ATIS-0100036.2013(R2018)

Media Plane Security Impairments for Evolving VoIP/Multimedia Networks

AMERICAN NATIONAL STANDARD FOR TELECOMMUNICATIONS

As a leading technology and solutions development organization, the Alliance for Telecommunications Industry Solutions (ATIS) brings together the top global ICT companies to advance the industry's most pressing business priorities. ATIS' nearly 200 member companies are currently working to address the All-IP transition, 5G, network functions virtualization, big data analytics, cloud services, device solutions, emergency services, M2M, cyber security, network evolution, quality of service, billing support, operations, and much more. These priorities follow a fast-track development lifecycle — from design and innovation through standards, specifications, requirements, business use cases, software toolkits, open source solutions, and interoperability testing.

ATIS is accredited by the American National Standards Institute (ANSI). The organization is the North American Organizational Partner for the 3rd Generation Partnership Project (3GPP), a founding Partner of the oneM2M global initiative, a member of the International Telecommunication Union (ITU), as well as a member of the Inter-American Telecommunication Commission (CITEL). For more information, visit www.atis.org.

AMERICAN NATIONAL STANDARD

Approval of an American National Standard requires review by ANSI that the requirements for due process, consensus, and other criteria for approval have been met by the standards developer.

Consensus is established when, in the judgment of the ANSI Board of Standards Review, substantial agreement has been reached by directly and materially affected interests. Substantial agreement means much more than a simple majority, but not necessarily unanimity. Consensus requires that all views and objections be considered, and that a concerted effort be made towards their resolution.

The use of American National Standards is completely voluntary; their existence does not in any respect preclude anyone, whether he has approved the standards or not, from manufacturing, marketing, purchasing, or using products, processes, or procedures not conforming to the standards.

The American National Standards Institute does not develop standards and will in no circumstances give an interpretation of any American National Standard. Moreover, no person shall have the right or authority to issue an interpretation of an American National Standard in the name of the American National Standards Institute. Requests for interpretations should be addressed to the secretariat or sponsor whose name appears on the title page of this standard.

CAUTION NOTICE: This American National Standard may be revised or withdrawn at any time. The procedures of the American National Standards Institute require that action be taken periodically to reaffirm, revise, or withdraw this standard. Purchasers of American National Standards may receive current information on all standards by calling or writing the American National Standards Institute.

Notice of Disclaimer & Limitation of Liability

The information provided in this document is directed solely to professionals who have the appropriate degree of experience to understand and interpret its contents in accordance with generally accepted engineering or other professional standards and applicable regulations. No recommendation as to products or vendors is made or should be implied.

NO REPRESENTATION OR WARRANTY IS MADE THAT THE INFORMATION IS TECHNICALLY ACCURATE OR SUFFICIENT OR CONFORMS TO ANY STATUTE, GOVERNMENTAL RULE OR REGULATION, AND FURTHER, NO REPRESENTATION OR WARRANTY IS MADE OFMERCHANTABILITY OR FITNESS FOR ANY PARTICULAR PURPOSE OR AGAINST INFRINGEMENT OF INTELLECTUAL PROPERTY RIGHTS. ATIS SHALL NOT BE LIABLE, BEYOND THE AMOUNT OF ANY SUM RECEIVED IN PAYMENT BY ATIS FOR THIS DOCUMENT, AND IN NO EVENT SHALL ATIS BE LIABLE FOR LOST PROFITS OR OTHER INCIDENTAL OR CONSEQUENTIAL DAMAGES. ATIS EXPRESSLY ADVISES THAT ANY AND ALL USE OF OR RELIANCE UPON THE INFORMATION PROVIDED IN THIS DOCUMENT IS AT THE RISK OF THE USER.

NOTE - The user's attention is called to the possibility that compliance with this standard may require use of an invention covered by patent rights. By publication of this standard, no position is taken with respect to whether use of an invention covered by patent rights will be required, and if any such use is required no position is taken regarding the validity of this claim or any patent rights in connection therewith. Please refer to [http://www.atis.org/legal/patentinfo.asp] to determine if any statement has been filed by a patent holder indicating a willingness to grant a license either without compensation or on reasonable and non-discriminatory terms and conditions to applicants desiring to obtain a license.

ATIS-0100036.2013(S2018), *Media Plane Security Impairments for Evolving VoIP/Multimedia Networks*

Is an American National Standard developed by the ATIS **Network Performance, Reliability, and Quality of Service Committee (PRQC)**.

Published by
Alliance for Telecommunications Industry Solutions
1200 G Street, NW, Suite 500
Washington, DC 20005

ATIS-0100036.2013(R2018)

American National Standard for Telecommunications

Media Plane Performance Security Impairments for Evolving VoIP/Multimedia Networks

Alliance for Telecommunications Industry Solutions

Approved: January 23, 2013

(Republished April 2022 with an administrative edit)

American National Standards Institute, Inc.

Abstract:

This ATIS Standard is intended to provide awareness and information regarding the use of security mechanisms in support of Next Generation Network (NGN) National Security and Emergency Preparedness (NS/EP) Services. When introducing network security mechanisms (e.g., IPSec) into Evolving Voice over Internet Protocol (VoIP)/Multimedia Networks, one may encounter impairments introduced or exacerbated by those network security mechanisms. One may need to explore tradeoffs between security and QoS to achieve the necessary communication channel during NS/EP conditions.

Foreword

The information contained in this Foreword is not part of this American National Standard (ANS) and has not been processed in accordance with ANSI's requirements for an ANS. As such, this Foreword may contain material that has not been subjected to public review or a consensus process. In addition, it does not contain requirements necessary for conformance to the Standard.

The Alliance for Telecommunication Industry Solutions (ATIS) serves the public through improved understanding between providers, customers, and manufacturers. The Network Performance, Reliability, and Quality of Service Committee (PRQC) develops and recommends standards, requirements, and technical reports related to the performance, reliability, and associated security aspects of communications networks, as well as the processing of voice, audio, data, image, and video signals, and their multimedia integration. PRQC also develops and recommends positions on, and foster consistency with, standards and related subjects under consideration in other North American and international standards bodies.

ANSI guidelines specify two categories of requirements: mandatory and recommendation. The mandatory requirements are designated by the word *shall* and recommendations by the word *should*. Where both a mandatory requirement and a recommendation are specified for the same criterion, the recommendation represents a goal currently identifiable as having distinct compatibility or performance advantages.

Suggestions for improvement of this document are welcome. They should be sent to the Alliance for Telecommunications Industry Solutions, PRQC, 1200 G Street NW, Suite 500, Washington, DC 20005.

At the time of consensus on this document, PRQC, which was responsible for its development, had the following leadership:

P. Tarapore, PRQC Chair (AT&T)
J. Colombo, PRQC Vice-Chair (Verizon)
A. Webster, Technical Editor (US Department of Commerce)
A. Nguyen, Technical Editor (National Communications Systems)
C. Underkoffler, ATIS Chief Editor

Active Participants:

K. Biholar
J. Colombo
C. Dvorak
G. Linnell
A. Nguyen
P. Tarapore
A. Webster
O. Lima
S. Makris

Table of Contents

Table of Figures

Table of Tables

Executive Summary

Voice over Internet Protocol (VoIP) Quality of Service (QoS) may worsen during times of congestion. There are many factors affecting QoS, including: packet loss, packet retransmission, packet delays, jitter, out-of-sequence packets, codec selection, and noise levels. Network congestion at routers leads to increased delays and potentially lost packets that may become compounded when the need for cryptographic services is present. Network security mechanisms such as IP Security (IPSec) and Secure Real-Time Transport Protocol (SRTP) introduce overhead to packets and may cause new points of potential congestion, such as the ingress/egress of an IPSec tunnel.

In order to understand the impact of adding network security services to VoIP calls, network modeling was conducted using a variety of voice codecs across three scenarios: 1) no congestion (baseline); 2) congestion at three times the baseline traffic; and 3) congestion at three times the baseline traffic while security services are used for calls. Mean Opinion Score (MOS) as estimated by the E-Model is used extensively to measure QoS throughout the modeling efforts for a select number of the modeled calls in the system. MOS is a subjective measure of call quality while the E-Model is a numerical approximation of MOS. It was demonstrated that low-bandwidth codecs are more easily able to withstand the effects of congestion, as well as the security measures, at the cost of lower initial call quality. In fact, the highest quality modeled codec, G.711, had the lowest final MOS when using security under congestion. Additional modeling showed that with increased priority weight at routers and priority queuing for IPSec cryptography engines, even G.711 could meet the Class 1 QoS requirements listed in ITU-T Y.1541. Class 1 QoS is defined as having an upper delay bound of no greater than 400 ms and packet loss no greater than 0.1%.

While network security mechanisms are not always pertinent for VoIP users, they may be important for users such as those in the United States National Security and Emergency Preparedness (NS/EP) community. Providing mechanisms which may improve voice QoS for certain high-priority calls during national security events and other periods of high congestion is essential for ensuring that those calls complete successfully. Three solutions are identified that could provide acceptable QoS to specific calls:

1. Increasing queuing priority at routers over the value held for public VoIP traffic.
2. Establishing priority queuing at both the IPSec ingress and egress.

   NOTE: This may require the copying of the Differentiated Services Code Point (DSCP) from the inner header to the IPSec header when entering an IPSec tunnel, so that the priority markings may be accessed by the egress cryptographic engine.
3. Adapting codec selection and the use of security services depending on the immediate conditions of the network. For instance, it may be worthwhile to switch to a low-bandwidth codec when facing serious congestion.

Implementing just the first two solutions in the modeling showed that packet loss could be reduiced from over 58% to 0.10%.

.

American National Standard for Telecommunications on –

Media Plane Performance Security Impairments for Evolving VoIP/Multimedia Networks

1 Introduction

Service quality in packet-based networks can be negatively affected by numerous conditions such as congestion, link bit error rates, or use of various security mechanisms. This document discusses the impacts of implementing security standards, including an evaluation by estimated mean opinion score (MOS) in a network model.

An overview of the purpose of this document is given in section 2, while section 3 lists all references from standards and provides a table of acronyms. Section 4 provides light background on the role that Quality of Service (QoS) plays in media services. Section 5 highlights modeling results from a simulation involving MOS and cryptography services for a standard voice call both with and without congestion. Finally, section 6 discusses Next Generation Network (NGN) National Security and Emergency Preparedness (NS/EP) services and highlights several proposed changes to support them. In addition, there are four annexes designed to provide a good set of background information on QoS and security services.

2 Scope, Purpose, & Application

This document focuses primarily on media flow performance, specifically the transfer of voice, all forms of video, high fidelity audio, and other information that is time sensitive. Voice over Internet Protocol (VoIP) and multimedia applications are evolving to use NGNs that are packet-based. These services use multiple broadband QoS-enabled transport technologies where service related functions are part of the service stratum and are independent of technologies in the transport spectrum.

Section 6 discusses NGN NS/EP services and how new proposed security features could improve their performance and survivability during heavy network congestion.

The NGN architecture is defined in ITU-T Y.2011 and ATIS-1000018, *NGN Architecture*. The performance objectives for Internet Protocol (IP)-based services are described in ITU-T Y.1541. With the use of resource and admission control functions defined in ITU-T Y.2111, QoS-related decisions are made based upon service-level agreements, service priority and multiple domains.

ATIS-0100014, *Information and Communications Security for NGN Converged Services IP Networks and Infrastructure*, discusses the security requirements and mechanisms to be considered in offering NGN services. The use of the security mechanisms discussed in the ATIS-0100014, such as various forms of encryption, impact the QoS objectives and requirements for offered services. This document addresses the selection of appropriate security support while meeting the performance (QoS) objectives, especially with consideration for NGN NS/EP services, which aim to provide priority telecommunications capabilities over NGN networks to personnel involved with national security or public safety. It is based off of the legacy Government Emergency Telecommunications Service (GETS), which was created in the mid-90s to provide the same capability over the Public Switched Telephone Network (PSTN).

3 Definitions, Acronyms, & References

3.1 Definitions

This document uses terms as defined in ATIS-0100014. In addition, the following definitions are used in this document:

Media Plane	The data plane in which one or more media streams and their associated media control protocols are exchanged after a media connection has been created by the exchange of signaling messages in the Signaling Plane. (IETF Internet Draft, *Session Initiation Protocol Benchmarking Terminology*, 2011)
Quality of Service	1. The performance specification of a communications channel or system. NOTE: QOS may be quantitatively indicated by channel or system performance parameters, such as signal-to-noise ratio (S/N), bit error ratio (BER), message throughput rate, and call blocking probability. 2. A subjective rating of telephone communications quality in which listeners judge transmissions by qualifiers, such as excellent, good, fair, poor, or unsatisfactory. (*ATIS Telecom Glossary*).

Also, this document makes use of the term “Mean Opinion Score” as defined in ITU-T G.107. More information regarding ITU-T G.107 can be found in Section 4.2 of this document.

3.2 Glossary of Acronyms

3DES	Triple Data Encryption Standard
AES	Advanced Encryption Standard
ANSI	American National Standard Institute
ATIS	Alliance for Telecommunications Industry Solutions
CA	Certificate Authority
CCITT	International Telegraph and Telephone Consultative Committee
DS	Differentiated Services
DSCP	DiffServ Code Point
DES	Data Encryption Standard
DH	Diffie-Hellman
DTLS	Datagram Transport Layer Security
ETS	Emergency Telecommunications Service
GETS	Government Emergency Telecommunications Service
HTTP	Hypertext Transport Protocol
ICMP	Internet Control Message Protocol
IETF	Internet Engineering Task Force
IKE	Internet Key Exchange
IP	Internet Protocol
IDPS	Intrusion Detection and Prevention Systems
IPSec	Internet Protocol Security

ITU-T	International Telecommunication Union Telecommunication Standardization Sector
LDAP	Lightweight Directory Access Protocol
MOS	Mean Opinion Score
MPEG	Moving Picture Experts Group
NAT	Network Address Translation
NCS	National Communications System
NGN	Next Generation Network
NIST	National Institute of Science and Technology
NNI	Network-Network Interface
NRS	Network Reliability and Security
NSA	National Security Agency
NS/EP	National Security and Emergency Preparedness
NTIA	National Telecommunications and Information Administration
PKI	Public Key Infrastructure
PRQC	Performance, Reliability and Quality of Service Committee
PRSSC	Performance, Reliability, and Signal processing Standards Committee
PSTN	Public Switched Telephone Network
QoS	Quality of Service
RADIUS	Remote Authentication Dial In User Service
RFC	Request for Comments
RSA	Rivest, Shamir, and Adleman
RTP	Real-time Transport Protocol
S/MIME	Secure Multipart Internet Mail Extensions
SBC	Session Border Control
SHA	Secure Hash Algorithm
SIP	Session Initiation Protocol
SSH	Secure Shell
SSL	Secure Socket Layers
SRTP	Secure Real-Time Transfer Protocol
TCP	Transmission Control Protocol
TLS	Transport Layer Security
UDP	User Datagram Protocol
UNI	User-Network Interface
VoIP	Voice over Internet Protocol
VPN	Virtual Private Network
ZRTP	Zimmermann Real-time Transport Protocol

3.3 Normative Reference Documents

The following standards contain provisions which, through reference in this text, constitute provisions of this Standard. At the time of publication, the editions indicated were valid. All standards are subject to revision, and parties to agreements based on this Standard are encouraged to investigate the possibility of applying the most recent editions of the standards indicated below.

3.3.1 ATIS Documents[1]

ATIS-0100009, *Overview of Standards in Support of Emergency Telecommunications Service (ETS)* (2006).

ATIS-0100014, *Information and Communications Security for NGN Converged Services IP Networks and Infrastructure* (2007).

ATIS-1000018, *NGN Architecture* (2007).

ATIS-1000020, *ETS Packet Priority for IP NNI Interfaces – Requirements for a Separate Expedited Forwarding Mechanism* (2007).

3.3.2 Internet Engineering Task Force (IETF) Documents[2]

RFC 2267, *Network Ingress Filtering: Defeating Denial of Service Attack which employ IP Source Address Spoofing* (1998).

RFC 2631, *Diffie-Hellman Key Agreement Method* (1999).

RFC 2865, *Remote Authentication Dial In User Service (RADIUS)* (2000).

RFC 3174, *US Secure Hash Algorithm 1 (SHA1)* (2001).

RFC 3588, *Diameter Base Protocol* (2003).

RFC 3711, *The Secure Real-Time Transport Protocol (SRTP)* (2004).

RFC 4251, *The Secure Shell (SSH) Protocol Architecture* (2006).

RFC 4301, *Security Architecture for the Internet Protocol* (2005).

RFC 4347, *Datagram Transport Layer Security* (2006).

RFC 4566, *Session Description Protocol* (2006).

RFC 5246, *The Transport Layer Security (TLS) Protocol Version 1.2* (2008).

RFC 5750, *Secure/Multipurpose Internet Mail Extensions (S/MIME) Version 3.2 Certificate Handling* (2010).

RFC 5751, *Secure/Multipurpose Internet Mail Extensions (S/MIME) Version 3.2 Message Specification* (2010).

RFC 5853, *Requirements from Session Initiation Protocol (SIP) Session Border Control (SBC) Deployments* (2010).

RFC 6189, *ZRTP: Media Path Key Agreement for Unicast Secure RTP* (2011).

3.3.3 ITU-T Documents[3]

G.107, *The E-Model: a Computational Model for use in Transmission Planning, Edition 7* (2009).

G.108, *Application of the E-Model: A Planning Guide, Edition 1* (1999).

G.113, *Transmission Impairments Due to Speech Processing* (2007).

---

[1] These documents are available from the Alliance for Telecommunications Industry Solutions (ATIS), 1200 G Street N.W., Suite 500, Washington, DC 20005. < https://www.atis.org/docstore/default.aspx >

[2] These documents are available from the Internet Engineering Task Force (IETF). < http://www.ietf.org >

[3] These documents are available from the International Telecommunications Union. < http://www.itu.int/ITU-T/ >

G.711, *Pulse Code Modulation of Voice Frequencies* (1988).

G.723, *Dual Rate Speech Coder for Multimedia Communications Transmitting at 5.3 and 6.3 Kbit/s* (1988).

G.728, *Coding of Speech at 16 Kbit/s using Low-Delay Code Excited Linear Prediction* (1992).

G.729, *Coding of Speech at 8 Kbit/s using Conjugate-Structure Algebraic-Code-Excited Linear Prediction* (CS-ACELP) (2007).

P.800, *Methods for Objective and Subjective Assessment of Quality* (2003).

Q.SUP62, *Overview of the Work of Standards Development Organizations and Other Organizations on Emergency Telecommunications Service* (2011).

X.800, *Security Architecture for Open Systems Interconnection for CCITT Applications* (1991).

X.902, *Information Technology – Open Distributed Processing – Reference Model: Foundations* (2009).

Y.1541, *Network Performance Objectives for IP-Based Services* (2006).

Y.2011, *General Principles and General Reference Model for NGNs* (2004).

Y.2111, *Resource and Admission Control Functions in NGNs* (2008).

3.4 Informative Reference Documents

3.4.1 ATIS Documents[1]

ATIS-0100022.2008, *Priority Classification Levels for Next Generation Networks*.

ATIS-0100024.2009, *User to Network (UNI) Media Plan Security Standard for Evolving Network*.

ATIS-0300074.2009, *Guidelines and Requirements for Security Management Systems*.

ATIS-0300276.2008, *Operations, Administration, Maintenance and Provisioning Security Requirements for the Public Telecommunications Network: A Baseline of Security Requirements for the Management Plane*.

ATIS-1000001, *Security for NGN – An End User Perspective* (2004).

ATIS-1000007.2006 (R2011), *Generic Control and Signaling Plane Security Requirements for Evolving Networks*.

ATIS-1000012.2006 (R2011), *Signaling System #7 Network and NNI Interconnection Security Requirements and Guideline*.

ATIS-1000019.2007 (R2012), *NNI Standard for Signaling and Control Security for Evolving Networks*.

ATIS-1000024, *Security Roadmap* (2008).

ATIS-1000025.2008, *UNI Interface Standard for Signaling and Control Security Requirements of Evolving Networks*.

ATIS-1000029.2008, *ATIS NGN Security Requirements*.

ATIS-1000030.2008, *Authentication and Authorization Requirements for NGN*.

3.4.2 Internet Engineering Task Force Documents[2]

RFC 0768, *User Datagram Protocol* (1980).

RFC 0791, *Internet Protocol* (1981).

RFC 0792, *Internet Control Message Protocol* (1981).

RFC 0793, *Transmission Control Protocol* (1981).

RFC 2560, *Internet Public Key Infrastructure Online Certificate Status Protocol* (1999).

RFC 2198, *RTP Payload for Redundant Audio Data* (1997).

RFC 3545, *Enhanced Compressed RTP for Links with High Delay, Packet Loss and Reordering* (2003).

RFC 3550, *RTP: A Transport Protocol for Real-Time Applications* (2003).

RFC 3984, *RTP Payload Format for H.264 Video* (2005).

RFC 4120, *The Kerberos Network Authentication Service (V5)* (2005).

RFC 4298, *RTP Payload Format for BroadVoice Speech Codecs* (2005).

RFC 4302, *IP Authentication Header* (2005).

RFC 4303, *IP Encapsulating Security Payload (ESP)* (2005).

RFC 4306, *Internet Key Exchange (IKEv2) Protocol* (2005).

RFC 4346, *The Transport Layer Security (TLS) Protocol Version 1.1* (2006).

RFC 4383, *The Use of Timed Efficient Stream Loss-Tolerant Authentication (TESLA) in the SRTP* (2006).

RFC 4421, *RTP Payload Format for Uncompressed Video: Additional Colour Sampling Modes* (2006).

RFC 4425, *RTP Payload Format for Video Codec 1 (VC-1)* (2006).

RFC 4513, *Lightweight Directory Access Protocol (LDAP): Authentication Methods and Security Mechanism* (2006).

RFC 4568, *Session Description Protocol Security Descriptions for Media Streams* (2006).

RFC 4587, *RTP Payload Format for H.261 Video Streams* (2006).

RFC 4588, *RTP Retransmission Payload Format* (2006).

RFC 4598, *Real-time Transport Protocol (RTP) Payload Format for Enhanced AC-3 (E-AC-3) Audio* (2006).

RFC 4629, *RTP Payload Format for ITU-T Rec. H.263 Video* (2007).

RFC 4771, *Integrity Transform Carrying Roll-Over Counter for the SRTP* (2007).

RFC 5027, *Security Preconditions for Session Description Protocol Media Streams* (2007).

RFC 5124, *Extended Secure RTP Profile for Real-time Transport Control Protocol* (2008).

3.4.3 ITU-T Documents[3]

X.509, *Information Technology – Open Systems Interconnection – The Directory: Public-Key and Attribute Certificate Frameworks* (2000).

Y.1271, *Framework(s) on Network Requirements and Capabilities to Support Emergency Telecommunications over Evolving Circuit Switched and Packet Switched Networks* (2004).

Y.2001, *General Overview of NGN* (2004).

Y.2171, *Admission Control Priority Levels in NGNs* (2006).

Y.2172, *Service Restoration Priority Levels in IP Networks* (2007).

Y.2201, *NGN Release 1 Requirements* (2007).

Y.2701, *Security Requirements for NGN Release 1* (2007).

4 Performance Measures & Quality of Service (QoS) Requirements

QoS is a performance measurement that encompasses the many aspects of an end-to-end connection that influence the service quality as perceived by the end user(s). It is an important measurement when considering interactive traffic such as VoIP and videoconferencing, since it is difficult for such services to buffer their stream without introducing intolerable delays. Codec selection, network congestion, network topology, network services (such as encryption and authentication), and even the room noise at either end when considering audio may all have a measureable impact on QoS.

4.1 Quality of Service Impairment Overview

There are a number of impairments in a packet network that can impact QoS. These are:

- Packet Loss/Rejection
- Packet Retransmission
- Packet Delays
- Jitter
- Packet Out-of-Sequence
- Codec Selection
- Noise Levels

Each of these types of impairments can render media traffic unusable, thus impacting the end user experience. The cause of these impairments may stem from various reasons such as the network issues based on the technology, traffic congestion in an overloaded network, loss of facilities, and implementation choices (including security mechanisms). For a complete description of each of these impairments, please refer to Annex A.

This document addresses only the QoS impacts from the implementation of security mechanisms. In addition, the current document only quantifies impact for IPSec and SRTP implementations, which are foremost protocols that can provide security to streaming voice and video services.

4.2 Mean Opinion Score (MOS) & the E-Model (ITU-T G.107, G.108, & G.113)

When considering voice communications, a leading measure of QoS is Mean Opinion Score (MOS), which is a scaled average judgment of call quality. While subjective by nature, there have been large amounts of interest in creating analytical approximations of MOS. ITU-T G.107 describes the E-Model, which is one of the most accepted computation models for MOS. ITU-T G.108 further provides instructions and parameters for tailoring the model to specific scenarios and ITU-T G.113 provides up to date E-Model parameters for a variety of speech processing codecs. There are five primary components to the E-Model:

1) *Basic Signal-to-Noise Ratio*: This is essentially a measurement of the noise present in an end-to-end connection, including the room noise at both ends and circuit noise.

2) *Equipment Impairment*: The codec selected for the voice transmission will have a static impairment factor as stated in ITU-T G.108. This impairment accounts for most of the quality loss due to codec delay, poor sampling rates, compression artifacts, and so forth.

3) *Simultaneous Impairment Factor*: This impairment consists of all impairments which occur in step with the voice transmission. Certain aspects of this factor may instead be incorporated into the Equipment Impairment value.

4) *Delay Impairment Factor*: Delays between the speaker and listener contribute to a decrease in call quality. The delay of the echo-path is important to consider in addition to typical end-to-end delay since it may lead to voice artifacts. There are many types of delays in an end-to-end voice transmission, including but not limited to queuing delay of packets at routers, time on wire, encryption delay, and codec delay.

5) *Advantage Factor*: There is currently work being done to address the advantage factor, which accounts for reasonable expectations for a type of service. For instance, one could expect better performance on a wireline connection than on a multi-hop satellite connection.

The E-Model is best used as a tool to predict worst-case MOS instead of as a process to analytically determine an MOS score. The E-Model linearly adds the effects of codec impairment and delay, which humans cannot do in subjective MOS measurements, but is still an appropriate tool. There are several proposed improvements to the E-Model, such as bursty packet loss that has been shown to be a major contributor to poor quality since it is difficult for any codec or service to mask.

When considering practical modeling and network mechanism operation, it is also useful to consider QoS classification bounds. ITU-T Y.1541 defines eight classes of QoS for use with varied telecommunications services. These eight classes describe packet loss, jitter, and delay bounds for service provider networks, not including user networks or user devices.

5 Security Mechanism Considerations & Modeling

It is imperative to provide security for some, but not all, NS/EP traffic flows. In some cases, security can be relaxed, if necessary, when the communication is vital. In other cases, the need for security is uncompromising. There are often trade-offs that can be made with QoS to maintain security and still achieve adequate communication. Clearly the principles we are applying to NS/EP traffic should also be applicable to non-NS/EP traffic flows that utilize QoS and/or security mechanisms. The use of any security protocol will, at a minimum, cause an increase in end-to-end delay and jitter; this delay is caused by the time it takes for cryptographic processing and also the queuing for that processing at the cryptographic endpoints. In certain cases, packet loss and possible retransmission will occur due to overflows at buffers caused by congestion at cryptographic engines. There should be no artifacts formed due to cryptography since it does not create packet loss like some compression methods.

Modeling work was performed to produce a quantitative assessment of the impact on MOS that using security could have when dealing with voice calls. A sample network comprised of a simple core and two wireless access networks was produced along with enough mixed discrete traffic to represent a standard busy hour load on each link in order to establish a baseline. Two additional scenarios show how public voice calls would perform during a congestion event (three times the amount of mixed traffic). For more information regarding modeling parameters, please refer to Annex E.

Six codecs with various characteristics were evaluated in order to capture a wide range of varying performance. Using the E-Model in scenarios with and without security allows one to see how call quality is affected by the services. Up to date codec impairment values are pulled from ITU-T G.113 and all other codec parameters come from their respective standards. Codec impairment values, algorithmic delays, payload lengths, and packets per second are included in Table 1 for each of the six codecs. End-to-end delay and packet loss are the primary alterations when moving from no security to its inclusion in the model under congestive circumstances, since there are no other artifacts added by encryption and authentication unless traffic is dropped by the cryptography engine. The time to encrypt a packet is primarily a function of its cryptographic payload and congestion present at the cryptographic engine.

Codec impairment is not the only way that codecs differ amongst themselves. Delay is also introduced algorithmically by the sampling, compression, decompression, and reconstruction processes. For instance, the algorithmic delay difference between G.711 and G.723.1 is over 60 ms, which is a very large portion of end-to-end delay for voice. The effect of additional delay is more severe as total delay

increases; therefore, codecs with high algorithmic delay may perform poorly for reasons other than their codec impairment. Encryption involves a portion of processing that is independent from payload length; therefore, codecs which produce a large number of packets per second tend to be more easily impacted by cryptographic congestion.

5.1 Modeling Description & Results

Table 1 shows codec specifics and the modeling results for each of the three scenarios:

1. Scenario 1 is the baseline, which features a typical busy hour traffic volume that does not cause congestion of any kind within the network. The volume of traffic is best defined as a mix of traffic types and codec selections that attempts to best replicate the conditions on a live network at the packet level. Network entities and links are between thirty and seventy percent utilized at this baseline traffic volume.
2. Scenario 2 shows the effects of adding moderate packet congestion to the baseline network. The additional packet congestion amounted to a number of calls equal to three times the baseline figure. Each traffic type was increased by the same proportion. The increased traffic crosses the network at a steady rate and any loss should be considered random.
3. The final scenario adds both IPSec and SRTP security services to a relatively small sample of calls in order to see if there is a noticeable performance decrease from enabling the use of cryptography. The IPSec tunnel makes use of 3DES encryption and SRTP includes the Authentication Header. Both of the IPSec cryptography points in the network are designed to be moderately congested at onset at approximately forty percent. Security mechanisms were added to enough voice calls to incur long processing delays and cryptographic loss.

Table 1 - Various Codec Performance with and without Security Mechanisms

	G.711	G.728	G.729	G.729a	G.723.1	GSM FR
Codec Impairment	0	7	10	11	19	20
Payload Bandwidth (bps)	64000	16000	8000	8000	5300	13200
Payload Length (Bytes)	160	20	20	20	20	33
Average Encryption Time (ms)	14	7.5	7.5	7.5	7.5	7.8
Packets per Second	50	100	50	50	33	50
Scenario 1 - Baseline Traffic without Security Mechanisms						
Baseline MOS without Security	4.35	4.15	4.05	4.01	3.67	3.63
Baseline Grade without Security	Very Satisfied	Satisfied	Satisfied	Satisfied	Some Dstfd.	Some Dstfd.
Scenario 2 - Congestion (3 Times Load) without Security Mechanisms						
Congested MOS without Security	2.64	4.15	4.04	4.00	3.67	3.62
Congested Grade without Security	All Dstfd.	Satisfied	Satisfied	Satisfied	Some Dstfd.	Some Dstfd.
MOS difference from Scenario 1 (%)	-39.46%	*negligible*	*negligible*	*negligible*	*negligible*	*negligible*
Scenario 3 - Congestion (3 Times Load) with sRTP (with Authentication) & 3DES IPSec						
Congested MOS with sRTP & IPSec	2.12	3.06	3.88	3.84	3.65	3.20
Congested Grade with sRTP & IPSec	Not Rcmnd.	Most Dstfd.	Some Dstfd.	Some Dstfd.	Some Dstfd.	Most Dstfd.
MOS difference from Scenario 2 (%)	-19.53%	-26.31%	-4.79%	-4.17%	*negligible*	-11.83%

The differences in MOS between the various codecs at baseline busy hour traffic levels is primarily due to the varying codec impairments, since there is no congestion present in the network that would cause long delays or packet loss. Environmental issues could lead to a decrease in MOS, although they are not taken into account for the modeling effort.

Scenario 2 demonstrates that high bandwidth codecs may be the first to degrade as congestion increases, since routers that treat traffic fairly will slow them down considerably compared to lower bandwidth ones. In this case, the highest MOS codec (G.711) at baseline traffic volumes becomes poorly performing once routers have to start fairly queuing traffic amongst the active sessions.

G.711 continues to perform very poorly when security services are enabled. Excluding G.711, G.728 has a significantly lower MOS than the other codecs since it has the highest number of packets per second by a multiple of two. The number of packets per second is an important consideration when using security services, since not only will the amount of overall overhead increase, but static processing delays at cryptographic endpoints will compound quickly.

Of all the codecs modeled, G.723.1 is the only one that is not affected by either the packet or cryptographic congestion. G.723.1 performs well in that regard since it is the lowest bandwidth codec in the group and it produces the fewest packets per second. That said, other codecs still had a more favorable MOS in each of the scenarios since G.723.1 begins with a high codec impairment.

5.2 Priority Service Considerations

Further modeling was performed in order to establish measures needed to ensure that voice QoS for priority calls (including ETS calls) meets the requirements set forth in ATIS-1000020, which suggests Class 1 QoS as listed in ITU-T Y.1541 for ETS calls specifically. Class 1 QoS according to ITU-T Y.1541 is defined as having an upper delay bound of no greater than 400 ms and packet loss no greater than 0.1%.

The results from the scenarios discussed in section 5.1 demonstrate that QoS parameters must grant greater priority to priority calls in certain cases of congestion in order to meet both the delay and packet loss thresholds. For instance, the explicit minimum configuration would depend entirely on the congestion scenario – G.711 suffered an average packet loss of 68.22% in Scenario 3, which is substantially greater than the Class 1 threshold. The severe packet loss can be attributed to the lack of priority at cryptography engines as well as the insufficient priority at each hop as there were many other VoIP packets to contend with at router interfaces.

In order to reduce packet loss for VoIP calls involving cryptography (Scenario 3 from section 5.1) to less than 0.1%, two changes had to occur within the network. First, queuing priority at routers had to approximately triple from the original point established for public VoIP traffic in the model. Second, priority queuing had to exist at both the IPSec ingress and egress, which could also require implementation of a mechanism such as the one discussed in section 6.2 of this document. Table 2 shows the modeling results from the addition of these two changes to the sample calls of Scenario 3.

Table 2 - Modeling Results for G.711 to Meet Class 1 QoS Requirements

Statistic Classification	Statistic Name	Y.1541 Class 1 QoS Threshold	Scenario 3 Result (Public)	Result to Meet Class 1 QoS	Improvement (%)
Y.1541 QoS	**Packet Loss (%)**	0.10%	59.14% (Over Threshold)	0.10%	99.83%
Classificiation	**Maximum Delay (ms)**	400.00	573.97 (Over Threshold)	117.19	79.58%
Statistic	**Average Jitter (ms)**	50.00	26.39	3.78	85.68%
Other	**Average Delay (ms)**	*N/A*	466.48	24.36	94.78%
QoS	**Mean Opinion Score**	*N/A*	2.12	4.36	105.66%
Measurement	**MOS Grade**	*N/A*	Not Recommended	Very Satisfied	*N/A*

Modeling results demonstrate that with sufficient queuing priority, thresholds for Class 1 QoS are able to be met even by the G.711, which was the worst-performing codec in Scenario 3 when considering MOS. If maximum delay was still above the upper bound of 400.0 ms after ensuring packet loss constraints are met, then further increasing queuing priority would reduce it appropriately. In addition, other network configurations would require a different set of parameters and it would be up to the service provider to configure their network properly. Section 6.2 further shows the QoS impact of excluding IPSec egress priority queuing.

6 Priority Services

NGN National Security and Emergency Preparedness (NS/EP) services are part of a United States directive to allow for the provisioning of telecommunications services for critical personnel under all circumstances. The United States National Communications System (NCS) manages these services and is in charge of ensuring that certain criteria for service acceptability are met. Part of this process involves researching, modeling, and suggesting mechanisms to industry that are needed for assured quality of experience.

6.1 National Security & Emergency Preparedness Overview

NGN NS/EP services make use of the IP Multimedia Subsystem (IMS), which is an architecture designed for the interoperability of media services over an IP-based core. IMS uses Session Initiation Protocol (SIP) for session negotiation and involves many functional entities for session setup, accounting services, QoS management, and interoperability. Points of possible congestion have been identified in the architecture and the NCS has written Government Industry Requirements (IR) to include necessary priority mechanisms for NGN NS/EP calls. The Government IR does not provide implementation details for any of the mechanisms, although it can be assumed that some form of priority queuing at routers would be necessary in order to meet QoS requirements during congestion as is illustrated in the section 5 modeling results.

6.2 IP Security Cryptography with DSCP Support

IP Security (IPSec) tunnel decryption presents a unique challenge for NGN NS/EP calls, since the priority markings on the packet are encrypted and unable to be used in order to facilitate priority queuing at the cryptography engine. While RFC 2475 states that the inner header's Differentiated Services (DS) Code Point (DSCP) must not be changed at any point within a DS domain, it only says that the network may optionally copy the inner DSCP to the outer IPSec header.

In order to allow priority queuing to take place at the cryptographic end-point, it is necessary to copy the DSCP from the inner packet to the IPSec header at the initial tunnel entry point. With proper network security services, a DSCP in the IPSec header copied from the inner header will allow the cryptography engine to prioritize the queuing of NGN NS/EP traffic during congestion periods. While this is currently optional in the RFCs (2474 and 2475), text revisions to make it mandatory for all IPSec tunnels supporting NGN NS/EP traffic would serve as a great benefit for priority voice services and later video and data services. This method makes no promise of securely transmitting the DSCP via the IPSec header and such a responsibility falls to the security procedures of the DS domain itself as it does for a DSCP in a plain IP header.

A consideration for DSCPs within IPSec tunnels is that the inner DSCP remains unchanged so that the tunnel may operate as a single virtual hop. In order to operate across DS domains, the outer header DSCP is permitted to be changed at the ingress point currently for all traffic (section 6 of RFC 2475). There must be additional provisions for NGN NS/EP traffic to use the same (or comparable) DSCP across domains so there needs to be a mechanism for DSCP translation for NGN NS/EP traffic unless all tunnels carrying NGN NS/EP media reside within the same DS domain.

Table 3, below, shows modeling results from a sample scenario where the NS/EP priority calls receive the copied DSCP at the ingress of the IPSec tunnel while others do not. When these calls do receive the copied DSCP value, their performance reflects the priority performance shown in Table 2, which makes use of the same priority queuing and traffic configuration across the network. As before, three times the normal amount of cryptographic traffic is delivered to the egress point while it is engineered to be at approximately 40% load. This configuration causes moderate packet loss of approximately one-tenth of all packets for priority calls that do not receive the DSCP treatment. Changing the processing capabilities of the cryptography engine or its queue length could cause varying results, as would changing the input load for the scenario. The actual copying of the DSCP is assumed to take minimal processor resources at the ingress point.

Table 3 - G.711 with and without IPSec Egress Priority During Congestion

Statistic Classification	Statistic Name	Y.1541 Class 1 QoS Threshold	Result without DSCP Copying	Result with DSCP Copying	Improvement (%)
Y.1541 QoS	Packet Loss (%)	0.10%	8.69% (Over Threshold)	0.10%	98.85%
Classificiation	Maximum Delay (ms)	400.00	268.21	117.19	56.31%
Statistic	Average Jitter (ms)	50.00	13.38	3.78	71.75%
Other	Average Delay (ms)	*N/A*	105.80	24.36	76.98%
QoS	Mean Opinion Score	*N/A*	4.23	4.36	3.14%
Measurement	MOS Grade	*N/A*	Satisfied	Very Satisfied	*N/A*

The results for G.711 show that granting the proper DSCP to packets transmitted through an IPSec tunnel can improve call quality even in cases of light congestion. In this case, the MOS grade increased from “Satisfied” to “Very Satisfied” when DSCP copying occurred.

There may be scenarios where a realistic priority queuing configuration would not be able to grant sufficient QoS to priority calls in order to meet requirements. The next section discusses a possible mechanism for selecting a codec that could meet Class 1 delay and packet loss constraints in light of network congestion; this would allow meeting QoS goals in a larger variety of situations without dynamically altering priority weights according to present congestion.

6.3 Quality of Service Adaptation

In order to meet established QoS goals, NGN networks may need to adjust portions of their operation depending on congestion. Despite many priority mechanisms which should allow for most priority calls to obtain a desirable QoS, there may arise congestion conditions which result in calls performing worse than needed. In such cases, the network could enact several different policies, which could each lessen the severity of the degraded service. Apart from possible queuing or retrying until a QoS target is obtainable, media codecs and security services could be altered for one session or a group of sessions to facilitate the agreed upon QoS.

6.3.1 NGN NS/EP QoS Threshold

Comparing the selected QoS threshold to the current conditions for a call requires a model for MOS, since it is actually a subjective score. While the E-Model is a generally accepted method for calculating QoS, it is expensive to compute and requires information which is difficult to collect, such as the noise levels at the sending and receiving locations. An alternative to establishing an actual MOS for voice quality would be to have individual thresholds for such metrics as packet loss, end-to-end delay, and jitter, and then send alerts when one or more of those are breached. An additional benefit is that such a method could be expanded to suit video and data services as well and not just voice.

Ideally, MOS for a phone call should not drop below 4.0, which is the boundary between “satisfied” and “some users dissatisfied”, although it should be suitable to lower the threshold to 3.5 for cases of heavy network congestion. 3.5 is more practical since it would permit the use of the G.729, G.729a, and G.723.1 codecs, all of which have codec impairments that would result in a MOS lower than 4.0 even in cases of no or light congestion. While an MOS of 3.5 is still in the range of “Most Users Dissatisfied”, it would be acceptable for priority users during periods of extreme congestion.

While an E-Model QoS goal of 3.5 is desirable, it is impractical to derive such a figure in a deployed network as stated above, so for the purposes of this document, the target QoS goals for NGN NS/EP voice traffic are as follows: no more than a 400 ms maximum delay bound end-to-end, no more than 50 ms of jitter, and no more than 0.1% packet loss, which follow the requirements set by Y.1541 for Class 1 QoS. The delay budget accounts for the delay within the transit networks, ignoring any delay that may reside within user networks or in their handset. User-specific delays are not taken into

account, since service providers are unable to provide performance guarantees for network entities outside of their network and those networks that they have peering agreements with. In reality, handset and codec delays could add up to about forty milliseconds of delay, depending on the codec.

For illustrative purposes, the following text assumes a 400 ms mouth-to-ear delay budget, which is the boundary for Y.1541 Class 1 QoS. When there is no congestion present in the network, the budget is primarily spent on propagation delay and codec delay, while other delays such as processing and emission delays are minimal. Using encryption will cause additional delay dependent on packet size and cryptography algorithms. Heavy cryptographic congestion will cause additional delays when using encryption. Heavy network packet congestion will cause processing and queuing delays to increase. All types of congestion may lead to packet loss, which can cause large drops in call quality even if delays remain low. The following table provides a notional representation of delay budget for a variety of scenarios.

Table 4 - Example Delay Budget (400 ms budget)

Delay Type	No Encryption (Baseline Values)	Encryption			
	No Congestion		Cryptographic Congestion	Packet Congestion	Overall Congestion
Cryptographic	0	15	**60**	15	50
Queuing	20	*no change*	*no change*	**210**	**190**
Propagation	100	*no change*	*no change*	*no change*	*no change*
Emission	10	*no change*	*no change*	*no change*	*no change*
Processing	10	*no change*	*no change*	15	15
Remaining	**260**	**245**	**200**	**50**	**35**

The table shows representative results for a session over a relatively slow network. Even so, it is difficult to approach the allotted 400 ms for Class 1 QoS. Delay is unlikely to increase too significantly under congestion unless queues are sized very large.

It is recognized that under severe emergency conditions with the likelihood of significant loss of network resources and bandwidth, it may not be possible to achieve the desired network performance for NS/EP services as defined in this section. Under such conditions, service completion is considered to be vital even under degraded network performance levels.

6.3.2 QoS Measurement

QoS monitoring is generally the role of a Session Border Controller, and the various roles of an Session Border Controller are divided between several functional entities in the IMS architecture. Generally in the IMS, Session Border Controller roles are performed by the Call Session Control Functions (CSCF), Policy Decision Function (PDF), Interconnection Border Control Function (IBCF), and servers in the access networks. QoS profile commitment is made by the PDF, which is generally a part of the Proxy CSCF (P-CSCF), although the policies are monitored and enforced at the Policy Enforcement Point (PEP), which may change depending on the access technology.

While there are several places in an IMS network that can measure performance data, capturing true end-to-end QoS figures is a much more difficult task. Even with data compiled in a central location

from all possible measurement points, there would still be many pieces of information missing. Tracer packets could help alleviate some of the difficulty as the cost of greater network and processing congestion. Interoperability over provider boundaries adds even more complexity, although this could be solved by aggregating captured information. Using a standard approach for QoS measurement would be beneficial for NGN NS/EP, especially since much of the task is left up to access networks which may each behave differently.

6.3.3 Methods to Ensure QoS Threshold

This proposed methodology for meeting QoS goals for NGN NS/EP voice calls over an IMS architecture has identified three possible decisions the network could make, depending on the state of the network. The first decision is whether or not to use security services. By default, they should always be used; however, if there is heavy cryptographic congestion present in the network, then the use of them may cause QoS to dip below the established threshold defined in section 6.3.1. Second, during extreme packet congestion, codecs could be switched over to lower bandwidth ones. In an IMS-based network, this decision is the responsibility of the PDF, usually a part of the P-CSCF. Lower bandwidth codecs generally have a greater codec impairment value, so switching to a lower bandwidth codec might often hurt QoS despite causing less aggregate congestion. It is likely that congestion levels would have to be well above those of expected standard fluctuations in busy hour call arrivals before the QoS tradeoffs tilt in favor of the lower bandwidth codecs. Finally, the system could queue NGN NS/EP calls if there is no other way to meet QoS targets. Each user profile should have markings indicating whether or not security services are essential, so that the system knows to queue a call if there is heavy cryptographic congestion instead of simply turning off the service.

A guiding assumption for the codec selection process is that once a call is established, then the codec will not have an opportunity to change unless there is a natural handover. Given that there is likely a fair amount of congestion on the signaling plane at the time that media resources are fully utilized, re-establishing codec selection during media plane congestion could further aggravate the possible congested state of related signaling entities. Monitoring the status of all signaling entities in the network to ensure that congestion is minimal before negotiating a codec alteration would add significant complexity to the system. It is for this reason that queuing new calls until resources are available is the desired approach when media plane resources are scarce.

6.4 Use Cases

This section describes three use cases for desired NGN NS/EP functionality under differing levels of network congestion. Results from a generic IMS model are provided to demonstrate impact. The actors in the use cases will be the originating user, terminating user, and the network for simplicity.

6.4.1 No Congestion

When the originating user placed an NGN NS/EP call attempt to the terminating user, the Session Description Protocol (SDP) was used during setup to discuss codec usage between the two parties and the network. Since there was no network congestion at the time, the network allowed the use of the G.711 codec and both the originating and terminating user agreed upon its use. In addition, an IPSec tunnel was established to safely transport media plane packets through the core network with SRTP used to secure application layer packets at the edge. Without the security mechanisms, the call would have achieved an end-to-end delay of 83 ms, well below the 400 ms threshold for NGN NS/EP voice described in section 6.3.1. With security mechanisms, the call still achieved an end-to-end delay of roughly 100 ms once the cryptographic delays were accounted for. In addition, jitter was very low and there was only a single packet dropped, which was caused by uncorrectable interference on the wire in the terminating access network. The predetermined QoS goals for the NGN NS/EP call were met since there was no congestion and the call continued at an acceptable level until completion.

6.4.2 Cryptographic Congestion

Again, when the originating NGN NS/EP user attempted to call the terminating user, the negotiation decided upon using G.711 in order to achieve performance goals since there was little packet congestion present in the network. However, the use of an IPSec tunnel depends on several factors, highlighted in the next two sub-sections.

6.4.2.1 Heavy Public Cryptographic Congestion

If the cryptographic congestion is only facing public traffic and the priority queue at the cryptography points is short or empty, then the DSCP associated with the traffic will be copied to the IPSec header, which will allow it to achieve low-latency performance amidst the public cryptographic traffic. This way, IPSec security may be used and the call can still meet its established goals. In this case, the priority media packets were able to bypass the congestion entirely (due to the existence of priority low-latency queuing at cryptography points) and achieved an end-to-end delay of 105 ms along with jitter and packet loss very similar to the no congestion use case. QoS goals were met due to a low-latency queue at cryptographic congestion points, which allowed the NGN NS/EP traffic to bypass all traffic; the decryption point utilized the copied DSCP marking on the IPSec header in order to guide the media streams packets to the correct queue.

Recommended action: DSCP markings should be copied to the IPSec header.

6.4.2.2 Heavy NGN NS/EP Cryptographic Congestion

However, if many other NGN NS/EP calls are also queuing for the cryptography engine, then it may not be feasible to use IPSec at all. Depending on the nature of the call, it may be suitable to proceed without security services entirely. If that is not the case, then the call will have to queue until resource use is low enough to permit it through.

In this use case, the priority queues at the cryptographic start and end point were so saturated that adding an addition calls demands would have resulted in an additional 60 ms end-to-end delay, which would have violated the QoS threshold. In addition, the queue was randomly dropping packets in the queue in accordance to its congestion avoidance algorithm, which would have resulted in moderate packet loss. While an IPSec tunnel would have been welcome for the call, the profile markings indicated that it was not essential; therefore, there was no need to queue the call and it was allowed through immediately and achieved an end-to-end delay of 85 ms with low jitter and negligible packet loss. QoS goals were met immediately without the need for an IPSec tunnel, which would have been the only point of congestion for the call.

Recommended action: User profile should contain indication of whether or not security is essential.

6.4.3 Heavy Network Congestion Affecting NGN NS/EP

In this final use case, there was very heavy packet congestion that affected NGN NS/EP calls and also included moderate cryptographic congestion. Priority queuing at routers was so congested that a new call using G.711 would have faced delays of almost 430 ms and packet loss of at least 8%, far above the 0.1% threshold. The packet congestion is so bad that the network decides to use the G.729 codec instead of G.711. G.729 has a much higher impairment than G.711, but the use of G.711 would have caused packet loss outweighing the impairment. Since fair queuing was being used within the network, the lower bandwidth G.729 codec was able to be used without any packet loss resulting from packet congestion. However, the profile markings for this user indicated that security was essential for this call so the system has to queue it until cryptographic resources are available. Luckily, the cryptographic load is lighter due to the lighter weight codec so the queuing only lasted for 4 seconds. Once the call was permitted, it achieved an end-to-end delay of 127 ms and only 11 ms of jitter with no packet loss. Despite very heavy network congestion, the QoS goals were met after codec negotiation and brief queuing.

Recommended action: QoS goals should be agreed upon and maintained.

7 Recommendations

This document describes research and modeling for QoS impact of adding security services to voice media flows. The modeling determined several points of possible failure within IP networks and then demonstrated the possible improvement from a selection of recommendations. The three recommendations are:

1. Providing appropriate priority queuing treatments at network routers so that media flows with priority markings may receive queuing preference over that of public traffic. This should not impact standards, but instead be recognized in service provider's best practices
2. Providing priority queuing treatments for priority users at IPSec tunnel ingress and egress points. The recommended method is the copying of DSCP values from the inner packet header to the IPSec header at the tunnel ingress.
3. Dynamically adapting codec selection and security services for priority users during periods of high congestion. This may require mechanisms within existing standard network management controls.

Annex A
(informative)

A Impairments Affecting QoS

Annex A defines many of the impairments that affect the QoS of VoIP media flows consistent with their usage throughout the rest of the document. In addition, the specific effects that each of these impairments may have on a single media flow is described. There are other types of impairments that may affect service as a whole, including service availability (downtime) and responsiveness to new requests such as call setup for voice calls.

A.1 Packet Loss/Rejection

Packet loss throughout this document is used as a percentage of the total packets involved in a single media flow that are rejected, lost, or corrupted beyond repair at some point within the transit from one user to another. Rejection often happens when queues or buffers are full at network entities and queuing logic must decide to drop certain packets. Loss, as used here, occurs when packets are routed incorrectly. Furthermore, if bit errors or tampering occurs during transit, then a packet may not be able to be read at the destination or even at hops before the destination.

When the media traffic is composed of streaming video (such as IPTV, video conferencing, or real-time networked video gaming) packet loss/rejection will cause video artifacts (i.e., image skipping, frame freezing, etc.) to occur. User acceptance of video artifacts is not high, yet will vary depending on whether the media traffic is from a subscription service or not. The effects of packet loss are higher when multiple packets get dropped in succession rather than randomly over time since it will create a very noticeable gap in the media.

When the media traffic is composed of streaming high fidelity audio (such as Internet radio or the audio portion of real-time networked video gaming) packet loss/rejection will cause audio artifacts (i.e., audio segments being repeated) to occur. User acceptance of audio artifacts is not high, yet will vary depending on whether the media traffic is from a subscription service or not.

A.2 Packet Retransmission

Whenever a packet is lost, there may be a retransmission attempt depending on the protocols used for packet transmission. These retransmission attempts may be a complete retransmission of the packet from the point of origination or even retransmission from a previous hop. Many interactive media flows do not make use of packet retransmission, since the delay introduced would greatly affect media quality at the receiving end.

Packet retransmission is often the result of lost or dropped packets and it can contribute greatly to increased delays and jitter in a network. When the media traffic is composed of streaming video (such as IPTV, video conferencing, or real-time networked video gaming) packet retransmission will cause video artifacts (i.e., tiling, frame freezing, etc.) to occur. User acceptance of video artifacts is not high, yet will vary depending on whether the media traffic is from a subscription service or not. Video media, that is downloaded prior to presentation, will not be impacted by packet retransmission since the downloading process will properly handle any retransmitted packet prior to display of the downloaded video media.

When the media traffic is composed of streaming high fidelity audio (such as Internet radio or the audio portion of real-time networked video gaming) packet retransmission will often cause audio artifacts (i.e., audio segments being repeated) to occur unless the stream is buffered or quality is otherwise ensured. User acceptance of audio artifacts is not high, yet will vary depending on whether the media traffic is from a subscription service or not. Audio media, that is downloaded prior to presentation, will not be

impacted by packet retransmission since the downloading process will properly handle any retransmitted packet prior to replay of the downloaded audio media.

A.3 Packet Delays

Two types of packet delays are discussed throughout the document: 1) average delay; and 2) maximum delay bound. Y.1541 performance bound criteria state that these values should not involve user terminals or user networks, since providers cannot feasibly affect performance within those bounds; however, use of the E-Model in the modeling effort requires that the full mouth-to-ear delay be taken into account. Average delay refers to the mean end-to-end delay of all successful packets. Maximum delay bound is the determination of the highest end-to-end delay that any single packet experiences in transit.

When the media traffic is composed of streaming video (such as IPTV, video conferencing, or real-time networked video gaming) packet delays will sometimes cause video artifacts (i.e., tiling, frame freezing, etc.) to occur on an interactive transmission such as videoconferencing. Plenty of other types of video services – such as streaming internet video – can be buffered to minimize the appearance of artifacts. User acceptance of video artifacts is not high, yet will vary depending on whether the media traffic is from a subscription service or not. Video media, that is downloaded prior to presentation, will not be impacted by packet delays since the downloading process will properly handle any delayed packet reception prior to display of the downloaded video media.

When the media traffic is composed of streaming high fidelity audio (such as Internet radio or the audio portion of real-time networked video gaming) packet delays will cause audio artifacts (i.e., audio gaps) to occur. User acceptance of audio artifacts is not high, yet will vary depending on whether the media traffic is from a subscription service or not. Audio media, that is downloaded prior to presentation, will not be impacted by packet reception delays since the downloading process will properly handle any packet reception delays prior to replay of the downloaded audio media.

A.4 Jitter

According to the *ATIS Telecom Glossary*, jitter is the “abrupt and unwanted variations of one or more signal characteristics, such as the interval between successive pulses, the amplitude of successive cycles, or the frequency or phase of successive cycles”. For the purposes of modeling in this document, jitter is calculated using a moving average of the weighted absolute value of the difference between the arriving packet’s delay and the last arriving packet’s delay.

When the media traffic includes streaming video, the packet delay variations referred to as jitter will cause artifacts where the audio may not be synchronized with the video. User acceptance of such artifacts is not high particularly when the user is paying for it (video on demand, pay-per-view). As with other causes of video artifacts, the type of service will affect their existence.

When the media traffic includes streaming high fidelity audio or a high definition VoIP call (HD Voice or wideband voice), jitter can cause degradation in the quality of the calls. The user acceptance of garbled calls making it unintelligible is not high even with non-HD voice calls and is less tolerated when a higher subscription rate is paid for the expected improved quality.

A.5 Packet Out-of-Sequence

The packets belonging to a stream may take different routes and this may result in out of sequence packets. These will have video and audio artifacts (loss of a video frame, audio drop out) with audio and video streaming.

A.6 Codec[4] Selection

For both voice and video streams, codec selection can be one of the most influential QoS factors. Different codecs exist to fit various needs in regards to sampling rate and delay, packet size and rate, bandwidth requirements, delay, and compression ratio and loss. Typically codecs with a large emphasis on compression will have a higher algorithmic delay since they must wait for additional samples and thus quality will degrade both from the increased delay and possibly lossy compression. However, low bandwidth connections may require a high-compression codec. When dealing with cryptography, codecs with large packets or high bandwidth requirements may stress the cryptographic engine at certain points in the network. Selecting a proper codec based on QoS requirements is an important step in ensuring end-to-end service quality is acceptable.

A.7 Noise Levels

According to the *ATIS Telecom Glossary*, "noise level" is "the noise power, usually relative to a reference.

> NOTE*:* Noise level is usually measured in dB for relative power or picowatts for absolute power. A suffix is added to denote a particular reference base or specific qualities of the measurement. Examples of noise-level measurement units are dBa, dBa(F1A), dBa(HA1), dBa0, dBm, dBm(psoph), dBm0, dBm0P, dBrn, dBrnC, dBrn(f_1-f_2), dBrn(144-line), pW, pWp, and pWp0".

Sending and receiving noise levels, as well as interference affecting transmission of electronic data, can have a severe impact on QoS for both voice and video streams. While there is often nothing that can be done to limit the impact that outside noise can have on a media session, QoS may be improved by moving either endpoint to a more quiet area or one with less electric interference.

---

[4] *Codec*: Acronym for coder-decoder. An electronic device that converts analog signals, such as video and voice signals, into digital form and compresses them to conserve bandwidth on a transmission path.

> NOTE: Codecs in this sense are used in this sense for video conferencing systems. (*ATIS Telecom Glossary*, 2011, Definition 3)

Annex B
(informative)

B Security Services

ITU-T Recommendation X.800 (1991) defines a security service as a mechanism which ensures adequate security of a part of a telecommunications system. Furthermore, more specific services are described that include Authentication, Authorization, Confidentiality, Integrity, and Non-Repudiation.

B.1 Authentication [5] & Authorization[6]

Authentication in telecommunications is the process of verifying that the supposed source of data is the actual source. Authorization differs in that it describes what an authenticated individual is allowed to do within a system and which resources he or she is allowed to affect. Authorization is commonly referred to as access control, since it essentially blocks access to certain processes within a system unless permission is explicitly given for the request.

B.2 Data Confidentiality

Data confidentiality is an important and wide security service which protects data from unintended disclosure. For many systems, such as voice streaming, it is often important that only the intended sender and recipient involved in a session are capable of accessing the information generated or used during that session. Confidentiality typically involves the use of encryption to conceal selective data. In some cases, the source and destination of data will also be encrypted in order to provide privacy.

B.3 Integrity[7]

This service focuses on the validity of information and the detection and prevention of attacks against it. For example, a man-in-the-middle attack may intercept a packet and change key information before forwarding it to its rightful recipient. Integrity services at the receiving end might be able to detect this attempted deception by comparing a checksum or running some other defensive procedure. As it pertains to media streaming, it is important to know that a third party is not eavesdropping on a communications session.

---

[5] *Authentication*: 1. Security measure designed to establish the validity of a transmission, message, or originator, or a means of verifying an individual's authorization to receive specific categories of information. 2. A security measure designed to protect a communications system against acceptance of a fraudulent transmission or simulation by establishing the validity of a transmission, message, or originator. 3. The verification of a claimed identity. Example: By the use of a password. (*ATIS Telecom Glossary*, 2011, Definitions 1, 2, and 4)

[6] *Authorization*: 1. The process by which an access control decision is made and enforced. 2. The granting of rights, which includes the granting of access based on access rights. 3. A mechanism that defines and controls what services or resources an entity can access. (*ATIS Telecom Glossary*, 2011, Definitions 4, 5, and 6)

[7] *Integrity*: 1. Condition existing when data is unchanged from its source and has not been accidentally or maliciously modified, altered, or destroyed. 2. The property that data has not been altered or destroyed in an unauthorized manner. 3. The state that exists when computerized data is the same as that in the source documents and has not been exposed to accidental or malicious alteration. (*ATIS Telecom Glossary*, 2011, Definitions 1, 2, and 5 of “Data Integrity”)

B.4 Nonrepudiation[8]

While less important with media streaming, nonrepudiation services notify either the sender or recipient that data was received or sent, respectively. The goal is to protect against future attempts to say that data was not sent or received as believed.

[8] *Nonrepudiation*: 1. The capability, in security systems, that guarantees that a message or data can be proven to have originated from a specific person. 2. Assurance the sender of data is provided with proof of delivery and the recipient is provided with proof of the sender's identity, so neither can later deny having processed the data. 3. The property of a receiver being able to prove that the sender of some data did in fact send the data even though the sender might later deny ever having sent it. (*ATIS Telecom Glossary*, 2011)

Annex C
(informative)

C Cryptography Algorithms

Cryptography algorithms are an essential part of a security solution that requires data confidentiality or integrity services. Most algorithms involve encryption at one point through the use of a key and then decryption at another point using the same or a different key. Any algorithm that uses the same or similar key is said to be a symmetric algorithm while others are asymmetric algorithms, which involve the public sharing of a key that creates cryptography that only the other end of the connection can decrypt using a private key. This section highlights important algorithms and considerations that pertain directly to security protocol suites that media streaming applications may use for security.

C.1 Asymmetric Cryptography

Asymmetric Cryptography involves encryption and decryption with separate keys[9]. Typically, a large random number will be the starting place for the derivation of both keys, which will be used to secure a single session and then discarded. If the public key is used to encrypt plaintext, then the paired private key will decrypt it on the other end. Often, a keying authority will be needed in order to ensure the proper use of key pairs, as is the case with Public Key Infrastructure (PKI). The major downside to asymmetric keying is that the computation is expensive compared to symmetric keying due to the key lengths used and as such it is generally not appropriate for the encryption of media data even if it is often used for establishing symmetric keys as part of a security protocol suite.

C.1.1 Asymmetric Key Exchange Algorithms

There are many widespread key exchange algorithms that make use of asymmetric keys[10]. These include Diffie-Hellman (DH) Key Agreement (RFC 2631); Rivest, Shamir, and Adleman (RSA); and Elliptical Curve algorithms. For the purposes of establishing a secure key exchange for signaling and symmetrical keys in the confines of a voice or video streaming session, DH is perhaps the most prevalent, since it is secure against eavesdroppers and utilizes potent cryptography without the need for a certification authority. DH also provides the basis for both IPSec Key Exchange (IKE) and ZRTP, which is a recently adopted keying exchange protocol in support of SRTP.

C.1.2 Public Key Infrastructure (PKI)

According to the *ATIS Telecom Glossary*, Public Key Infrastructure (PKI) is, "[a] framework established to issue, maintain, and revoke public key certificates accommodating a variety of security technologies, including the use of software. [INFOSEC-99]".

Certain encryption algorithms may make use of PKI in order to distribute certificates for the signing of encrypted data. If this is the case, then a Certificate Authority (CA) will be in charge of issuance. While

---

[9] Includes use of an *asymmetric cryptographic algorithm*, which is "a cryptographic formula that uses two related keys – a public key and a private key – each of which has the characteristic algorithm that, given the public key, it is computationally infeasible to derive the private key. [After X9.62]" (*ATIS Telecom Glossary*, 2011).

[10] Also known as *asymmetric encryption*, which is "an encryption system that utilizes two keys, one called a public key (which is known to both the sender and the recipient of encrypted data), and the other, called a private key (known only to the individual sending the data).

NOTE: Data are encrypted with the private key and decrypted with the public key. Asymmetric encryption allows for the secure transfer of data. [After Bahorsky]" (*ATIS Telecom Glossary*, 2011).

the use of such a structure is fundamentally sound, it does add additional complexity to a security scheme, which can cause longer delays, introduce new bottlenecks, and potentially expose the system to new threat opportunities. All of these issues can lead to a decrease in QoS if not handled properly.

C.2 Symmetric Cryptography

When both cryptographic endpoints use the same key or two keys related trivially, the algorithm is said to be symmetric[11]. These algorithms are generally faster than asymmetric algorithms, although there is a need for asymmetry in the sharing of keys. Symmetric cryptography is appropriate for the encryption and decryption of media packets in a streaming session. Popular symmetric algorithms include Triple Data Encryption Standard (3DES) and Advanced Encryption Standard (AES), both of which are sufficient for encrypting up to the Top Secret level according to the National Institute of Science and Technology (NIST).

[11] Also known as the *symmetric authentication method*, which is the "method for demonstrating knowledge of a secret, in which both entities share a common authentication information [10181-2]" (*ATIS Telecom Glossary*, 2011).

Annex D
(informative)

D Security Mechanisms

Three important topics within NGN security are authentication, confidentiality, and integrity. Authentication ensures connection integrity and origin, which can protect against falsified credentials and man-in-the-middle attacks where data is intercepted by a third party and manipulated. Confidentiality, depending on the strength of the encryption protocol, guards data from being viewed by parties other than the intended. Confidentiality may also mask the sender and receiver in a session. Finally, integrity detects and prevents outside alterations made to data. However important security protocols are for certain streams, there is an associated overhead, both in terms of packet size and processing time, which could lead to poor QoS.

There are many current viable frameworks and algorithms for supplying media streams with security services as well as new ones undergoing research. More information on security services is provided in Annex B. This section will attempt to highlight some of the foremost security protocols outlined by ATIS, IETF, and ITU-T.

D.1 Major Security Protocols Found in ATIS-0100014

ATIS-0100014, Section 5.4.3, provides insight into the state of standards for security services as of October, 2007. This section summarizes the major protocols presented in the document and how they could potentially have an impact on voice and video services. The major protocol suites listed consist of IP Security (IPSec), Transport Layer Security (TLS), and Datagram Transport Layer Security (DTLS). The document also discusses three authentication protocols: Remote Authentication Dial In User Service (RADIUS), Diameter, and Keyed Digest. In addition, Public Key Infrastructure (PKI) is listed as a major protocol, although this document discusses that particular mechanism in Annex C.

D.1.1 IP Security (IPsec)

IPSec, defined in RFC 4301, is a comprehensive Internet security suite that can offer encryption, authentication, and confidentiality services for traffic flows running over both User Datagram Protocol (UDP) and Transmission Control Protocol (TCP). When in “tunnel mode” through a Virtual Private Network (VPN), IPSec can encrypt the entire packet from the network layer, which greatly increases confidentiality and authenticity while additionally offering privacy. IPSec can carry packets securely from one gateway to another, typically beginning at the edges of a private network. IPSec may reliably protect packets from outsider attacks (attacks originating from outside the private networks of the end points), although it does nothing to protect against insider attacks when operating only through a central VPN, in which case it should be used in conjunction with an application layer security protocol. Tunnel mode is important for protecting traffic flow privacy during a session. On the other hand, transport mode would be used if layer 3 packet inspection is important, which is likely not the case for voice or video over IP.

The key exchange algorithm used by IPSec is based off of Diffie-Hellman (DH) keying and is called the IPSec Key Exchange (IKE) algorithm. However, a PKI is used in order to establish a shared session secret. The use of PKI introduces extra complexity and possible congestion points during the keying signaling.

IPSec encrypts not only payload data, but also header information, appending new headers to allow for VPN tunneling between cryptographic end-points. The packet length overhead for IPSec varies depending upon the encryption algorithm, but using standard 3DES encryption, the overhead is 54 bytes for an average 120-byte voice packet. Secure hash functions may also be used to further protect information. The primary concern with using VPN tunnels is that cryptographic congestion may occur

at the end points of the tunnel. Just as with packets, gateways have a finite capacity for processing encryption traffic and traffic loads in excess of the capacity will buffer until there is overflow. Cryptography engine implementations may differ from device to device, including both software and hardware solutions. Hardware solutions are almost always faster but more expensive than software ones. An additional concern is that priority markings cannot be seen at the far end of the tunnel until the entire packet has been decrypted and therefore priority queuing is not an option for decryption traffic, which is a recurring issue for all secure encryption systems.

D.1.2 Transport Layer Security (TLS & DTLS)

TLS, defined in RFC 5246, is the successor to Secure Socket Layers (SSL), which is a fairly ubiquitous technology for offering security over the Internet. Both use asymmetric keys for privacy and public keys are used in order to initiate a three-way handshake that establishes secure signatures to use for the exchange of data. Historically, this security protocol has been used for TCP traffic – such as hypertext transport protocol (HTTP) – only, although Datagram Transport Layer Security (DTLS) extends support to other protocols such as UDP which is used to transmit certain voice and video data primary. These protocols rest above the transport layer, offering encryption and confidentiality to application and session data. Additionally, a VPN tunnel may be established with TLS. TLS offers robust authentication possibilities through digital certificates and secure hash algorithms (SHA).

The key exchange used in TLS uses a three-way handshake that involves not only the asymmetric exchange of encryption keys, but also the guided selection of symmetric ciphers and hash functions, depending on the highest shared capabilities of both the client and server. Key exchange is conducted with a managed PKI system, which tends to be expensive in both complexity and processing time. Rivest, Shamir, and Adleman (RSA) algorithm is an option for key exchange within TLS, although it is not recommended today as it is weak security.

Since TLS has its roots in a client-server model, it is not extremely straightforward to shoehorn it into peer-to-peer telecommunications. One way around the classical limitation is to use the NGN intermediate servers in order to establish a secure connection between users, although this adds even more additional complexity to the security solution.

D.1.2.1 Datagram Transport Layer Security (DTLS)

Datagram Transport Layer Security (DTLS), defined in RFC 4347, is a set of extensions provided for TLS in order to supply the same authentication and confidentiality framework for transport protocols other than TCP, most notably including UDP which is what would be most used for voice and video streaming sessions. The same set of complications exists for DTLS as for TLS.

D.1.3 Secure Shell (SSH)

Secure Shell (SSH), defined in RFC 4251, and later expanded through various other RFCs, is another client-server protocol that establishes a secure tunnel for a variety of applications including telnet and common file transfer programs and protocols. Similar to TLS in many ways, SSH uses public certificates in order to authenticate during the key exchange process. As primarily a remote terminal and file transfer protocol, it has little practical application in telecommunications, although it is notable for its commonality with other protocol suites. SSH may be used in order to configure routers remotely during network engineering sessions.

D.1.4 Authentication Protocols

In order to validate entities wishing to communicate with one another securely, an authentication protocol should be used. Many authentication protocols also perform authorization and accounting, which is important for the logging of supposedly secure sessions in order to fulfill non-repudiation and maintain the tenants of security services outlined by ITU-T.

D.1.4.1 RADIUS

RADIUS is a low layer protocol for performing authentication, authorization, and network resource accounting for network services; it was brought into IETF standards under RFC 2865. Though it used sufficiently strong encryption for the time it was developed, the chosen symmetric encryption and hashing function are weak by today's standards. It has been mostly replaced by Diameter.

D.1.4.2 Diameter

As the successor to RADIUS, Diameter (RFC 3588) also provides authentication, authorization, and network resource accounting services for computer networks in an extendable container. Diameter is an important protocol for functional network entities to communicate with other network servers during signaling. Used in conjunction with IPSec or TLS, Diameter offers sufficient security for client-server interactions. As primarily a client-server protocol, Diameter is not appropriate for the end-to-end encryption of streaming traffic, since such traffic comes in a peer-to-peer ad hoc package.

D.1.4.3 Keyed Digest

Keyed Digests are a type of cryptographic hash function that produces a message digest based on a set of keys. Keyed Digests can provide both authentication and confidentiality services to a degree, although most older designs have been proven to have flaws. Keyed Digests rely on the use of pre-shared secret keys, which are simplified through the use of a PKI. Secure Hash Algorithm (SHA), defined in RFC 3174, is a family of such algorithms that is discussed more in-depth in section 5.2.3.

D.2 Other Major Security Protocols

While section 5.4.3 of the ATIS-0100014 document listed many important security protocols for telecommunications, there are other major protocols, which are discussed in this section, including Secure Real-time Transport Protocol (STRP), Zimmermann Real-time Transport Protocol (ZRTP), Secure Multipart Internet Messaging Extensions (S/MIME), and SHA. There are countless other low-layer protocols such as Point-to-Point Protocol (PPP) that were deemed to be out of scope for this paper.

D.2.1 Secure Real-time Transport Protocol (SRTP)

SRTP, defined in RFC 3711, encrypts application layer data at session end points using symmetrical master-key encryption, although it leaves headers exposed, which could undermine the privacy of the cryptographic traffic. Despite that shortcoming, SRTP is able to efficiently encrypt and protect call data without relying on a network resource for cryptography since all of the encryption will happen at the handsets on each end. Whether or not confidentiality is a concern is dependent on the security needs of the associated traffic. In addition, SRTP is relatively low packet length overhead compared to protocols which encrypt the full packet since it only affects payload data. Derivation of session keys is from a single master key which is exchanged using an outside protocol.

> NOTE: The term Master can be interpreted to have unfortunate connotations in current usage. This terminology has been used in the approved IETF References and is being retained solely to assist the reader in understanding the principles and when referring to the referenced text.

Since encryption is performed at the session endpoints under SRTP, processing relies on the equipment situated there. Therefore, each user may have a different experience when using the protocol. That said, due to SRTP's relatively low overhead, most modern devices should be able to keep up with the processing demands for at least low bandwidth voice media.

It is beneficial to have a protocol such as SRTP to secure the trusted network before traffic reaches a cryptographic tunnel such as IPSec. Using multiple security mechanisms can build up resiliencies to more types of attacks. SRTP does support the Advanced Encryption Standard (AES), which provides a very large cryptographic search space with no known attacks that reduce it to a space that is searchable using current technology. Like the use of 3DES, the National Institute of Science and Technology (NIST) approves of the use of AES for the foreseeable future.

D.2.2 ZRTP Key Exchange for SRTP

Session Description Protocol (RFC 4566) is a typically used key exchange protocol, although ZRTP (RFC 6189) is an up and coming protocol for the exchange for the SRTP master key included in IETF standards in early 2011. It defines a VoIP-centric key negotiation scheme for SRTP and uses a DH handshake without the need to rely on PKI, but also has man-in-the-middle attack detection capabilities.

D.2.3 Secure Multipart Internet Messaging Extensions (S/MIME)

Secure Multipart Internet Messaging Extensions (S/MIME), defined in RFCs 5750 and 5751, is a public key signing scheme for primarily MIME data, although it has applications for the securing of handshake traffic for other security protocols as well. The protocol does require a PKI and there are additional challenges involving non-ubiquitous support for the protocol.

D.2.4 Secure Hash Algorithm (SHA)

Cryptographic hash functions such as SHA-1, developed by the National Security Agency (NSA), provide additional security and form an optional part of security suites such as IPSec and TLS. They involve minimal effort from a processing perspective and offer advantageous authentication services as well as data confidentiality. Older hash algorithms, such as Message Digest algorithm MD5 (and even more recently SHA-1), have been proven to be non-resistant to certain attacks and as such should only be used with other more secure technologies.

D.3 Packet Filtering[12] Mechanisms

Packet filtering mechanisms provide optional network services, both involving security and not, that may have an impact on all types of traffic including peer-to-peer voice and video streaming. The inclusion of most filters will introduce new possible congestion points, although not all services are mandatory and may scale down in the face of congestion. The mechanisms inspected in this document are stateful firewalls, intrusion detection and prevention systems (IDPS), and Session Border Control (SBC).

D.3.1 Stateful Firewalls

Stateful firewalls are devices that perform Stateful Packet Inspection, which is a process that maintains connection and session information in memory and then scans each incoming packet in order to ensure that they belong to a registered connection. If an incoming packet does not match a listed profile, then it is dropped at the firewall and potentially logged. The maintained connection information may include

[12] *Packet filtering* is "the action a device takes to selectively control the flow of data to and from a network. Packet filters allow or block packets, usually while routing them from one network to another (most often from the Internet to an internal network, and vice-versa). To accomplish packet filtering, you set up rules that specify what types of packets (those to or from a particular IP address or port) are to be allowed and what types are to be blocked" (*ATIS Telecom Glossary*, 2011).

IP addresses, ports, and sequence numbers for retransmission mechanisms such as those built into TCP. Scanning each packet is relatively simple processing for the firewall, although setting up resources is more processing-intense, but that only occurs once per connection. Stateful firewalls can go a long way in preventing unauthorized denial of service attacks against key resources in public and private networks and they are typically worth the small per-packet processing cost. However, as with any congestible resource, there exists a risk that it will act as a bottleneck for the communications network.

A similar technology is Deep Packet Inspection which can search packet payloads for signatures that match profiles for viruses, spam, or intrusion attempts. It is worth noting that such a process is very processing-intense and not suitable for streaming services. Additionally, it is not viable to use Deep Packet Inspection with encrypted payloads.

D.3.2 Intrusion Detection & Prevention Systems (IDPS)

IDPSs, mentioned in RFC 2267 and defined in NIST 800-94, monitor networks for outside attacks using three methods: signature-based protection, statistical anomaly-based detection, and stateful protocol analysis detection. Signature-based protection is the most common and involves matching known attack patterns to current network traffic. Statistical anomaly-based detection establishes an acceptable baseline definition and reports fluctuations that result in levels outside a certain threshold. Finally, stateful protocol analysis detection forms a set of acceptable benign profiles and reports any traffic profile that acts outside of that subset. Each of these search approaches is very computationally expensive, although there is room to simplify many aspects of these systems.

While potentially computationally expensive, IDPSs can be vital in preventing intrusions or alerting networks engineers about problems that could cause interferences with legitimate traffic flows, including voice or video streaming. IDPSs may be designed so that once they reach a congested state, they simply reduce their search space in order to avoid congestion; of course, this type of approach reduces their effectiveness, but that may be acceptable compared to the alternative of introducing network congestion depending on the importance of security in the system in question.

D.3.3 Session Border Control (SBC)

SBC, defined in RFC 5853, has many purposes, including but not limited to allowing disparate network parts to communicate, shaping and policing traffic, acting as a point of lawful interception, providing Network Address Translation (NAT) services, and acting as a point of data and fax interconnection. In regards to voice or video media, the use of SBCs may extend the length of the media path, which lowers QoS potentially. SBCs classically also have control over codec use which could cause QoS changes throughout a call. As with any server with a fluctuating load, they are able to congest and as such should be engineered properly for the target system. Under some circumstances and in order to hide topology, SBCs may act as proxy endpoints for connections, which dramatically increases signaling duration.

Annex E
(informative)

E Network Parameters Used in Modeling

The primary modeling scenarios of the analysis presented in section 5 of this document use a model involving two wireless access networks connected by a simple core network. Mixed traffic consisting of approximately 80% Best-Effort and 20% Expedited Forwarding is added in order to meet typical assumed engineering guidelines of 70% access network utilization and 30% core utilization. Access congestion is comprised of nearly entirely voice calling for this study so that voice calls must compete against other voice calls and not against traffic with a lower traffic class. On top of the baseline traffic volume, 10 handset-to-handset priority voice calls are added to account for the trace group that drives the analysis figures. When sRTP sessions are involved, they run from the starting handset to the terminal handset. An IPSec VPN tunnel is set up between two gateway routers and is used by any voice call in the model that requires IPSec. Figure E.1 (below) shows the complete layout of the network used in modeling.

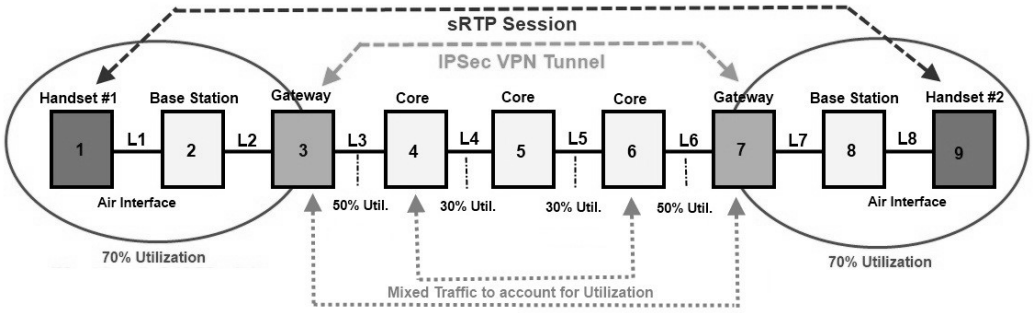

Figure E.1 - Modeling Network Layout

The individual parameters for each of the routers and links in the model shown above are included in Table E.1 (below). In addition, sRTP is modeled as a set cryptographic delay of 0.004 seconds when the Authentication Header is used and codec delay is naturally dependent on the codec used.

Table E. 1 - Modeling Network Parameters

Modeling Parameter	Router / Link (References Table A.1 Above)							
	1 (Handset) / L1	2 / L2	3 / L3	4 / L4	5 / L5	6 / L6	7 / L7	8 / L8
Baseline Link Utilization (%)	70%	70%	50%	30%	30%	50%	70%	70%
Link Capacity (Mbps)	8.72	8.72	147.82	2384.19	2384.19	147.82	8.72	8.72
Link Distance (km)	Wireless	5.0	5.0	1000.0	1000.0	5.0	5.0	Wireless
Processing Time (s)	*N/A*	0.0001	0.0001	0.00001	0.00001	0.00001	0.0001	0.0001
Static IPSec Encryption Time (s)	*N/A*	*N/A*	0.0005	*N/A*	*N/A*	*N/A*	0.0005	*N/A*
IPSec Encryption Time per Byte (s)	*N/A*	*N/A*	0.000005	*N/A*	*N/A*	*N/A*	0.000005	*N/A*
Number of IPSec Processors	*N/A*	*N/A*	1	*N/A*	*N/A*	*N/A*	1	*N/A*
Baseline IPSec Utilization (%)	*N/A*	*N/A*	40%	*N/A*	*N/A*	*N/A*	40%	*N/A*

Cryptographic traffic in the network is flagged so that the baseline utilization of the IPSec tunnel ingress and egress cryptographic engines is 40%. When congestion is included in the second and third scenarios of section 5, it is a straight three-fold increase of each type of traffic (best-effort and expedited

forwarding) at each stage of the network. Additional cryptographic traffic is only flagged when cryptographic congestion is needed specifically, as it is for the modeling in section 6.2.