European Digital Cinema Audio Security

European Digital Cinema Audio Security

European Digital Cinema Audio Security

F. Frescura, L. Verducci, & P. Micanti,
Digilab2000 s.r.l.,
University of Perugia – DIEI,
and the EDCINE consortium

UCL PRESSES
UNIVERSITAIRES
DE LOUVAIN

Registration of copyright: D/2009/9964/20

ISBN: 978-2-87463-160-3

Printed in Belgium

Distribution: www.i6doc.com, on-line university publishers.

This book is available on order from bookshops or at:

CIACO University Distributors

Grand-Place, 7

1348 Louvain-la-Neuve, Belgium

Phone: 32 10 47 33 78

Fax: 32 10 45 73 50

e-mail: duc@ciaco.com

Overview

EDCine Project purposes

The Enhanced Digital Cinema (EDCine) project is focusing on the optimisation, enhancement and interoperability issues of JPEG 2000 based Digital Cinema standards. JPEG 2000 compression and the "DCI specifications" document [DCISPEC] are the cornerstones of the project's efforts.
The "DCI specifications" document published in April 2007 confirms the SMPTE decision to choose JPEG 2000 instead of MPEG for digital cinema coding. This choice clearly does not take account of many European issues. The MXF file format and security tools specifications are being adopted following the DCI studios' wishes. They lead to an unquestionable increase in the quality of the theatre experience. Nevertheless the specifications' rigidity, due mainly to the total absence of choice about the compression format, will force theatres willing to be able to project digital content to comply with those requirements within a few years to make great financial efforts.
If this aspect can be considered not to be an issue for large multiplex theatres, which are often directly tied to a major, the same cannot be asserted for the great majority of European cinemas, which consist mainly of a number of small theatres distributed in small or medium-sized towns and for which conforming to DCI specifications could be beyond their own investment capabilities.

EDCine will optimise and improve the Digital Cinema Initiative (DCI)-SMPTE specifications by including quality optimisation, robustness to transmission errors, content security tools,

stereoscopic imaging, interactive access, forensic marking and metadata for indexing and retrieval. The project will also supply interoperability with equipment using transcoding to/from MPEG and ensure end-to-end security from production to projection and to the other content distribution markets (e.g., television, mobile devices and th Internet). Networked access to cinema archives will allow users optimised access to video data for retrieval and downloading.

The EDCine project will demonstrate that convergence of the DCI model, the European needs and networked audio-visual systems while keeping the level of perceived quality as high as possible and ensuring security of content along the distribution chain is possible. The project's ultimate aim is actively to push achieved developments at the ITU, SMPTE, JPEG and MPEG standard bodies where the partners are already active, either directly or through the European Digital Cinema Forum (EDCF).

EDCine Project activities' organisation

In order to reach its objectives, EDCine project has been organised in several activity packages called "work –packages". The work packages themselves are organised by activity area. The following activity areas have been identified:

- Administrative management (WP9)
- Technical coordination (WP1, 7, 8 and 11)
- Technology exploration and development (WP2 to WP5)
- Application of technologies developed in the specific contexts (WP6).
- All the other WPs are linked to WP6, in which all the

technology modules developed are going to be integrated in the final systems and subsystems.

The project can be summarised by a matrix organisation, shown in the following figure, where the project's core is represented by the research and development work packages.

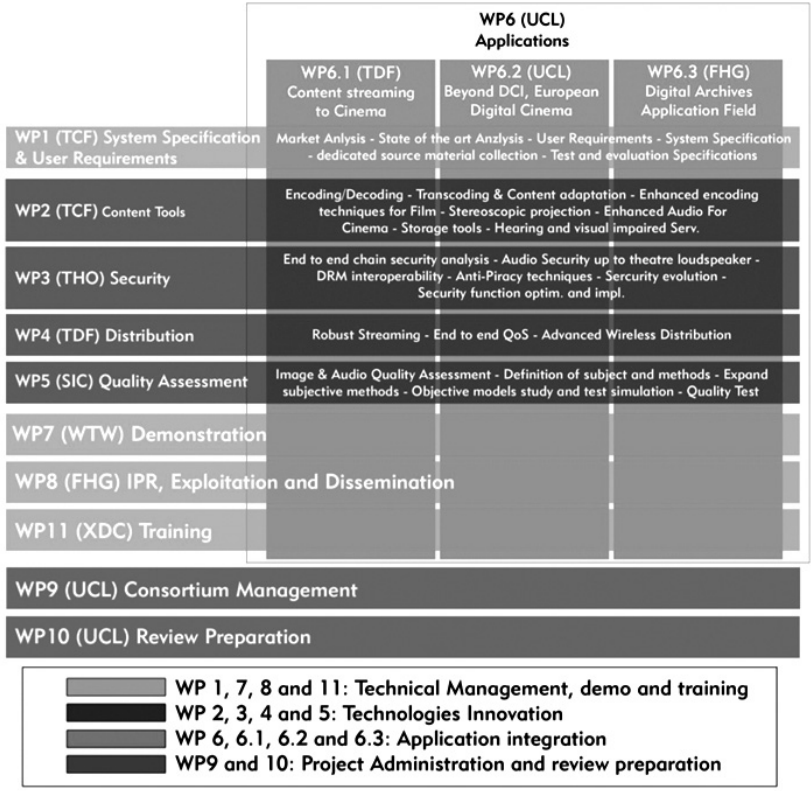

Figure 1. – Work package organisation of the EDCine project

EDCine WP3: Digital Cinema security

The objectives of Work Package 3 is to evolve from DCI system specifications version 1.0 to security function implementations. It will focus on several specific issues that are not necessarily tackled by DCI, such as audio security, DRM interoperability and anti-piracy techniques.

The Audio Distribution Equipment is designed within Task 3.2 of WP 3. The design effort is aimed at setting the specifications of and implementing a secure digital audio streaming solution for in-theatre playback of multichannel, high fidelity audio content over IP networks, providing streaming content copy protection by the use of strong encryption algorithms and authentication schemas. DSP based hardware will be developed and both wired and wireless architectures will be analysed.

Scope of the document

This document describes the fundamental system specifications and requirements for the Audio Distribution Equipment system and presents a possible system architecture implementing the fundamental requirements. The details are in the following sections:

System Requirements Specifications: Provides fundamental system specifications. The Interfacing section describes system interfaces available for communication with other digital cinema equipments and infrastructure. The Audio Section defines audio formats supported in terms of bit depth, sample rate and minimum channel count. The Synchronisation Section

is dedicated to requirements involving the delicate aspects of audio-video synchronisation and inter-channel synchronisation. The Security Section focuses on the content protection, reliability and security renewability requirements.

Proposed System Architecture: Describes a system architecture capable of implementing the fundamental specifications. The description mainly focuses on system security aspects. The Principal System Components Section identifies the main parts constituting the system. The interoperation scenarios section presents in details two possible operation modes for the system: The communication and security protocols used and state machines for the involved entities are described for both modes.

It should be noted that this is a first draft of the System Requirements Specifications document. Most of the requirements presented herein are therefore liable to undergo further development and modification. Again, a System Implementation Requirement Specifications Section is not present, as its function is fulfilled by the Proposed System Architecture Section, which is also intended to be a primer for the testing and development of such implementation specifications.

Document Language

This document consists of normative text and optional informative text. Normative text is text that describes the elements of the design that are indispensable or contains the conformance language keywords: “shall”, “should” or “may”. Informative text is text that is potentially helpful to the user but not indispensable and can be removed, changed or added editorially without affecting interoperability. Informative text does not contain any conformance keywords. All text in the document is, by default, normative except any section titled “Introduction”. Normative references are those external documents referenced in normative text and are indispensable to the user. Informative, or bibliographic, references are those references made from informative text or are otherwise not indispensable to the user.

The keywords “shall” and “shall not” indicate requirements that must be followed strictly in order to conform to the document and from which no deviation is permitted.

The keywords “should” and “should not” indicate that among several possibilities one is recommended as particularly suitable, without mentioning or excluding others; or that a certain course of action is preferred but not necessarily required. In the negative form, a certain possibility or course of action is deprecated but not prohibited.

The keywords “may” and “need not” indicate a course of action permissible within the limits of the document.

The keyword “reserved” indicates that a condition is not defined and shall have no meaning. However, it may be defined in the future. The keyword “forbidden” is the same as reserved, except that the condition shall never be defined in the future.

A compliant implementation is one that includes all mandatory provisions (“shall”) and, if implemented, all recommended provisions (“should”) as described. A compliant implementation need not implement optional provisions (“may”).

Requirements are indicated with the key phrases “is required to”, “is encouraged to” and “can”, which represent “shall,” “should” and “may”. This is necessary in order to distinguish requirements from the specification conformance language.

Sentences with the following keywords are in *italics*: shall, shall not, should not, is required, is not required, is not encouraged and is encouraged.

The names of standards publications and protocols are placed in [bracketed text]. International and industry standards contain provisions that, through a reference in this text, constitute provisions of this specification. At the time of publication, the editions indicated were valid. These referenced standards are subject to revision, and parties to agreements based upon this specification are encouraged to investigate the possibility of applying the most recent editions of the referenced standards.

System Objectives

Certain fundamental requirements have been acknowledged while writing the specifications for an Audio Distribution Equipment within the Enhanced Digital Cinema project. They are as follows:

The Audio Distribution Equipment shall have the capability to present a cinema audio experience that is better than or equal to what one could achieve nowadays with a traditional cinema audio presentation system.

The hardware and software used in the system should be easily upgraded as advances in technology are made and cinema

presentation standards changes over time.

The Audio Distribution Equipment shall provide a reliability and availability that are equal to or better than those of current film presentation systems.

Protection of intellectual property is a critical aspect of the design of the system. The security system shall provide means to keep the audio content encrypted while streaming to the loudspeakers so that only trusted receivers will be able to decode the content and deliver it to the audience.

The security system shall be renewable in the event of a breach in any part of the system.

References

[DCISPEC] Digital Cinema Initiatives, LLC, "Digital Cinema System Specification," Version 1.2, March 07, 2008.

[BALLOU-91] G. Ballou, Ed., "Handbook for Sound Engineers," 2nd ed., Howard Sams, 1991.

[EVEREST-01]F. A. Everest, "The Master Handbook of Acoustics," 4th ed., McGraw-Hill, 2001.

[EBU3285] EBU Technical Specification 3285, "BWF - a format for audio data files in broadcasting," Version 1, July 2001.

[SRTP-rfc3711] M. Baugher, E. Carrara, D. McGrew, M. Naslund, and K. Norrman, "The Secure Real-time Transport Protocol," IETF RFC 3711, March 2004.

[RTP-rfc3550] H. Schulzrinne, S. Casner, R. Frederick, V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications," IETF RFC 3550, July 2003.

[SIP-rfc3261] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley, E. Schooler, "SIP: Session Initiation Protocol," IETF RFC 3261, June 2002.

[SIP-Overview]Radvision Ltd., "SIP Technical Overview," 2005, [Online], Available: http://www.sipcenter.com/sip.nsf/html/WEBB5YP4SU/$FILE/RADVISION_SIP_Overview_04-05.pdf

[SIP-rfc3856] J. Rosenberg, "A Presence Event Package for the Session Initiation Protocol (SIP)," IETF RFC 3856, August 2004.

[SDP-rfc4566] M. Handley, V. Jacobson, C. Perkins, "SDP: Session Description Protocol," IETF RFC 4566, July 2006.

[SDP-rfc4567] J. Arkko, F. Lindholm, M. Naslund, K. Norrman, E. Carrara, ”Key Management Extensions for Session Description Protocol (SDP) and Real Time Streaming Protocol (RTSP),” IETF RFC 4567, July 2006.

[MIKEY-rfc3830] J. Arkko, E. Carrara, F. Lindholm, M. Naslund, K. Norrman, “MIKEY: Multimedia Internet KEYing,” IETF RFC 3830, August 2004.

[MIKEY-RSA-R] D. Ignjatic, L. Dondeti, F. Audet, P. Lin, “MIKEY-RSA-R: An Additional Mode of Key Distribution in Multimedia Internet KEYing (MIKEY),” IETF RFC 4738, November 2006.

[SAP-rfc2974] M. Handley, C. Perkins, E. Whelan, “Session Announcement Protocol,” IETF RFC 2974, October 2000.

[PTP-IEEE1588] International Standard IEC 61588, IEEE 1588, “Precision clock synchronization protocol for networked measurement and control systems,” IEC, September 2004

Chapter 1
System Requirements Specification

The Audio Distribution Equipment delivers the sound of the theatrical presentation to the audience. It is responsible for receiving the uncompressed and unencrypted digital audio from the Media Block and distributing it over a secure IP channel to the loudspeakers subsystem where it will be converted to analogue, amplified and translated to acoustic energy.

The following draft specifications cover only the essential Audio Distribution Equipment requirements, focusing on the emerging aspects of the system. In the following parts of the document we have specified the fundamental system requirements, followed by interfaces and audio characteristics specification. These requirements sections are very close to the corresponding sections in the DCI system requirements specification document for compatibility and interoperability reasons. The next sections, 2.5 and 2.6, address the system's synchronisation and security requirements. In these contexts these EDCine specifications goes beyond DCI, raising the bar for digital audio distribution in theatres.

Although a detailed implementation specification is left to further developments, a proposal for a possible implementation is presented in the next chapter.

Fundamental system requirements

This section describes the main Audio Distribution Equipment goals.

Interfacing

The Audio Distribution Equipment shall provide suitable interfaces to connect to other components of the distribution system for audio input and distribution and for system control purposes.

Audio

The audio experience provided to the audience by the system shall be better or comparable in quality to state-of-the-art cinema audio distribution systems. The audio format shall be seamlessly compatible with current DCI specifications.

Security

The system shall prevent major threats for the copyrighted distributed audio content and for the presentation session itself. The security shall be promptly renewable in the case of a breach.

Reliability and Availability

The system shall offer a reliability and an availability better than or comparable to state-of-the-art cinema audio distribution systems. The system shall provide means to easily and promptly recover from failures.

Interfaces

The Audio Distribution System input, distribution and control interface are required to be a Fast (100Base-T) or Gigabit (1000Base-T) Ethernet interface adopting an IP protocol stack.

Chapter 2
Audio Characteristics

The standardised audio characteristics meet the current DCI Digital Cinema System Specification for bit depth, sample rate and channel count to assure seamless compatibility of the ADE with the DCI distribution and presentation infrastructure.

Bit Depth

The bit depth shall be 24 bit per sample. Audio material having other bit depths shall be justified to the most significant bit.

Sample Rate

The audio sample rate shall be either forty-eight or ninety-six thousand samples per second per channel, commonly expressed as 48,000 or 96,000 kHz. This means that at 24 FPS playback there are exactly 2000 audio samples per frame per channel for 48,000 kHz and exactly 4000 audio samples per frame per channel for 96,000 kHz. At 48 FPS playback, there are exactly 1000 audio samples per frame per channel for 48,000 kHz and 2000 audio samples per frame per channel for 96,000 kHz.

Channel Count

The Audio Distribution System shall support up to sixteen full-bandwidth channels.

Streaming Format

The streaming audio format shall comply with the Broadcast Wave file format (.wav) audio content, per [EBU3285]. The audio content shall not be compressed.

Chapter 3
Synchronisation

The objective to keep the audio content stream encrypted until it is played back by the loudspeakers imposes strict requirements regarding audio-video and inter-channel synchronization. The main issue to be tackled is the fact that every loudspeaker is actually a completely independent computer device having enough processing power to handle session management and complex authentication and decryption operations. IP streaming environments do not guarantee fixed or uniform time delays, therefore a stream re-synchronisation before playback is needed. Not taking care of these aspects would lead to a very poor audio quality from the system.

Audio-video Synchronisation

General

The Audio-video synchronisation shall be frame based. The audio stream shall be organised in audio data fragments representing audio information for a single frame of video.

Programmable delay adjustment

In order to permit calibration of the sound sources in the auditorium and consequently to provide the best spatial sound experience to the audience, the system shall provide a way to manage the audio stream playback delay as a whole and per channel basis. This features shall also allow recovery of the correct audio-video synchronisation after the machine dependent processing delay introduced by any digital cinema projector

attached to the system.

Master delay adjustment

The system shall provide a way to adjust the playback delay of the audio stream as a whole. The adjustment interval shall be in the 0÷1.5 second range (at least). The zero delay requirement means that the audio contents will be played back at the minimum delay permitted by the system (this may need further specification). The delay measurement shall be taken as the difference between the actual playback instant and the minimum delay.

Channel delay adjustment

The system shall provide a way to adjust an audio channel playback delay independently of each other. The adjustment interval shall be in the 0÷500 milliseconds range (at least).The delay measurement shall be taken as the difference between the actual playback instant of the considered audio channel and the predicted playback instant of the audio stream as a whole.

Granularity

The system shall have an audio sample level delay adjustment granularity. This means having delay steps of about 0.01 milliseconds for a 96,000 kHz sample rate and of about 0.02 milliseconds for a 48,000 kHz sample rate.

Chapter 4
Inter-channel synchronisation requirements

The inter-channel synchronisation requirements are aimed at ensuring that the surround sound image produced by the loudspeakers array in the exhibition environment will have the same or better sound quality and fidelity than that offered by available theatre audio distribution systems based on analogical baseband signal distribution systems.

According to experiments ([BALLOU-91], [EVEREST-01]) time delays between audio sources of the order of 0.1 millisecond can affect sound imaging. This means that the maximum allowable misalignment between different audio channels in a surround sound system needs to be reasonably lower than this. Hence the sample-level synchronisation accuracy requirement.

Maximum inter-channel misalignment

The system shall present a maximum channel misalignment of the order of 0.01 millisecond, equivalent to the order of magnitude of single sample alignment at 96,000 kHz sample rate.

In any case the maximum channel misalignment shall not be higher than 0.1 millisecond.

Stream Metadata

Streamed audio content is required to contain metadata that indicate the first sample of audio data for each frame to be played

back and, depending on the implementation, for each channel. The embedded stream metadata shall also univocally associate each frame-based audio fragment with the unique time instant it is scheduled to be played back.

Chapter 5
Security

This section defines the requirements for the Audio Distribution Equipment security. Security is an end-to-end process, where the Audio Distribution Equipment represents one of the ends of such process and surely the one more exposed to potential attacks and threats. The following specifications are focused on extending the security process to include the final playback in the loudspeakers, with the purpose of offering a dramatic improvement in the protection of the copyrighted audio content exhibited in theatres.

Fundamental security system requirements

The goal of the Audio Distribution Equipment security system is to protect the copyrighted audio content in the exhibition environment up to its playback on authorised loudspeakers. Furthermore, the security system shall avoid unauthorised content sources to access the exhibition environment's loudspeaker system.

Content protection and piracy prevention

The security system shall provide a means for securing content against unauthorised access, copying, editing and playback. Protection shall be standardised primarily through the application of strong encryption technology, public key encryption and management of content key access.

Resist threats

The security system shall support prevention of the following threats:

Content theft (piracy) – as noted above

Unauthorised exhibition and theft of service

Manipulation of content

Denial of service

Reliability

The security system shall recognise that "the show must go on" except in extreme circumstances. Intelligent means to locate failures expediently shall be supported along with field replaceable – but not field serviceable – security devices.

Renewability

The security system shall provide a means for the interdiction or renewal of the system's device authorisation and authentication credentials.

Chapter 6
Proposed System Architecture

In the following paragraphs a possible system architecture, suitable to implement previously exposed fundamental system requirements, is proposed. The various system elements are identified first, and then two possible interoperation scenarios are presented. The purpose of the proposed architecture is to be a starting point for further developments in system requirements' specifications and related system implementation specifications. An early system implementation following the proposed architecture is now undergoing development and testing.

Principal system components

There are several elements of the ADE system. Each element plays a specific role in the system.

Audio Distribution Server

The Audio Distribution Server (ADS) is the source of the audio stream. It is responsible for the management of every aspect related to audio distribution. It can be thought of as embedded in the MediaBlock or as a separate entity, directly connected to it. Whichever the case, we will assume that the ADS is housed in a secure environment so that we can neglect security issues related to the MediaBlock-ADS Gigabit Ethernet link.

Active Loudspeakers

Active Loudspeakers (ALs) are responsible for final playback of the audio material, thus performing the decoding, buffering and

final scheduling for playback of audio frames received from the ADS. They also play an active role in session set-up, establishing communication and security parameters in conjunction with the ADS. The ALs consist of an active (amplified) loudspeaker and an embedded computer system required to perform authentication, communication and decoding tasks, enclosed within the loudspeaker casing.

Dedicated switched Fast/Gigabit Ethernet

The audio content stream flows on a dedicated switched 1000Base-T Ethernet equipped with IEEE1588-compliant switches.

Clock Synchronisation protocol

The IEEE1588 protocol (PTP – Precision Time Protocol) has been chosen to meet the synchronisation requirements between the ALs clocks and ADS clock: [PTP-IEEE1588] shows that sub-microsecond synchronisation accuracy can be achieved using PTP.

Authentication and streaming protocols

The authentication protocol is the element through which the ADS and ALs mutually verify each other's authorisation to transmit or receive audio content to or from each other. Once the legitimacy of the two devices is determined, the audio content is sent from the ADS to the ALs using a secure streaming protocol, which session encryption key is derived from shared secrets established during the authentication protocol phase. This prevents unauthorised devices both utilising the content and performing Theft of Service

attacks. Finally, in the event that legitimate devices are compromised to the point that they could allow unauthorised use of audio content, renewability allows ADE devices to identify such compromised devices and prevent the audio content's transmission.

Interoperation scenarios

The presented interoperation scenarios are focused on the security aspects of the system specifications. The descriptions revolve around the authentication and secret key distribution protocol used to establish a secure communication channel between the parties. Both of the presented scenarios adopt Public Key encryption (RSA) techniques for the authentication phase and symmetric key strong encryption algorithms (AES) for audio content stream encryption. The proposed implementations adopt protocols and techniques standardised in IETF RFCs to meet the system's ease of upgradeability requirements. The use of public key encryption requires setting up a Public Key Infrastructure (PKI), which is not covered in this document.

Scenario 1

In the first scenario the ADS uses a Session Announcement Protocol (SAP) in conjunction with a Session Description Protocol (SDP) to announce and distribute the session description information (e.g., media type and format, transport protocol, session timing, and bandwidth constraints). The proposed authentication and key management protocol is RSA mode MIKEY [MIKEY-rfc3830] tunnelled over SDP, as described in [SDP-rfc4567]. Once both parties are mutually authenticated and know the shared session key, the ADS can start the SRTP session.

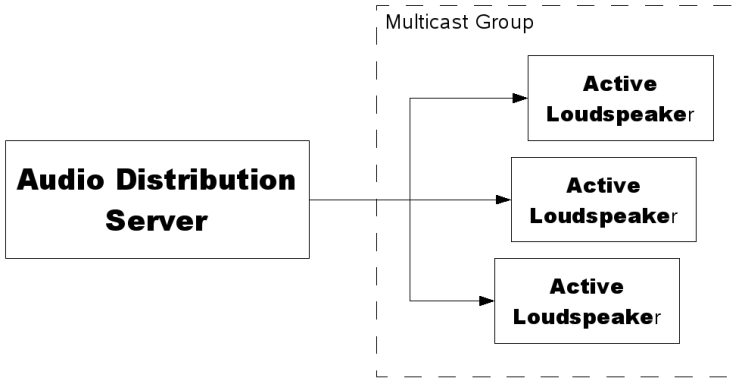

Figure 2. Scenario 1, block diagram

The session key is the same for every AL and the SRTP stream is sent over a multicast address. This approach makes session set-up and management very simple for both the ADS and ALs.

Phase 1 – authentication and session key distribution

Mutual authentication and secret session key material distribution is achieved by sending an SAP announcement containing session, authentication and keying information to the ALs multicast group.

Preconditions

Both the Audio Distribution Server and the Active Loudspeakers own each other's certificates signed by a trusted Certification Authority and a Certificate Revocation List. These certificates can be statically pre-configured or pre-exchanged between the parties by other means. The CRL shall be periodically updated. The required procedures are not covered in this scenario.

Active Loudspeaker Certificate check

The ADS checks for the validity of the ALs' certificates before session announcement, verifying for both the CA sign correctness and its absence from the Certificate Revocation List. If the certificate can be trusted the ADS begin the session announcement.

Session Announcement and key distribution

The ADS generate a SAP announcement carrying as a payload a SDP packet containing all the required information about the session, such as session ID, time and duration information, and bandwidth, along with adopted transport protocol (SRTP), ports and addresses.

Furthermore, the ADS generates a MIKEY message containing the ADS certificate and the MIKEY envelope key (from which the ALs can derive the effective session key (: see Appendix E for a

MIKEY overview)) encrypted with the AL public key and signed by the ADS. The MIKEY message is then tunnelled over the SDP packet, using an extended attribute field (cfr. [SDP-rfc4566], [SDP-rfc4567]), to the ALs.

Upon receiving the SAP announcement packet, the ALs use their private key to extract the envelope key from the MIKEY message and – after a successful ADS certificate verification – they check the integrity and authenticity of the message. If verification succeeds, session parameters and key material received are accepted. Finally the ALs prepare the SRTP session environment and start waiting for packets on specified ports.

Phase 2 – SRTP streaming

Once the parties are mutually authenticated and session keys have been distributed, the ADS can start SRTP (cfr. [RTP-rfc3550], [SRTP-rfc3711]) streaming using connection and stream format parameters previously communicated to the ALs.

Phase 3 – authentication and session key renewal

To limit the amount of cryptographic material that could eventually be collected by an eavesdropper monitoring the network connection between the ADS and ALs for the purpose of attempting a cryptanalytic attack, the authentication and session key distribution phase shall be periodically performed. The ADS will generate a new MIKEY envelope key and distribute it again to the ALs using an SAP session modification message.

Phase 4 – session expiration

Upon the session's expiration the ADS will be responsible for cancelling the session by sending a SAP session delete message, so

that ALs will be able to free the resources and become available for the next session.

In any event, a watchdog mechanism shall be implemented by the ALs to avoid the possibility of session lock-ins after a long time without receiving SRTP or SRTCP packets.

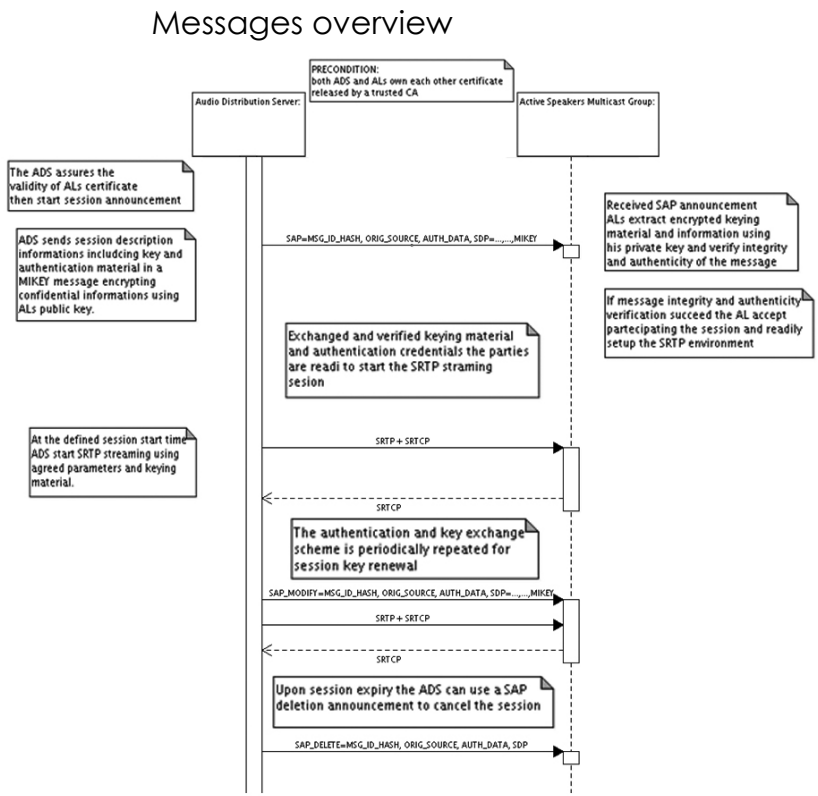

Figure 3. Scenario 1, authentication protocol messages overview

Audio Distribution Server state diagram

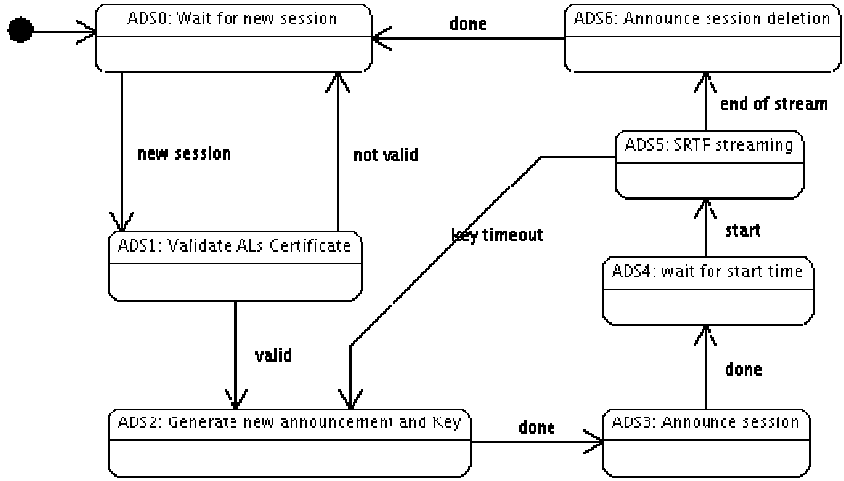

Figure 4. Scenario 1, Audio Distribution Server state diagram

State ADS0: Wait for new session. In this state the ADS is waiting for a new audio session to be scheduled.

Transition ADS0:ADS1. The scheduling of a new audio presentation session from the mediablock triggers this transition and initiates the session announcement and authentication procedure.

State ADS1: Validate ALs Certificate. In this state the ADS

performs the validation of ALs certificate requested for the next scheduled session. The process involves the verification of the certificate CA sign and validation against the CRL.

Transition ADS1:ADS0. If the ALs' certificate is found to be invalid, the ALs cannot be trusted and the session set-up is aborted. The ADS falls back to the ADS0 waiting state.

Transition ADS1:ADS2. If the ALs' certificate is valid, the ALs are trusted and can be authorised to receive the protected audio content.

State ADS2: Generate new announcement and key. In this state the ADS computes the generation of the SAP session announcement, including the session description information and MIKEY message with encrypted key and authentication credential, embedded in a SDP packet.

Transition ADS2:ADS3. The completion of the SAP announcement computation triggers this transition to the ADS3 state.

State ADS3: Announce session. In this state the ADS send the session announcement over the SAP predefined multicast network address and port. The announcement is repeated for a finite and defined amount of time.

Transition ADS3:ADS4. After the session announcement, the ADS shifts to a state where it waits for the session start time.

State ADS4: Wait for start time. In this state the ADS waits for the session start time previously announced in the session description information.

Transition ADS4:ADS5. When the session start time expires, the ADS is triggered in the ADS5 state.

State ADS5: SRTP streaming. In this state the ADS streams SRTP the audio content, encrypted with the shared session key previously distributed, to the ALs over the SRTP. The transmission goes on for a limited amount of time, then a session key renewal is required.

Transition ADS5:ADS2. When the current session key expires, the ADS switches back to the ADS2 state. A new session key is generated and distributed.

Transition ADS5:ADS6. When the presentation ends an “end of stream” event switches the ADS to the ADS6 state.

State ADS6: Announce session deletion. In this state the ADS sends a session deletion announcement to inform the ALs of the end of the session. This allows the ALs to prepare themselves for the next session.

Transition ADS6:ADS0. After performing the session deletion announcement, the ADS falls back to the initial ADS0 state, waiting for the next session to come.

Transition Any State:ADS0. In whichever state the ADS is, an unpredicted or unmanaged event or a user triggered event (e.g., abort) causes the ADS to abort current operations and to switch back to the ADS0 initial waiting state.

Active Loudspeaker state diagram

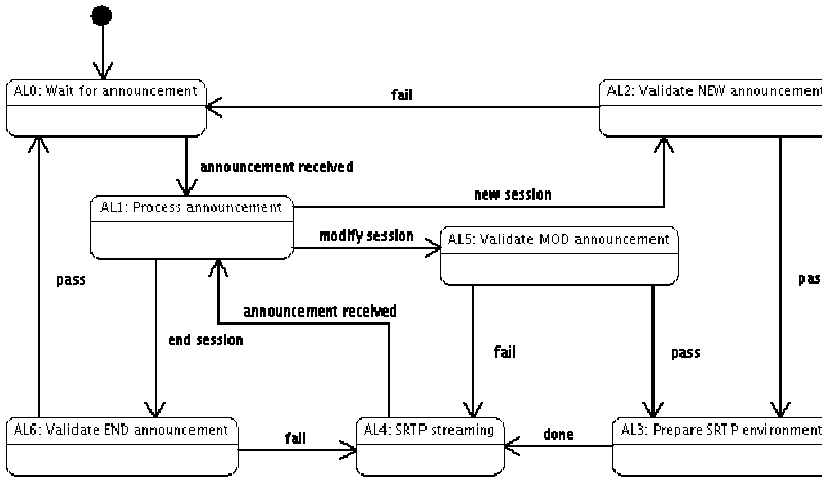

Figure 5. Scenario 1, Active Loudspeakers state diagram

State AL0: Wait for announcement. In this state the AL is waiting for a new session announcement from the ADS. This is the initial AL state.

Transition AL0:AL1 and AL4:AL1 After receiving a session announcement, the AL switches to the AL1 state to process it.

State AL1: Process announcement. In this state the AL processes the received SAP announcement, extracting both session information from the SDP packet and authentication and

keying material from the MIKEY message.

Transition AL1:AL2. If a new session announcement is received, the AL switches to state AL2.

State AL2: Validate NEW announcement. In this state the AL verifies the integrity and authenticity of the received announcement, checking the packet digital signature using the ADS certificate included in the MIKEY message. The ADS certificate validity is also checked against the trusted CA and the CRL.

Transition AL2:AL0. If the validation process fails, the sender of the received announcement is neither authenticated nor authorised to stream the content over the AL, so the procedure is aborted and the AL switches back to state AL0.

Transition AL2:AL3. If the validation process succeeds, the ADS is authenticated and the AL switches to state AL3.

State AL3: Prepare SRTP environment. In this state the AL uses the information previously extracted from the session announcement to set up or update the SRTP cryptographic context and the required communication sockets.

Transition AL3:AL4. Once the SRTP environment set-up is completed, the AL switches to the AL4 state.

State AL4: SRTP streaming. In this state the session set-up is done, and the AL can receive the audio stream over SRTP from the ADS. In any event the AL shall be listening on SAP dedicated ports to receive session modification and session deletion announcements. In the case of an unexpected session interruption (i.e., when the ADS ceases sending SRTP packets without any session deletion packet sent before), the AL shall fall back to the AL0 state in a limited amount of time.

Transition AL1:AL5. If a modify session announcement is received, the AL switches to state AL5.

State AL5: Validate MOD announcement. In this state the AL verifies the integrity and authenticity of the received announcement, checking the packet digital signature using the ADS certificate included in the MIKEY message. The ADS certificate validity is also checked against the trusted CA and the CRL.

Transition AL5:AL4. If the validation process fails, the sender of the received announcement is neither authenticated nor authorised to stream the content over the AL, so the procedure is aborted, the AL switches back to state AL4 and no session parameter is modified.

Transition AL5:AL3. If the validation process succeeds, the ADS is authenticated and the AL switches again to state AL3 to perform the update of session parameters and session key

renewal.

Transition AL1:AL6. If an end session announcement is received, the AL switches to state AL6.

State AL6: Validate END announcement. In this state the AL verifies the integrity and authenticity of the received announcement, checking the packet digital signature using the ADS certificate included in the MIKEY message. The ADS certificate validity is also checked against the trusted CA and the CRL.

Transition AL6:AL4. If the validation process fails, the sender of the received announcement is neither authenticated nor authorised to stream the content over the AL, so the procedure is aborted, the AL switches back to state AL4 and no session parameter is modified.

Transition AL6:AL0. If the validation process succeeds the ADS is authenticated and the AL closes the SRTP communication and discards the cryptographic context, falling back to the initial state AL0.

Transition Any State:AL0. In whichever state the AL is, an unpredicted or unmanaged event or a time-out event causes the AL to abort current operations and to switch back to the AL0 initial waiting state.

Scenario 2

In the second scenario the ADS and the ALs use Session Initiation Protocol [SIP-rfc3261], [SIP-rfc3856] in conjunction with the Session Description Protocol (SDP) to start the session and distribute session description information (e.g., media type and format, transport protocol, session timing, and bandwidth constraints). The proposed authentication and key management protocol for this second scenario is RSA reversed mode MIKEY [MIKEY-RSA-R] tunnelled over SDP [SDP-rfc4567]. Once both parties are mutually authenticated and the shared session key and connection properties are known, the ADS can start the SRTP session.

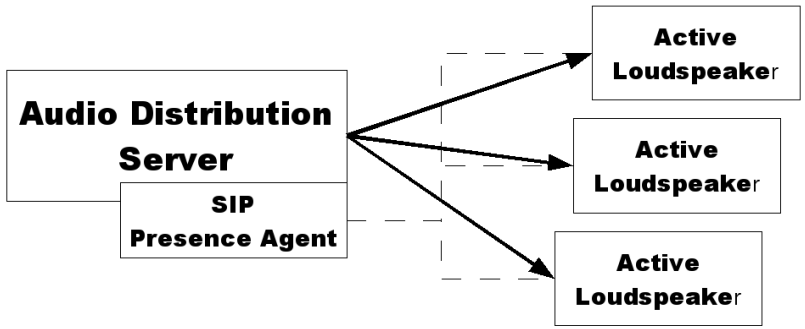

Figure 6. Scenario 2, block diagram

In this scenario the ADS directly implements an SIP Presence Agent [SIP-PRES-rfc3856] that notifies the ALs' presence to the ADS. This avoids the session announcement requirement and allows the ADS to establish an individual unicast SRTP stream with each AL, thus using a separate session key for each unicast link.

Phase 1 – presence notification

Upon start-up the ALs communicate their availability to receive incoming connections to the SIP presence agent implemented by the ADS. The presence signalling allows the ADS to know exactly which and how many ALs are available for the audio session presentation.

Phase 2 – authentication and session key exchange

In this phase, mutual authentication and secret session key material distribution are achieved using the SIP, SDP and MIKEY protocol stack as described in [MIKEY-RSA-R].

Preconditions

Both the ADS and the ALs are configured to trust a common CA, and they shall own or retrieve an updated CRL for the certificates emitted by this CA. The required procedures are not covered in this scenario.

SIP invitation from ADS

At the time of session start, the ADS sends an SIP INVITE message trying to establish a separate SRTP session with each AL that previously notified its presence. The SIP INVITE message contains the SDP packet carrying session characteristics and information and the MIKEY initiator message. The MIKEY message is signed with the ADS private key and carries the ADS's own certificate and other MIKEY-specific information. The whole SDP packet is also signed with the ADS's private key.

Invite authentication

Upon receiving the SIP INVITE message from the ADS, the AL shall verify both invite integrity and authenticity using the certificate carried in the MIKEY message. At first, the validity of the certificate itself is verified against the CA and the CRL. If the certificate is valid and trustworthy the ADS's public key is extracted from it and used to verify the signatures on the SDP and MIKEY messages. If the messages are found to be unmodified and authentic, the information embodied in the SDP packet payload shall be used to prepare the SRTP session environment, otherwise the AL will negate session establishment to the ADS.

Session key generation and invitation response

Assured of the invitation's trustworthiness, the AL will generate an envelope key (from which the real session encryption key will be derived, see [MIKEY-rfc3830] and Appendix E) to be incorporated into the MIKEY responder message, together with its own certificate. The MIKEY message containing the envelope key shall be sent to the ADS within the SIP acknowledge message that is to be sent to accept the session's establishment. Both the SDP packet and the MIKEY message carried in the SIP response are signed by the AL and the envelope key is encrypted using the ADS's public key.

Response authentication

After receiving the SIP acknowledge response message from the AL, the ADS shall authenticate it. It will extract from the MIKEY message the AL certificate and check its validity against the CA and the CRL. If the AL certificate is found to be trustworthy it shall be used to verify the integrity and authenticity of both the SDP and MIKEY messages. If the verification process succeeds, the ADS will extract the MIKEY envelope key from the MIKEY message, decrypting it with its private key, and derive from it the actual

session key. At this point, the session establishment is accepted. If any part of the response message is found to be corrupted or the AL certificate cannot be trusted, then the ADS will reject the received envelope key and the session establishment will fail.

Phase 3 – SRTP streaming

Once a session with one or more ALs has been established, the ADS will start the content streaming over the network using the agreed stream and connection parameters and session encryption key. It is worthwhile recalling that, in this scenario, the ADS will manage a separate SRTP stream for each AL with which it has established a session.

Phase 4 – authentication and session key renewal

To limit the amount of cryptographic material that could eventually be collected by an eavesdropper monitoring the network connection between the ADS and the ALs for the purpose of attempting a cryptanalytic attack, the authentication and session key exchange phase shall be periodically performed. The ADS will solicit a new authentication phase and the generation of a new envelope key from the ALs by resending the SIP invitation message.

Phase 5 – session expiration

Upon session expiration, the ADS shall cancel all the sessions previously established with the ALs, by sending a SIP BYE message, so that the ALs will be able to free the resources and become available for the next session.

In any event, a watchdog mechanism shall be implemented by the ALs to avoid the possibility of session lock-ins after a long time without receiving SRTP nor SRTP packets.

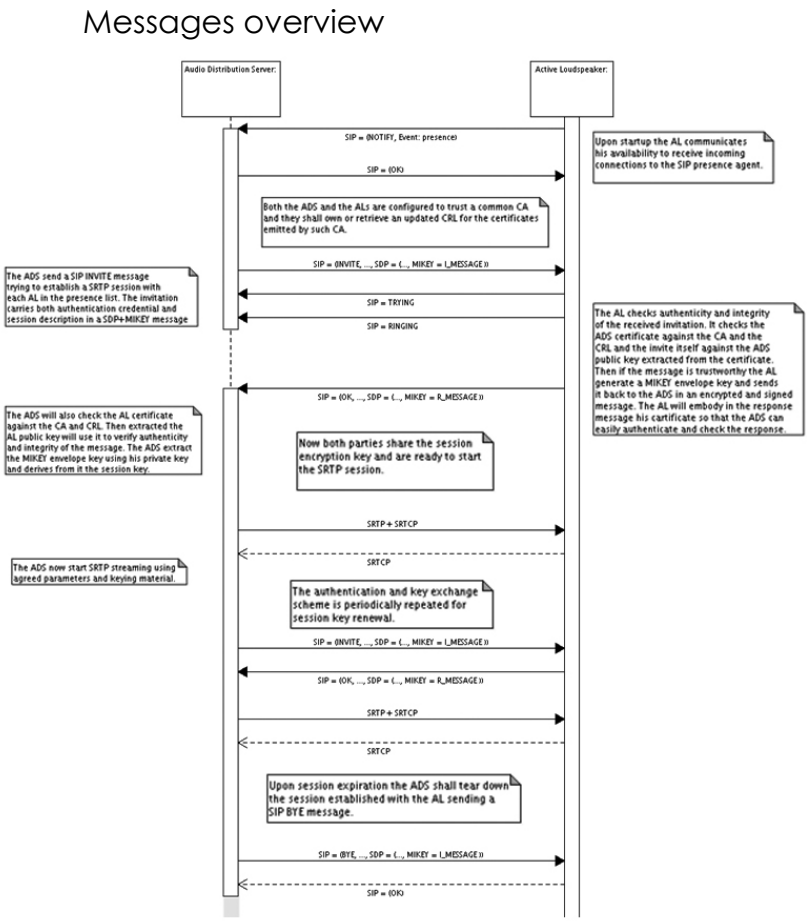

Figure 7. Scenario 2, authentication protocol messages overview

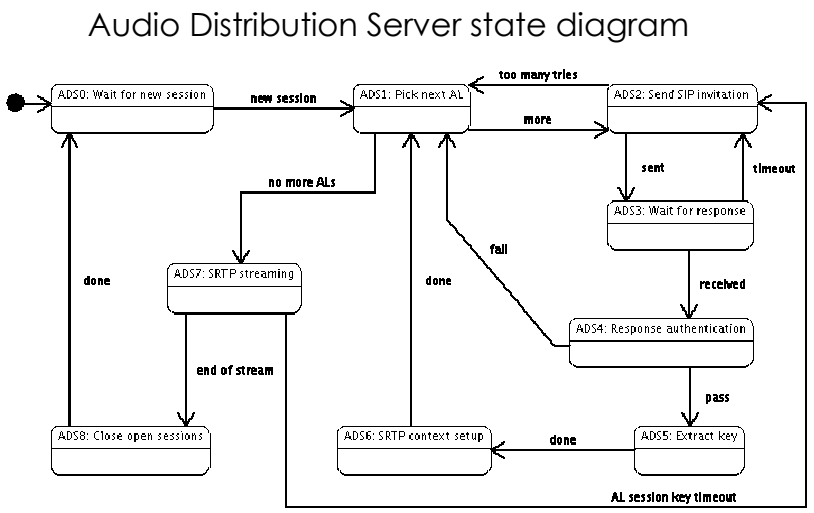

Figure 8. Scenario 2, Audio Distribution Server state diagram

State ADS0: Wait for new session. In this state the ADS is waiting for a new audio session to be scheduled.

Transition ADS0:ADS1. The scheduling of a new audio presentation session from the mediablock triggers this transition and initiates the SIP session initiation with the ALs.

State ADS1: Pick next AL. In this state the ADS picks from the Presence Agent database the next AL with which it will try to establish (or modify) an SRTP session.

Transition ADS1:ADS2. This transition is triggered if there are more ALs with which establishing a connection can be attempted.

State ADS2: Send SIP invitation. In this state the ADS generates the SIP invitation message containing the SDP session description information and the MIKEY message carrying the ADS certificate required for the invitation to be authenticated from the AL. Then the invitation is sent to the previously selected AL to initiate the session. If the AL does not respond promptly, the ADS can retry to send the invitation for a maximum number of times.

State ADS3: Wait for response. In this state the ADS is waiting for a response to the SIP invite message from the AL to which it sent the invitation.

Transition ADS2:ADS3 and ADS3:ADS2. Once the SIP invitation is sent, the ADS switches to the state ADS3, waiting for a response. If the response does not arrive in a maximum amount of time, a time-out event triggers the ADS to fall back to the state ADS2.

Transition ADS2:ADS1. If the maximum number of invitation messages are sent to a AL with no response from it, the ADS switches back to the state ADS1, where it can select the next AL to try to connect to.

Transition ADS3:ADS4. The AL response reception triggers the transition of the ADS in the ADS4 state.

State ADS4: Response authentication. In this state the ADS checks the integrity and authenticity of the AL response using the AL certificate carried in the MIKEY message. If the certificate is emitted by a trusted CA, meaning that the AL is authorised to receive and play back protected content, the ADS extracts the AL public key from the certificate and uses it to verify the signature of the response following the MIKEY scheme.

Transition ADS4:ADS1. If the AL certificate is not issued by a trusted CA or the response message is found to be corrupted or not authentic, then the authentication fails and the ADS aborts any further transaction with the current AL. This triggers the ADS's transition to the state ADS1, where the next AL in the presence database will be selected.

Transition ADS4:ADS5. If the AL certificate is emitted by a trusted CA and the response message is found to be unmodified and authentic, then authentication is successful and the session initiation can continue. The ADS will then switch to the ADS5 state.

State ADS5: Extract key. In this state the ADS extracts from the response message the MIKEY traffic generation key, from which the session key for the SRTP stream will be derived.

Transition ADS5:ADS6. After extracting the TGK, the ADS switches to the ADS6 state.

State ADS6: SRTP context set-up. In this state the ADS has acquired all the necessary elements and can prepare the SRTP session context.

Transition ADS6:ADS1. Once the SRTP environment set-up has been completed, the ADS switches back to the ADS1 state to select the next AL to start a session with.

Transition ADS1:ADS7. If there are no more ALs in the presence list to attempt to establish a session with, the transition to the ADS7 state is triggered.

State ADS7: SRTP streaming. In this state the ADS streams the audio content over the SRTP channels it has previously established with the registered ALs.

Transition ADS7:ADS2. On the periodic session key time-out for a certain SRTP session, the ADS falls back to the ADS2 state and renegotiates a new session key with the associated AL. To renew the session key, all the authentication process steps are performed again.

Transition ADS7:ADS8. When the end of presentation is reached, the ADS ceases to stream the audio content over the SRTP and switches to the ADS8 state.

State ADS8: Close open sessions. In this state the ADS closes all the sessions opened with the ALs, sending a SIP BYE message,

and frees all of the resources used by those sessions.

Transition ADS8:ADS0. After closing all the open sessions, the ADS is switched back to the ADS0 state, waiting for new scheduled sessions.

Transition Any State:ADS0. Whatever the state, an unpredicted or unmanaged event or a user-triggered event (e.g., abort) causes the ADS to abort current operations and switch back to the ADS0 initial waiting state.

Figure 9. Scenario 2, ADS Present Agent state diagram

State PA0: Wait for SIP SUBSCRIBE. In this state, the ADS-PA is waiting for subscriptions to come from the ALs. This is the initial state to which the ADS-PA switches at start-up.

Transition PA0:PA1. When a SIP SUBSCRIPTION message is received or a subscription time-out has expired, the ADS-PA switches to the PA1 state.

State PA1: Update presence database. In this state the ADS-PA performs the presence database update on the basis of the event that has just triggered the transition. If the transition was solicited by a subscription message, a new entry is created in the presence database or, if the associated subscriber is in the database already, his subscription is renewed. If otherwise the transition was solicited by a subscription timeout expiration or by a subscription expiration message (a subscription message requesting immediate expiration, see [SIP-PRES-rfc3856]), the associated subscriber is removed from the presence database.

Transition PA1:PA0. When the presence database has been updated, the ADS-PA switches back to the PA0 state.

Active Loudspeaker state diagram

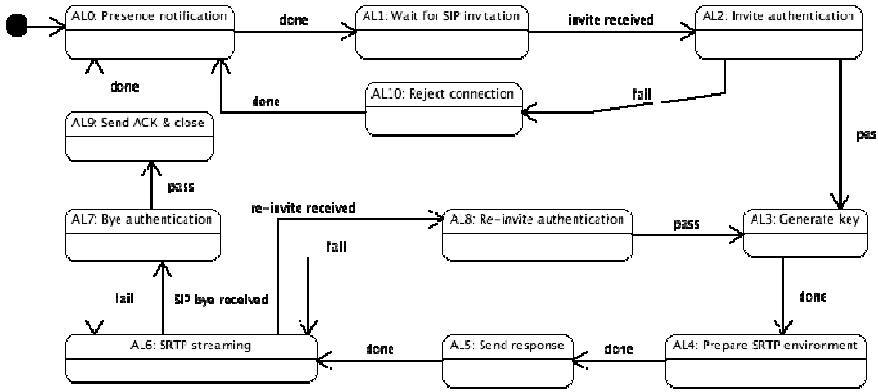

Figure 10. Scenario 2, ADS Present Agent state diagram

State AL0: Presence notification. In this state the AL communicates with the ADS Presence Agent to notify its presence and availability to participate in the audio presentation session.

Transition AL0:AL1. Once its presence has been notified to the ADS, the AL switches to the AL1 state.

State AL1: Wait for SIP invitation. In this state the AL is waiting for an SIP invitation message from the ADS.

Transition AL1:AL2. Upon the SIP invitation's arrival, the AL changes its state to AL2.

State AL2: Invite authentication. In this state, the AL checks the integrity and the authenticity of the invite received from the ADS, using the ADS certificate carried in the MIKEY message. If the certificate is emitted by a trusted CA, meaning that the ADS is authorised to establish a session and stream protected content, the AL extracts the ADS's public key from the certificate and uses it to verify the sign of the invite following the MIKEY schema.

Transition AL2:AL10. If the invite authentication fails, meaning that the received invite is corrupted or is not authentic, the AL switches to AL8 state.

State AL10: Reject connection. The AL rejects the session initiation attempt, sending back to the ADS an SIP unauthorised response message.

Transition AL2:AL3. If the invite is found to be both intact and authentic, the AL switches to the AL3 state.

State AL3: Generate key. In this state the AL generates the MIKEY envelope key to be transmitted to the ADS encrypted with its public key, previously extracted from the certificate and the MIKEY Traffic Generation Key (TGK) – see Appendix E for further details.

Transition AL3:AL4. Once the envelope key has been generated, the AL switches to the AL4 state.

State AL4: Prepare SRTP environment. The AL uses the MIKEY TGK to derive the SRTP session key and then sets up the SRTP session environment.

Transition AL4:AL5. Once the SRTP environment has been prepared, the AL switches to the AL6 state.

State AL6: SRTP streaming. In this state the AL is ready to receive and playback the audio content SRTP stream, by using the parameters previously agreed with the ADS.

Transition AL6:AL7. If during SRTP streaming an SIP bye message is received, this means that the ADS wants to close the session. The AL then switches in the AL7 state.

State AL7: Bye authentication. In this state, the AL checks the integrity and the authenticity of the bye message received from the ADS, using the certificate previously obtained for the current session.

Transition AL7:AL6. If the SIP bye message authentication fails this means that the message is corrupted or that the sender is not authorised to close the current session. In such cases, the AL switches back to the AL6 state, ignoring the session close request.

Transition AL7:AL9. If the SIP bye message authentication succeeds, the AL switches to the AL9 state.

State AL9: Send ACK & close. In this state the AL sends an SIP ACK message to the ADS, closes the session and frees all the used resources.

Transition AL9:AL0. After closing the session, the AL switches back in the AL0 state.

Transition AL6:AL8. If, during SRTP streaming, the AL receives a new SIP invite message, the ADS is requesting the session key's renewal. The AL switches to the AL8 state.

State AL8: Re-invite authentication. In this state the AL checks the integrity and the authenticity of the invite received from the ADS, using the ADS certificate carried in the MIKEY message. If the certificate is emitted by a trusted CA, meaning that the ADS is authorised to establish a session and stream protected content, the AL extracts the ADS's public key from the certificate and uses it to verify the signature of the invite following the MIKEY schema.

Transition AL8:AL6. If the SIP invite message authentication fails, meaning that the message is corrupted or that the sender is not authorised to update the current session parameters, the AL switches back to the AL6 state, ignoring the re-invite request.

Transition AL8:AL3. If the SIP invite message authentication succeeds, the AL switches to the AL3 state to renew the session key.

Transition Any State:AL0. Whatever the state, an unpredicted or unmanaged event or a time-out event causes the AL to abort current operations and to switch back to the AL0 initial state.

Considerations

In the first scenario presented, the ALs must share the same certificate and therefore the same public/private key pair. Furthermore, both the ALs and ADS must have both their and the other party's certificates before any communication can start. This requires a pre-configuration process that can be performed manually or programmatically, by means not covered in this overview.

In any event, the session and security set-up prove to be extremely simple using the first approach. In fact, they are accomplished in half –a –round trip, for all the required information for session and security set-up are carried in the session announcement.

In the second scenario, every AL receives its own certificate and the associated public/private key pair. The ADS and the ALs mutually exchange their certificates during the session set-up process, so that there is no need for certificate pre-configuration, although setting up a Public Key Infrastructure, with a common trusted CA, is still needed.

The establishment of multiple unicast SRTP sessions, following the second approach, complicates both the session set-up phase and the streaming phase, requiring the ADS to manage all the sessions independently. This could require a much larger amount of processing power for the ADS, and its feasibility has to be assured.

The downside of its simplicity is that the first approach may be more prone to cryptanalytic attacks from eavesdroppers than the second. This is due to the fact that the ADS streams the content to the ALs over a multicast channel, thus using a unique session encryption key for all of the audio channels. This means that an eventual eavesdropper can collect much more cryptographic material to attempt a cryptanalytic attack, in comparison with the second

approach, where each audio channel is streamed over a separate session, thus having its own session key. In addition, if a session key is compromised, the attacker can gain access to the whole protected content – i.e., all of the channels – (streamed with that key). It should be noted, however, that this risk could easily be lowered by opportunely reducing the session key expiration time-out, leading to a more frequent session key renewal.

In the event that one of the private keys (of the ADS or one of the ALs) is compromised, an eavesdropper could gain access to all of the protected material (ADS key compromised or ALs key compromised – in the first scenario), or to a substantial part of it. To regain system security, the certificate related to the compromised key must be added to the CRL and the owner of the compromised key must be provided with a new key pair and certificate issued by the trusted CA.

Appendix A
SRTP overview

Secure Real-time Transport Protocol: main features

SRTP is defined in [SRTP-rfc3711] as an RTP profile designed to grant confidentiality, authenticity and protection from replay attacks to RTP and RTCP traffic. Confidentiality is achieved by adopting the Advanced Encryption Standard's (AES) strong encryption algorithm to encode the RTP payload while at the same time the HMAC/SHA1 hash function is computed over both the header and the payload of the packet, accounting for sender authentication and replay attack protection.

Another emerging SRTP characteristic is found in the use of one Master Key and multiple Salting Keys:

The Master Key is the secret key from which the session encryption keys are derived by applying a key derivation function.

The Salting Keys are derived from the Master Key, too, with a cryptanalytic safe algorithm. They are used in the encryption scheme to lower the threats posed by some cryptanalytic attacks.

SRTP packet format and fields description

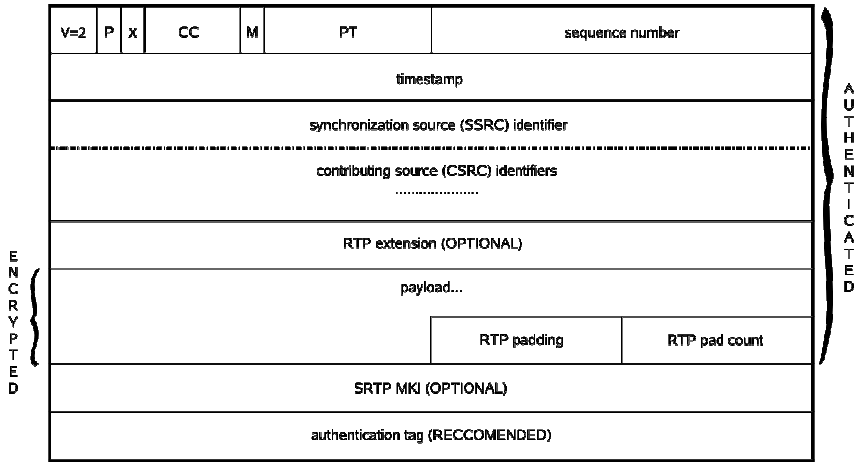

Figure 11. SRTP packet format

The SRTP packet is constituted by RTP packet fields, with two additional fields:

marker (M – 1 bit): Its interpretation is profile-dependent. It can be used, for example, to mark frames' limits.

payload type (PT – 7 bits): It identifies the payload format type and helps payload interpretation in the application layer.

sequence number (16 bits): It increases with every packet emitted from the source. The receiver uses this field to detect lost packets and to reorder those that are received out of sequence.

timestamp (32 bits): This represents the sampling instant of the first octet of the RTP packet. It allows synchronisation and jitter measurement in packet arrivals.

synchronization source ID (SSRC 32 bits): It identifies the synchronisation source (the main stream source) and is randomly chosen.

The following are the only SRTP fields not yet defined in RTP:

master key identifier (MKI): This is an optional field and has a variable length. It is used for key management and session key renewal purposes. It identifies a master key in SRTP cryptographic context.

authentication tag: This field has a variable length and its use is recommended in the standard definition. It provides header and payload authentication allowing for sequence number authentication and thus reply attack protection.

For further details on the SRTP please refer to [SRTP-rfc3711]

Appendix B
SAP overview

Session Announcement Protocol, main features

The Session Announcement Protocol is designed to assist in multicast multimedia session publication and to communicate relevant information about session set-up to potential session participants.

In the SAP an entity-defined announcer periodically sends an announcement packet to a multicast address and a known port. In the SAP there is no handshake mechanism and the announcer is unaware of the presence of any listener.

The session announcement packet may contain an authentication header. The session description data can also be encrypted.

The SAP packet contains the following fields:

A: address type – It is 0 if the IP address in the originating source field is an IPv4 address or 1 if it is an IPv6 address.

T: message type – It is 0 for a session announcement and 1 for a session deletion.

E: cryptography – It is set to 1 if the payload is encrypted

C: compression – It is set to 1 if the payload is compressed

Authentication length: This is the length of the authentication field

Message identifier hash: Along with the announcer address it represents an unequivocal session announcement identifier.

Originating source: This is the source IP address.

Payload type: This is a MIME identifier. It describes the payload format (for example SDP).

For further details on the SAP please refer to [SAP-rfc2974].

SAP packet format and fields description

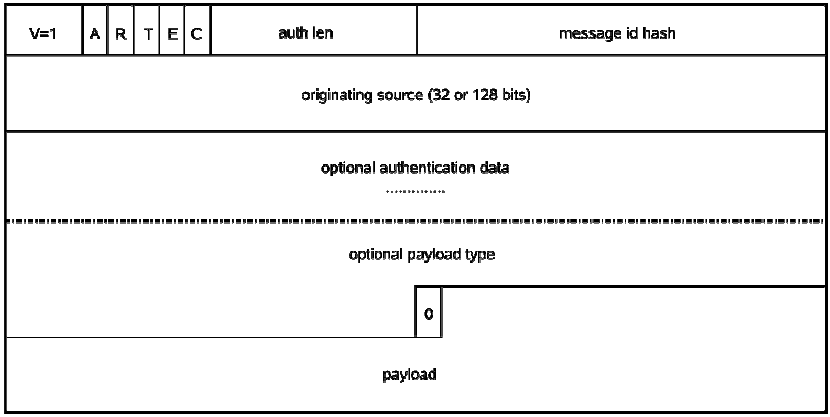

Figure 12. SAP packet format

Appendix C
SIP overview

Session Initiation Protocol, main features

The Session Initiation Protocol is a control and signalling application layer protocol. Its purpose is the instauration, modification and termination of multimedia communication session on packet switched networks. SIP invites used for session instauration carry session information allowing the participants to establish an agreement on compatible transmission fomats.

SIP makes use of proxy servers and other entities in order to facilitate the routing of requests. Nevertheless, its simpler implementation allows direct interaction of the session endpoints incorporated in the User Agents.

A User Agent is defined as an application capable of acting as both a User Agent Client and a User Agent Server:

The User Agent Client (UAC) is a client application that advances SIP requests.

The User Agent Server (UAS) is a server application that handles SIP requests and gives responses to the clients.

SIP messages are composed of three main sections: a start line, a message header and the message body. The start line defines the message type (method type in requests and status code in the answers) and the protocol version, while the message header carries message attributes describing additional information in the format <name>:<value>. The message body contains the information about the session to start (for example codec information). A possible body

format could be SDP.

SIP User Agents communication example

In the following figure an example of session establishment and cancellation between two User Agents is depicted.

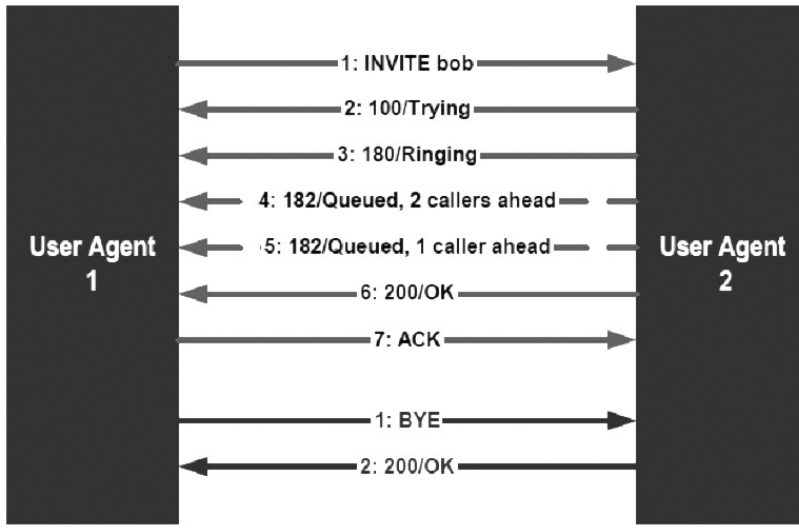

Figure 13. SIP communication example between User Agents (from [SIP-Overview])

For further details on the SIP please refer to [SIP-rfc3261], [SIP-rfc3856].

Appendix D
SDP overview

Session Description Protocol, main features

SDP is a protocol aimed at multimedia session description for the purpose of session announcement (SAP) or session invitation (SIP), or any other form of invitation to multimedia sessions participation. In this context the standard defines a multimedia session as an aggregate of communication streams existing in the network for a determined amount of time.

Typically, SDP packets carry the following information:

Session information:

- Session name and scope

- Session duration

Information on available/required bandwidth

Streams information:

- Kind of stream: audio, video, etc.

Transport protocol

Payload format: MPEG, PCM...

Multicast address and port number to join the multicast group participating in the session

SDP protocol structure

SDP packets are constituted by plain text messages describing the session's characteristics. In the following list the fields marked with an asterisk are optional:

Session Description:

v = protocol version

o = owner/creator and session identifier

s = session name

i =* session information

u =* session's description URI

e =* e-mail address

p =* telephone number

c =* connection information

b =* bandwidth information

Time description:

t = time the session is active

r =* zero or more repeat times

Media description:

m = media name and transport address

i =* media title

c =* connection information

b =* bandwidth information

k =* encryption key

a =* media attributes

For further details of the SDP please refer to [SDP-rfc4567].

Appendix E
MIKEY overview

Multimedia Internet KEYing, main features

MIKEY is a protocol designed for secret key and security parameter exchange and distribution in real time communication applications. Its services are requested by other security oriented protocols such as the SRTP. SRTP support in MIKEY is described in detail in [MIKEY-rfc3830].

This protocol can be employed in several scenarios:

peer to peer: unicast, for example a SIP session between two participants who have to agree on session security setup.

one to many – simple: multicast, real time presentations. The sender is in charge of set-up security.

many to many: interactive groups, without a centralised control unit. Each transmitting party may set up the security for its own outgoing media.

MIKEY design goals can be summarised as follows:

To offer end-to-end security

Simplicity

Efficiency: low bandwidth requirements, low computational cost, few lines of code and minimum number of round trips.

Tunnelling: Possibility of "tunnelling" or integrating MIKEY in session establishment protocols, such as the SDP.

Independence from any specific security functionality of the underlying transport protocol.

MIKEY operation mode overview

As stated above, MIKEY's aim is to generate and distribute the information required to set up one or more secure communication sessions (Crypto Sessions – CS or Crypto Sessions Bundle – CSB) with the chosen security protocol. More in detail, MIKEY defines and distributes an aggregate of security parameters called Data Security Association (DSA) and a group of Traffic encryption keys Generation Keys (TGK). This is achieved thanks to one of three available key transport/exchange strategies.

The TGK keys will be then used by the participants to derive, in a cryptanalytically secure way, a Traffic Encryption Key (TEK) for every session of the CSB. The TEKs together with the security protocol parameters constitute the Data SA. These are then used as input for the security protocol.

Now the security protocol can use the resulting TEK directly or further derive its own session keys from it (as is done in the SRTP).

The exchange distribution scheme can be reiterated to refresh the TGKs and TEKs while the session is going.

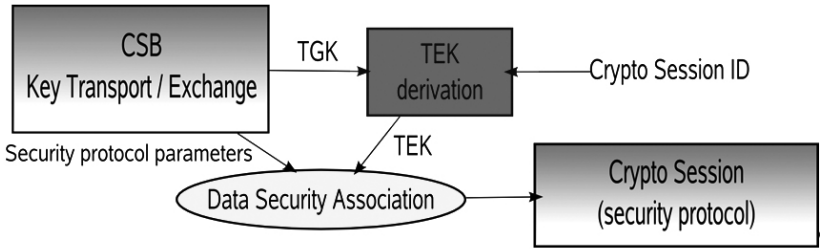

Figure 14. MIKEY operation overview

MIKEY keys exchange/distribution

MIKEY supports three different schemas for TGK exchange, namely,

use of a pre-shared key

public-key encryption

Diffie-Hellman (DH) key exchange

Furthermore, in [MIKEY-RSA-R] is defined a variant of the public-key method that is the fittest for the purposes of the system described herein, guaranteeing secure session key exchange and allowing client authentication. A simple description of the MIKEY RSA Reverse mode key exchange method described in [MIKEY-RSA-R] follows.

MIKEY RSA Reverse mode key exchange method

The method depicted in the following figure is started by the Initiator.

The Initiator sends a message to the Responder signed with his own private key. The message asks the Responder to send back a TGK.

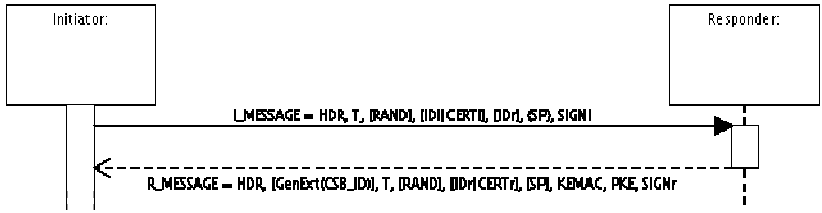

Figure 15. MIKEY RSA Reverse mode key exchange

Table of abbreviations and acronyms

ADE
Audio Distribution Equipment: the apparatus specified in this System Requirements Specification document.

ADS
Audio Distribution Server: the device responsible of distributing encrypted audio streams to Active Loudspeakers.

AES
Advanced Encryption Standard: a symmetrical key block cipher adopted as DES (Data Encryption Standard) substitute by the U.S. government. It has been extensively analysed and is now used worldwide.

AL
Active Loudspeaker: the device receiving and decoding the encrypted audio stream sent from the Audio Distribution Server.

CA
Certification Authority: the entity that issues digital certificates for use by other parties. It is an example of a trusted third party. CAs are characteristic of many Public Key Infrastructure schemes.

CRL
Certificate Revocation List: a list of certificates (or more accurately: their serial numbers) that have been revoked, are no longer valid, and should not be relied on by any system user.

DCI
Digital Cinema Initiatives: the consortium of studios formed to establish a standard architecture for digital cinema systems. Its work

produced the overall system requirements and specifications for digital cinema, now released under version 1.1.

DRM
Digital Rights Management: term that refers to access control technologies used by content publishers and copyright holders to limit usage, copy and distribution of digital media content.

DSP
Digital Signal Processor: a specialised microprocessor designed specifically for digital signal processing tasks, generally to be carried on in real time.

EBU
European Broadcasting Union.

EDCine
Enhanced Digital Cinema: The EDCine project is focusing on the optimisation, enhancement and interoperability issues of JPEG 2000-based Digital Cinema. The project is organised in several work packages (WPs).

EDCF
European Digital Cinema Forum: EDCF functions as a network for European cooperation on e- and d-cinema activities. It identifies key issues, gathers information and creates models to encourage private investments and public support schemes. It liaises with other relevant bodies to assist in the establishment of appropriate world-wide standards for e- and d-cinema.

FPS
Frames Per Second: the measurement of the frequency (rate) at which an imaging device produces unique consecutive images called frames.

HMAC

Key-Hashed Message Authentication Code: a type of message authentication code (MAC) calculated using a specific algorithm involving a cryptographic hash function in combination with a secret key. It may be used to verify simultaneously both the data integrity and the authenticity of a message.

IEC
International Electrotechnical Commission: a non-profit non-governmental international standards organisation that prepares and publishes international standards for all electrical, electronic and related technologies.

IEEE
Institute of Electrical and Electronic Engineers: an international non-profit professional organization for the advancement of theory and practice of electrical, electronics, communications and computer engineering, as well as computer science. IEEE serves as a major publisher of scientific journals and a conference organiser.

IETF
Internet Engineering Task Force: develops and promotes Internet standards. It is an open standards organisation with no formal membership or membership requirements.

IP
Internet Protocol: a network layer protocol. It is a data oriented protocol used for communication across a packet-switched internetwork. IP provides the service of communicable global addressing amongst computers.

ISO
International Organization for Standardization: an international standard-setting body composed of representatives from various national standards organisations. ISO promulgates worldwide industrial and commercial standards.

ITU
International Telecommunication Union: an international organisation established to standardise and regulate international radio and telecommunications. Its main tasks include standardisation, allocation of the radio spectrum, and organising interconnection arrangements between different countries to allow international phone calls.

JPEG
Joint Picture Experts Group: a joint committee between ISO and ITU-T. It created the JPEG and JPEG 2000 image codec standards.

MIKEY
Multimedia Internet KEYing: a key management protocol that is intended fo use with real-time applications. It can be specifically used to set up encryption keys for multimedia sessions that are secured using the SRTP.

MIME
Multipurpose Internet Mail Extension: an Internet standard that extends the format of e-mail to support text in character sets other than US-ASCII, non text attachments, multi-part message bodies and header information in non-ASCII character sets. However its use has grown beyond this to describing content type in general.

MPEG
Moving Picture Experts Group: an ISO/IEC working group charged with the development of video and audio encoding standards.

MXF
Material eXchange Format: a container format for professional digital and audio media defined by a set of SMPTE standards.

PCM
Pulse Code Modulation: a digital representation of an analogue signal where the magnitude of the signal is sampled regularly at uniform intervals, then quantized to a series of symbols in a digital (usually binary) code.

PKI
Public Key Infrastructure: an arrangement that binds public keys to respective user identities by means of a certification authority (CA). The user identity must be unique for each CA.

PTP
Precision Time Protocol: a time-transfer protocol defined in the IEEE 1588-2002 standard that allows precise synchronisation of networks (e.g., Ethernet). Accuracy within the nanosecond range can

be achieved with this protocol when using hardware-generated timestamps.

RFC
Request For Comments: a series of memoranda encompassing new research, innovations and methodologies applicable to Internet technologies. The Internet Engineering Task Force adopts some of the proposals published in RFCs as Internet standards.

RSA
RSA is an algorithm for public-key cryptography. It was the first algorithm known to be suitable for signing as well as encryption. It is widely used in electronic commerce protocols and is believed to be secure given sufficiently long keys. Its name derives from that of its inventors: Ron Rivest, Adi Shamir and Leonard Adleman.

RTP and **RTCP**
Real-time Transport Protocol defines a standardized packet format for delivering audio and video content over the Internet. Real-time Transport Control Protocol is its companion control protocol.

SAP
Session Announcement Protocol is a protocol for broadcasting multicast multimedia session information. It typically uses Session Description Protocol (SDP) as the format of session descriptions, and the multicast session typically uses the Real-time Transport Protocol (RTP).

SDP
Session Description Protocol is a format for describing streaming media initialisation parameters. It is intended for describing multimedia sessions for the purposes of session announcement, session invitation and other forms of multimedia session initiation.

SHA

Secure Hash Algorithm. The acronym identifies a family of cryptographic hash functions designed by the National Security Agency. Hash algorithms compute a fixed-length digital representation (known as a message digest) of an input data sequence (the message) of any length. They are called “secure” when:

- It is computationally unfeasible to find a message that corresponds to a given message digest.
- It is computationally unfeasible to find two different messages that produce the same message digest.
- Any change to a message (including single bit changes) will, with an exceeding high probability, result in a completely different message digest.

SIP
Session Initiation Protocol is a signalling protocol, widely used for setting up and tearing down multimedia communication sessions such as voice and video calls over the Internet. Other feasible application examples include video conferencing, streaming multimedia distribution, instant messaging and presence information.

SMPTE
The Society of Motion Pictures and Television Engineers is an international professional association of engineers working in the motion imaging industries. As internationally-recognised standards developing organization, SMPTE has over 400 standards, recommended practices and engineering guidelines for television, motion pictures, digital cinema, audio and medical imaging.

SRTP and **SRTCP**
The Secure Real-time Transport Protocol defines a profile of RTP intended to provide encryption, message authentication and integrity, and replay protection to the RTP data in both unicast and multicast applications. SRTCP is the secure companion control protocol.

Table of Contents