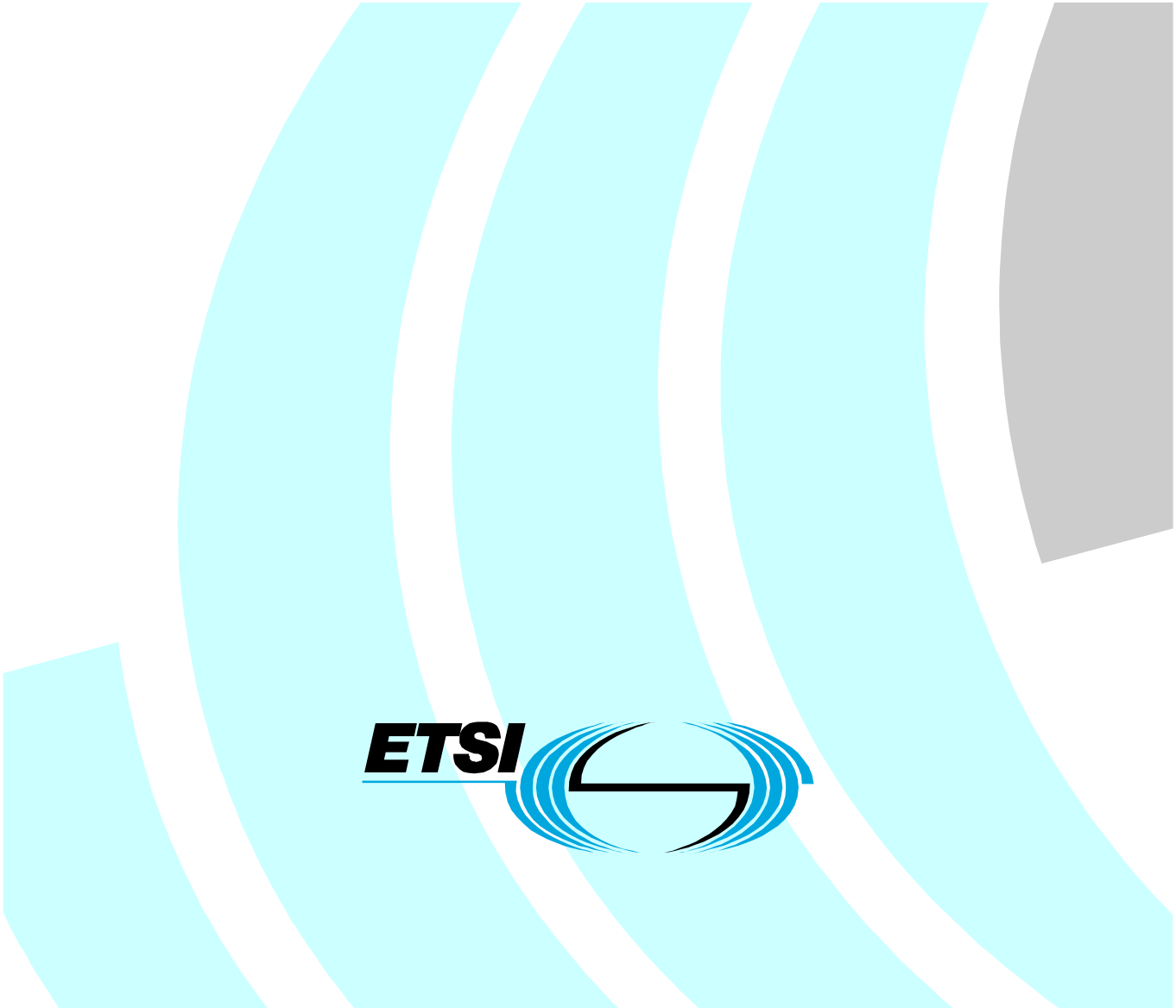
ETSI TR 102 287 V1.1.1 (2004-05)

Technical Report

Satellite Earth Stations and Systems (SES);
Broadband Satellite Multimedia (BSM);
IP Interworking over satellite;
Security aspects



Reference
DTR/SES-00082

Keywords
broadband, interworking, IP, multimedia, satellite, security

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - NAF 742 C
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° 7803/88

Contents

Intellectual Property Rights

IPRs essential or potentially essential to the present document may have been declared to ETSI. The information pertaining to these essential IPRs, if any, is publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (http://webapp.etsi.org/IPR/home.asp).

Pursuant to the ETSI IPR Policy, no investigation, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Foreword

This Technical Report (TR) has been produced by ETSI Technical Committee Satellite Earth Stations and Systems (SES).

Introduction

The present document establishes a framework for specifying security requirements for Broadband Satellite Multimedia (BSM) Networks based on the Internet Protocol (IP) suite of protocols and standards developed in ETSI and other bodies. It investigates how security standards can be adapted, translated or made transparent to satellite transmission protocols and equipment. It identifies new specifications to improve the security of BSM systems in transporting Internet and multimedia traffic.

The present document presents a threat analysis for the BSM system and defines the security services against these threats at various layers of the BSM protocol stack. In addition, it also investigates the security management architecture with focus on IP Security (IPSec) key management for unicast and multicast traffic. As such, it will enable satellite and Internet service providers to establish secure and trusted connectivity for IP traffic.

The recommended ETSI specifications will ensure end-to-end security and in turn wide acceptance of BSM in the Internet community.

1 Scope

The present document reviews the threats and applicable security services that are relevant to BSM systems and leads to recommendations for new technical specifications in this area.

The present document only considers geostationary satellites and fixed terminals.

2 References

For the purposes of this Technical Report the following references apply:

[1] D McGovern BT Broadcast and Satellite Communications: "Pushing the internet direct to the user - Security issues", IEE publication 2003.

[2] ETSI TR 101 984: "Satellite Earth Stations and Systems (SES); Broadband Satellite Multimedia; Services and Architectures".

[3] ETSI TR 101 985: "Satellite Earth Stations and Systems (SES); Broadband Satellite Multimedia; IP over Satellite".

[4] ETSI TS 102 292: "Satellite Earth Stations and Systems (SES); Broadband Satellite Multimedia (BSM) services and architectures; Functional architecture for IP interworking with BSM networks".

[5] ATM Forum Technical Committee af-sec-0100.002 (March 2001): "ATM Security Specification Version 1.1".

[6] ETSI TS 103 197: "Digital Video Broadcasting (DVB); Head-end implementation of DVB SimulCrypt".

[7] ETSI EN 301 192: "Digital Video Broadcasting (DVB); DVB specification for data broadcasting".

[8] ETSI EN 301 790: "Digital Video Broadcasting (DVB); Interaction channel for satellite distribution systems".

[9] Open Mobile Alliance OMA-Download-DRM-v1-0-20020905-C: "Digital Rights Management Version 1.0" Version 05-September-2002, http://www.openmobilealliance.org/tech/docs/index.htm#DRM.

[10] ETSI TS 102 293: "Satellite Earth Stations and Systems (SES); Broadband Satellite Multimedia (BSM) services and architectures; IP Interworking over satellite; Multicast group management; IGMP adaptation".

[11] ETSI TS 102 189-3: "Satellite Earth Stations and Systems (SES); Regenerative Satellite Mesh - A (RSM-A) air interface; MAC/SLC layer specification; Part 3: SLC layer".

[12] ETSI. DTS/SES-00081: "Satellite Earth Stations and Systems (SES); BSM Air Interface Specification; Common Air interface specification; Satellite Independent Service Access Point SI-SAP".

[13] BRAHMS IST project publications on satellite security: http://brahms.tilab.com.

[14] GEOCAST IST project publications on satellite security: http://www.geocast-satellite.com/index2.phtml.

[15] SATIP6 IST project publications on satellite security: http://satip6.tilab.com.

[16] ESA Contract 16996/02/NL/US (Octalis 2003): "Next Generation Conditional Access Systems for Satellite Broadcasting".

[17] H. Cruickshank, S. Iyengar, M Howarth, Z. Sun, F. Zeppenfeldt and G. Kenny: "Secure IP multicast over satellites". ETSI Broadband Satellite Multimedia (BSM)" working group Meeting 13 ETSI; Sophia Antipolis; France February 2003.

[18] Z. Sun, M.P. Howarth, H. Cruickshank, S. Iyengar and L. Claverotte: "Networking Issues in IP Multicast over Satellite" International Journal for Satellite Communications, Special Issue on QoS for Satellite IP Paper number: 2002-Q08, 2003.

[19] IETF RFC 3547: "The Group Domain of Interpretation".

[20] H. Cruickshank, S. Iyengar and M. P. Howarth and Z. Sun: "Key management and multi-layer IPSec for in satellite multicast" Joint COST 272-280 Workshop, European Space Agency (ESA), Holland, June 2003.

[21] Z. Sun, and H Cruickshank, S. Iyengar and M. P. Howarth: "IP multicast over satellite" proceedings of the 21st AIAA International Communication Satellite Systems Conference and Exhibit, April 2003.

[22] S. Josset, L. Duquerroy, M. Önen, M. Annoni, G. Boiero, N. Salis: "Satellite IP SEc: An optimized way of securing multicast wireless communications", Information Security Solutions Europe - ISSE 2002, Disneyland Paris, 2-4 October 2002.

[23] Z. Sun, H. Cruickshank, S. Iyengar and M. Howarth: "IP Multicast over Satellites - Technology Challenges", proceedings of the 20th AIAA International Communication Satellite Systems Conference and Exhibit, Montreal, Canada, 12-15 May 2002.

[24] IETF draft-ietf-msec-ipsec-multicast-issues-01.txt: "IP Multicast issues with IPSec".

[25] U. Reimers: "Video Broadcasting. The International Standard for Digital Television" published by Springer, ISBN 3-540-60946-6. 2001.

[26] S. Iyengar, H. Cruickshank and Z. Sun: "Security issues in IP Multicast over GEO Satellites", AIAA 19th conference in Toulouse, April 2001.

[27] S. Setia.: "Kronos: A Scalable Group Re-Keying Approach For Secure Multicast", Proceedings IEEE Symp. on Research in Security and Privacy 2000, pp.215-228.

[28] H. Cruickshank, Z. Sun, S, Iyengar: "Secure IP multicast over GEO satellites", IEE Colloquium on Broadband Satellite: the Critical Success Factors Technology, Services & Markets, 16-17 October 2000.

[29] IETF RFC 2401: "Security Architecture for the Internet Protocol".

[30] IETF RFC 2402: "IP Authentication Header".

[31] IETF RFC 2406: "IP Encapsulating Security Payload (ESP)".

[32] IETF RFC 2408: "Internet Security Association and Key Management Protocol (ISAKMP)".

[33] IETF RFC 2409: "The Internet Key Exchange".

[34] IETF RFC 2627: "Key Management for Multicast: Issues and Architectures".

[35] D. Matthew J. Moyer, Josyula R. Rao, and Pankaj Rohatgi: "A Survey of Security Issues in Multicast Communications", IEEE Network Magazine, November/December 1999.

[36] R. Canetti, J. Garay, G. Itkis, D. Micciancio, M. Naor and B. Pinkas: "Multicast security: A taxonomy and some efficient constructions", in Proceedings of the 18th Annual Joint Conference of the IEEE Computer and Communications Societies - INFOCOM 1999. Volume 2, pages 708-716. March 1999.

[37] ETSI ETR 289. "Digital Video Broadcasting (DVB); Support for use of scrambling and Conditional Access (CA) within digital broadcasting systems".

[38] B. Schneier: "Applied cryptography".

[39] M. Howarth, S. Iyengar, Z. Sun and H. Cruickshank: "Dynamics of key management in satellite multicast". Accepted for publication in the IEEE JSAC special issue on Broadband IP Networks via Satellites, which will be published in the first quarter of 2004.

[40] IETF draft-ietf-msec-mikey-05.txt: "MIKEY: Multimedia Internet KEYing".

[41] IETF draft-duquer-fmke-00: "The Flat Multicast Key Exchange protocol", L. Duquerroy, S. Josset.

[42] IETF draft-irtf-gsec-lifecycle-00.txt: "Life cycle key management costs in secure multicast".

[43] IETF draft-ietf-msec-gsakmp-sec-03.txt: "Group Secure Association Key Management Protocol (GSAKMP)".

[44] ITU-T Recommendation X.509: "Information technology - Open Systems Interconnection - The Directory: Public-key and attribute certificate frameworks".

[45] IETF RFC 2246: "The TLS Protocol Version 1.0".

[46] IETF RFC 3135: "Performance Enhancing Proxies Intended to Mitigate Link-Related Degradations".

[47] IETF RFC 822: "Standard for the format of ARPA Internet text messages".

[48] IETF RFC 2015: "MIME Security with Pretty Good Privacy (PGP)".

[49] IETF RFC 2407: "The Internet IP Security Domain of Interpretation for ISAKMP".

[50] IETF draft-duquer-fmke-00.txt: "The Flat Multicast Key Exchange protocol".

[51] draft-ietf-msec-gsakmp-sec-04.txt: "GSAKMP".

3 Definitions, symbols and abbreviations

3.1 Definitions

For the purposes of the present document, the following terms and definitions apply:

authentication: communicating parties are assured of each other's identity

authorization: checking that an authentic user is authorized to perform specific tasks

cipher key: value that is used to determine the transformation of plain text to cipher text in a cryptographic algorithm

cipher text: data produced through the use of encipherment

confidentiality: only trusted users can access the content of the data

decipherment: reversible encipherment

hash/message digest: mathematical formula that converts a message of any length into a unique fixed-length string of digits (typically 160 bits) known as "message digest" that represents the original message

NOTE: A hash is a one-way function - that is, it is infeasible to reverse the process to determine the original message. Also, a hash function will not produce the same message digest from two different inputs.

digital signature: electronic signature that can be used to authenticate the identity of the sender of a message, or of the signer of a document

NOTE: It can also be used to ensure that the original content of the message or document that has been conveyed is unchanged.

digital certificates: electronic document that establishes your credentials when doing business or other transactions on the web

NOTE: They are issued by a certificate authority and contain a user's name, expiration dates, a copy of the certificate holder's public key, and the digital signature of the certificate-issuing authority so that a recipient can verify that the certificate is real. Some digital certificates conform to a standard such as ITU-T Recommendation X.509 [44].

encipherment: cryptographic transformation of data to produce cipher text

integrity: prevention of unauthorized modification of information using secret key message authentication code or public key digital signature

non-repudiation: a user cannot deny the fact that it has accessed a service or data

plain text: unencrypted source data

3.2 Symbols

For the purposes of the present document, the following symbols apply:

~	Concatenation
C/N	Carrier to Noise ratio

3.3 Abbreviations

For the purposes of the present document, the following abbreviations apply:

ACF	Access Control Field
AES	Advanced Encryption Standard
AH	Authentication Header
ATM	Asynchronous Transfer Mode
BoD	Bandwidth on Demand
BSM	Broadband Satellite Multimedia
BSM-SM	BSM Security Manager
CA	Certification Authority
CA	Conditional Access
CAM	Content Addressable Memory
CAT	Conditional Access Table
CPU	Central Processing Unit
CW	Control Word
DCF	DRM Content Format
DES	Data Encryption Standard
DOI	Domain of Interpretation
DoS	Denial of Service
DRM	Digital Rights Management
DSL	Digital Subscriber Line
DSM-CC	Digital Storage Media - Command and Control
DSS	Digital Signature Standard
DVB	Digital Video Broadcast
DVB-RCS	DVB, Return Channel Satellites
DVB-S	Digital Video Broadcast by Satellite
DVB-T	Digital Video Broadcasting - Terrestrial
ECM	Entitlement Checking Message
EKE	Explicit Key Exchange
EMM	Entitlement Management Message
ESP	Encapsulated Security Payload
FEC	Forward Error Correction
FMKE	Flat Multicast Key Exchange
GC	Group Controller
GCKS	Group Controller and Key Server

GDOI	Group Domain of Interpretation
GEO	Geostationary Earth Orbit
GM	Group Member
GSAKMP	Group Secure Association Key Management Protocol
GTEK	Group TEK
IC	Integrated Circuit
IKE	Internet Key Exchange
IP	Internet Protocol
IPSec	Internet Security
ISAKMP	Internet Security Association Establishment and Key Management Protocol
ISL	Inter-Satellite Link
KEK	Key Encrypting Keys
LAN	Local Area Network
LKH	Logical Key Hierarchy
MAC	Message Authentication Code
MD5	Message Digest 5
MIKEY	Multimedia Internet KEYing
MIME	Multipurpose Internet Mail Extensions
MKE	Main Key Exchange
MPEG	Moving Picture Experts Group
MPEG-TS	MPEG Transport Stream
MSEC	Multicast security group in the IETF
MUX	Multiplexer
NACK	Negative ACKnowledgement
NAT	Network Address Translation
NCC	Network Control Centre
NOCC	Network Operations Control Centre
OBC	On-Board Controller
OBP	On-Board Processor
OBS	On-Board Switch
OMA	Open Mobile Alliance
OS	Operating System
OSI	Open Systems Interconnection
PEP	Performance Enhancing Proxy
PES	Program Elementary Stream
PGP	Pretty Good Privacy
PHY	PHYsical
PID	Packet Identifier
PKI	Public Key Infrastructure
PMT	Program Map Table
POP	Proof-of-Possession
PPV	Pay-Per-View
QKE	Quick Key Exchange
QoS	Quality of Service
RCST	Return Channel Satellite Terminal
RSA	Rivest, Shamir and Adleman
RSM-A	Regenerative Satellite Mesh – Type A
RTCP	Real time Transport Control Protocol
RTJ	Request To Join group
RTP	Real time Transport Protocol
RTSP	Real Time Streaming Protocol
SA	Security Association
SACK	Selective Acknowledgement
SAM	Security Access Module
SAR	Segmentation And Reassembly
SAS	Subscriber Authorization System
SAT-RMTP	Satellite Reliable Multicast Transport Protocol
SCH	Secure Hash Algorithm
SDP	Session Description Protocol
SDR	Session Directory, Revised
SI	Service Information
SIP	Session Initiation Protocol

SLC	Satellite Link Control
SLC	Satellite Link Control
SMAC	Satellite Medium Access Control
SMS	Subscriber Management System
SP	Service Provider
SPHY	Satellite PHYsical
SPI	Security Parameter Index
SRTCP	Secure RTCP
SRTP	Secure RTP
SSH	Secure Shell
SSL	Secure Socket Layer
ST	Satellite Terminal
STB	Set Top Box
STF	Special Task Force
TC	Transmission Convergence
TCP	Transmission Control Protocol
TDMA	Time Division Multiple Access
TEK	Traffic Encryption Keys
TLS	Transport Layer Security
TS	Transport Stream
TT&C	Telemetry Tracking and Command
UDP	User Datagram Protocol
USB	Universal Serial Bus
UTOPIA	Universal Test & Operations Physical Interface for ATM
VC	Virtual Connection
VoD	Video-on-Demand
VPN	Virtual Private Network
XML	eXtensible Markup Language

4 Overview

Security has many aspects in satellite and IP networks. In order to focus on BSM specifics, the present document addresses only those aspects that impact BSM architectures or are impacted by BSM architectures.

In addition, the present document combines the contributions of a number of standardized bodies in each clause. Each topic is discussed in more detail in a separate clause. Table 1 lists all the clauses together with a brief description of the contents of each clause. The reader is assumed to know the basics of security and its implementation in communication and data networks. This note is also organized by layers, as proposed by the BSM stack of protocols. Another organization could have followed a more functional approach based on authentication, data privacy and integrity. All these topics are nevertheless addressed as layer 2 or above functions in clauses 7 and 8.

Table 1: Security overview

Topic	Clause	Description
Introduction to security	5	Relationship between BSM architecture and security
Threats and countermeasures	6	Analysis of threats to BSM system and definition of security services that can be used as countermeasures
Security layering	7	Security solutions are presented in the link, network, transport and application layers
Security management	8	Various security management protocols for IPSec, DVB-RCS and DVB-S system and presented
Recommendation	9	What specifications should ETSI develop in order to ensure that BSMs can support security

5 Introduction to BSM security

Although the Internet has developed as a point-to-point network service, there are some limitations in its adaptation to large volumes, maintaining user performance expectations, and driving down costs. The broadcast media offer a means of tackling these problems, and also of increasing the take-up of multicast services in a way never fully exploited by the terrestrial Internet.

A general broadcast-based system configuration is shown in figure 1 (see [1]), whether terrestrial or satellite, has the further advantages of service coverage to rural and remote areas, and a simple topology with the minimum of intermediate switching nodes between core network and user. A return channel is still generally required, and this can be provided via the dial-up copper or wireless network which already supports access to the narrowband Internet.

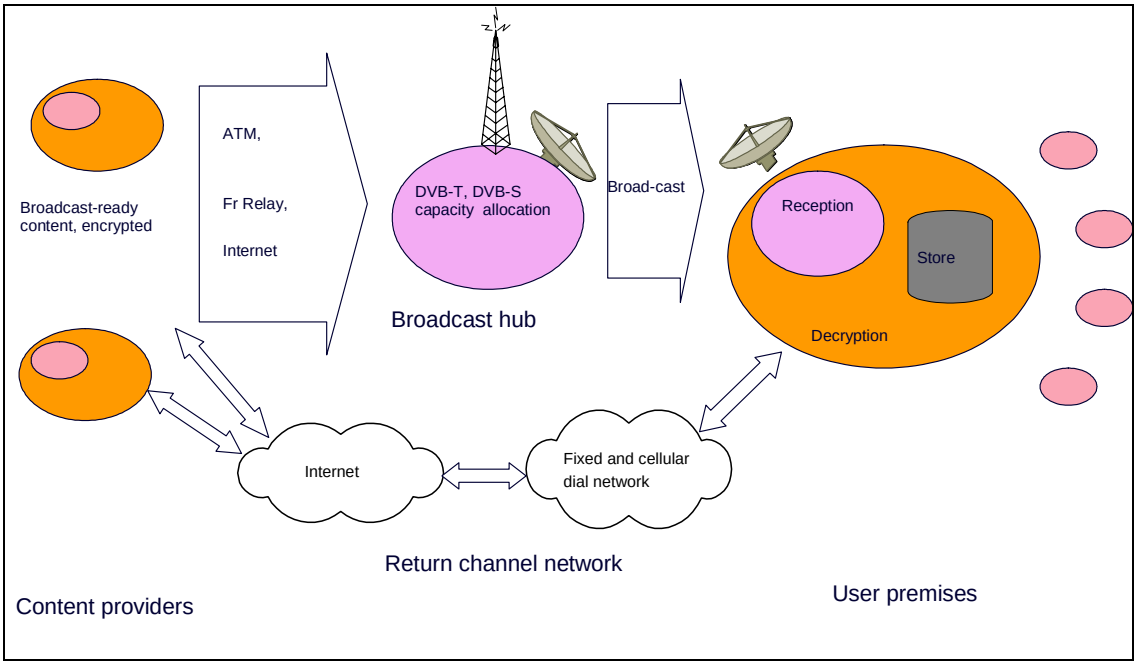


Figure 1: Broadcast broadband system

Therefore satellite revenues in the next few years are increasingly likely to come from the delivery of IP-based applications and services, either to complement terrestrial broadband services, or to offer added-value services in some niche markets. The challenge for the next generation of satellite systems is therefore to define a common basis for efficient integration of satellites in IP-centric secure telecommunication networks. Satellite access, rather than long-distance transport, is seen as a particularly convenient element of the overall telecommunication infrastructure, since it provides ubiquitous broadband access to anyone deploying a satellite terminal, both for single residential users and SOHO/corporate networks.

In addition, security is becoming an important issue for the success of any services offered by satellite networks. However satellite environment has its own security challenges such as:

- eavesdropping and active intrusion is much easier than in terrestrial fixed or mobile networks because of the broadcast nature of satellites;
- satellite channels may require robust security synchronization. This demands a careful evaluation of encryption systems to prevent the Quality of Service (QoS) degradation because of security processing.

Therefore the BSM security system must be efficient and have minimum impact on BSM system performance. Another reason for security provisioning is that protection of Content may be regarded as a significant part of the BSM value, and may most conveniently be done within the BSM solution.

Provision of a BSM security structure does not mean a complete security solution. It is expected that certain elements including encryption algorithms, keys and key generation processes, and some hardware elements, will remain external or proprietary. BSM topologies are very diverse, and support a wide range of applications for which there is an equally wide range of security requirements. Examples of services that can be carried over BSM networks are:

- a videoconference application with security implemented only in the application software, emails protected via SMIME, and Internet connection to credit card purchasing protected by Secure Socket Layer (SSL);
- a dedicated VPN connection to an employer corporate network using IPSec and the employer's nominated key generator agent;
- delivery of subscription-based Content, in file or streaming form, protected by CA.

The structure of the present document is organized in a way to provide an introduction to the security issues in the BSM architecture in clause 5. Clause 6 provides a brief analysis of security threats and counter measures in BSM network. Clause 7 presents security solutions in the various layers of the BSM protocol stack such as link layer security (DVB-S, DVB-RCS and ATM), network layer security (IPSec), transport layer (such as SSL) and application layer security (such as secure XML and DRM). Clause 8 presents security and key management issues and possible solutions in BSM networks. Clause 9 provides the recommended ETSI specifications for security.

5.1 Relation between satellite link characteristics and security

There are aspects of BSM links, which differ from their terrestrial equivalents. The Service Providers (SP) should be aware of these and take steps to ensure that any security solution selected by the user will work over the particular BSM service offered.

Delay

Each BSM service is designed with a topology with a particular delay and delay variation. No upper limit of delay or delay variation range is specified in BSM, this being a matter for individual service designers and operators.

Regarding security, the processing delay of encryption must be kept to minimum.

Number of hops

There is no restriction in principle to the number of satellite hops in a particular BSM service, or an upper limit to the total link delay. It is in an SP's interest to minimize the number to ensure minimum delay, but the flexibility of BSM means that in some cases several hops may be the best solution. The SP must therefore be aware of the number of hops in a service to ensure that suitable provision can be made to verify security operation. Hop-by-hop encryption or authentications is not desirable if the number of hops is large.

Error performance

BSM links may generally be assumed to be Quasi-Error-Free during the period of link-available. The use of concatenated or turbo Forward Error Correction (FEC) means that there is an approximate 8 decades per dB of BER variation per dB of C/N ratio, meaning that the duration of periods with a significant error rate is very small, and can be discounted. Fades due to atmospheric or other causes will occur and may cause significant outage periods. However, BSM link errors can lead to loss of security synchronization, which can impact the BSM network throughput performance.

Fade events

BSM has no particular fade performance specification, this being a subject of individual service definition. Some links may be designed to 99,99 % or better, while others may be designed to work in a frequent-fade mode, for example to accommodate very small fixed dishes, or mobile terminals. The appropriate degree of tolerance to outages needs to be designed into the service application, and into any security system, which it uses.

Network topology and connectivity options

There are only a small number of topologies and that the Satellite Independent security will be BSM agnostic; where the system type counts is in the satellite dependent part; the SI-SAP should be used.

Performance Enhancing Proxies (PEPs) and other Layer 4 network functions

PEPs are used in some high-speed satellite systems to recover the effect of link delay on TCP operation. BSM does not specify particular PEPs or make any restriction on their operation. Since they involve the interference in TCP operation at end and intermediate network nodes, it is possible that they may interfere with the operation of some security systems. The SP must therefore verify the operation of his BSM system with regard to any use of PEPs, and confirm the operation of a security overlay accordingly.

Regarding security, there are problems for using PEP with Internet security (IPSec), where end to end IPSec will prevent the proper working of satellite PEP.

5.2 BSM system architecture

BSM systems are composed of a space segment, of one or more satellites, and of a ground segment made of a Network Control Centre (NCC), of gateways and of individual Satellite Terminals (STs) (see figure 2). Depending on the type of payload in the satellites different network architectures are made possible.

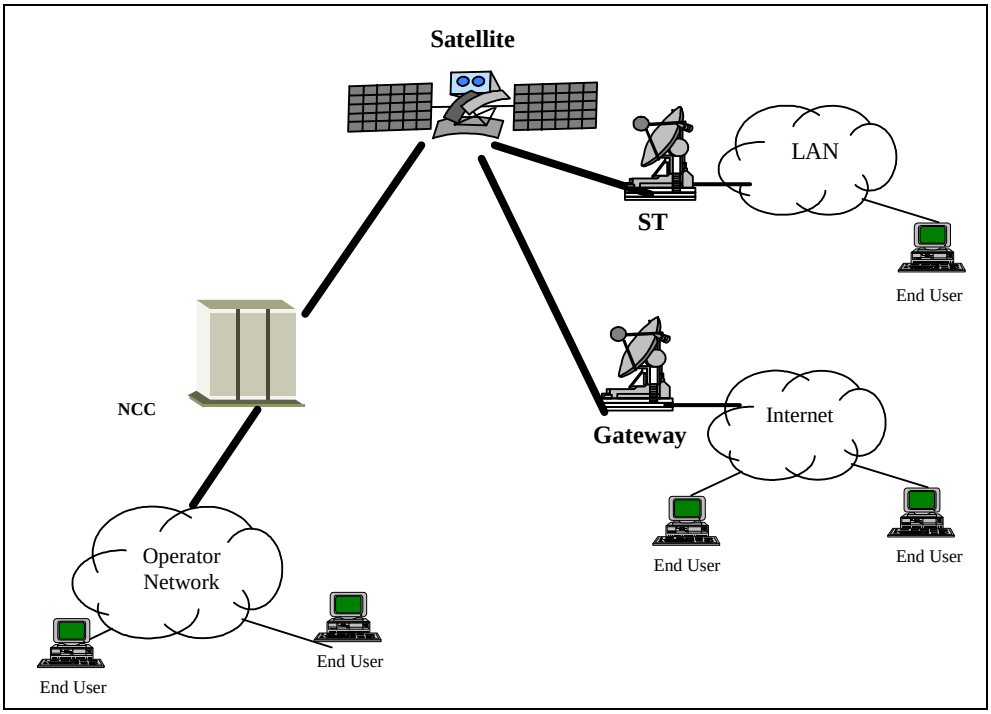


Figure 2: Generic BSM system

As defined in TR 101 984 [2], the BSM offers a general IP service that needs to evolve with the changes in the Internet. In the context of present document the use cases include but are not limited to:

- point-to-point connectivity;
- Internet access;
- content distribution; and
- real time multimedia streaming.

These use cases are shared amongst all current projects of the Special Task Force (STF) as well as having been addressed in both the Architectures [2] and Internet Protocol [3] TRs. The use cases are provided by BSMs using three main network architectures that support point to point, multicast and broadcast services, namely:

- access network;
- content distribution to the edge;
- core network.

A BSM network can support all three scenarios. However, the present document will give priority to issues related to the first two scenarios, namely access network scenarios and content distribution to the edge. This is because it does not specifically address multicast services where the BSM can play an important core role. In unicast services the data rates usually associated with core networks are above those offered by BSMs by orders of magnitudes (terabits per seconds on optical core networks).

The BSM satellite systems can be divided into the following types:

Transparent system

A BSM with a non-regenerative payload (a repeater) is commonly called a "bent-pipe system" or transparent system. This system does not terminate any layers of the BSM protocol stack in the satellite. The satellite simply repeats the signals from the user links to the feeder links transparently. With this system (which can use global and spot beams) the communications between a Satellite Terminal (ST) and the Internet are done via a gateway terminal attached to the Internet. The forward channel uses the satellite transmission. However, the return channel can use a number of technologies (e.g. satellite, phone or DSL network etc). This system is mainly used for access as it requires double satellite hops for ST to ST communications. In this system, all network functions are performed by the NCC.

Regarding security, transparent satellites link should be secured either at the satellite link layer or at the higher layer.

Regenerative satellites (OBP)

A regenerative satellite offers bridging or network functionality in the satellite. Usually, this added functionality is to maximize the efficiency of multi-beam satellites and to improve allocation of spectrum resources on the uplink. In general the On-Board Processor (OBP) uses an On-Board Switch (OBS) to send BSM cells from beam to beam (digital switching). An On-Board Controller (OBC) manages the uplink and downlink resources as well as some performance management onboard. In this system the Network Control Centre (NCC) is used for overall coordination, non real time resource management and network management. This system enables single-hop ST-to-ST (peer to peer) communications while still enabling access when required.

Regarding security, OBP satellites requirements are similar to transparent plus there is a need to secure communications between the NCC and the OBP.

Each of the above use case has some security requirements that have to be addressed. Examples of such requirements are confidentiality and integrity of data from source to end-users, source authentication, protection of the management of the infrastructure from unauthorized people and protection against denial of service attacks. More detailed analysis is presented in clause 6. One important requirement is transparency, where BSM networks should support any data encryption service, which operates in a manner that does not make specific demands or constraints on the network. Any such service must be able to operate in the transmission conditions of the type of BSM service selected, and must comply with any legal provisions, which apply.

6 BSM security threats and countermeasures

This clause introduces basic security concepts. A good place to begin is by defining the basic concepts involved in securing any object. The key words in the security lexicon are vulnerability, threat, attack, and countermeasure.

Vulnerability is the susceptibility of a situation to being compromised. It is a potential, a possibility, and a weakness, an opening. Vulnerability in itself may or may not pose a serious problem, depending on what tools are available to exploit that weakness. For example, the use of the public Internet to carry user data and network management traffic is vulnerability.

A threat is an action or tool, which can exploit and expose vulnerability and therefore compromise the integrity of a given system. Not all threats are equal in terms of their ability to expose and exploit the vulnerability.

An attack defines the details of how a particular threat could be used to exploit vulnerability. It is entirely possible that situations could exist where vulnerabilities are known and threats are developed, but no reasonable attack can be conceived to use the specific threat upon a vulnerability of the system.

Countermeasures are those actions taken to protect systems from attacks, which threaten specific vulnerabilities. In the network security world, countermeasures consist of tools such as virus detection and cleansing, packet filtering, password authentication, and encryption.

Any security scheme must identify vulnerabilities and threats, anticipate potential attacks, assess whether they are likely to succeed or not, assess what the potential damage might be from successful attacks, and then implement countermeasures against those defined attacks which are deemed to be significant enough to counter. Threats to BSM systems could be broadly categorized into network, software, hardware and human (user) threats.

6.1 Network threats

The simplest type of network threats are called passive attacks, the passive attackers are involved in eavesdropping on, or monitoring of, transmissions, with a goal to obtain information that is being transmitted. The two types of passive attacks are release of message contents and traffic analysis. Passive attacks are very difficult to detect because they do not involve any alteration of the data, also the attacker does not need sophisticated tools or knowledge to perform such attacks.

Active attacks are in general more difficult to implement successfully than passive attacks, and usually require more sophisticated resources. Examples of active attacks are given below:

Masquerade (impersonation)

A masquerade is an entity that pretends to be a different entity. A masquerade is usually used with some other forms of active attack, especially replay and modification of messages. For instance, authentication sequences can be captured and replayed after a valid authentication sequence has taken place. An authorized entity with few privileges may use a masquerade to obtain extra privileges by impersonating an entity that has those privileges. Typically this is achieved by stealing other people's passwords or credentials. For example replay attacks are a form of masquerade where a valid message containing authentication information may be replayed by another entity in order to authenticate itself to the authenticating entity.

Modification of Messages (manipulation)

Modification of messages occurs when the content of a data transmission is altered without detection and results in an unauthorized effect. This is the case when, for example, a message "Allow Alice to read confidential file Accounts" is changed to "Allow Bob to read confidential file Accounts".

Repudiation by a party

Repudiation by a party can be either of origin or destination. Repudiation of origin occurs when a party denies being the originator of a message and repudiation of destination occurs when a party denies the reception of a message. For example when a client downloads a movie from a pay per view service, then the client should not be able to denying the downloading of the movie in order not to pay the charges.

Denial of Service (DoS) attacks

Denial of Service attacks occurs when an entity fails to perform its proper function or acts in a way that prevents other entities from performing their proper functions.

The DoS attack may be general, as when an entity suppresses all messages, or there may be a specific target, as when an entity suppresses all messages directed to a particular destination, such as the security audit service. The attack may involve suppressing traffic as described in this example or it may generate extra traffic. It is also possible to generate messages intended to disrupt the operation of the network, especially if the network has relay entities that make routing decisions based upon status reports received from other relay entities.

6.2 Software threats

Many systems fail because of mistakes in implementation. Some systems do not ensure that plaintext is destroyed after it is encrypted. Other systems use temporary files to protect against data loss during a system crash, or virtual memory to increase the available memory; these features can accidentally leave plaintext lying around on the hard drive. In extreme cases, the operating system can leave the keys on the hard drive. Moreover confidential information of a company or clients should be stored securely at the provider's site or it creates a serious threat since the provider will have all confidential information of clients, which could be the misused.

Some examples of software threats are:

Trapdoor

When an entity of a system is altered to allow an attacker to produce an unauthorized effect either on command or on a predetermined event or sequence of events, the result is called a trapdoor. For example, a password validation could be modified so that, in addition to its normal effect it also validates an attacker's password.

Trojan horse

When introduced to the system, a Trojan horse produces an unauthorized function in addition to its authorized function. For example, a relay that also copies messages to an unauthorized channel is a Trojan horse.

Threat from cryptographic designs

A cryptographic system can only be as strong as the encryption algorithms, digital signature algorithms, one-way hash functions, and message authentication codes it relies on. Break any of them, and you have broken the system.

Systems are often found "void the warranty" of their cryptography by not using it properly: failing to check the size of values, reusing random parameters that should never be reused, and so on. Encryption algorithms do not necessarily provide data integrity. Key exchange protocols do not necessarily ensure that both parties receive the same key. Random-number generators are another place where cryptographic systems often break. Good random-number generators are hard to design, because their security often depends on the particulars of the hardware and software. The cryptography may be strong, but if the random-number generator produces weak keys, the system is much easier to break.

Threat from complete Operating System (OS) dependence

If the security of software application depends only on the security of the operating system as opposed to complementing it also ensures a threat.

6.3 Hardware implementation threats

Some systems, particularly commerce systems, rely on tamper-resistant hardware for security: smart cards, electronic wallets, dongles, etc. Hardware security is an important component in many secure systems; security rests solely on assumptions about tamper resistance is not enough. When design systems that use tamper resistance, it is better to build in complementary security mechanisms just in case the tamper resistance fails.

All hardware systems including client stations, satellite terminals and network equipment like routers and firewalls can provide a way of attack if not properly configured, since they will become the entry point of attack. Unauthorized access to these machines also possesses a threat since it means access to the system. If all the major hardware systems are not backed up in case of emergency like power outage or denial of service attacks then it possess a serious threat as the data stored in these systems as well as the availability of the service as a whole is disrupted.

6.4 Human (user) threats

Insider attacks

Insider attacks occur when legitimate users of a system behave in unintended or unauthorized ways. Most known computer crime has involved insider attacks that compromised the security of the system. For example the users of the system should not divulge secret information of the company resources by giving their username/passwords to other people who are not part of this company.

One of the important insider attacks is piracy threat, where one of the legitimate members of the group can give the BSM application to others without the groups knowledge.

Many systems break because they rely on user-generated passwords. Left to them, people do not choose strong passwords. Even when a system is secure if used properly, its users can subvert its security by accident, especially if the system is not designed very well. The classic example of this is the user who gives his password to his co-workers so they can fix some problem when he is out of the office. Users may not report missing smart cards for a few days, in case they are just misplaced. They may not carefully check the name on a digital certificate. They may reuse their secure passwords on other, insecure systems. They may not change their software's default weak security settings. If there are no trained staff (administrators) to monitor and configure the systems and network then this could become a major threat.

Outsider attacks

Outsider attacks are carried out in order to gain entry into the system since they are not members of that organization or clients of a provider. They may use techniques such as wiretapping (active), intercepting and replaying or modifying messages, disrupting services using denial of service attacks etc. in order to carry out the network attacks as mentioned earlier.

6.5 Security services definition

Examining clauses 6.1 to 6.4 shows that eavesdropping (passive attacks) can be considered as a major threat to BSM networks, especially for broadcast services. Also there are other major network threats to BSM networks such as impersonation, message modifications and denial of service attacks. These threats will require appropriate security counter measures. Other issues such as software, hardware and human threats will need some other measures such as good software and hardware design and maintenance, proper satellite equipment testing, and proper training of satellite personnel and customers regarding basic security issues.

In order to counter the major threats mentioned earlier, we could apply some fundamental security services, including authentication, integrity, and confidentiality authorization and non-repudiation:

- **Confidentiality/privacy/secrecy/service** are used to create a private session. Although encryption is typically used to provide this service, a weaker form of confidentiality may be achieved by limiting the routing of session datagrams. Confidentiality can be used as a countermeasure against eavesdropping, masquerading, traffic analysis and leakage of sensitive information by exploiting processes with legitimate access to the data.

- **The integrity service** guarantees that the messages are received with no modification by unauthorized entities. In order to provide this service the mechanisms used are encipherment, digital signatures, and Message Authentication Codes (MAC). This helps to prevent someone from re-injecting previously authenticated packets into a traffic stream. It also prevents manipulation of messages such as messages may be deliberately modified, inserted, replayed, or deleted by an intruder.

- **Authentication** is a service offered used to verify the identity of entities involved in a communication. The simplest technique is user ID and password. Authentication can be mutual (both communicating entities) or one way (only the originator). The appropriate mechanisms to provide these services are encipherment of information and digital signature. Digital signatures schemes, such as the Digital Signature Standard (DSS); Rivest, Shamir and Adleman (RSA) are examples of a strong authentication mechanisms based on public key technology. Certificates signed by a trusted Certification Authority (CA) are used to bind the identity of an entity to its public key. The certification process provides the necessary assurance to enable the proper identification of entities during the authentication process.

- **Authorization and access control** is a service where each individual user privileges are verified. This service is normally needed in conjunction with authentication in order to provide access control. This prevents an unauthorized use of a resource such as intruders can access services by masquerading as users or network entities. Also prevents denial of service attacks such as disturbing, misusing network services or resource exhaustion and overloading.

- **Non-repudiation** is a service that prevents a transmitter or receiver from denying its acts. The main mechanism used for this service is digital signatures.

7 Security layering in the BSM protocol stack

The BSM protocol stack is defined in TS 102 292 [4] and is shown in figure 3.

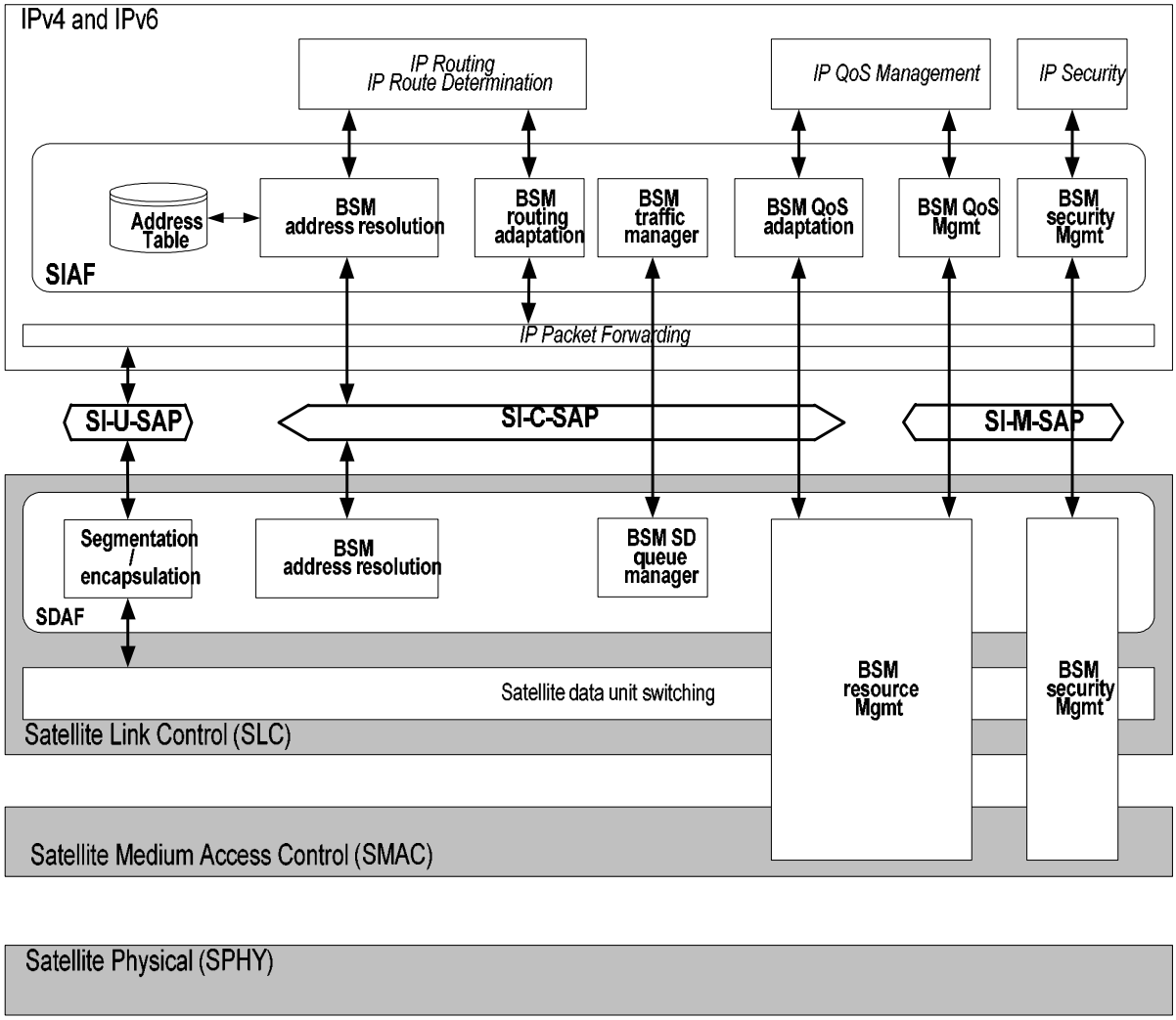


Figure 3: BSM protocol stack

Security may be provided at any level of the BSM protocol stack such as link, network, transport or application layers. In this discussion for simplicity we ignore presentation and session layers in the classic OSI model and focus on the model provided by Internet protocols. In general, there is a need to establish a trust relationship between users of the end-to-end security system through a security management system. The security operations may be visible to end users and applications if they are implemented at the application level, or it can be transparent if implemented in the lower layers. This clause discusses the advantages and disadvantages of each security approach.

7.1 Link level security

Security services can also be provided at the link layer such as Asynchronous Transfer Mode (ATM) cell level and MPEG-TS for DVB-S and DVB-RCS systems. Link layer security has the following advantages:

- security is provided independently of upper layer protocols (whether IP, TCP, UDP, RTP or reliable multicast);
- it can protect satellite link against traffic analysis and illegal changes to satellite network configuration;
- it can provide protection to all real time and non real time applications.

The disadvantages of link layer security are as follows:

- only satellite terminals are authenticated;
- only satellite link traffic can be encrypted and digitally signed.

7.1.1 ATM security

ATM Forum has defined four security services, in the ATM Security specifications [5] as follows:

- user plane security: the user plane security defines the mechanisms to allow for secure communication between nodes in an ATM network. The user plane security, as defined by the ATM Forum, can be subdivided into access control, authentication, data confidentiality, and data integrity;
- control plane security: the control plane defines the call control signalling needed to establish, maintain and close a certain Virtual Connection (VC). Thus, authentic signalling has been defined as the main target of control plane security for any endpoint to endpoint, switch to switch, or endpoint to switch signalling communication;
- support services: the support services define the certification infrastructures, the key exchange mechanisms, and the basic negotiation of security requirements and capabilities;
- management plane security: the management plane is responsible for both performing management functions for the system as a whole (plane management), and for performing network and system management functions such as resource management (layer management).

The ATM Forum's Security Specification states that the ATM cell payload is encrypted and the cell header is unchanged. A survey of available ATM Integrated Circuits (ICs) shows that state-of-the-art Segmentation And Reassembly (SAR) controllers integrate both the AAL and the ATM layer into one unit. Thus, to maintain compatibility between existing ATM hardware and encryption hardware, access to the ATM cell can only be made at the hardware interface between the SAR controller and the Transmission Convergence (TC) unit. This interface has been standardized by the ATM Forum as the Universal Test and Operations Physical Interface for ATM level 2 (UTOPIA). By intercepting the UTOPIA interface a standard compliant key agile ATM cell payload encryption is feasible up to high transmission rates (i.e.155 Mbps). In addition to the high transmission rates possible, a further advantage of intercepting the cell stream at the UTOPIA is that the solution is independent of the hardware since most ATM hardware manufacturers support UTOPIA. Intercepting standardized UTOPIA decouples the encryption hardware from the physical media and meets the objective of being applicable to different media. Even if this hardware architecture seems to be a simple one, there are two important performance related considerations to be made:

- ATM throughput: the encryption unit has to handle the full bi-directional bandwidth;
- statistical multiplexing: a per-VC encryption scheme with unique session keys for each user connection is to be supported. This requires that the cryptographic unit must be capable of changing the keys rapidly (a key agile system). Research in key agility has shown that one encryption unit for each direction is sufficient, if the key memory is integrated in the encryption unit using fast Content Addressable Memory (CAM) techniques.

ATM Forum specifications address the security issues in terrestrial fixed networks only. There is very limited work on done on securing satellite ATM. There are several technical challenges need to be evaluated carefully for securing ATM satellites such as the encryption synchronization in high bit error rates environment, where errors are of bursty nature. Therefore it is important to examine the impact of such errors on ATM cell payload encryption performance. Another issue is the transmission rate and encryption key updating, where ATM has been designed for the high data rates. Therefore, there is a need for a mechanism to change the encryption key frequently. This challenge is not specific to satellites and includes terrestrial ATM networks as well.

7.1.2 DVB-S Conditional Access

Conditional Access (CA) is a service that allows broadcasters to restrict certain programming products to certain viewers. The CA does this by encrypting the broadcaster's programs. Consequently, the programs must be decrypted at the receiving end before they can be decoded for viewing. CA offers capabilities such as Pay-Per-View (PPV), interactive features such as Video-on-Demand (VoD) and games, the ability to restrict access to certain material (adult movies, for example) and the ability to direct messages to specific set-top boxes (perhaps based on geographic region).

The Conditional Access system used in the DVB system [6] and [7] includes three main functions: scrambling/descrambling, entitlement checking and entitlement management.

The scrambling/descrambling function aims to make the service incomprehensible to unauthorized users. Descrambling can be achieved by any receiver having an appropriate descrambler and holding a secret Control Word (CW). Scrambling can be applied to service components, either using a common Control Word or using separate Control Words for each component.

The entitlement checking function consists of broadcasting the conditions required to access a service, together with encrypted secret codes to enable the descrambling for authorized receivers. These codes are sent inside dedicated messages called Entitlement Checking Messages (ECMs) and these are carried in the ensemble.

The entitlement management function consists of distributing entitlements to receivers. There are several kinds of entitlements matching different means of subscribing to a service: subscription per theme, level or class, pre-booked pay-per-programme or impulse pay-per-programme, per service or per time. This information is sent inside dedicated messages called Entitlement Management Messages (EMMs) and these may be carried in the same ensemble as the scrambled services or by some other means. The control and management functions require the use of secret keys and cryptographic algorithms.

To understand how CA is used, we first need to look at the data it encrypts. Each individual program that a broadcaster provides is composed of many elements, such as video, audio and text. In digital television, these elements are converted into digital form using the MPEG-2 codec. The MPEG-2 data associated with each program are broken up into many packets, and the sum total of these packets for each program is called the Program Elementary Stream (PES). The PES for each program is then multiplexed together with those of other programs. This stream of multiplexed programs is then broken up into 188-byte packets for transmission, at which point it is called the Digital Video Broadcast (DVB) MPEG-2 Transport Stream (TS). The CA service can scramble the programming data either at the PES level or the TS level. The preferred option is scrambling at the TS level.

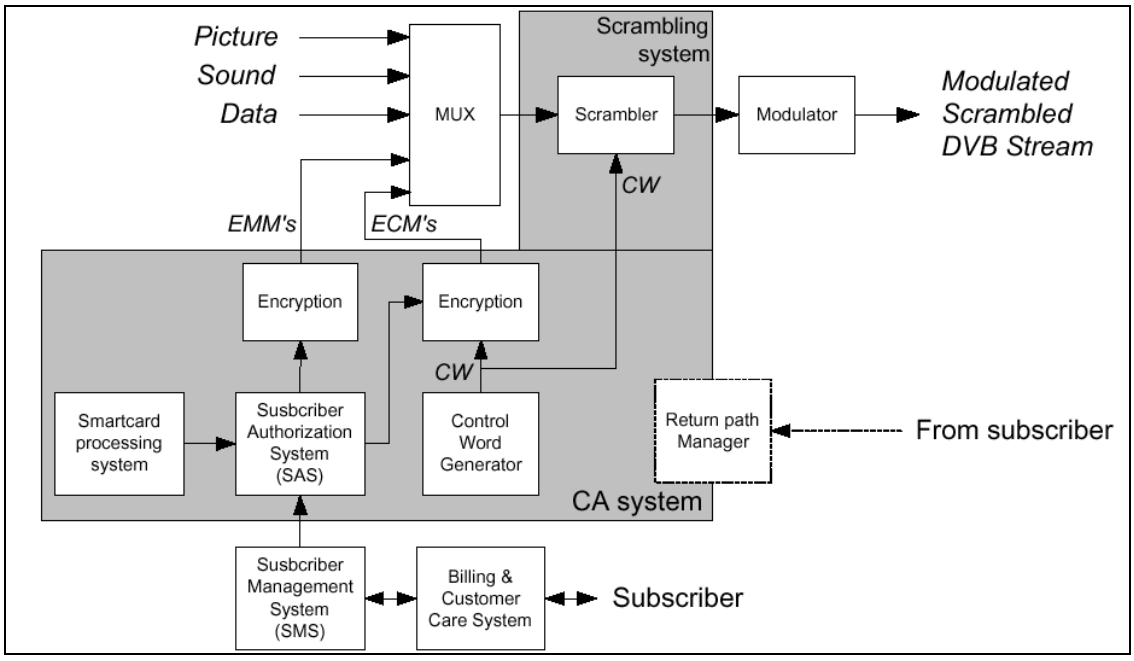


Figure 4: General architecture for conditional access system

A general architecture is shown in figure 4. The main system components are: a Multiplexer (MUX) that combines the video stream, audio stream, data stream and the EMMs and ECMs into a single DVB stream. This multiplexer usually is a dedicated off-the-shelf device. Another component is the Modulator that takes the resulting signal and modulates it for its transmission to the satellite. The third component is the conditional access system that is composed of several specific modules:

- the scrambler: Scrambles the payload of the packets composing the transport stream, using a Control Word generated by the Control Word generator. The scrambler usually scrambles the packets containing the picture and audio information and sometimes some packets containing data. Packets containing EMMs and ECMs are not scrambled. The preferred implementation of the scrambler is in the multiplexer device. Stand-alone scramblers also exist;

- the Subscriber Authorization System (SAS): Processes the different viewing authorizations given to the subscribers and uses them to generate adequate EMMs and ECMs;
- the Control Word Generator that creates the control words: Two encryption engines (often implemented by the same software) are used to encrypt the content of the EMMs and the Control Words stored in the ECMs;
- the smart card processing system: Contains information about the secret information stored into consumer smart cards or set-top boxes. This module is sometimes integrated in the SAS.

The conditional access system needs information from other modules of the system such as:

- the Subscriber Management System (SMS) holds all the data related to subscribers, running subscriptions and payments. This system interacts with the billing and customer care system to generate revenues. The SMS tells which programs subscribers are authorized to view;
- the return path manager (if a return path exists): this module can be used by the conditional access system to perform verification operations and to get feedback on the set-top box status and behaviour.

At the receiving end, it is the job of the Set-Top Box (STB) to descramble the CA encryption and decode the MPEG-2 data for viewing. Figure 5 is a block diagram of a typical STB. The main areas of the STB that are involved with conditional access are shown in grey. The block labelled CA might be a dedicated, embedded CA module, or it might be one of the standard descramblers.

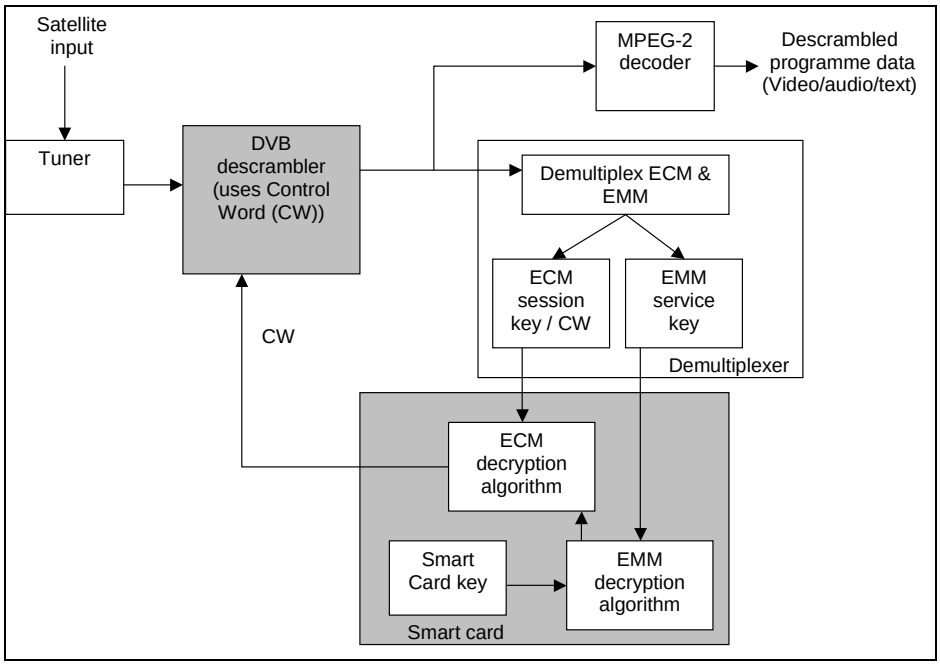


Figure 5: Conditional access in a typical set-top box

The tuner portion of the STB receives the incoming signal, demodulates it and sends the resulting data to the transport stream generator. This part of the STB reconstitutes the transport stream, which contains many packets of information. Each packet has associated with it (in its header) a Program Identifier (PID). All packets with PID value hex 1 have not been encrypted and are used by the demux processor to construct the Conditional Access Table (CAT). This table identifies all the PID values of the transport packets containing the EMMs. The demux processor also constructs the Program Map Table (PMT) from non-encrypted packets and gives the PID values of all the transport streams associated with a particular program. Private data associated with the program can also be included in this table. For example, the PID value of the packet containing the Entitlement Control Message (ECM). The data contained in these two messages (the EMM and the ECM) are vital in descrambling the encrypted programming material.

However, it should be noted that the standards do not specify the smart-card electronics or algorithms. Therefore, the system described here is a typical example. The EMM acquired by the demux processor is related to the authorization of services. It allows a particular set-top box, or a particular geographic region, to access services. It contains the encrypted service key. Typically, this key is changed every few months to discourage hackers.

The encrypted multi-session key, carried by the ECM, is related to particular programming material. This key, once decrypted, actually becomes the control word that is fed into the DVB descrambler, allowing the transport stream to be descrambled so that the viewer can see a particular program or view the programming material for a particular session. As figure 5 shows, the service key (EMM) is sent to the smart card, where it is decrypted with the help of the user key held inside the smart card. The descrambled service key is then used as the key to descramble the session key (ECM). This descrambling yields the Control Word (CW). It is this CW that is the key to the DVB transport-stream descrambler.

The main weakness of DVB-S CA is the one-way (broadcast) transmissions. Therefore it is very difficult to stop fraud and cloning pay TV smart cards without an efficient return channel and an efficient way to update smart card keys.

7.1.3 DVB-RCS security

As specified in [8], security is intended to protect the user identity including its exact location, the signalling traffic to and from the user, the data traffic to and from the user and the operator/user against use of the network without appropriate authority and subscription. Three levels of security can be applied to the different layers:

- DVB common scrambling in the forward link (could be required by the service provider);
- satellite interactive network individual user scrambling in the forward and return link;
- IP or higher layer security mechanisms (could be used by the service provider, the content provider).

Although the user/service provider could use its own security systems above the data link layer, it may be desirable to provide a security system at the data link layer so that the system is inherently secure on the satellite section without recourse to additional measures. Also, since the satellite interactive network forward link is based on the DVB/MPEG-TS Standard, the DVB common scrambling mechanism could be applied, but is not necessary (it would just add an additional protection to the entire control stream for non-subscribers). This concept is shown in figure 6.

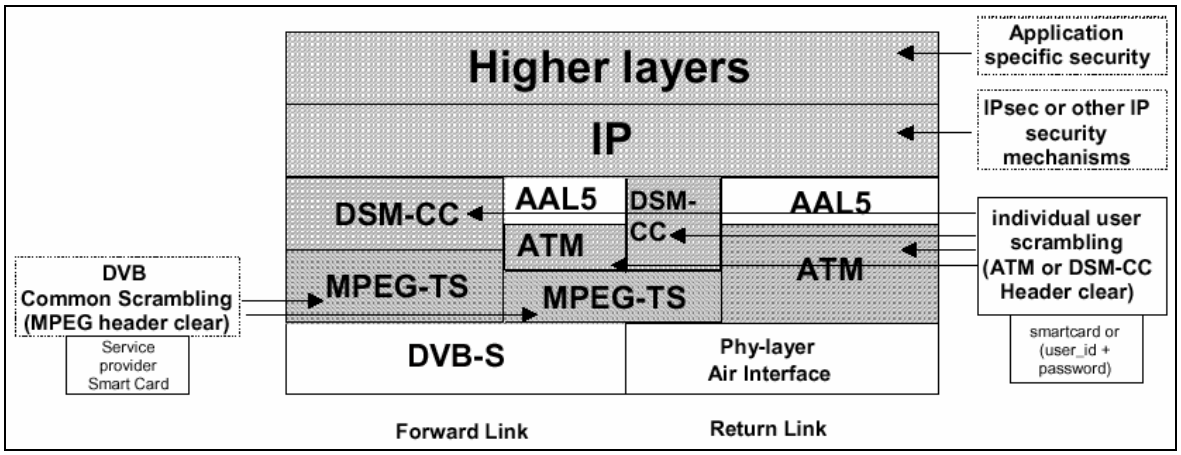


Figure 6: Security layers for satellite interactive network

In the following it is assumed there can be more than one user per Return Channel Satellite Terminal (RCST) and that such users will have security in their own right. The term RCST and ST (Satellite Terminal) have the same meaning in the present document. Security is thus defined at a level higher than the individual ST. On a user basis, an authentication algorithm may either check for user name and password on the client device or may use a Smart Card within the ST. All data and control to and from each user may be scrambled on an individual user basis. Each user may have a control word for the return and the forward link that does not allow anybody other than the NCC/Gateway or the user himself to descramble the data, except for lawful interceptors such as country authorities. An optional security mechanism derived from the one used in is also considered.

7.2 Network layer security

Security services can also be provided at the network layer. IPSec (RFCs 2401 [29], 2402 [30] and 2406 [31]) is a protocol that operates "above" IP and below layer 4 protocols such as TCP and UDP.

Network layer security has the following advantages:

- security is provided independently of upper layer protocols (whether TCP, UDP, RTP or reliable multicast);
- it can protect against network traffic re-routing and illegal changes to the network configuration;
- it can provide protection to all real time and non real time applications.

The disadvantages of network layer security are as follows:

- only the remote end network address (e.g. IP address) is authenticated;
- in the case of IPSec, applying security services at the IP layer can cause interworking problems with related protocols. Two examples are: Network Address Translations (NAT) can not be used (since IP addresses cannot be changed en route); and PEPs (RFC 3135 [46]) used to enhance performance on links such as mobile and satellite will fail, since the datagram contents (e.g. a TCP segment) are encrypted.

7.2.1 Internet Security (IPSec)

The security architecture of the Internet Protocol known as IP Security (IPSec) is the most advanced effort in the standardization of Internet security. The IPSec protocol suite is used to provide inter-operable cryptographically based security services (i.e. confidentiality, authentication, integrity, and non-repudiation) at the IP layer. It is composed of an authentication protocol: Authentication Header (AH), a confidentiality protocol: Encapsulated Security Payload (ESP) and it also includes an Internet Security Association Establishment and Key Management Protocol (ISAKMP). These security protocols are designed for both IP version 4 (IPv4) and IP version 6 (IPv6) environments.

As shown in figure 7, the IP Authentication Header (AH) provides connectionless integrity and data origin authentication for IP datagrams. It can also provide protection against replays. The authentication header may be used, alone or in combination, with the ESP. AH authenticates slightly more information in the IP datagram than does the ESP authentication (the IP datagram header is not included in the computation of the cryptographic integrity checksum of ESP). The authentication header protocol has two modes: transport or tunnel.

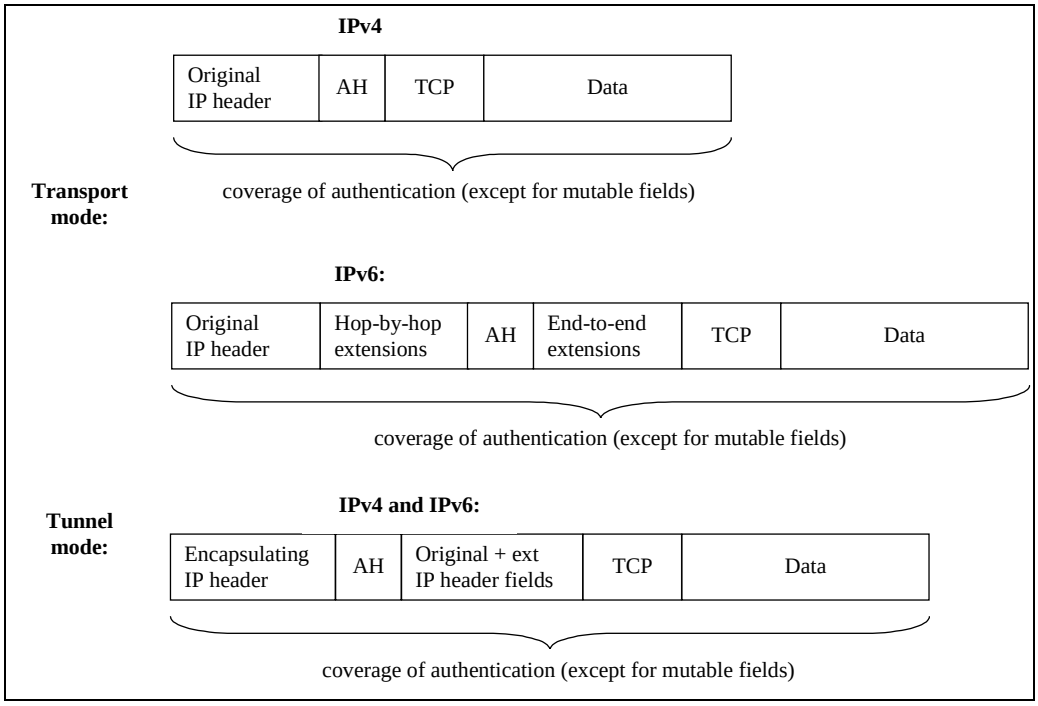


Figure 7: Authentication Header (AH) in transport and tunnel modes

Transport mode is used only in host-to-host authentication while tunnel mode can be used between two hosts, a host-to-gateway and gateway-to-gateway. The tunnel allows the host to delegate the security service to the gateway. This is especially interesting for companies with two private distant networks connected through the public Internet. In this mode, the IP header of the host/gateway responsible for computing/checking the AH is added while the old IP header is kept in the new IP datagram and moved after the AH.

The authentication header contains several fields to identify the authentication service being provided plus the cryptographic checksum. The header field "Next" identifies the next payload following the AH header (IANA IP protocol numbers). The "Length" field gives the length of the whole AH header - 2 (in 32 bits unit). The Security Parameter Index (SPI) field identifies the security association for this datagram (unique value for a given IP destination. SPI and destination uniquely identifies a security association). Finally, the sequence number is an optional field. It is included only if the anti-replay service is selected. The AH does not protect mutable fields of IP datagrams (e.g. record route, timestamp, loose source routing and strict source routing options).

As shown in figure 8, the Encapsulating Security Payload (ESP) header provides a mix of security services: data confidentiality, data origin authentication, connectionless integrity, anti-replay, and a limited traffic flow confidentiality. The set of services depends on the options selected during security association establishment. ESP may be used alone or in combination with AH. It is designed to work in transport mode or in tunnel mode.

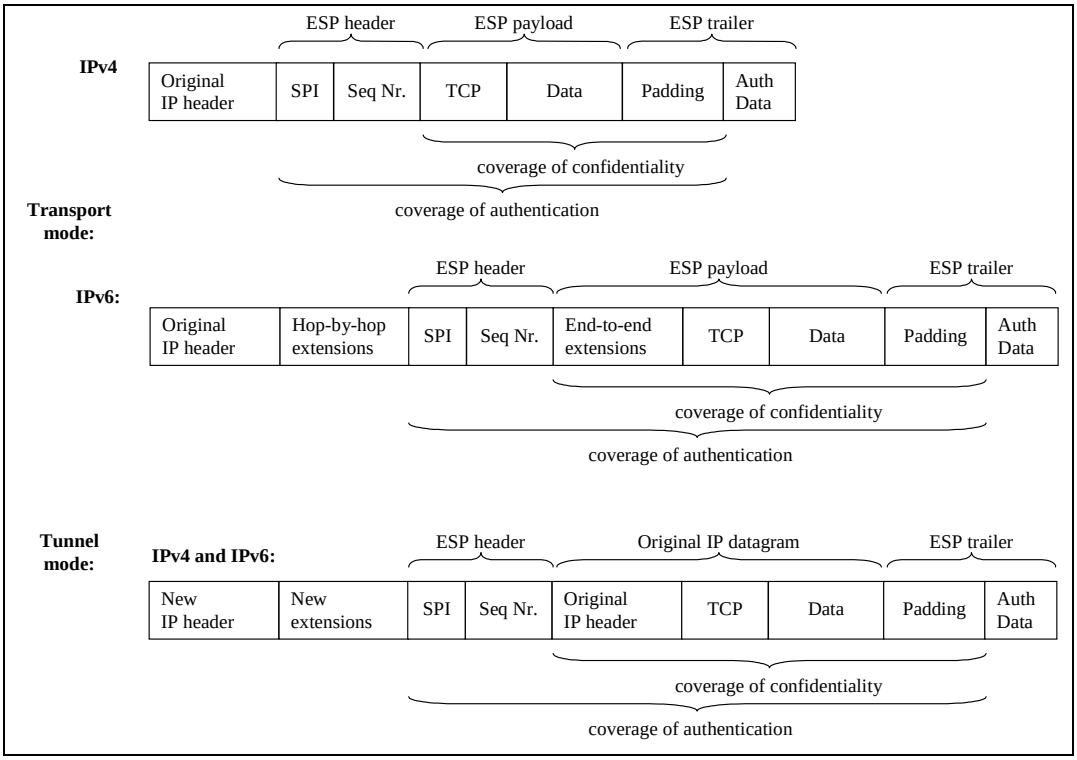


Figure 8: Encapsulated Security Payload (ESP) in transport and tunnel modes

7.3 Transport layer security

Security services may alternatively be provided at the transport layer. Examples of this are TLS (RFC 2246 [45]), or a reliable multicast protocol that includes security services. A protocol such as TLS assumes a reliable transport protocol such as TCP, and therefore effectively operates "above" layer 4 in the ISO protocol stack. Transport layer security that is embedded in the transport layer, such as a reliable multicast protocol has the following advantages:

- in the case of Unix-based environments, security is implemented in the user space rather than in the kernel, simplifying configuration of hosts;
- keys can be common for each host, simplifying key management.

Corresponding disadvantages are as follows:

- the endpoint IP host addresses are known and therefore susceptible to traffic analysis;
- in comparison to the operation of TLS with TCP, there is no generic security system for unreliable transport protocols such as UDP, which is widely used to carry multicast and real time traffic.

7.3.1 Transport Layer Security (TLS)

The TLS protocol provides communications privacy over the Internet. The protocol allows client/server applications to communicate in a way that is designed to prevent eavesdropping, tampering, or message forgery. The protocol is composed of two layers: the TLS Record Protocol and the TLS Handshake Protocol. At the lowest level, layered on top of some reliable transport protocol (e.g. TCP), is the TLS Record Protocol. The TLS Record Protocol provides connection security that has two basic properties:

- the connection is private. Symmetric cryptography is used for data encryption (e.g. DES (Digital Encryption Standard, AES (Advanced Encryption Standard) etc.). The keys for this symmetric encryption are generated uniquely for each connection and are based on a secret negotiated by another protocol (such as the TLS Handshake Protocol). The Record Protocol can also be used without encryption;
- the connection is reliable. Message transport includes a message integrity check using a keyed MAC. Secure hash functions (e.g. SHA, MD5, etc.) are used for MAC computations. The Record Protocol can operate without a MAC, but is generally only used in this mode while another protocol is using the Record Protocol as a transport for negotiating security parameters.

The TLS Record Protocol is used for encapsulation of various higher level protocols. One such encapsulated protocol, the TLS Handshake Protocol, allows the server and client to authenticate each other and to negotiate an encryption algorithm and cryptographic keys before the application protocol transmits or receives its first byte of data.

The TLS Handshake Protocol provides connection security that has three basic properties:

- the peer's identity can be authenticated using asymmetric, or public key, cryptography (e.g. RSA etc.). This authentication can be made optional, but is generally required for at least one of the peers;
- the negotiation of a shared secret is secure: the negotiated secret is unavailable to eavesdroppers, and for any authenticated connection the secret cannot be obtained, even by an attacker who can place himself in the middle of the connection;
- the negotiation is reliable: no attacker can modify the negotiation communication without being detected by the parties to the communication.

7.3.2 Secure Real Time Transport Protocol (SRTP)

SRTP is a profile of the Real Time Transport Protocol (RTP), which can provide confidentiality, message authentication, and replay protection to the RTP/RTCP (control) traffic.

SRTP can achieve high throughput and low packet expansion. SRTP proves to be a suitable protection for heterogeneous environments, i.e. environments including both wired and wireless links. To get such features, default transforms are described, based on an additive stream cipher for encryption, a keyed-hash based function for message authentication, and an "implicit" index for sequencing/synchronization based on the RTP sequence number for SRTP and an index number for Secure RTCP (SRTCP). The main security features of SRTP are to ensure:

- the confidentiality of the RTP and RTCP payloads; and
- the integrity of the entire RTP and RTCP packets, together with protection against replayed packets.

Other functional goals for the protocol are:

- a framework that permits upgrading with new cryptographic transforms;
- low bandwidth cost, i.e. a framework preserving RTP header compression efficiency;
- low computational cost;

- small code size and data memory for keying information and replay lists;
- limited packet expansion to support the bandwidth economy goal;
- independence from the underlying transport, network, and physical layers used by RTP, in particular high tolerance to packet loss and re-ordering, and robustness to transmission bit-errors in the encrypted payload;
- in addition, SRTP provides for some additional features. They have been introduced to lighten the burden on key management and to further increase security. They include:
  - a single so-called master key provides keying material for confidentiality and integrity protection, both for the SRTP stream and the corresponding SRTCP stream. This is achieved due to a key derivation function, providing so-called session keys for the respective security primitive, securely derived from the master key;
  - in addition, the key derivation can be configured to periodically "refresh" the session keys, which limits the amount of ciphertext produced by a fixed key, available for an adversary to cryptanalise;
  - so-called salting keys are used to protect against off-line pre- computation attacks.

These properties ensure that SRTP is a suitable protection scheme for RTP/RTCP in both wired and wireless scenarios.

7.4 Application layer security

In principle, the security system should be as close as possible to the end user or entity and therefore application level security can provide a good solution. In application layer security, the security services are provided within each application, and are embedded within application code. Application layer security has the following advantages:

- the security services are independent of the underlying protocols;
- the security services provide a level of assurance that is independent of the ownership of the underlying networks (for example, public Internet, VPNs, other departments in a corporation);
- data is not compromised if it is incorrectly delivered to the wrong host or application.

However, application layer security has the following disadvantages:

- security has to be individually built into each application, increasing software development and test timescales, with potentially reduced levels of software assurance;
- keys are consequently separate for each application, again with duplication of effort in key management;
- traffic analysis can be easily performed by a potential adversary: the endpoint addresses (e.g. TCP port and IP host address) are visible in clear text. Consequently, an adversary knows who is communicating, even if they can not determine what is being said;
- denial of service attacks are possible, where an active attacker injects a large number of rogue packets which the application level security system will check and reject, consuming a large amount of CPU time at the end system.

7.4.1 Security for eXtensible Markup Language (XML)

Over the last two years, the eXtensible Markup Language (XML) has rapidly emerged as the standard for electronic data exchange in business applications. In parallel, Public Key Infrastructure (PKI) and digital certificates have continued their expanding adoption by net marketplaces, Internet merchants, and suppliers as the de facto strong foundation for authenticating users, Web sites, and business partners.

As XML has gained momentum as the preferred format for exchanging business information on the Web, the need for standard mechanisms for applications to provide entity authentication and privacy services for XML documents has grown. Today's marketplaces are therefore eager for XML and PKI to work together in fulfilling the widely held expectations for cryptographically secure, XML-coupled business applications. All e-commerce applications require trust and security, making it mission-critical to devise common XML mechanisms for authenticating merchants, buyers, and suppliers to each other, and for digitally signing and encrypting XML documents like contracts and payment transactions.

In addition, a joint IETF-W3C working group has just completed a proposed standard for XML Digital Signatures, but its charter expressly limits its work to that specific area. Recently, a new working group within the W3C has also been chartered to develop standards for XML Encryption.

7.4.2 Digital Rights Management (DRM)

File transfer is currently the largest reported consumer of Internet capacity, in the form of peer to peer connections for sharing of music files. The music publishing industry appears to be reluctant to commit to multicast for music distribution, because of the potential for files immediately to be "shared" with other users. These are considerations for the developing Digital Rights Management (DRM) industry [9], to enable copyright owners to realize revenue from each transfer, whether via central distribution or peer to peer. The BSM broadcast model for the Internet should enable control to revert to the rights owners, so that peer to peer connection can be minimized.

DRM is a means of encrypting files before transmission and user local storage, so that the files can be decrypted by the holder of a valid key under defined commercial conditions. It therefore comprises both an encryption/decryption process, and a process of key management linked to subscription management.

The main purpose of DRM is to enforce purchase of content in the mass market, by enabling legal purchase, and making illegal copying or forwarding uneconomic or unsafe. In the realm of content owners, distributors and carriers it is expected that legal processes alone will be used to protect material at source.

Each file is encrypted with a key. The key, that enables users to decrypt, is provided in a licence package, implemented in software, which includes the conditions of use, for example the dates of validity, the number of times it may be played, or the feature set of a computer program. The conditions can extend to stipulation of a process whereby, if a file is copied or forwarded, a fee can be collected and transferred to the rights owner. This could be a means to legitimise peer to peer forwarding and may be useful for cases where material is not available directly from a distributor.

7.4.2.1 Open Mobile Alliance (OMA) DRM

The objectives of Open Mobile Alliance (OMA) DRM is to enable the controlled consumption of digital media objects allowing content providers to express usage rights, e.g. the ability to preview DRM content, to prevent downloaded DRM content from being illegally forwarded (copied) to other users. OMA membership includes major software and mobile network providers such as Microsoft, IBM, Nokia and Ericsson.

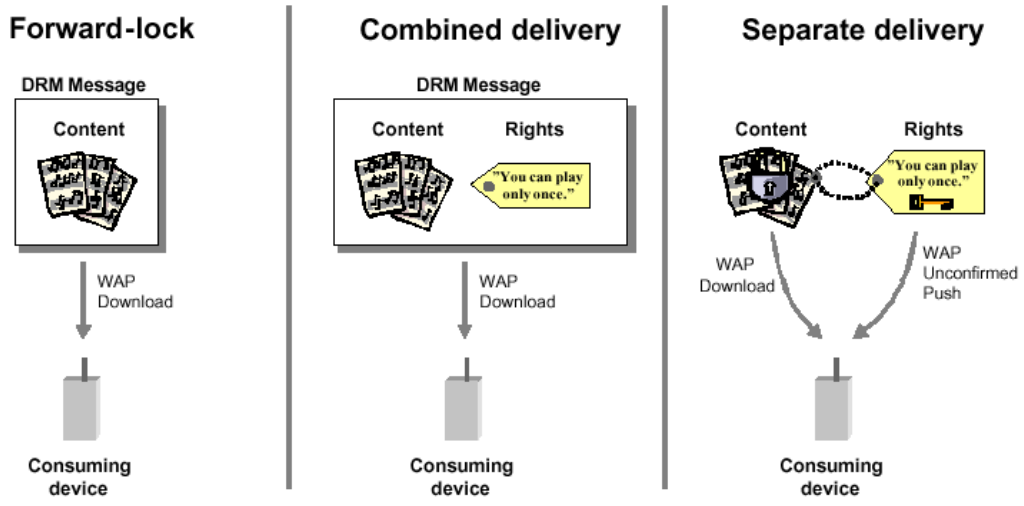


Figure 9: OMA DRM delivery methods

The rights can be delivered to the consuming device by downloading them together with the content or by sending the rights object separately from content (see figure 9). The former case (combined delivery) is simpler whereas the latter case (separate delivery) provides more security by making it more difficult to steal the content.

For forward lock and combined delivery content provider needs to package content, optionally with a rights object, into a DRM message. That message may be delivered to the device using e.g. the OMA Download mechanism. In the separate delivery method the content provider needs to convert the plaintext media object into DRM Content Format (DCF) defined in the "DRM Content Format" specification. This conversion includes symmetric encryption of the content making the DRM protected content object useless to parties not having access to the content encryption key. Thus content in DRM format may be distributed via an insecure transport whereas a more secure transport (from DRM point of view) is used to deliver the rights object with the encryption key.

For handling streaming media, the OMA DRM mechanisms can be used to indirectly control media objects via control of the meta-data. For example, it can be a Session Description Protocol (SDP) record as a description of a media streaming session. In order to allow for the same security level for streams as for downloaded objects, it is recommended that the streaming player is not allowed to store the media streams. Furthermore, the real-time media streams should be protected using a robust stream encryption mechanism suitable for a wireless environment, such as for example the Secure Real-Time Protocol (SRTP), described in clause 7.3.2.

7.4.3 Secure Shell (SSH)

Secure Shell is developed by SSH Communications Security http://www.ssh.com, which has a range of features including protection of all passwords and data, fully integrated secure file transfer and file copying, and automatic authentication of users. It is a program to log into another computer over a network, to execute commands in a remote machine, and to move files from one machine to another. It provides strong authentication and secure communications over unsecured channels. It is intended as a replacement for rlogin, rsh, and rcp.

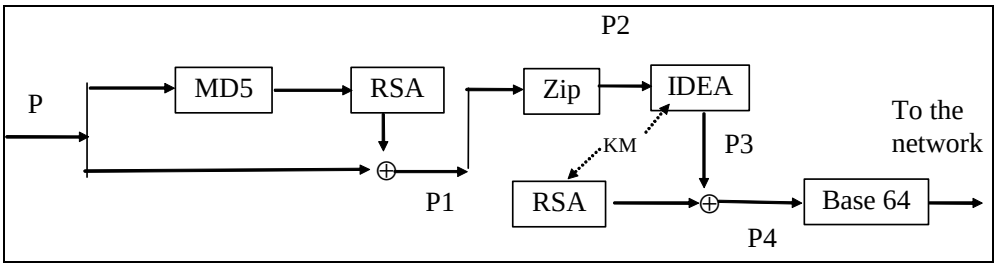
SSH is a protocol for secure remote login and other secure network services over an insecure network. The Internet Draft describes the architecture of the SSH protocol, as well as the notation and terminology used in SSH protocol documents. It also discusses the SSH algorithm naming system that allows local extensions. The SSH protocol consists of three major components: The Transport Layer Protocol provides server authentication, confidentiality, and integrity with perfect forward secrecy. The User Authentication Protocol authenticates the client to the server. The Connection Protocol multiplexes the encrypted tunnel into several logical channels.

7.4.4 Pretty Good Privacy (PGP)

Pretty Good Privacy (PGP) is a complete email security package that provides privacy, authentication, digital signatures and compression. PGP is based on RSA and MD5 public key system and IDEA secret key system.

As shown in figure 10, here Alice wants to send plaintext message P to Bob in a secure way. Assuming that both Alice and Bob know each other's public keys, Alice first hashes her message P using MD5 and then encrypts the resulting hash with her private RSA key. The encrypted hash and the original message are now concatenated into message P1 and compressed into message P2 using the ZIP program. Next PGP prompts Alice for a random input to generate a 128-bit IDEA key KM. The session key KM is used to encrypt P2, and KM is now encrypted using Bob's public key. These two components are concatenated and converted to base64 to make it compatible with RFC 822 [47] and MIME.

When Bob gets the message, he reverses the base64 encoding and decrypts the IDEA key using his private RSA key. Using this key, he decrypts the message to get P2. After decompressing it, Bob decrypts the hash using Alice's public key. If the hash agrees with his own MD5, he knows that P is the correct message from Alice.



P: Plaintext
P1: P + signed hash of P (using source's RSA private key)
P2: Ziped P1
P3: Encrypted P2 (with IDEA"s secret key KM)
P4: P3 + secret key KM for IDEA (encrypted with destination's RSA public key)

Figure 10: PGP email security system

RFC 2015 [48] describes how Pretty Good Privacy (PGP) can be used to secure emails in the Multipurpose Internet Mail Extensions (MIME) format.

7.4.5 Tailor made security for satellite applications

It is possible to develop secure and reliable transfer applications that take the satellite link characteristics into account. One example is the secure Satellite Reliable Multicast Transport Protocol (SAT-RMTP) that was developed as part of in GEOCAST project demonstrator (A European IST project). SAT-RMTP provides network delivery of the content over a multicast enabled network. This transfer protocol has been developed to mitigate the demands on the server/network, and to tolerate packet loss due both to congestion at network bottlenecks, introduced by receiver performance bottlenecks, and from link impairments (fading).

Encoded multimedia clips are stored at the content server. Using the content tool, these files are encrypted and session information is created. An MPEG-1, MPEG-2 or QuickTime format can be used. At the client, each user uses a client session tool to select the required multimedia content. A session protocol (Session Directory Revised (SDR)) is used, which has a menu-driven interface. The session protocol describes the objects (files) available for download.

The security protocol identifies individual end-users and associates security state (keys) with the clients. The encryption and decryption is provided at the application level, and when encryption is required the transfer protocol only distributes encrypted content.

Figure 11 shows the interaction between the session server and its clients. The application has been constructed in a way that allows the session server process and a number of transfer/security protocol instances to be located on a common server platform. These may alternatively be distributed among a number of server platforms.

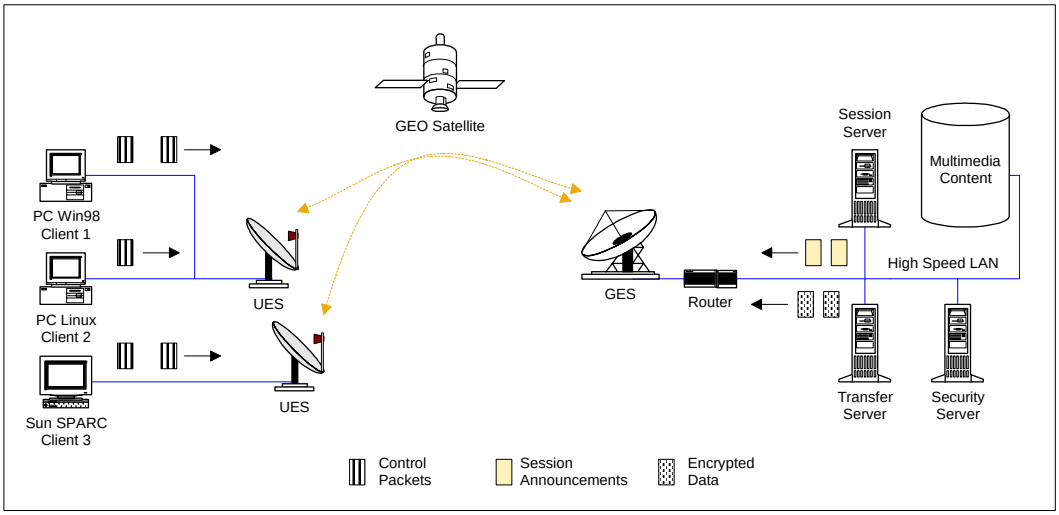


Figure 11: SATRMTP illustration of transfer and session protocols communicating

Regarding security, all clients will have a Secure Association (SA) with the security server. This SA may either be secret key based or public key based. In a secret key SA, each client shares a secret key with the server. In a public key SA each client has a certificate, which is stored at the server, and its own private key, which is stored locally. In the public key SA each client also stores the public certificate of the security server.

In the SATRMTP application, the security server maintains two databases: a user information base (containing client access permissions) and a security database (where keys are stored). A client/server security protocol is developed and implemented to provide authentication and access functionality, and securely distribute keys to clients (using the secure association. Following transfer of the file, the content decryption tool, on the client side, will decrypt each object. The tool also checks the digital signature of the sender to authenticate the sender. The decrypted file is then passed to a multimedia player. Following successful decryption the security client could send an acknowledgement to the security server to confirm successful decryption to the server.

7.5 End-to-end and satellite network security

End-to-end security may be provided at any level of the protocol stack such as application, transport or network layers, as presented in clauses 7.2 to 7.4. In general, there is a need to establish a trust relationship between users of the end-to-end security system through a security management system. The security operations may be visible to end users and applications if they are implemented at the application level, or it can be transparent if implemented in the lower layers.

In contrast, satellite network security focuses on access control and data encryption/integrity mechanisms within the BSM satellite network boundaries. Link layer security is the best solution here, as presented in clause 7.1. The satellite network can star and mesh configurations with regenerative or bent pipe satellites. DVB and ATM security procedures can be used to secure satellite links. IPSec can be used to provide satellite network security by implementing IPSec tunnels.

7.6 Security services in BSM protocol layers

Table 2 provides a summary of the major advantages and disadvantages of security in each layer of the BSM protocol stack.

Table 2: Security layers comparison

	Link layer	Network layer	Transport layer	Application layer
Major advantages	Complete control of satellite link security	IPSec is the best solution for Internet security	Widely used for securing TCP connections	Can satisfy applications requirement very well
Major disadvantages	Only the satellite hop is secure	IPSec works only for IP networks	No security for UDP and multicast	No transparency, where applications need modification to fit security

Also the security services that can be provided in layer of the BSM protocol stack are summarized as follows:

Table 3: Security services at various protocol layers

	Link layer	IP Network layer	Transport layer	Application layer
Satellite terminal authentication	Yes	Yes (IP address)	No	No
User terminal authentication	No	Yes (IP address)	No	No
User authentication	No	No	Yes	Yes
Satellite link privacy	Yes	Yes (IPSec IP tunnel)	No	No
End to end privacy	No	Yes	Yes	Yes
Satellite link data integrity	Yes	Yes (IPSec IP tunnel)	No	No
End to end data integrity	No	Yes	Yes	Yes

Examining table 3, show that implementing network layer security such as IPSec, provides the flexibility of closer integration with the Internet and satisfy the requirement of some multimedia services for satellite and/or end to end security.

8 Security management survey

Any security system that requires two parties to have the same key material needs a key management system to ensure that the two parties are agreed on which encryption algorithms they are going to use and which encryption keys are needed. In all cases, the most difficult security problem is the security management and key distribution. Five key management examples are given in the following clauses: DVB-S conditional access, DVB-RCS security, IP unicast security, access control using firewalls and IP multicast security management.

8.1 DVB-S Conditional Access Key Management

See clause 7.1.2 for details.

8.2 DVB-RCS Key Exchange Protocols

Main Key Exchange

Main Key Exchange (MKE) uses Diffie-Hellman to develop a shared secret between the NCC and ST, which is independent of the cookie value. Furthermore, it uses the cookie value to authenticate the ST to the NCC. It optionally uses the newly developed shared secret to update the cookie value. Finally, it derives a shared secret key used for the security context that is used to process payload stream data.

The exchange is initiated by the NCC sending a message containing the Diffie-Hellman values, m, g, X, and a random nonce string, nonce1. The ST responds with a message containing its Diffie-Hellman value, Y, a random nonce string, nonce2, and an authentication string, auth. The NCC and ST each use the same formula to calculate the authentication string:

auth = H (cookie, nonce1 ~ nonce2)

Which is communicated by the ST and checked by the NCC. This proves the identity of the ST, since it requires knowledge of the cookie to calculate the correct value of auth. The ST and NCC each use the Diffie-Hellman values to arrive at the same secret value, s:

S = g^(x × y)modm

This unsigned integer value is encoded as a byte string, of length specified by the Diffie-Hellman parameter size, using big-endian byte ordering. It is then used to calculate a temporary shared secret string, temp:

temp = H (encode (s), nonce2 ~ nonce1)

If the cookie is to be updated, the new value is computed in sections for n = 1, 2, …:

newcookie(n) = H (temp ~ (unsigned char)1 ~ (unsigned char)n, "")

These string values are computed and concatenated until the total length matches or exceeds the length of the cookie. The cookie is then obtained by taking the first 20 bytes out of the concatenated sections, starting from the beginning. The session key used for payload stream encryption is likewise computed in sections:

key(n) = H (temp ~ (unsigned char) 2 ~ (unsigned char) n, "")

Where, again, a sufficient number of sections are calculated to produce enough bytes to cover the length of the key. The session key is obtained "in the same manner as the cookie" by taking the required number of bytes out of the concatenated sections, starting from the beginning.

Quick Key Exchange

Quick Key Exchange (QKE) uses the existing cookie value to authenticate the ST to the NCC, and then derive a shared secret key used for the security context that is used to process payload stream data. The exchange is initiated by the NCC sending a message containing a random nonce string, nonce1.

The ST responds with a message containing a random nonce string, nonce2, and an authentication value, auth. The value of auth is calculated in the same way as for Main Key Exchange, and can be used to verify the identity of the ST. The ST and NCC then each calculate a temporary shared secret string, temp:

temp = H (cookie ~ (unsigned char)3 , nonce2 ~ nonce1)

This value is used to produce the payload encryption key in the same way as for Main Key Exchange.

Explicit Key Exchange

Explicit Key Exchange (EKE) is used by the NCC to deliver a pre-determined session key to the ST. The session key is encrypted under a temporary key derived from the cookie value, and is used for the security context that is used to process payload stream data.

The delivery is performed by the NCC sending a message containing a random nonce string, nonce1, and a byte string value, encryptedkey, which has the same length as a key used for payload encryption. The ST responds with a message containing a random nonce string, nonce2, and an authentication value, auth. The value of auth is calculated in the same way as for Main Key Exchange, and can be used to verify the identity of the ST. Both the NCC and ST calculate a temporary shared secret string, temp:

temp = H (cookie ~ (unsigned char)4, nonce1)

Which is used to produce sections of a temporary key, in the same way as for Main Key Exchange. The NCC uses these temporary key string sections to XOR with the session key to obtain the encryptedkey value, and the ST performs a second XOR operation to decrypt the session key value.

8.3 IPSec management

IP Security (IPSec) provides confidentiality, data integrity, access control, and data source authentication to IP datagrams. These services are provided by maintaining shared state between the source and the sink of an IP datagram. This state defines, among other things, the specific services provided to the datagram, which cryptographic algorithms will be used to provide the services, and the keys used as input to the cryptographic algorithms. Establishing this shared state in a manual fashion does not scale well. Therefore a protocol to establish this state dynamically is needed.

For unicast traffic the IPSec Key Management Protocol is called the Internet Key Exchange (IKE). IKE has been developed to ensure that when two parties wish to communicate securely they can agree all the necessary information in a secure manner even if they have not communicated before. One building block is the Internet Security Association and Key Management Protocol (ISAKMP) (RFC 2408 [32]). ISAKMP defines procedures and packet formats to establish, negotiate, modify and delete Security Association (SA). It is not bound to any specific cryptographic algorithm, key generation technique or key exchange technique. ISAKMP is designed to be key exchange independent and can support several key exchange protocols. The Internet Key Exchange (IKE) Internet draft describes a specific key exchange protocol.

IKE is the protocol, which performs mutual authentication and establishes security associations (SAs) for IPSec. The base protocol of the first version of IKE was documented in RFCs 2407 [49], 2408 [32] and 2409 [33]. The purpose of IKE is to negotiate and provide authenticated keying material for, security associations in a protected manner. It can be used for negotiating Virtual Private Networks (VPNs) and also for providing a remote user from a remote site (whose IP address need not be known beforehand) access to a secure host or network. Client negotiation is supported. Client mode is where the negotiating parties are not the endpoints for which security association negotiation is taking place. When used in client mode, the identities of the end parties remain hidden.

Furthermore IKE incorporates a mechanism to counter denial of service attacks in which servers are flooded with bogus request messages. The goal of the attacker perpetrating these attacks is to keep a server busy with the verification of a large number of bogus requests in order to cause abnormal CPU usage and consequently degrade the service provided by the server to legitimate users. The IKE mechanism to prevent such denial of service attacks is based on the anti-clogging technique. The principle of anti-clogging is to perform the exchange of a pair of "cookies" at the beginning of each client-server connection before initiating any resource-intensive verification. The initial exchange provides a weak authentication and allows for the verification of the client's presence at the claimed IP address thus thwarting all flooding attempts using bogus IP addresses from a single host. The computation of the cookie by the server is based on a simple hash function requiring low CPU usage, in comparison with CPU-intensive strong authentication and key generation operations and no resource reservation takes place before the completion of the successful cookie exchange.

IKE is sometimes called IKEv1 and its implementations incorporated additional functionality including complex features for Network Address Translation (NAT) traversal, legacy authentication, and remote address acquisition, which were not documented in the base documents. A new version of IKE has been finalized by the IPSec WG in the IETF. The goal of the IKEv2 specification is to specify all that functionality in a single document, as well as simplify and improve the protocol, and fix various problems in IKEv1 that had been found through deployment or analysis. IKEv2 preserves most of the features of the original IKE, including identity hiding, perfect forward secrecy, two phases, and cryptographic negotiation, while greatly redesigning the protocol for efficiency, security, robustness, and flexibility.

8.4 IP multicast security

In BSM architecture, IP multicast will be an efficient way to distribute data from a server to a group of clients. IP multicast is an Internet protocol that enables transmission of data packets to a group of receivers. IP multicast makes efficient use of bandwidth by setting up a mid-point between unicast traffic (one-to-one) and broadcast IP traffic (one-to-all in a network). This is well suited for one-to-many or many-to-many bulk data transfer or multimedia (audio/video) streaming transmission to a large number of heterogeneous receivers. IP multicast efficiently supports this type of transmission by enabling sources to transmit a single copy of a message to a group of interested receivers.

The anonymous-receiver model underlying IP Multicast is attractive precisely because the distribution tree is easily extendible, subject to the resources available to the multicast routing protocol. Any host in a subnet can join a multicast group without its subnet router passing identification information about the host to other routers upstream in the distribution tree. This allows IP Multicast to scale to a large number of participating hosts. The extendibility of the distribution tree in IP Multicast makes the IP Multicast model very attractive from the perspective of scalability.

However, from the perspective of security, additional mechanisms and services must be built atop the basic IP Multicast model. This decoupling of security from the IP Multicast model is advantageous, since it allows differing security models and architectures to be deployed, without affecting the multicast distribution tree which delivers the multicast data end-to-end. This decoupling is also important from the application's perspective, since each application requires different forms of host information and other security parameters, and may deploy differing user-identification and user-authentication mechanisms. In the Functional BSM Multicast Architectures document [10], BSM multicast group management and multicast routing is specified. From the security prospective, the group management and routing should be transparent to security. As such, the BSM security model should work with any group management or routing protocols.

As shown in figure 12, there are several interrelated factors or aspects of IP Multicast that influence the approaches and mechanisms used to secure it. Of these, some broad and most relevant factors include:

- multicast application type;
- group dynamics;
- salability issues;
- underlying trust model.

The combinations of these factors will be the deciding factor on the group security policy parameters such as:

- procedures for establishing, running and terminating the group;
- key distribution methods. For example flat key or LKH distribution system can be chosen to fit the requirements of that group. For one-to many applications with large and dynamic groups, the rekeying policy must be defined carefully in order to reduce the bandwidth need to key management traffic;

- the strength of the encryption algorithms and digital signature needed, such as using DES, triple DES or AES for privacy and RSA or DSS signature systems;
- what to do when things go wrong: Such as network failure, DoS attacks, problems with entity authentication and authorization. The procedures are usually defined in the security policy. There must be a clear way for creation, dissemination and enforcement of security policies.

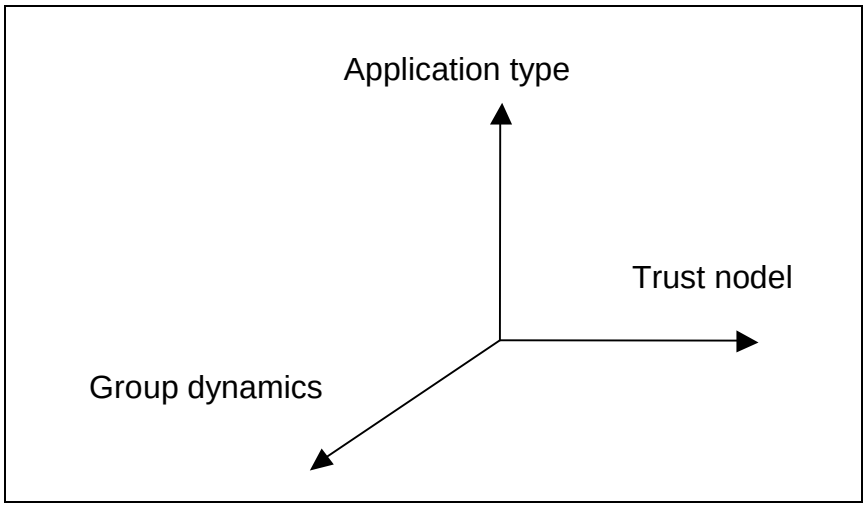


Figure 12: Factors affecting secure multicast system design

It is expected that the BSM multicast groups can be large and very dynamic. Therefore it is essential to examine various methods to manage security for such large groups. In following clauses, present two concepts: one is efficient architecture for security key distribution (dissemination) in large and dynamic groups, the other concept is security management systems that define the rules for establishing secure groups, controlling members join and leave and terminating the group securely.

8.4.1 Scalable key distribution architecture

Logical Key Hierarchy (LKH) is a mechanism for security key management within a group of entities, providing the ability to initialize the group with a common key and then to rekey the group as required (RFC 2627 [34]). It is thus of particular application in secure multicast communications. In the IETF, this architecture is the preferred way for key dissemination to large groups. LKH can be used with any of the security management protocols mention in clause 8.5.2.

LKH requires two types of entity as shown in figure 13: a Group Controller (GC) and one or more Group Members (GMs). The former is responsible for creating and distributing keys and rekeying (to maintain security) as appropriate; the group members are entities with access to the group keys. To support LKH, the GC communicates with each GM, but GMs do not need to communicate with each other. In a secure multicast communication it is not necessary for the GC to be co-located with any multicast data source. LKH provides the following advantages:

- scalable: the resources required to manage keys within a group grow more slowly than the group size, *N*; these resources include network transmission requirements, GC storage, GM storage, GC encryption effort, and GM decryption effort;
- collusion-proof: no set of entities together can obtain any key unless one or more of them could have obtained it individually.

The benefit of LKH is particularly apparent when a group needs to be rekeyed.

LKH does not specify mechanisms for transmitting keys between a group controller and group members: this is the function of a group key management protocol such as GSAKMP. LKH is independent of any particular encryption or decryption algorithms.

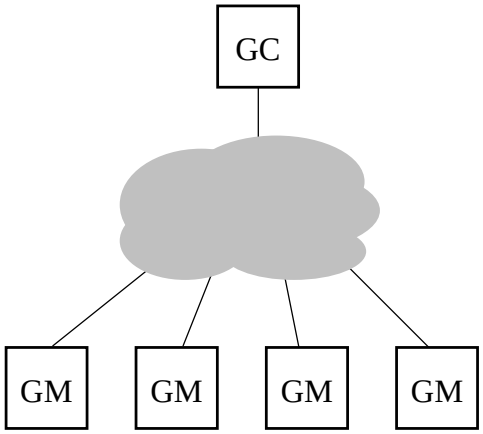


Figure 13: LKH entity logical structure

Each GM is assumed to have an initial pairwise secure association with the GC, that is to say the GM and the GC share a secret key known only to the two entities. The mechanism by which this secure association occurs is outside the scope of LKH, but typically it could involve using a technique such as Diffie-Hellman to create a shared secret known only to the two parties; or a pre-shared secret; or a secret exchange using a public key system.

We introduce LKH by considering a simple flat key management system that can be used to share a single key, "A", so that it is known to the GC and all GMs but to no other entities. The flat key management system consists of *N* pairwise keys each shared between the group controller and one of the *N* group members (see figure 14). Each of these pairwise secure associations are represented by a circle and the group key is represented by the box labelled "A". If the group key is changed the new group key has to be encrypted with each user's unique pairwise key and then unicast to that user; each of these encrypted keys is represented by one of the lines drawn in figure 14. Thus for *N* users a total of *N* encrypted keys are generated and transmitted across a network.

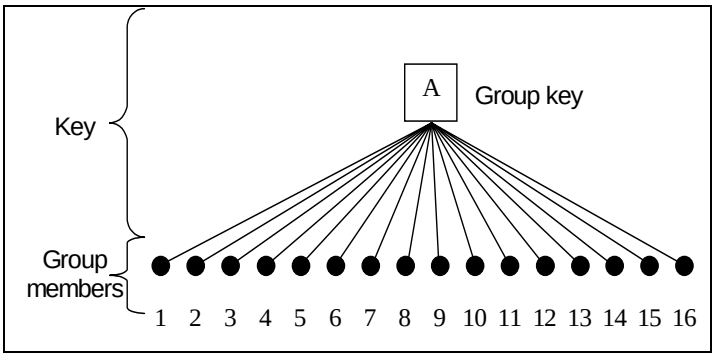


Figure 14: Key hierarchy: *N* pairwise keys

We contrast this with LKH, where a tree of keys is used to share a single key "O" so that it is known to the GC and all GMs but to no other entities. In figure 15 the keys are labelled A through O, the circles again represent the pairwise keys, and the lines each represent encrypted keys sent across the network, as we shall now see. Suppose now that User 11 needs to be deleted from the multicast group. Then all of the keys held by User 11 (keys F, K, N, O) must be changed and distributed to the users who need them, without permitting User 11 to obtain them or anyone else, who is not entitled to them. To do this, we must replace the keys held by User 11, proceeding from the bottom up.

The server chooses a new key for the lowest node (not the leaf, for which a unicast secure association exists between the GC and the GM), and then transmits it encrypted with the appropriate daughter keys. Thus for this example, the first key replaced is Key F, and this new key will be sent encrypted with User 12's unique pairwise key. The second key replaced is Key K, which is sent encrypted with the newly replaced Key F (for User 12) and also sent encrypted with key E (for Users 9 and 10). Key N is then sent encrypted in the newly replaced Key K (for Users 9, 10, and 12) and also encrypted in key L (shared by Users 13 through 16). Finally, Key O is replaced, and this new key is sent encrypted in the newly replaced Key N (for Users 9, 10, and 12 through 16) and also separately is encrypted in key M (shared by Users 1 to 8). Since we are proceeding from the bottom up, each of the replacement keys will have been replaced before it is used to encrypt another key.

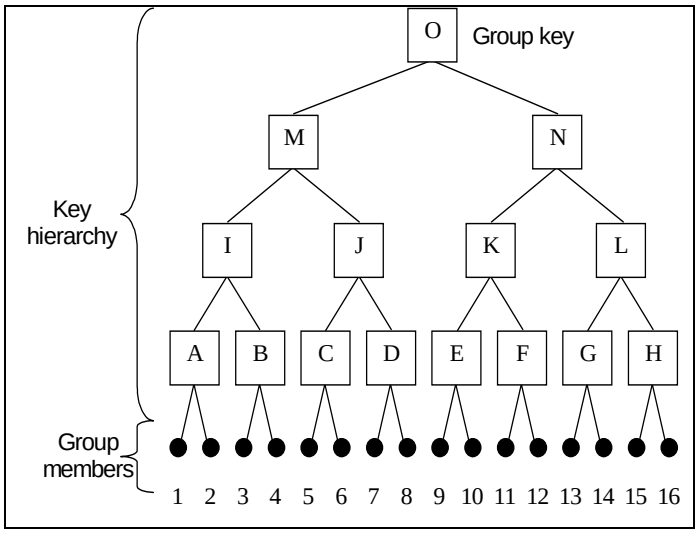


Figure 15: Logical key hierarchy

The seven keys sent represent a significant saving on the 16 keys that would need to be transmitted using the flat key system of figure 14. We briefly write these keys as $\{F\}_{12}$ $\{K\}_{E}$ $\{K\}_{F}$ $\{N\}_{K}$ $\{N\}_{L}$ $\{O\}_{M}$ $\{O\}_{N}$. In general, the number of transmissions required is the sum of the degrees of the replaced nodes. In a *k*-ary tree of depth *d*, this is a total of $kd - 1 = k \log_k N - 1$ transmissions.

The Group Traffic Encrypting Key (GTEK), used to encrypt data traffic, may, depending on the group security policy, either be key O (see figure 15), or it may be separately encrypted using key O and transmitted to all group members.

8.4.2 Multicast key management protocols

8.4.2.1 Group Secure Association Key Management Protocol (GSAKMP)

GSAKMP (draft-ietf-msec-gsakmp-sec-04.txt) is a three message multicast key management protocol that does not require an underlying secure unicast security association, it is assumed that a non-secure mechanism is used to transmit non-sensitive information about the security mechanisms to be used for group establishment.

The GSAKMP protocol includes mechanisms for group policy dissemination, group key dissemination, and group rekey operation. The transmission of a fully specified policy token to all joining group members is what allows GSAKMP to support distributed architectures and multiple data sources within a single cryptographic group. Distributed architectures are supported because the policy token allows rule-based allocation of Group Security Association actions to network resources. Multiple data sources are supported because the inclusion of a policy token and policy payloads allow group members to review the group access control and authorization parameters.

In a GSAKMP group there are two types of entity a Group Member and a Group Controller these are defined below.

Group Member: A Group Member (GM) is any entity with access to the group keys. Regardless of how a member becomes a part of the group or how the group is structured, GMs will perform the following actions:

- validate the authorizations for security relevant actions;
- accept group keys from the Group Controller;
- request group keys from the Group Controller;
- maintain local Certificate Revocation Lists (CRLs);
- enforce the cooperative group policies as stated in the group policy token;
- perform peer review of key management actions; and
- manage their local key.

Group Controller: The Group Controller (GC) is a group member with authority to perform any critical protocol actions including:

- creating and distributing keys;
- maintain the Rekey infrastructure;
- building and maintaining the Rekey arrays.

The sequence of events for GSAKMP is straightforward. The sequence is:

- security suite definition is transmitted outside the protocol;
- group establishment phase which consists of:
  - Request To Join (RTJ);
  - key download;
  - acknowledgement;
- group Security Association (SA) up and running;
- group management.

The announcement will contain at a minimum the security suite, which is the types of encryption algorithms to use and the format of the keys to expect.

While the initial sequence an entity uses to join a group are unicast between the entity and the GC, group management messages are multicast in nature and include rekey, policy changes, and group deletion.

GSAKMP life cycle

The life cycle of a GSAKMP group secure association can be divided into three phases, and these are now briefly discussed. The discussion is illustrated with the message flows shown below; the left side of the diagram represents the actions of the GC, and the right side of the diagram represents the actions of the GMs.

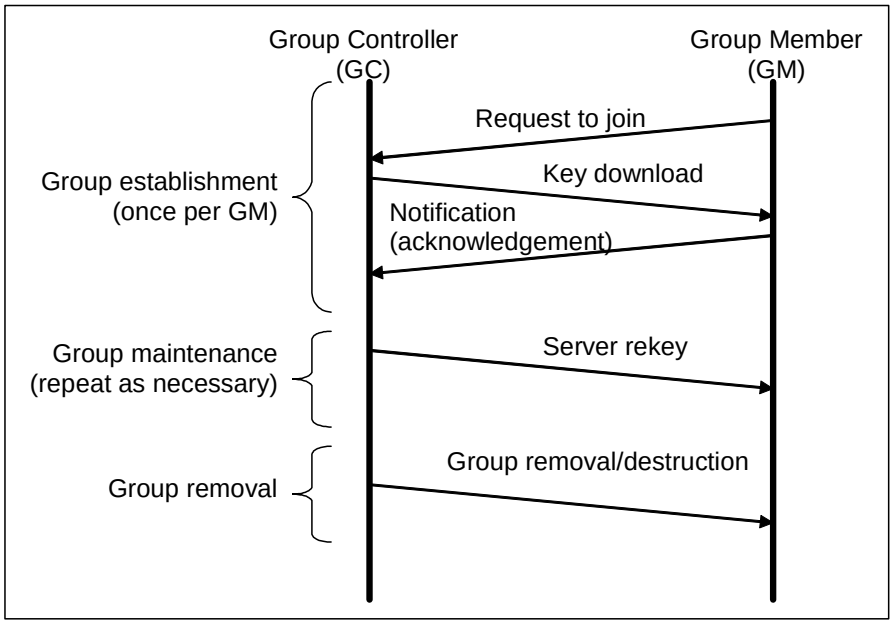


Figure 16: GSAKMP message exchange

GSAKMP group establishment

Potential GMs may join a group in one of two ways:

- invitation (push);
- request (pull).

Figure 16 illustrates a Request to Join group (RTJ), a "pull" message sent from a potential GM. On receiving the RTJ, the GC must either accept or deny the request. If accepted the GC checks the RTJ message and following successful authorization and verification creates the key download payload. This includes the policy token payload, the traffic encryption key payload and also the rekey event payload, which helps in scalable, group maintenance. The GM returns an acknowledgement to the GC on successful receipt of the key download message.

The member will initiate the following series of 3 messages for group establishment:

- Request To Join (RTJ) initiates the GSAKMP group establishment portion of the protocol. RTJ contains a key creation field for use in-group establishment;
- key download contains a key creation field and encrypted Policy Token and Key Download payloads;
- the Acknowledgement message completes the authentication of the member.

GSAKMP group maintenance

The Group Maintenance phase includes the following:

Member joins and leaves

The addition of group members to a previously established group will closely follow the processing presented in clause 6.4.4. With the exception of the pure group establishment tasks (e.g. creation of policy token, GTEK, and Rekey array), an entity becomes a GM using the same message exchanges described in clause 6.4.4.

A group member that elects to voluntarily leave the group is responsible for destroying their own key(s). Any further action for a voluntary leave must be specifically addressed in the group's security policy.

Rekey Events

A Rekey event is any action, including compromises, that involves the creation and dissemination of a new group key and/or Rekey information. Once it has been identified, using the group's security policy, that a Rekey event has occurred, the GC must create and send a signed message containing the GTEK and Rekey array to the group.

Each GM who receives this message must verify the signature on the message to ensure its authenticity. If the message signature does not verify, the processing is terminated. GSAKMP sends a properly authenticated message with a Notification Payload of type NACK to indicate termination by the GM. Upon verification the GM will find the appropriate Rekey download packet and decrypt the information with a stored Rekey key.

GSAKMP group removal/destruction

The final phase in the group's life cycle is group removal. If a decision is made to destroy the group, the notification may either be broadcast on a key management channel or through a directory service.

Rekey Management

The use of advanced re-key management techniques such as the Logical Key Hierarchy (LKH) reduces the overhead created by key management during re-key events. This allows extremely large groups (100 000 - 1 000 000 users) to be established and controlled in an efficient manner.

8.4.2.2 Multimedia Internet KEYing (MIKEY)

The MIKEY protocol is used for peer-to-peer, simple one-to-many, and small-size (interactive) groups, and is intended for use in real-time applications. One of the main multimedia scenarios is the conversational multimedia scenario, where users may interact and communicate in real-time. In these scenarios it can be expected that peers set up multimedia sessions between each other, where a multimedia session may consist of one or more multimedia streams (e.g. SRTP streams).

The MIKEY protocol is designed to have the following characteristics:

- end-to-end security: only the participants have access to the generated key(s);
- simplicity;
- efficiency: designed to have:
  - low bandwidth consumption;
  - low computational workload;
  - small code size;
  - minimal number of round-trips;
  - tunnelling, possibility to "tunnel" MIKEY in session establishment protocols (e.g. SIP and RTSP);
  - independent of specific security functionality of the underlying transport.

MIKEY can be used with symmetric keys, public keys, or shared key development using the Diffie-Hellman protocol. When MIKEY messages are tunnelled using SIP , the SIP security features could be used to secure MIKEY messages.

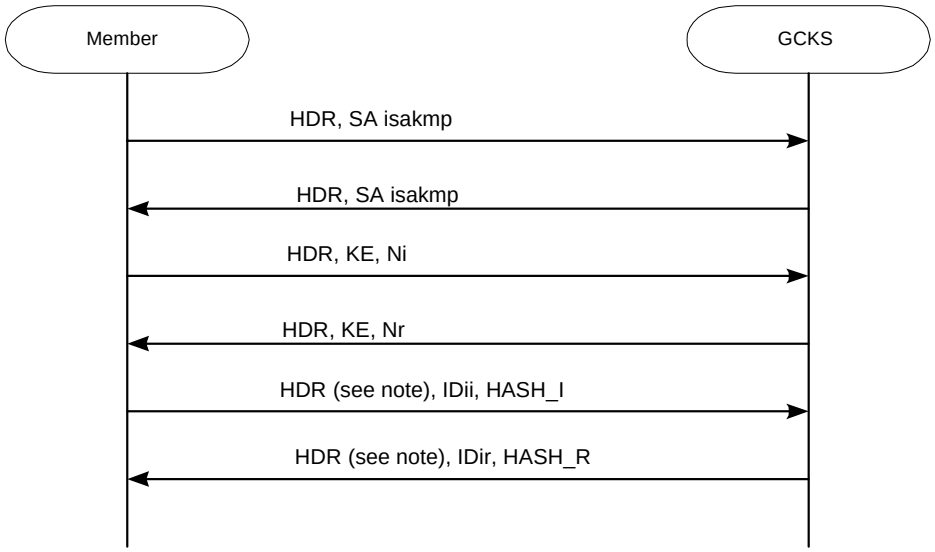
MIKEY only provides registration of users and initial key distribution. Unlike GSAKMP, it does not provide rekeying mechanisms or allow group security policy dissemination. The protocol is optimized primarily for real-time performance, but is not scalable to large groups.

8.4.2.3 Group Domain of Interpretation (GDOI)

GDOI (RFC 3547 [19]) is an ISAKMP Domain of Interpretation (DOI) for group key management to support secure group communications. It proposes new exchanges according to the ISAKMP and IKE standard. All GDOI messages are used to create, maintain, or delete security associations for a group. These security associations protect one or more Key-Encrypting Keys (KEK), Traffic-Encrypting Keys (TEK), or data shared by group members. GDOI is composed of 3 different exchanges.

IKE phase 1

GDOI re-uses the IKE Phase 1 to create an ISAKMP SA between a potential member and the Group Controller and Key Server (GCKS), which provides confidentiality and integrity. That secure channel is used to protect a new Phase 2 exchange called "GROUPKEY-PULL", which is defined below.



NOTE: Protected by the IKE Phase 1 SA, encryption occurs after HDR.

Figure 17: GDOI "IKE Phase 1" exchange

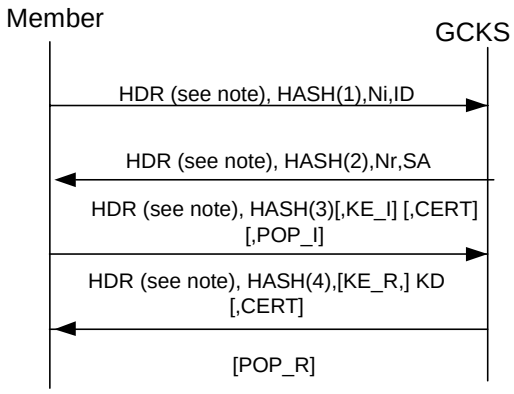
GROUPKEY-PULL exchange

The GDOI proposes a new Phase 2 exchange, called GROUPKEY-PULL, whose goal is to establish a Re-Key SA and/or Data Security SAs at the Member for a particular Group. This exchange is initiated by a potential member towards the GCKS. The IKE Phase 1 SA protects the GROUPKEY-PULL; there may be multiple GROUPKEY-PULL exchanges for a given Phase 1 SA.

The potential member sends the initial message with the identifier of the group (contained in the ID payload) it wants to join.

The GCKS Responder informs the potential member of the cryptographic policies of the Group in the SA payload. A Re-Key SA can be specified and one ore more Data Security SAs are specified.

The last message downloads the KD (Key Download) payload. If a Re-Key SA is defined in the SA payload, then KD will contain the KEK; if one or more Data SAs are defined in the SA payload, KD will contain the TEKs. Besides, if a Re-Key SA is defined in the SA payload, the SEQ payload must be present (the SEQ payload contains the current value of the sequence number which orders the Group-Key Push datagrams).



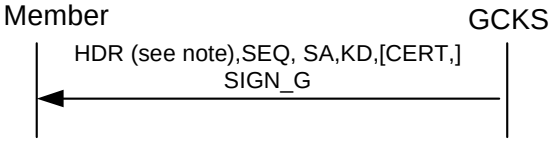
NOTE: Protected by the IKE Phase 1 SA, encryption occurs after HDR.

Figure 18: GDOI "GROUPKEY-PULL" exchange

The HASH payloads prove that the peer knows the Phase 1 secret key, and guarantee "liveliness" (prevents someone from replaying a recent GROUPKEY-PULL exchange). As a matter of fact, parameters of hashing functions always depend on random numbers (Ni and Nr), which guarantees a unique HASH payload for each message. With the HASH of the second message, the Group Member is sure that the message comes from the GCKS, and is not a replayed message. The HASH of the third message guarantees that the GROUPKEY-PULL exchange was initiated by the potential Group Member, and not by an intruder replaying a recent GROUPKEY-PULL exchange.

In the two last messages, a "Proof-of-Possession" (POP) and a "Certificate" (CERT) payloads can be included (optional). The POP payload contains the digital signature of the message, which has been computed with the private key of the sender. The certificate contains the corresponding public key certified by the group owner. If these payloads are sent by the member, this proves that it is authorized to get access to the group keys (the group owner gave him a public/private key pair). If they are sent by the GCKS, this proves that it is authorized to give the keys for this group.

GROUPKEY PUSH



NOTE: Protected by the IKE Phase 1 SA, encryption occurs after HDR.

Figure 19: GDOI "GROUPKEY-PUSH" exchange

The GDOI uses the GROUPKEY-PUSH datagram to replace a Re-Key SA KEK, and/or create a new Data Security SA. The GROUPKEY-PUSH message is "pushed" from the GCKS to the Members (in multicast). GROUPKEY-PUSH messages are encrypted with the Re-Key SA, and signed by the GCKS.

The KD payload contains the new keys.

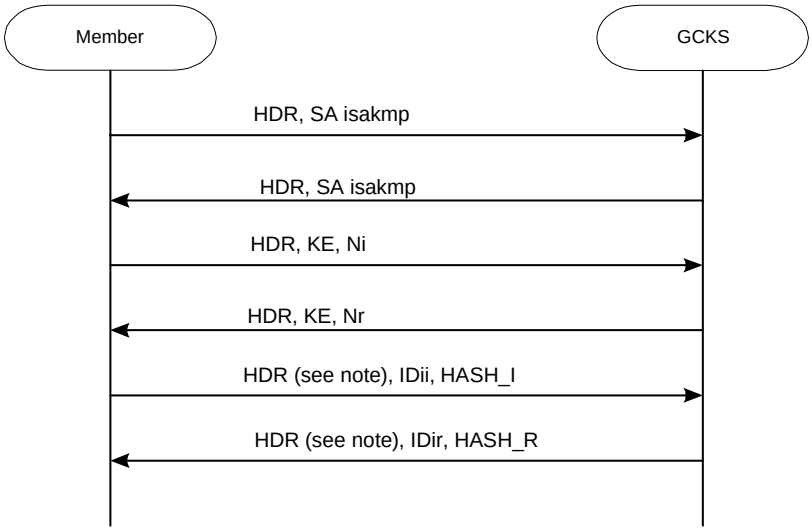
The SEQ payload contains a sequence number, which corresponds to the sequence number, which has been initialized in the GROUPKEY-PULL exchange. The Member accepts only the GROUPKEY-PUSH messages with a sequence number superior to the previous message's ones (this mechanism avoids replay attacks).

8.4.2.4 Flat Multicast Key Exchange (FMKE)

FMKE (Flat Multicast Key Exchange, (IETF draft-duquer-fmke-00.txt [50]) is a new group key management protocol. Its objective is to manage securely group Security Associations (SA), i.e. to establish and update Data-Security and Re-Key SAs in Group Members. FMKE is composed of 3 different exchanges.

FMKE Phase 1 (IKE phase 1)

FMKE re-uses the IKE Phase 1 to create an ISAKMP SA between a potential member and the Group Controller and Key Server (GCKS), which provides confidentiality and integrity. That secure channel is used to protect a new Phase 2 exchange, which is defined below.



NOTE: Protected by the IKE Phase 1 SA, encryption occurs after HDR.

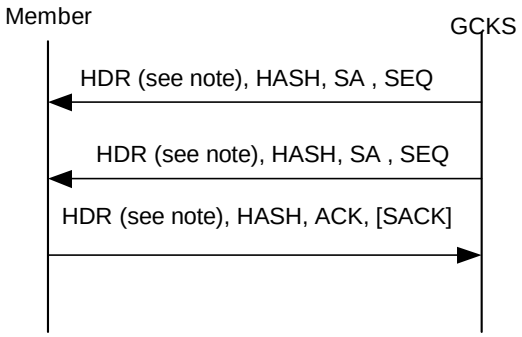
Figure 20: FMKE phase 1 exchange

FMKE Phase 2

The goal of the Phase 2 exchange is to establish Re-keys SAs and/or Data-security SAs at the member. The transmitted SAs belong to the same or to different groups. During this phase, the member can receive all the Data security SAs and Re- key SAs it is authorized to get access to .The Phase 2 exchange takes place once, after the Phase 1; and is initiated by the GCKS. It is protected by the Phase 1 SA.

Phase 2 exchange is composed of two types of messages: the messages sent by the GCKS, which contain the SA attributes the member is authorized to get access to, and the messages sent by the member, which are used to acknowledge the previous messages.

Acknowledgement messages are sent periodically by the member.



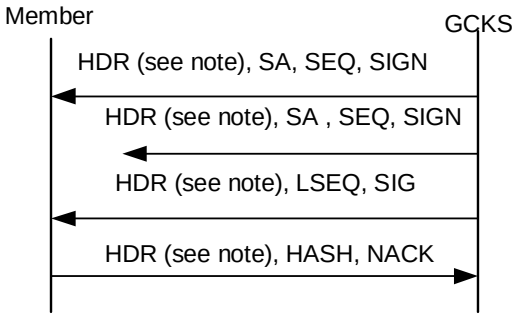
NOTE: Protected by the phase 1 SA, encryption occurs after HDR.

Figure 21: FMKE phase 2 exchange

In the GCKS messages, the SEQ payload contains the current value of the sequence number, which orders these messages. In the member messages, the value of the ACK (Acknowledgement) payload mentions the sequence number of the last correctly received message. The SACK (Selective Acknowledgement) payload can be optionally included in the Phase 2 member messages. When one or several GCKS messages are missing, it allows mentioning of some of the already received next messages numbers.

The HASH payload proves that the messages have not been modified during transmission and that the peer knows the Phase 1 secret. The HASH payload guarantees also that messages are not replayed from an old session establishment, as it required the secret of the last Phase 1. The SEQ payload in the GCKS messages allows the member to delete all messages already received in the Phase 2, as it has to check that the sequence number is greater than in the previous SEQ payloads.

FMKE Phase 3



NOTE: Protected by the Re-key SA, encryption occurs after HDR.

Figure 22: FMKE phase 3 exchange

The goal of the Phase 3 is to create, update, and/or replace Data security SAs, and/or to update the Re-key SA into Group Members (belonging to the same group). The Phase 3 is protected by the Re-key SA of the group. The GCKS pushes the new SA attributes.

In the Phase 3, the GCKS sends two types of messages. The first type contains the SA payload, and is used to create, update and/or replace new SAs (TEKs and KEKs are included). In these messages, as in Phase 2, a SEQ payload is included containing a sequence number which orders them. The second type messages are sent periodically. They contain a LSEQ (Last Sequence Number) payload whose value mentions the last sequence number used by the GCKS. They are required to ensure reliability.

The Group Member sends only one type of messages. These messages are used as Non-acknowledgement messages, that is to say they request the retransmission, in multicast, of the message(s), whose sequence number(s) is(are) mentioned in the NACK (Negative Acknowledgement) payload(s). A Group Member can determine that it has not received one or several messages thanks to the SEQ and LSEQ payloads included in the messages sent by the GCKS. For a group, a Phase 3 exchange can be initiated by the GCKS as soon as the Re-key SA is established in at least one Group Member.

The SIG payload contains the digital signature of the entire message before encryption (including the header and excluding the SIG payload itself). The signature guarantees source data authentication, i.e. the source of these messages is the GCKS.

The SEQ payload of the GCKS messages allows Group Members to detect replayed messages: they delete all already received messages.

The HASH payload of the Group member messages proves that the messages have not been modified during transmission, and that the source is one of the Group members.

Rekeying:

Two Rekeying modes are defined: one in multicast, based on the phase 3 (same messages except that a Key Download payload is added in the GCKS messages), and one in unicast based on the phase 2 (same messages except that a Key Download payload is added in the GCKS messages).

FMKE can be seen as a use case of GDOI. The main differences between GDOI and FMKE are:

- FMKE offers reliable key distribution exchange;
- FMKE Phase 2 requires fewer messages for an equal number of SAs to transmit than GDOI GROUPKEY PULL;
- in FMKE, members do not request to get access to a group. They receive directly all the SAs they are authorized to know.

8.5 Access control

Functionally, access control can be divided into two distinct issues: The first issue is to protect a cluster of more loosely administered machines hidden behind it from hackers. The solution is a dedicated gateway machine with special security precautions on it, used to service outside network, especially Internet, connections and dial-in lines. This called a firewall.

The second issue is limited to more advanced satellite systems, which feature satellites with increasing processing responsibilities. This is called capacity protection and is described in clause 8.5.2. Also protocol issues related to bandwidth request/assignment are presented in clause 8.5.3.

8.5.1 Firewalls

Firewalls are a basic means for providing network security, behaving like a moat around a medieval castle (see figure 23). Firewalls restrict information entering and leaving at carefully controlled points, and prevent unacceptable attempts at accessing resources within the firewall. While an important use of firewalls is to enable secure Internet access to corporate networks, they are also used to restrict access to departmental private and mission critical information.

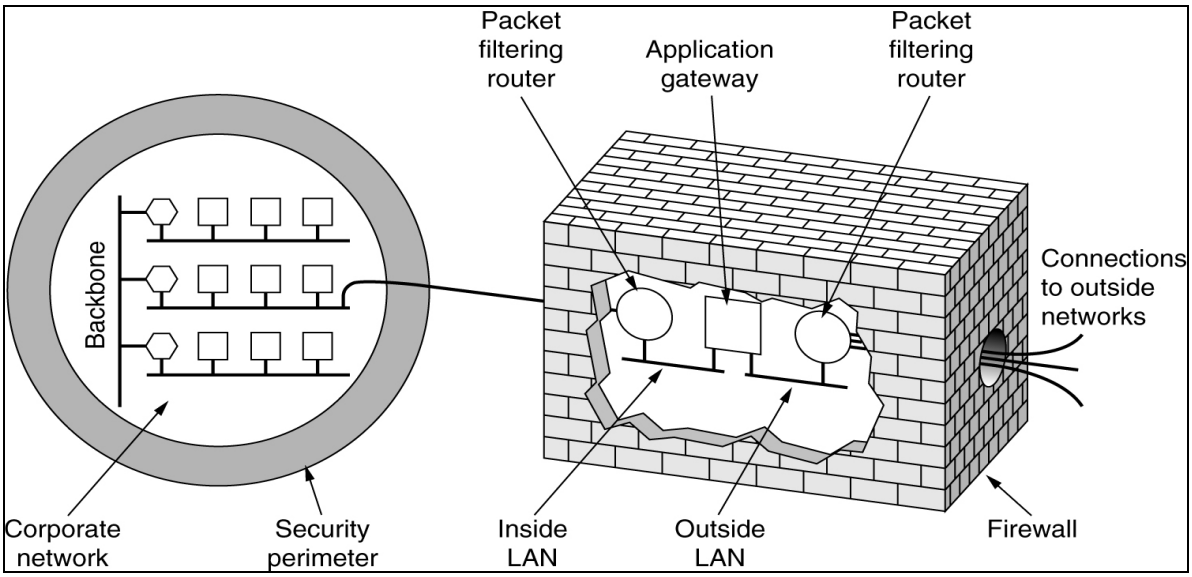


Figure 23: Firewall protection from untrusted networks

Firewalls are not the answer to all security problems an organization faces; Firewalls themselves are vulnerable to security violations. In general firewall can perform the following:

- packet filtering is one of the key operations performed by a firewall. This is the operation where certain packets or datagrams are allowed to pass through the firewall and others are not. Filtering may be done on;
- the source IP address, b) the destination IP address, c) the TCP or UDP source port number, d) the TCP or UDP destination port number, e) the IP protocol (ID) number;
- Network Address Translations (NAT) translates internal or private IP addresses to public or globally routable IP addresses. This translation allows users to hide their internal network's address, and yet communicate with networks that use public IP address;
- IP Security (IPSec) as described above;
- proxy servers or application firewalls is an approach to have more than one firewall in which some of the filtering are performed in what is referred to as a proxy firewall or an application gateway. For example all TELNET and FTP packets to a trusted network specific host (destination server) are passed by the application gateway only if compliant with firewall access control list. In such a way, the TELNET/FTP application source (i.e. the remote client attached to an untrusted network) never connects to the destination server directly.

8.5.2 Capacity protection in a regenerative satellite system

Regenerative Satellite systems with extensive on-board processing and a complex switch fabric supporting a mesh topology with a large number of antenna beams and inter-satellite links share a security concern unknown to conventional and simpler satellite system designs. Within the BSM family structure, these form the regenerative satellite mesh (RSM) family of satellite systems.

Locating the Bandwidth on Demand (BoD) function on-board the satellite is desirable to reduce the delay associated with uplink resource assignments. The ST requests bandwidth and the network grants the request from the pool of available resources, thus incurring at least two round trip delays in the conventional system but only one round trip delay when the processing is performed on-board. However, the Satellite has limited processing and memory capability and cannot have access to ST subscription information or service level agreements when processing bandwidth requests and cannot store capacity usage data. On the other hand, the hub has complete access to all subscriber information and can record all usage information if BoD requests are processed there.

The problem is compounded by the Mesh topology of the RSM satellite systems with hundreds of spot beams. Since user data can flow in a single hop between any two STs rather than through a HUB, as is the case in a star topology, the HUB or the Network Operations Control Centre (NOCC) does not see user data flows and cannot record usage data. This would be a burden on the satellite's resources to require the satellite to record usage information.

Capacity protection is the security mechanism responsible for all aspects of protecting satellite capacity against misuse. It is specifically responsible for protecting against the following ST transgressions:

- transmitting in unauthorized slots or frequencies;
- transmitting with falsified Source IDs;
- impersonating another ST;
- accessing unauthorized services;
- accessing downlink beams and inter-satellite links for which it is unauthorized;
- under-reporting its usage;
- exceeding its authorized usage.

Capacity protection should be strong enough so that fraud is not cost effective. That is, the service obtained via fraud must be worth less than the cost of bypassing system security.

A centralized and very simple capacity protection solution works well in a star topology and when the BoD function is performed at a HUB since all requests come to the HUB and all user data traffic flows through the HUB. A distributed and relatively more complex capacity protection solution may be preferred in a mesh topology and when the BoD is performed on board the satellite.

8.5.2.1 Problems, risks and threats

From a security standpoint the Satellite System is in many ways similar to earlier, wireless, public networks. All such networks have requirements to protect against unauthorized usage. The system must be protected against carrying traffic from either unauthorized STs, carrying traffic which in some way exceeds a Set's authorized limits or carrying traffic from STs in such a way that the actual end-user is not billed for the usage.

RSM satellite systems differ from earlier wireless, public networks in that it is a single-hop, packet-switched network where the packet switch's resources are located in space and are extremely limited. Without special measures, an unauthorized ST might easily be able to use the system by simply transmitting packets through TDMA slots into the switch and have it route them to the desired destination, potentially without the knowledge of the NOCC. An ST, which transmits at higher than normal power would be able to transmit through any TDMA slots and overpower the STs to which the slots have been allocated.

8.5.2.2 Dependency on other security mechanisms

Capacity protection is dependent upon three other security mechanisms:

a) space segment security;

b) ST security;

c) signalling security.

Space segment security is responsible for protecting the payload against unauthorized access. It is also responsible for protecting the spacecraft bus against unauthorized access such as protection of Telemetry Tracking and Command (TT&C) communications with the Spacecraft.

Capacity Protection depends on Space Segment security in that the Space Segment may receive from the NOCC key material necessary for the Space Segment to implement its part of Capacity Protection. Space Segment security provides for the protection of this key material during transit to the payload.

Space Segment Security mechanisms are implemented by conventional means and are internal to the Space and NOCC Segments respectively. Space segment security is outside the scope of the present document.

ST security is responsible for protecting an ST against unauthorized use.

Capacity Protection depends on ST Security in that tampered ST software offers some limited ability to exceed authorized capacity by bypassing ST software enforced admission control, uplink power management and other capacity limiting policies.

Capacity Protection depends on ST Security to a larger degree in that a compromised ST allows some security compromises that are only detectable via usage auditing.

Signalling security is responsible for:

a) the security aspect of a Set's entering the system, that is, ST registration;

b) providing privacy and tamper resistance for network management communications between the NOCC and ST;

c) the secure distribution of capacity protection key material to STs.

Capacity Protection depends on Signalling Security to securely distribute its key material to STs.

8.5.3 Capacity protection protocol issues

The remainder of this clause describes Capacity Protection in terms of:

a) packet authentication – how a satellite authenticates traffic packets received from STs on traffic channels and how a satellite determines that the requested routing or switching is allowed;

b) bandwidth requests - how the Satellite authenticates a bandwidth request;

c) bandwidth grants - how the ST authenticates a bandwidth grant.

8.5.3.1 Packet authentication

The packet authentication and checking algorithms:

- should minimize the satellite resources necessary to perform the checking;
- should not require rigid key distribution synchronization;
- should provide key material to the ST that can be used to authenticate that ST's identity, limiting the system damage caused by a compromised ST;
- should provide a transmission mechanism whereby an unregistered or unauthorized ST, or an ST, which does not have valid keys, may still access the NOCC for management functions, such as registration, request for system information or security key distribution, through the satellite;
- should support the system ability to enforce payload bandwidth constraints such as flow-control, quality of service metrics and congestion management policies;
- should allow authorization enforcement that limits an ST to resource utilization only for a particular purpose;
- the NOCC should be able to control an ST's access to the following resources:
  - traffic-bearing transmission through contention channels;
  - special downlink beams (region, global, etc. which might use higher satellite transmission power) and resources;
  - multicast transmission, i.e. access to satellite packet replication resources;
  - Inter-Satellite Link (ISL) transmission;
  - dedicated channel groups, i.e. uplink resources which are dedicated for use by a sub-set of STs or users (frequencies or time slots).

8.5.3.2 Bandwidth requests

Conventionally, when the Medium Access Control (MAC) sublayer has packets to transmit, it sends a bandwidth request to the network. If the bandwidth request is processed at a HUB or NOCC, the bandwidth request can be readily compared with any SLA or subscriber information and granted or denied as appropriate. Fraud or misuse can be readily detected. When the BoD function is moved to the satellite, access to subscriber information is impractical. Therefore, a different bandwidth request mechanism is required.

This bandwidth request mechanism should be able to deny any bandwidth requests, which violates an SLA or the ST's subscription. This mechanism should also protect against a single or small group of STs flooding the spacecraft with bandwidth requests. Since initial bandwidth requests would likely use contention channels, the bandwidth request mechanism should limit access to these resources. A bandwidth request protection algorithm should provide a mechanism that allows the ST to approve each bandwidth request before it is transmitted to the spacecraft. The algorithm should provide the spacecraft with a means of authenticating requests to ensure that they have not been tampered with after approval. This algorithm should be designed to be strong enough and flexible enough to support a wide range of Capacity Protection designs and bandwidth-on-demand algorithms over the life of the system.

8.5.3.3 Bandwidth assignment

Within the Capacity Protection security mechanism, the ST should be able to authenticate bandwidth assignments received from the payload. To do this, the ST should be able to verify that a specific bandwidth assignment (including the frame number the assignment applies to) was received, untampered with, from the payload.

The algorithms used to achieve this should minimize satellite resources such as power, processing and memory.

8.5.4 The RSM-A solution

The SAM is the security component of a satellite terminal. Physically it is a secure chip embedded in the terminal. The SAM contains secret key material and authenticates every Regenerative Satellite Mesh - Type A (RSM-A) packet sent out by the terminal by generating an Access Control Field (ACF), which can be verified by other authorized components of the system. The SAM will only sign requests that are valid within the policies set forth for that particular ST. On the receive side it verifies that management messages are authentic messages from the NOCC. This system is described in
TS 102 189-3 [11] and shown in figures 24 and 25.

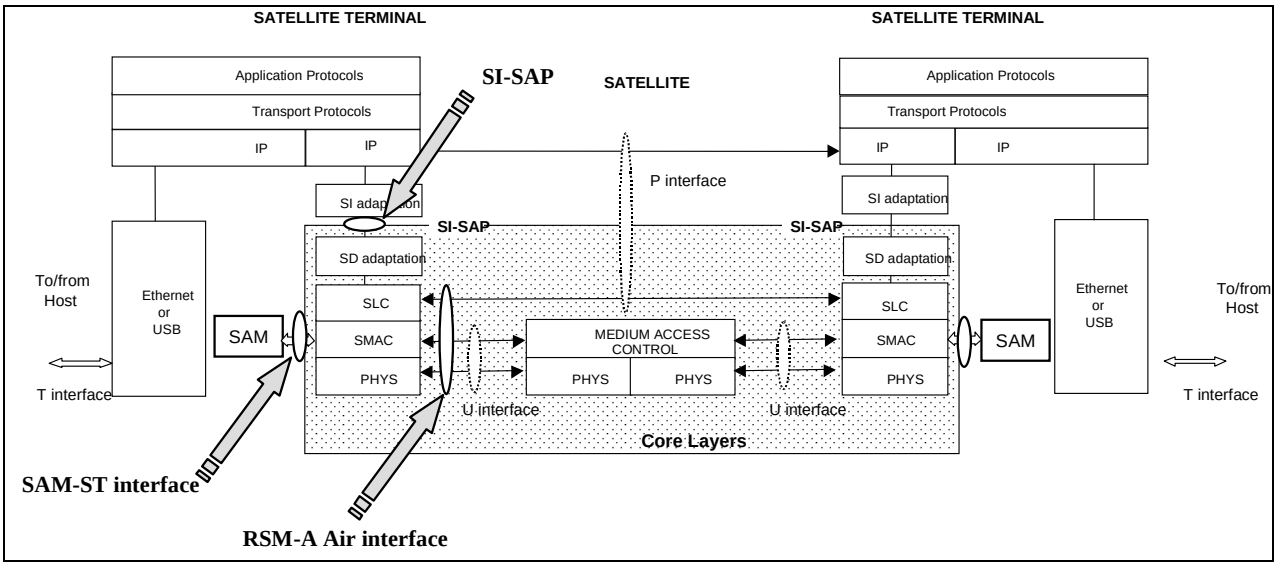


Figure 24: Protocol Architecture and SAM-ST interface

The SAM's areas of responsibility to the RSM-A system are as follows:

- authentication;
- authorization protection;
- registration;
- usage audit.

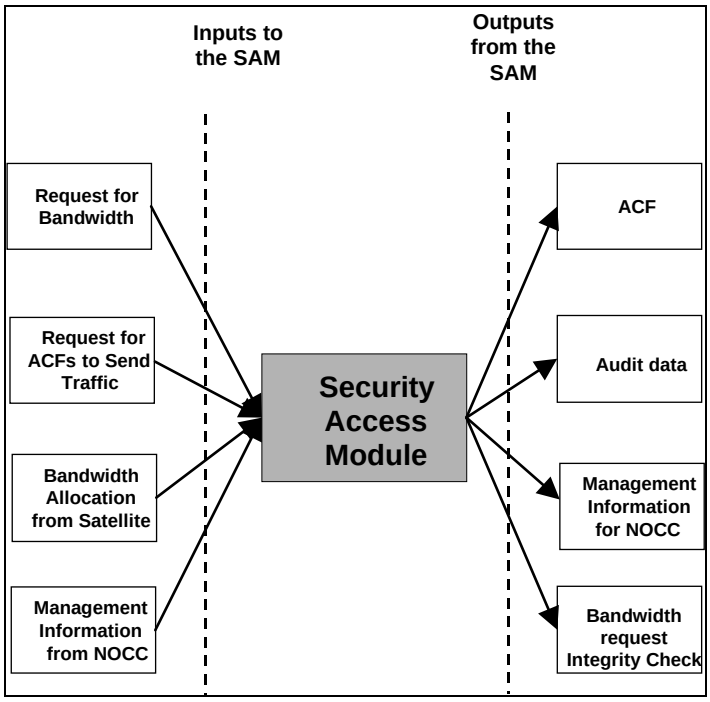


Figure 25: Security function interactions between the SAM and the ST

An important functionality for the SAM is bandwidth request verification. The ST requests Bandwidth from the satellite based on the policies that it signed up for and received from the NOCC during the registration process. The SAM ensures that the ST adheres to these policies and therefore will not sign requests that are not within these guidelines. The SAM will approve requests that agree with the policy permitting the ST to send a message to the satellite requesting a bandwidth allocation. The SAM will return an Integrity Check code within the bandwidth request response message if the ST is permitted to request the bandwidth. The SAM also returns a bandwidth request failure message to the ST with a failure code if the SAM rejects the ST's bandwidth request. The ST does not request bandwidth from the satellite without a valid integrity code from the SAM.

Also the SAM communicates some management messages to the NOCC through the ST. These messages are typically transparent to the ST and are always encrypted by the SAM. The NOCC sends messages to the SAM through the ST and these messages are also always encrypted. The ST will strip the ST-SAM header from all messages received from the SAM and destined for the NOCC. The ST will add UDP/IP headers and management protocol headers as required.

8.6 Key management options for BSM systems

As presented in clause 8.1 to 8.5, the security management and access control solutions depending on the layering implementation of security.

IPSec and its related unicast and multicast key management/distribution architectures are very convenient for implementation in BSM networks.

Satellite terminal access control using a scheme like SAM is another convenient tool to prevent access to unauthorized users.

If DVB-RCS or DVB-S systems are used then a combination of IPSec and DVB-S CA or DVB-RCS security can be implemented to cover various aspects of satellite and end to end security requirements. However combined solutions will incur extra costs in terms of computational power requirements, throughput restrictions and delay.

9 Recommended Specifications to be produced by ETSI

9.1 Discussion

In the present document, clause 6 has presented the general threats to BSM network, service providers and users with an analysis of the required security services to prevent these threats. Also clause 7 has given an overview of the choices of security solutions available to BSM networks in the protocol stack with a comparison of merits of end-to-end and satellite only security solutions. In addition, clause 8 has presented various solutions to the most difficult problem in security i.e. "security management and access control". The security management solutions included multicast, DVB-S (broadcast), DVB-RCS (interactive satellite) and the IPSec solutions.

The IPSec and its related unicast and multicast key management/distribution architectures are very convenient for implementation in BSM networks and this solution will be mainly implemented in the Satellite Independent (SI) part of the protocol stack. Some of BSM security management will relate to securing unicast and on-to-one connections. However, a large part of the security will relate to multicasting and broadcasting. The process of securing and performing key management for multicast and broadcast is more complex than unicast. Clause 8.4, has introduced some factors that influence the group security system design for BSM. Examples of these factors are multicast application type, group dynamics, scalability and the underlying trust model. The combinations of these factors will be the deciding factor on the group security policy parameters such as:

- procedures for establishing, running and terminating the group;
- key distribution methods. For example flat key or LKH distribution system can be chosen to fit the requirements of that group;
- the strength of the encryption algorithms and digital signature needed, such as using DES, triple DES or AES for privacy and RSA or DSS signature systems;
- what to do when things go wrong: such as network failure, DoS attacks, problems with entity authentication and authorization. The procedures are usually defined in the security policy. There must be a clear way for creation, dissemination and enforcement of security policies.

Also these factors will influence the choice of key management protocol such as GDOI, FMKE, GSAKMP or MIKEY as presented in clause 8.4.

For broadcasting services a careful attention to the security requirement and network configuration will be the deciding factors in choosing security management systems such as DVB-S CA, DVB-RCS or IPSec and its related group management protocols.

Satellite network access control has been presented in clause 8.5 with the focus on two types of systems: One is firewall and the other is chip/smartcard such as the SAM. Firewalls mainly protect the private network from outside attacks. The chips/smartcards protect from insiders attacks and misuse, where capacity protection issues are described in clause 8.5.2 and protocol issues related to bandwidth request/assignment are presented in clause 8.5.3. Therefore, access control will play a major part in satisfying some of the security requirements for many services over BSM networks.

Also there are many other issues to be considered when implementing a security system for BSM networks such as:

- transparency: such as detailing the conditions under which IPSec can be supported transparently over BSM networks. Also application layer security can work transparently over BSM networks such as the DRM scheme mention in clause 7.4.2;
- DVB-S conditional access: critical examination of the DVB-S conditional access and its usefulness for BSM services;

- multicast security solutions: critical analysis of the multicast security solutions in the IETF Multicast SECurity (MSEC) group and the associated key distribution architectures and recommend a few concrete solutions for BSM networks;
- security features: recommend security features to be provided by BSM systems; e.g. features to be provided at link layer, network layer etc.;
- security policy: define BSM security policy for specific threats (such as Denial of service, masquerading and replay attacks);
- interaction between layers: the mapping between security management functions and the lower layers such as the relationship between multicast security management and IPSec (IP layer), DVB-S or DVB-RCS (link layer, such as MPEG-TS) messages.

9.2 BSM Security Manager (BSM-SM)

From the previous clauses it is easily inferred that for maintaining security and evaluating data privacy/integrity performance of the BSM world in the Internet world, there is a need for some "manager". The protocol stack from the BSM drives the development of the BSM Security Manager (BSM-SM). The "manager" resides above the SAP and defines how IP protocols and packet are secured through the BSM, which Satellite Independent (SI) protocols are used and how they in turn might trigger the Satellite Dependent (SD) functions.

The BSM-SM must have the flexibility to support various types of BSM operators from a wholesale bandwidth provider to a network operator. It is located both at ingress and egress of the BSM and will have ST implementation as well as NCC implementations. Therefore the BSM-SM can be described as an "intelligent" manager that supervises security procedures with reference to security policies and access control procedures.

Managing security can go from "do not care" (i.e. full transparency) to full capabilities for managing security and policing. Hence the BSM-SM architecture must be modular and easily upgradeable. Also, because of the nature of the BSM, the security manager might need to interact with different layers of the OSI stack. The BSM-SM could be designed as a standalone, however in reality it will have to relate to other BSM control and management entities. Hence its implementation will have to rely on module and established module communications.

The BSM-SM has a lot of commonality with the IPSec and MSEC security management functions, such as unicast key management, multicast group establishment/maintenance/removal and multicast group key distribution. However, BSM-SM might have some specific functions to satisfy the BSM service requirements. For example, if the SAM is used for access control, then interactions between the BSM-SM and SAM must be defined. Also if link layer security is chosen (such as the DVB-RCS security recommendations), then the BSM-SM will interact through the SAP with lower layers. An overview of the system architecture is described in [3] and shown in figure 3 The BSM-SM might communicate at different levels of the BSM stack.

9.3 Recommended security TSs

Figure 26 presents the structure of the recommended technical specifications (TSs). They are separated to multicast and unicast. All recommended TSs share the following characteristics, namely:

- the use of open specifications: Where appropriate, the proposed TSs will be based on available open specifications (e.g. the outputs of the IETF MSEC and IPSec groups);
- all specifications should be located in the satellite independent layers and should be independent of the specifics of the satellite dependent lower layers where appropriate;
- all specifications should interwork with IP functions;

In the following is a brief description of all the recommended TSs:

- TS1: General security architecture: This specification covers the security manager architecture as described in detail in clause 9.2 based on the BSM architecture [3]. The architecture will do a functional decomposition showing which functions are positioned above and below the SI-SAP [12];
  - TS 1.1: Performance parameters: This specification defines a BSM specific set of security performance parameters that are measurable at the ST, the NCC/gateway and the service provider. This can be based on IETF current work in the IPSec and MSEC groups. Examples of such parameters are processing power needs for security, extra data overheads, recovery time from minor/major failure situations, the choice of security algorithms and protection measures against denial of service attacks.
- TS2: Unicast security. This specification covers aspects of unicast security policy data handling and key exchange and storage;
  - TS2.1: Unicast security policies. This specification covers aspects of policy in the context of unicast security, taking into consideration the fact that policies may be expressed in different ways and may exist at different levels in a given unicast security architecture;
  - TS2.2: Unicast data handling. This specification covers the security-related treatments of unicast data by the sender and the receiver such as data encryption/decryption and digital signature signing/verification.
- TS 3: Multicast security. This specification provides an overview and rationale of the multicast security architecture used for small and large multicast groups. It will specify the multicast security reference Framework and the security services that may be part of a secure multicast solution;
  - TS3.1: Multicast security policies. This specification covers aspects of policy in the context of multicast security, taking into consideration the fact that policies may be expressed in different ways and may exist at different levels in a given multicast security architecture;
  - TS3.2: Multicast data handling. This specification covers the security-related treatments of multicast data by the sender and the receiver such as data encryption/decryption and digital signature signing/verification;
  - TS3.3: Group Key Management. This specification is concerned with the secure distribution and refreshing of keying material;
    - TS3.3.1: Registration protocol. This is a protocol between the group controller/key server and a joining group member to mutually authenticate each other;
    - TS3.3.2: Rekey protocol. This is a mechanism that is used by the group controller/key server to periodically update or change the data Security Association (SA) by sending rekey information to the group members. Rekey messages may result from group membership changes, changes in a group security policy, the creation of new traffic-protection keys or from key expiration.
- TS4: Capacity protection security: This specification covers a coherent access control and secure bandwidth request/assignment to satellite terminals;
- TS5: DRM security. DRMs (Digital Rights Management) are application level security systems. This specification covers interaction aspects between the BSM-Security manager and DRM.

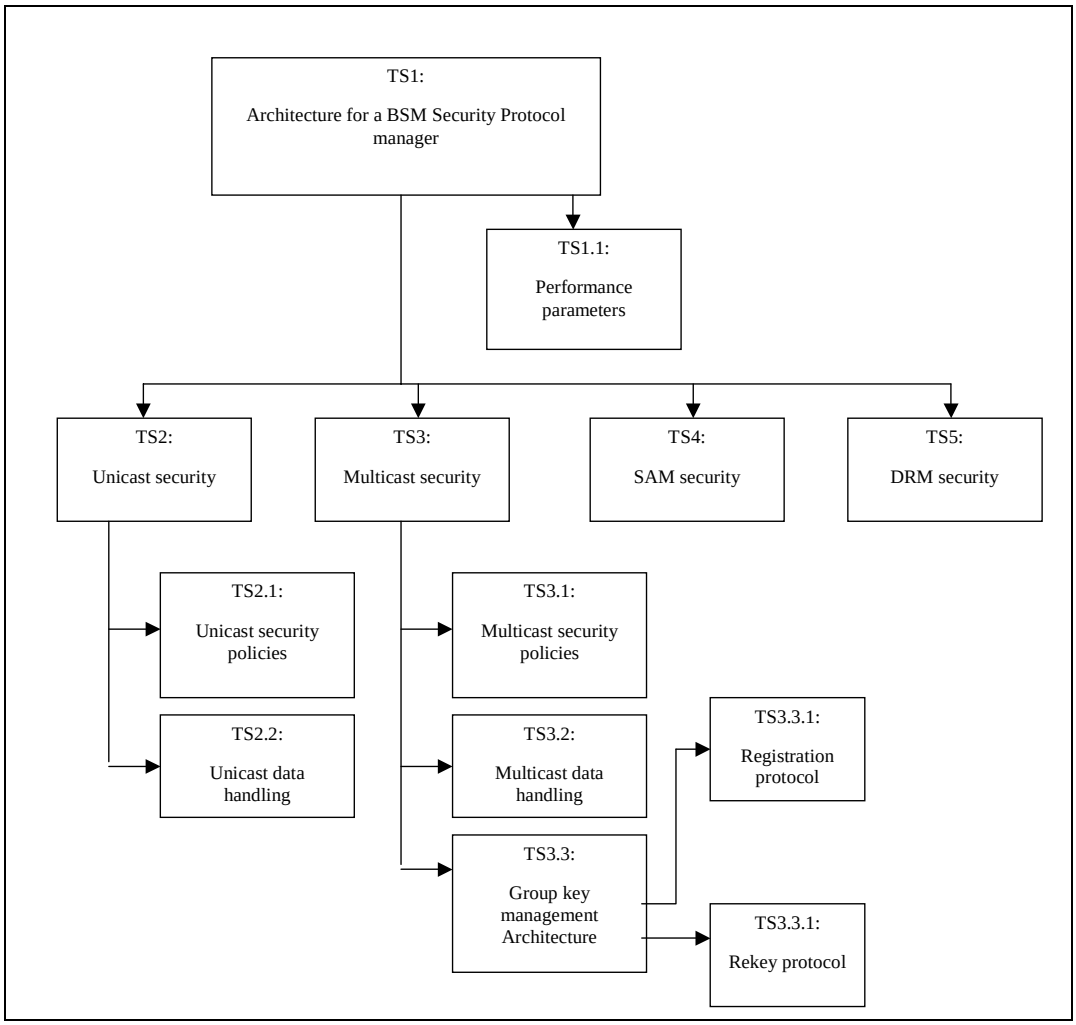


Figure 26: Recommended Technical Specifications

List of figures

List of tables

History

Document history		
V1.1.1	May 2004	Publication