VOICE VIDEO ENDPOINT SECURITY REQUIREMENTS GUIDE (SRG) TECHNOLOGY OVERVIEW

Version 1, Release 7

26 January 2018

Developed by DISA for the DoD

Trademark Information

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

1. INTRODUCTION

1.1 Executive Summary

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

Voice Video Endpoints communicate with Voice Video Session Managers over DoD networks using Session Initiation Protocol (SIP), H.323, and proprietary protocols such as Skinny Client Control Protocol (SCCP) and Unified Networks IP Stimulus (UNIStim) to register endpoint devices, initiate sessions, and share status information. Voice Video Endpoints communicate directly as peers using Real-time Transport Protocol (RTP), RTP Control Protocol (RTCP), Secure RTP (SRTP), Secure RTP Control Protocol (SRTCP), Session Description Protocol (SDP), and proprietary protocols. Products providing endpoint functionality for Voice over IP (VoIP), Video Enhanced VoIP (VVoIP), Unified Capabilities (UC), and Videoconferencing (VC) are within the scope of this SRG. Session managers, border elements, and supporting infrastructure are outside the scope of this document, to include enterprise session controllers, local session controllers, gatekeepers, media and signaling gateways, session border controllers, and proxies.

To produce complete guidance for Voice Video Endpoints, additional resources may be required in conjunction with this document. Products using this endpoint guidance will also rely on the Network Device Management (NDM) SRG requirements to provide guidance for management of endpoint network devices. Other endpoints, such as multipoint control units, may require additional considerations to develop guidance.

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 Endpoint SRG is based on the Network SRG. This Voice Video Endpoint SRG contains general check and fix information that can be used 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 provides 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: Database, Web Server, and Domain Name System (DNS) 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 Information Assurance Support Environment (IASE) website. This site contains the latest copies of any STIGs, SRGs, and other related security information. The address for the IASE site is http://iase.disa.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 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 SIP, H.323, and proprietary protocols such as SCCP and 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.

2.4 On-Hook Audio Security

All unclassified voice video endpoints deployed within a Sensitive Compartmented Information Facility (SCIF) must be National Telecommunications Security Working Group (NTSWG) approved devices in accordance with the CNSSI 5000 and CNSSI 5001. Compliance with federal standards helps establish a common security baseline across DoD that reflects the most restrictive security posture consistent with operational requirements. In common configurations, voice video endpoints can transmit conversations in secure areas over unclassified networks. Voice video endpoint microphones, speakers, and supporting electronics may pick up nearby conversation audio and conduct it over the network connection, even when the endpoint is on-hook, powered or not. The Technical Surveillance Counter-Measures (TSCM) program protects sensitive government information, to include classified information, through the establishment of on-hook audio security standards. Voice video endpoints certified by NTSWG are modified to prevent this behavior or limit it to within acceptable levels.

3. CONCEPTS AND TERMINOLOGY CONVENTIONS

3.1 Overview

Systems incorporating voice 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 endpoints rely on various supporting components. The most common protocols used for media communication by voice video endpoints is the RTP/RTCP pair and the SRTP/SRTCP pair. For signaling, voice video endpoints use SIP, H.323, and proprietary protocols such as SCCP and UNIStim.

3.2 Unified Capabilities (UC)

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 Media Protocols

Media protocol pairs provide the method to transport voice video streams and control information across networks while maintaining suitable quality and efficiency. The media protocols streams also contain information identifying CODECs used, encryption details, and the data itself. The sections below describe the common protocols used by voice video peers for media streams.

3.3.1 Real-Time Protocol (RTP)

The most common protocol for streaming audio and video packets is RTP as defined by the Internet Engineering Task Force (IETF) in Request For Comment (RFC) 3550 over User Datagram Protocol (UDP). The payload format for a number of CODECs is defined in RFC 3551, although payload format specifications are also defined in documents published by the International Telecommunication Union (ITU) and in other IETF RFCs. Real-time multimedia streaming requires timely delivery of information but tolerates some loss of packets by using suitable error concealment algorithms. RTP is used in conjunction with H.323, SIP, and most proprietary session management systems. Information provided by RTP includes timestamps for synchronization, sequence numbers for packet loss and reordering detection, and the payload format, which indicates the encoded format of the data.

3.3.2 Real-Time Control Protocol (RTCP)

RTP addresses additional issues through RTCP, which is also defined in RFC 3550, to provide mechanisms to help address delay and jitter. RTCP periodically sends control information and

Quality of Service (QoS) feedback and synchronization between the media streams. RTCP traffic bandwidth is small compared to RTP, typically around 5 percent.

3.3.3 SRTP and SRTCP

One area of concern when communicating over public networks like the Internet is the potential for eavesdropping. To address these security concerns, Secure RTP (SRTP), defined in IETF RFC 3711, is a security profile for RTP that adds confidentiality, message authentication, and replay protection to that protocol. SRTP is ideal for protecting media traffic because it can be used in conjunction with header compression with no effect on IP QoS. Secure RTCP (SRTCP) provides the same security-related features to RTCP as the ones provided by SRTP to RTP. For DoD networks, the use of SRTP and SRTCP are recommended for all communications.

3.4 Signaling Protocols

Signaling protocols provide the method by which voice video systems communicate among devices to register with a session manager, initiate and terminate sessions, and communicate with other devices. The most common signaling protocol families are SIP and H.323, but several proprietary protocols are also used, such as SCCP and UNIStim. The sections below describe the common protocols used by voice video peers for signaling. A more in-depth discussion of these protocols is available in the Voice Video Session Management SRG Overview.

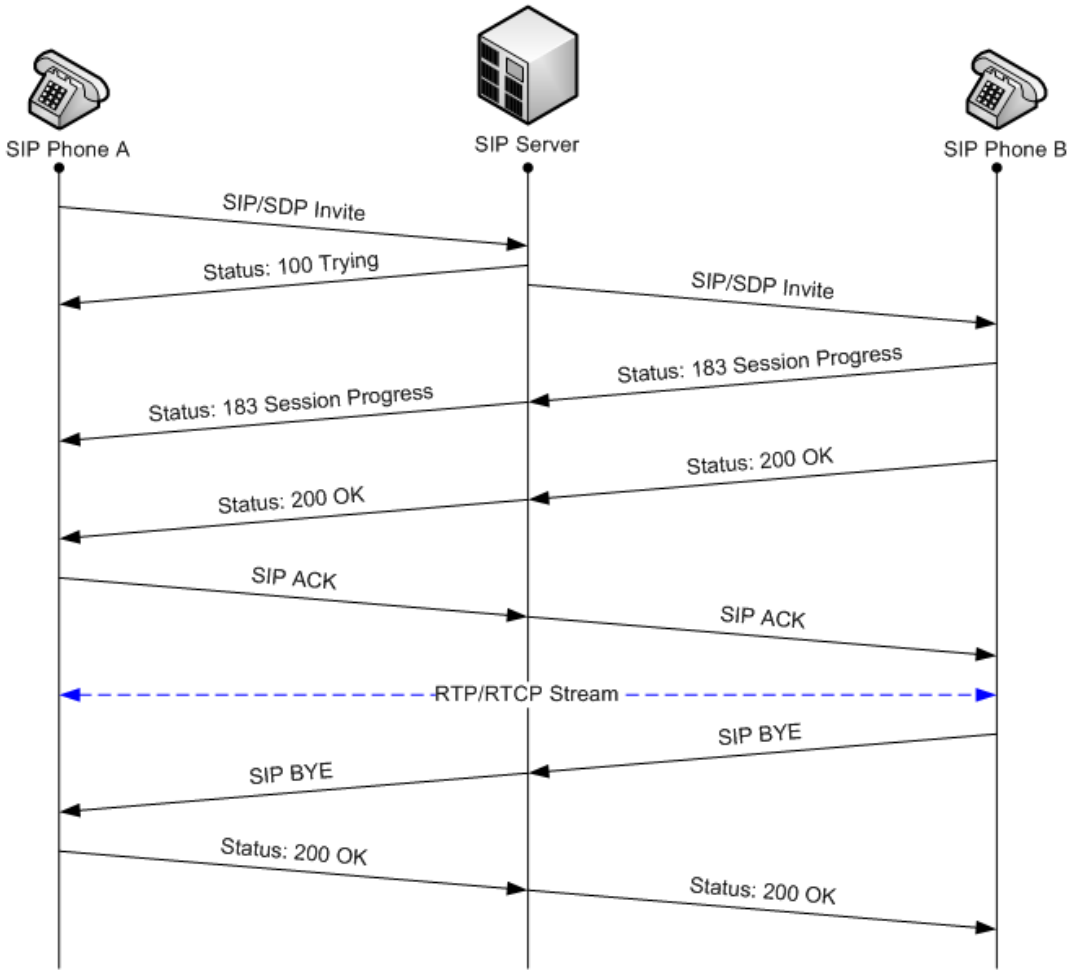
3.4.1 Session Initiation Protocol

SIP is a communications protocol for signaling and controlling multimedia communication sessions defined in IETF RFC 3261. 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 provides mechanisms for registration of endpoints, call progression, and streaming to gateways.

3.4.2 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.

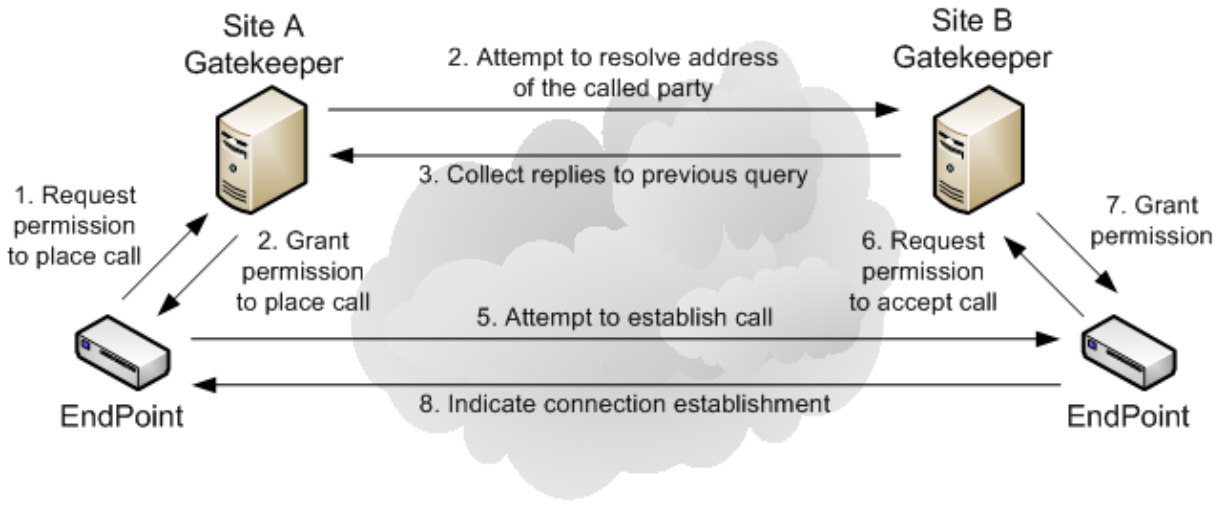
Figure 3-1: SIP Request, Acceptance, Setup, and Termination



3.4.3 H.323 System Specification

H.323 is a standard approved by the ITU to promote compatibility of videoconference transmissions over IP networks. H.323 is a recommendation that sets standards for multimedia communications over LANs that do not provide a guaranteed QoS. Although it was unknown if manufacturers would support H.323, it is now widely implemented by voice video 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. 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, such as registration, call signaling, security, and multimedia communications.

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



3.4.4 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 Fire and Emergency Services

The FCC requires that 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 callback 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 callback numbers not automatically transmitted with 911 calls.

3.6 Voice Video Endpoints in Secured Spaces

The use of unclassified voice video endpoints within secured spaces, where classified and sensitive conversations may be overheard requires additional considerations. Requirements are more stringent for deploying voice video endpoints in a SCIF than for deploying in, non-SCIF classified discussion areas.

3.6.1 Deployment in SCIFs

The use of unclassified voice video endpoints within secured spaces, where classified and sensitive conversations may be overheard requires additional considerations. Requirements are Dependent on the level of conversation within a secure area, Media protocol pairs provide the method to transport voice and video streams and control.

3.6.2 Deployment in Other Classified Areas

The use of unclassified voice video endpoints within secured spaces, where classified and sensitive conversations may be overheard requires additional considerations. Requirements are Dependent on the level of conversation within a secure area, Media protocol pairs provide the method to transport voice and video streams and control.