SECURITY AND DSPs IN VoIP (Voice over IP) NETWORKS

M.Tolga SAKALLI [*], Ercan BULUŞ [*], Ion TUTANESCU [**]
[*]*Computer Engineering Department, Trakya University, Turkey*
[**]*Electronics Department, University of Pitesti, Romania*
E-mail: tolga@trakya.edu.tr, ercanb@trakya.edu.tr, tution@acasa.ro

Key words: Security, DSP, VoIP, Delay, Performance

Abstract

In this study, we examine the security in VoIP and new generation and low-cost DSPs, which can be used to handle complex calculations of encryption algorithms, and we present performance results and encryption/decryption processing delays for DES CBC and 3DES CBC, which are used in IPSec protocol. These performance results were derived from TI's C6711 DSP device.

I. Introduction

Voice over IP (VoIP) is briefly transmission of voice over IP protocol. Since more people have been online, voice, video and data traffic on the Internet are increasing and this resulted the need to expand trunk networks. Apart from that there are some constraints in VoIP applications. These are bandwidth, latency, jitter, echo cancellation, reliability, and security. Because security is concerned with this study Security constraints will be discussed for VoIP protocols. From the point of security, delay is the main constraint and it is desired to minimize delay in real-time communications. The real-time nature of the problem poses some constraints. In the case of voice transmission, the maximum acceptable delay in packet delivery for optimal voice quality is 150 ms, which can be extended up to 200 ms in case of encrypted communications. Thus in a standard VoIP application, after the signal has been digitized, there are 150 ms to code the signal using some standard scheme [e.g., ITU standards [9] G.729, G.723, etc.], divide into packets, then route the packets on the Internet, and reconstruct the original traffic stream at the destination, where it usually is buffered in order to smooth the jitter. Because of such a timing constraint, voice packets are small (10-50 bytes long payload) [6].

When the security is concerned with VoIP, IPSec and PGP protocols are used. From these protocols, IPSec provides security services for IP traffic. The services provided by IPSec are based on two protocols: an authentication protocol (AH) and a combined encryption and authentication protocol (ESP). The first protocol provides services such as connectionless integrity and sender authentication, while second protocol is in charge of guaranteeing confidentiality among other services. When the encryption algorithms are concerned with IPSec, DES CBC and 3DES CBC are used to provide confidentiality. In this study, performance results of these algorithms on new generation, low-cost TI's C6711 device and encryption/decryption process delays for this device were given. From these results it can be suggested that new generation DSPs can be used efficiently to handle these encryption algorithms which have complex calculations.

II. Voice over IP

Voice over IP is a technique for transmitting voice data over the Internet. The following steps can be stated as: (1) Analog signal is digitized. (2) Packet generation of digital signal is performed according to TCP-UDP/IP protocols. (3) Packets are transmitted on the network. (4) Packet reception and analog signal reconstruction are performed at the destination. From the point of voice quality, there are some factors which degrade the voice quality. Among the factors which degrade voice quality are packet loss, delay variation, or jitter, voice compression schemes, transducers, echo cancellation algorithms. In this paper we focused on encryption process delay and the encryption algorithms IPSec use.

VoIP is a typical real-time application as the original signal has to be reproduced at the destination as close as possible to the instant when it was generated; therefore the signal delay is a qualifying parameter for VoIP application. The signal delay can be divided to three parts; (1) digitization process, may vary between 0.75-30 ms (2) queuing delay (packet waiting at router), 30 ms. (3) jitter delay (buffering arriving packets), 40-70 ms.

One of the most important constraints for VoIP is bandwidth. If the comfort noise feature is used to improve voice quality, net bandwidth consumption averaged over a 20-30 second period for an active speech conversation is 10,8 Kbps.

The other important constraint for VoIP is delay and delay jitter. Codec delay, serialization delay, queuing delay, propagation delay, etc. all contribute to the overall network packet delays, which could vary dynamically depending on the network dynamics. On the aspect of delay If security in VoIP is concerned, encryption, authentication and key exchange algorithms are by nature computationally intensive and when employed for VoIP only add up to the overall packet delay. More specifically

public key encryption requires very high computation power.

III. Security Characteristics of Some Existing VoIP Protocols

VoIP protocols which are discussed in this paper are: H.323, SIP (Session Initiation Protocol), MGCP (Media Gateway Control Protocol). All VoIP protocols are application layer protocols and these protocols operate on top of IP, i.e. Internet Protocol. VoIP protocols are not limited to use any specified transport layer protocols, such as TCP (Transmission Control Protocol) or UDP (User Datagram Protocol), but in practice that is the case, because they are transport layer protocols that are used in the Internet to a large extent. In figure 1 Protocol architecture is shown. These protocols have similarities since they all address the same need of VoIP communication. One area where protocols differ from each other is the location of intelligence (control), where heavy computation is done. In figure 2 Intelligence in VoIP is shown.

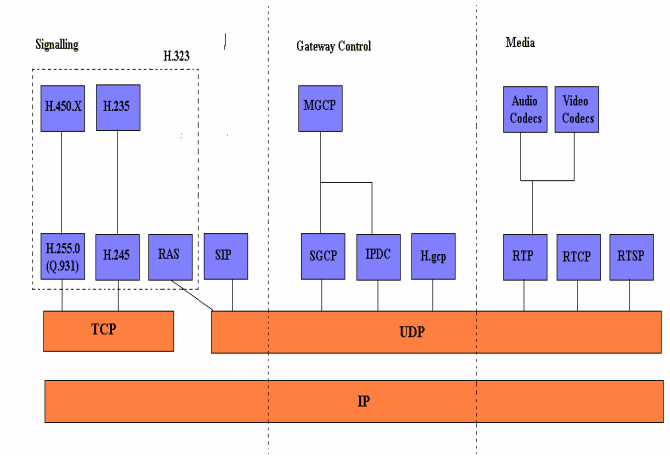

Figure 1-) The Protocol Architecture of VoIP

Protocols	*Intelligence*
H.323	Everywhere
SIP	Intelligent Endpoints Dump Network
MGCP	Intelligent Network Dump Endpoints

Figure 2-) Intelligence In VoIP

Voice needs to be encoded by using appropriate codec at the time. Encoded audio stream is then passed to another layer application protocol RTP (Real Time Transport Protocol) that runs over UDP and is used to transfer real time information streams over the Internet as the name implies. RTCP provides status and control information for the use of RTP.

There are two different ways to provide security in VoIP. First one is to use built-in internal mechanisms of VoIP protocols. Second one is to use external, application or network layer protocols (e.g. IPSec).

3.1. ITU-T's H.323

H.323 [11] specification defines how voice, data, and video traffic will be transported over IP based local area networks. This recommendation is based on the real-time transport protocol / RTP control protocol (RTP /RTCP) for managing audio and video signals.

Authentication: It can be said that there are two types of authentication: (1) symmetric encryption-based that requires no prior contact between the communicating entities and (2) an ability to have some prior shared secret (in the H.235 recommendation referenced to as "subscription" based). This second way to authenticate can either be symmetric or asymmetric.

Encryption: Encryption can be done in RTP-layer which is used to transport information streams. Encryption can be utililized packet-by-packet basis, which means that it is up to the specified policy how encryption capabilities are applied to each packet.

3.2. IETF's SIP

SIP [7,8] is a text-based protocol, similar to HTTP and SMTP, for initiating interactive communication sessions between users. Such sessions include voice, video, chat, interactive games. In figure 3 A basic SIP session is shown. Apart from that SIP is not dependent of the service of the any specific transport protocol. In the SIP RFC, TCP and UDP are suggested but it also states that SIP also may be used with protocols such as ATM AAL5 [RFC 1483], IPX, frame relay, or X.25.

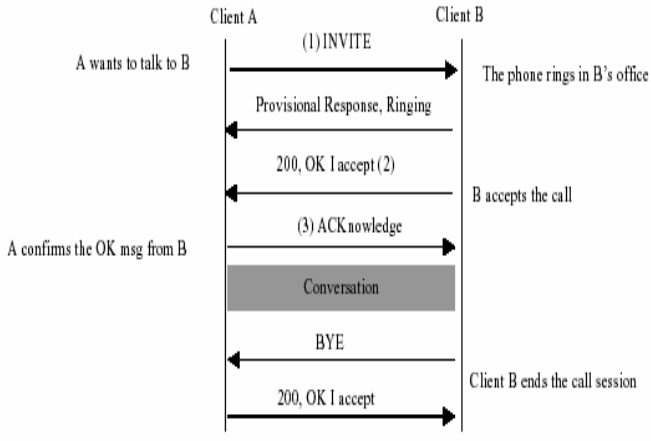

Figure 3-) Basic SIP Session

Authentication: All authentication mechanisms that SIP provides are challenge-response based. There exist three alternatives, which are Basic authentication, Digest authentication, and PGP authentication.

Encryption: Encryption can take place either end-to-end between user agents, or hop-by-hop between any two SIP entities. Hop-by-hop encryption encrypts the whole SIP message and is supposed to work on the transport level or the network layer. The algorithm used is not specified but IPSec [RFC2401] is suggested.

3.3. IETF's and ITU-T's MGCP /MEGACO/H.248

In MGCP/MEGACO/H.248 [7], authentication and encryption is clear and straightforward. According to RFC2885, IPSEC "should" be used for the authentication and encryption of protocol connections. IKE should also be used to provide more robust keying options

IV. VoIP security requirements

Because VoIP networks are essentially an IP network, VoIP network and terminals face the same security threats inherent with any IP network. Because of the open nature of the underlying IP network, some VoIP security requirements could be stated as:

(1) Protection of privacy of the call conversation. It is provided with encrypting all connections between network elements.
(2) Authentication of call end entities.
(3) Protection of servers and clients from well-known threats, such as "denial of service".
(4) Protection of caller behavior and statistical information from unauthorized access.
(5) Access control by service provider

4.1. Delay Sensitivity of VoIP

As it was mentioned before VoIP is real-time application and sensitive to network packet delays and delay jitter. The maximum acceptable delay for voice transmission is 150 ms.

Encryption, authentication and key exchange algorithms are computationally intensive and from the point of security there are two options to handle packet-processing delay. One option is to employ end-to-end encryption to distribute computational power requirement. This option has a disadvantage if the system management, key exchange management and maintenance costs are considered. The other option is to employ security services at edge routers, gateways but this option is a bottleneck when the network throughput is very high.

Our study shows that when we use new generation and low-cost DSP hardware, encryption/ decryption processing delay is very low and it can be ignored in VoIP communication.

4.2. IPSec

IPSec[5] is a protocol and uses a collection of protocols to provide security at the network layer for any application. It was standardized by the IETF to address the security problems inherent in IP. It is independent of underlying network topologies and provides transparent services to the application layer. In addition, IPSec is independent of cryptographic algorithms. That means it allows integration of new and stronger encryption algorithms into the IPSec architecture.

The services provided by IPSec are based on two protocols: an authentication protocol (AH) and a combined encryption and authentication protocol (ESP). The first protocol provides services such as connectionless integrity and sender authentication, while second protocol is in charge of guaranteeing confidentiality among other services. ESP can be configured to provide confidentiality alone, or authentication alone, or both. Encryption algorithms DES in CBC mode, 3DES in CBC mode, IDEA, CAST can be used with ESP for providing confidentiality. For authentication, message digest algorithms, such as HMAC-MD5 can be employed.

V. DES and 3DES Algorithms

IBM has developed DES in cooperating with National Securities Agency (NSA) in 1974 and it has been the worldwide encryption standard for more than 20 years. The predominant weakness of DES is its 56-bit key. When it was developed, it was more than sufficient. Nowadays it has become insufficient to protect against brute-force attack by modern computers. As a result of the need for greater encryption strength, DES evolved into triple-DES. Triple-DES encrypts using three 56-bit keys, for encryption strength equivalent to a 168-bit key. This implementation, however, requires three times as many rounds for encryption and decryption and highlights a second weakness of DES- speed. DES was developed for implementation on hardware and DES implementations in software are often less efficient than other standards which have been developed with software performance in mind.

5.1. Implementation of DES

DES [1,3,4] is a block cipher; it uses 64 bit blocks. Confusion and diffusion are the encryption techniques, which are used to accomplish this algorithm. In this algorithm, chosen sections of data are substituted for corresponding sections from the original data. The choice of the substituted data is based upon the key and the original plaintext. Diffusion is accomplished through permutation. Rearranging the order of the various sections permutes the data. These permutations, like the substitutions, are based upon the key and the original plaintext. The substitutions and permutations are specified by the DES algorithm. Chosen sections of the key and the

data manipulated mathematically and then used as the input to a look-up table. In DES these tables are called the S boxes and the P boxes, for the substitution tables and the permutation tables, respectively. In software these look-up tables are realized as arrays and key/data input as the index to the array. Usually the S and P boxes are combined so that the substitution and following permutation for each sound can be done with a single look-up.

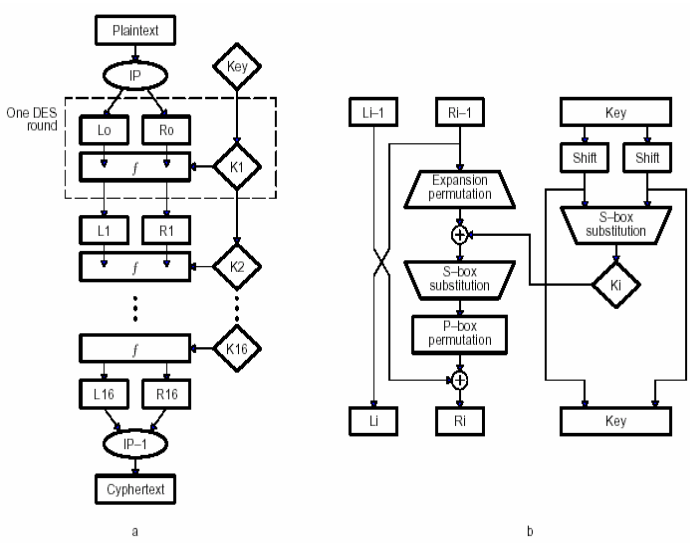

Figure 4-) a-) DES
b-) One round of DES

In order to calculate the inputs to the S and P box arrays, portions of the data are XORed with portions of the key. One of the 32-bit halves of the 64-bit data and the 56-bit key are used. Because the key is no longer than the data half, the 32-bit data half is sent through an expansion permutation, which rearranges its bits, repeating certain bits, to form a 48-bit product. The S and P box look-ups and the calculations upon the key and data, which generate the inputs to these table look-ups, constitute a single round of DES.

This same process of S and P box substitution and permutation is repeated sixteen times, forming the sixteen rounds of the DES algorithm. There are also initial and final permutations, which occur before and after sixteen rounds.

VI. DSP Hardware

In this study, DSP hardware used is TI's TMS320C6711 DSP device. Texas Instruments TMS320C6000 generation of high performance includes the TMS320C6711. C6711 is a low cost version of the original C6000 floating-point device, the C6701. It provides 900 MFLOPs (million floating-point operations per second) at 150 Mhz. C6000 generations is based on TI's VelociTI architecture an advanced very long instruction word (VLIW) architecture for DSPs. VelociTI provides eight execution units, including two multiplexers and six arithmetic logic units (ALUs). These units operate in parallel and can perform up to six floating-point instructions during a single clock cycle. C6711 has 72 Kbytes of on-chip memory [3].

VII. DES Performance Results on DSP

Speech signal was sent to DSP through serial port of computer. This signal is analog signal. After it was converted to digital form, encryption, decryption and verification were carried out on this digital form. After words, CCS, DSP/BIOS statistic objects and DES source codes were used to gather statistics, cycles for 1024 byte data. For DES and triple-DES, the equivalent data rates for each mode on the C6711 (150 Mhz) were measured in billions of bits per second. Data rates were calculated directly using cycle counts.

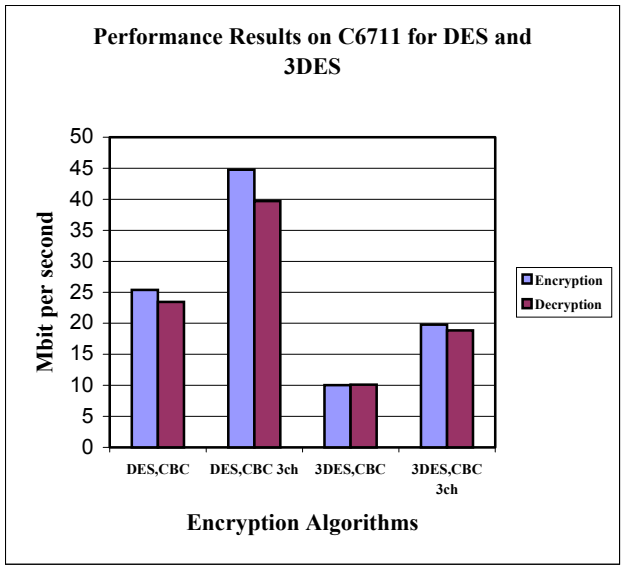

Figure 5-) DES and 3DES performance results on C6711 for 1024 byte data

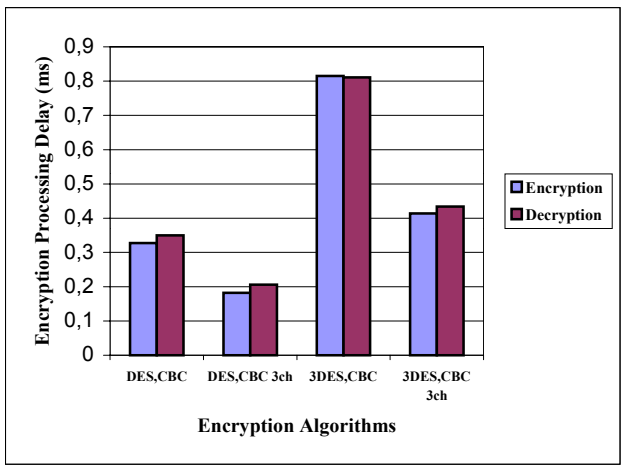

Figure 6-) Delay (ms) according to performance results for DES and 3DES algorithms

7.1. Encryption/decryption Delay

DES CBC, 3DES CBC and their 3 channels performance results for encryption and decryption which is for 1024 byte data is shown in figure 5. In addition, encryption/decryption processing delays for 1024 byte data according to these results is shown in figure 6. As it is shown, DES CBC encryption/decryption processing delay is 0,327 ms and 0, 35 ms respectively. For 3DES CBC encryption/decryption processing delay is 0,414 ms and 0,434 ms respectively. These results show that encryption/decryption processing delays are very low and it can be suggested that low-cost and new generation DSPs can be used for encryption/decryption, which has more complex calculations, in VoIP networks. Data rates in figure 5 were calculated using cycles counts.

$$Data\ Rate = \frac{150\ Mhz}{(cycles\ /\ 8192\ bits)}$$

Encryption/decryption processing delays in figure 6 were calculated using performance results.

$$Encryption/decryption\ processing\ delay = \frac{1024\ byte\ (8192\ bits)}{Performance\ results\ on\ C6711\ for\ encryption\ algorithms}$$

VIII. Conclusion

This study shows that IPSec protocol is widely used in VoIP networks. DES and 3DES in CBC mode for encryption algorithms are used in this protocol. In our study, we presented performance results and encryption/decryption processing delays for DES and 3DES in CBC mode. These performance results were derived from TI's C6711 DSP device which is low-cost and new generation DSP. According to the results, it can be suggested that when developing VoIPSec protocol and developing new and stronger encryption algorithms for IPSec, we can use these low-cost and new generation DSPs which provide efficient encryption/decryption processing delays (less than 1 ms).

REFERENCES

[1] **R. Stephen Preissig**, 2000. "Data Encryption Standard (DES) Implementation on the TMS320C6000", Literature Number SPRA702, Texas Instruments

[2] **Bruce Schneier**, 1996. "Applied Cryptography, Second Edition", John Wiley & Sons, Inc. , New York, Ny

[3] **M. Tolga Sakallı, Ercan Buluş**, 2002. "Speech Encryption with DSP", Scientific Bulletin of Electronics and Computer Science, Pitesti University, Romania

[4]**M. Tolga Sakallı, Ercan Buluş**, 2002. "DES'İN TMS320C6711 DSP Cihazı üzerindeki Uygulaması, Performansı ve Karşılaştırılması", ELECO 2002

[5] **Mohan Krishna Ranagathan, Liam Kilmartin**, 2002. "Performance Analysis of secure session initiation protocol based VoIP networks", Computer Communications, Elsevier.

[6] **Roberto Barbieri, Danilo Bruschi, Emilia Rosti**, 2002. "Voice Over IPsec" : Analysis and Solutions, 18th Annual Computer Security Applications Conference December 9-13, Las Vegas, Nevada

[7] **M. Marjalaakso**, 2001. "Security Requirements and Constraints of VoIP". Technical Report, Dept. of Electrical Engineering and Telecomunications, Helsinki University of Technology.

[8] **Fredric Thernelius**, 2000. "SIP, NAT and Firewalls", Master's Thesis, Kungl Tekniska Högskolan,Stockholm,http://www.cs.columbia.edu/~hgs/sip/drafts/Ther0005_SIP.pdf

[9] **Demir Öner**, 2002. "Internet Protokolu Üzerinden Ses İletişimi (VoIP)" : Standartlari, Servis Kalitesi, Trafik ve Kanal Hesaplamaları, ELECO 2002

[10] "Security in SIP-Based Networks", White Paper, http://www.cisco.com/warp/public/cc/techno/tyvdve/sip/prodlit/sipsc_wp.pdf

[11] "H.323", http://www.iec.org/online/tutorials/