and the others in the message body of the 200 OK message (in the case that some called terminals answered with a 200 OK and other called terminals sent 4xx final responses),
- a 499 response message (new message, not exists in the SIP standard) listing all called terminals with their response type (if 4xx final response messages have been created for/by all called terminals).

The NCC sends this final response to the calling terminal. As specified in the SIP Invite Client Server Transaction, the NCC sends an ACK message to each terminal which has sent/created a negative final response. After receiving the final response, the calling terminal sends an ACK to the NCC. Then the NCC distributes the ACK request to the successful called terminals (the others have already received an ACK from the corresponding Client Server Transaction).

An other extension to the SIP standard is to address a satellite spot beam in the To header field instead of a specific terminal. This means that all terminals in this spot beam shall be invited to the session.

3.6.3.2. Point to Point Multipoint Parameter Re-Negotiation

The re-INVITE request is used for changing the QoS parameters of an active ptm session as well used in the ptp case. The re-INVITE for a ptm session contains all called terminals of the session in the To header field. The usage is the same as in the ptp case, except for the QoS parameter "connection symmetry" which stays "ptm" in any case. (The connection symmetry may change from uni- to bi-directional or reverse.). See description for re-INVITE in subsection 3.6.2.2.1.

3.6.3.3. Point to Multipoint Connection Add Party

The add-party procedure in a ptm session is handled by the add-party-INVITE request, which is sent by the calling terminal, as well used and described in the ptp case, see subsection 3.6.2.3.1 for more details. The add-party-INVITE contains the desired terminal or terminals in the To header field. The request body is empty; therefore the change of QoS parameters is not foreseen. However,

if through the add-party-procedure more bandwidth or other traffic parameters are required, the re-INVITE procedure can be used.
The add-party-INVITE is not successful, if either the called terminal(s) do not accept the invitation or the NCC cannot fulfill the request.

3.6.3.4. Point to Multipoint Connection Drop Party

Only the calling terminal may initiate the drop party procedure in a point-to-multipoint connection. The calling terminal sends a BYE request to the NCC. The To header field contains a list of terminals or only one terminal which informs the NCC which terminal(s) should be dropped from the session. If the To header field contains all called terminals, consequently all terminals will be dropped, which means that the session is terminated by the calling terminal, see subsection 3.6.3.6.
If the NCC receives a BYE request and the To header contains some of the called terminals, the NCC knows that the given terminal or the listed terminal(s) have to be dropped from the session. The NCC sends the BYE request to the called terminal(s) (containing in the To header field only the corresponding terminal) to release these terminals from the session. The called terminal(s) have to answer with 200 (OK). Afterwards the NCC informs the calling terminal with 200 (OK) where the To header field contains again all the dropped terminals.
In a drop party procedure the message body of the BYE request is empty and therefore no changes of the QoS parameters are foreseen. After the drop party procedure the calling terminal can change the QoS parameters through a re-INVITE request, if it wishes to do so.

If the drop-party-procedure is not successful, it can have several causes. The NCC can respond to the BYE request with 4xx or 5xx. The called terminal can respond with 4xx. The behavior of the calling terminal or of the NCC with respect to drop the called party depends on the reason code.
If the NCC gets a 4xx from the called terminal, the NCC has to try again if the response code is 400, 408 or 416. If the response for the BYE is a 481 (Call/Transaction Does Not Exist) or a 408 (Request Timeout) or no response at all is received for the BYE (that is, a timeout is returned by the client transaction), the NCC must consider the session as terminated and must answer with 200 (OK) to the calling terminal.
If the calling terminal gets a 400, 408, 416, 500 or 505 response from NCC, the calling terminal has to try again. If the response code is 481, 487 or no response (timeout occurs), the calling terminal should not try again, but consider the session is terminated.
A 499 response is sent back from the NCC if the BYE transaction fails for *all* called terminals. The response type of each terminal is contained in the body of the 499 message so that the calling terminal knows to which terminals it has to send the BYE request again.

3.6.3.5. Point to Multipoint Connection Leave Party

In the point-to-multipoint case a BYE request sent by a called terminal does not terminate the session, but the called terminal leaves the session. In this case the called terminal sends a BYE request to the NCC. The NCC is responsible for the following two tasks: first to inform the calling terminal with a BYE request, that one of the called terminals has left the session, and second to release the connection to the called terminal through a 200 OK response.

If the calling terminal gets a BYE request from one of the called terminals, the calling terminal has to answer with 200 (OK), because the BYE request is only a notification for the calling terminal, that the specific called terminal has left the session.

If a called terminal leaves the session, the QoS parameters are not changed. The calling terminal can change the session parameters through a re-INVITE, if it is necessary.

If the called terminal that leaves the session is the last one, then the CAC has also to be informed and the RM has also to be configured accordingly, then in this case the session is terminated.

If the leave-party-functionality is not successful then following behaviors has to be considered. The response to the BYE request is unsuccessful if the called terminal gets a 4xx or 5xx from the NCC. The called terminal should try to leave the session again if the response code is: 400, 408, 416, 500 or 505. Otherwise, if the response is 481, 487 or no response (timeout occurred), the called terminal must consider the session as terminated.

If the NCC receives a non-200 response from the calling terminal during the notification that a called terminal left, the NCC should only try again, if the response code is 400, 408 or 416.

3.6.3.6. Point to Multipoint Connection Termination

In the point-to-multipoint case a BYE request sent by the calling terminal terminates the session, only if all invited terminals are listed in the To header field of the request. If not all called terminals are present in the To header field, the NCC supposes that the listed terminals have to be dropped from the session (and consequently one or more called terminals stay in the session). Therefore the calling terminal has to put all called terminals to the To header field in order to terminate the session. The session termination means that:

- all called terminals are dropped from the session,
- the QoS parameters are released and available again for future sessions, and
- the connection to the calling terminal is released.

The response to the BYE request is unsuccessful, if the called terminal receives

- a 5xx response from the NCC (NCC error) or
- a 499 response from the NCC which contains the reponses of all called terminals (4xx).

The request is partly successful, if the called terminal receives

- a 200 (OK) with a non-empty body; the body contains the responses (4xx) of the terminals for which the BYE has failed.

The calling terminal should try again to terminate the session for all terminals where the response code is 400, 408, 416 or if the response is 500 or 505. Otherwise, i.e. if the response is 481, 487 or no response (timeout occurred), the calling terminal must consider the corresponding terminal to be dropped. The session can be considered as terminated if all called terminals can be considered as dropped.

Once the NCC gets a BYE request from the calling terminal with all called terminals in the To header field, the NCC has to drop each called terminal through one BYE request as described in the drop party procedure. If the NCC gets a non-200 (OK) from any of the called terminal(s) during the drop party procedure, the NCC should only try again, if the response code is 400, 408 or 416.

If no final response for the BYE request is sent or if the final response is 481 or 487, the NCC informs the CAC about not needed resources and configures the RM accordingly.

4. Summary and Conclusion

The purpose of this thesis was to investigate the applicability of VoIP protocols for connection control in a satellite environment and to find to most applicable protocol for multimedia satellite purposes. The reason why VoIP protocols were chosen is seen in the functionality of these protocols to support connection control in a telephony environment. The second, much more important reason was to find a simpler protocol than the already used powerful B-ISDN, which has a high complexity and functionality, needed in the terrestrial network, but not really necessary in a satellite environment. Thus, the investigation and analysis of the existing VoIP protocols, H.323, SIP and Megaco began. The goal was to find the most applicable protocol for satellite purposes. It was realized very early that none of these protocols were created to operate in a satellite environment and therefore some modification, adaptation and extension of the chosen protocol will be necessary. For this reason a protocol was required which is mostly compliant without many changes to fulfill the requirement of the satellite system. First a set of criteria' was identified in order to perform a comparative analysis between the candidate protocols. The criteria outlined the requirement of a satellite network. Each criteria was subdivided in sub-criteria which allows an examination and comparison of the protocols.

The criteria applicability outlined that SIP and H.323 are both suitable in the same manner. Both support the basic functionality, such as establishing and releasing the connection and also the services such as registration and authentication, and the supplementary services. Megaco does not support these services in a suitable way. The same way was valid for the criteria's network element and communication model: H.323 and SIP have the same structure of the network topology as needed in a satellite system, while Megaco does not. In total SIP fulfills only two criteria better than H.323: QoS Parameter transport and adaptability, because SIP has a structure that can be simply modified.

The second criteria complexity was a quantitative analysis because complexity can be high or low and the best way to measure it was a comparison in a quantitative manner. The result was as expected. H.323 is the oldest and also the most developed protocol with the highest functionality, this results in the highest complexity between the candidate protocols. The H.323 umbrella recommendation includes much more functionality than is needed for connection-related signaling. SIP has a considerably lower complexity. Merano has the lowest complexity. This is due to the simplicity of the protocol, which does not fulfill the requirement of call control for a satellite environment.

The criteria performance was very important due to the fact that call control signaling in a satellite environment occurs a signaling overhead on the broadband channel. Therefore criteria such as the required signaling channels and the uncompressed size of the protocol messages were analyzed to calculate the signaling load at the RASC (Random Access Channel) and DSC (Dedicated Signaling Channel) channels that are used by satellite terminals to signal call establishment and release procedures. The result of this calculation outlined that H.323 and SIP has a similar signaling characteristics and load. For Megaco it was not useful, because the Megaco architecture cannot be mapped to the satellite network architecture.

The criteria extensibility examined the compatibility between protocol versions and easiness of integrating advanced services. Due to the reason that Megaco hardly supports the required basic functionality, the extension would require a high effort, therefore is not recommended. The comparison of SIP and H.323 outlined that SIP protocol structure allows additionally message header and parameters without changing the existing functionality, during H.323 hardly possible to extend, because the extension of H.323 messages is only possible, where placeholders were defined.

The criteria resource and bandwidth efficiency analyzed the possibility of bandwidth management before und during the call and the possibility and kind of resource reservation/allocation mechanism used by the candidate protocols. The result outlined, that H.323 and SIP support the required criteria in similar way, although both fall back upon other protocols, such as RSVP and by itself they do have suitable mechanisms, as well, too Megaco does not offer any mechanism for bandwidth management.

The final result of the analysis shows the advantages and the disadvantages of each protocol. While Megaco has been ruled out very easy due to the fact, that is does not fulfill the basic requirement, SIP and H.323 stayed in competition. H.323 pointed positive on the similarity to the already used B-ISDN for satellite system. Further advantage is the binary representation of the messages and its involved small message size for performance purposes. The disadvantages of H.323 are the high complexity and the difficult protocol extensibility, protocol reduction to fulfill the satellite requirement. SIP allows to extend and to reduce easily and has a low complexity. The disadvantage of SIP is the textual encoding of the messages and consequently the large message size.
One important requirement of satellite network is the bandwidth management, this means the possibility to allocate a bandwidth for the call. This requirement is not fulfilled by any protocol, consequently the chosen protocol has to be extensible to supports the resource reservation mechanism. SIP and H.323 both support the basic functionality for call control signaling, but only SIP can be extended for resource management. SIP wins this race. SIP was considered the most attractive candidate protocol for call control in a MM satellite network. The additional conclusion was that certain adaptations and optimizations was however necessary.

The general intension was to adhere as far as possible to the SIP standard, any deviation was motivated and justified. Two aspects for deviation was quite simple to justify: Extension and

Simplification. Extension was quite necessary to support satellite specific services, such as transport of QoS parameters for resource reservation and bandwidth management. Simplification was reasonable due to the fact, that only a subset of the SIP services, outlined from the SIP standard are needed and will be used in satellite network. One example is the routing functionality, which is not needed in a one-hop satellite system. The third section of this thesis concentrates how to perform this two task: extension and simplification.

The first obstacle was the fact, that SIP was originally designed for IP network and not for a satellite network, therefore the first task was to found how to integrate SIP in a satellite environment. The result of this investigations outlined more then one possibilities, during the preferred solution was to include an adaptation layer between SIP and Satellite-Layer-2. This solution complies the requirements of the Satellite Layer-2 and also of SIP, because it offers possibilities to consider the requirement of both in one layer between of them.

The second task was to find the way how to solved the problem of the big message size of the SIP protocol, which was given through the textual encoding. Two solutions have been introduced: either message compression or using a binary encoding. The selection of the strategy can be done most efficiently for a given satellite system and is part of the implementation. One the Satellite-SIP protocol standard is defined and specified in detail the message size reduction strategy can be measured.

The third important task of the adaptation of the SIP protocol standard was to optimize the SIP message and the SIP parameter set. This optimization was divided in two parts: first the removal of messages and parameters from the SIP standard, which are not needed in satellite system. This results a simplification for the satellite-specific-SIP protocol through minus 32 response messages (32 of 50) and minus 32 parameters (32 of 44). The second part of the optimization discusses the required extension of the messages and parameters to support satellite specific services. The following two services are not part of the SIP standard and had to be added: support of point-to-multipoint services and support of transport of QoS parameters to control resource and bandwidth management.

The last task was to analyze the satellite specific services and to describe how will the SIP protocol standard including the satellite specific extensions fulfill the each service. First registration and deregistration including authentication was described. This was followed by the point-to-point services and later by the point-to-multipoint services. The point-to-point services outlined following tasks: ptp connection initiation, ptp connection parameters re-negotiation for QoS purposes, ptp add party to the connection and at last ptp connection termination. The ptm services result the same tasks as the ptp, and additionally following two tasks: ptm drop party from the connection and ptm leave party from the connection.

At the end it has to be considered, that this thesis select the most suitable protocol from the three given VoIP protocols, but the most important realization was the point that it does not exist a protocol, which is create for satellite purposes and therefore the successfully usage of a VoIP protocol can only be performed though modifications and adaptation for satellite purposes. One goal was to adhere as far as possible to the SIP standard. At the end the next Table 37 gives an overview of changes and deviations from the SIP standard, which was necessary to fulfill the satellite specific requirements.

Table 37: Deviations from the SIP standard

	SIP Standard	**Satellite-SIP**
Usage of Protocol	SIP is used to create session consciousness in the end user devices.	Modification: Satellite-SIP is used as signaling protocol for a payload connection with guaranteed QoS. This means that Satellite-SIP
Users	SIP establishes sessions between end-user devices such as e.g. SIP phones.	Modification: Satellite-SIP is used for the signaling of payload connections over the satellite between satellite terminals. End-user devices such as e.g. PCs are connected to these satellite terminals. The end points of a Satellite-SIP session are always satellite terminals.
	SIP foresees human/user interaction for the called side.	Reduction: Satellite-SIP does not foresee the possibility of human/user interaction in the signaling process. The called terminal is supposed to respond automatically to an incoming request. If there are not any technical problems, the called terminal must accept an INVITE request.
Entities / Architecture	SIP supports entities such as end-user devices (e.g. SIP phones), registrar servers, proxy servers, etc.	Modification: Satellite-SIP only supports satellite terminals (correspond to end-user devices in SIP) and the NCC (which combines some proxy and registrar functionality and extends it).
Functionalities		Reduction: Ringing: Satellite-SIP does not support the ringing functionality. Ringing is not necessary as no human/user interaction is possible during the session signaling.
		Extension: Satellite-SIP explicitly foresees the integration of Connection Admission Control and Resource Management subsystems during session establishment, release or QoS parameter changing during a session.
		Extension: Satellite-SIP supports ptm sessions.
Messages	SIP supports the complete set of SIP-messages	Reduction: Satellite-SIP supports requests and responses as described in the SIP standard with the restriction that not all messages are included in Satellite-SIP protocol, due to the reason, that not the whole functionality from the SIP standard is adapted.
Message Parameters	SIP supports the complete set of SIP message header parameters	Reduction: Satellite-SIP supports a reduced set of message header parameters
		Extension: Satellite-SIP supports an extended set of message body parameters, which are inserted to carry QoS parameters for Connection Admission Control and Resource Management.

Annex A Multimedia Satellite Network

A.1 The Euroskyway Program

The EuroSkyWay program is a next generation satellite system, which provides broadband services in Ka-band with full digital payload, guaranteed service quality and supports two-way communications for business and private market for the European and Mediterranean area. Further a point to point, point to multipoint connectivity, bi-directional bandwidth on demand, dynamic bandwidth resource allocation, secure communications, satellite dishes about 80 cm, simple and mobile terminals.

The main services of the Euroskyway program are following:

1. Internet Services with guaranteed download rate 128 kbps (residential) or 640 kbps (business). The Internet services support a Quality of Services with guaranteed bandwidth and bi-directional connectivity and allow differentiated services for classes of users. The bi-directional transfer allows the up and the download over the same connection. All internet services, which are also supported through terrestrial network, are provided. Some of them are: direct connection to Websites/WebTV/Web Radio, upload/download from web serves, Videoconferencing, direct email delivery. The differences to the terrestrial internet are the QoS (Quality of Services) service support and the guaranteed bandwidth.
2. Interactive TV and Streaming Video in Ku-band or Ka-band. The new generations of TV services are interactive, digital and operate on Ka-band and Ku-band. The user terminals are able to receive TV program on demand, additionally to the traditional and the pay-TV.
3. Virtual Private Network for LAN-to-LAN meshed connectivity on demand with 2 mbps. The EurpSkyWay program allows the creation of Virtual Private Networks directly through individual companies, organizations etc. (for example a connection between two branch offices).
4. Voice Services based on Voice over IP standard with 6 kbps bandwidth per connection. Voices over IP (VoIP) services are an alternative to PSTN voice services and an efficient solution of voice and data integration in private and public networks. The following scenarios can be considered: Voice between PC and PC, Voice between PC and Phone, and at last Voice between Phone and Phone.

A.2 Technical Introduction

This subsection introduces a satellite system and also a GEO satellite from the technical point of view. A Multimedia Satellite Network, such as the EusoSkyWay Programm, uses a GEO satellite.

A.2.1 The Satellite Functionality

Each satellite system consists of some network components, which are required for operation and to fulfill the main functionality, these are follow:

- Communication capability with earth: communication antenna, radio receiver and transmitter, which enable the communication between the satellite and the ground station (also called Network Operation Center). The ground station uplinks messages to the satellite, while the satellite downlink messages back to the earth.
- Power source: most of the satellites use battery power combined with solar recharging. Other satellites have fuel cells that convert the chemical energy to electrical energy. Only a small number of satellites (the GEOs) use the altitude, gravity and rotation to control and stabilize the satellite position in the orbit.
- Body with control system: the body of the satellite, also called the bus of the satellite, holds all of the scientific equipment and necessary components of the satellite.

These different components accomplish the tasks assigned to the satellite. These include optical systems, sensors in a range of wavelengths. The satellite receives data from one satellite dish and transmits this data in bit sequence to the ground station. A satellite works similarly to a mirror, because it reflects the data from one earth station to the other. In a similar way also works environmental satellites, which transfer or reflect an image. The data allows the calculation of the positioning information (in case of GPS) or the imaging information (in case of environmental satellites). The design of the complexity of the satellite and the ground station decides how the complexity is divided. For examples, when the satellite handles complex calculations, the ground station can be relatively simple. In the case of a "dump pipe", where a satellite has only a reflect functionality, the ground station has to be more complex.

In case of a Multimedia Satellite Network, such as the EuroSkyWay, the complexity is distributed through the complexity of the ground station and the complexity of the satellite. The satellite contains an on-board modulation/demodulation and on-board processing capability. The distinct advantage of an on-board processing architecture is that it can provide global access to the network and guaranteed QoS via on-demand bandwidth assignment mechanisms. The ground segment consists of a control center, which provides the means and resources to manage and control the satellite connection. Furthermore, it provides an interface to the users to allow optimum utilization of the system resources according the users requirements.

A.2.2 The Satellite Frequency

The range of frequency in which a satellite communicates is called a band. There are three commonly used bands in telecommunication: C-band, Ku-band and Ka-band. The next Table 38 shows the different frequencies of these three bands.

Table 38: Range of Frequency – commonly used bands

	Earth to Space	Space to Earth
C-band	5.850 – 6.425 GHz	3.6 – 4.2 GHz
Ku-band	12.75 - 13.25,13.75 - 14.8 GHz	10.7 - 12.75, 17.3 - 17.7 GHz
Ka-band	27.5 - 30.0 GHz	17.7 - 21.2 GHz

The C-band is able to reach several continents, but the satellite dishes have to be three meters in diameter. The Ku-band reaches only one continent and has significant smaller dishes than C-band; television stations mostly use it. The Ka-band provides sufficient bandwidth to support multimedia application and is therefore mostly used by multimedia companies. The wide range of the Ka-band allows the transmission of data in multiple frequencies simultaneously and also allows two way broadcast services, bidirectional communication. The Ka-band allows the transmission of 1.5 Megabytes per second, which is 150 times faster than data transmission over phone lines. The GEO satellite of the EuroSkyWay program operates in Ka-band. The next Figure 10 shows all satellite frequencies which exist.

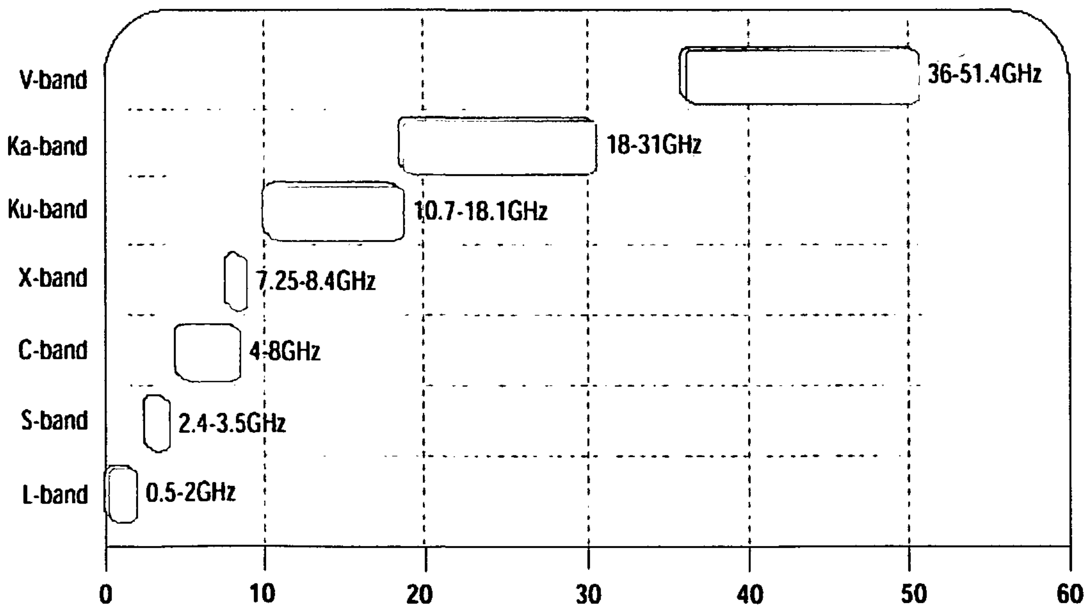


Figure 10 The Satellite Frequencies

A.2.3 MEO, GEO, LEO

A.2.3.1 LEO – Low Earth Orbit

LEOs – Low Orbit Satellite – are in orbit on an altitude of 700 to 1,400 km. From this altitude a lower coverage is possible. For global coverage 40 satellites are needed. The satellites have no fixed position but they rotate around the earth and the rotation speed allows getting around of the earth in only 1.5 hours (90 minutes), this follows that LEOs reaches a speed of 27,359 kilometers per hour (17,000 miles per hour). Because LEOs are so close to the earth, they must travel very fast so the gravity will not pull them back to the atmosphere. The weight of the satellite is less than 1 ton. Through the high rotation speed the lifetime of the satellite is restricted to only 3 to 7 years. The reason for this low lifetime is the high energy consumption through the high rotation. One rocket can transport 3 to 4 satellites from the earth to the orbit. The disadvantage of this transport is the cost for the second transport if the first rocket has a false start and it is destroyed. The antenna for receiving is small and it can also be used in mobile phones. Through the low altitude the signal in one way, from earth to the satellite and reverse, needs only 0.05 second, which is a significant advantage of LEOs.

A.2.3.2 MEO – Medium Earth Orbit

A MEO - Medium Earth Orbit - satellite is in the earth orbit on an altitude of 10,000 to 15,000 km. The number of required MEOs for global coverage is 10 to 15 satellites. One MEO has two rotations in a day around the planet, therefore the rotation time is 12 hours. The lifetime of a MEO is 10 to 15 years. A signal, from the earth to the satellite and from the satellite to the earth, takes 0.10 second.

A.2.3.3 GEO – Geostationary Earth Orbit

GEO broadband satellite system consist the high power and the high potential for multimedia application, such as conventional broadcasts, interactive TV (game shows, gambling, talk shows, home shopping), interactive services (banking, shopping, e-commerce, education, information services, database access) and internet access. Therefore a GEO satellite is used in a Multimedia Satellite Network, as well in EuroSkyWay program.

From the technical point of view GEO satellites are in the earth orbit at on the altitude of 36,000 km. In this altitude the gravity is neutralized through the rotation and in this way the satellite stays always on the same position directly above the equator. The rotation speed of the satellite is the same as the rotation speed of the earth 30 km in an hour, which allows exactly one rotation in 24 hours. The satellite weighs between 7 and 8 tone and one rocket transports only one satellite from the earth to the orbit. For global coverage 3 to 4 satellites are needed, for Europe one satellite is sufficient. GEOs have a lifetime of 15 years. The antenna for receiving is a satellite dish with a diameter of 80 cm and therefore it can only be used as stationary receiver attached on a building. Through the high altitude GEOs have the highest delay, which is a significant disadvantage. The signal from earth to satellite upstairs and from the satellite to the earth downstairs takes 0.25 second.

A.2.3.4 Comparison between LEO, MEO and GEO

The next Figure 11 shows a comparison of the three satellite systems and also the differences in altitude, in coverage and in satellite dimension.

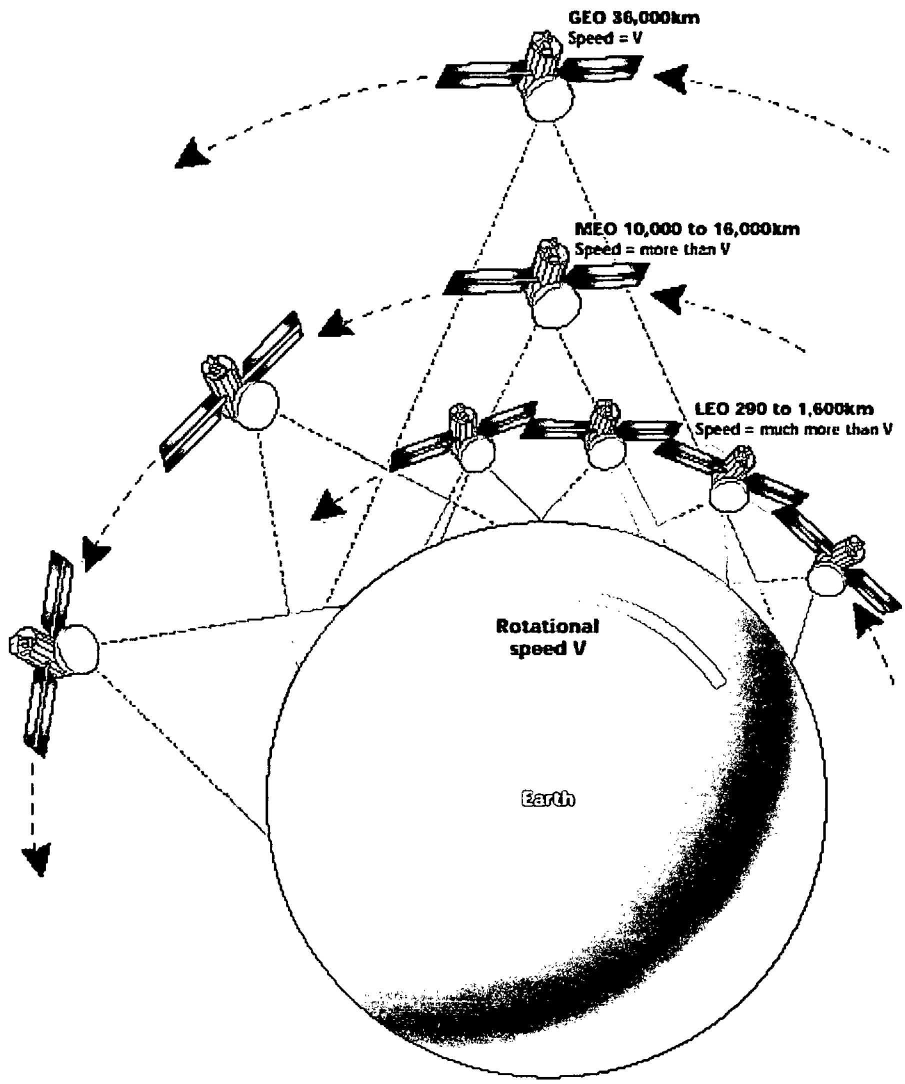


Figure 11: Differences in altitude, coverage and dimension

A.3 Architecture of the Multimedia Satellite Network

A.3.1 Network Elements

The network architecture of a Multimedia Satellite Network, such as the EuroSkyWay program contains the following three components also shown in the Figure 12 below:

1. Space Segment – a GEO satellite
2. Control Segment – the control station which consist the Satellite Operation Centre, the Network Operation Center and the Service Centre
3. Ground Segment – are the terminals and gateways (the actual user)

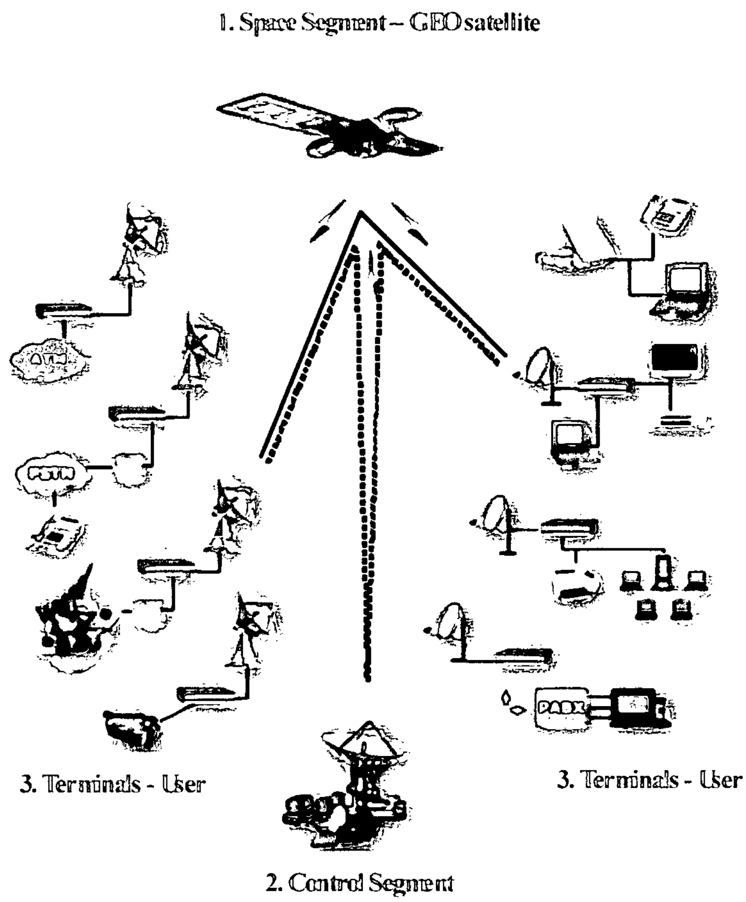


Figure 12: The Satellite Network Architecture

A.3.1.1 Space Segment

The Space Segment consists the satellite, which is a geostationary - GEO - satellite, in orbital position from 5° East to 39° East, using Ka band (20-30 GHz) frequency. The coverage of the satellite spot, shown in Figure 4, will include Europe, the Mediterranean Basin and surrounding seas and North Africa for a lifetime of 15 years, shown in the next Figure 13.

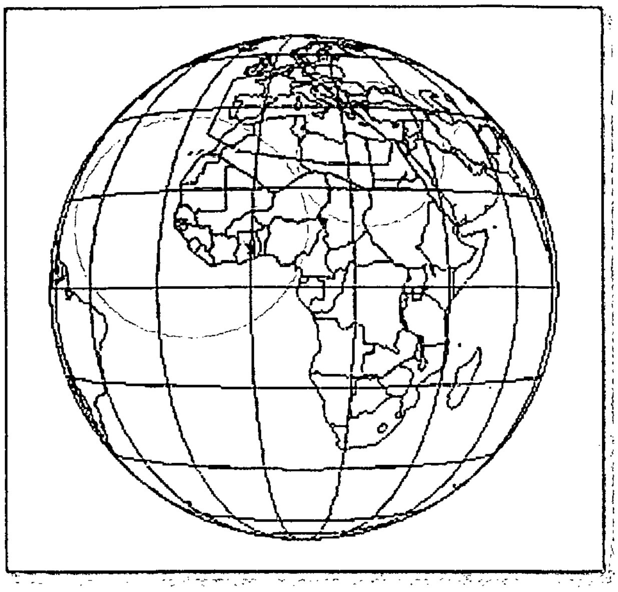


Figure 13: Satellite Coverage of EuroSkyWay

The payload traffic is composed of the following two components:

- a regenerative component with 30 channels, to support the governmental services and the commercial market with frequencies for uplink 27.5÷30.0 GHz and downlink 17.7÷20.2 GHz. The total on board processed traffic capacity is of about 1 Gbps (30 channels x 32.768 Mbps).
- a transparent component with 6 channels for special application with frequencies for uplink 30÷31 GHz and the downlink 20.2÷21.2 GHz. The total on board processed traffic capacity is of about 40 Mbps.

A.3.1.2 Control Segment

The ground control segment consists of Satellite Operation Center (SOC) and Network Operation Center (NOC). The SOC has the basic function of monitoring and controlling the whole satellite, including the payload and utilizing data. The NOC manages the connections and monitors and controls the network. It is composed of the following subsystem:

- NCC - Network Control Center - the NOC subsystem, which provide functionality related to the real-time communication session management, supporting the connectivity service.
- NMC - Network Management Center - provides the typical network management functions, required in order to manage the entire system, allowing correct and continuous system operability.
- CCC Customer Control Center - responsible for everything concerning the subscriber and billing management.
- DBCS Data Base Control System - in charge of managing the data in the Data Base present in the NOC.

The main internal component for connection, admission and bandwidth control is the NCC with the following architectural components:

CCM - Connection Control Manager

CAC - Connection Admission Control

LM - Localization Management

SCH - Synchronization Channel Handling

SSM - Secure Service Management

DBAM - Database Access Module

These components are connected through an interface called MH, message handler. The second interface is called SH, signaling handler, which allow the connection between the NCC and an external components.

A.3.1.3 Ground Segment

The ground segment contains the satellite terminals. There are three kinds of terminals: Gateways, Service Provider and User Terminal:

- The Gateways Terminal (GTW) are the interconnection points between the satellite and the terrestrial Telecommunications Network.
- The Service Provider Terminal (PrT) delivers their services directly to end-users via either satellite or landlines.
- The Satellite User Terminals (UT) can be portable, mobile or fixed. They have been equipped to operate in both transmission and reception in Ka-band. They are able both to receive (down-link) a signal in Digital Video Broadcasting (DVB) format from a satellite in geostationary orbit and to transmit (up-link) a Ka-band signal via the same satellite. The next two tables: Table 39 and Table 40 describe the terminal types and their characteristics.

Table 39: Terminal Types

Terminal Type	**User Class**	**Tx / Rx Rate (Mbps)**	**Frequency Tx / Rx (GHz)**
SaT-A	Residential User	0.16 / 32.768	29.5÷30.0 / 19.7÷20.2
SaT-B	Corporate/Professional User	0.512 / 32.768	29.5÷30.0 / 19.7÷20.2
SaT-C	Corporate Headquarters	2.048 / 32.768	29.5÷30.0 / 19.7÷20.2
PrT-A	Internet Access Point (POP, NAP)	8.192 / 32.768	27.5÷28.6 / 17.7÷18.8
PrT-B/GTW	Internet Exchange Point (IEP)	32.768 / 32.768	27.5÷28.6 / 17.7÷18.8

Table 40: Terminal Characteristics

Terminal Characteristics	**SaT – A**	**SaT – B**	**SaT - C**	**PrT-A**	**PrT-B/GTW**
Availability %	99.5	99.5	99.7	99.9	99.9
Type	Fixed	Fixed	Fixed	Fixed	Fixed
EIRP dBW	43	48	56	64	70
G/T dB/K (clear sky)	18.2	18.2	22.5	25.5	29.5
RF Power (W)	1	3	7	25	35
Antenna Size (m)	0,75	0,75	1,20	1,70	2,70

A.3.2 Connection between the Network Element

In the next Figure 14 two kinds of connections are available, one of them is marked with dots and the second uses a continuous line.

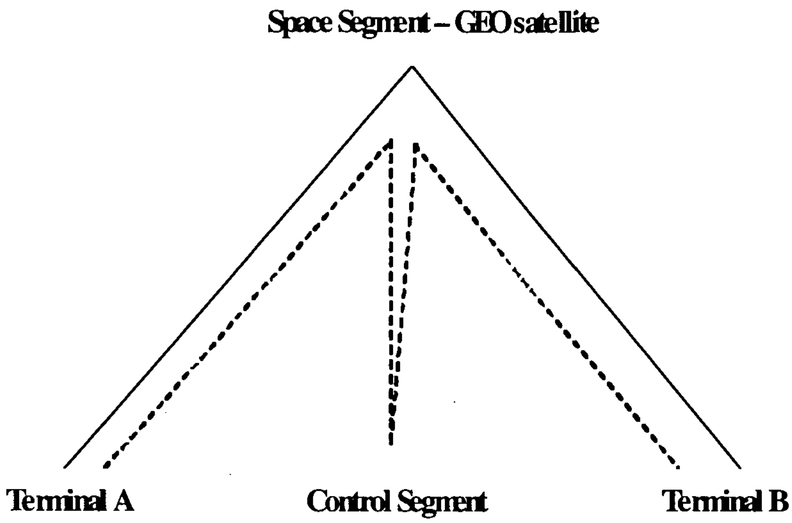


Figure 14: Signaling and Payload Connections

The dotted line is the signaling connection and the continuous line represents the payload connection. The signaling connection is between a terminal and the control segment, and serves for connection admission and for connection control, which contains the connection setup and release. After the connection admission and the connection setup is successful, the terminal can exchange any kind of data, the actual payload, between itself and the second terminal, this connection is the continuous line of the Figure.

The whole available bandwidth has to be divided into traffic, payload data and signaling data. The signaling data is used by the users of multimedia satellite system in order to transmit call specific signaling in order to establish and release calls. The signaling bandwidth has to be dimensioned correctly to avoid bottlenecks. The total amount of signaling data depends on the number and size of signaling messages per call establishment and release and the call attempt rate per satellite terminal. Therefore it is immense important for the analysis of VoIP Protocols to consider the message size and the message number of the protocol in order to keep the signaling data low.

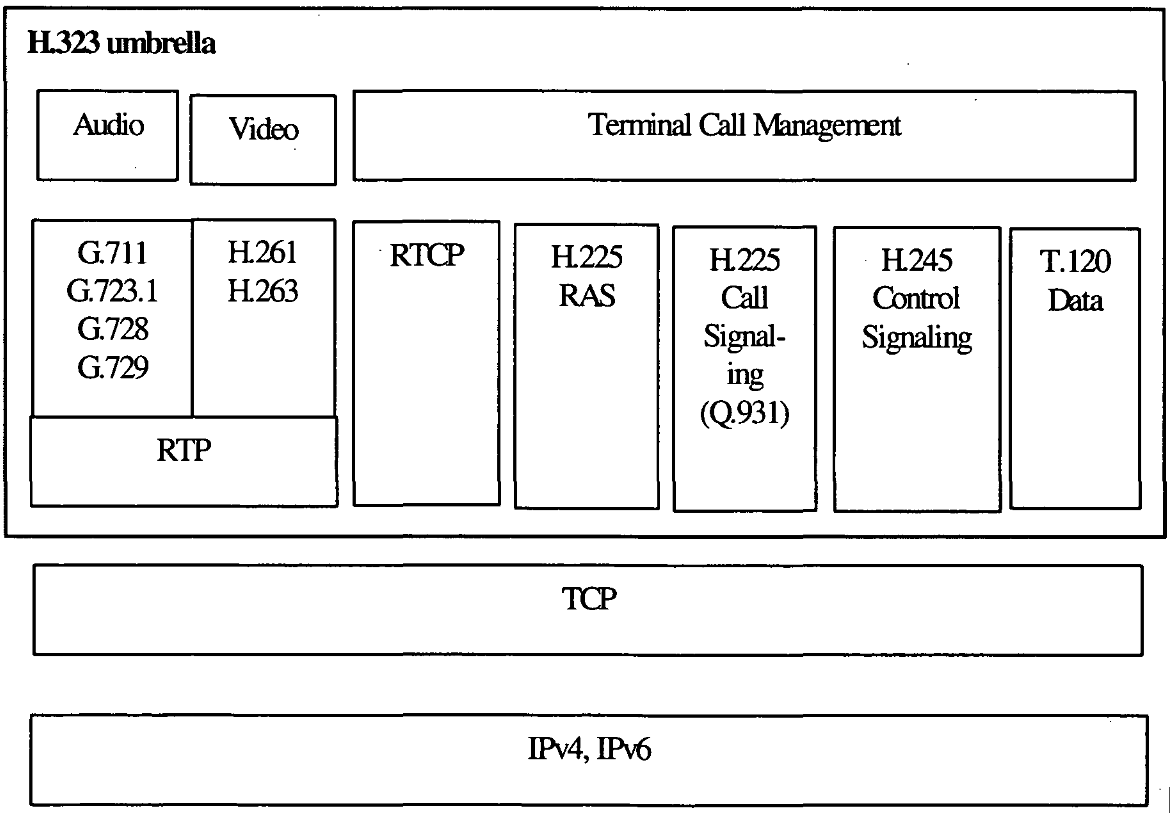
Annex B H.323 Details

B.1 H.323 Protocol Stack

H.323 is an umbrella recommendation and defines how audio and videoconferencing systems communicate over packet-switched networks that do not guarantee Quality of Service (QoS), such as the Internet and Intranets. It includes parts of

- H.225.0 [49], Q.931 [68], H.245 [50] for H.323 [54] signaling, which is transported reliably over TCP [22], because H.323 has no reliability mechanism.
- RTP/RTCP [26]/[43] for payload transport. Media streams are transported on RTP/RTCP. RTP carries the actual media and RTCP carries status and control information.
- Audio/video codecs, such as the audio codecs (G.711 [44], G.723.1 [46], G.728 [47], etc.) and video codecs (H.261 [51], H.263 [53]) that compress and decompress media streams.

The following Figure 15 present the H.323 umbrella recommendations.



[54]

Figure 15: H.323 umbrella recommendation architecture

B.2 H.323 Protocol Components and Architecture

B.2.1 H.323 Protocol Components

The H.323 Protocol has following component:

Terminals: Used for real-time bi-directional multimedia communications, a H.323 terminals can either be an application running on a personal computer (PC) or a stand-alone device, running a H.323 application and the multimedia applications. It supports audio communications and can optionally support video or data communications.

Gatekeepers: Gatekeepers provide call control services for H.323 endpoints, such as addressing, authorization and authentication of terminals and gateways, bandwidth management, accounting, billing, and charging. Gatekeepers may also provide call-routing services. A gatekeeper can be considered the brain of the H.323 network. Gatekeepers in H.323 networks are optional. If they are present in a network, however, endpoints must use their services. The H.323 standards define mandatory services that the gatekeeper must provide and optional functionality that it can provide. The mandatory Gatekeeper functions are following:

1. Address translation: translation of alias addresses to Transport Addresses using a table that is updated with registration messages. Other methods of updating the translation table are also allowed.
2. Admission Control: authorization of LAN access, which may be based on call authorization, bandwidth, or some other criteria. Admissions Control may also be a null function which admits all requests.
3. Bandwidth Control: based on bandwidth management. Bandwidth Control may also be a null function which accepts all requests for bandwidth changes.
4. Zone Management: the Gatekeeper provides the above functions for terminals, MCUs, and Gateways which have registered within its Zone of control.

An optional feature of a gatekeeper is the routing of call signaling messages between H.323 endpoints. Endpoints may send call-signaling messages to the gatekeeper, and the gatekeeper routes them to the destination endpoints. Alternatively, endpoints may send call-signaling messages directly to the peer endpoints. Routing calls through gatekeepers provides better performance in the network, as the gatekeeper can make routing decisions based on a variety of factors, for example, load balancing among gateways.

Gateways: A gateway connects two dissimilar networks. An H.323 gateway provides connectivity between an H.323 network and a non–H.323 network. This connectivity of dissimilar networks is achieved by translating protocols for call setup and release, converting media formats between different networks, and transferring information between the networks connected by the gateway. A gateway is not required for communication between two terminals on an H.323 network.

Multipoint Control Units (MCUs): MCUs provide support for conferences of three or more H.323 terminals. All terminals participating in the conference establish a connection with the MCU. The MCU manages conference resources and negotiates between terminals.

The gatekeepers, gateways, and MCUs are logically separate components of the H.323 standard but can be implemented as a single physical device.

B.2.2 H.323 Protocol Architecture

An H.323 architecture can be seen as a zone, with only one gatekeeper. A zone may be independent of network topology and may be comprised of multiple network segments that are connected using routers or other devices. The H.323 architecture runs on top of common network architectures and independently from any hardware or operating system. The Figure 16 below illustrates these entities and their logical relationship.

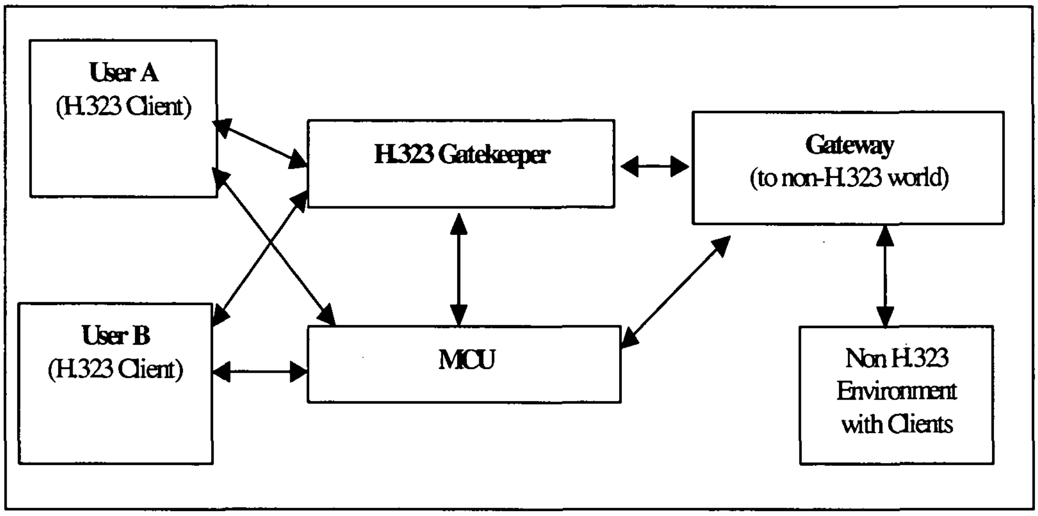


Figure 16: H.323 Protocol Architecture

B.3 H.323 Protocol Services and Messages

B.3.1 H.323 Protocol Services

B.3.1.1 H.225 RAS Services

The RAS (Registration, Admission and Status) Channel is a signaling channel, which is opened between endpoints (terminals, gateways, MCUs) and gatekeepers. It is used to carry messages used in the Gatekeeper discovery and endpoint registration processes. Furthermore, the RAS channel is used for other kinds of control mechanisms, such as admission control, to restrict the entry of an endpoint into a zone, bandwidth control, and disengagement control, where an endpoint is disassociated from a gatekeeper and its zone. The RAS Channel is an unreliable channel, which allows the usage of a protocol such as UDP. In case of packet lost the H.323 peer will not get an answer for it sent RAS message but a timeout. Consequently the H.323 peer has to send the RAS message again.

H.225, RAS services:

1. Gatekeeper discovery: the gatekeeper discovery process is used by the H.323 endpoints to determine the gatekeeper with which the endpoint must register. The gatekeeper discovery can be done statically or dynamically. In static discovery, the endpoint knows the transport address of its gatekeeper a priori. In the dynamic method of gatekeeper discovery, the endpoint multicasts a Gatekeeper Request (GRQ) message on the gatekeeper's discovery multicast address: "Who is my gatekeeper?" One or more gatekeepers may respond with a Gatekeeper Confirmation (GCF) message: "I can be your gatekeeper." and returns the Transport Address of the Gatekeeper's RAS Channel. If a Gatekeeper does not want the endpoint to register to it, it will return Gatekeeper Reject (GRJ).
2. Endpoint registration: registration is a process used by the endpoints to join a zone and inform the gatekeeper of the zone's transport and alias addresses. All endpoints register with a gatekeeper as part of their configuration. Registration has to occur before any calls are attempted and may occur periodically as necessary. An endpoint shall send a Registration Request (RRQ) to a Gatekeeper. This is sent to the Gatekeeper's RAS Channel Transport Address. The endpoint has the Network Address of the Gatekeeper from the Gatekeeper discovery process. The Gatekeeper shall respond with either a Registration Confirmation (RCF) or a Registration Reject (RRJ).
3. Endpoint location: an endpoint which has an alias address and would like to determine its contact information, should issue a Location Request (LRQ) message. This message may be sent to a specific Gatekeeper's RAS Channel or may be multicast to the Gatekeeper's well-known Discovery Multicast Address. The Gatekeeper with which the requested endpoint is registered will respond with the Location Confirmation (LCF) message containing the contact information of the endpoint. Contact information shall include the Call Signaling Channel and RAS Channel addresses and optionally additional destination information with dialing and extension information concerning the requested endpoint. All Gatekeepers with which the requested endpoint is not registered will return Location Reject (LRJ). Any Gatekeeper with which the requested endpoint is not registered will not respond to the LRQ, if it received the LRQ on the Discovery Multicast address.
4. Admissions, bandwidth change, status and disengage: the RAS Channel is also used for the transmission of Admissions, Bandwidth Change, Status and Disengage messages. These messages take place between an endpoint and a Gatekeeper and are used to provide admissions control and bandwidth management functions. The Admissions Request (ARQ) message specifies the requested Call Bandwidth. This is an upper limit on the aggregate bit rate for all transmitted and received, audio and video channels excluding any RTP headers, RTP payload headers, network headers, and other overhead. Data (T.120) and control (H.245) channels are not included in this limit. The Gatekeeper could reduce the requested Call Bandwidth in the Admissions Confirm (ACF) message. An endpoint or the Gatekeeper may attempt to modify the

Call Bandwidth during a call using the Bandwidth Change Request (BRQ) message. The Disengage procedure uses the following messages: Disengage Request (DRQ), Disengage Confirm (DCF), Disengage Reject (DRJ) for call clearing between Terminal and Gatekeeper, because the Gatekeeper needs to know about the release of bandwidth, when a call is released between the terminals.

B.3.1.2 H.225 Call Signaling Services

H.225 call signaling is used to set up connections between H.323 endpoints (terminals and gateways), over which the real-time data can be transported. Call signaling involves the exchange of H.225 protocol messages over a reliable call-signaling channel. For example, H.225 protocol messages are carried over TCP in an IP–based H.323 network.

H.225 messages are exchanged between the endpoints if there is no gatekeeper in the H.323 network. If a gatekeeper exists in the network, the H.225 messages are exchanged either directly between the endpoints or between the endpoints after being routed through the gatekeeper. The first case is called direct call signaling. The second case is called gatekeeper-routed call signaling. The gatekeeper decides about the method to be used during the RAS-admission message exchange.

The Call Signaling Channel may carry signaling for many concurrent calls, using the call reference value to associate the message with the call. The Recommendation H.225.0 specifies the mandatory Q.931/H.450 messages that are used for call signaling, which are: Alerting, Call Proceeding, Connect, Setup, Release Complete, Status and Facility.

B.3.1.3 H.245 Control Signaling Services

H.245 control signaling consists of the exchange of end-to-end H.245 messages between communicating H.323 endpoints. The H.245 control messages are carried over H.245 control channels. The H.245 control channel is the logical channel 0 and is permanently open, unlike the media channels. The H.245 Control Channel carries end-to-end control messages governing operation of the H.323 entity, including capabilities exchange, opening and closing of logical channels, mode preference requests, flow control messages, and general commands and indications. Recommendation H.245 specifies a number of independent protocol entities which support endpoint-to-endpoint signaling. H.323 endpoints shall support the syntax, semantics, and procedures of the following protocol entities: Master/slave determination, Capability Exchange, Logical Channel Signaling, Bi-directional Logical Channel Signaling, Close Logical Channel Signaling, Mode Request, Round Trip Delay Determination and Maintenance Loop Signaling. The two most important procedures are:

1. Capabilities Exchange: capabilities exchange is a process using the communicating terminals exchange messages to provide their transmit and receive capabilities to the peer endpoint. The related messages are: TerminalCapabilitySet, TerminalCapabilitySetAck.
2. Logical Channel Signaling: a logical channel carries information from one endpoint to another endpoint (in the case of a point-to-point conference) or multiple endpoints (in the case of a point-to-multipoint conference). H.245 provides messages to open or close a logical channel; a

logical channel is unidirectional. The related messages are: OpenLogicalChannel, OpenLogicalChannelAck.

B.3.1.4 Basic Call Services

Depending on whether a gatekeeper is used or not, there are the following different variants of call setup procedures:

- Neither endpoint registered at Gatekeeper: the two endpoints communicate directly. Endpoints use the H.255 call signal with Q.931 messages, like Setup, Call Proceeding, Alerting and Connect, which contains an H.245 Control Channel Transport Address for use in H.245 signaling. No RAS messages will be used.
- Both endpoints registered to the same Gatekeeper: in this variant the gatekeeper can choose between direct call signaling or route the call signaling. In first case only the RAS messages will be exchanged between the Gatekeeper and the terminal, but the H.225 and H.245 messages will be exchanged directly between the two terminals. In the second case all messages (RAS, H.225, H.245) will be sent to the gatekeeper and the gatekeeper will route the messages to the second terminal.
- Only the calling endpoint has Gatekeeper: as in the scenario above there are two options: direct call signaling or route the call signaling. In both cases the calling terminal exchanges RAS messages with gatekeeper. But in the first case the calling endpoint sends the H.225 and H.245 messages directly to the called endpoint. In the second case the messages are routed through the gatekeeper. There is also the possibility that only the called endpoint has Gatekeeper.
- Both endpoints registered to different Gatekeepers: this scenario leads to four different ways how to communicate: (1) both gatekeepers assign direct call signaling, or (2) the callee endpoint gatekeeper chooses direct call signaling and the called endpoint gatekeeper routes call signaling, or (3) the called endpoint gatekeeper chooses direct call signaling and the callee endpoint gatekeeper routes call signaling, or (4) both gatekeepers assign routed call signaling.
- Fast Connect Procedure: H.323 endpoints may establish media channels in a call using the procedure defined in the H.245 Recommendation or the "Fast Connect" procedure. The Fast Connect procedure allows the endpoints to establish a basic point-to-point call with as few as one round-trip message exchange, enabling immediate media stream delivery upon call connection. The calling endpoint initiates the Fast Connect procedure by sending a Setup message containing the “fastStart” element to the called endpoint. The “fastStart” element consists of a sequence of OpenLogicalChannel structures describing media channels which the calling endpoint proposes to send and receive, including all of the parameters necessary to immediately open and begin transferring media on the channels. Using the Fast Connect procedure spares the H.245 call control signaling and the connection can be connected faster. Thus, the called endpoint must be prepared to immediately receive media on the channels it accepted in the Q.931 setup message containing fastStart. After receiving the connect message, the media conversation can begin.

After the conversation is finished a call termination procedure is needed. The process of call termination procedure is different using FastConnect or using the usual setup procedure. If the Fast Connect procedure is used, the call completion is done without initiation of H.245 procedures, the call will be terminated by either endpoint sending a Q.931 Release Complete message. If a usual setup procedure is used with H.245 messages, then the call is terminated with the transmission of H.245 closeLogicalChannel and endSessionCommand message in the H.245 Control Channel, indicating to the far end that it wishes to disconnect the call. The Q.931 Release Complete messages ends the disconnect procedure between the terminals. It is followed by the disengage through the RAS messages Disengage Request (DRQ) and Disengage Confirmation (DCF) from the gatekeeper.

B.3.1.5 Advanced Services

The H.323 advance services are composed of the H.450 supplementary services with following operations:

1. Call Transfer [56]: Call Transfer (SS-CT) enables a user to transfer an established call to another user.
2. Call Diversion (Call Forwarding) [57]: comprises the services Call Forwarding Unconditional (SS-CFU), Call Forwarding Busy (SS-CFB), Call Forwarding No Reply (SS-CFNR) and Call Deflection (SS-CD). These are supplementary services which apply during call establishment providing a diversion of an incoming call to another destination endpoint.
3. Call Hold [58]: the Call Hold (SS-HOLD) services allows the served user, which may be the originally calling or the called user, to interrupt communications on an existing call and then subsequently, if desired, re-establish (i.e. retrieve) communications with the held user. While having put the held user into a hold condition, the served user may perform other actions.
4. Call park and Pick up [59]: Call Park (SS-PARK) is a supplementary service that enables a user A (Parking User) to place an existing call with user B (Parked User) to a Parking Position (parked-to endpoint). Call Pickup (SS-PICKUP) is a supplementary service that enables a user (Picking-up User) to either retrieve a parked call (also from a different location) or to pick up an alerting call.
5. Call waiting [60]: the CW supplementary service (SS-CW) permits a busy user B to be informed of an incoming call while being engaged with one or more other calls.
6. Message Waiting Indication [61]: the Message Waiting Indication supplementary service provides a general purpose mechanism by which a user can be advised that messages intended for that user are available. A variety of message types are supported, such as voice mail, fax and teletex.

B.3.2 H.323 Protocol Messages

B.3.2.1 H.225 RAS Messages

The H.225, RAS messages are described in the H.225 recommendation [49]. As a summary, the completed list of mandatory messages is provided here:

- GRQ (Gatekeeper Request), GCF (Gatekeeper Confirm), GRJ (Gatekeeper Reject)
- RRQ (Registration Request), RCF (Registration Confirm), RRJ (Registration Reject)
- URQ (Unregistration Request), UCF(Unregistration Confirm), URJ (Unregistration Reject)
- ARQ (Admission Request), ACF (Admission Confirm), ARJ (Admission Reject)
- BRQ (Bandwidth Request), BCF (Bandwidth Confirm), BRJ (Bandwidth Reject)
- DRQ (Disengage Request), DCF (Disengage Confirm), DRJ (Disengage Reject)
- LRQ (Location Request), LCF (Location Confirm), LRJ (Location Reject)
- IRQ (Info Request), IRR (Info Request Response), IACK (Info Request Ack)
- RIP (RAS timers and Request in Progress)
- RAI (Resources Available Indicate), RAC (Resources Available Confirm)

B.3.2.2 H.225 Call Signaling Messages

The H.225, call signaling messages are described in the H.225 [49], Q.931 [68] and H.450 [55] recommendations. As a summary, the completed list of mandatory messages is provided here:

- Call Establish Messages:
    - Alerting - This message may be sent by the called user to indicate that called user alerting has been initiated. In everyday terms, the "phone is ringing."
    - Call Proceeding - This message may be sent by the called user to indicate that requested call establishment has been initiated and no further call establishment information will be accepted.
    - Connect - This message shall be sent by the called entity to the calling entity (gatekeeper, gateway, or calling terminal) to indicate acceptance of the call by the called entity.
    - Setup - This message shall be sent by a calling H.323 entity to indicate its desire to set up a connection to the called entity.
- Miscellaneous messages
    - Status - The Status message shall be used to respond to an unknown call signalling message or to a Status Inquiry message.
    - Status Inquiry - The Status Inquiry message will be used to request call status.
- Q.932/H.450 messages
    - Facility - The Facility message shall be used to provide information on where a call should be directed (FacilityReason = routeCallToMC), or for an endpoint to indicate that the incoming call must go through a gatekeeper (FacilityReason = routeCallToGatekeeper).
- Call Clearing messages
    - Release Complete - This message shall be sent by a terminal to indicate release of the call if the reliable call signaling channel is open. Afterwards, the Call Reference Value (CRV) is available for reuse. The disconnect message Release Complete has the Cause IE

parameter, with the following releaseCompleteReason codes: noBandwidth, gatekeeperResources, unreachableDestination, destinationRejection, invalidRevision, noPermission, unreachableGatekeeper, gatewayResources, badFormatAddress, adaptiveBusy, inConf, undefinedReason, facilityCallDeflection, securityDenied, calledPartyNotRegistered, callerNotRegistered.

B.3.2.3 H.245 Control Signaling Messages

The H.245, control signaling messages are described in the H.245 recommendation [50]. As a summary, the completed list of mandatory messages is provided here:

- Master-Slave Determination Messages: Determination, Determination Acknowledge, Determination Reject, Determination Release.
- Terminal Capability Messages: Capability Set, Capability Set Acknowledge, Capability Set Reject, Capability Set Release.
- Logical Channel Signaling Messages: Open Logical Channel, Open Logical Channel Acknowledge, Open Logical Channel Reject, Open Logical Channel Confirm, Close Logical Channel, Close Logical Channel Acknowledge, Request Channel Close, Request Channel Close Acknowledge, Request Channel Close Reject, Request Channel Close Release.
- Request Mode Messages: Request Mode, Request Mode Acknowledge, Request Mode Reject, Request Mode Release
- Round Trip Delay Messages: Round Trip Delay Request, Round Trip Delay Response
- Maintenance Loop Messages: Maintenance Loop Reject, Maintenance Loop Command Off
- Commands: Send Terminal Capability Set, Encryption, Flow Control, End Session, Multipoint Mode Command, Cancel Multipoint Mode Command, Video Freeze Picture, Video Fast Update Picture, Video Fast Update GOB, Video Fast Update MB
- Indications: Function Not Understood, Function Not Supported, Terminal Number Assign, Multipoint Conference, Cancel Multipoint Conference, MC Location Indication, H2250MaximumSkewIndication

Annex C SIP Details

C.1 SIP Protocol Stack

SIP is designed as part of the overall IETF multimedia data and control architecture, currently incorporating protocols such as

- the Resource Reservation Protocol (RSVP, RFC 2205) [28] for reserving network resources
- the Real-time Transport Protocol (RTP, RFC 1889) [26] for transporting real-time data and providing QoS feedback
- the Real-time Streaming Protocol (RTSP, RFC 2326) [30] for controlling delivery of streaming media
- the Session Announcement protocol (SAP, RFC 2974) [38] for advertising multimedia sessions via multicast,
- the Session Description Protocol (SDP, RFC 2327) [31] for describing multimedia sessions.

However, the functionality and operation of SIP does not depend on any of these protocols. SIP can also be used in conjunction with other call setup and signaling protocols. In that mode, an end system uses SIP exchanges to determine the appropriate end system address and protocol from a given address that is protocol-independent. The following Figure 17 shows the described protocols placed in the Open Systems Interconnection (OSI) Model.

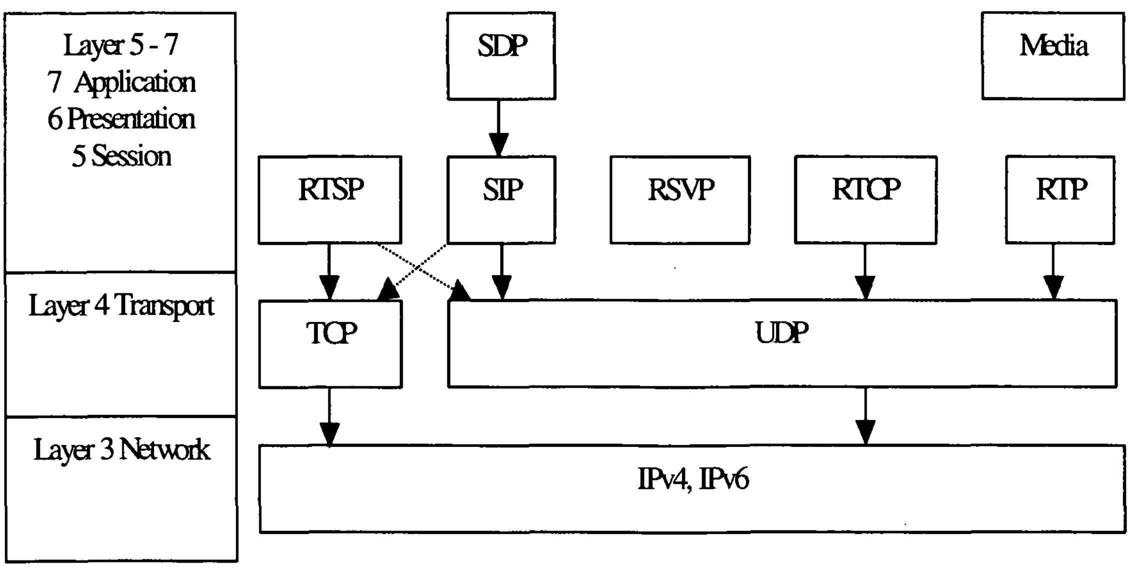


Figure 17: SIP Protocols placed in OSI Model

C.2 SIP Protocol Components and Architecture

C.2.1 SIP Protocol Components

The SIP Protocol has following component:
Client: an application program that sends SIP requests and receives a SIP response from the SIP server.

Location Server, Location Service: a location service is used by a SIP redirect or proxy server to obtain information about a callee's possible location(s). Location services are offered by location servers. Location servers may be co-located with a SIP server, mostly as a database, where the location information of each user is saved.

Proxy, Proxy Server: an intermediate device that can act as both, a server and a client, for the purpose of making requests on behalf of other clients. Requests are serviced internally or by passing them on, possibly after translation, to other servers. A proxy interprets, and, if necessary, rewrites a request message before forwarding it. The right proxy server does not pass the received request to the next proxy, but forwards the request to the current location of the callee. The Proxy server can forward the invitation to multiple servers at once, trying to contact the user at one of the locations. Summarizing, proxies can provide functionality such as authentication, authorization, network access control, routing, reliable request retransmission and security.

Redirect Server: a redirect server is a server that accepts a SIP request, maps the address into zero or more new addresses and returns these addresses to the client. Thus it provides the client with information about the next hop(s) that a message should take. Unlike a proxy server, it does not initiate its own SIP request. Unlike a user agent server, it does not accept calls.

Registrar, Registration Server: a registrar is a server that accepts REGISTER requests from UACs for registration of their current location. A registrar is typically co-located with a proxy or redirect server.

User Agent Client (UAC), Calling User Agent: a user agent client is a client application that initiates the SIP request.

User Agent Server (UAS), Called User Agent: a user agent server is a server application that contacts the user when a SIP request is received and that returns a response on behalf of the user. The response accepts, rejects or redirects the request.

C.2.2 SIP Protocol Architecture

From an architectural point of view the physical components of a SIP network are clients and servers. From a logical point of view SIP is a peer-to-peer protocol, whereby the peers are called User agents (UAs). A user agent can adopt one of the following roles:

- User agent client (UAC): a client application which initiates a SIP request
- User agent server (UAS): a server application which contacts a user, when a SIP request is received or responds on behalf of the user

A SIP end point can take both roles and function as UAS as well as UAC, depending on the UA that initiated the request. The SIP architecture contains the following components: User Agent Client, User Agent Server, Registrar (Registration Server), Location Server, Proxy Server and Redirect Server. The following Figure 18 shows the relationships between these components and a typical SIP network architecture.

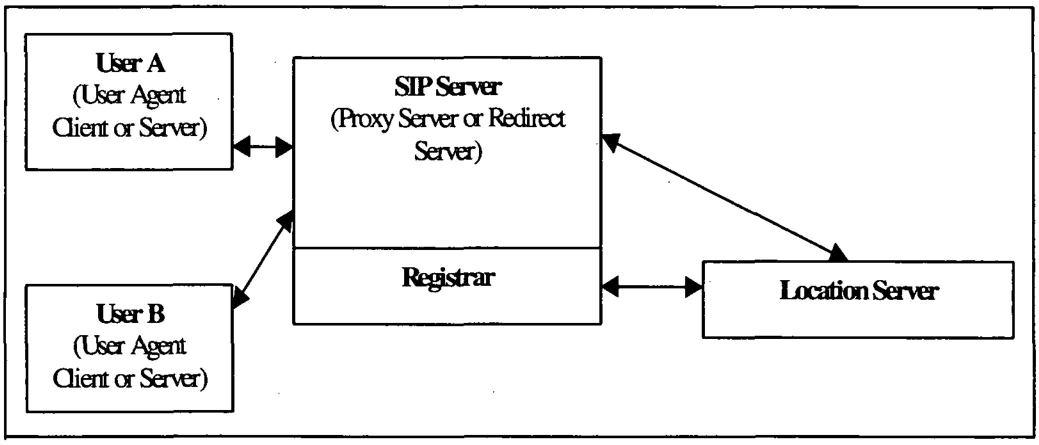


Figure 18: SIP Network Architecture and Components

SIP clients may be

- Phones: acting as UAS or UAC initiating or responding to SIP requests
- Gateways: providing call control, translation of transmission formats between SIP conferencing endpoints and between communication procedures, translation between audio and video codecs, call establishment and clearing on the LAN side and the circuit-switched network side

SIP servers comprise the following types:

- Proxy servers
- Redirect servers
- Registrar servers
- Location Server

Based upon the architecture the following two scenarios are possible: an architecture with a proxy server and an architecture with a redirect server. Figure 19 and Figure 20 show these two variants:

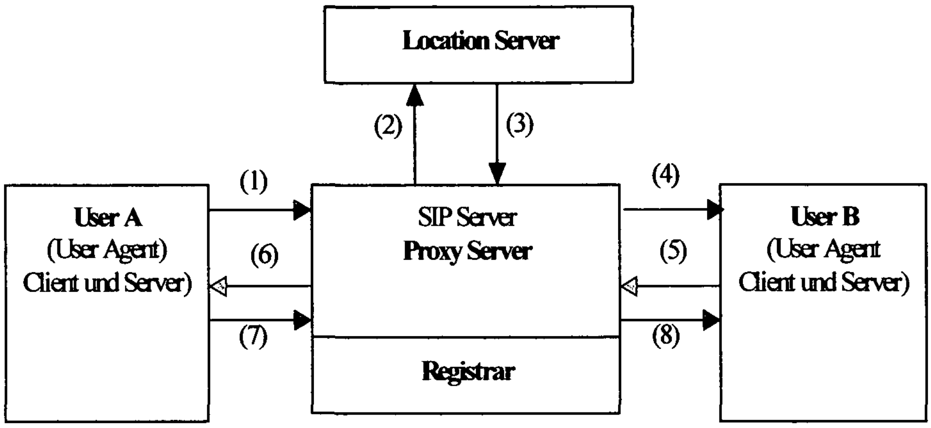


Figure 19: SIP Basic Call with Proxy Server

In Figure 19, User A sends an INVITE request to his proxy server (1). The proxy server accepts the invitation and contacts his location server for a more precise location (2). The location server returns the locations of User B (3). The proxy server issues an INVITE request to the address given by the location server (4). The user agent server at User B alerts the user B, who is willing to accept the call. The acceptance is returned to the proxy server, by a 200 OK response (5). A success response is then sent by the proxy to the original caller (6). This message is confirmed by the caller with an ACK request (7). The ACK request is forwarded to the callee (8).

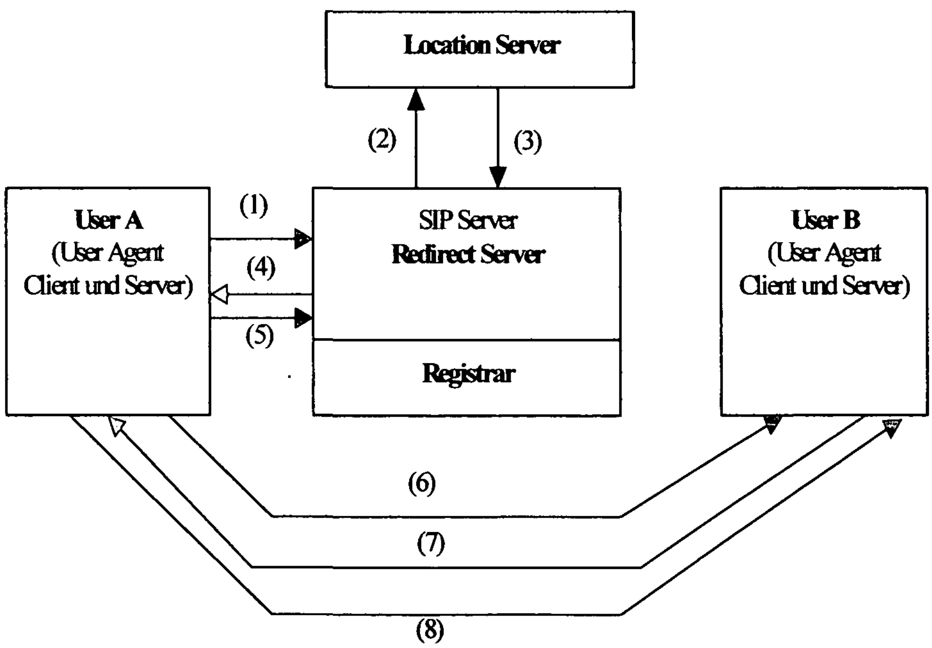


Figure 20: SIP Basic Call with Redirect Server

In Figure 20, User A sends an INVITE request to its redirect server (1). The redirect server accepts the invitation and contacts his location server for a more precise location (2). The location server returns the locations of User B (3). The redirect server returns a redirection response of class 3xx[v] storing the address to contact in the Contact header field (4). The caller acknowledges the response with an ACK request to the server (5). The caller issues a new INVITE request with the same Call-ID but a higher Command Sequence (CSeq) number. This is sent to the address given by the server to the location of User B (6). User B is willing to accept the call. The acceptance is returned to User A, by a 200 status response (7). Finally, the caller confirms the successful response with an ACK request and the call is connected (8).

The communication between the users and the SIP Server (Proxy or Redirect) is based on SIP messages, defined in the SIP standard (RFC3261) [40]. SIP messages are also used between User and Registrar. The SIP Server and Registrar are mostly co-located and have their own communication interface. For the communication between the SIP server and the location server the following protocols can be used: Lightweight Directory Access Protocol (LDAP) (RFC 1777) [24], finger (RFC 1288) [20] or rwhois (RFC 2167) [27].

C.3 SIP Protocol Services and Messages

C.3.1 SIP Protocol Services

C.3.1.1 Basic Services

Basic Services contains the registration of a user on the SIP Registrar, the Localization of a user through the SIP server and the establishment of a call between two users. Summery of basic services:

- Addressing and Naming: to be invited to a SIP session and identified, the called party has to be named. Since it is the most common form of user addressing in the Internet, SIP choses an email-like identifier of the form “user@domain”, “user@host”, “user@IP address” or “phone-number@gateway”.
- Locating a Server: when a client wishes to send a request, it first obtains the address of the participant to be contacted. If the address consists of a numeric IP address, the client contacts the SIP server there. If otherwise the address is of the form name@domain, the client has to translate the domain part to an IP address where a server may be found. This is done with a Domain Name System (DNS) lookup.
- Locating a User: when the SIP server receives a request it has to locate the user in its domain. The user’s location could be of different kind. He could for example be logged in at zero to

[v] A class 3xx response can be "300" - Multiple Choices, "301" - Moved Permanently, "302" - Moved Temporarily, "305" - Use Proxy, "380" - Alternative Service.

many hosts or at a different domain. To find the user's current location there is an outside SIP entity, a location server. Being asked for a user's location the location server returns a list of zero to many locations where the user could be found. The location can dynamically be registered with the SIP server. This is done by sending a REGISTER request, a registration procedure.

- Registration: the REGISTER request is used to convey location information to a SIP server. The REGISTER request allows a client to inform the proxy or redirect server about the address(es) where it can be reached. Clients can register from different locations. A server (Registrar) can choose any duration for the registration lifetime. Registrations not refreshed after this period of time will be silently discarded. Responses to a registration will include an Expires header or expires Contact parameters, indicating the time when the server will drop the registration. If none is present, one hour is assumed per default. Clients can request a registration lifetime by indicating the time in an Expires header in the request. A server will never use a higher lifetime than the one requested, but may use a lower one. A client can cancel an existing registration by sending a REGISTER request with an expiration time (Expires) of zero seconds.
- Basic Call: a successful SIP call, also called invitation consists of two requests, INVITE followed by ACK. The INVITE request asks the callee to join a particular conference or establish a two-party conversation. After the callee has agreed to participate in the call, the caller confirms that it has received that response by sending an ACK request. If the caller no longer wants to participate in the call, it sends a BYE request instead of an ACK.
- Changing an Existing Session: under some circumstances, it is desirable to modify the parameters of an existing session. This is done by re-issuing the INVITE request or with OPTIONS request, using the same Call-ID header of the request, but a new or different body or header fields to convey the new information. This re-INVITE must have a higher Cseq header in the request than any previous request from the client to the server. For example, two parties may have been conversing and then want to add a third party, switching to multicast for efficiency. One of the participants invites the third party with the new multicast address and simultaneously sends an INVITE to the second party, with the new multicast session description, but with the old call identifier.

C.3.1.2 Advanced Services

Advanced services include group invitation, call forwarding and user location, call transfer, call hold and conference. Summery of basic services:

- Group Invitation: existing telephony signaling protocols typically only support the invitation of a single individual. SIP aims at also allowing calls that reach either the first available individual from a group (also called reach-first), similar to automatic call distribution (ACD), or a whole group of callees (also called reach-all), e.g., the whole department. This can be done in two ways: unicast: each call reaches a single user individual; this simplifies error reporting and the client state machine; or multicast: instead of calling sequential each callee party, a multicast invitation can be sent.

- Call Forwarding: SIP generalizes call forwarding in these and other observable conditions, e.g. the result of a user location query or user preference. User preference can be expressed as rules or manual response to a call, e.g., by clicking on a "do not disturb" button when a call arrives. The forwarding location can be determined in a way, that a SIP user agent informs the SIP server of the domain of its presence via a REGISTER message. If the user preference is expressed as a rule, the possibility of using Call Processing Language (CPL) [37] for call forwarding is given. The callee's user can create and placed a redirection response, in form of a script, for the SIP server, based on any number of reasons such as the caller, the time of day or availability of callee. When a call arrives for that callee (user), the SIP server executes the script to arrive at a forwarding decision.
- Conference: If a two party call is established and one of the participants wants to invite a third party to the call, this can be done in a way that every participant has a signaling session with each other. This type of conference is called fully meshed unicast conference. Any participant can leave the conference at any time without terminating it. The conference terminates when the last two parties terminate their call. If a conference starts as fully meshed unicast conference it may become cumbersome as the number of participants grows. In this case it is possible to transform the fully meshed unicast conference to a multicast conference by redirecting all the single conference connections to a conference bridge. After this it is still possible for each participant to invite new parties to the conference, even during the transition period. Still everybody can leave the conference at any time without doing any harm to it including the transition period.

Other advance services such as call hold or call transfer are not defined in the SIP Standard, but there exits several Internet Draft which gives a proposal how to implement such services.

C.3.2 SIP Protocol Messages

In principle, there are two kind of SIP messages: request and response. Clients issue requests and servers answer with responses. In both types of SIP messages different headers describe the details of the communication. Request and response use a generic-message format, which consists of a start-line, one or more header-fields ("headers"), an empty line indicating the end of the header fields, and an optional message-body.

C.3.2.1 SIP Request

The request is characterized by the Start-Line, which is also called Request-Line. It starts with a method token followed by a Request-URI and the protocol version. There are six different kinds of requests in the current version of SIP (version 2.0). They are referred to as methods and are listed here with their functionality.

- INVITE: The INVITE method indicates that the user or service is being invited to participate in a session. For a two-party call, the caller indicates the type of media it is able to receive as well as parameters such as the network destination. A success response indicates in its message body which media the callee wishes to receive.
- ACK: The ACK request confirms that the client has received a final response to an INVITE. It may contain a message body with the final session description to be used by the callee. If the message body is empty, the callee uses the session description in the INVITE request. This method is only used with the INVITE request.
- BYE: The user agent client (UAC) uses BYE to indicate to the user agent server (UAS) that it wishes to release the call.
- CANCEL: The CANCEL request cancels a pending request, but does not affect a completed request. (A request is considered completed if the server has returned a final response.)
- OPTIONS: The OPTIONS method solicits information about capabilities, but does not set up a connection.
- REGISTER: Conveys information about a user's location to a SIP server.

C.3.2.2 SIP Response

After receiving and interpreting a request message, the recipient responds with a SIP response message, indicating the status of the server, which may be success or failure. The responses can be of different kinds and the type of response is identified by a status code, a 3-digit integer. The first digit defines the class of the response. The other two have no categorization role. The six different classes that are allowed in SIP are listed below with their meaning. These classes can be categorized by provisional and final responses. A provisional response is used by the server to indicate progress, but does not terminate a SIP request. A final response terminates a SIP request.
Response Codes:

- *1xx: Informational (provisional responses)*
  "100" - Trying
  "180" - Ringing
  "181" - Call Is Being Forwarded
  "182" - Queued
  "183" - Session Progress
- *2xx: Success*
  "200" - Ok
- *3xx: Redirection*
  "300" - Multiple Choices
  "301" - Moved Permanently
  "302" - Moved Temporarily
  "305" - Use Proxy
  "380" - Alternative Service
- *4xx: Client Error*

"400" - Bad Request
"401" - Unauthorized
"402" - Payment Required
"403" - Forbidden
"404" - Not Found
"405" - Method Not Allowed
"406" - Not Acceptable
"407" - Proxy Authentication Required
"408" - Request Timeout
"409" - Conflict
"410" - Gone
"411" - Length Required
"413" - Request Entity Too Large
"414" - Request-URI Too Large
"415" - Unsupported Media Type
"420" - Bad Extension
"480" - Temporarily not available
"481" - Call Leg/Transaction Does Not Exist
"482" - Loop Detected
"483" - Too Many Hops
"484" - Address Incomplete
"485" - Ambiguous
"486" - Busy Here
"487" - Request Cancelled
"488" - Not Acceptable Here

- *5xx: Server Error*

"500" - Internal Server Error
"501" - Not Implemented
"502" - Bad Gateway
"503" - Service Unavailable
"504" - Gateway Time-out
"505" - SIP Version not supported

- *6xx: Global Failure*

"600" - Busy Everywhere
"603" - Decline
"604" - Does not exist anywhere
"606" - Not Acceptable

C.3.2.3 Message Body – Session Description Protocol (SDP)

The purpose of SDP is to convey information about media streams in multimedia sessions to allow the recipients of a session description to participate in a session. Indeed SDP is a data format rather than a protocol, because is contains a format to describe a session and it is used in the message body of a real signaling protocol. Thus SDP includes the Session name and purpose, the Time(s) the session is active, the media comprising the session, the media destination (e.g. IP addresses, port numbers, formats and so on). Additional information may be: information about the bandwidth to be used by the conference and contact information for the person responsible for the session.

An SDP session description consists of a number of lines of text of the form <type>=<value>. The <type> is always exactly one character and is case-significant. The <value> is a structured text string whose format depends on the <type>. It will also be case-significant unless a specific field defines otherwise. The mandatory types are:

- Session description: v (Protocol version), o (Owner/creator and session identifier), s (Session name).
- Time description: t (Time the session is active)
- Media description: m (Media name and transport address)

An SDP parser must completely ignore any announcement that contains a <type> letter that it does not understand.

Annex D H.248 Details

D.1 Megaco Protocol Stack

The transport mechanism for the protocol should allow reliable transport of transactions between a MGC and MG. Transport shall remain independent of particular commands which are being sent and shall be applicable to all application states. For transport of the protocol over IP, MGCs shall implement both TCP and UDP, a MG shall implement TCP or UDP or both. The following Figure 21 shows the Megaco protocol placed in the OSI Model.

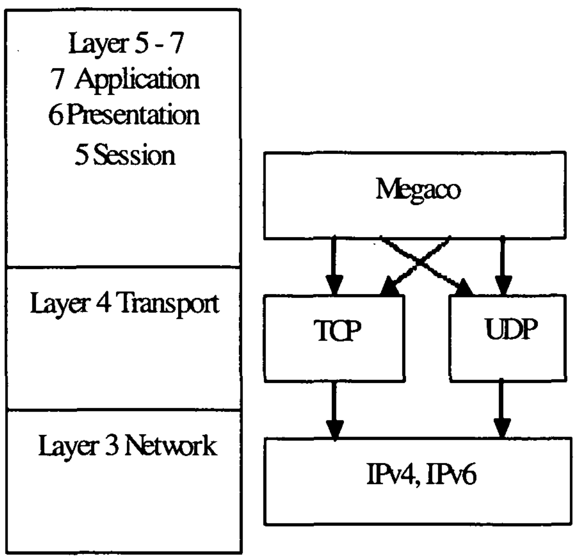


Figure 21: MEGACO Protocol Stack

D.2 Megaco Protocol Components and Architecture

D.2.1 Megaco Protocol Components

Media Gateway (MG): the MG is the entity responsible for monitoring and controlling the media endpoints, media connections and media resources at the edge of a Switched Circuit Network (SCN) and a Packet Data Network (PDN). Typically, the MG interfaces with the SCN via E1/T1 digital trunks and with the PDN via a LAN or WAN interface. Essentially the MG converts media provided in one type of network to the format required in another. For example, a MG could terminate bearer channels from a switched circuit network (e.g., DS0s) and media streams from a packet network (e.g., RTP streams in an IP network). This gateway may be capable of processing audio, video and T.120 alone or in any combination, and is capable of full duplex media translations. The MG may also play audio/video messages and perform other Interactive Voice

Response (IVR) functions, or may perform media conferencing. This layer has no knowledge of call level features and acts as a simple slave.

Media Gateway Controller (MGC): the Media Gateway Controller controls the Media Gateway and parts of the call state that pertain to connection control for media channels in a MG. The MGC contains all call control intelligence and implements call level features such as forward, transfer, conference and hold. This layer also implements any peer-level protocols for interaction with other MGCs or peer entities, manages all feature interactions, and manages any interactions with signaling such as SS7, H323 or SIP.

Signaling Gateway (SG): the signaling gateway is responsible for signaling termination and transport of signaling information to the MGC.

D.2.2 Megaco Protocol Architecture

The communication interface and the architectural decomposition of Voice over IP (VoIP) gateways into Signaling Gateway (SG), Media Gateway Controller (MGC) and Media Gateway (MG) are shown in Figure 22.

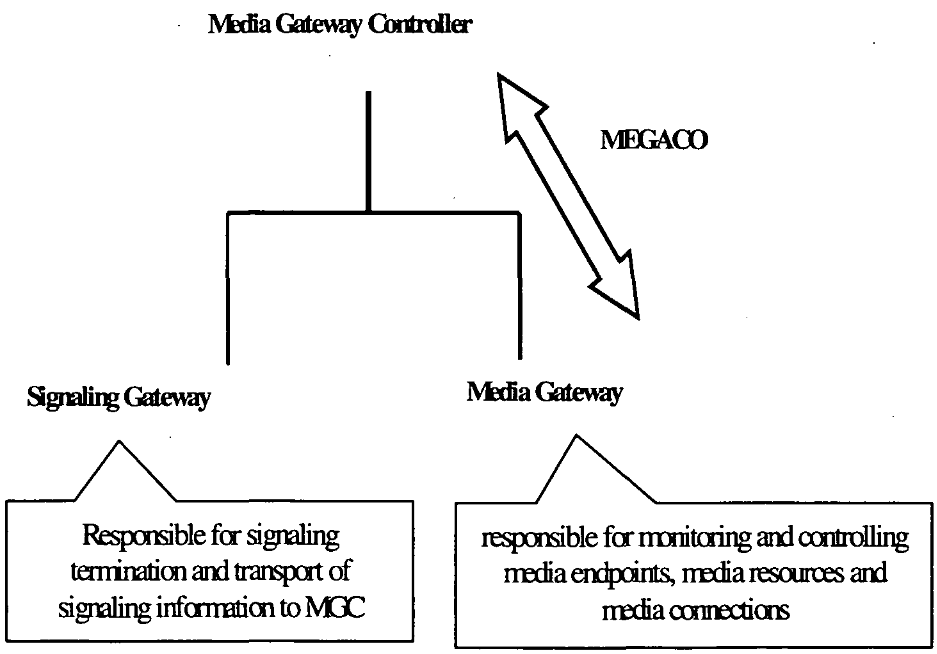


Figure 22: MEGACO Communication Architecture

Recommendation H.248 defines the protocols used between elements of a physically decomposed multimedia gateway, used in accordance with the architecture as specified in Recommendation H.323. There are no functional differences from a system view between a decomposed gateway with distributed sub-components potentially on more than one physical device, and a monolithic gateway.

The following Figure 23 shows the architecture in a simply way. The call signaling is SS7 and will go from the user terminal (telephone 1) through the Central Office and the Signaling Gateway, which communicates with the Media Gateway Controller to apply for media channels, to the second user terminal (telephone 2). The Central Office can be a PBX or a private telephone exchange. The MGC controls the MG through Megaco/H.248 protocol and supports the channel through the MG for the media payload. For Call Signaling any other protocol can be used, depending on the network type. In our scenarios SS7 is used, other protocols could be DSS1, H.323 or SIP. In our scenario on both sides SS7 is used. Using the same signaling protocol on both side simplifies the scenario, because no signaling translation from one protocol (for example DSS1) into the other (H.323) is necessary, which is also no functionality of Megao.

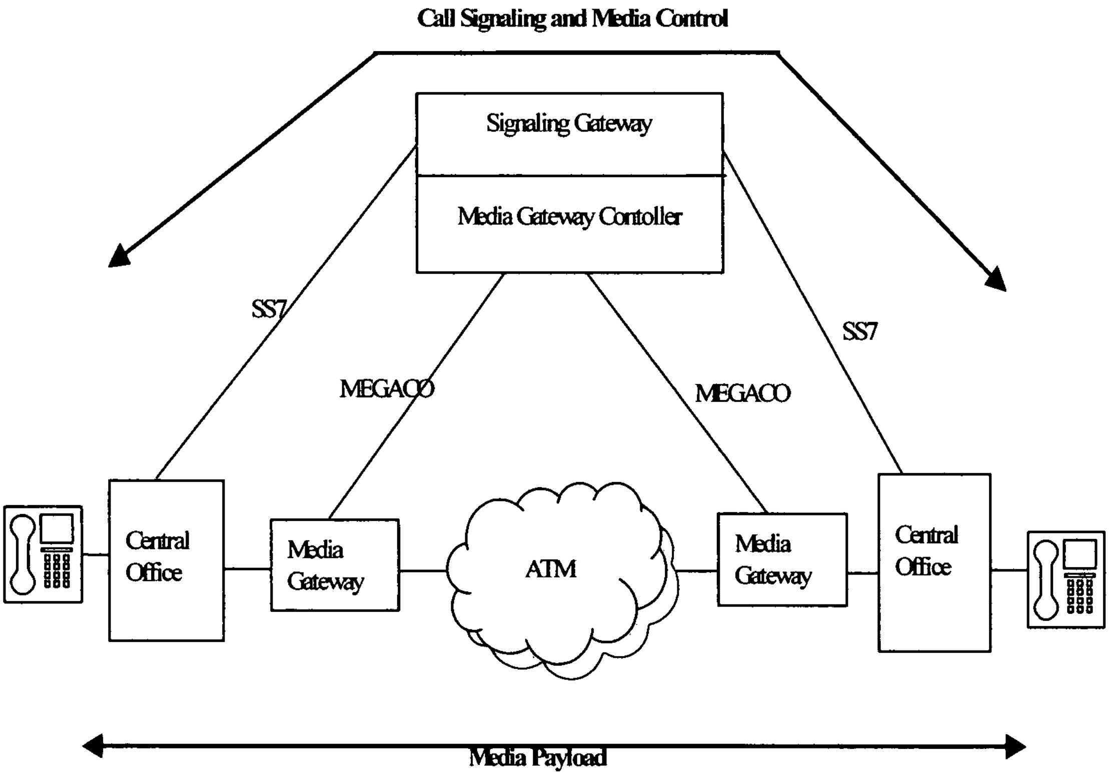


Figure 23: MEGACO Architecture

D.3 Megaco Protocol Services and Messages

D.3.1 Megaco Protocol Services

D.3.1.1 Basic Services

The basic service and key features of Megaco is the connection model. The connection model for the protocol describes the logical entities or objects within the MG that can be controlled by the MGC. The key terms used in the connection model are Terminations and Contexts.

A Termination is a logical entity in the MG that origins and/or terminates one or more media streams. In a multimedia conference, multiple media streams can be originated and/or terminated. A Termination is described by a number of characteristics, which are grouped in a set of Descriptors that are included in commands. Terminations have unique identities assigned by the MG at the time of their creation. Terminations may have signals applied to them. Signals are MG generated media streams - tones and announcements - as well as line signals such as on/off hook. Terminations may be programmed to detect Events which can trigger notification messages to the MGC, or actions by the MG.

A Context is an association between a collection of Terminations. The Context describes the topology and the media mixing and/or switching parameters if more than two Terminations are involved in the association. A Context is created by the *Add* of its first Termination and is destroyed by the *Subtract* of its last Termination. A Termination that is associated to a Context can be moved to another Termination using the command *Move*. Contexts are deleted implicitly when the last remaining Termination is subtracted or moved out. A typical Context for a two-party call in Megaco would contain two Terminations, one representing a PSTN trunk (DSO) and the other an RTP Stream Termination, as shown in Figure 24. Both Terminations have to be explicitly added to the Context. There is a special type of Context, the Null Context, which contains all Terminations that are not associated to any other Terminations. For example, all idle lines are represented by Terminations in a Null Context.

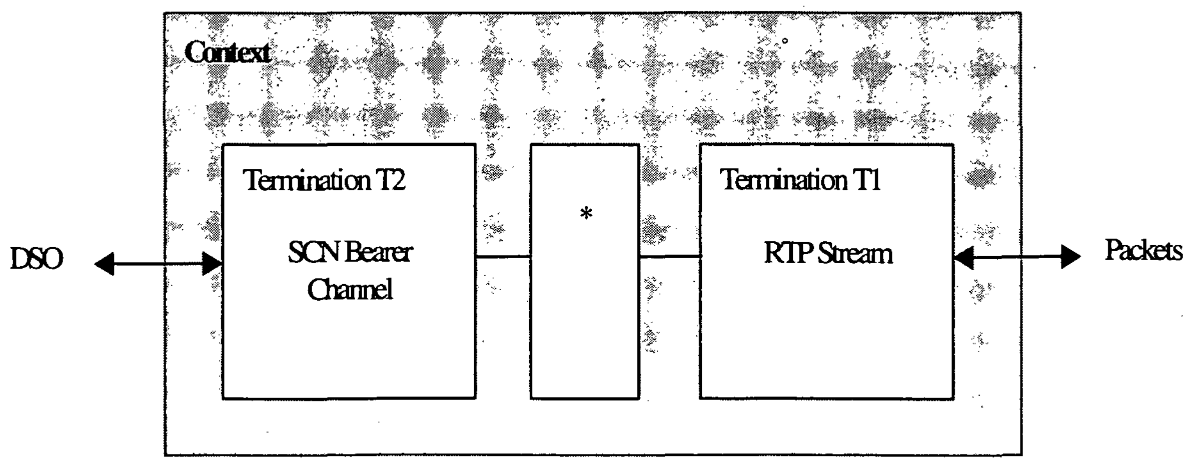


Figure 24: MEGACO Connection Model

The maximum number of Terminations in a Context is a MG property. MGs that offer only point-to-point connectivity might allow a maximum of two Terminations per Context. MGs that support multi-point conferences might allow three or more Terminations per Context. Context creation is dynamic in nature in the sense that a Context is created as soon as a Termination is added, whereas a Context gets deleted implicitly when the last Termination is subtracted from it.

D.3.1.2 Advanced Services

MGC contains all call control intelligence and manages all call scenarios, like from basic call till transfer, forward conference and hold. The Megaco/H.248 Protocol does not contain its own supplementary (advanced) services. Instead, these services are supported by H.323 or SIP. The Megaco/H.248 protocol only describes the communication between a MGC and a MG and the translation of media streams from one type of network to the format required in another.

D.3.2 Megaco Protocol Messages

Megaco Messages are split up in two groups – commands and descriptors. This sections gives an overview of this concept and introduces the most important commands, desciptors and error codes.

D.3.2.1 Megaco Commands

The protocol provides Commands for manipulating the logical entities – Contexts and Terminations – of the protocol connection model. For example, Commands exist to add Terminations to a Context, modify Terminations, subtract Terminations from a Context and audit properties of Contexts or Terminations. Commands provide complete control of the properties of Contexts and Terminations. This includes

- the specification of events which are reported to a Termination,
- the specification of signals/actions which have to be applied to a Termination
- and the specification of the topology of a Context (who hears/sees whom).

Most Commands are used by the MGC, which acts as command initiator, to control MGs, which act as command responders. However, there are some Commands for MGs as command initiators. The Commands are sent to the MG by the MGC except the case that the MG sends 'Notify' to the MGC. ServiceChange is a Command that can be sent by either entity to the other.

Below is an overview of the commands:

- Add: Adds a Termination to a Context.
- Subtract: Disconnects a Termination from its Context and returns statistics on the Termination's participation.
- Modify: Modifies the properties, events and signals of a Termination.
- Move: Automatically moves a Termination to another Context.
- Audit Value: Returns the current state of properties, events and signals of Terminations.
- AuditCapabilities: Returns all the possible values for Termination properties, events and signals allowed by the MG.

- Notify: Allows a MG to inform the MGC of the occurrence of events in the MG.
- ServiceChange: Allows the MG to notify the MGC that a Termination or group of Terminations is about to be taken out of service or has just been returned to service. The MG also uses this command to announce its availability to the MGC.

Commands between the Media Gateway Controller and the Media Gateway are grouped into Transactions, each of which is identified by a TransactionID. Transactions consist of one or more Actions. An Action consists of a series of Commands that are limited to operate within a single Context. Consequently, each Action typically specifies a ContextID.

D.3.2.2 Descriptors

The parameters to a command are called Descriptors. A descriptor consists of a name and a list of items. Some items may have values. Many Commands share common descriptors. Command parameters are structured into a number of descriptors. In general, the text format of descriptors is DescriptorName=<someID>{parm=value, parm=value....}. Parameters may be fully specified (only one value is given), overspecified (supported values are given) or underspecified (list of potential values is given).

D.3.2.3 Error Codes

Errors consist of an IANA registered error code and an explanatory string. Sending the explanatory string is optional. When a MG reports an error to a MGC, this is done using an error descriptor. An error descriptor consists of an error code and optionally the associated explanatory string. The identified error codes are:

400 – Bad Request
401 – Protocol Error
402 – Unauthorized
403 – Syntax Error in Transaction
406 – Version Not Supported
410 – Incorrect identifier
411 – The transaction refers to an unknown ContextId
412 – No ContextIDs available
421 – Unknown action or illegal combination of actions
422 – Syntax Error in Action
430 – Unknown TerminationID
431 – No TerminationID matched a wildcard
432 – Out of TerminationIDs or No TerminationID available
433 – TerminationID is already in a Context
440 – Unsupported or unknown Package
441 – Missing RemoteDescriptor
442 – Syntax Error in Command

443 – Unsupported or Unknown Command
444 – Unsupported or Unknown Descriptor
445 – Unsupported or Unknown Property
446 – Unsupported or Unknown Parameter
447 – Descriptor not legal in this command
448 – Descriptor appears twice in a command
450 – No such property in this package
451 – No such event in this package
452 – No such signal in this package
453 – No such statistic in this package
454 – No such parameter value in this package
455 – Parameter illegal in this Descriptor
456 – Parameter or Property appears twice in this Descriptor
471 – Implied Add for Multiplex failure
500 – Internal Gateway Error
501 – Not Implemented
502 – Not ready
503 – Service Unavailable
504 – Command Received from unauthorized entity
505 – Command Received before Restart Response
510 – Insufficient resources
512 – Media Gateway unequipped to detect requested Event
513 – Media Gateway unequipped to generate requested Signals
514 – Media Gateway cannot send the specified announcement
515 – Unsupported Media Type
517 – Unsupported or invalid mode
518 – Event buffer full
519 – Out of space to store digit map
520 – Media Gateway does not have a digit map
521 – Termination is "ServiceChanging"
526 – Insufficient bandwidth
529 – Internal hardware failure
530 – Temporary Network failure
531 – Permanent Network failure
581 – Does Not Exist

References

Patent:

[1] P9024: Verbindungssteuerung in einem Transit-Telekommunikationsnetz (Österreichische Patentanmeldung 2003)

Books:

[2] M. Hein, M. Reisner, A. Voß: Voice über IP - Sprach-Daten-Konvergenz richtig nutzen; Franzis Verlag 2002; ISBN: 3-7723-6686-4

[3] Black, Uyless: „Voice over IP" - 2000 by Perentice Hall PTR, Upper Saddle River, New Jersey - ISBN 0-13-022463-4

[4] Black, Uyless: „Internet-Technologien der Zukunft : [paketvermittelte Internetkommunikation ; Audio und Video im Internet]" / Uyless Black. - 1. [Dr.]. - München [u.a.] : Addison-Wesley, 1999. - ISBN 3-8273-1546-8

[5] Taylor, Ed: „TCP-IP ohne Geheimnis" / Ed Taylor [Dt. von: G & U, Flensburg]. - 1. [Dr.]. - Hannover : Heise, 1994. - ISBN 3-88229-042-0

[6] Comer, Douglas: „Computernetzwerke und Internets" / Douglas E. Comer. - München : Pearson Studium, 2000. - ISBN 3-8273-7012-4

[7] Helgert, Hermann J.: „Integrated services digital networks: architectures, protocols, standards" / Hermann J. Helgert. - Reading, Mass. [u.a.] : Addison-Wesley, 1991. - ISBN 0-201-52501-1

Article and Papers:

[8] QSIP = Extending SIP for QoS support / D. Papalilo, S. Salsano, L.Veltri, "Extending SIP for QoS support", Università di Roma "La Sapienza" (Italy)

[9] PSE-COSIM.01

[10] SIP-ROHC = Hans Hannu, „Signaling compression, Application signaling over cellular links" (based on <draft- hannu- rohc- signaling- cellular- 01. txt>)

[11] SIP-WIRE = Jouni Korhonen "SIP Signaling Compression for 3G Wireless Network"

[12] SIP-3G = Dean Willis, "SIP and 3G Wireless"

List of Internet Drafts

[13] SIP Extension for QoS support in Diffserv Networks

[14] DRAFT 1 = IETF, Internet Draft, draft-ietf-sip-manyfolks-resource-07.txt, G. Camarillo, W. Marshall, Jonathan Rosenberg "Integration of Resource Management and SIP

[15] DRAFT-2 = IETF, Internet Draft, draft-camarillo-sip-sigcomp-00.txt, G. Camarillo, "SigComp discovery for SIP"

[16] DRAFT-3 = IETF, Internet Draft, draft-ietf-rohc-sigcomp-algorithm-00.tx, Richard Price, Jonathan Rosenberg, Abigail Surtees, Mark A West, Lawrence Conroy, "Universal Decompression Algorithm"

[17] DRAFT-4 = IETF, Internet Draft, draft-ietf-rohc-signaling-req-assump-06.txt, Hans Hannu, "Signaling Compression Requirements & Assumptions"

[18] DRAFT-5 = IETF, Internet Draft, draft-ietf-sip-compression-01.txt, G. Camarillo, "Compressing the Session Initiation Protocol"

[19] DRAFT-6 = IETF, Internet Draft, draft-rosenberg-rohc-sip-udpcomp-00.txt, Rosenberg, "Compression of SIP"]

List of RFCs – IETF Standards:

[20] RFC 791: Internet Protocol (IP)

[21] RFC 768: User Datagram Protocol (UDP)

[22] RFC 793: Transmission Control Protocol (TCP)

[23] RFC 1288: finger

[24] RFC 1777: Lightweight Directory Access Protocol (LDAP)

[25] RFC 1945: Hyper Text Transfer Protocol (HTTP)

[26] RFC 1889: Real-time Transport Protocol (RTP)

[27] RFC 2167: rwhois

[28] RFC 2205: Resource Reservation Protocol (RSVP)

[29] RFC 2234: Augmented BNF for Syntax Specifications (ABNF)

[30] RFC 2326: Real-time Streaming Protocol (RTSP)

[31] RFC 2327: Session Description Protocol (SDP)

[32] RFC 2960: Stream Control Transmission Protocol (SCTP)

[33] RFC 2976: The SIP INFO Method

[34] RFC 2402: IP Authentication Header

[35] RFC 2705: MGCP

[36] RFC 2821: Simple Mail Transfer Protocol (SMTP)

[37] RFC 2824: Call Processing Language (CPL)

[38] RFC 2974: Session Announcement protocol (SAP)

[39] RFC 3015: MEGACO

[40] RFC 3261: Session Initiation Protocol (SIP)

[41] RFC 3265: Session Initiation Protocol (SIP)-Specific Event Notification

[42] RFC 3312: Integration of Resource Management and SIP

[43] RFC 3350: Real-time Transport Control Protocol (RTPC)

Standards:

[44] G.711 = CCITT Recommendation G.711, Pulse Code Modulation (PCM) of voice frequencies

[45] G.722 = CCITT Recommendation G.722, 7 kHz audio-coding within 64 kbit/s.

[46] G.723.1 = ITU-T Recommendation G.723.1, Speech coders: Dual rate speech coder for multimedia communications transmitting at 5.3 and 6.3 kbit/s.

[47] G.728 = CCITT Recommendation G.728, Coding of speech at 16 kbit/s using low-delay code excited linear prediction.

[48] G.729 = ITU-T Recommendation G.729, Coding of speech at 8 kbit/s using Conjugate Structure Algebraic-Code-Excited Linear-Prediction (CS-ACELP)

[49] H.225.0 = ITU-T Recommendation H.225.0, Call signaling protocols and media stream packetization for packet based multimedia communication systems

[50] H.245 = ITU-T Recommendation H.245, Control protocol for multimedia communication

[51] H.248 = ITU-T Recommendation H.248, Gateway control protocol

[52] H.261 = ITU-T Recommendation H.261, Video codec for audiovisual services at p × 64 kbit/s.

[53] H.263 = ITU-T Recommendation H.263, Video coding for low bit rate communication

[54] H.323 = ITU-T Recommendation H.323, Packet-Based Multimedia Communications Systems

[55] H.450.1 = ITU-T Recommendation H.450.1, Generic functional protocol for the support of supplementary services in H.323.

[56] H.450.2 = ITU-T Recommendation H.450.2, Call transfer supplementary service for H.323

[57] H.450.3 = ITU-T Recommendation H.450.3, Call diversion supplementary service for H.323

[58] H.450.4 = ITU-T Recommendation H.450.4, Call hold supplementary service for H.323

[59] H.450.5 = ITU-T Recommendation H.450.5, Call park and call pickup supplementary services for H.323

[60] H.450.6 = ITU-T Recommendation H.450.6, Call waiting supplementary service for H.323

[61] H.450.7 = ITU-T Recommendation H.450.7, Message waiting indication supplementary service for H.323

[62] H.450.8 = ITU-T Recommendation H.450.8, Name identification supplementary service for H.323

[63] H.450.9 = ITU-T Recommendation H.450.9, Call Completion Supplementary Services for H.323

[64] H.450.10 = ITU-T Recommendation H.450.10, Call offering supplementary services for H.323

[65] H.450.11 = ITU-T Recommendation H.450.11, Call intrusion supplementary services

[66] H.450.12 = ITU-T Recommendation H.450.12, Common Information Additional Network Feature for H.323

[67] T.120 = ITU-T Recommendation T.120, Data protocols for multimedia conferencing.

[68] Q.931 = ITU-T Recommendation Q.931, ISDN user-network interface layer 3 specification for basic call control

List of Acronyms

ABNF	Augmented BNF for Syntax Specifications
ACD	Automatic Call Distribution
AM	Authentication Management
ASN.1	Abstract Syntax Notation 1
BER	Bit Error Rate
BHCA	Busy Hour Call Attempts
B-ISDN	Broadband Integrated Services Digital Network
CAC	Connection Admission Control
CAR	Call Acceptance Rate
CBR	Call Blocking Rate
CCC	Customer Control Center
CCM	Connection Control Manager
CDV	Cell Delay Variation
CPL	Call Processing Language
CTD	Cell Transfer Delay
DBAM	Database Access Module
DBCS	Data Base Control System
DNS	Domain Name System
DSC	Dedicated Signaling Channel
DSP	Session Description Protocol
DSS1	Digital Subscriber Signaling 1
DVB	Digital Video Broadcasting
ESA	European Space Agency
ESW	EuroSkyWay
GEO	Geostationary Earth Orbit
GPS	Global Positioning System
GTW	Gateway Terminal
HM	Message Handler
HTTP	Hyper Text Transfer Protocol
IANA	Internet Assigned Authority
IANA	Internet Assigned Numbers Authority
IETF	Internet Engineering Task Force
IP	Internet Protocol
ISDN	Integrated Service Digital Network
ITU-T	Telecommunication Standardization Sector of the International

	Telecommunications Union
L2	Layer 2
L3	Layer 3
LAN	Local Area Network
LDAP	Lightweight Directory Access Protocol
LEO	Low Orbit Satellite
LM	Localization Management
LM	Localization Management
MBS	Max Burst Size
MC	Multipoint Controller
MCU	Multipoint Controller Unit
MDR	Mean Data Rate
Megaco	Media Gateway Control Protocol
MEO	Medium Orbit Satellite
MG	Media Gateway
MGC	Media Gateway Controller
MGCP	Media Gateway Control Protocol
MIME	Multi-Purpose Internet Mail Extensions
MM	Multimedia
MMUSIC	Multiparty Multimedia Session Control
MTU	Maximum Transmission Unit
NCC	Network Control Center
NMC	Network Management Center
NOC	Network Operation Center
OSI	Open System Interconnection
PC	Personal Computer
PDN	Packet Data Network
PDR	Peak Data Rate
PER	Packed Encoding Rules
PrT	Service Provider Terminal
PSTN	Public Switched Telephone Network
ptmp	point to multipoint
ptp	point to point
QoS	Quality of Services
RAS	Registration Admission and Status
RASC	Random Access Channel
RFC	Request for Comments
RSVP	Resource Reservation Protocol

RTCP	Real Time Control Protocol
RTP	Real Time Protocol
RTSP	Real Time Streaming Protocol
SAP	Session Announcement Protocol
SCH	Synchronization Channel Handling
SCN	Switched Circuit Network
SCTP	Stream Control Transmission Protocol
SG	Signaling Gateway
SH	Signaling Handler
SIP	Session Initiation Protocol
SMTP	Simple Mail Transfer Protocol
SOC	Satellite Operation Center
SS7	Signaling System 7
SSM	Secure Service Management
TCP	Transmission Control Protocol
TLS	Transport Layer Security
UAC	User Agent Client
UAS	User Agent Server
UDP	User Datagram Protocol
UF	Utilization Factor
URI	Uniform Resource Identifier
UT	Satellite User Terminal
VoIP	Voice over Internet Protocol
WAN	Wide Area Network
WWW	World Wide Web

List of Figures

List of Tables

Lebenslauf

Dipl.-Ing. Edith DUSCH, geb. Kemeny am 28.12.1974 in Vukovar, Kroatien
Ungarische Abstammung mit österreichischer Staatsbürgerschaft
verheiratet mit Dipl.-Ing. Alexander Dusch, keine Kinder
wohnhaft in 1100 Wien, in der Leibnizgasse 55/16

1981 – 1989	Allgemeinbildende Schule in Sombor (ehem. Jugoslawien) entspricht Volksschule und Unterstufenrealgymnasium
1989 – 1991	Naturwissenschaftliches Gymnasium in Sombor Zweig: Mathematik-Informatik Unterbrechung im dritten Jahr wegen Umsiedlung nach Wien
1991 – 1995	Oberstufenrealgymnasium BRG in 1220 Wien, Polgarstrasse Zweig: Informatik
1995 – 2001	Technische Universität Wien Studienrichtung Informatik 881
2002 – 2004	Technische Universität Wien Doktoratsstudium der technischen Wissenschaften An der Fakultät für Technische Naturwissenschaften und Informatik
seit 2001	Siemens AG Österreich Abteilung: PSE ECT PSS2