Global Information Assurance Certification Paper

Voice over Internet Protocol: A Discussion on How to Securely Implement on an Existing Data Network

Kevin Larson
July 18, 2004
GSEC Pratical Requirements (v.1.4b) (August 2002**)**

Abstract

Voice-over-Internet-protocol (VoIP) refers to the transmission of voice data over a packet-switched network. VoIP promises lower communication costs and greater flexibility while using a corporation's current infrastructure to perform everyday business. Current examples of VoIP applications are Microsoft's NetMeeting and Instant Messenger. 3Com, Hewlett Packard, and Siemens are only a few examples of vendors offering VoIP systems and equipment. But, like any other new technology, significant security challenges are presented. Assumptions can be easily made that since digitized voice data are placed into packets, VoIP equipment and applications can be plugged into an already secure network with out taking any extra precautions. When the truth is combining voice and data on one network also combines their threats as well.

This document will present an overview of VoIP in hopes the reader will gain a better understanding of the technology and what preparation is required for a secure implementation. The paper is divided into two sections. The first section will provide a VoIP overview with an emphasis on the popular Session Initiation Protocol (SIP) and Quality of Service (QOS) issues which are inherent in VoIP networks. The second section of the paper reviews some techniques and recommendations that can be used to secure a VoIP network.

Table of Contents

1 VoIP Overview

The first step to any new implementation, in this case VoIP, is to learn about the technology. Adding new software and hardware can have a significant impact to a network. Some questions to be asked while reviewing the technology is how will my existing network change, if the network has a firewall in place what rule sets will have to be implemented, and what has to be tested to insure the implementation happens smoothly. Section 1 will cover the following:

- Differences between circuit-switched and packet-switched voice systems
- Data handling
- Session Initiation Protocol (SIP)
- Quality of Service (QoS)
- Application and System Testing

Understanding the technology will minimize the risk of incorrectly configured devices and negative impact to existing applications.

1.1 Differences and Benefits of Packet-Switched over Circuit-Switched networks

Telephone calls have been traditionally made on a circuit-switched network. Circuit-switched refers to when a telephone call is initiated, a circuit is established between the sender and receiver until the conversation is completed [1]. The circuit acts as a specific tunnel which includes the bandwidth and processing time to accommodate the call. Because the circuit is constant throughout conversation, circuit-switched networks is bandwidth intensive and limits the amount of calls a circuit-switch can handle.

Networks utilizing VoIP systems and applications are based on packet-switched networks. Packet switching improves efficiency by using a connection just long enough for a sender to send a chunk of data to a receiver. The sender's computer segments the data into packets. The receiver's address is annotated in specific fields within the packet. When the receiver's computer gets the packets, it reassembles them into the original data. Chunks of data on IP networks cause bursts of activity which in comparison to circuit-switched networks experience constant activity. These bursts of network traffic can impact the quality of the voice and lead to Quality of Service issues which will be discussed later.

VoIP systems can come in a variety of styles and are becoming widely accepted. Vendors are currently offering traditional telephone VoIP handsets. An example is Siemens optiPoint 600, a universal telephone that supports both circuit-switched and VoIP communication. The optiPoint 600 office includes a built-in headset jack, WAP browser, LDAP interface and miniswitch [2]. Workstations can have multiple VoIP applications loaded. Microsoft's NetMeeting uses the H.323 protocol and Instant Messenger uses the SIP protocol to allow voice data

[1] http://computer.howstuffworks.com/ip-telephony.htm/

[2] http://www.sipcenter.com/sip.nsf/html/Sponsors+Siemens

handling. Linux, Windows, and Macintosh operating systems all support various VoIP applications. The SIP center website at URL http://www.sipcenter.com provides a good resource for applic[3]ations, hardware, and white papers.

Bottom line is VoIP provides several advantages over traditional circuit switching:

- VoIP allows several calls to occupy the amount of space occupied by only one call in a circuit-switched network
- VoIP can utilize data compression
- VoIP is compatible with existing data networks infrastructure

1.2 VoIP Data Handling

A call first must be initiated from a sender to a receiver. When the receiver answers the call a session is created. The voice data to be transmitted from the sender is converted into a digitized format and then segmented into a stream of packets. An analog-to-digital converter, also known as an audio codec, on the sender's side handles converting voice data from analog to digital. Compression algorithms are used to shrink down the number of bits required to be transmitted. Compression is an important piece to the efficiency of VoIP networks. Reducing the amount of data to be sent will decrease network bandwidth usage and help increase the speed in which the data can be sent. Next, the compressed voice data is inserted into data packets to be carried on the Internet. Real Time Protocol (RTP) is typically the protocol used for handling voice packets. RTP packets have specific fields for holding data required to correctly re-assemble the packets into the proper order at the receiver's end. User Datagram Protocol (UDP) will be responsible for carrying the voice data through multiple network nodes and networks within the Internet [3].

Once the packets are delivered to the receiver's end the process is reversed. The packets are disassembled and put into the proper order, digitized voice data is extracted from the packets and uncompressed. Then the digitized voice is processed by the audio codec to render it into analog signals allowing the receiver to understand.

Figure 1 displays the VoIP stack. IP is carried over the network layer which is also known as layer 3 in the OSI model. Layer 3 is primarily responsible for end-to-end communication between computers located on different networks. The IP protocol identifies the destination and the best path for routing the packet to its specific destination. UDP and TCP are used at the transport layer, or layer 4. The transport layer protocols are responsible for getting the data to the destination in a timely manner[4]. UDP and TCP are encapsulated into the IP packets.

---

3 http://csrc.nist.gov/publications/drafts.html

[4] Zacker, Craig. "Windows 2000 Network Infrastructure Administration"

When a packet is sent via TCP the sender sends a packet to the receiver and waits for a response. Once the receiver receives the packet the receiver sends back a reply to the sender that the packet was received successfully. If the sender does not receive a response another packet will automatically be resent. UDP is TCP's exact opposite. There are no checks in place to verify a packet has reached its destination. UDP is typically used in streaming audio and video, where some packet loss is acceptable. VoIP specific protocol packets are encapsulated into UDP or TCP depending on the application.
The SIP signaling protocol is responsible for locating the receiver and determining the receiver's availability and capability. Once availability has been established SIP will perform a session setup and manage the setup. The Real Time Transport Control Protocol (RTCP) helps administrators to maintain the VoIP quality by gathering the following Quality of Service characteristics: delay, jitter, and packet loss[5]. The data gathered can be used to troubleshoot problem networks. The Real Time Transport Protocol (RTP) is responsible for packet reassembly and UDP/TCP is responsible for getting the voice data to its destination. IP routes the data between networks to its destination. All these protocols need to work together to ensure voice integrity and quality.

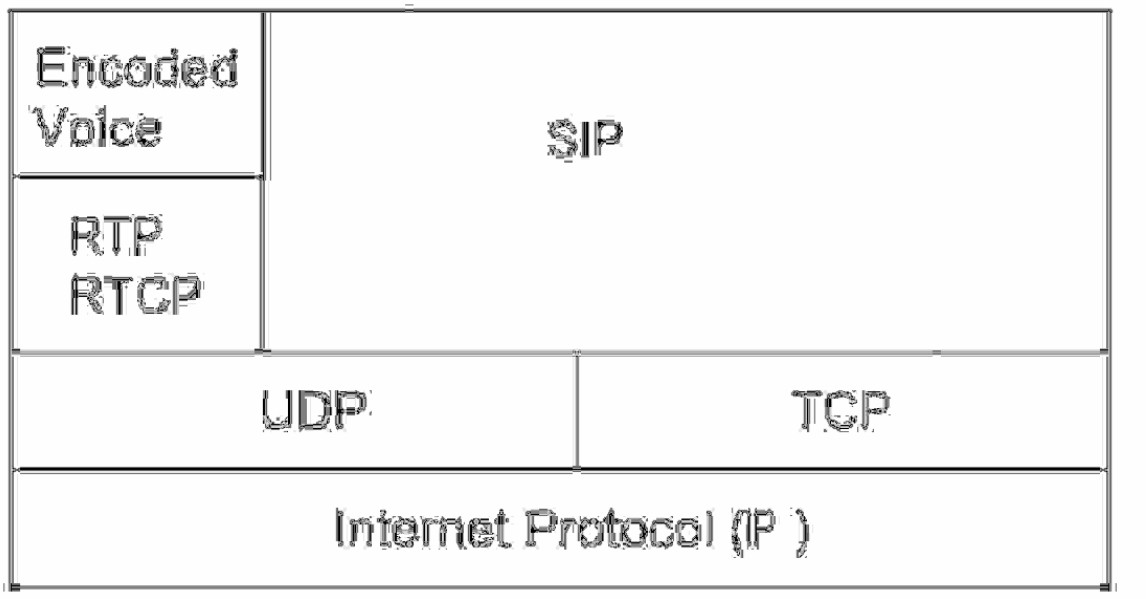

Figure 1 VoIP Protocol Stack [5}

1.3 Session Initiation Protocol

In a public network environment, products from different vendors need to operate with each other if voice over IP is to become common among users. To achieve interoperability, standards are being devised and the most common standard for VOIP is SIP.

[5] http://itresearch.forbes.com/detail/RES/1080147951_179.html

A protocol is an agreed-upon format for transmitting data between two devices. Protocols specify what type of error checking and data compression will be used. How the sending device will signal it is finished sending data and how the receiving device will verify it received all the data is also determined in the protocol [6].

SIP was developed by the Internet Engineering Task Force (IETF). A SIP compliant device or application has a user agent which allows VoIP calls to be made. The SIP architecture is based upon a client-server architecture and is modeled after TCP's three way handshake. RFC 3261 defines SIP as an application layer control (signaling) for creating, modifying, and terminating sessions with one or more participants[7] . For locating perspective session participants, and for other functions, SIP enables the creation of an infrastructure of network hosts (called proxy servers) to which user agents can send registrations, invitations to sessions, and other requests. Requests are generated by the client and sent to the server. The server processes the request and then sends a response back to the client. A transaction is made up of client requests and server responses. SIP includes INVITE and ACK messages which define the process of opening a reliable channel to which a call control message may be passed. An important characteristic of SIP is that SIP provides the application reliability and does not rely on TCP for reliability.

There are several types of servers used to support SIP with in a network. The registration server receives updates concerning the current location of users. A redirect server receives requests, determines the next-hop server, and returns the address of the next hop server to the client instead of forwarding the request. The next hop server has more information about the location of the called party. The client takes this information and sends the data to the specific address. A SIP gateway server translates between the VoIP and the Public Switched Telephone Networks.

Figure 2 displays an example of a conversation between a host using XTen Networks Inc. Free World Dialup X-Lite application [8] and a server. The user can dial a number using the X-Lite soft phone application and an Invite is sent from the host (192.168.X.XXX) to the SIP server @fwdpulver.com (85.39.XXX.XXX). To start the test the user dialed 411. "Status: 407 Proxy Authentication Required" indicates the client must first authenticate with the proxy. The client then forwards an authentication and "Request: ACK" from the proxy server. Once authentication is accomplished another invite is sent to the server. The server responds with a "Status: 100 Trying" which means that the request has been received and that some unspecified action is being taken. "Status: 180 Ringing"

[6] http://www.webopedia.com/TERM/p/protocol.html

[7] http://www.faqs.org/rfcs/rfc3261.html

[8] http://www.freeworldialup.com/

signals the server's user agent is trying to alert the server that there is an incoming call.

A user agent is an end system acting on behalf of a user. There is a client and server portion. The client portion is called the User Agent Client (UAC) while the server portion is called the User Agent Server. The UAC is used to initiate a SIP request while the UAS is used to receive request and return responses on behalf of the user.

When the request has succeeded a "Status: 200 OK" is sent from the server to the originating client. The session is established and the protocol changes from SIP to UDP for transferring the voice data. When the client terminates the call a"Request: BYE" is received at the client and a "Status: 200 Ok" is sent back to server when the call has been terminated.

Source	Destination	Protocol	Info
192.168.X.XXX	65.39.XXX.XX›	SIP/SDP	Request: INVITE sip:613@fwd.pulver.com, with session description
65.39.XXX.XX›	192.168.X.XXX	SIP	Status: 407 Proxy Authentication Required
192.168.X.XXX	65.39.XXX.XX›	SIP	Request: ACK sip:613@fwd.pulver.com
192.168.X.XXX	65.39.XXX.XX›	SIP/SDP	Request: INVITE sip:613@fwd.pulver.com, with session description
65.39.XXX.XX›	192.168.X.XXX	SIP	Status: 100 trying -- your call is important to us
65.39.XXX.XX›	192.168.X.XXX	SIP	Status: 180 Ringing
65.39.XXX.XX›	192.168.X.XXX	SIP/SDP	Status: 200 OK, with session description
192.168.X.XXX	65.39.XXX.XX›	SIP	Request: ACK sip:PPC1054684@65.39.XXX.XXX:50 82
192.168.X.XXX	65.39.XXX.XX›	UDP	Source port: 8000 Destination port: 35448
192.168.X.XXX	65.39.XXX.XX›	UDP	Source port: 8000 Destination port: 35448
192.168.X.XXX	65.39.XXX.XX›	UDP	Source port: 8001 Destination port: 35449
192.168.X.XXX	65.39.XXX.XX›	UDP	Source port: 8000 Destination port: 35448
65.39.XXX.XX›	192.168.X.XXX	SIP	Request: BYE sip:426943@192.168.X.XXX:5060
192.168.X.XXX	65.39.XXX.XX›	SIP	Status: 200 Ok

Figure 2 SIP conversation

1.5 Quality of Service

IP was originally designed to carry data, so it does not provide real time guarantees but only provides best effort service. For voice communications over IP to become acceptable, the delay needs to be minimal to ensure quality of voice, we can use Echo Cancellation, Packet Prioritization (giving higher priority to voice packets) or Forward Error Correction[9].

Quality of Service (QoS) is guaranteed on a circuit switched network by dedicating a constant tunnel between the callers. The guarantee on a circuit switched network does not exist on a IP network. Dropped and retransmitted packets on an IP network is a common occurrence. Packets may be routed

[9] http://www.cse.ohio-state.edu/~jain/cis788-99/ftp/voip_protocols/index.html

through different paths. These alternate paths can cause packets to arrive late or out of sequence at their destination.

Delay of packets is a QoS that can cause echo and talker overlap. One source of Delay is caused by applications requiring a collection of voice samples for processing. This type is called Accumulation delay or Algorithmic delay. The physical medium of the network used to transmit packets the voice data can cause another source of delay known as Network Delay or Latency.150 ms is the upper bound for one way traffic in domestic calls across Public Switch Telephone Network lines in the continental United States. Time constraints create a very small window of time for packet delivery.

Jitter is a type of delay that is a by product from packets arriving to a host out of sequence. Jitter can be caused by network congestion, timing drift, or route changes[10]. Impact caused by Jitter can lead to packets arriving late and be processed out of sequence. VoIP systems use RTP to transfer voice media within UDP packets. UDP is a connectionless protocol and does not allow out of order packets to be reassembled and placed in the correct order at the protocol level. But, RTP does allow applications to reorder the packets by using the sequence number and time stamp fields. Due to the tight time restrictions, network traffic overhead caused by reordering packets can become a problem. Buffers are used at endpoints to smooth out delays of packet arrivals to their destination.

Lost-packets can pose a more severe problem to VoIP applications than delay. IP networks do not guarantee service. Since all voice frames are treated like data under peak loads and heavy congestion, voice frames will be dropped along with data. Since voice packets are time sensitive they are not as easily fixed by retransmission like dropped data packets. Following are examples of two ways a network can compensate for lost voice packets:

- Automatically replay the last packet received when a lost packet was supposed to play. This method will provide a fill in when there is no new voice data is available. This method works well when there are not a large amount of packets lost but, the effectiveness will degrade as the number of lost packets increase.
- Send redundant voice data. This method can compensate for lost packets but the compensation is at the expense of network bandwidth.

Network bandwidth is another concern when implementing VoIP into a network. Bandwidth congestion can cause several QoS problems. Implementers must strive to allocate necessary bandwidth for both data and voice when sharing the same network originally designed for one. RTP header compression will reduce the size of the voice stream traffic and Voice Activation Detection prevents the transmission of empty voice packets. Both are methods for reducing traffic congestions on the network and reduce the impact to the network bandwidth.

[10] http://searchnetworking.techtarget.com/sDefinition/0,,sid7_gci213534,00.html

1.6 Application and Systems Testing

Testing is one of the most important steps of the implementation. Any impacts to the network which could cause security or performance issues should be caught during testing. It is in this step where the implementer can learn more about VoIP applications and how they will react within the existing network. Pilots and Alpha testing are usually performed in a lab or on a test network. Here the implementer can learn about the installation of the products. Test new applications to ensure they are compatible with the existing and verify support documentation is adequate. Locking down or securing the new application via network devices or operating system can be tested as well. It is a challenge to all administrators and implementers to obtain usability with security.

The beta testing phase can be done on a production network. End users get the chance to use the product and provide feedback. Administrators get the chance to verify the implementation is secure and the user is able to use the application as well. There should not be a major impact to production because everything has been previously tested during the pilot and alpha phases.

2 VOIP Security

The previous section of the paper was dedicated in giving a high level overview of VoIP. We have discovered that vendors are starting to develop equipment and applications. The SIP protocol and QoS issues have been reviewed as well. Now that we have an understanding of VoIP and discussed some of its characteristics, we can start looking into some techniques of securing it. Hackers use familiar data communication threats such as Denial-of-Service attacks, viruses, worms, Trojan horses, packet sniffing, and password attacks to compromise IP network security. Phone Phreaks have exposed telephone networks to Denial-of-Service attacks, fraud, and eavesdropping. Because VoIP combines data and voice on the same network, the complexity of securing the network increases as well.

2.1 Voice Security Policy

The first step of implementation is to develop a security policy. Before VoIP can be implemented and used to its fullest potential the policy needs to be supported by management and adhered to by end users. The policy contents should include its background and reason. It is in this section where the purpose and extreme need of having a voice security measures in place is critical. The reason for the policy could be a sales pitch to the reader stating why it is important for security to be adhered to. The desired outcome will be achieving much needed management and user support to secure VoIP applications and systems.

The new policy should also list any related policies or documents. Phone etiquette, information security policies, IT department and vendor service level agreements can also be referred to. This will ensure the voice policy will work with any existing voice or network policies and procedures. If the new voice

policy does make any existing policies or procedures obsolete, this should be documented as well.

Scope will cover what the range of the policy is. Specific subnets and applications with in a network are an example of what could fall into scope of the security policy. The most important part of the voice security policy is the statement. It is in the statement where the actual guiding principles will be documented. User expectations for the VoIP applications should be documented. What type of information and data (ex. confidential, secret, or top secret) discussed on the IP network could be documented here as well. The contents of the policy should also include what actions are necessary when the policy is not adhered to and when they are to be exercised. If the policy did not cover a particular scenario a change control process to update it should be referred to. The policy should be a living document that grows with VoIP as the technology matures within the network.

2.2 Cisco's SAFE Blue Print

Earlier in this paper an example of the FWD X-Lite SIP application was reviewed. In this section Cisco's SAFE blue prints for IP a network carrying voice and system is reviewed[11]. SAFE focuses on four main voice systems: the IP phone, the call processing manager, the voice-mail system, and the voice gateway. One measure SAFE explains is to segment and separate voice calls from data, and to physically secure IP phones. Segmentation will help to prevent infiltration of voice platforms by viruses that have penetrated the data network, and to broker connections between the voice and data segments of the network. The procedure to segment the network between voice and data should be documented in the security policy referenced earlier. The reason for segmentation and who is responsible for the security of each segment should be documented in the policy as well.

Hardening layer 2 and 3 devices is the major part of SAFE. Locking down SNMP, turning off unneeded services, using SSH, and authenticating routing updates are examples of router hardening. SAFE recommends switch hardening through features such as ARP inspection or through the use of private VLANS. Using IP permit lists to restrict access to management ports, using dedicated VLAN ID for all trunk ports, and disabling unused ports are other means to harden switches.

[11] http://www.cisco.com/warp/public/784/packet/jan04/pdfs/PK16106C_SecureVOIP.pdf

2.3 Existing Security Features for SIP

Properly securing SIP on a network is a complex task. SIP has to be able to work on public networks and is expected to create a session between hosts that have no trust relationships defined. Examples of previous exploits are registration hijacking, server impersonation, and tampering with message bodies proves the urgent need for security.

The RFC 3261 [7] highlights the following fundamental security services as being required for the SIP protocol:

- Preserve the confidentiality and integrity of the message
- Prevent replay attacks or message spoofing
- Providing for the authentication and privacy of the participants in a session
- Prevent Denial of Service attacks

Instead of recreating security measures to handle the issue previously mentioned, administrators can reuse security mechanisms from existing security models when possible. Encrypting messages is an important technology to preserve the confidentiality of the messages with in a session. TLS and IPSec are other examples of technology at the Network and Transport layers.

There is a challenge capability provided by SIP. It is based on the HTTP authentication model which relies on 401 and 407 response codes. The 401 (Unauthorized) response message is used by a server to challenge the authorization of the user agent. The 407 (Proxy Authentication Required) is used by the proxy to challenge the authorization of a client and must include a Proxy-Authentication header field containing at least one challenge applicable to the proxy for the requested resource.

2.4 Security Watch

There are numerous internet sites that reports security threats. Carnegie Mellon Software Engineering Institute operates the CERT Coordination Center which can be accessed by going to URL http://www.cert.org/. This site provides advisories on security threats as they are reported from various sources. CERT Advisory CA-2003-06 “Multiple Vulnerabilities in Implementations of SIP” reports a flaw that can lead to DoS attacks or cause network instability. The advisory reports that the Oulu University Secure Programming Group (OUSPG) performed research on a subset of SIP related to the Invite message. The Invite message which SIP user agents and proxies are required to accept for set up sessions to be established. Links to the testing suite used, system impacts, vendors affected, and solutions are documented within the advisory.

SANS Institute offers the Internet Storm Center which can be accessed at URL http://isc.sans.org/. On line personnel called “Handlers” monitor network logs submitted from numerous parties to keep track on what is happening with Internet network traffic. The information from the logs is reported out in a daily Handler’s diary. Archived diaries, reports and security news can also be found on the site.

Anti-virus vendors such as Symantec and McAfee have web sites dedicated to report out on security threats as well. Symantec’s Security Response page can be accessed at URL http://securityresponse.symantec.com/ . Latest virus threats, security advisories, updates, and virus definitions are posted at this site. McAfee Security has a similar web site that post current threats and virus information located at URL http://us.mcafee.com/virusInfo/default.asp?cid=9043.

Operating System vendors have security sites that list threats and post fixes as well. Microsoft TechNet offers a security web page that can be accessed by going to URL http://www.microsoft.com/technet/Security/default.mspx. Security threats posted against Microsoft products are posted here. Red Hat maintains a similar page for security threats and fixes which can be accessed by going to URL http://www.redhat.com/security/. Keeping operating systems up to date is just as crucial to network and application stability as keeping anti-virus software up to date.

2.5 Conclusion

In the previous pages of this paper we have discussed a lot of information concerning implementing VoIP on an existing network. The first section gave a general over view of VoIP. The benefits of combining data and voice on an individual network were presented. We reviewed the VoIP stack and discussed how voice data is converted from analog to digital, compressed, and inserted into packets. The many protocols used to support transmission of voice packets was discussed as well.

The second section of the paper discussed some techniques on how to secure VoIP applications and systems. We looked at Cisco's Safe blue print which could be used to secure a network carrying voice. Some of the techniques listed in the Safe included segmenting networks to allow only voice or data. Hardening layer 2 and 3 devices is another technique discussed in Safe as well. How to apply other existing security such as encryption and authentication were discussed next. The final discussion was web sites and resources that implementers and network administrators could reference on a daily basis to monitor new security threats and keep their applications current with released security fixes.

VoIP can reduce corporation's communication costs significantly. Vendors are developing equipment and applications to embrace this technology. Combining voice and data on the same network can ease network manageability as long as the network administrator remembers voice and data threats are combined as well. After reading this paper I hope implementers and network administrators can see VoIP can be easily implemented on an existing network as long as the proper understanding, preparation, and security measures are in place.

3 References

[1] Tyson, Jeff. “How IP Telephony Works” 30 May 2004.
URL: http://computer.howstuffworks.com/ip-telephony.htm/

[2] Siemens Products. “The optiPoint 600 office”
URL: http://www.sipcenter.com/sip.nsf/html/Sponsors+Siemens

[3] Kuhn, Walsh, Fries. “Security Considerations for Voice Over IP Systems”
NIST Special Publication 800-58
URL: http://csrc.nist.gov/publications/drafts.html

[4] Zacker, Craig. “Windows 2000 Network Infrastructure Administration”

[5] Hardman, Dennis. “Noise and Voice Quality in VoIP environments”
URL: http://itresearch.forbes.com/detail/RES/1080147951_179.html

[6] Definition of Protocol
URL: http://www.webopedia.com/TERM/p/protocol.html

[7] SIP RFC 3261
URL: http://www.faqs.org/rfcs/rfc3261.html

[8] Free World Dialup X—Lite
URL: http://www.freeworldialup.com/

[9] Arora, Rakesh. “Voice over IP: Protocols and Standards”
URL: http://www.cse.ohio-state.edu/~jain/cis788-99/ftp/voip_protocols/index.html

[10] Definition of Jitter
URL:
http://searchnetworking.techtarget.com/sDefinition/0,,sid7_gci213534,00.html

[11] Kreiling, Janet. “Securing Voice in an IP Environment: Defense-In-Depth Strategies for Making Voice as Secure as any other Mission-Critical Application”
URL:
http://www.cisco.com/warp/public/784/packet/jan04/pdfs/PK16106C_SecureVOIP.pdf