AFRL-RI-RS-TR-2023-108

GOVERNMENT SECURE VOICE ARCHITECTURE

TEXAS A&M UNIVERSITY

JUNE 2023

FINAL TECHNICAL REPORT

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

STINFO COPY

AIR FORCE RESEARCH LABORATORY
INFORMATION DIRECTORATE

■ **AIR FORCE MATERIEL COMMAND** ■ **UNITED STATES AIR FORCE** ■ **ROME, NY 13441**

NOTICE AND SIGNATURE PAGE

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data does not license the holder or any other person or corporation; or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

This report is the result of contracted fundamental research deemed exempt from public affairs security and policy review in accordance with SAF/AQR memorandum dated 10 Dec 08 and AFRL/CA policy clarification memorandum dated 16 Jan 09. This report is available to the general public, including foreign nations. Copies may be obtained from the Defense Technical Information Center (DTIC) (http://www.dtic.mil).

AFRL-RI-RS-TR-2023-108 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

FOR THE CHIEF ENGINEER:

/ S /
TODD CUSHMAN
Work Unit Manager

/ S /
JAMES PERRETTA
Deputy Chief,
Information Warfare Division
Information Directorate

This report is published in the interest of scientific and technical information exchange, and its publication does not constitute the Government's approval or disapproval of its ideas or findings

REPORT DOCUMENTATION PAGE

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ORGANIZATION.

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED	
		START DATE	END DATE
JUNE 2023	FINAL TECHNICAL REPORT	SEPTEMBER 2020	APRIL 2023

4. TITLE AND SUBTITLE

GOVERNMENT SECURE VOICE ARCHITECTURE

5a. CONTRACT NUMBER	5b. GRANT NUMBER	5c. PROGRAM ELEMENT NUMBER
FA8750-20-2-1005	N/A	
5d. PROJECT NUMBER	**5e. TASK NUMBER**	**5f. WORK UNIT NUMBER**
		R32L

6. AUTHOR(S)

Walt Magnussen, Michael E. Fox, Eman Hammad, Henning Schulzrinne

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Texas A&M University
400 Harvey Mitchell PKY S STE 300
College Station TX 77845-4375

8.PERFORMING ORGANIZATIONREPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10.SPONSOR/MONITOR'S ACRONYM(S)	11.SPONSOR/MONITOR'S REPORT NUMBER(S)
Air Force Research Laboratory/RIGA 525 Brooks Road Rome NY 13441-4505	AFRL/RI	AFRL-RI-RS-TR-2023-108

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited. This report is the result of contracted fundamental research deemed exempt from public affairs security and policy review in accordance with SAF/AQR memorandum dated 10 Dec 08 and AFRL/CA policy clarification memorandum dated 16 Jan 09.

13. **SUPPLEMENTARY NOTES**

14. ABSTRACT

While the basic tools to secure the voice network have been there from the beginning, there has been little to implement them. One reason for this could be the fact that until now there has been no clear documentation of a comprehensive architecture defined to do so. The Department of Homeland Security, Science and Technology Division initiated this study in the fall of 2020 with the purpose of designing, developing testing and documenting a secure voice architecture that could be implemented within and between all federal agencies.

15. SUBJECT TERMS

Secure voice network, mobile-to-mobile secure communications, secure voice, legacy and emerging 5G cellular networks, securing new networks devices, architecture, testbed

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES
a. REPORT	b. ABSTRACT	C. THIS PAGE		
U	U	U	SAR	27

19a. NAME OF RESPONSIBLE PERSON	19b. PHONE NUMBER *(Include area code)*
TODD CUSHMAN	**N/A**

PREVIOUS EDITION IS OBSOLETE.

STANDARD FORM 298 (REV. 5/2020)
Prescribed by ANSI Std. Z39.18

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1.0 SUMMARY

This project, intended to design, install, test and document Secure Voice Architecture kicked off November 12th, 2020. This project brings together a set of Principle Investigators known to be leaders in the voice services, testing and security. Dr. Walt Magnussen, who has 25 years of experience in VoIP systems and Mr. Michael Fox who leads all the operational aspects of the Internet2 Technology Evaluation Center (ITEC) led this project. Dr. Henning Schulzrinne from Columbia University is one of the principal authors of the SIP standards and Dr. Eman Hammad from TAMU (Texas A&M University) Commerce provided the Cybersecurity expertise. The project was managed by Ms. Anjuli "A.J." Renold, also from ITEC.

The project required the installation of a testbed designed to emulate three separate agencies using voice systems used by federal agencies today. After discussion and review of federal agency systems, the team identified Cisco Call Manager, Avaya Aura and Cisco/BroadSoft solutions. Many agencies use service provided VoIP services, also referred to as Centrex. The BroadSoft solution provides core voice networks for AT&T, Verizon, and Lumen. Utilizing the testbed, the team installed and assessed each of the services in the architecture.

The lab platform was extended over an IPSec tunnel to Columbia University and to TAMU Commerce to allow access for the Co-PIs and their teams.

The team hosted a virtual and in-person stakeholder meeting in Washington DC to present findings. A summary of the findings is recorded and available on the ITEC website. The team also briefed FCC (Federal Communications Commission) on STIR/SHAKEN, part of the architecture.

Meetings between Texas A&M University, TAMU Commerce, Columbia University and AFRL have been held monthly, with a technical deep dive into one of the features each quarter. Quarterly reports have been published describing development progress and highlighting next steps for the coming quarter.

2.0 INTRODUCTION

About the turn of the last century, we began to carry voice traffic over networks. For almost the entire century this traffic was circuit switched with dedicated resources being allocated to each caller's traffic. It began with manual connections using cord boards and operators patching the connections. The first major enhancements were that of analog step offices that used a set of relays to make the connections. The 1960s and 1970s brought in the Digital Offices and the first out of band signaling system (Signaling System Seven (SS7). During this time, there was little thought given to securing the voice traffic through encryption or other means with a few exceptions such as secure voice used by the department of defense.

The 1980s saw a revolutionary change in data communications with the invention and rapid adoption of the Internet and the Internet Protocol (IP). While the predominant use of the early Internet was data the idea of using this newly found network for voice services was not far behind. In 1996 the International Telecommunication Union (ITU) which is a specialized unit within the United Nations published the H.323[1] specification. This standard is a multimedia specification which was widely used for video conferencing but did not find much success in providing voice services due to it's lack of features such as call transfer. Cisco did add the required functionality to H.323 in their proprietary Skinny Client Control Protocol (SCCP)[2]. Skinny and other vendor specific IP protocols dominated the Voice over IP market for the next several years. Still there was little to no attention given to the thought of securing the voice networks.

In June of 2002 the Internet Engineering Task Force (IETF published RFC 3261 which defined the Session Initiation Protocol (SIP)[3]. SIP was also a multimedia protocol but it and it's associated protocols did define enough features to support a very feature rich standards based voice platform. With that the transition to an all standards-based ecosystem began,

From the outset, SIP took security into consideration, RFC 3261 sections 26.4.3 and 26.4.4 described how to secure SIP or the signaling portion of VoIP using RFC 5246 Transport Layer Security (TLS)[4]. Shortly thereafter, the IETF defined a method for securing the media portion of the VoIP call in RFC 3711 which describes the use of the Secure Real Time Protocol (SRTP)[5]. While the basic tools to secure the voice network have been there from the beginning, there has been little to implement them. One reason for this could be the fact that until now there has been no clear documentation of a comprehensive architecture defined to do so.

---

[1] https://www.itu.int/rec/T-REC-H.323

[2] https://en.wikipedia.org/wiki/Skinny_Client_Control_Protocol

[3] https://www.rfc-editor.org/rfc/pdfrfc/rfc3261.txt.pdf

[4] https://www.ietf.org/rfc/rfc5246.txt

[5] https://www.rfc-editor.org/rfc/rfc3711

2.1 Towards Zero Trust Voice

In November of 2022 the Department of Defense CIO Zero Trust Portfolio Management Office released their DoD CIO Zero Trust Strategy report[6]. The Executive Summary begins with the following statement

> "*Incremental improvements will not give us the security we need; instead, the Federal Government needs to make bold changes and significant investments in order to defend the vital institutions that underpin the American way of life." — Executive Order on Improving the Nation's Cybersecurity 12 May 2021".*

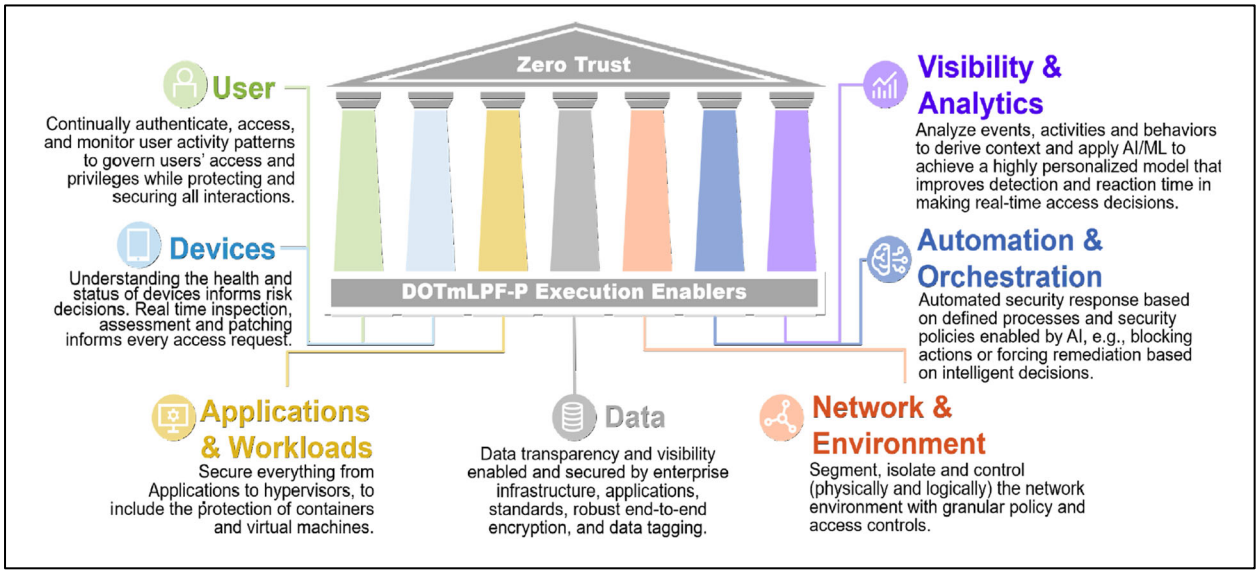

Figure 1: DoD Zero Trust Pillars

The above diagram describes the seven pillars of the DoD Zero Trust architecture. These tenants are as much a part of the voice services as they are of any other service that rides on our network infrastructure. Throughout the document several elements that are a part of the VoIP infrastructure such as devices, applications, data, network, and analytics are discussed. In fact, they all are significant aspects of one of the seven pillars of the Zero Trust Capabilities. In spite of VoIP being implicitly included in the architecture, it was never explicitly mentioned.

On November 16, 2018, President Trump signed into law the Cybersecurity and Infrastructure Security Agency Act of 2018[7], which elevated the mission of the former NPPD within DHS, establishing the Cybersecurity and Infrastructure Security Agency (CISA). Whereas the DoD is responsible for ZTA and cybersecurity within the Department of Defense, CISA holds the same responsibility for all civilian infrastructure within the United States. Several actions have been taken in this space including Office of Management and Budget Executive Order 22-09 entitled "*Moving*

[6] https://dodcio.defense.gov/Portals/0/Documents/Library/DoD-ZTStrategy.pdf

[7] https://en.wikipedia.org/wiki/Cybersecurity_and_Infrastructure_Security_Agency_Act

the U.S. Government Toward Zero Trust Cybersecurity Principles [8] '. This document establishes a mandate for all federal agencies to establish a zero trust plan and specifically calls for the encryption of DNS and HTTP, two protocols widely deployed in voice services and covered in detail in this study. Another good report is the "*CISA Zero Trust Maturity Model Version 2.0*" released in April of 2023[9] which is closely related to the Zero Trust Architecture document released by the Department of Defense.

2.2 The Purpose of This Study

The Department of Homeland Security, Science and Technology Division awarded this study in the fall of 2020 with the purpose of designing, developing testing and documenting a secure voice architecture that could be implemented within and between all federal agencies. There were a few assumptions that had to be made at the outset. The first and most important was that the proposed had to be built tested on standards-based voice systems that were commonly used by agencies today. Any solution that was going to require the wholesale replacement of all the equipment in place today was not considered feasible.

2.3 A Path Forward

It is unclear as to why securing voice services has garnered so little attention. It could be a false sense of trust that we seem to inherently have in the voice services, it could be that there is not much information dissemination about voice services vulnerabilities, or it could be that there was just not a plan on how to do so before this document.

While there are few if any requirements for securing voice traffic today there is significant reason to at least consider doing so. Today the navy does require TLS and SRTP encryption using CAC cards for teleworking. This covers Confidential Unclassified Information (CUI) at the IL4 level. These levels are shown below[10];

[8] https://www.whitehouse.gov/wp-content/uploads/2022/01/M-22-09.pdf

[9] https://www.cisa.gov/sites/default/files/2023-04/zero_trust_maturity_model_v2_508.pdf

[10]

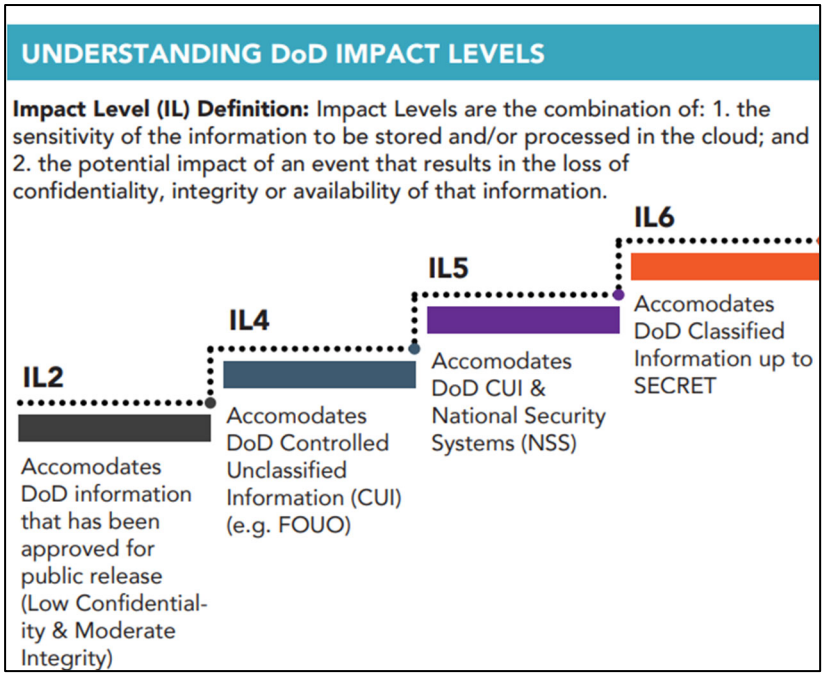

Figure 2: DoD Impact Levels

This study lays out a clear step by step plan for rolling out a secure voice network. This will not come at no cost but then none of the rest of the path towards zero trust will either. As noted in CISA's. Zero Trust Maturity Model V2.0

> The path to zero trust is an incremental process that may take years to implement.

There are several paths forward from this point. One option would be adding voice services to the existing Zero Trust discussions, another would be another OMB or other such mandate and a third would be to find one or two agencies with a more distinct requirement for secure voice services and have them be the trail blazers. In any case it will be interesting to follow this incremental process along its path.

3.0 METHODS, ASSUMPTIONS, AND PROCEDURES

3.1 Assumptions

In developing the architecture, the researchers were guided by three assumptions. The first was that the solution had to be 100% standards driven. The second assumption was that it could only use features implemented in VoIP systems widely utilized today. The final assumption was that the architecture be made up of sub-elements that could be implemented individually without an agency being required to provide the entire architecture. Each of the sub-elements would have to provide an increased level of protection against cyber threats if only that sub-element were implemented.

3.1.1 Industry Standards

The first key assumption was that the recommended architecture should make use of only industry standards. Proprietary protocols should not be used. The industry standards used in the architecture include:

- IPv4 (Internet Protocol), IPsec (Internet Protocol Security), TCP (Transmission Control Protocol), UDP (User Datagram Protocol), TLS (Transport Layer Security)
- SIP (Session Initiation Protocol), SIP-TLS (SIP over TLS)
- RTP (Real Time Protocol), SRTP (Secure RTP)
- DNS (Domain Name System), DNSSEC (DNS Security), DNS/TLS (DNS over Transport Layer Security), DNS/HTTPS (DNS over Hypertext Transfer Protocol Secure), ENUM (E.164 NUMber mapping in DNS),
- X.509 Certificates, DANE (DNS-based Authentication of Named Entities), SIP Domain Signatures, STIR (Secure Telephony Identity Revisited), SHAKEN (Signature-based Handling of Asserted information using toKENs)
- 5G-NSA (5^{th} Generation wireless network – Non-StandAlone), IMS (IP Multimedia Subsystem)

The reports submitted for each task included a list of the standards documents use, including IETF RFCs and 3GPP specifications.

3.1.2 Voice over IP Phone System Features

The second key assumption for this project was that the feature set used to implement the architecture must be available in VoIP systems that are widely used today. The most common VoIP phone systems in use today in government agencies are the Avaya Aura system and the Cisco Unified Communications Manager (CUCM) system. The most common service provider VoIP system in use today is the Cisco BroadWorks (formerly BroadSoft) system. All three of the platforms were used in the lab testbed network.

3.1.3 Individual Sub-elements

The final key assumption was that the architecture be implemented using sub-elements that could be implemented individually, depending on the capabilities present in each agency. These sub-elements included:

- Secure signaling and media transport within agencies
- ENUM call routing between agencies
- Secure ENUM infrastructure
- Certificate authority for X.509 certificates
- Secure caller identities
- Secure wireless calls
- Secure trunks to voice service providers

Each of these sub-elements were explored as individual tasks, and a detailed report exists for each task. These detailed reports are listed in Appendix A.

3.2 Methods

3.2.1 Lab Testbed Network

A lab testbed was constructed to emulate a multi-agency environment. The high-level architecture is shown in the following diagram and described below.

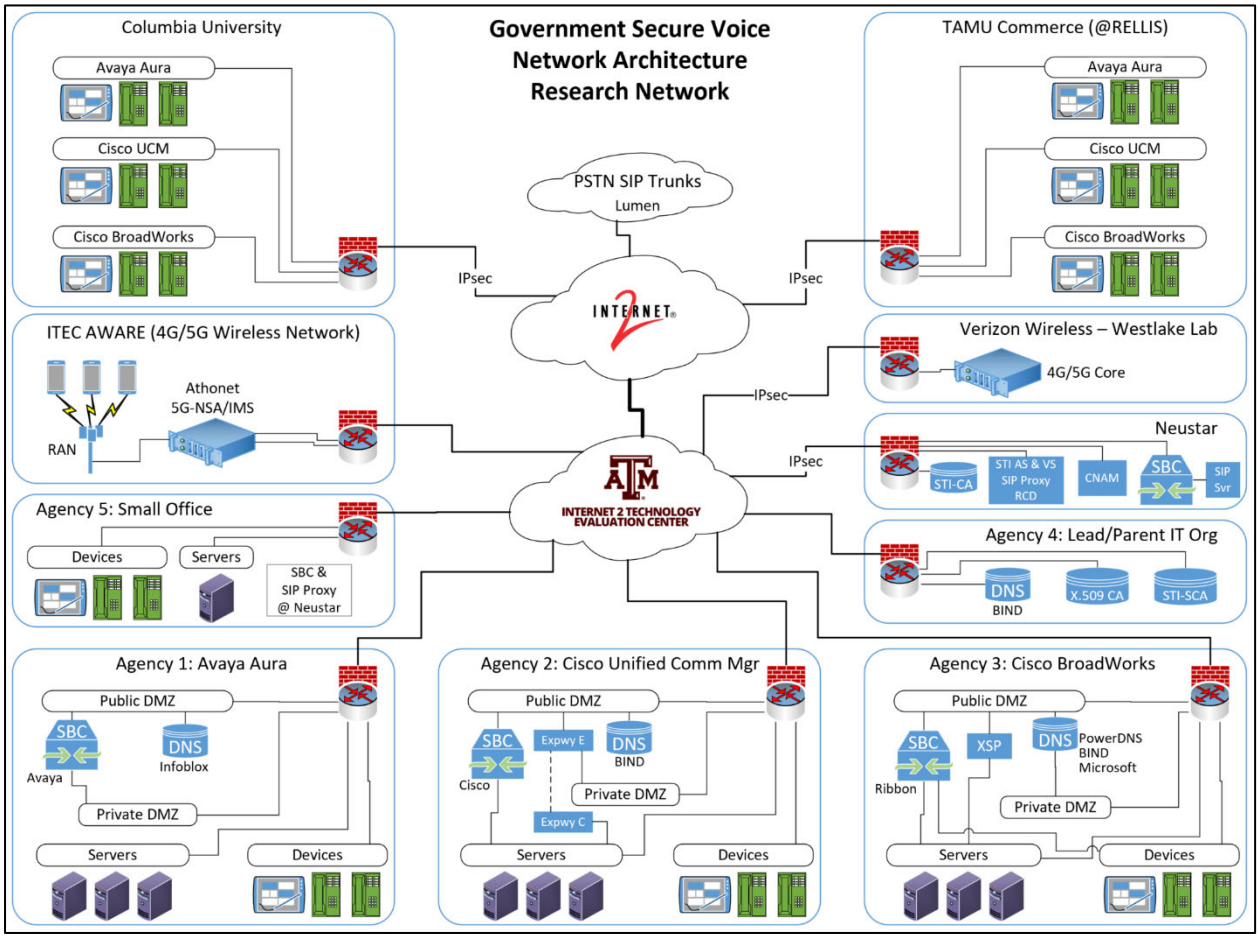

Figure 3: Lab Testbed

Key components of the testbed were:

- Agency 1: represented an agency using the Avaya Aura VoIP system. It included the Avaya Aura phone system servers, an Avaya SBC, Infoblox DNS servers, local Avaya phones, and remote Avaya phones at Columbia University and TAMU Commerce (at the RELLIS campus).
- Agency 2: represented an agency using the Cisco CUCM VoIP system. It included the Cisco CUCM phone system servers, a Cisco CUBE SBC, a Cisco Expressway system for remote user access, BIND DNS, local Cisco phones, and remote Cisco phones at Columbia University and TAMU Commerce (at the RELLIS campus).
- Agency 3: represented an agency using the Cisco BroadWorks VoIP system. It included the BroadWorks phone system servers, a Ribbon SBC, the BroadWorks XSP system for remote user access, PowerDNS, BIND, and Microsoft DNS, local Polycom and Mitel phones, and remote Polycom and Mitel phones at Columbia University and TAMU Commerce (at the RELLIS campus).
- Agency 4: represented an optional central/lead/parent IT organization that provides IT services to the other agencies. It included BIND DNS, two X.509 Certificate Authorities (CAs), and a Secure Telephone Identity Certificate Authority (STI-CA). The STI-CA was hosted at Neustar.
- Agency 5: represented a small, remote office with minimal capabilities. It included a Kamailio open-source SIP server and phones from Polycom and Mitel. It used SIP proxies and an SBC at Neustar.
- ITEC Aware Network: represented a 4G wireless carrier using the IP Multimedia Subsystem (IMS) to provide Voice over LTE (VoLTE) service.
- Neustar: is a leader in the secure telephone identity market. They provided a number of services related to STIR/SHAKEN, including: Authentication Service, Verification Service, Call Session Registry, CNAM Service, SIP Proxies, an SBC, and some publicly routable test numbers for testing calls to and from the PSTN.
- Verizon Wireless lab: provided assistance with real-world public carrier deployment best practices.
- ITEC lab network: the ITEC core lab network provided IP connectivity between all agencies, allowing for packet captures at any point.
- Firewalls: were implemented at each agency. The firewall policies were representative of what would typically exist at an agency. Policies included intra-agency restrictions as well as inter-agency restrictions.
- IPsec Tunnels: were used for site-to-site VPN connections to locations outside of the ITEC lab.
- IP addressing: Each agency had a combination of public and private IP addresses. The public addresses were used in the public DMZ for services that needed to be reachable

from other agencies, such as the public-facing SBC interfaces, remote access servers, and public-facing DNSs. Private IP addresses were used elsewhere.

3.2.2 Multi-vendor Implementation

To validate that the features used to implement the architecture are widely available, a multi-vendor implementation was used. The testbed was designed and constructed using multiple vendors in many key areas related to VoIP call handling and routing, including:

- VoIP systems: Avaya Aura, Cisco Unified Communications Manager, Cisco BroadWorks (formerly BroadSoft), Kamailio
- VoIP telephones: Avaya, Cisco, Polycom, Mitel
- Session Border Controllers: Avaya, Cisco, Ribbon
- DNS servers: Infoblox, BIND, PowerDNS, Microsoft

3.2.3 Traffic Captures

Traffic captures were performed using Wireshark and its command line equivalent, tshark. Depending on the architectural sub-element being tested, it was often necessary to capture traffic in multiple locations at the same time. For example, to analyze both the signaling and media flow at all segments of a voice call between two agencies can involve up to five or more network taps.

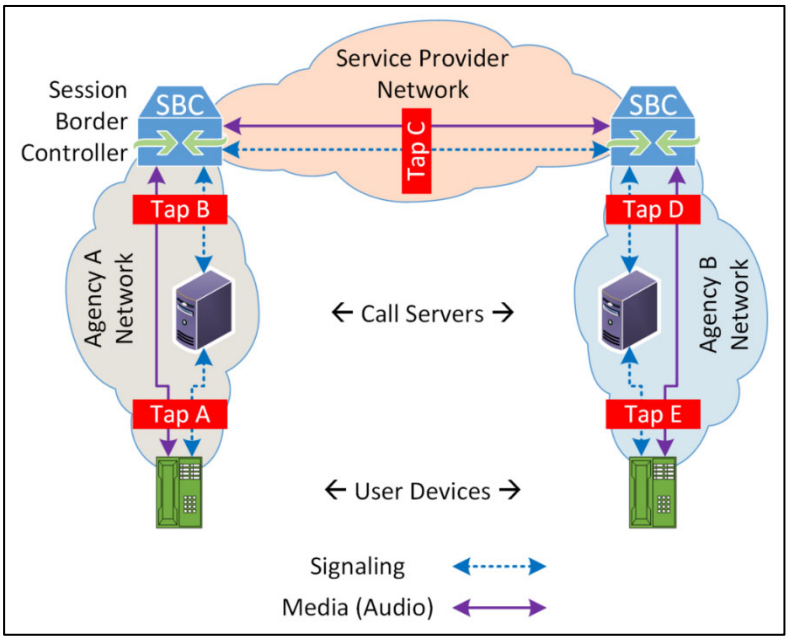

Figure 4: Tapping VoIP Calls

So, a series of virtual machines, each with a set of virtual interfaces, were created to allow multiple users to capture traffic on multiple subnets at one time.

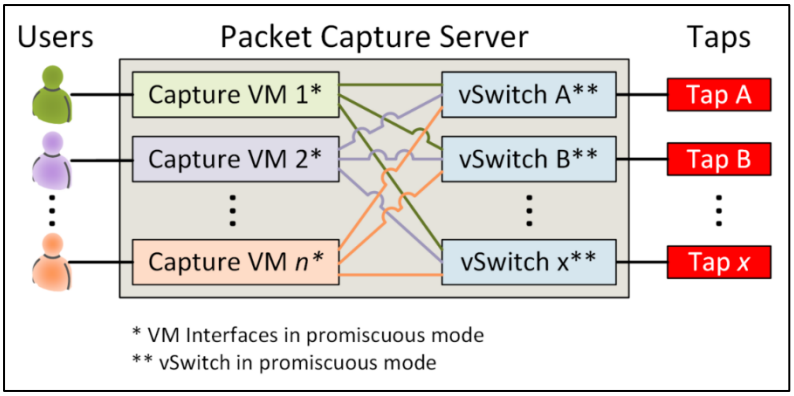

Figure 5: Virtual Packet Capture Configuration

3.2.4 Traffic Visualization

Viewing packet captures in Wireshark or tshark is useful for analyzing specific fields in each packet. But it can be more difficult to visualize the overall communications flow when viewing one packet at a time. So, a tool called VisualEther was used to display the traffic captures in a "ladder" diagram format. This type of diagram shows more clearly the traffic flows between each of the network elements. The output can be customized to show specific fields of interest for the particular type of traffic flow being examines. An example of the start of an ENUM DNS query is shown in the following diagram.

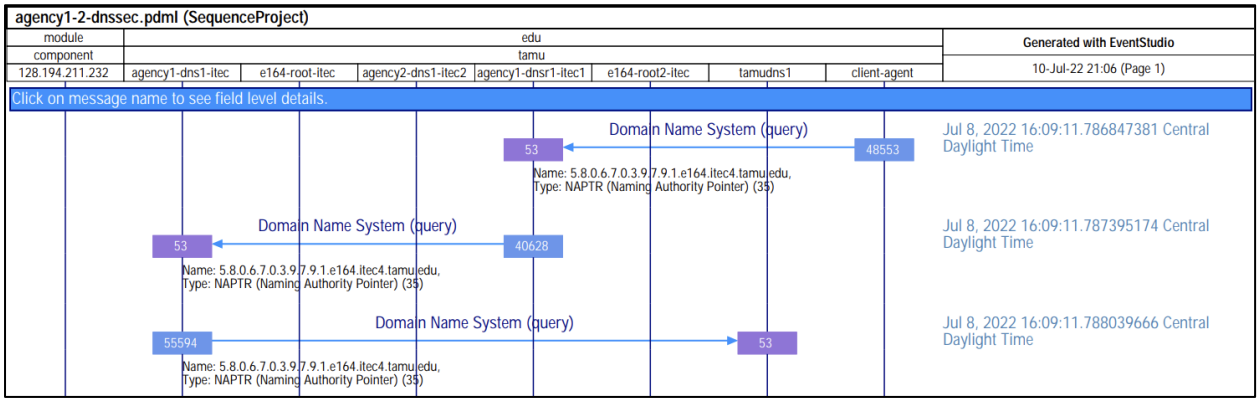

Figure 6: Packet Flow Ladder Diagram using VisualEther

4.0 RESULTS AND DISCUSSION

4.1 Overview

Interim project reports were submitted containing detailed information for each of the project tasks. A list of the reports and the tasks that they cover is provided in Appendix B.

Each of the interim reports contained all details about the task, including:

- Definition and scope of the task
- Technology background
- Network diagrams, IP addresses, configuration options
- Test methodology
- Test results, including traffic captures and packet flow visualizations
- References used

Rather than repeat that information here, only a high-level summary of the results of each task is listed here. For complete details, consult the appropriate report.

4.2 Task 0: Build the lab testbed

To complete this project, the researchers implemented a testbed that emulated three larger agencies, some remote offices, a central IP support organization, a wireless network, and some service providers.

The team selected VoIP systems most commonly in use today. The systems were the Avaya Aura (Agency 1), Cisco Universal Call Manager (Agency 2), and the BroadSoft (now Cisco BroadWorks) platform (Agency 3). Avaya Aura and Cisco UCM are the most widely used platforms in government agencies. BroadWorks is the most widely used platform for cloud service providers.

Some observations made during the completion of this task include:

- Achieving basic IP and SIP connectivity between vendors was straightforward
- Each product vendor and integration engineer had their own preferred configuration options. For example, address formats in the SIP INVITE packet varied:
  - sip:<10-digits>@<IPaddr> Ex: sip:6132581234@11.22.33.44
  - sip:<country-code><10-digits>@<IPaddr> Ex: sip:16132581234@11.22.33.44
  - sip:<E164>@<IPaddr> Ex: sip:+16132581234@11.22.33.44
  - sip:<E164>@<domain> Ex: sip:+16132581234@example.com
  - sip:<E164>@<domain>;user=phone Ex: sip:+16132581234@example.com;user=phone

Recommendations for implementation include:

- Picking a standard value or format for all configuration options will make troubleshooting and documentation MUCH easier
  - Example: SIP Connect 2.0 recommends: sip:<E164>@<domain>;user=phone

4.3 Task 1: Security Calls within the Enterprise with TLS-SIP, SRTP

Unsecured VoIP systems utilize the Session Initiation Protocol (SIP) to control the call (i.e., signaling) and the Real Time Protocol (RTP) to convey the voice audio (i.e., media). The SIP standard also supports the use of Transport Layer Security (TLS) for encrypted signaling and Secure Real Time Protocol (SRTP) for encrypted media. Both options are available in all three of the systems implemented in the testbed. SIP-TLS and SRTP are selected and configured using the management interfaces for the respective VoIP system configuration. TLS relies on X.509 digital certificates, covered in task 4 below.

Some observations made during the completion of this task include:

- All systems had compatible implementations of SIP-TLS/SRTP
- Each vendor had its preferred encryption settings, but it is easy to find common ground
- Older phones may need CA cert downloaded to their trust store

Recommendations for implementation include:

- Choosing a standard set of cypher suites can help as technology progresses
- Be sure to encrypt phone configuration downloads
    - Phones typically download configs when they boot up
    - Use the DHCP option defined by the vendor for encrypted config transfer
- Initially allow downgrade to unencrypted connection to prevent service disruption
- After testing, disable unencrypted flows to prevent downgrading connection

4.4 Task 2: Routing calls between agencies with ENUM (E.164 NUMber mapping)

The Internet Engineering Task Force (IETF) described an alternate call routing process in a series of standards called E.164 Number to URI Mapping (ENUM). This process maps telephone numbers into DNS NAPTR records, with the DNS returning the URI of the SIP call manager that owns the destination number. This allows calls to be routed in a scalable way. Each provider maintains its own DNS records that other providers can query for routing information.

Calls today between agencies are primarily sent to the Public Switched Telephone Network or PSTN. This made routing simple for the agency: any calls destined for outside of the agency are sent to the service provider. The service provider either uses the well-established legacy Signaling System 7 or SS7 to route the calls over legacy networks, or a service-provider implementation of ENUM.

To route calls directly to other agencies, each agency will need a scalable way to manage the routing information. This section describes the implementation and testing of ENUM for direct agency-to-agency call routing. It describes some of the challenges involved and some ways to overcome those challenges.

Some observations made during the completion of this task include:

- Storing and sharing a private, multi-agency telephone number directory using DNS can be done in a scalable and distributed way.
- SBC (Session Border Controller) support and configuration complexity for ENUM-based routing varies significantly between products.
- Vendors of enterprise-class voice systems do not prioritize full ENUM functionality.
    - Limited ENUM lookup format
    - ENUM lookup failed

Recommendations for implementation include:

- Use a hybrid approach for agency-to-agency call routing

4.5 Task 3: Securing ENUM

Most computer-related activity is dependent on the Domain Name System (DNS) to convert hostnames, such as host.example.com, to IP addresses. If ENUM is implemented, call routing would also be dependent on DNS. That makes the integrity of the DNS responses that much more critical. To ensure that the DNS responses can be trusted, several security mechanisms were deployed, including Domain Name System Security Extensions (DNSSEC), Transaction Signatures (TSIG), secure replication, DNS over TLS, and DNS over HTTPS.

DNSSEC uses a set of keys to establish a chain of trust between DNS servers. TSIG and secure replication are used between primary and secondary servers within an organization. And DNS over TLS and DNS over HTTPS are used to encrypt the query/response communications between DNS clients and servers.

Some observations made during the completion of this task include:

- DNSSEC requires significant careful planning and execution
    - Automation of DNSSEC is improving
    - Multivendor DNSSEC works but there is significant management overhead
- Enterprise-class VoIP systems may not have configuration options for using encrypted DNS (DoT, DoH) queries, DANE, or SIP Domain Signatures
- Public CAs may not support the extendedKeyUsage attributes required to create SIP Domain Signatures

Recommendations for implementation include:

- The previously suggested hybrid approach can mitigate most issues found

4.6 Task 4: Certificate Authorities (CAs)

TLS relies on X.509 certificates to authenticate the other end of a connection in the same way that a web browser relies on certificates for accessing a secure web site with HTTPS. Several options are available for generating X.509 certificates, including commercial services that charge by the certificate, federated services that charge a flat fee for a group of organizations, and private certificate authorities created by agencies. These models were reviewed and two examples of setting up a private certificate authority were explored.

Some observations made during the completion of this task include:

- There are three main types of arrangements between a certificate user and the Certificate Authority:
    - Public CA: the certificate user obtains purchases certificates from a CA in the same way that other users do
    - Consortium CA: a group of certificate users make an agreement with a CA for a specific set of services at a custom price
    - Private CA: the certificate user creates their own CA

Recommendations for implementation include:

- There is no one best answer for all situations. Some key comparison areas are summarized in the table below. These will help with an implementation decision.

Table 1: Comparison of Certificate Authority Types

Feature/Benefit	Public CA	Consortium CA	Private CA
Initial implementation effort	Low	Medium	High
Sustained effort	Low	Low-Medium	High
Charge per certificate	Yes	No	No
Certificates for private hosts	No	Usually	Yes
Hide existence of hosts or domains from outside entities	No	No	Yes
Must distribute CA certs to VoIP system and network devices	Usually	Usually	Yes
Must distribute CA certs to end user devices	Usually Not	Usually Not	Yes

4.7 Task 5: Secure Caller Identity with STIR/SHAKEN, CNAM, RCD

A large percentage of today's VoIP cyber-attacks involve manipulation of the caller's telephone number, (Caller ID) or caller name (CNAM), allowing attackers to pretend that they are someone that they are not. These attacks vary from nuisance robocalls to life endangering "swatting" calls. A combination of services which are Secure Telephone Identity Revisited (STIR) and Signature-based Handling of Asserted information using toKENs (SHAKEN) are mandated by the FCC for implementation by US telephone service providers. These help to reduce the risk for

calls coming from US service providers. Unfortunately, international calls are beyond the control of the FCC. And direct enterprise-to-enterprise calls are outside the scope of the current FCC mandate. This section of the project describes how agencies could utilize STIR/SHAKEN, CNAM and Rich Call Data (RCD) to further secure their VoIP traffic from cyber-attacks. It also describes how one agency, the U.S. Marine Corps Recruiting Section used Rich Call Data and STIR/SHAKEN to increase the effectiveness of their call center software.

Some observations made during the completion of this task include:

- STIR/SHAKEN and CNAM have been predominately carrier technologies; enterprise personnel may be unfamiliar with the concepts and technology
- STIR/SHAKEN ecosystem is currently designed exclusively for service providers
- These processes will need to be expanded to accommodate government agencies
- Service providers currently strip enterprise-added Identity headers
- This should change as the technology is more widely deployed
- Enterprise-class VoIP systems may not include STIR/SHAKEN, CNAM or RCD capabilities
- SIP Proxies can be used to perform the AS, VS and CNAM lookups
- Examples: carrier-class SBCs, service provider offerings
- Rich Call Data can significantly improve call answer rates
- Many desktop phones may not be capable of displaying Rich Call Data
- The integration effort is significant; an interoperability test bed is important

Recommendations for implementation include:

- Use hosted STI-CA services to manage the added complexity of the SHAKEN certificate process until the organization is ready to build its own
- Use certificate delegation to enable multiple agencies to use a single Service Provider Code
- Use carrier-class SBCs to perform the Authentication Service (AS), Verification Service (VS) and CNAM lookups
- Use hosted identity services (AS, VS, RCD) until the organization is ready to build its own
- Use Rich Call Data for outbound call centers (and inbound with call back)
- Establish a government interoperability testbed for STIR/SHAKEN
- Example: Neustar operates the “ATIS Robocalling Testbed” (90+ carriers, multple countries)

4.8 Task 6: Secure IMS-based (wireless) Calls

The IP Multimedia Subsystem (IMS) is the SIP functionality behind how VoIP calls are made over a wireless network. IMS is a Global 3GPP standard that is required for wireless service providers to assure call persistence as a mobile user moves between networks within the service provider, and to support billing information and authentication as these same mobile users roam between service provider’s networks. This section, which is still under development, will describe how calls

from service provider and private LTE networks will be able to provide secure voice communications to the rest of the architecture.

Some observations made during the completion of this task include:

- There is little that can be done to protect IMS call signaling or media (audio) from threat actors.
    - IMS call signaling passes through numerous IMS core network elements. Anyone with administrative access to those core elements would be able to spy on the call signaling.
    - The DTLS-SRTP and SDES options for IMS media security depends on secure signaling for its key exchange. A threat actor could disable secure signaling to prevent secure media.
    - The KMS option for IMS media security does not depend on IMS network signaling. But it does depend on deploying a key management infrastructure, which a threat actor may not provide or may compromise. This would leave the media unsecured or could cause the audio connection to fail, depending on how the IMS client is configured.
    - End-to-end media security depends on being able to terminate the call at an SRTP-capable UE (User Equipment, i.e., a mobile phone) or AS (Application Server). For example, if the IMS core routes to the remote end through a PSTN gateway, such as if a leg of the route must travel over the circuit switched PSTN, then the IMS will terminate the media path at a local PSTN gateway, which may result in call setup failure or media failure. A threat actor could force end-to-end media security requests to fail in this way, so that clients are forced to send unencrypted media, thereby allowing the threat actor to snoop on their calls.
- Existing mainstream service providers are capable of deploying IMS signaling and media security in the IMS core networks. However, they tend to opt not to do so, due to the difficulty it creates for troubleshooting connectivity issues. To compensate, they rely on strong physical and administrative access security controls.
- Affordable and manageable IMS network cores that support both IMS signaling and media security, and that include technical support, training, and professional services are commercially available for private implementations.
- Further investigation is needed to understand the IMS client capabilities for private IMS deployments.

Recommendations for implementation include:

- Investigate IMS client security capabilities for the specific version of the client installed on the phone. This can vary, depending on the public carrier involved.
- Works with the private or public service provider to implement security mechanisms in the wireless and IMS networks.

4.9 Task 7: Secure Trunking Between Enterprise and Service Providers

Calls that cannot be secured end-to-end will always have a default route to a public voice service provider. Today these calls are usually connected without the benefit of encryption. This section describes how to establish a secure voice trunk service from the secure architecture described in this document and the service provider use of TLS and SRTP. Although this call may be transcoded at the far end of the call from TLS to SIP and from SRTP to RTP before it is connected to the unsecured party, it will be secured further than it would be otherwise. This feature was tested and documented with the service provider that provides the commercial SIP trunks to the testbed.

Some observations made during the completion of this task include:

- Most connections to service providers today are not encrypted
- Service providers are just now productizing TLS connections
- SIP Connect 2.0 seeks to standardize the various parameter choices for enterprise-to-service provider SIP connections
- Ultimately, it is up to whatever service provider has defined in their product offering
- Connections within the service provider's network are usually not encrypted

Recommendations for implementation include:

- Check with service provider first, before planning this change
- Review SIP Connect 2.0 with the service provider; standardize where possible

4.10 Task 8: Feasibility of Implementation

The overall goal of this portion of the project was to evaluate the feasibility of applying the findings of this project in government agencies. This task developed several tools and content to communicate with stakeholders and enable them to reflect and evaluate their current Voice over IP (VoIP) implementations against the outputs and recommendations of this project. The main objective was to assess the feasibility (from different perspectives) of project outputs implementation at government agencies.

Some observations made during the completion of this task include:

- Secure voice services for government agencies is a critical step to ensure that sensitive VoIP communications are properly protected against common vulnerabilities and threat vectors.
- The feasibility of implementing secure voice services depends on multiple factors including the agencies' technical resources and expertise, existing policies, and current security requirements in contracts with voice service providers, in addition to other competing priorities within the agencies' limited resources.
- This project tasks provided tools, content, detailed instructions, and recommendations that would help interested agencies better understand and prioritize implementation steps regarding their own secure voice services.

Recommendations for implementation include:

- Establish a pilot project to assist the first few agencies in implementing the parts of the architecture that are appropriate for them. Include documentation of the process to assist other agencies in the future.

5.0 CONCLUSIONS

The project team has successfully completed all of the tasks assigned. A voice testbed was deployed and used to do the testing and detailed documentation describes how one would implement parts or all the secure voice architecture. This document is close to 800 pages and can be a roadmap towards securing the voice networks in the future.

This architecture document is broken down into sections, each of which deals with a part of architecture. These sections discuss how to enable secure protocols withing existing systems, how to route calls to keep them on net, how to secure resources that VoIP depends on such as DNS, how to authenticate caller ID information, and how to secure voice services in a private LTE network. Many of these capabilities can be implemented individually or an agency could choose to implement the entire architecture.

When this project was awarded in the fall of 2020, there was already sensitivity to the transmission of Controlled Unclassified Information (CUI) over the public switched telephone network with no security layered into that transmission. Fast forward to 2023 and the Department of Homeland Security as well as the Department of Defense have both made the implementation of a Zero Trust Architecture (ZTA) a priority for all Defense and Critical Infrastructure communications. This project has laid out a roadmap that would allow voice services to be added to the federal ZTA initiatives.

Possible next steps could include:

1. Select field trials to serve as a proof of concept within a few federal agencies.
2. An analysis of the elements of the secure voice architecture to determine which functions are best served in a distributed manner at the agency or department level and which of them should be centralized.
3. An analysis of the costs/benefits of adding the secure voice architecture to the federal ZTA mandates.

This project would not have been possible without the donations of major systems by equipment manufacturers and the consulting services of firms that were obviously experienced in dealing with complex configurations.

To the extent that other funding will allow, it is our intention to keep the testbed as current as we can for any potential follow-up work and other research opportunities.

6.0 REFERENCES

Interim project output reports were submitted containing detailed information for each of the project tasks. A list of the reports and the tasks that they cover is below. For much more detail about each of the product tasks, including technical background and implementation steps, see the appropriate report.

Output 1 – Detailed Lab Architecture

- Task 0: Partition pieces of the existing testbed and extend it into Columbia University and the TAMU RELLIS labs

Output 2 – SIP-TLS and SRTP Configuration and Testing

- Task 1: Secure call setup using TLS and SRTP across enterprise networks

Output 3 – ENUM

- Task 2: Routing calls using ENUM
- Task 3: Security ENUM

Output 4 – Certificate Authority

- Task 4: Standardizing the Certificate Authority

Output 5 – STIR/SHAKEN

- Task 5: Implementing STIR/SHAKEN to avoid caller ID spoofing

Output 6 – IMS

- Task 6: Securing mobile IMS-based calls

Output 7 – TLS Trunks

- Task 7: Secure SIP Trunking

Output 8 – Feasibility

- Task 8: Feasibility Assessment

APPENDIX A – Publications and Presentations

List the dates, times, title, event, and speakers of any presentations made under this effort and the title author and publication information for any publication made under this effort.

LIST OF SYMBOLS, ABBREVIATIONS, AND ACRONYMS

3GPP	3rd Generation Partnership Program
5G	5th Generation wireless network
AFRL	Air Force Research Laboratory
AS	Authentication Service
ATIS	Alliance for Telecommunications Industry Solutions
BIND	Berkeley Internet Name Domain
CA	Certificate Authority
CIO	Chief Information Officer
CISA	Cybersecurity and Infrastructure Security Agency
CNAM	Caller Name
CUCM	Cisco Unified Communications Manager
CUI	Confidential Unclassified Information
DANE	DNS-based Authentication of Named Entities
DHCP	Dynamic Host Configuration Protocol
DHS	Department of Homeland Security
DMZ	De-Militarized Zone
DNS	Domain Name System
DNSSEC	Domain Name System Security
DoD	Department of Defense
DoH	DNS over HTTPS
DoT	DNS over TLS
DTLS	Datagram TLS
ENUM	E.164 NUMber mapping in DNS
FCC	Federal Communications Commission
HTTPS	Hypertext Transfer Protocol Secure
IETF	Internet Engineering Task Force
IL	Impact Level
IMS	IP Multimedia Subsystem
IP	Internet Protocol
IPv4	Internet Protocol version 4
IPsec	Internet Protocol Security

ITEC	Internet2 Technology Evaluation Center
ITU	International Telecommunication Union
KMS	Key Management System
LTE	Long Term Evolution (4th generation wireless)
NPPD	National Protection and Programs Directorate
NSA	Non-StandAlone
OMB	Office of Management and Budget
PSTN	Public Switched Telephone Network
RCD	Rich Call Data
RFC	Request for Comments
RTP	Real Time Protocol
SBC	Session Border Controller
SCCP	Skinny Client Control Protocol
SDES	SDP Security Descriptions for Media Streams
SDP	Session Description Protocol
SHAKEN	Secure handling of Asserted information using toKENs
SIP	Session Initiation Protocol
SRTP	Secure Real Time Protocol
SS7	Signaling System 7
STI-CA	Secure Telephony Identity – Certificate Authority
STIR	Secure Telephony Identity Revisited
TCP	Transmission Control Protocol
TLS	Transport Layer Security
TAMU	Texas A&M University
UDP	User Datagram Protocol
UE	User Equipment
URI	Uniform Resource Identifier
VoIP	Voice over Internet Protocol
VoLTE	Voice over LTE
VPN	Virtual Private Network
VS	Verification Service
ZTA	Zero Trust Architecture