VOICE VIDEO SESSION MANAGEMENT SECURITY REQUIREMENTS GUIDE (SRG) OVERVIEW

Version 2, Release 1

23 October 2020

Developed by DISA for the DoD

Trademark Information

Names, products, and services referenced within this document may be the trade names, trademarks, or service marks of their respective owners. References to commercial vendors and their products or services are provided strictly as a convenience to our users, and do not constitute or imply endorsement by DISA of any non-Federal entity, event, product, service, or enterprise.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

1. INTRODUCTION

1.1 Executive Summary

This Voice Video Session Management Security Requirements Guide (SRG) provides the technical security policies and requirements for applying security concepts to voice and video systems.

The protocol suites used for voice and video session management products are Session Initiation Protocol (SIP), H.323, and proprietary protocols such as Skinny Client Control Protocol (SCCP) and Unified Networks IP Stimulus (UNIStim). Within DoD, SIP, SCCP, and UNIStim are associated with Voice over IP (VoIP), Video enhanced VoIP (VVoIP), and Unified Capabilities (UC), while H.323 associates with Videoconferencing (VC). Only products providing session management functions (registration, session initiation, session records, etc.), such as enterprise session controllers, local session controllers, soft-switches, and gatekeepers, are within scope of this SRG. Session border controllers, proxies, media and signaling gateways, Multipoint Control Units, and endpoints are outside the scope of this document.

Products using this session management guidance will also rely on Network Device Management (NDM) SRG requirements to provide guidance for management of network devices containing the session manager network application.

1.1.1 Security Requirements Guides (SRGs)

SRGs are collections of requirements applicable to a given technology family. SRGs represent an intermediate step between Control Correlation Identifiers (CCIs) and Security Technical Implementation Guides (STIGs). CCIs represent discrete, measurable, and actionable items sourced from Information Assurance (IA) controls defined in a policy, such as the National Institute of Standards and Technology (NIST) Special Publication (SP) 800-53. STIGs provide product-specific information for validating and attaining compliance with requirements defined in the SRG for that product's technology area.

There are four core SRGs: Application, Network, Operating System, and Policy. Each addresses the applicable CCIs in the context of the technology family. Subordinate to the core SRGs, there are Technology SRGs developed to address the technologies at a more granular level.

This Voice Video Session Management SRG is based on the Network SRG. This Voice Video Session Management SRG contains general check and fix information that can be utilized for products for which STIGs do not exist.

The STIGs based on this SRG will provide the product-specific technical implementation guidance for that product. The STIG will contain the specific check and fix information for the product it covers.

SRG Hierarchy example:

Application SRG
|__Database SRG
|__MS SQL Server 2005 STIG

The SRG relationship and structure provides the ability to identify requirements that may be considered not applicable for a given technology family and provide appropriate justification. It also provides the structure to identify variations in specific values based on the technology family. These variations will be captured once and will propagate down to the Technology SRGs and then to the STIGs. This will eliminate the need for each product-specific STIG to address items that are not applicable.

1.1.2 SRG Naming Standards

In an effort to establish consistency across the SRGs, a naming standard for the Group Title and STIGIDs has been established.

Technology SRG Naming Standards

For Technology SRG Group Title and STIGIDs the following applies:

{Core SRG value}-{*Technology SRG*}-{5- or *6-digit numeric sequence number*}

Examples:

SRG-NET-000001-RTR-000001
SRG-APP-000001-COL-000001
SRG-NET-000001-VVSM-00001
SRG-OS-000001-UNIX-000001

Checks/Fixes will be included at this level in a general form. These checks and fixes will apply for any STIGs that are created for products that do not have product-specific check and fix guidance.

1.2 Authority

DoD Instruction (DoDI) 8500.01 requires that "all IT that receives, processes, stores, displays, or transmits DoD information will be […] configured […] consistent with applicable DoD cybersecurity policies, standards, and architectures" and tasks that Defense Information Systems Agency (DISA) "develops and maintains control correlation identifiers (CCIs), security requirements guides (SRGs), security technical implementation guides (STIGs), and mobile code risk categories and usage guides that implement and are consistent with DoD cybersecurity policies, standards, architectures, security controls, and validation procedures, with the support of the NSA/CSS, using input from stakeholders, and using automation whenever possible." This document is provided under the authority of DoDI 8500.01.

Although the use of the principles and guidelines in these SRGs/STIGs provide an environment that contributes to the security requirements of DoD systems, applicable NIST SP 800-53 cybersecurity controls need to be applied to all systems and architectures based on the Committee on National Security Systems (CNSS) Instruction (CNSSI) 1253.

1.2.1 Relationship to STIGs

The SRG defines the requirements for various technology families, and the STIGs are the technical implementation guidelines for specific products. A single SRG/STIG is not all-inclusive for a given system, which may include, but is not limited to: Network Device Management (NDM), Application Layer Gateway (ALG), Database, Web Server, and Voice Video Services Policy SRGs/STIGs. For a given system, compliance with all (multiple) SRGs/STIGs applicable to a system is required.

1.3 Vulnerability Severity Category Code Definitions

Severity Category Codes (referred to as CAT) are a measure of vulnerabilities used to assess a facility or system security posture. Each security policy specified in this document is assigned a Severity Category Code of CAT I, II, or III.

Table 1-1: Vulnerability Severity Category Code Definitions

	DISA Category Code Guidelines
CAT I	Any vulnerability, the exploitation of which will **directly and immediately** result in loss of Confidentiality, Availability, or Integrity.
CAT II	Any vulnerability, the exploitation of which **has a potential** to result in loss of Confidentiality, Availability, or Integrity.
CAT III	Any vulnerability, the existence of which **degrades measures** to protect against loss of Confidentiality, Availability, or Integrity.

1.4 SRG and STIG Distribution

Parties within the DoD and Federal Government's computing environments can obtain the applicable STIG from the Cyber Exchange website at https://cyber.mil/. This site contains the latest copies of STIGs, SRGs, and other related security information. Those without a Common Access Card (CAC) that has DoD Certificates can obtain the STIG from https://public.cyber.mil/.

1.5 Document Revisions

Comments or proposed revisions to this document should be sent via email to the following address: disa.stig_spt@mail.mil. DISA will coordinate all change requests with the relevant DoD

organizations before inclusion in this document. Approved changes will be made in accordance with the DISA maintenance release schedule.

1.6 Other Considerations

DISA accepts no liability for the consequences of applying specific configuration settings made on the basis of the SRGs/STIGs. It must be noted that the configuration settings specified should be evaluated in a local, representative test environment before implementation in a production environment, especially within large user populations. The extensive variety of environments makes it impossible to test these configuration settings for all potential software configurations.

For some production environments, failure to test before implementation may lead to a loss of required functionality. Evaluating the risks and benefits to a system's particular circumstances and requirements is the system owner's responsibility. The evaluated risks resulting from not applying specified configuration settings must be approved by the responsible Authorizing Official. Furthermore, DISA implies no warranty that the application of all specified configurations will make a system 100 percent secure.

Security guidance is provided for the Department of Defense. While other agencies and organizations are free to use it, care must be given to ensure that all applicable security guidance is applied both at the device hardening level as well as the architectural level due to the fact that some of the settings may not be able to be configured in environments outside the DoD architecture.

1.7 Product Approval Disclaimer

The existence of a STIG does not equate to DoD approval for the procurement or use of a product.

STIGs provide configurable operational security guidance for products being used by the DoD. STIGs, along with vendor confidential documentation, also provide a basis for assessing compliance with Cybersecurity controls/control enhancements, which supports system Assessment and Authorization (A&A) under the DoD Risk Management Framework (RMF). DoD Authorizing Officials (AOs) may request available vendor confidential documentation for a product that has a STIG for product evaluation and RMF purposes from disa.stig_spt@mail.mil. This documentation is not published for general access to protect the vendor's proprietary information.

AOs have the purview to determine product use/approval IAW DoD policy and through RMF risk acceptance. Inputs into acquisition or pre-acquisition product selection include such processes as:

- National Information Assurance Partnership (NIAP) evaluation for National Security Systems (NSS) (http://www.niap-ccevs.org/) IAW CNSSP #11

- National Institute of Standards and Technology (NIST) Cryptographic Module Validation Program (CMVP) (http://csrc.nist.gov/groups/STM/cmvp/) IAW Federal/DoD mandated standards
- DoD Unified Capabilities (UC) Approved Products List (APL) (http://www.disa.mil/network-services/ucco) IAW DoDI 8100.04

2. ASSESSMENT CONSIDERATIONS

2.1 NIST SP 800-53 Requirements

All applicable baseline technical NIST SP 800-53 requirements and security best practice requirements are included in this SRG.

CNSSI 1253 defines the required controls for DoD systems, based on confidentiality, integrity, and availability (baseline) of the given information system. In all cases, CNSSI 1253, along with required baselines, will serve as the policy requirement for any given asset or information system.

2.2 General Procedures

This SRG has procedures that are intended to provide appropriate evaluation and remediation functions for a typically configured system. These procedures are not product specific and are intended for use when a product-specific STIG is not available.

2.3 Voice Video Assessment Guidance

To assess Voice Video components and systems, the following resources apply.

2.3.1 Video Services Policy

Policy and architectural guidance for implementing systems on the DoDIN is contained in two documents. The Voice Video Services Policy STIG provides the policy and architectural guidance for VoIP systems (also referred to as UC systems or implementations) used to support the DoD. The Video Services Policy STIG contains the policy and architectural guidance for Video Conference (VC) systems in use within DoD. Some overlap exists between the documents. When VoIP session managers are fielded, the Voice Video Services Policy STIG is applicable. When VC session managers are fielded, the Video Services Policy STIG is applicable.

2.3.2 Network Device Management (NDM)

Network devices usually contain a management component to handle administration of the network device itself. NDM security practices and procedures applicable to the management of all DoD network devices are contained in the NDM SRG. The NDM guidance works with the technical requirements in other SRGs. Vendors of session management products will use the NDM SRG for the management plane and the Voice Video Session Management SRG for the control and data planes of the device.

2.3.3 Voice Video Session Management

Session managers for voice and video systems will rely on the Voice Video Session Management SRG for technical guidance. The protocol suites used for voice and video session management

products include Session Initiation Protocol (SIP), H.323, and proprietary protocols such as Skinny Client Control Protocol (SCCP) and Unified Networks IP Stimulus (UNIStim). For DoD, SIP, H.323, SCCP, and UNIStim are associated with VoIP and VC sessions. Currently, most session managers handle multiple protocols.

2.3.4 Voice Video Border Elements

Voice video border elements are products providing services at the border and within enclaves in support of the voice video system. These products often work in parallel with the data firewalls, providing routing and conversion of voice video transmissions. Border elements rely on the Back-to-Back User Agent (B2BUA) function of the enterprise network Session Border Controller (SBC). SBCs perform inspection and proxy functions for specific ports and protocols used by voice and video signaling and media. Gateways enable communication between voice video networks and other networks, such as PSTN or ISDN networks. Border elements will use the guidance in the Application Layer Gateway (ALG) SRG for the technical implementation of these devices and devices with this functionality.

2.3.5 Voice Video Endpoints

Voice video endpoints include VoIP hardware phones, VC desktop terminals, UC and VC soft clients, and VC Coders/Decoders (CODECs) used in conference rooms with multiple cameras, microphones, and displays. The guidance for these devices is contained in the Voice Video Endpoint SRG.

3. CONCEPTS AND TERMINOLOGY CONVENTIONS

3.1 Overview

Systems incorporating voice and video services have evolved from circuit-switched analog systems to efficient digital packet-switched networks as bandwidth and reliability improved. Addressing security concerns for current voice video systems operating on the DoD Information Network (DoDIN) requires an understanding of the architecture and how the various components rely on session management. The most common protocols used for communications of voice and video are SIP, H.323, and proprietary protocols such as SCCP and UNIStim, which are discussed in this section. Many session managers serve SIP, H.323, proprietary protocols, or a combination of these.

3.2 Unified Capabilities

UC are the integration of voice, video, and/or data services delivered ubiquitously across a secure and highly available network infrastructure, independent of technology, to provide increased mission effectiveness to the warfighter and business communities. The DoD Unified Capabilities Requirements (UCR) specifies the technical requirements for certification of approved products to be used in DoD networks to provide end-to-end UC and is available for download from http://www.disa.mil/network-services/UCCO/Policies-and-Procedures.

3.3 Session Initiation Protocol

SIP is a communications protocol for signaling and controlling multimedia communication sessions defined in Request For Comment (RFC) 3261 maintained by the Internet Engineering Task Force (IETF). The protocol defines the messages sent between endpoints and servers, controlling the establishment, termination, and other essential elements of a call. SIP is an application layer text-based protocol designed to be independent of the underlying transport layer.

SIP servers create, modify, and terminate sessions consisting of one or several media streams. SIP relies on several application layer protocols that identify and carry the session media. Media identification and negotiation is achieved with the Session Description Protocol (SDP). For the transmission of voice and video media streams SIP typically employs the Real-time Transport Protocol (RTP) or Secure Real-time Transport Protocol (SRTP). The protocol may be encrypted with Transport Layer Security (TLS) to secure transmissions of SIP messages.

3.3.1 Registration of Endpoints

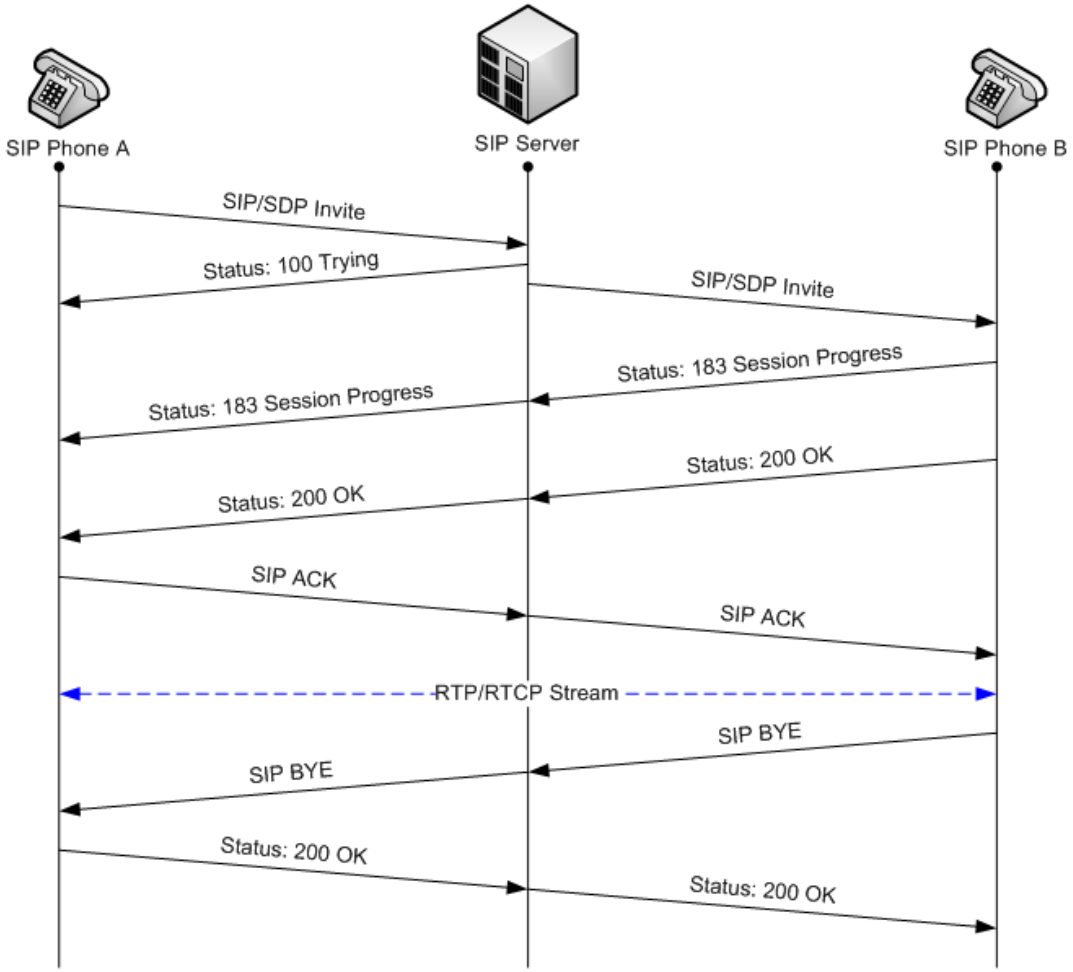
Registration is the process in which the endpoint authenticates to the SIP server (registrar) to let it know the location, availability, and user capabilities. The process includes the transfer of configuration files and the user's contact list from the server to the endpoint. Upon receipt, the SIP server sends back a challenge and upon receiving back a correct response, the SIP server validates the user's credentials and registers the user in its contact database. It then responds with an acknowledgement, which includes the user's current contact list in contact headers.

Registration will then update on a regular schedule with the endpoint, sending the list of addresses where the SIP server will redirect or forward requests.

3.3.2 Session Progression

The SIP session progresses at the application layer when a caller initiates establishment of the session through a SIP server to another endpoint as diagramed in Figure 3-1. Unless both the call imitator and receiver are on the same enclave, the SIP proxy in the diagram is usually a series of SIP servers acting as local and enterprise session managers. SIP sends parameters using SDP (Session Description Protocol) in the invitation to identify the form the audio or video will use. Once both endpoints agree and are ready to start exchanging media or data, RTP (Real-time Transport Protocol) is used to exchange the data or voice packets. The RTP streams will be carried on a port from a range of ports, which are assigned to each endpoint after they negotiate and accept a particular port on each side. The SIP session terminates after an endpoint sends the termination request and then receives acknowledgement from the distant end.

Figure 3-1: SIP Request, Acceptance, Setup, and Termination



3.3.3 Common Log Format

Logging of session events for SIP is best performed according to RFC 6873, which identifies the Common Log Format (CLF) for SIP. The CLF is analogous to the Call Detail Records (CDRs) used in packet-switch networks and provides a non-proprietary framework for logging essential session information. This CLF mimics the successful event logging format found in well-known web servers like Apache, which provides familiarity for administrators. It allows session correlation across diverse processing elements. In operational SIP networks, a request will typically be processed by more than one SIP server and the CLF allows the administrator to trace the progression of the requests as they traverse through the different servers, establishing a concise diagnostic trail of a SIP session.

3.3.4 Assured Services Session Initiation Protocol

Building on SIP and providing necessary end-to-end assured service for the DoDIN is Assured Services Session Initiation Protocol (AS-SIP). AS-SIP provides support for Multi-Level Precedence and Preemption (MLPP) that establishes communications priorities based on user authorizations. DoD supporting Command and Control (C2) communications relies on the implementation of AS-SIP and MLPP to ensure that flag officers and senior staff are provided higher priority VoIP communications than other users. The specification requires implementing SRTP, TLS, and Differentiated Services Code Point (DSCP), and how the protocol is to be implemented on both unclassified and classified networks.

3.4 H.323 System Specification

H.323 is a standard approved by the International Telecommunication Union (ITU) to promote compatibility in videoconference transmissions over IP networks. H.323 is a recommendation that sets standards for multimedia communications over LANs that do not provide a guaranteed Quality of Service (QoS). Although it was unknown if manufacturers would support H.323, it is now widely implemented by voice and videoconferencing equipment manufacturers because of its call control and management for both point-to-point and multipoint conferences as well as gateway administration of media traffic, bandwidth, and user participation.

Media streams are transported using RTP and RTP Control Protocol (RTCP). RTP carries the actual media and RTCP carries status and control information. The signaling is transported reliably over TCP. The H.323 standard relies on a number of other standards and protocols to provide supplementary services and functionality.

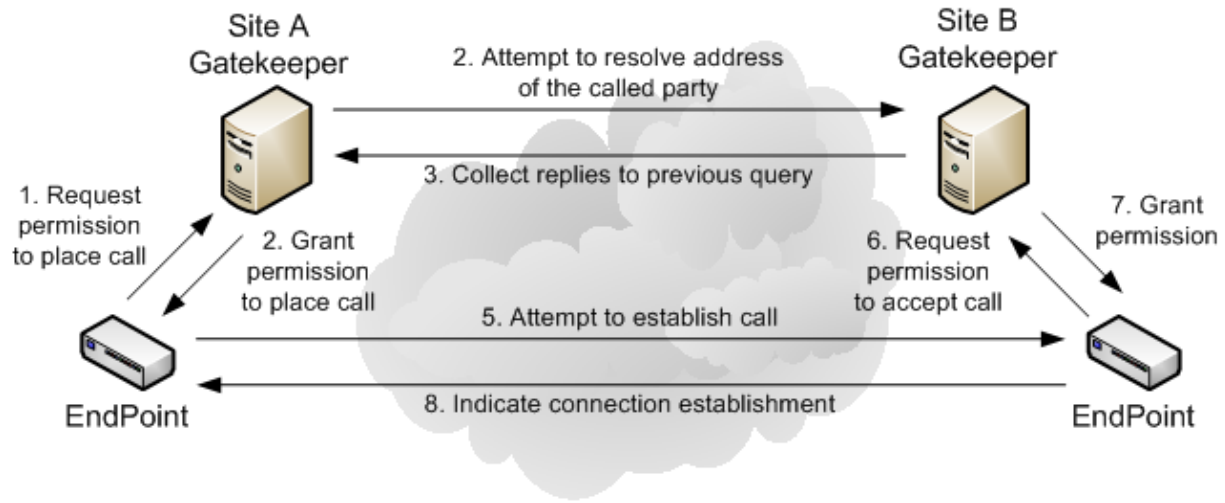
3.4.1 H.225 Registration, Admission, and Status

H.225 defines a Registration, Admission, and Status (RAS) channel to carry messages used in the gatekeeper discovery and endpoint registration processes. An endpoint at power on will send a request message to locate gatekeepers willing to provide service. Gatekeepers then respond with a confirmation message and the endpoint then selects a gatekeeper to work with. Once the endpoint determines the gatekeeper to work with, it will attempt to register and the gatekeeper will respond. At this point, the endpoint is known to the network and can make and place calls.

3.4.2 H.225 Call Signaling

Once the endpoint address is resolved, it uses H.225 Call Signaling to establish communication with a remote entity. For an endpoint to place a call, it requests admission to the gatekeeper. The gatekeeper resolves the address locally, by consulting another gatekeeper, or by querying another network service. Then the gatekeeper returns the address of the remote endpoint in the admission confirm message and the endpoint can then place the call.

Figure 3-2: H.323 Call Establishment Using a Gatekeeper



Upon receiving a call, a remote endpoint will also send an Admission Request (ARQ) and receive an Admission Confirm (ACF) in order to get permission to accept the incoming call. This is necessary, for example, to authenticate the calling device or to ensure that there is available bandwidth for the call.

3.4.3 H.245 Multimedia Communication

The H.245 control protocol for multimedia communication specifies master/slave determination, multimedia capability exchange, multimedia logical channel opening and closing, and other control functionality. H.245 conveys information needed for multimedia communication, such as encryption, flow control, jitter management, preference requests, and the opening and closing of logical channels used to carry media streams. It defines separate send and receive capabilities and the means to send these details to other devices that support H.323. Additionally, H.245 offers the possibility to be tunneled within H.225 call signaling messages to ease firewall traversing.

3.4.4 H.235 Security

H.235 series describes security within H.323, including security for signaling and media. H.235 provides enhancements to incorporate security services such as Authentication and Privacy (data

encryption). H.235 works with other H series protocols that use H.245 as their control protocol. An H.235 aware gatekeeper can assure that trusted H.323 endpoints are granted access to the gatekeeper's services, which include RAS and call control. NIST Special Publication 800-58 Section 4.2 discusses H.235 security in detail.

3.5 Other Session Management Protocols

A number of protocols are used in addition to the SIP and H.323 families. Many of these are proprietary, such as SCCP and UNIStim. Others are specific to device functionality, such as Media Gateway Control Protocol (MGCP). Many manufacturers of session managers, such as Cisco and Avaya, developed their own proprietary protocols because standard protocols were still in their infancy. Over time, these proprietary protocols are being replaced or complemented by standardized protocols, including H.323 and SIP. The greatest concern with proprietary protocols is a distinct lack of security in most. For DoD, only secured protocols or unsecure protocols securely encapsulated may be used.

3.5.1 Skinny Call Control Protocol (SCCP)

SCCP is a lightweight IP based protocol for session signaling. An SCCP client uses TCP/IP to communicate with one or more Cisco Unified Communications Manager (CUCM) applications in a cluster. It uses the RTP over User Datagram Protocol (UDP)-transport for media traffic with other Skinny clients or an H.323 terminal. SCCP is a stimulus-based protocol and a communications protocol for hardware endpoints and other embedded systems having significant CPU and memory constraints. Some Cisco analog media gateways register and communicate with a CUCM using SCCP. Other vendors have implemented SCCP in VoIP terminals and IP phones, media gateways, and softswitches. An open source version of the protocol is also available.

3.5.2 Unified Networks IP Stimulus (UNIStim)

UNIStim is a communications protocol originally developed by Nortel (now Avaya) for IP phone and IP PBX communications. The protocol works using a master/slave mode of operations. UNIStim communicates user actions from a terminal and the commands sent to the terminal. Like SCCP, UNIStim is a stimulus-based protocol, and can implement new phones without modifying software embedded in the terminals. Efforts to standardize UNIStim draw on work with pre-standard UNIStim and already developed work on the Nortel IP PBX systems and its IP Centrex platforms from 1996. The UNIStim protocol is currently implemented on Avaya systems.

3.5.3 Media Gateway Control Protocol (MGCP)

MGCP controls media gateways on IP networks connected to the public switched telephone network (PSTN). The protocol architecture and programming interface is described in RFC 2805 and the current definition is RFC 3435 which overrides RFC 2705. MGCP is a successor to the Simple Gateway Control Protocol (SGCP) developed by Bellcore and Cisco. MGCP is a text-

based signaling and call control communications protocol used in VoIP systems. MGCP uses the SDP for specifying and negotiating the media streams to be transmitted in a call session and the RTP for framing of the media streams.

3.5.4 Additional Proprietary Protocols

A number of vendors use proprietary protocols for session management and communications among voice video components. Cisco's SCCP is used by a number of other vendors, to include Asterisk and Digium. Avaya's UNIStim is used by a limited number of vendors, including Asterisk. Asterisk's own Inter-Asterisk Exchange (IAX) protocol establishes connections between clients and Asterisk servers or between two Asterisk PBX units. Microsoft's Skype protocols are closed source and therefore less is known about their operation. Ports, Protocols, and Services Management (PPSM) Category Assurance List (CAL) and Vulnerability Assessments (VAs) specify how protocols may be used on DoD networks.

3.6 Fire and Emergency Services

The FCC requires interconnected VoIP telephone services using the Public Switched Telephone Network (PSTN) meet Enhanced 911 (E911) obligations. Fire and Emergency Services (FES) rely on E911 systems to automatically provide to emergency service personnel a 911 caller's call back number through Automatic Number Identification (ANI) and, in most cases, location information through extended Automatic Location Identification (ALI) information or access to an extended ALI database. Providing 911 service is mandatory and cannot be opted out.

To reduce possible risks to public safety, functionality supporting FES and E911 must be implemented for voice systems. Customers must have a clear understanding of the limitations, if any, of their 911 service. Labels warning customers must be used if 911 service is limited or not available and customers must place the labels on or near equipment used with VoIP service. Calls must be routed to the Public Safety Answering Point (PSAP) in areas where emergency service providers are not capable of receiving or processing the location information or call back numbers not automatically transmitted with 911 calls.