Cooperative Firewall Signaling over SCION

Dheeraj Chandrashekar

School of Electrical Engineering

Thesis submitted for examination for the degree of Master of Science in Technology.
Espoo 17.12.2021

Supervisor

Prof. Raimo Kantola

Advisor

Maria Riaz

Aalto University
School of Electrical
Engineering

Aalto University
School of Electrical
Engineering

Aalto University, P.O. BOX 11000, 00076 AALTO
www.aalto.fi
Abstract of the master's thesis

Author Dheeraj Chandrashekar

Title Cooperative Firewall Signaling over SCION

Degree programme Masters in Computer, Communication and Information Sciences

Major Communications Engineering **Code of major** ELEC3029

Supervisor Prof. Raimo Kantola

Advisor Maria Riaz

Date 17.12.2021 **Number of pages** 80 **Language** English

Abstract

5th Generation (5G) and Internet of Things (IoT) have contributed to the rise in connected devices, which in turn has exhausted the available set of IP addresses. NAT is a popular solution that solves the IP exhaustion problem, but suffers from a reachability issue. Customer Edge Switching (CES) is a firewall solution intended to replace the traditional NAT by enforcing cooperative behavior. While CES solves the reachability issue, it is still troubled by some of the typical attacks present on the current Internet. Scalability, Control, and Isolation on Next-Generation Networks (SCION) is a new Internet architecture designed to provide effective point-to-point packet delivery. Realizing the SCION network would require changes to infrastructure and the protocol stack. However, SCION provides an application for the end-hosts in an IP network to connect to SCION using SCION-IP-Gateway (SIG).

SCION does not provide any defensive mechanism for application-layer DoS attacks, while CES does. Having an end system focused on trust solution over SCION would provide defense against trivial attacks and application-layer DoS attacks. End-domains can benefit from the integration of CES and SCION, where CES provides host-level authenticity by cooperative behavior concept, and SCION can provide network-level security by design.

In this thesis, the control plane/signaling traffic between the two CES nodes is switched from routed IP to SCION whenever available using SIG. The implementation is carried out in three phases: Proactive, Reactive, and Monitor phases, and verified with a range of tests such as design verification, delay calculation of CETP optimization, and SIG performance. The evidence suggests that the solution has no change from an end-user perspective. SCION's SIG is stable and provides good performance. The solution is the first prototype of an end-to-end, client-to-server trustworthy communication and service solution over the wide-area network.

Keywords NAT, Customer Edge Switching, SCION, CES signaling over SCION

Acknowledgements

This master thesis has been carried out under the Network Security research group of the Department of Communications and Networking at Aalto University.

I would like to thank my supervisor, Professor Raimo Kantola, for realizing the potential in me to carry out such a challenging job. His constant support along with his vast knowledge and experience guided me through some of the tough spots. I would also like to thank my advisor, Maria Riaz, for her help from the beginning of the thesis until the end. CES, being a new concept for me, your support in understanding it is very much appreciated. Thank you, for making sure that I was not blocked with any item and ensuring I could continue smoothly.

I would like to thank my entire family for believing in me and supporting me throughout. Amma and Divya, you are the pillars of strength for me. Sumukh, thanks for keeping me grounded and sane by making me get back into some of my lost habits.

I would also like to thank my friends who were with me throughout this journey. Your good wishes made me accomplish many feats. Special thanks to Greeshma and Anjana for checking on my mental health regularly. I owe you guys big time.

Finally, I want to thank the almighty mother, who is the reason I have achieved so many things in life and will continue to achieve in the future. Lastly, I dedicate this work to my buddy and best friend, *Spotty*. No one knows me better than you do. Thank you for being in my life. I miss you and love you so much, always.

17 December 2021, Espoo, Finland

Dheeraj Chandrashekar

Table of Contents

List of Figures

List of Tables

List of Acronyms

5G	5th Generation
ALG	Application Layer Gateway
API	Application Programming Interface
AS	Autonomous System
ASE	AS Entry Fields
BGP	Border Gateway Protocol
C2C	CES-to-CES
CCN	Content-centric Networking
CDN	Content Distribution Network
CES	Customer Edge Switching
CETP	Customer Edge Traversal Protocol
CN	Customer Network
CP	Control Plane
DDNS	Dynamic Domain Name System
DDoS	Distributed Denial of Service
DHCP	Dynamic Host Control protocol
DNAT	Destination NAT
DNS	Domain Name System
DNSSEC	DNS Secure
DOS	Denial of Service
DP	Data Plane
FQDN	Fully Qualified Domain Name
FTP	File Transfer Protocol
H2H	Host-to-Host
HF	Hop Fields
HTTP	Hyper Text Transfer Protocol
ICE	Interactive Connectivity Establishment
ICMP	Internet Control Message Protocol
ICN	Information-centric Networking
ID	Identity
IDS	Intrusion Detection System
IETF	Internet Engineering Task Force
INF	Info Fields
IP	Internet Protocol
IPv4	Internet Protocol version 4
IPv6	Internet Protocol version 6
ISD	Isolation Domain

ISO	International Organization for Standardization
ISP	Internet Service Provider
ITU	International Telecommunication Union
IoT	Internet of Things
LAN	Local Area Network
MAC	Medium Access Control
MTU	Maximum Transmission Unit
NAPT	Network Address/Port Translator
NAPTR	Name Authority Pointer
NAT	Network Address Translation
NCC	Network Coordination Centre
NDN	Named Data Networking
PCB	Path-segment Construction Beacons
PCFS	Packet-carried Forwarding State
PKI	Public Key Infrastructure
PRGW	Private Realm Gateway
RAINS	Another Internet Naming Service
RGW	Realm Gateway
RIPE	Regional Internet Registry for Europe
RLOC	Routing Locator
SCION	Scalability Control and Isolation on Next-Generation Networks
SCMP	SCION Control Message Protocol
SCTP	Stream Control Protocol
SIG	SCION-IP-Gateway
SNAT	Source NAT
SPN	Service Provider Network
STUN	Session Traversal Utilities for NAT
TCP	Transmission Control Protocol
TLV	Type
TRC	Trust Root Configuration
TURN	Traversal Using Relays around NAT
UDP	User Datagram Protocol
UPnP	Universal Plug and Play
VM	Virtual Machine
VPN	Virtual Private Network
WAN	Wide Area Network

1 Introduction

The current Internet we see and experience today has evolved at an unprecedented scale over the past few decades. It has exceeded all initial expectations in terms of capacity, usage, applications, and services. According to the International Telecommunication Union (ITU), there is no doubt that these expectations will continue to grow exponentially in the near future[29]. The emergence of new technologies under the 5th Generation (5G) and Internet of Things (IoT) has contributed to the rise in connected devices. These connected devices would require an IP Address for communication, which would create scarcity for the available IP addresses. For example, on 25th Nov 2019, RIPE NCC, the regional Internet registry for Europe, announced that they had allocated the last block of IPv4 addresses and have officially exhausted the available set of IPv4 addresses[26]. This exhaustion of IP addresses was an anticipated event, and multiple solutions proposed are implemented to address the issue. Two prominent solutions are Internet Protocol version 6 (IPv6)[3] and Network Address Translation (NAT)[4]. While IPv6 is proposed as the long-term and recommended solution despite implementation challenges which are out of scope in this thesis, NAT provides a logical solution with minimal implementation efforts, hence is more popular than IPv6.

NAT enables hosts in a private network to connect to public realms using a shared public IP address. Due to its minimum implementation efforts, NAT was successful in overcoming the IPv4 address exhaustion to some extent. Nonetheless, it suffered from a reachability problem when a host located in the public network tried to reach another in a private network when there is no explicit mapping in the NAT device to route these packets[1]. Many solutions such as User Datagram Protocol (UDP) Hole Punching, STUN[5], TURN[6], and ICE[7] were proposed but failed to be promising.[15]

Customer Edge Switching (CES) [8]is a solution proposed and developed at Aalto University to overcome this reachability problem of NAT. This solution is realized at the edge of the customer network and intends to replace the traditional NAT. It acts as a cooperative firewall, allowing both sender and receiver networks to cooperate and counter the actions of the compromised hosts. It also promotes cooperative security among administrators and provides defense against spoofing and Distributed Denial of Service (DDoS). CES implements a Realm Gateway (RGW) which acts as a Source NAT (SNAT) for outbound connections from the private network hosts

and a Destination NAT (DNAT) for traversing inbound connections from the clients in the public realm towards the private hosts [15]. This solution does not require host or network infrastructure changes; instead, it reuses the existing protocols and operations. All in all, CES provides an innovative solution to the reachability problem of NAT. However, it is still prone to the menaces plaguing the current internet. BGP route hijacking, DDoS attacks, network congestion are some of the issues that can disrupt the normal functioning of the network[9].

Exploring further into the root cause of these problems, it is understandable that the active development of any application or service deals with only the protocol stack and its modification. Among the protocols available, the core ones have remained the same; for example, Internet Protocol (IP) and the Border Gateway Protocol (BGP) are the core technologies that are still running the backbone of Internet Communication and have been in use for more than 30 years. IP and BGP have many shortcomings such as source address spoofing, DoS and Distributed DoS, route hijacking, man in the middle attacks after route hijacking, lack of trustworthy names[10], [19].

No single solution caters to all of the issues mentioned. There have also been proposals of having a completely new Internet architecture that is better and more capable than the current one. Information-centric networking (ICN) or Content-centric networking (CCN) architecture, Named Data Networking (NDN) architecture[11], and Mobility first architecture[12] are few examples that focus on a specific set of requirements. SCION [27](Scalability, Control, and Isolation on Next-Generation Networks) [2] is one such Internet architecture that is interesting as it focuses on security and high availability for point-to-point communication, which are critical for the CES solution mentioned above [27].

1.1 Background and Motivation

Scalability, Control and Isolation on Next-Generation Networks (SCION) is a clean-state Internet architecture for end-to-end communication, designed to offer efficient point-to-point packet delivery and high availability even in the presence of adversaries and failures[27]. By design, SCION provides defense against trivial issues such as spoofing, DDoS defenses. It also provides enhanced security options, multipath communication, and path-aware networking[27]. Many ISPs and banks are already using the production version; even the Swiss Government utilizes SCION for some of its

services, promoting SCION's trust[30]. SCIONLab Network is a global research network designed to test and experiment with the SCION architecture. The SCIONLab [28]infrastructure comprises a network of globally connected Autonomous Systems (ASes). Realizing CES over SCION ideally would require a complete change in the CES code base as SCION's native language is "Go lang," and CES's is "Python." However, SCION has an interworking solution for these kinds of scenarios, referred to as a way of SCION deployment meant for End Domains - SIG, SCION-IP-Gateway. SIG enables seamless integration of SCION capabilities in end-domain networks. Thus, running CES over SCION using SCION's native SIG would be a quick solution compared to re-writing the CES code base [27].In this scenario, CES would add host and application-level security controls to the end to end communication solution, while SCION would handle the wide-area segment of communication in a trustworthy manner.

CES supports the idea of control plane/data plane separation and uses Customer Edge Traversal Protocol (CETP), designed for signaling about packet delivery through a tunnel from one CES to another. Once CETP signaling detects a policy match and establishes a base level of trust between CES to CES node, several different tunneling protocols can be triggered to deliver the end system data plane packets. In practice, any tunneling protocol supported by Open Flow such as GRE, VXLAN, Geneva, IPSEC can be used. CETP provides a range of tools meant for defensive mechanisms for the private networks and places the responsibility of using those tools on the receiver's network. In other words, CETP allows the inbound edge to specify which ID types must be used to identify a remote host, whether protection against source address spoofing is required, whether the CETP control signaling needs to be signed to ensure non-repudiation, and also whether the integrity of edge-to-edge communication has to be preserved by encryption or not[13]. Since SCION itself eliminates spoofing and guarantees the authenticity of addresses and names[27], running CES to CES signaling over SCION would improve trust. Also, many of the low-level heuristic algorithms used in CES for routed IP networks, such as spoofing elimination on the CES to CES communication and mutual authentication of the CES nodes, would be unnecessary, leading to short session up times with enhanced security. The added security that CES can still provide in SCION network relates to the hosts that are communicating and the applications that are allowed to send and receive traffic.

1.2 Research Problem

The purpose of this thesis is to add new functionality to the CES firewall solution where the signaling traffic is switched from routed IP (normal Internet) to the SCION network whenever available, using SCION's native SCION-IP-Gateway application.

CES would support the switching of signaling traffic from routed IP over to SCION upon receiving the SIG information from the remote end. Once the switch-over begins, SCION's SIG encapsulates the CETP packet with the SCION header and sends the packet over the SCION network. Upon reaching the destination SIG, the packet header would be stripped and the packet would be forwarded to the normal inbound CES for processing. In addition to the switch-over capability, monitoring the status of the SIG, fall back to default behavior (traffic over routed IP) during failover, or any other discrepancies are part of the new functionality.

1.3 Objective and Scope

Both networks: SCION and routed IP must be deployed under the same orchestration to realise the switch-over functionality. Hence, the initial step would include designing the topology and setting up the required network connections. The SCION infrastructure resources are gained from the SCIONLab team, and the hosts and nodes running CES code would be realized via Linux containers. The whole topology would run over Virtual machines (VM), including the SCION's Autonomous systems (AS).

Following set of objectives are defined as part of the functionaility:

- Verify that CES and SIG can co-exist under a single orchestration and operate without interfering with each other.
- Design and verify sending CETP traffic over SCION using SIG and that we are able to deploy such a communication solution smoothly.
- Verify in failure cases, that communication can be restored following a suitable policy.
- To define and use suitable security policies allowing CES to optimize the security of communications with the minimum number of round trips edge to edge before data starts to flow.

- Verify, that in the case of CETP signaling over SCION, CES can send the data plane traffic either over SCION or using a tunnel over an IP network.

In addition, despite topics such as containerization, testbed, deployment being out of scope, a summary of the mentioned topics is included, which was crucial towards achieving the final goal. Which is creating the prototype of an end-to-end, client-to-server trustworthy communication and service solution over the wide area by combining Customer Edge Switching with SCION.

1.4 Structure

The thesis is divided into the following chapters.

Chapter 2 provides general knowledge on the Internet Protocol Suite by discussing Internet protocols and the issues present in the current Internet along with Network Address Translation.

Chapter 3 discusses Network security by introducing the concept of trust, security threats and basics of firewall solutions.

Chapter 4 involves introducing the topic of Customer Edge Switching, its architecture, and its operation, described in detail.

Chapter 5, along with the introduction, explains the fundamental aspects of SCION Internet Architecture in relevant detail.

Chapter 6 details the CES signaling switch-over solution. The proposed solution and design along with setup configurations of both SCION and CES are described. Finally, code implementation and its phases are explained in detail.

Chapter 7 includes all the evaluations performed over the solution. Design verification, packet visualization, CETP optimization, and SIG performance are all verified and the results are presented.

Chapter 8 provides a conclusion on the solution proposed and implemented. It also provides an insight into the evaluations performed on the said solution. Future work related to the presented work is also discussed in brief.

2 Internet Protocol Suite

This chapter provides a basic understanding of the concepts related to computer communication in general. An overview of the protocol stack and its structure is presented, followed by a detailed discussion on the Internet Protocol (IPv4), its characteristics, and implementation is elaborately explained. NAT is introduced as a solution to the address depletion problem, which is further discussed with scenarios. Working and implementation of the NAT and its drawbacks are mentioned. Briefly, the solutions proposed for the NAT issues are discussed as well. A quick overview of the Domain Name System (DNS) is included as part of this chapter.

2.1 Overview

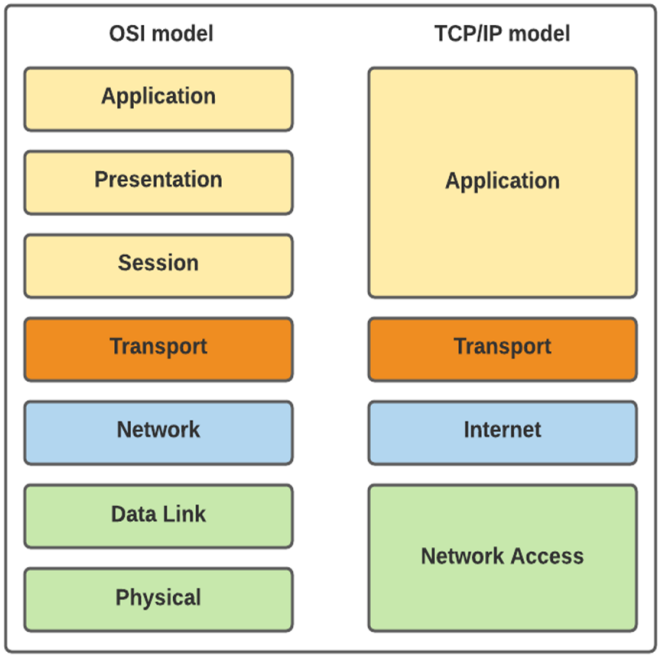

Figure 1: Protocol stack in OSI and TCP/IP models

Protocols are a set of rules required for communication between computers. Group of protocols following a hierarchy of layers form a protocol stack. The protocols within a stack determine the rules for a network model such as the OSI or TCP/IP models. The OSI model developed by the International Organization for Standardization (ISO) has seven layers required for communication. Similarly, the TCP/IP

model developed by Internet Engineering Task Force (IETF) has four layers defined within the Internet Protocol Suite. Both models follow a standard rule that any layer in the stack must serve the layer above it and be served by the layer below it.[15]

Figure 1 presents the layers for both models; the OSI model has an Application layer at the top and a Data link layer at the bottom. Each layer has specific functions to perform: the physical layer is responsible for the movements of bits from one node to the next, the data link layer is responsible for moving frames from one node to the next, the network layer is responsible for the delivery of individual packets from source host to the destination host, the transport layer is responsible for the delivery of a message from one process to another, the session layer is responsible for dialog control and synchronization, the presentation layer is responsible for translation, compression and encryption, finally application layer is responsible for providing services to the user.[14]

TCP/IP model, when compared to OSI, the network access layer is equivalent to the combination of the physical and data link layers. The internet layer is equal to the network layer, and the application layer is roughly identical to the session, presentation, and application layer.[14]

2.2 Internet Protocol

As mentioned in the previous section, communication at the network layer is host-to-host; a device somewhere in the world wants to communicate with another device elsewhere in the world. This communication happens over the Internet. The packet transmitted by the sending device passes through several nodes like the Local area network (LAN) or Wide area network (WAN) before reaching the destination device. A global addressing scheme required for this level of communication is achieved through IP addresses.

2.2.1 IPv4

IP address or IPv4 (Internet Protocol version 4) address is a 32-bit address that uniquely and globally identifies an interface on a device on the Internet. Each IP address can define one, and only one, connection to the outlet or point of attachment. Protocol IPv4 defines addresses; hence it has an address space which is the total number of addresses used by the protocol. If a protocol uses N bits to represent an

address, the address space is 2^N because a bit can have two values (0 or 1), and N bits can have 2^N values. IPv4 uses 32-bit addresses, which means the complete list of addresses is equal to 2^{32} or 4,294,967,296. Theoretically, without restrictions, there could be more than 4 billion devices connected to the Internet.

IPv4, at its inception, followed a classful addressing scheme to distribute and allocate the existing IP addresses using the concept of classes. The address space is split into five classes (A, B, C, D, E). Each class is categorized into a fixed number of blocks with a fixed size. This type of scheme wasted a large part of the available addresses. Also, the fast growth of the Internet led to the depletion of the available IP addresses.[14]

To overcome these issues classless addressing scheme was designed and implemented. In this scheme, addresses are still given in blocks, but the concept of classes is ignored. When an organization (small or large) needs to be connected to the Internet, it is granted a range of IP addresses. The size of the range varies based on the nature and the size of the organization. The responsibility of IP address allocation is given to global authorities specific to the countries with mutual consent. However, these authorities do not normally allocate addresses to individual entities. They assign rather a large block of addresses to an Internet Service Provider (ISP). The ISP divides the obtained block into smaller subblocks and issues these subblocks to its customers.[14]

2.3 Network Address translation (NAT)

2.3.1 NAT traversal protocols

Range	Total
10.0.0.0 to 10.255.255.255	2^{24}
172.16.0.0 to 172.31.255.255	2^{20}
192.168.0.0 to 192.168.255.255	2^{16}

Table 1: Addresses for private networks

The exponential growth of the Internet posed itself as a challenge and led to the problem of address exhaustion. Previously an address was assigned to a user when needed, but the situation is different today. With the explosion of connected devices, the shortage of IP addresses required for communication becomes a severe problem.

A quick and effective solution to this problem was Network Address Translation (NAT). NAT solution enables hosts in private domain to have a large set of addresses internally called the private addresses, and one address or a small set of addresses externally called public addresses. To separate the IP addresses used within an organization and those used for the external Internet, three address groups are reserved for the private network, shown in Table 1.

Any organization is allowed to utilize an address out of this group without permission from the Internet authorities. It is common knowledge that these reserved addresses are for private networks. They are unique within the organization, but they are not unique globally. Any router will not forward packets with these addresses outside the organization. The organization must have at leadt one single connection to the global Internet through a router that runs NAT software. Figure 2 shows a simple NAT implementation.[14]

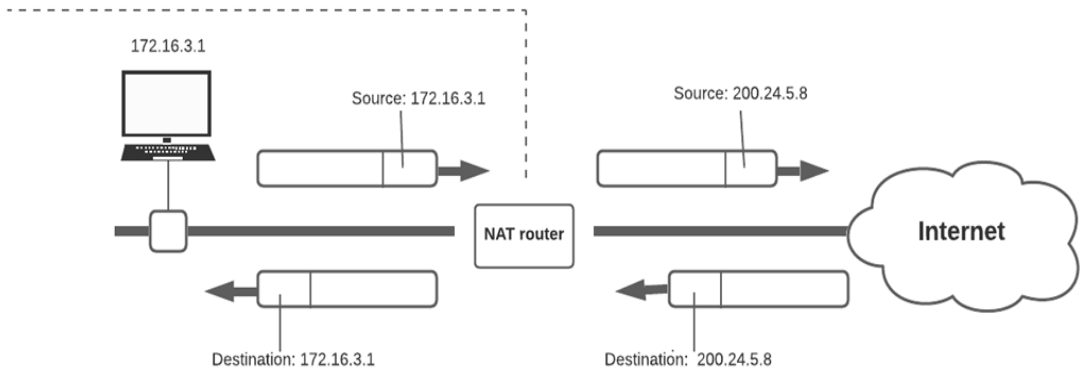

Figure 2: NAT implementation

The router that connects the network to the worldwide Internet uses a set of private addresses and at least one global address. All the outgoing packets pass through the NAT router, which replaces the source address with a global NAT address in the packet header [14]. All packets incoming pass the NAT router, which replaces the destination address with the correct private address. For the incoming packet, the replacement of the IP address is done with the help of a translation table. A simple translation table has two columns: the private address and the external address. When the NAT router translates the source address of the outgoing packet,

it notes the destination address. When the response is received from the destination, the router uses the packet's source address to find the private address of the packet from the table.[14]

In this NAT method, communication is always initiated by the private network or by a box owned by the customer (such as ADSL router). It requires that the private network start the communication. NAT is used by ISPs which assign one single address to a customer. The host may be a part of a private network that has many private addresses. NAT router can use a pool of global IP addresses in case a set of private hosts want to connect to the Internet at the same time. However, no private host can access two external server programs (HTTP and FTP) at the same time. In order to eliminate this and to allow many-to-many relationships between private network hosts and external server programs, port numbers are utilized and added to the translation table referred to as Network Address/Port Translator (NAPT)[14]

2.3.2 NAT issues

Regardless of its variants, NAT performs the address binding statically with a fixed address assignment or dynamically at session initiation. In the address binding stage, all fields in the packet headers like checksum dependent on the source address must be updated. Hence, it is clear that NAT modifies packet headers at layers 3 and 4 to perform its function successfully. Even with its effectiveness NAT still has several problems that require proper solutions.[13]

NAT suffers from a reachability problem. An inbound packet can only be delivered to the private network if its addressing entry is present in the NAT table. Otherwise, the packet will be dropped by the NAT. If the packet payload includes IP addresses, then traditional NAT becomes useless as it does not operate above layer 4. As discussed previously, a client would need to initiate the connection in a NAT scheme. However, this may not be the case always. In a peer-to-peer application, the connection initiation can be established by any of the parties involved. This results in a serious problem in the NAT functionality. Also, NAT does not support all protocols. For example, Stream Control Protocol (SCTP), developed recently and used widely in all media platforms, is not supported by NAT. Encryption protocols limit access to layers 3 and 4, thereby becoming another issue to NAT.[13]

Over the years, there have been many practical solutions to address problems

observed in the NAT. The techniques and protocols proposed as part of the solutions can be classified as NAT Traversal Protocols. STUN - Session Traversal Utilities for NAT, TURN - Traversal Using Relays around NAT, ICE - Interactive Connectivity Establishment, ALG - Application Layer Gateway, UPnP - Universal Plug and Play are few example solutions proposed for NAT problems. These proposals do not act as an effective solution that caters to the scenarios a typical NAT is exposed to. The need for an effective solution that could cater to all the scenarios faced by NAT still exist. From the security perspective, the major issues inherent in IP networking include the lack of stable and trustworthy identities or addresses of the interfaces or devices. This allows using of a fake address when sending a packet and changing the device address or name just after executing some malicious action. Because the network does not keep the state of past communications, this makes it easy for a malicious actor to hide when it so decides. Tracing back to the resources used in the attack is far from trivial. Unfortunately, NAT can make this problem worse while it focuses on solving the address space scalability problem.

2.4 Domain Name System

2.4.1 Introduction

Many services in the application layer of the Internet model follow the client/server model. These programs are divided into two groups: those that can be directly used by the user (like mail) and those that support other services. The Domain Name System (DNS) is a supporting program. DNS provides the translation from a human-entered named destination, for example, website address, to its IP address. As a network service, DNS has evolved from host name-to-IP address lookup utility to enabling very sophisticated lookup applications supporting voice, data, multimedia, and security applications[16].

DNS is effectively a distributed hierarchical database where stored data is available throughout the Internet. A domain name space was designed to organise names into a hierarchy, where names are defined in an inverted tree structure with the root at the top. The tree can have 128 levels (0-127). Each node in the tree has a label and domain name. A full domain name is a sequence of labels separated by dots.[14]

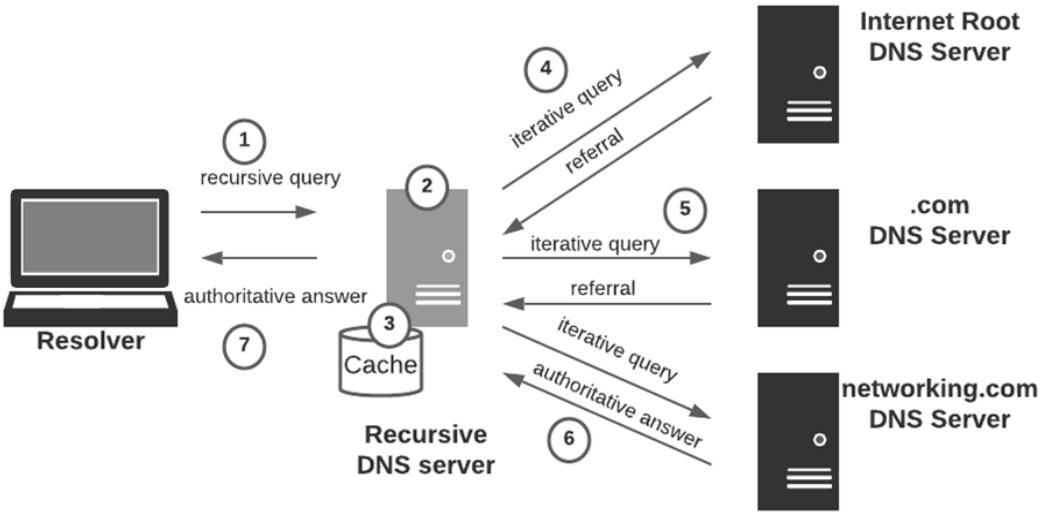

Figure 3: Recursive and iterative queries in name resolution

2.4.2 DNS resolution

The example shown in Figure 3 illustrates how domain information is organized and how a DNS server leverages the hierarchical data structure. The numbers represent the steps followed when a host enters the host domain name "pc10.networking.com." as its intended destination in any application (such as email or web browser). The Application Programming Interface (API) then communicates with the TCP/IP stack called a resolver. A resolver translates the webserver name entered into an IP address that can be used for further communication. The resolver will send a query to the local DNS server, requesting it to provide an answer. The local DNS server will then attempt to answer the query in a recursive way. Recursive DNS query would require DNS server to find the answer to its query if its not avaliable locally. This entire process is termed name resolution.[16]

The resolution process consists of two steps: finding a name server that has the authoritative information to resolve the query and querying that server for the desired information. In Figure 3, the desired information was the IP address corresponding to the domain name "pc10.networking.com." This translation, mapping the domain name to an IP address is saved in the DNS server as a resource record. Different types of resource records are defined for different types of lookups. Each resource record contains a key or lookup value and a corresponding resolution or answer value.

A few examples of DNS records are listed below [16].

- A: Maps a hostname to an IPv4 address of the host.
- PTR: Maps an IPv4 or IPv6 address to a hostname.
- NS: Indicates the authoritative name server for a delegated zone.
- NAPTR: Naming authority pointer that allows expressions encoded as Uniform Resource Identifier (URI).

3 Network Security

This chapter focuses on the concepts of network security. It explains how trust is critical in securing a network and how it can help achieve common security objectives. Security threats are discussed in more detail by understanding the hacking process followed by the list of different types of attacks observed over the Internet and their preventive measures. Finally, the chapter introduces the basics of a Firewall solution.

3.1 Overview

Network security refers to the prevention or mitigation of unwanted intrusion into, use of, or damage to communications on a network. It also includes monitoring for abuses, checking for errors in the protocol, blocking unauthorized transmissions, and responding to any issues promptly. Network security should consist of elements that should ideally prevent undesired activities and support desired actions. This is hard to achieve efficiently, cost-effectively, and transparently. Efficient network security provides easy and quick access to resources for users. Cost effectiveness refers to controlling user access to resources and services without added expense. Transparency supports the enforcement of network security policies without interrupting the way authorized users perform legitimate tasks.[18]

Technology is developing and improving at an unprecedented rate. Connectivity is a part of every individual or companies daily routine. But, at the same time, malicious hackers are becoming more capable of disrupting critical services, stealing identities, information and money by every means possible. Businesses spend more time, money, and effort protecting their assets than they do on the initial installation of the network. Internal or external threats can cause devastating system failure or compromise the whole network. To summarize, network security should support users in doing their tasks while protecting against compromise, maintain high performance, and reduce costs. These goals are tediously challenging, but many solutions exist that aid in achieving the same. All these solutions have one thing in common "Trust."

3.1.1 Trust

Trust is the expectation from the other user that they will act in the best interest. It is the confidence that the user will act in accordance with the security policies and will not attempt to violate the stability, privacy, or integrity of the network and its resources. Trust sometimes is violated, and these violations can be accidents,

oversight, or ignorance. In some cases, it can be deliberate. A user or connected device can be termed as trustworthy based on past experiences and behaviors. Once a set of rules are defined, and every user agrees to abide by those rules, a conditional trust is established. Over time, as users demonstrate that they are willing to follow the rules and meet the expected conduct, then that user can be listed as trustworthy.

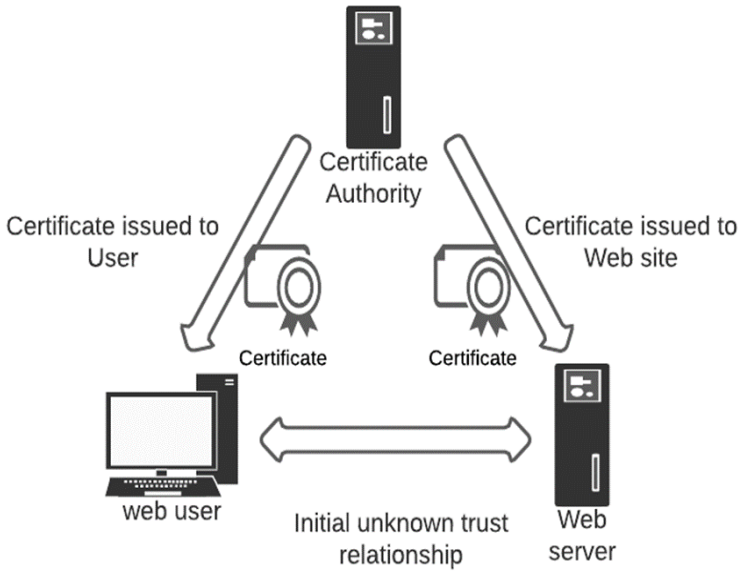

Figure 4: An example of a third-party trust system

Trust can also be relayed to a third party. Suppose a trustworthy third party knows both the parties involved in the communication and that the third party states the parties involved are trustworthy. In that case, it is assumed that the conditional trust is established. A simple example of a third-party trust system is that of the digital certificate issued by the public certificate authority. As shown in Figure 4, a host communicates with a web server. Initially, the host does not know the webserver that it is trying to contact is what it claims to be or if its identity is spoofed. Once, the host, examines the digital certificate issued to the webserver from the same certificate authority that issued the host's digital certificate, the host can then trust the identity of the webserver. This occurs due to the host and web server both having a common trustworthy third party.

3.1.2 Security objectives

In network security, trust is complex. A network is only as secure as its weakest link. Every aspect of the network such as software, hardware, configuration, communication patterns, content, and users need to work together to maintain network security. There are three security objectives namely:

- **Confidentiality/privacy** - protection against unauthorized access. It ensures that data is not disclosed to any unauthorized party.
- **Integrity/nonrepudiation** - protection against unauthorized changes. It ensures that data remain consistent and is not and cannot be changed by unauthorized parties..
- **Availability/uptime** - protection against downtime, loss of data, and locked access, while providing stable uptime, protecting data, and supporting authorized access to resources. [18]

3.2 Security Threats

3.2.1 Introduction

The current Internet is critical for many organizations, including companies, universities, and government entities. Individuals also rely on the Internet for their personal and professional activities. But behind these utilities, there is a dark side where adversaries attempt to create havoc by damaging infrastructures, invading privacy, and blocking critical services on which a user is dependent. With the frequency and variety of the attacks, along with the threat of new and more disruptive planned attacks, network security has become a hot topic in the field of networking.[17]

Devices are connected to the Internet to exchange data. This can include all kinds of services such as web pages, calls, email, streaming, and so on. But these devices can come in contact with malicious content or software, collectively known as 'malware' that can infect the devices. Once infected, the device can be triggered to perform unwanted actions, deleting files, installing spyware, collect personal data, passwords are some of the examples of such activities. The compromised device can also be enlisted in a network of similarly compromised devices often referred to as 'botnet,' which later can be controlled and leveraged for further attacks.

Malware that is present on the current Internet is capable of self-replicating. Once a device gets infected with malware, the malware would then seek entry into other devices over the Internet, and this continues as a chain process affecting as many devices as it can. Malware can spread in the form of a virus or a worm. Viruses are malware that require user interaction to infect the user's device. A simple example is that of an email attachment with a malicious link; upon clicking the same, the virus would then get activated and start spreading. Worms are malware that can perform their actions without user interaction—for example, connecting to a public network or downloading untrustworthy applications. The more popular class of security threats are the denial-of-service (DOS) attacks, which do not intend to steal any information but rather try and disrupt network or another bit of infrastructure rendering resources unusable to users. Another critical threat is that of IP spoofing, where packets with the wrong source address are injected into the Internet, with the intention of posing as another user. More attacks are discussed further. [17]

3.2.2 Hacking process

A hacker or a hackers group may attack an individual or an establishment by either choosing to steal the data or make it unavailable. It is important to learn how these attacks are performed and propagated throughout the Internet. Their activities appear chaotic and random where no fixed procedures are followed. They seek vulnerabilities on a specific target using any or all means at their disposal. Hacking can be grouped into five main sub-groupings of events, representing only the hacking process, not the mechanism to prevent them. Figure 5 depicts the hacking process sub-groupings, where the order of stages occurs if an attack is successful; else, the hacker can attempt a fall-back attack.

- **Reconnaissance** - The act of exploring and inspecting that can be referred to as footprinting, research, discovery, and information gathering, where the hackers learn as much information on the target as possible.
- **Scanning** - The act of using various tools to verify the information gathered in step one. The main aim is to discover active systems using methods like ping sweeps and port scanning.
- **Enumeration** - The process of identifying sufficient details about a specific target and learning whether a vulnerability that can be attacked exists. Operating system identification, application identification followed by extraction of information are the steps followed in this stage.

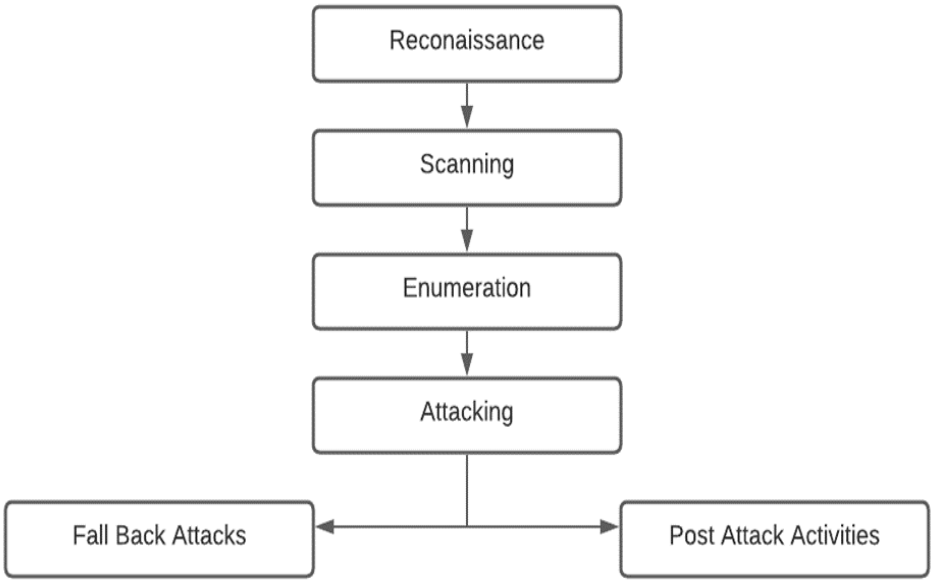

Figure 5: Five stages of hacking.

- **Attacking** - The attacking phase is usually the most hyped, as it is the briefest stage in the whole process. A successful attack with solid research can take just seconds. Even if the attack fails due to the host's defenses, repeated attempts of the attack can sometimes frustrate the defenses. If the attack succeeds, then the attacker goes ahead with the post-attack activities else with fall-back attacks.
- **Post-attack activities** - In this stage, the hacker has successfully breached the system and can perform the intended damage.
- **Fall back attacks** - If the attack was unsuccessful, the hacker could still perform alternative attacks such as DoS, eavesdropping, session hijacking, man-in-the-middle attacks, and more.

3.2.3 Types of security attacks

In this section some of the most common attacks along with thier preventive solutions are are discussed.

Eavesdropping - Eavesdropping refers to the listening in on communications, which can include recording of network traffic with the help of a packet sniffer. Any form

of plain and directly usable data is often subjected to interception and recording. Eavesdropping can be prevented by using encrypted protocols.

Replay attacks - Replay attacks, also called playback attacks, refer to the retransmission of captured communications with a goal of gaining interactive or session access to the target system. Replay attacks can be prevented by using session-based encryption, one-time pad, timestamps.

Insertion attacks - Insertion attacks have many forms, and all of them include the introduction of unauthorized content or host to a secure infrastructure. SQL injection, IDS (Intrusion Detection System) insertion, and rogue devices are the commonly observed insertion-based attacks.

- **SQL insertion** lets the attacker inserts code into a script hosted on a website, thereby giving access to the back-end database of the web application. SQL insertion can be prevented by following defensive programming and filtering input.
- **Intrusion Detection System (IDS)** insertion lets the attacker exploit the nature of the IDS to collect and examine every packet to deceive the IDS into believing an attack took place when it really hasn't. IDS insertion can be prevented by using modern IDS techniques with an anomaly, heuristic, and behavioral detection.
- **Rogue device** insertion is a kind of physical insertion attack where the attacker inserts an imposter device into the infrastructure. Rogue device insertion can be thwarted using encrypted communications, pre-configured network access, and regular site surveys. [18]

Fragmentation attacks - Fragmentation attacks are the direct exploitation of the fragmentation offset feature of the IP packets. Fragmentation occurs when specific network segments cannot support large datagrams. These large datagrams are fragmented into a more node-compatible size. When the fragmented frames reassemble, manipulations of the datagrams can occur, which can cause several malicious reconstructions, such as overlapping- overwriting of datagrams and Overrun- results in very large datagrams. Protection against fragmentation attacks can include IDS detection, smallest Maximum Transmission Unit (MTU) acceptable over the network, and firewalls.

Buffer overflows - Buffer overflow attack allows an attacker to inject more additional data into a buffer than it can hold, leading to the additional data taking the next area of the memory. This overflow can cause a crash, freeze, or arbitrary code execution. Defensive programming can prevent buffer overflow attacks.

Man-in-the-middle-attacks - MitM attacks happen when a hacker meddles in a communication session between a client and a server by tricking the client into initiating the session with the hacker's server instead of the original server. This attack involves a pre-attack element, in which the client is fed wrong information that leads it to start a session with the hacker's server. The hacker can achieve this in several ways listed below:

- **ARP spoofing** - ARP broadcast service can be spoofed to send false MAC addresses to the host.
- **MAC spoofing** - The hacker's server impersonates the MAC address of the real server, wherein the traffic would be directed to the fake server when the real server is loaded with traffic.
- **DNS poisoning** - An attacker compromises a DNS server by placing incorrect FQDN to IP mapping records.
- **DNS spoofing** - A hacker can host a rogue DNS, sending incorrect DNS responses to the client's DNS query.
- **ICMP redirect** - A host alters its routing table when ICMP redirects occur on a subnet with multiple routers. This attack could allow the attacker to redirect the traffic along a different route than the expected one.
- **Proxy manipulation** - A client's proxy configurations are altered in a way that the request for the services goes via the hacker's system that acts as a MitM proxy.
- **Rogue DHCP** - A compromised DHCP server can provide IP address configuration leases for a subnet and define a gateway since the hacker's host acts as a router/proxy.
- **Rogue access point** - A hacker can configure a rogue wireless access point that can trick users into connecting, serving as a proxy.

Defenses against MitM attacks include IDS and IPS methods that can monitor for network attacks or abnormal activities.

Session hijacking - Session hijack occurs when an attacker takes over the connection after a client has authenticated with a server. Initially, to perform the attack, the attacker must eavesdrop on the session to learn details, such as the address of the session end-points and the following sequence numbers. Armed with this information, the attacker can desynchronize the client, take on its address, and inject modified packets into the data stream. If the server accepts the initial modified packets as valid, then the session has been hijacked. The attacker can employ some of the MitM attacks to alter the route of a session. Any device that uses TCP/IP without encryption is exposed to session hijacking. Virtual Private Network (VPN) is a robust protective measure against session spoofing.

Spoofing attacks - In general, spoofing means the falsification of the original information. E-mail addresses, MAC addresses, and IP addresses can all be spoofed, and the user is tricked into believing a communication originated from somewhere other than its actual source. Spoofing is quite challenging to prevent and only moderately detectible. Hence, there is no direct preventive mechanism for spoofing rather than monitoring the traffic for any anomalies.

Denial of Service (DoS) DoS attack interrupts the normal functioning of traffic and communications. DoS can be of two types: flaw exploitation and traffic generation.

- **Flaw exploitation DoS** attacks exploit a programming bug or convention, resulting in system freezing, crashing, rebooting, or failing to respond to communications. Flaw exploitation can be mitigated with the help of patches and IDS or IPS solutions.

  An example of this type is the vulnerability found in Linux and FreeBSD kernels uncovered by Juha-Matti Tiili from Aalto University, Department of Communications and Networking / Nokia Bell Labs in 2018. This vulnerability allows a hacker to carry out denial of service attacks with low packet volumes. The attack is carried over a two-way TCP connection by sending specially designed packets within ongoing TCP sessions, thereby eliminating the need of spoofed addresses.[32]

- **Traffic generation DoS** attacks flood a target with huge traffic that can

consume bandwidth and processing, preventing authorized communications. Traffic filtering, especially upstream filtering, is the only practical way of avoiding this attack. Upstream filtering refers to a parent network (ISP), which provides traffic filtering before entering the internal network to which the entities are connected.

Distributed Denial of Service (DDoS) - The DDoS attack is an advanced version of the DoS attack, which attacks through massively distributed processing and processing with the help of agents such as a bot or a zombie. These agents can create a network of their own known as the botnet army or the zombie army. Figure 6 depicts the DDoS attack where the attacker remotely regulates the botnet and guides it to perform several malicious activities.

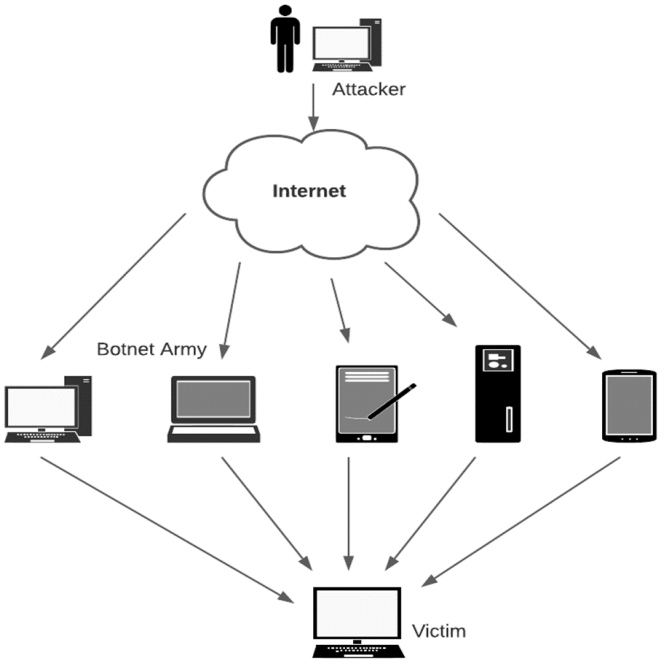

Figure 6: Distributed denial of service flooding attack against a primary target.

Generally, the foremost victims of the botnet are known as primary victims, while the compromised hosts hosting the botnet's agents are known as secondary victims. These botnets distributed throughout the internet can perform many malicious activities, including flooding, spamming, eavesdropping, intercepting, MitM, session hijacking, spoofing, packet manipulating, malware distributing, phishing site hosting, passwords stealing, encryption cracking, and more[18]. Firewall, antivirus,

anti-malware scanning, and IDS/IDP solutions are some of the preventive measures against DDoS. ISP-level response to a DDoS attack is engaging a brush-up service on-demand when under a DDoS. A brush-up process is executed on a powerful cloud platform involving a large number of CPUs to process the packet flow, identify and filter the DDoS packets while letting the legitimate traffic pass through.

BGP route hijacking - Border Gateway Protocol (BGP) is an old and still in use protocol that runs at the backbone of today's Internet. BGP is mainly used to communicate between different Autonomous Systems (AS), referred to as Inter-AS communication, conveying routing information between ASes. A small issue in the protocol might cause a devastating impact across the global Internet. BGP lacks peering authentication by default. Also, there is no verification method of prefix or routing information received from an authorized peer. By default, BGP will not check whether the prefixes advertised in the update message are owned by the advertiser or authorized by the owner to advertise these prefixes[23]. Hence, a small error can navigate across the network over peer by peer. All the above reasons support the fact that BGP is highly vulnerable to security issues.

BGP prefix hijacking cases are widely available online, one such example is mentioned here. On June 2019, European mobile traffic was rerouted through China Telecom when an AS announced more than forty-thousand IPv4 routes that had been learned from other peers and providers to its provider China Telecom. [33]

3.3 Firewalls

A firewall is a device placed between the internal and external network designed to forward some packets and filter (not forward) others. Firewalls filter traffic using rules. Whether stateful inspection, static packet filtering, circuit proxy, application proxy, or content filtering, all firewalls use rules to filter traffic. Rules are described as lines of instruction that assess and take action on network traffic[18]. Generally, two rules define the foundation for rules administering traffic crossing the firewall: default permit or default deny.

- A **default deny** rule assumes that all traffic is possibly malicious or at least undesired or unauthorized, and thereby everything is obstructed by default. Valid traffic is granted access through an exception rule. This is referred to as deny by default, allow by exception.

- A **default allow** rule assumes that most traffic is harmless, thereby allowing everything by default. Unwanted and unauthorized traffic is blocked with the use of an exception rule. This is referred to as allow by default, deny by exception.

Firewall rules control what traffic enters or leaves a secured network. Depending on its position, it can protect the private network from the public Internet or even filter traffic between the internal network. A firewall is classified into a packet-filter firewall or a proxy-based firewall.

- A **packet-filter firewall** filters packets at the network or transport layer. It forwards or blocks packets based on the header information in the network layer and transport layers, such as IP addresses, port numbers, and protocol (TCP or UDP). A packet-filter firewall uses a filtering table to decide which packets must be dropped[18].

- A **proxy firewall** filters messages based on the information available in the message itself at the application layer. This is realized by placing a proxy application gateway between the client and the server. When the client sends a message, the proxy firewall runs a server that receives the request. The server later verifies the packet at the application level and checks if the request is legitimate. If the verification succeeds, the message is forwarded to the real server; else, an error message is sent back to the client[18].

Even though firewalls are an essential part of the network security infrastructure and much trust is imposed, they are not perfect solutions. Ultimately, firewalls are software code prone to bugs that can lead to vulnerabilities that are exploited to trigger an array of attacks such as buffer overflows, fragmentation attacks, overlapping, and overrun. Another limitation of firewall is Firewalking. It is a technique used to learn the configuration of a firewall from an external network using an internal host's valid IP address. A hacker then attempts to establish a communication session from an external network to an internal host over many different ports. The main goal is to discover the rules or filters on a basic packet filtering firewall.

4 Customer Edge Switching

This chapter introduces Customer Edge Switching; its architecture and operation are discussed in more detail. Furthermore, the CES-to-CES mode of communication is explained with an example, and the chapter ends with brief information on Customer Edge Traversal Protocol.

4.1 Overview

Customer Edge Switching (CES) is a firewall solution proposed as an extension and replacement to a traditional NAT—developed at Aalto University by the research group lead by Prof. Raimo Kantola under the department of Communication and Networking. CES follows a trust-to-trust communication model to provide global connectivity for hosts in private networks. CES's typical use case is providing and ensuring adequate security for devices such as IoT in the edge nodes, which are governed by the policies set by the users or administrators[20].

Security in CES relies on the trust relation created between the network nodes based on the parameters exchanged during the session establishment stage[20]. It implements network algorithms and functions that address security and the reachability problem introduced by a traditional NAT. It also enforces a cooperative behavior between hosts within a network served by CES nodes. These nodes act as a connection broker by executing host policies, such that the packet is forwarded to the receiver only if the sender meets the requirements of the receiver. Moreover, it is also responsible for managing the identities and addresses of all the served hosts[1]. A feature of CES is that it provides assurances to remote networks about the host identity. CES supports the idea of Identity (ID), routing locator (RLOC) split. End hosts can have a globally unique identifier, such as a Fully Qualified Domain Name (FQDN) or a Mobile Subscriber ISDN (MSISDN) number, while the packets are routed based on addresses (RLOC) and before communication can take place the ID of the destination must be translated to an address.

CES deployment is simple without any need for significant upgrades on hosts or protocols. Apart from replacing the NAT devices, CES relies on the Domain Name System (DNS) to map the FQDN of a host to the IP address of the edge node serving the host[20]. Changes to the DNS infrastructure such as DNS protocol or the DNS record types are unnecessary except for some special configurations.

4.2 CES Architecture

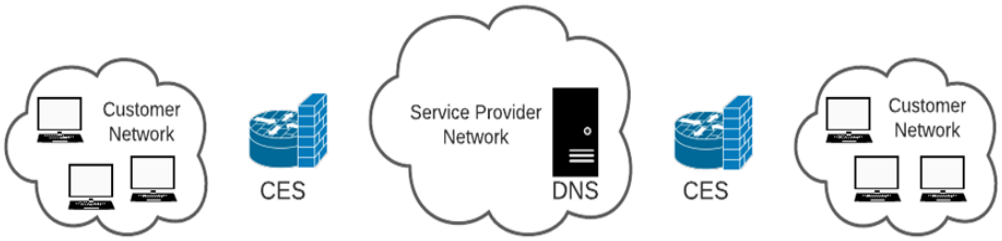

Figure 7: CES Architecture.

CES architecture divides the global network into the Customer Networks (CN) and the Service Provider Network (SPN). A set of End users are situated within a CN, and CES is deployed at the edge of the customer network, acting as a trust boundary. CNs are independent and isolated from each other with no direct communication. The CES device has at least two interfaces that connect to the CN and the SPN network[15]s. Figure 7 depicts a simple architecture of CES where a group of users together form a CN, separated from SPN by a CES node.

4.3 CES Communication Modes

Figure 8 shows the different modes in which a CES can operate. These signaling cases can be Inter-CES Communication and Intra-CES Communication. In Intra-CES, communication takes place between hosts of the same CES. In Inter-CES, communication takes place between the hosts of different networks served by other CES nodes. In the case of legacy Internet, a CES node will have a Private Realm Gateway (PRGW) which uses well-defined protocols to provide a source NAT (SNAT) and a destination NAT (DNAT) functionality. CES operation is divided among the Control Plane (CP)- responsible for initiation and establishment of signaling sessions and user connections and the Data Plane- concerned with routing user's data packets. Apart from using the standard protocols, CES nodes depend on a new protocol named Customer Edge Traversal Protocol (CETP) that is not standardized yet. The current version of CETP is the experimental version.[20]

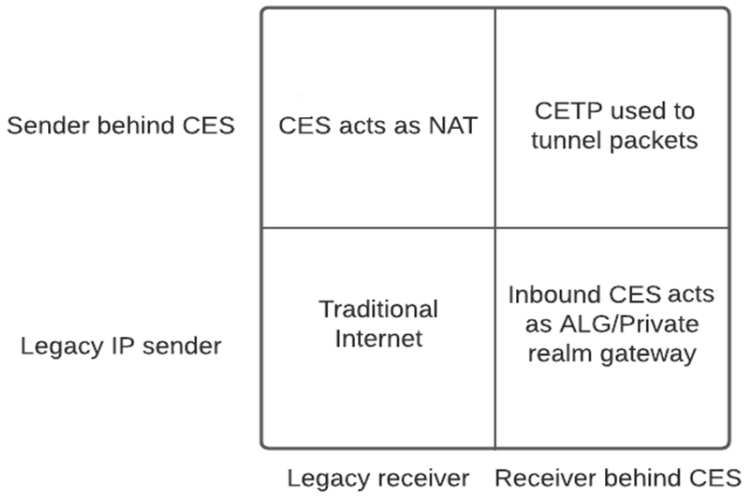

Figure 8: Modes of CES.

4.4 CES Operation

As mentioned earlier, each host has a unique FQDN through which it is reachable on the global internet. To resolve an FQDN to a reachable Internet address, CES uses DNS as a means of initiating communication, resulting in CES architecture to be tightly coupled with the Domain Name System (DNS). To figure out if the remote end supports CES, the originating CES node initiates a Naming Authority Pointer (NAPTR) DNS query towards the destination as part of CETP service discovery[1]. A DNS NAPTR response indicates CETP service availability at the remote end thereby confirming originating CES to perform cooperative firewall functions towards the destination, for host-to-host connection. If the response indicates the unavailability of the CETP service on the remote side then the original host DNS query is forwarded to the Internet.

Policies in CES are defined as a set of rules and parameters upon which a connection establishes, and data communication occurs using several keywords and parameters. CES holds three types of policies: CES-CES Policies (C2C), Host-Host Policies (H2H), and firewall policies. A C2C policy and H2H policy are involved in session creation. Firewall policies are concerned with packet filtering. C2C and H2H are grouped under the control plane and firewall policy under the data plane.[20]

CETP policy negotiation follows a successful CETP service discovery process.

When a NAPTR response indicating the availability of CES functions on the remote end is received, the outbound CES (oCES) node initiates a three-layered signaling channel towards remote CES, where the bottom layer maintains a set of transport connections with remote CES. One layer maintains CES-to-CES (C2C) relation, allowing the two CES nodes to exchange the network-specific policies, events, and rules for the underlying exchange of user-data connections. The last layer is responsible for host-to-host (H2H) or host-to-service policy negotiations.[20]

After the policy negotiations succeeds, both CES nodes allocate a proxy address to represent the remote host within their private network. They also insert flows to the OpenvSwitch, which aids in tunneling the data across the network. The originating CES node would respond to its host with the proxy address of the remote end and any futher communication would use these proxy addresses of the hosts.

4.5 CES-to-CES Communication

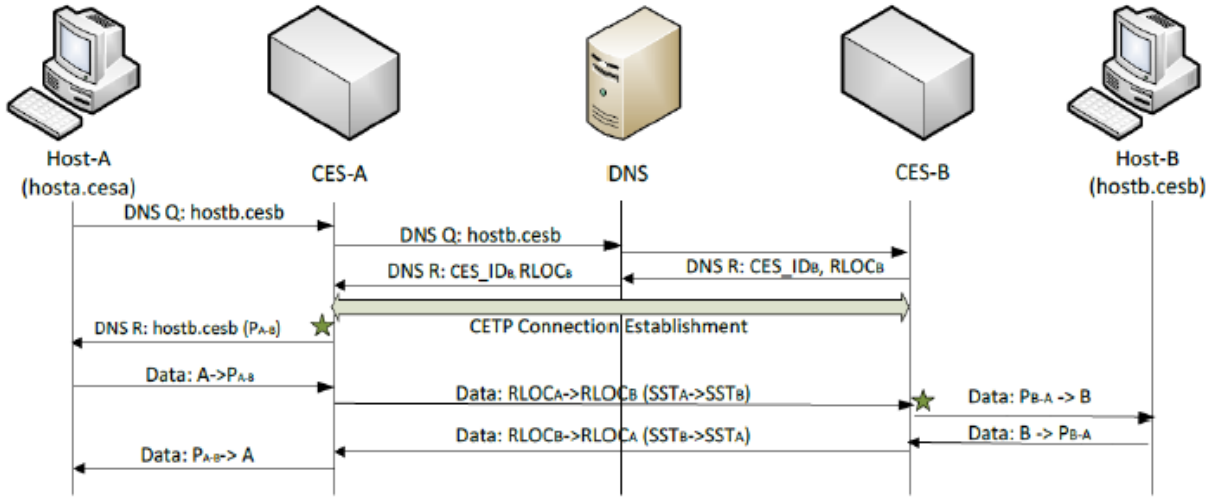

Figure 9: CES-to-CES Communication.

Figure 9 describes a CES-to-CES communication scenario where Host-A connected to CES-A tries to communicate with Host-B connected to CES-B. In the connection initiation phase, Host-A queries a DNS server for the name resolution of the FQDN of Host-B. CES-A receiving this request checks if the FQDN belongs to its list of local hosts. Once the CES-A knows that the host is not part of the local host list, it translates the DNS query to a NAPTR query and forwards it to the

DNS server. Upon receiving the request, the DNS server identifies the remote CES (CES-B) node based on its name server record and forwards the request to the DNS server of CES-B. CES-B replies with CES-B routing locator(S) in this case RLOC, which implies that the destination host is behind a CES and can be accessed via the RLOCs through the CES service.

CES-A receives DNS NAPTR response and initiates a policy negotiation phase where a connection state is established based on the principles and parameters agreed between the two CES nodes. Furthermore, both CES nodes create a state entry in their connection table and store the binding between the private address, the proxy address, local/remote RLOCs ports, and session tags. Also, CES-A modifies the DNS NAPTR response to carry the proxy address of destination Host-B and forwards it to Host-A. Host-A would then start sending the data packets using the proxy address, assuming it to be the Host-B address. When this packet reaches CES-A, CES-A matches it with the saved state entry in the connection table and forwards it to CES-B, complying with the negotiated parameters. CES-B receives the packet and checks the inbound packet according to its stored state, and delivers the translated packet to Host-B.[20]

4.6 Customer Edge Traversal Protocol (CETP)

```
{ "request": [
    { "ope":"query", "group": "id", "code": "fqdn" },
    { "ope":"query", "group": "control", "code": "caep" },
    { "ope":"query", "cmp":"optional", "group": "control", "code": "hard_ttl" }
  ],
 "offer": is a subset of the 'available' policy elements, and we skip showing it for brevity sake,
 "available": [
    { "ope":"info", "group": "id", "code": "fqdn", "value":"a1.cesa." },
    { "ope":"info", "group": "control", "code": "caep", "value":"195.148.124.145" },
    { "ope":"info", "group": "control", "code": "hard_ttl", "value":""}
  ] }
```

Figure 10: CETP policy format.

CETP's main role is to deliver signaling packets from one CES node to another. The introduction of the CETP in the network has no impact on the end host. The end host is oblivious to CETP in the network. The latest CETP protocol is represented

in a JSON-style format, which makes it possible to exchange the CETP policies as messages over transport protocols, such as TCP or TLS. The flexible CETP message format allows rapid prototyping and easy extension of protocol for any future purposes and mitigates the delay introduced by slow packetizing and de-packetizing of CETP, which was directly carried over IPv4 and IPv6, in the previous protocol version.[34]

Figure 10 shows an example CETP policy format. Each policy is expressed using three distinct sets: *offer*, *request*, and *available*. Each set carries the policy elements, which collectively define a policy. The *available* contains all the policy elements supported by a host or a network's policy. The *offer* is a subset of available and it contains the policy elements offered by the CES node (of the sender) at the start of the policy negotiation. The *request* contains the policy elements that are requested from the remote CES node. Each policy element within CETP is further structured into a set of fields namely: operation, compatibility, group, code and value fields each supporting standard values.[34]

4.6.1 CETP Security

CETP provides a range of tools and procedures available to the CES devices to effectively ward off some of the security issues observed on the Internet. Table lists some of the security mechanisms offered by CETP.

Security Threat	**Counter measures offered by CETP**
ID Theft	Signature, executing assurance query via CA address Type Length and Value (TLV)
Eavesdropping, MitM, Data modification,	Signature
IP(ID) spoofing, spamming, SYN flooding	Return routability check on naming and forwarding level
DDoS attack	Reporting attack via unexpected message TLV and embedded reputation system

Table 2: Security mechanisms of CETP.

Return routability checks

CETP proffers return routability checks on naming and forwarding level to detect packets with spoofed source address or ID by checking whether the sender is accessible at the claimed address or not. CETP exploits control TLVs, including cookie and

FQDN, to perform a return routability check mechanism.

State management

Each communication state depends on its timeout and can be active for a defined period of time. This entry will be removed from the table if no relevant packets are received within the timeout. When the timeout of state information expires, the CES deletes the state, informing the peer about the state expiry, and requests the removal of the corresponding connection state by sending a timeout TLV with a zero value.

ID management

Endpoints can negotiate the types of ID that must be used by the other end by performing ID negotiation. An edge node creates a query for the required ID type and directs it to the peer. On the remote end, the next packet is built with the requested source ID. The inbound edge creates a new connection state and delivers the data packet to the destination host upon the first packet.

Signature

To avoid hijacking of RLOCs, it can be signed cryptographically.

Reporting attacks

A trust domain can report a problem upon attack detection to the concerned entity if it is likely compromised.

Maintaining trust-to-trust relations in CES assumes that the remote CES is forced to cooperate because if a CES node is detected to be malicious, all traffic from the served network will suffer. To maintain this assumption, a CES node contains an embedded reputation system allowing an inbound CES to refuse communication with a remote outbound CES based on the reputation of the outbound CES. A CES reputation is formed based on all the evidence of the CES behavior and the behavior of the hosts served by it.

5 SCION

This chapter provides a brief description of SCION by explaining its network infrastructure and operation. It also provides details on the use-cases and deployment scenarios that SCION caters to.

5.1 SCION Architecture

Scalability, Control, and Isolation on Next-Generation Networks (SCION) is a clean-slate Internet architecture designed to provide highly available and effective point-to-point packet delivery, even in the presence of malicious attackers in the network. SCION is an open-source Internet platform project developed and managed by the Network Security Group, ETH Zürich. Currently, there are two active deployments of the SCION network: the research version and the production version. The production version is managed by a company called Anapaya systems[31]. The research version named "SCIONLab Network" is managed by the SCIONLab team and designed to test and experiment with the SCION architecture[31].

The SCIONLab infrastructure constitutes a network of globally connected Autonomous System (AS). The research version is used for this thesis and has certain restrictions. For example, SCIONLab centralizes management of the control plane Public Key Infrastructure (PKI), where the single point of failure is absent in the actual deployment. Another restriction is that the SCIONLab infrastructure uses overlay links over the publicly routed Internet. These restrictions can hamper the full potential of SCION. The production version of SCION is already in use by many ISPs, banks and the Swiss government.

5.1.1 Overview

SCION utilizes an isolation domain (ISD) as its fundamental building block for achieving high availability, transparency, scalability, and support for heterogeneous trust. It establishes a logical grouping of autonomous systems (ASes), as shown in Figure 11. An ISD can be governed by several ASes called core ASes. An ISD typically contains various regular ASes governed by a trust root configuration (TRC). The TRC represents the roots of trust used to validate bindings between names and public keys or addresses.[27]

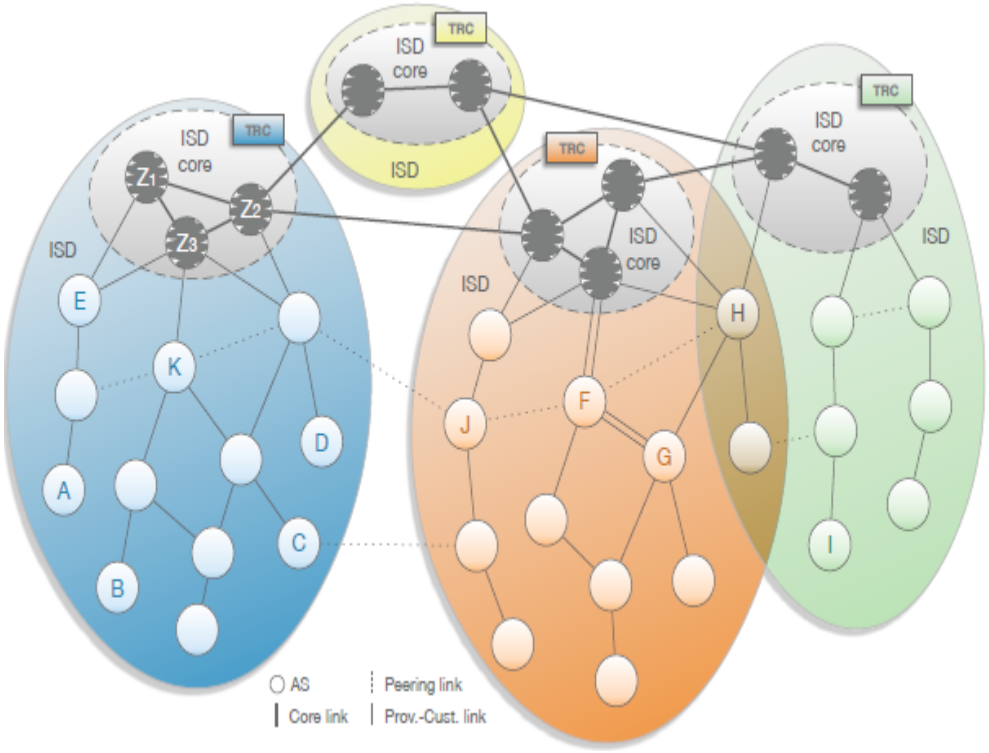

Figure 11: Grouping of Autonomous systems (ASes) into ISDs. [27]

An AS can join an ISD by obtaining connectivity from another AS in the ISD. This connection indicates an acceptance of the ISD's TRC. Generally, 3–10 ISPs constitute an ISD core, and their associated customers participate in the ISD. All ASes within an ISD also agree on the TRC, i.e., the entities that operate the trust roots and set the ISD policies. Even though an ISD ensures isolation from other networks, its primary purpose is to provide transparency and maintain heterogeneous trust environments.[27]

SCION is a "path aware" network. Every connected AS within the SCION network would perform a path exploration phase and find the cryptographically protected AS-level path information relating to its neighboring ASes. Path exploration is followed by a Path registration phase where the discovered path information is registered as Path segments with a dedicated SCION native Path server. End-to-end communication in a SCION network begins with the name resolution of a SCION address, followed by a path lookup phase where the end host would get the path segments from the path server and construct a forwarding path to the destination in the path combination phase. At the end of the steps mentioned above, the path

information from the source to destination is encrypted placed in the SCION packet's header that traverses the network.

5.1.2 Components of a SCION autonomous system

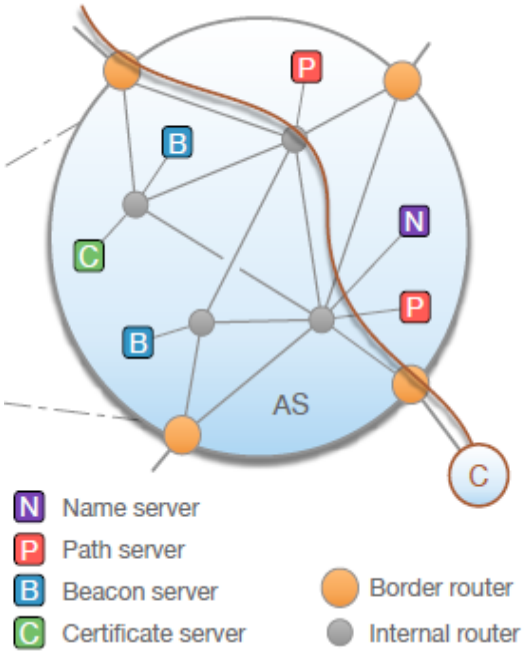

Figure 12: Components of a SCION AS.
[27]

Each component as shown in Figure 12 has been designed for a specific function that aids in the connectivity throughout the network.

Beacon servers discover path information. They are responsible for generating, receiving, and propagating path-segment construction beacons (PCBs) to explore paths and construct path segments. The entire process is called beaconing. SCION maintains two levels of beaconing: intra-ISD beaconing (construct path segments from a core AS to non-core ASes within an ISD) and inter-ISD beaconing (construct path segments amongst core ASes within an ISD and across ISDs).[27]

Non-core AS beacon servers receive PCBs and re-send them to the customer ASes, which provides AS-level path segments. At every AS, information about the ingress and egress interfaces of the AS is added to the PCB. The ingress and egress interfaces identify the link to a neighboring AS. Regularly, a beacon server generates

a set of PCBs, which it forwards to its customer ASes. Inter-ISD beaconing is similar to BGP's route-advertising process. However, in SCION, the process is periodic, and PCBs are flooded over policy-compliant paths to discover multiple paths between any pair of core ASes.[27]

Name servers in SCION are similar to DNS servers, translating a human-understandable name into a SCION address. SCION introduces the RAINS (Another Internet Naming Service) system for this purpose, and based on the (ISD, AS) tuple, an end-to-end path can be looked up and constructed. The end-host address and end-to-end path are placed in the SCION packet header to enable delivery to a given destination.[27]

Path servers cache mappings from AS identifiers to sets of announced path segments and are organized as a hierarchical caching system similar to DNS. ASes select the set of path segments through which they want to be reached and upload it to a path server in the ISD core.[27]

Certificate servers store cached copies of TRCs (retrieved from the ISD core), AS certificates and manage keys and certificates for securing inter-AS communication. Certificate servers are queried by beacon servers when validating the authenticity of PCBs.

Border routers connect different ASes supporting SCION and carry out the task of forwarding packets. Forwarding can be towards a server when a packet having a service address is received or towards a host within the AS or next border router in the case of a data packet. SCION can run using any communication framework inside an AS (e.g., OSPF, SDN, MPLS).

5.1.3 SCION packet format

A high-level format of a SCION packet is shown in Figure 13. SCION packet header consists of a common header, addresses, and forwarding path.

Common header - A compulsory header in the first 8 bytes of the packet holding encoded information of, the length of the packet, the types of the source and destination address, the current position in a path, and the type of the following header (an extension or layer-4 protocol).[27]

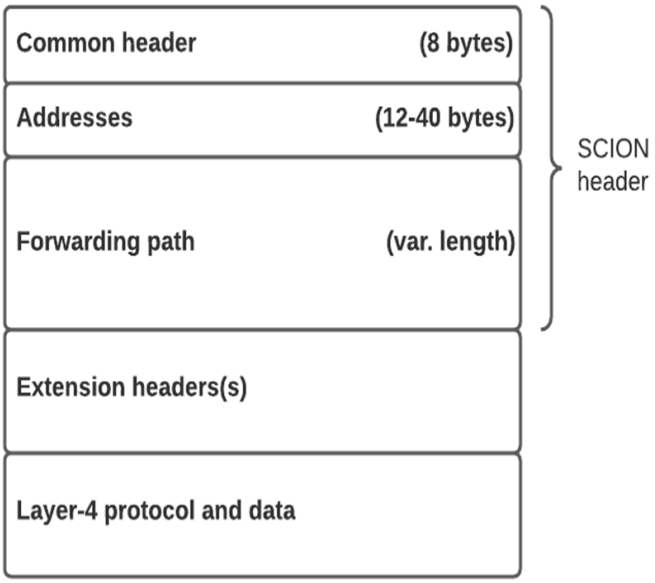

Figure 13: SCION packet Format(High-level).

Addresses - Source and destination addresses are arranged after the common header. A SCION address consists of an ISD identifier, an AS identifier, and an end-host address. SCION allows ASes to use different address spaces (e.g., IPv4, IPv6, or MAC addresses) to address their end hosts. It also permits hybrid addressing where a source and destination may have addresses of different types.[27]

Forwarding path - consists of a sequence of info fields (INF) and hop fields (HF) containing the information required by border routers for packet forwarding. The forwarding path is blank for SCION packets that do not leave the source AS.[27]

Extension headers - extensions point to the layer-4 header. There are two types of extension headers:

- **hop-by-hop extensions** - processed by source and destination end hosts, as well as by every border router on the path.
- **end-to-end extensions** - processed only by source and destination end hosts. [27]

Layer-4 protocol and data - encapsulated payload within a layer-4 protocol.

5.1.4 SCION addresses

SCION address comprises three values in a tuple: ISD Idententifier, AS identifier, and an end-host IP address. The ISD identifier is globally unique, the AS identifier is locally unique within the ISD, and the end-host address is routable within the AS.

The current SCION implementation uses 12 bits for the ISD identifier and 20 bits for the AS identifier, allowing up to 4,096 ISDs globally and 1,048,576 ASes per ISD. Value 0 for both ISD and AS identifiers is reserved. The size of the end-host address is variable and depends on the address type. The length of a complete SCION address is not fixed and depends on the type of the end-host. The sizes are determined as shown in table 3.[27]

Type	**Size**
IPv4	$8B$
IPv6	$20B$
Service	$4B$

Table 3: Size of a SCION address.

A combination of IPv4 or IPv6 as a source address and Service type for a destination address is used for sending control-plane requests to a SCION service. The type *Service* is introduced to inform the destination AS that a given packet should be sent as an anycast packet to an instance of the correct service.[27]

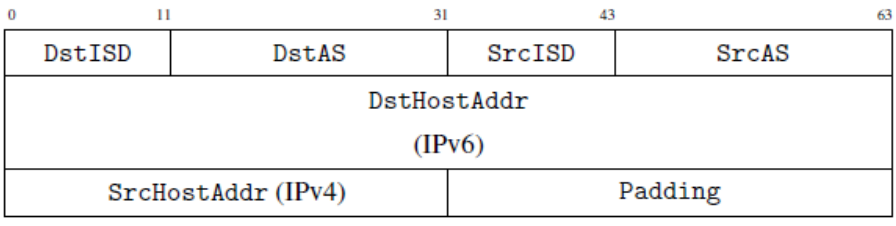

Figure 14: SCION address size for different types of end-host.

To enable border routers easy access to ISD-AS identifiers and the destination address (accessed more frequently than other values), placement of the values follows Figure 14. Destination and source ISD-AS identifiers are placed at the start, which is then concatenated with destination and source host addresses. Padding

may be required in case of a hybrid addressing where IPv4 and IPv6 are used together.

5.2 SCION Operation

5.2.1 Communication flow

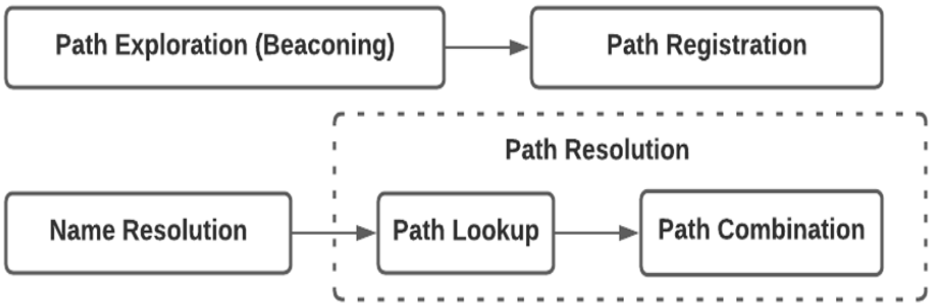

Figure 15: Process followed to create a forwarding path.

Figure 15 presents the sequence of phases required to obtain an end-to-end path for a SCION host to communicate within its network.

In the **Path exploration** phase, SCION ASes discover available paths to the core ASes. A core AS announces a PCB and pushes it as a policy-constrained multipath flood within an ISD to explore intra-ISD paths or amongst core ASes to explore inter-ISD paths. PCBs save encrypted AS-level path information and this entire process is named as beaconing.

In the **Path registration** phase, SCION ASes transform PCBs into path segments and register them with the Path servers situated within the network. This registration ensures that the path segment is available for any valid SCION host to communicate in the network.

In the **Name resolution** phase, domain names are translated into their corresponding SCION address.

In the **Path resolution** phase, the end host creates an end-to-end forwarding path to the required destination. This phase consists of two stages: path lookup,

getting the path segments from the path servers, and path combination, creating a forwarding path from the received path segments.

Each of the phases are explained in more detail in the following sections.

5.2.2 Path exploration and registration

Path-Segment Construction Beacons (PCBs)

PCBs are used for path exploration in both intra-ISD and inter-ISD. They contain topology and authentication information. Additionally, they may include metadata that assists with path management and selection. Generally, a PCB represents a single path segment that can be used to construct end-to-end forwarding paths.[22]

Each PCB consists of one info field (INF) and many AS entry fields (ASE). The info field (INF) provides basic information about the PCB. It provides:

- The type and the direction of the constructed end-to-end path.
- A timestamp that denotes when the PCB's propagation started.
- An identifier of the isolation domain.
- The length of the forwarding path's segment.

The AS entry field holds information about an AS, and it consists of many sub-fields such as certificate fields, meta fields that carry MTU size, and ISD identifiers. It includes a Hop Entry (HE) field, which in turn contains Hop Field (HF) used for the data plane packet forwarding which specifies the incoming and outgoing interfaces of the ASes on the forwarding path.

Intra-ISD path exploration and registration

The PCB formation process is started by each core AS once per propagation period. When an AS receives a PCB, its beacon server registers the contained path segment at the path servers, extends the PCB, and propagates the PCB further downstream. The ingress border router of the downstream AS receives the PCB packet and sends it to one of its beacon servers. The beacon server then verifies the structure and the signature of the PCB. Once verification is successful, the beacon server adds the PCB to its local database. Every propagation period, the beacon server selects the PCBs from its database and continues path exploration by sending

them to its downstream ASes. These selection criteria are set according to local AS policies.[27]

The paths need to be published during the registration phase to make them available to their own and remote end hosts. At a time interval called a registration period, a beacon server selects two sets of path segments: up-segments - to allow a local end host to contact core ASes, and down-segments - to allow remote end hosts to fetch paths from core ASes towards a target AS.[27]

Inter-ISD path exploration and registration

Path exploration in inter-ISD follows the beaconing process conducted only by core ASes to form core segments. The formation of inter-ISD beacons is identical to that of intra-ISD PCBs. The only difference is that every core AS periodically initiates core beaconing by sending beacons to all its neighbor core ASes (not to its customers, as in the intra-ISD case). Also, in inter-ISD, the PCB from each core AS is flooded to all other core ASes.[27]

5.2.3 Path lookup

Path lookup is a crucial step of SCION's path management architecture. It enables end hosts to receive path segments found during path exploration and construct end-to-end paths from a set of possible path segments[27]. SCION's path lookup infrastructure is designed based on requirements such as:

- **Low latency** - path lookup must be performed before a packet can be sent to its destination.
- **Scalability** - path lookup has to scale with respect to the number of users and the number of paths available in an ever-expanding network.
- **Availability** - to avoid outages and attacks, path lookup infrastructure should be distributed and replicated to guarantee high availability.

Host-to-host communication is enabled by a combination of up to three path segments that form a forwarding path. The path lookup process provides a source host with several path segments and at least one set of connecting path segments. Connecting path segments are the segments that can be combined towards the destination by joining their corresponding endpoints[27].

A source end host initiates a path lookup by issuing a path request including the destination ISD and AS identifiers to a local path server. The local path server then forwards the request to one of the core path servers. If the destination is within the same ISD as the source, the path server knows the destination and returns a certain number of segments to the local path server. If the destination is in a different ISD than the source, the core path server requests the segments from a core path server in the destination ISD before returning them to the local path server.[27]

5.2.4 Path combination

Name resolution and Path lookup are followed by Path combination, in which the end host forms a forwarding path from the obtained path segments from a path server. The construction of the path can vary based on the destination.

The forwarding path is encoded in the SCION packet header making routing tables unnecessary for border routers. Information is encoded as a packet-carried forwarding state (PCFS) in the packet header. By inverting the forwarding path from the packet header, the destination can respond to the source.

5.3 SCION Security Analysis

5.3.1 Security goals

SCION has security goals defined for two actors, namely autonomous systems (ASes) and hosts. Few goals like resilience to failures, availability, and support of heterogeneous environment apply to the entire architecture.

SCION achieves a high level of availability and provides tools to protect against DoS attacks effectively. SCION's heterogeneous trust environment is realized through isolation and transparency. Global connectivity is achieved for packets that need to traverse ISD bounds even when the routing within an ISD is independent.[27]

Autonomous systems

The main goal of an AS is to assuredly and efficiently accomplish network connectivity to other entities connected to the Internet. ISPs must provide uninterrupted connectivity to their customers. An AS verifies routing information received and

ensures that malicious entities cannot modify the propagated information in the control plane. In the data plane, an AS will forward packets only on valid segments. A segment is valid if the hop fields used for forwarding in the AS correspond to a PCB in the control plane.[27]

End hosts

Security goals defined for an end host are various and complex addressed on different levels of the network stack. Some of the properties present at the network layer are mentioned for both control and data plane.

Control plane security goals for End hosts are:

- **Reachability**: Connectivity between hosts on the Internet situated at any location.
- **Path diversity**: Since disjoint paths improve availability, a diverse set of paths should be available to choose from.

Data plane security goals for End hosts are:

- **Truthful forwarding**: During packet forwarding, the path selected by the source is the route traversed.
- **Path transparency**: End hosts should understand the path taken by a packet.
- **Packet integrity**: Receiving hosts verify a packet that includes its path, the same as the one sent by the source.
- **Source authentication**: The receiving host authenticates the origin of a packet.
- **Weak and robust detectability**: An on-path attacker cannot completely disguise his presence on the path, even with changing the path information in the packet's header.

5.3.2 Defence against attacks

Distributed denial-of-service attacks

ISPs that do not enable protection against source address spoofing eventually become a target for DoS attacks. The ISPs do not provide a quick switch-over to a

better path to the victims in an attack. Revenue models hinder ISPs from addressing DoS attacks contributing to the increasing occurrences of DoS attacks on the network infrastructure. DoS attacks are so common that they have boomed an industry of content distribution networks (CDNs) and cloud-based DoS mitigation systems to reduce the impact of DoS attacks. Even with these systems in place, a network can become unavailable when an adversary generates enough traffic to cause congestion in the network.[27]

SCION provides defense against DoS by enabling inter-domain traffic management and resource allocation. The five core methods are:

- path announcements with a short lifetime.
- non-registered (or hidden) paths.
- multipath communication.
- source authentication using OPT extension.
- COLIBRI extension for bandwidth allocation.

These methods guarantee communication for two communicating entities and can mitigate the network-level congestion even for the case of public services.

Packet replication attacks

In packet replication, a large number of packets can be used to bypass simple firewalls or intrusion detection systems consuming both the bandwidth and the computing power. In turn, this can be utilized to launch a DoS attack in the network. To prevent such attacks, SCION has an in-network replay detection system.

Resource exhaustion

Resource exhaustion can occur due to poor design decisions that no longer meet the requirements of scalability. This can include exhaustion of the IP space, BGP misconfiguration errors, or DoS attacks against unprotected services. To bypass resource exhaustion, SCION does not keep state in a performance-crucial infrastructure. For example, SCION routers do not hold any forwarding state.

5.4 SCION Deployments

5.4.1 ISP deployment

An ISP wishing to deploy SCION needs to connect with other SCION ISPs. This can happen either over an existing network link or through a dedicated link. The ISP also must obtain a certificate for their ASes. SCION deployment, in turn, requires the set up of border routers and services. Figure 16 depicts the three deployment scenarios SCION supports.

- **Minimal deployment**: can have a single SCION border router and services all deployed at a single location or even a single host. Since it places more reliance on the existing network infrastructure, fewer guarantees are accomplished. This is because an adversary might overload the legacy network with traffic, obstructing the links that SCION uses.
- **Intermediate deployment**: ISP would deploy many SCION border routers adjacent to existing legacy border routers, along with multiple servers distributed over their network to achieve tolerance to failures.
- **Ideal deployment**: SCION border routers would be directly connected to the neighboring ISPs' SCION border routers.

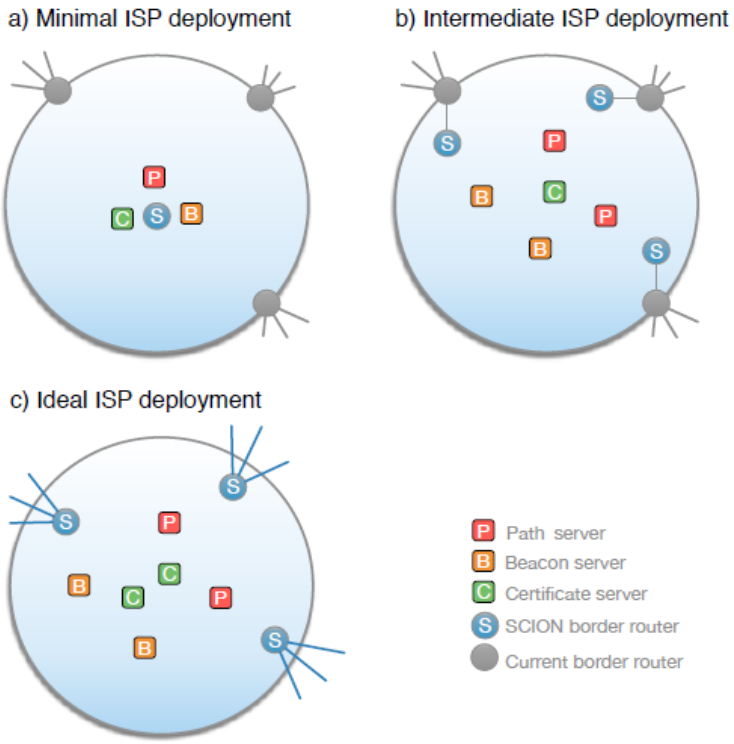

Figure 16: ISP Deployment scenarios.

5.4.2 End-domain deployment

End-host with SCION support

SCION native support is available through Ubuntu 16.04 OS for end hosts. SCION communication is achieved via SCION's built-in UDP, TCP, and SSP protocols.

End-host without SCION support

SCION proposes two methods in which the end hosts need not be upgraded; the HTTP(S) proxies and VPN tunneling.

HTTP(S) forward and reverse proxy

Figure 17: Deployment through SCION proxy.

The SCION HTTP(S) Proxy shown in Figure 17 enables a legacy host to browse the web over the SCION network infrastructure. It consists of two parts: Forward (Bridge) Proxy and Reverse Proxy. HTTP(S) requests are received from a standard web browser running on an end host in the forwarding proxy. These requests are communicated to the reverse proxy through the multipath socket. The reverse proxy retrieves the HTTP(S) requests from the website on the Internet or a SCION web server running on the same machine.

VPN-based deployment

Figure 19 depicts the VPN-based deployment method, which offers a gateway to route VPN traffic between two VPN servers over the SCION network. The gateway encapsulates a VPN's UDP packets and sends them over a SCION multipath-UDP (MPUDP) connection. At the other end, these packets are decapsulated back into UDP packets. In this method, the VPN server need not be changed, but the VPN connection can still benefit from SCION's higher availability and the dynamic route optimization of the multipath socket.[27]

Figure 18: Deployment through SCION VPN.

5.4.3 The SCION-IP-Gateway (SIG)

A common approach to deploying the End-domain in SCION, makes use of the SCION-IP gateway (SIG). It enables SCION to interoperate with the legacy IP end hosts benefitting them with SCION deployment by transparently obtaining improved security and availability properties.[27]

SIG requirements:

- **IP-in-SCION Encapsulation**: Carrying legacy IP traffic over a SCION network involves encapsulating the IP traffic in SCION packets. The encapsulation protocol should be specifically non-reliable to avoid problems with stacking retransmission timers.[27]
- **Routing and Connectivity**: Interoperability demands that legacy IP connectivity is transparently maintained (i.e., communicating hosts should not be aware that SCION is involved, nor should their connectivity be impacted). Traffic routing must be fully supported between two legacy IP hosts—one in a legacy (i.e., non-SCION) AS and one in a SCION AS. The same applies to traffic exchanged between two legacy IP hosts that both reside in SCION ASes.[27]
- **Addressing**: Legacy hosts do not support SCION's name resolution service (RAINS) and still rely on DNS. The latter does not provide routing information to the legacy host, as SCION information is not mentioned in DNS. As a result, interoperability requires that bare IP addresses be sufficient for hosts' legacy addressing in SCION ASes.
- **Support for Layer-4 Protocols**: Current Internet massively uses TCP and UDP along with various layers-4 protocols such as SCTP, L2TPv3, IPIP, and ICMP2. Therefore, the interoperability solution for SCION must be layer-4 agnostic.

Interoperability between IP and SCION

SCION ASes that need to facilitate legacy IP connectivity between its legacy hosts and those in other ASes deploy a SIG service. The SIG service is accountable for providing interoperability between SCION and the legacy IP. It is responsible for routing and encapsulation of IP traffic and handles all traffic between SCION

ASes. The sending side encapsulates the traffic, and the receiving side decapsulates it back to IP packets. All IP traffic into or out of an AS goes through the SIG service governed by IP routing rules.[27]

- **Routing**: SCION ASes with a direct connection to the Internet have their outgoing IP traffic sent via the SIG service, which is set as the default gateway. In the case of the incoming IP traffic, the AS advertises its local IP allocations via its IP border routers. SCION ASes without direct connection to the Internet use another SCION AS and, in turn, its SIG service for the connectivity.
- **Mapping legacy IP addresses to SCION ASes**: SIG service receiving an IP packet needs to determine the SCION AS to which the destination IP belongs. This mapping from public IP to SCION AS must be verifiable to prevent an AS from claiming IP space. This verifiability is performed by each SCION AS by exporting an IP allocation config (IAC) through its certificate service. The IAC contains a list of IP allocations the AS owns, together with the public key of the AS.[27]
- **Encapsulation**: The SIG encapsulation protocol is built on top of UDP/SCION by converting IP packets into a byte stream to the remote SIG service. The stream contains the original layer-3 and above contents of the encapsulated IP packet(s). The SIG service communicates with a single stream per remote SCION AS. Each SIG payload has a SIG header: a 4-byte sequence number, a 2-byte index field, and two unused bytes. The sequence number is utilized by a receiving SIG to detect packet reordering and loss. The index field is utilized by the receiver to resynchronize in the event of packet loss.[27]
- **SIG negotiation**: A SIG service that needs to send encapsulated traffic to another AS must determine which address and port to send the traffic to. Therefore, the sending SIG service must send a query to the remote SIG, and the remote instance would respond with its own address, SCION control port, and dedicated encapsulation port. The query is sent every 500ms, and if no response is received for two consecutive queries, the sending SIG service will use the remote SIG address again and start sending packets to the new address it receives in response. This allows failover in case the remote SIG instance becomes unreachable.[27]
- **Protocol mismatch**: When a client attempts to connect with a protocol the destination does not support, errors are generated. These errors must be

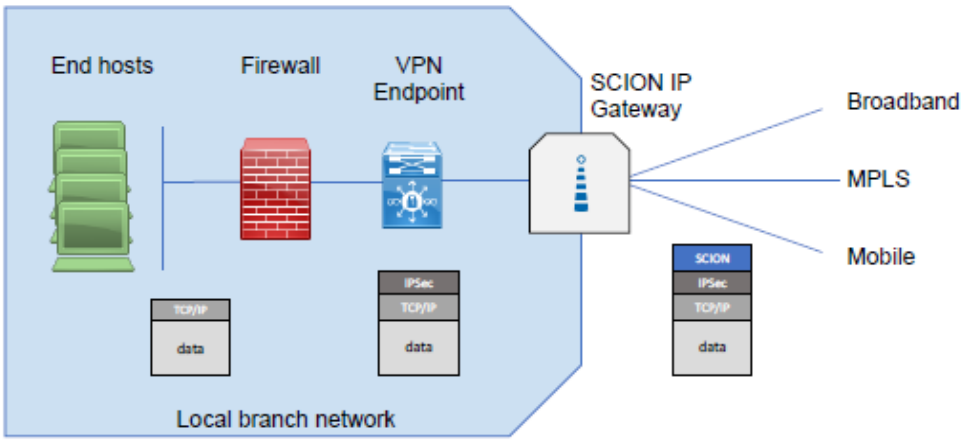

Figure 19: SIG Encapsulation.

handled distinctly, and the client needs to be informed rightly. Following cases are handled by SIG:

- **IP client to SCION service**: destination host generates an error based on the protocol used, and the reply is sent back to the source.
- **SCION client to IP service on IP host**: destination host will create an ICMP protocol unreachable reply and routes it to the client.
- **SCION client to IP service on SCION host**: destination host makes an SCMP port unreachable error and routes it back to the client as normal SCION traffic.

6 CES Signaling over SCION-IP-Gateway

6.1 Motivation

In the previous sections, the vulnerabilities present in the current Internet were presented and their remedies discussed in detail. Also, the concept of SCION, its capability in mitigating a wide range of IP vulnerabilities by design, were addressed. Table 4 presents the list of possible vulnerabilities in both IP and SCION.

Attacks	Routed IP	SCION
Source address authentication	✔	✗
Packet manipulation attacks	✔	✗
Man-in-the-middle attacks	✔	✗
Link DDoS	✔	✗
Address spoofing	✔	✗
Network layer DDoS	✔	✗
Prefix hijacking	✔	✗
Bandwidth exhaustion	✔	✗
Layer 3 DDoS	✔	✗
Outages due to unavailability	✔	✗
Packet replication attacks	✔	✔
Application-layer DoS attacks	✔	✔

Table 4: Possible attacks on IP and SCION.

It is important to note that Packet replication and Application layer DoS attacks are still possible on the SCION network. In packet replication attacks, the hacker replays the packets previously received by-passing simple firewalls or intrusion detection systems (IDS). Both bandwidth of the network and the computing power of the nodes are exhausted, when large numbers of packets are replayed which leads to DoS attacks. SCION's future version plans to integrate a high-speed in-network replay detection system [24] to prevent such replay attacks.

SCION does not provide any defensive mechanism for application-layer DoS attacks [25] and, the vulnerabilities found in a service can still be exploited. Application-layer attacks can be prevented by having an end-to-end or host-to-host based trust solution. Therefore, having an end system focused on trust solution over SCION would prevent many application-layer DoS attacks. CES, as seen earlier, runs on the principle of trust-based communication end-to-end. This solution which acts as a cooperative firewall can prevent application-layer attacks. Hence, the goal is to execute the CES solution over the SCION network, thereby achieving the combined benefits of the SCION network and CES's protection against the application-layer (DoS) attacks. Also, by adding new checks into H2H CETP, we could prevent also other application layer attacks.

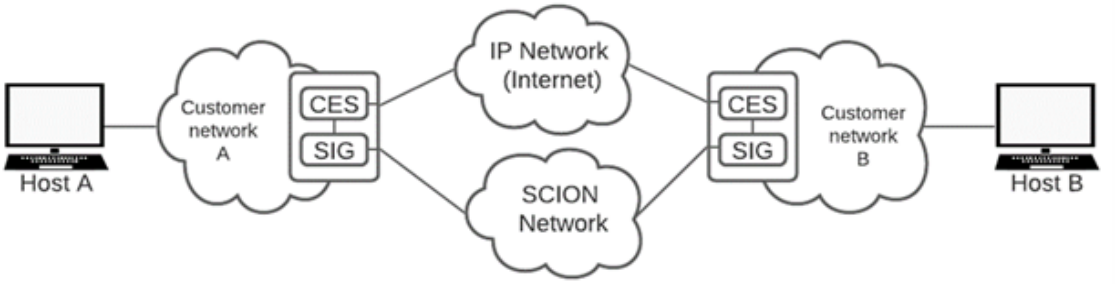

Figure 20: IP and SCION paths between two hosts.

SCION supports End domains to connect to the SCION network without any changes to the protocol stack or the hardware by providing a SCION-IP-Gateway (SIG). SIG acts as a tunnel between two ends of IP network. Both the originating and the receiving end must have the SIG configured and running. On the originating side, the SIG encapsulates the IP packet with the SCION packet header, while the SIG at the destination side strips the SCION packet header and forwards the IP packet to IP network. The traffic between the two SIGs traverses the SCION network comprising of SCION network infrastructure. Thus, CES is configured and executed alongside SCION's SIG. Furthermore, the CES code is updated to interact with SIG whenever available, effectively providing an additional path towards the destination that passes through the SCION network as shown in, Figure 20. The switch-over is performed when the originating CES has SIG configured and receives information of the remote SIG from the NAPTR response. If for any reason the switch-over fails, then the communication falls back to the routed IP.

6.2 Proposed solution and design

6.2.1 Solution

It is important to note that the traffic switched from routed IP to SCION will only be the Control plane/Signaling traffic exchanged between the two CES nodes. However, it is possible to switch the entire traffic from routed IP to SCION. The reason for selectively switching the signaling traffic is mentioned below:

- Signaling traffic is a critical transaction that involves the exchange and verification of certificates and policies from both ends. Naturally, protecting them from any attacks is a desired property.
- Once the CES-to-CES negotiation succeeds, the hosts use proxy addresses allocated by the CES for data plane communication. Thereby, eliminating the need to send them over SCION for added security.
- To showcase SIG as a filter that can be configured to selectively allow traffic of interest.

A byproduct of this design is that the open source SCION limitation in link bandwidth is alleviated for possible future experimentation. In the FUNET environment the data plane can easily use e.g. 10 Gbps links while the open source SCION links are limited e.g. to about 100Mbps. In addition, the option of data plane over IP allows to increase investments into SCION gradually assuming that the investor already has a powerful IP infrastructure. Hence, the solution presented in this thesis focuses on switching signaling traffic (TCP traffic) exchanged between the CES nodes, from routed IP to SCION network using SCION-IP-Gateway, whenever available.

6.2.2 Design

Both CES and SIG should be deployed under the same orchestration. For this purpose, the existing CES orchestration [13] that utilizes Linux containers to realize the network topology is updated to support SIG. The topology is shown in Figure 21. SCION AS that supports SIG is also deployed within the CES nodes. Deployment of SCION AS and setting up of SIG is discussed in the next section. Each node depicted in the topology is a Linux container running inside a Ubuntu 18 Virtual machine (VM). The specification of the VM is mentioned in, Table 5. Each container created inside the VM share the host's specification. There are 8 containers created each representing a node as shown in Figure 21.

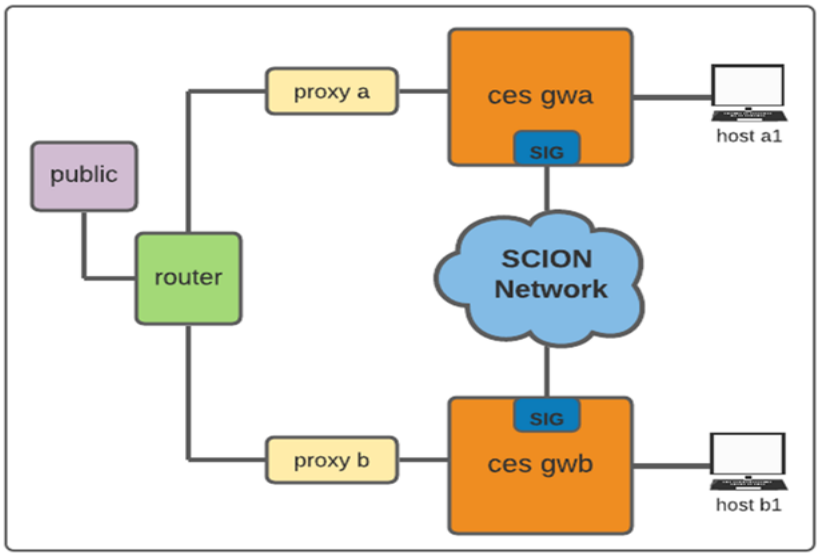

Figure 21: Proposed IP and SCION topology.

Operating System	Ubuntu 18 (64bit)
CPU	6
System memory	8GB
CPU MHz	2.5GHz

Table 5: Virtual machine specification.

6.3 Setup and configurations

6.3.1 CES

The setup configuration of CES orchestration is taken from Aalto 5G Github page [37]. The instructions provided are followed to create the required containers with necessary connections. CES code [38] is copied onto the deployed containers and executed. Note, at this stage the CES code does not have the proposed solution, but a simple CES-to-CES communication should work. After the CES orchestartion and the CES-to-CES communication succeeds, the SCION SIG setup can be configured as described in the next section.

6.3.2 SIG

In order to install SIG, a SCION AS must be running on the node. Hence, after deploying a SCION AS, SIG can be installed over it.

SCION AS Configuration

SCION AS can be created by using the service provided by the SCION Lab team named as *User AS*. A *User AS* can be created [35] by mentioning an attachment point connecting to an already existing SCION AS present in the SCION research network. Two SCION AS are required one for each CES node. Since the containers are created inside a single VM and the Internet access is blocked, the SCION AS must be connected to the attachment point via OpenVPN. The use of OpenVPN has impact on performance as the load on attachment point can make the SIG tunnels unstable [36]. The configuration file generated for the *User AS* with the above options can be deployed to the container.

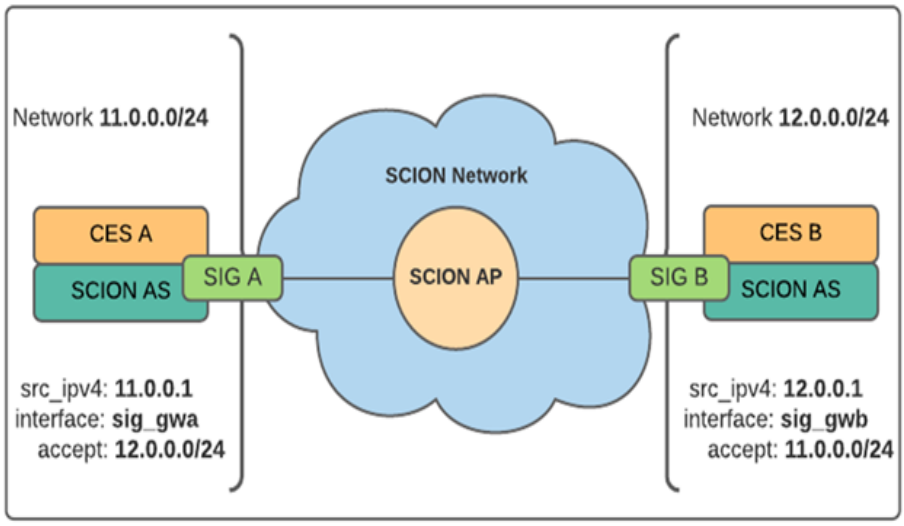

Figure 22: SCION's SIG Configurations.

SCION-IP-Gateway Configuration

Following the instructions for setting up the SIG [36], the configurations defined are shown in, Figure 22. For communication between the SIGs, any IP address can be assigned. Hence, two networks of the range 11.0.0.0/24 and 12.0.0.0/24 are defined for both SIGs, that are globally non-routable. Additionally, each SIG is configured with an interface name, source IPv4 address, and a rule to accept the traffic from the remote SIG network. Once all the configurations are loaded, a simple ping should succeed from 11.0.0.1 to 12.0.0.1. It is important to note that the *ip route* is added automatically by both SIGs to reach each other over the configured IP address. At the time of setting the above SIG, the SCION Attachment point (AP) was situated in Switzerland, as it was the available AP that was more stable compared to others. Hence, the traffic from SIG A would reach the SCION AP first and then back to

SIG B. This setup increases the RTT from SIG A to SIG B. However, selecting a nearby AP would significantly reduce the RTT.

6.3.3 Linux Container Topology

The Linux containers and their connections according to the proposed design are shown in Figure 23. The same set of IP addresses are used for the implementation, described in the next chapter.

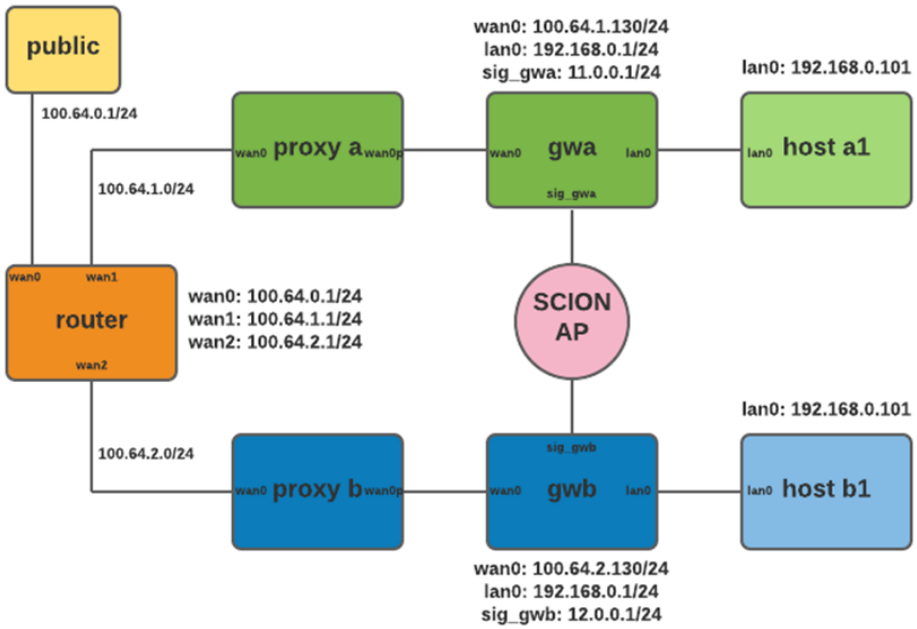

Figure 23: Linux Container Topology.

6.4 Implementation

The implementation of the proposed solution involves modifying the CES code base to recognize SIG and switch signaling traffic over to the SCION network. It is divided into three phases:

- **Proactive phase**: CES must receive host SIG IP from the configuration file. It must recognize SIG service running on the host and perform three actions namely: set MTU value on SIG interface, load the DNS NAPTR record with SIG-IP, and set an inbound rule to accept traffic from SIG.
- **Reactive phase**: Upon receiving a NAPTR response with SIG-IP, CES must check if a route exists to the remote SIG-IP. If a route exists, then CES must add an outbound rule pointing to the remote SIG-IP instead of remote CES-IP.
- **Monitor phase**: If the switch-over was successful, CES must monitor the traffic flowing over SIG and prompt CES at regular intervals.

Scenarios such as connection error, mismatch configurations and unavailability of service can occur at any point during the implemented phases. The standard response to these scenarios is to clear the previous configurations of the switch-over, and fall back to the default behaviour i.e. communication over routed IP.

6.4.1 Proactive Phase

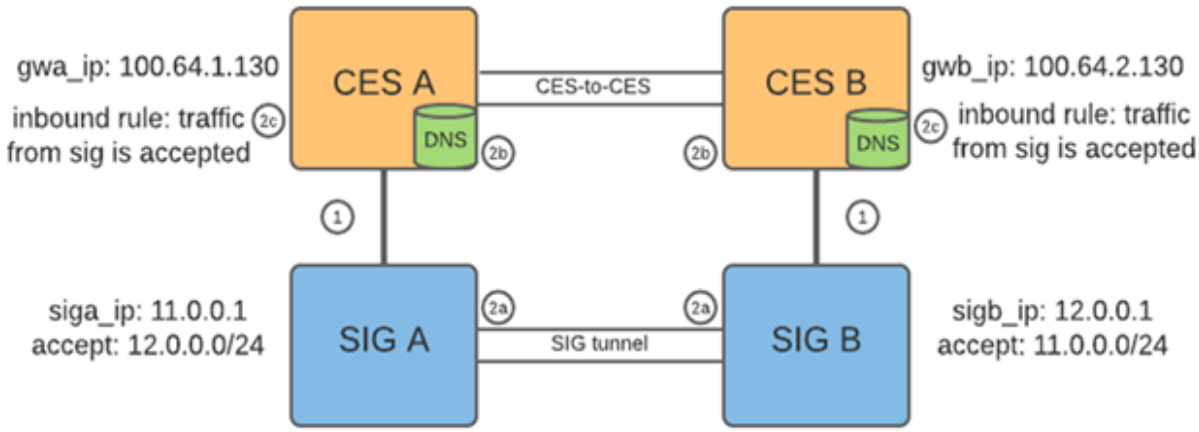

Figure 24: Proactive steps.

This phase is executed when CES starts and at regular intervals of time. Figure 24 shows the steps performed during this phase. For the explanation, only CES A and SIG A are considered, because the steps for the remote end are exactly the same.

- **Step 1**: When CES starts up or at a certain time interval, CES checks if the host SIG service is running on the IP address received from the config file. If the service is running, proceed to the next step else, trigger the cleaning process.
- **Step 2a**: Set MTU value for the SIG interface. Currently, the value being set is 1300. This step is needed as SIG cannot set a proper value [36].
- **Step 2b**: Load DNS NAPTR record with host SIG information (host SIG-IP). So that the remote end can know that the host CES node supports SIG switch-over.
- **Step 2c**: Add inbound traffic rule to accept traffic from SIG to CES. In the case of CES A and SIG A, an inbound rule is added to forward traffic from SIG A (11.0.0.1) to CES A (100.64.1.130).

Post performing the above steps, CES continues with its normal flow. Step 1 is repeated at regular intervals and, step 2 is performed only if previously not done. In cases such as SIG not configured on the host, SIG running on a different IP address, or SIG service abruptly going down, step 1 triggers a cleaning process where the following actions are performed:

- **Step 1**: Remove inbound and outbound rules added as part of the switch-over, if present.
- **Step 2**: Load DNS NAPTR record with host SIG-IP as 0.0.0.0. So that the remote end can know that the host CES node does not supports SIG switch-over.

Flow chart of the proactive phase

CES code base was modified to support the above mentioned steps. The same has been presnted as a flow chart in Figure 25.

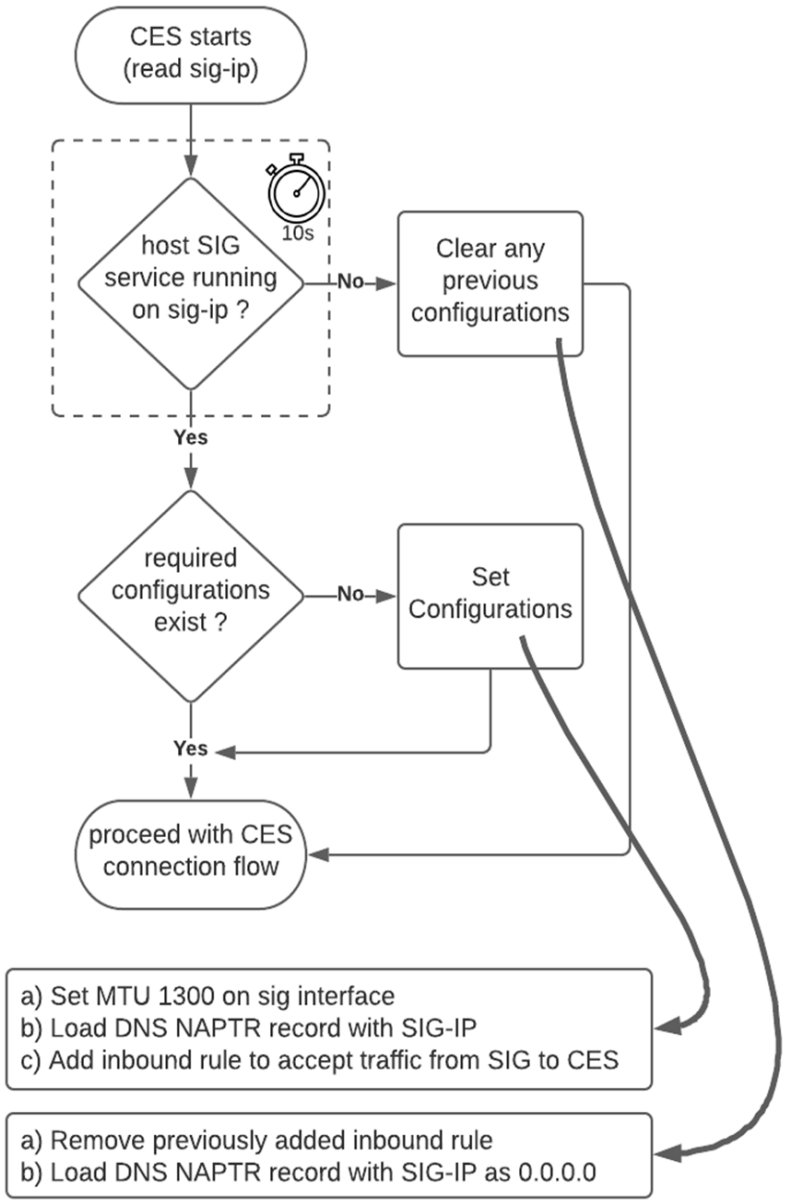

Figure 25: Proactive phase flow chart.

6.4.2 Reactive Phase

The reactive phase consists of steps performed by the originating CES when it receives a NAPTR response from the remote CES. Figure 26 shows the steps performed during this phase. For the explanation, CES A is the originating CES, receiving the NAPTR response from CES B.

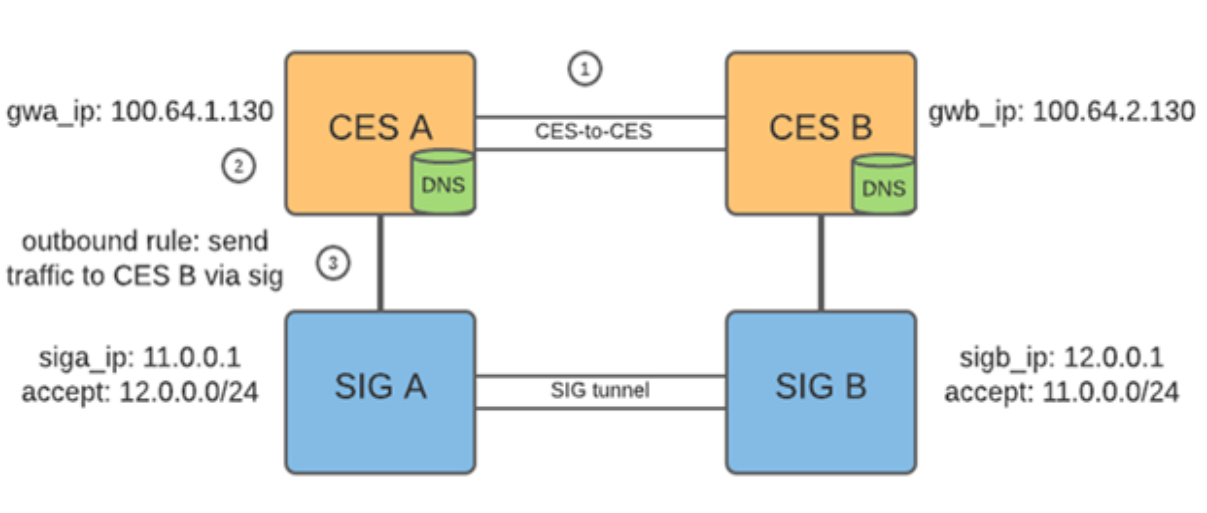

Figure 26: Reactive steps.

- **Step 1**: CES A receives the NAPTR response from the remote CES i.e. CES B. The response is parsed for a valid SIG-IP address (any IP address except 0.0.0.0).
- **Step 2**: CES A checks if the host has SIG configured and running. Additionally, CES A verifies if a route to the remote SIG exists, and proceeds to next step if successful.
- **Step 3**: Add outbound rule to send traffic to local SIG (SIG A) instead of CES B.

In case step 1 and step 2 fails for reason such as CES receives SIP IP as 0.0.0.0, host SIG not configured and no route to the remote SIG, step 3 is skipped altogether. Post step 3 the monitoring phase starts followed by normal CES flow.

Flow chart of the reactive phase

CES code base was modified to support the above mentioned steps. The same has been presnted as a flow chart in Figure 27.

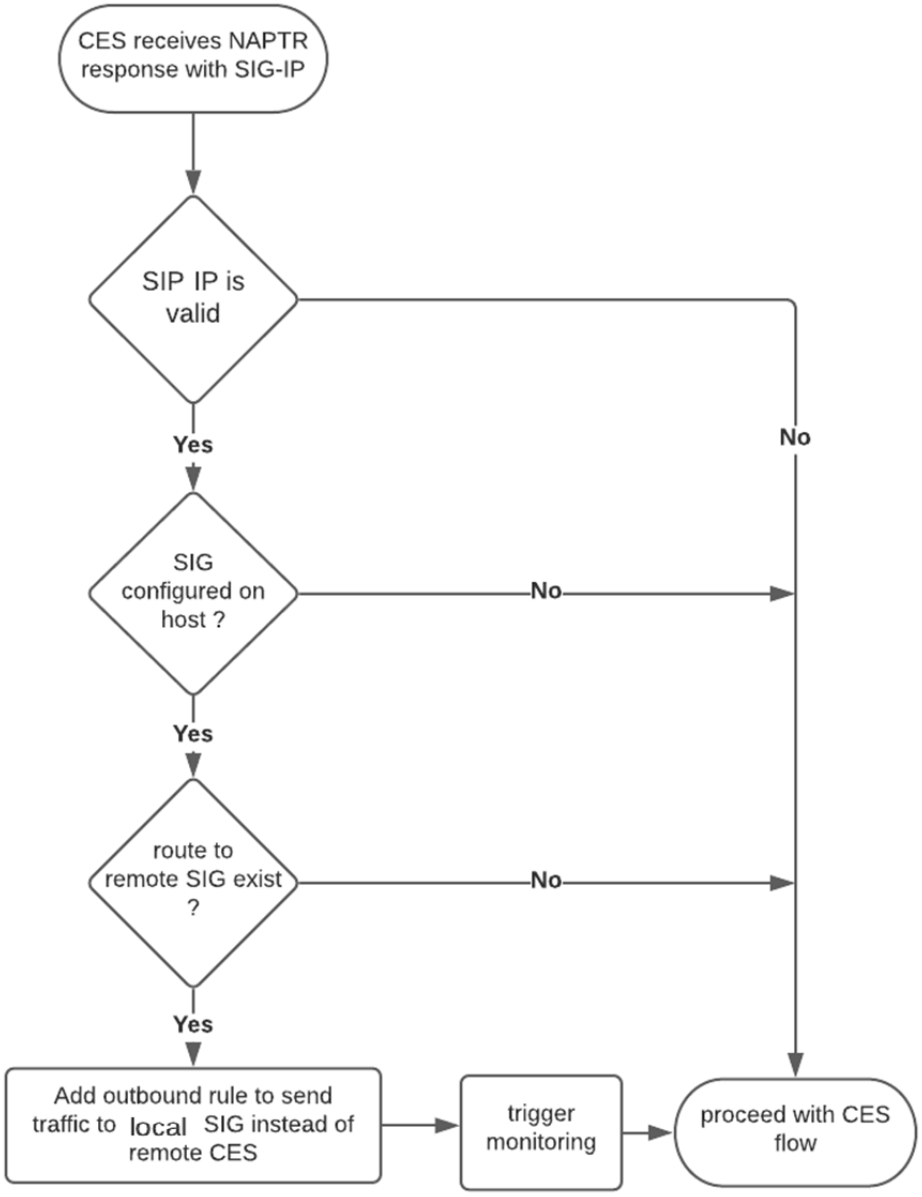

Figure 27: Reactive phase flow chart.

6.4.3 Monitoring Phase

The monitoring phase is started only after step 2 of the reactive phase is completed, and stays active as long as step 2 is valid. In this phase, the switch-over is monitored at a regular interval (10s), to prompt the CES that the traffic is being run over SCION.

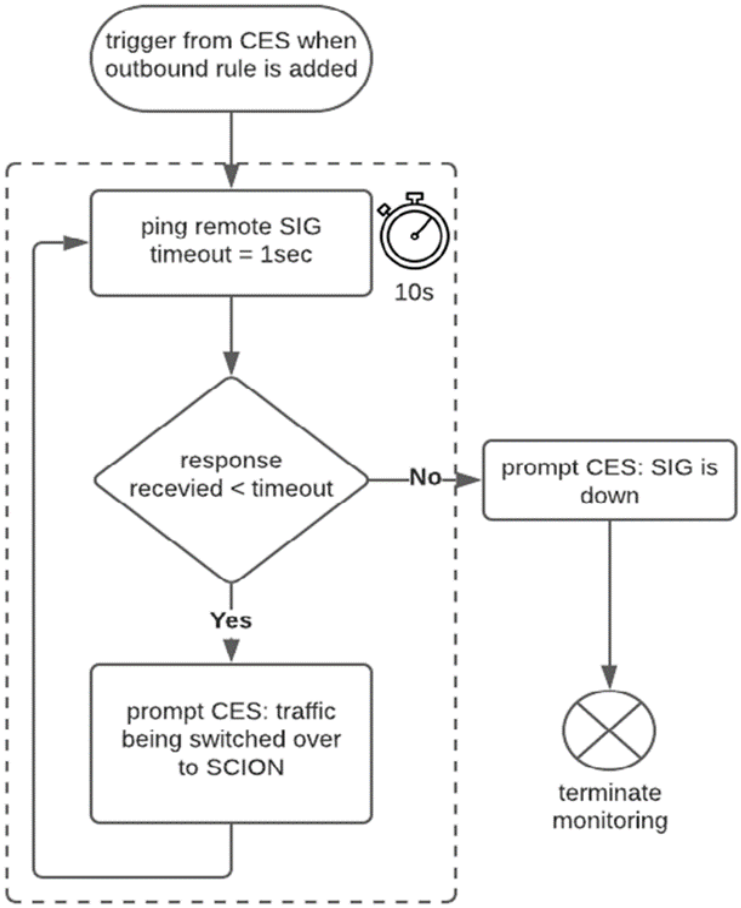

Figure 28: Monitoring phase flow chart.

Since the important point to monitor is the SIG-to-SIG tunnel, this is achieved by sending a ping from one SIG to another, with an acceptable timeout. If the pinging SIG receives the response within the timeout, then the tunnels are problem-free and the monitor can continue. Else, a prompt is sent to the CES conveying that SIG is down, effectively terminating the monitoring phase. Figure 28 shows the flow chart of the monitoring phase in case of a successful switch-over from IP to SCION. It is important to note that in the monitor phase, there is no cleaning/clearing up, as these are taken care of in the proactive phase.

6.4.4 CES Signaling over SCION's SIG - message flow

Figure 29 shows the complete sequence of message flow in a switch-over case, when hosta1 wants to communicate with hostb1.

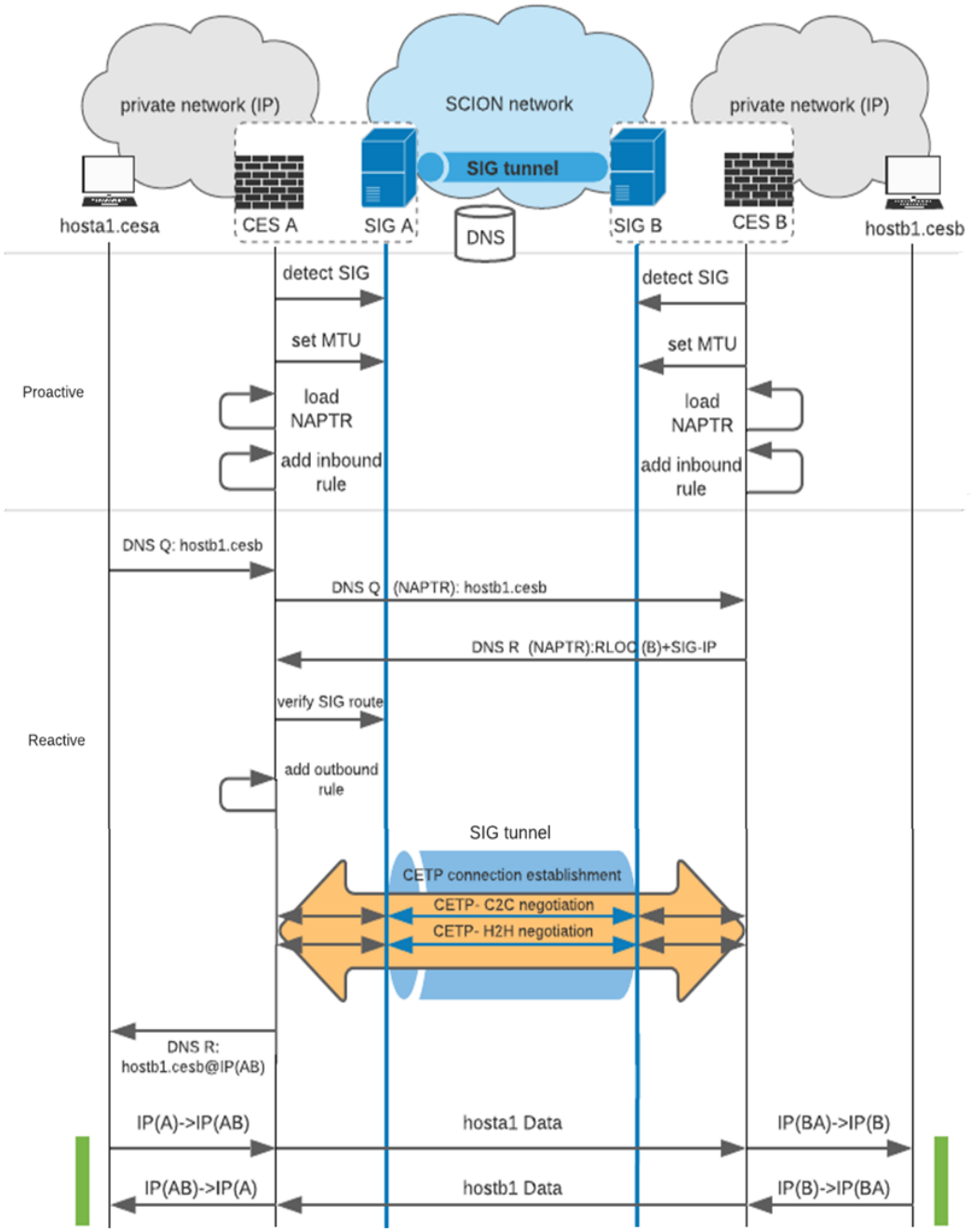

Figure 29: CES Signaling over SIG message flow.

7 Evaluation

7.1 Scenario based Design Verification

The implemented solution is expected to function smoothly and fall back to the default behavior in cases of discrepancies. Discrepancies can occur either from the host, remote end, or from the network.

In this section, some of the scenarios are simulated, and the behavior of the solution is verified. For all the scenarios described, *hosta* situated behind CES A (has SIG A) tries to communicate with *hostb* behind CES B (has SIG B). Hence, the evidence is collected from CES A, and can provide enough proof to verify the behavior.

7.1.1 Both SIGs are configured and reachable - (Sunny day scenario)

Case: SIG A and SIG B are configured and connected (Figure 30). *hosta* belonging to CES A, pings *hostb* connected to CES B.

Figure 30: SIG A and SIG B are configured and reachable.

Expected behavior: CES A translates the DNS query recevied from *hosta's* to NAPTR query and forwards it to CES B. CES B responds with NAPTR response including SIG B IP. CES A receives the response with SIG B IP and sends the signaling (TCP) traffic to SIG A, which inturn forwards the traffic to SIG B.
Evidence: Figure 31 shows the screenshot from CES A's console, which indicates the performed steps. The signaling traffic was successfully sent over SIG and the data traffic over routed IP. Solution works as expected.

```
SCION-SIG - INFO - ################################################################################
SCION-SIG - INFO - Received SIG IP is: 12.0.0.1
SCION-SIG - INFO - SCION-IP-Gateway is configured on Host. Verifying SIG configurations...
SCION-SIG - INFO - Verification Done.
SCION-SIG - INFO - Rule added to Switch Control Plane traffic from routed IP to SCIONs SCION-IP-GATEWAY..!!!
SCION-SIG - INFO - ################################################################################
```

Figure 31: CES A's console prompt when both SIGs are configured and reachable.

7.1.2 SIG A is NOT configured

Case: SIG A is not configured on host where CES A is configured, and SIG B is configured and running (Figure 32). *hosta* belonging to CES A, pings *hostb* connected to CES B.

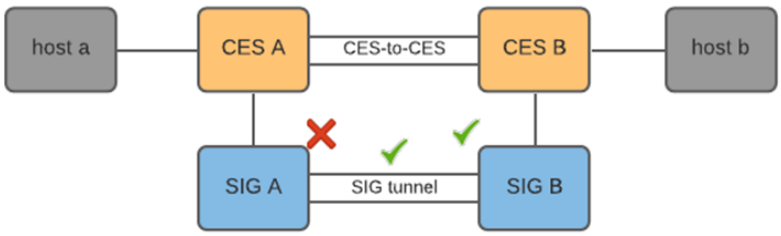

Figure 32: SIG A is not configured.

Expected behavior: CES A translates the DNS query received from *hosta's* to NAPTR query and forwards it to CES B. CES B responds with NAPTR response, including SIG B IP. CES A receives the response with SIG B IP and realizes SIG A is not configured, thereby skipping the SIG switch-over and proceeding with CES-to-CES flow.

Evidence: Figure 33 shows the screenshot from CES A's console, which indicates the performed steps. Signaling switch-over to SIG is skipped. Solution works as expected.

```
SCION-SIG - INFO - ################################################################################
SCION-SIG - INFO - Received SIG IP is: 12.0.0.1
SCION-SIG - INFO - SCION-IP-Gateway is NOT configured on Host. Ignoring SIG Step
SCION-SIG - INFO - ################################################################################
```

Figure 33: CES A's console prompt when SIG A is not configured.

7.1.3 SIG B is NOT configured

Case: SIG B is not configured on the remote end where CES B is configured, and SIG A is configured and running (Figure 34). *hosta* belonging to CES A, pings *hostb* connected to CES B.

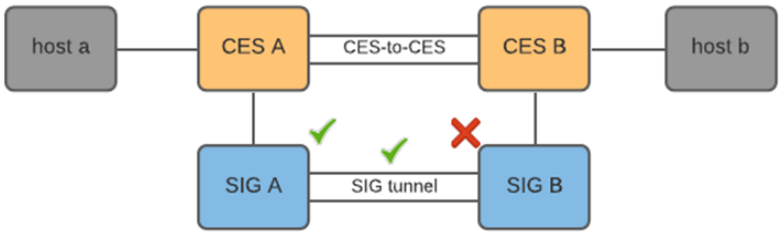

Figure 34: SIG B is not configured.

Expected behavior: CES A translates the DNS query received from *hosta's* to NAPTR query and forwards it to CES B. CES B responds with NAPTR response, including SIG B IP as 0.0.0.0. CES A receives the response with SIG B IP and realizes SIG B is not configured, thereby skipping the SIG switch-over and proceeding with CES-to-CES flow.
Evidence: Figure 34 shows the screenshot from CES A's console, which indicates the performed steps. Signaling switch-over to SIG is skipped. Solution works as expected.

```
SCION-SIG - INFO - ################################################################################
SCION-SIG - INFO - Received SIG IP is: 0.0.0.0
SCION-SIG - INFO - SCION-IP-GATEWAY is Skipped.
SCION-SIG - INFO - ################################################################################
```

Figure 35: CES A's console prompt when SIG B is not configured.

7.1.4 Both SIGs are configured but unreachable

Case: SIG A and SIG B are configured, but unreachable (Figure 36). *hosta* belonging to CES A, pings *hostb* connected to CES B.
Expected behavior: CES A translates the DNS query received from *hosta's* to NAPTR query and forwards it to CES B. CES B responds with NAPTR response, including SIG B IP. CES A receives the response with SIG B IP and realizes no route exists between SIG A and SIG B, thereby skipping the SIG switch-over and proceeding with CES-to-CES flow.

Figure 36: SIG A and SIG B are configured but unreachable.

Evidence: Figure 34 shows the screenshot from CES A's console, which indicates the performed steps. Signaling switch-over to SIG is skipped. Solution works as expected.

```
SCION-SIG - INFO - ################################################################################
SCION-SIG - INFO - Received SIG IP is: 12.0.0.1
SCION-SIG - INFO - SCION-IP-Gateway is configured on Host. Verifying SIG configurations...
SCION-SIG - INFO - No route exist between the SCION-IP-GATEWAYS
SCION-SIG - INFO - Skipping SIG.
SCION-SIG - INFO - ################################################################################
```

Figure 37: CES A's console prompt when both SIGs are configured but unreachable.

7.2 Packet Visualization

Packet capture at different nodes provides evidence of the working solution. It is enough to capture the packets at CES A, *hosta1* and SIG A to observe the message flows. Packets are captured for both the CES-to-CES scenario and the CES-to-SIG (signaling) scenario to understand the difference between them.

7.2.1 CES-to-CES

In the CES-to-CES scenario, the packets are captured at *hosta1* and CES A. *hosta1* behind CES A tries to communicate (ping) to *hostb1* behind CES B.

hosta1: From the packet capture, shown in Figure 38, it is clear that*hosta1* sends a DNS query for *hostb1* to CES A using its private IP address. CES A replies with the IP address (proxy) of *hostb1*. Data packets (ping) from *hosta1* are sent to the received IP address (proxy).

Source	Destination	Protocol	Info
192.168.0.101	192.168.0.1	DNS	Standard query 0xfe91 A hostb1.gwb.demo
192.168.0.101	192.168.0.1	DNS	Standard query 0x0da7 AAAA hostb1.gwb.demo
192.168.0.1	192.168.0.101	DNS	Standard query response 0xfe91 A hostb1.gwb.demo A 172.16.1.100
192.168.0.1	192.168.0.101	DNS	Standard query response 0x0da7 AAAA hostb1.gwb.demo A 172.16.1.100
192.168.0.101	172.16.1.100	ICMP	Echo (ping) request id=0x05fc, seq=1/256, ttl=64 (reply in 6)
172.16.1.100	192.168.0.101	ICMP	Echo (ping) reply id=0x05fc, seq=1/256, ttl=62 (request in 5)
192.168.0.101	172.16.1.100	ICMP	Echo (ping) request id=0x05fc, seq=2/512, ttl=64 (reply in 8)
172.16.1.100	192.168.0.101	ICMP	Echo (ping) reply id=0x05fc, seq=2/512, ttl=62 (request in 7)

Figure 38: Packet capture at *hosta1*.

CES A: From the packet capture, shown in Figure 39, it is observed that CES A translates the DNS query from *hosta1* to NAPTR query and forwards it to CES B. Upon receiving the NAPTR response, CES A initiates the CETP connection establishment over TCP. After the CETP connection succeeds, DNS response is sent to *hosta1*, and the data packets coming from *hosta1* are forwarded to CES B.

Source	Destination	Protocol	Info
100.64.1.130	100.64.0.1	DNS	Standard query 0xd5b8 NAPTR hostb1.gwb.demo
100.64.0.1	100.64.1.130	DNS	Standard query response 0xd5b8 NAPTR hostb1.gwb.demo NAPTR 100
100.64.1.130	100.64.0.1	DNS	Standard query 0xa1f5 NAPTR hostb1.gwb.demo
100.64.0.1	100.64.1.130	DNS	Standard query response 0xa1f5 NAPTR hostb1.gwb.demo NAPTR 100
100.64.1.130	100.64.2.130	TCP	30760 → 49001 [SYN] Seq=0 Win=64240 Len=0 MSS=1460 SACK_PERM=1
100.64.2.130	100.64.1.130	TCP	49001 → 30760 [SYN, ACK] Seq=0 Ack=1 Win=0 Len=0 MSS=1460 SACK_
100.64.1.130	100.64.2.130	TCP	30760 → 49001 [ACK] Seq=1 Ack=1 Win=64256 Len=0 TSval=21187293
100.64.2.130	100.64.1.130	TCP	[TCP Window Update] 49001 → 30760 [ACK] Seq=1 Ack=1 Win=65152
100.64.1.130	100.64.2.130	TCP	22690 → 49003 [SYN] Seq=0 Win=64240 Len=0 MSS=1460 SACK_PERM=1
100.64.2.130	100.64.1.130	TCP	49003 → 22690 [SYN, ACK] Seq=0 Ack=1 Win=0 Len=0 MSS=1460 SACK_
100.64.1.130	100.64.2.130	TCP	22690 → 49003 [ACK] Seq=1 Ack=1 Win=64256 Len=0 TSval=21187293
100.64.2.130	100.64.1.130	TCP	[TCP Window Update] 49003 → 22690 [ACK] Seq=1 Ack=1 Win=65152
100.64.1.130	100.64.2.130	TCP	23834 → 49002 [SYN] Seq=0 Win=64240 Len=0 MSS=1460 SACK_PERM=1
100.64.2.130	100.64.1.130	TCP	49002 → 23834 [SYN, ACK] Seq=0 Ack=1 Win=0 Len=0 MSS=1460 SACK_
100.64.1.130	100.64.2.130	TCP	23834 → 49002 [ACK] Seq=1 Ack=1 Win=64256 Len=0 TSval=21187293
100.64.2.130	100.64.1.130	TCP	[TCP Window Update] 49002 → 23834 [ACK] Seq=1 Ack=1 Win=65152
100.64.1.130	100.64.2.130	TLSv1.3	Client Hello
.... TCP packets			
100.64.2.130	100.64.1.130	TLSv1.3	Application Data
100.64.1.130	100.64.2.130	TLSv1.3	Application Data
100.64.2.130	100.64.1.130	TLSv1.3	Application Data
100.64.1.130	100.64.2.130	TCP	30760 → 49001 [ACK] Seq=5924 Ack=6205 Win=64128 Len=0 TSval=211872
40.10.181.107	20.229.20.145	ICMP	Echo (ping) request id=0x0612, seq=1/256, ttl=63 (reply in 64)
20.229.20.145	40.10.181.107	ICMP	Echo (ping) reply id=0x0612, seq=1/256, ttl=63 (request in 63)
100.64.1.130	100.64.2.130	TCP	22690 → 49003 [ACK] Seq=3176 Ack=4747 Win=64128 Len=0 TSval=211872
40.10.181.107	20.229.20.145	ICMP	Echo (ping) request id=0x0612, seq=2/512, ttl=63 (reply in 67)
20.229.20.145	40.10.181.107	ICMP	Echo (ping) reply id=0x0612, seq=2/512, ttl=63 (request in 66)

Figure 39: Packet capture at CES A.

7.2.2 CES over SCION

In the case of CES signaling over SIG, the packet capture is done at three nodes: *hosta1*, CES A, and SIG A. *hosta1* behind CES A tries to communicate (ping) to *hostb1* behind CES B.

hosta1: From the packet capture, shown in Figure 40, it is clear that *hosta1* sends a DNS query for *hostb1* to CES A using its private IP address. CES A replies with the IP address (proxy) of *hostb1*. Data packets (ping) from *hosta1* are sent to the received IP address (proxy). There is no change from the end-host perspective, when the switch-over occurs.

Source	Destination	Protocol	Info
192.168.0.101	192.168.0.1	DNS	Standard query 0xfe91 A hostb1.gwb.demo
192.168.0.101	192.168.0.1	DNS	Standard query 0x0da7 AAAA hostb1.gwb.demo
192.168.0.1	192.168.0.101	DNS	Standard query response 0xfe91 A hostb1.gwb.demo A 172.16.1.100
192.168.0.1	192.168.0.101	DNS	Standard query response 0x0da7 AAAA hostb1.gwb.demo A 172.16.1.100
192.168.0.101	172.16.1.100	ICMP	Echo (ping) request id=0x05fc, seq=1/256, ttl=64 (reply in 6)
172.16.1.100	192.168.0.101	ICMP	Echo (ping) reply id=0x05fc, seq=1/256, ttl=62 (request in 5)
192.168.0.101	172.16.1.100	ICMP	Echo (ping) request id=0x05fc, seq=2/512, ttl=64 (reply in 8)
172.16.1.100	192.168.0.101	ICMP	Echo (ping) reply id=0x05fc, seq=2/512, ttl=62 (request in 7)

Figure 40: Packet capture at *hosta1*.

CES A: From the packet capture, shown in Figure 41, it is observed that CES A translates the DNS query from *hosta1* to NAPTR query and forwards it to CES B. Upon receiving the NAPTR response, CES A initiates the CETP connection establishment over TCP via SIG (shown in the next packet capture). After the CETP connection succeeds, DNS response is sent to *hosta1*, and the data packets coming from *hosta1* are forwarded to CES B. It is important to note that the CETP messages are completely missed in the packet capture as they are sent over SCION's SIG.

Source	Destination	Protocol	Info
100.64.1.130	100.64.0.1	DNS	Standard query 0xee77 NAPTR hostb1.gwb.demo
100.64.1.130	100.64.0.1	DNS	Standard query 0x0f36 NAPTR hostb1.gwb.demo
100.64.0.1	100.64.1.130	DNS	Standard query response 0x0f36 NAPTR hostb1.gwb.demo
100.64.0.1	100.64.1.130	DNS	Standard query response 0xee77 NAPTR hostb1.gwb.demo
150.221.76.201	2.114.82.182	ICMP	Echo (ping) request id=0x058b, seq=1/256, ttl=63 (re
2.114.82.182	150.221.76.…	ICMP	Echo (ping) reply id=0x058b, seq=1/256, ttl=63 (re
150.221.76.201	2.114.82.182	ICMP	Echo (ping) request id=0x058b, seq=2/512, ttl=63 (re
2.114.82.182	150.221.76.…	ICMP	Echo (ping) reply id=0x058b, seq=2/512, ttl=63 (re

Figure 41: Packet capture at CES A.

SIG A: Figure 42 shows the packet capture at SIG A. The CETP connection messages over TCP are now sent through the SCION's SIG. It is important to note that the source and the destination addresses of the TCP packets use SIG's IP address.

Source	Destination	Protocol	Info
11.0.0.1	12.0.0.1	TCP	55188 → 49001 [SYN] Seq=0 Win=64240 Len=0 MSS=1460 SACK_PERM=1 TSva
11.0.0.1	12.0.0.1	TCP	19774 → 49003 [SYN] Seq=0 Win=64240 Len=0 MSS=1460 SACK_PERM=1 TSva
11.0.0.1	12.0.0.1	TCP	22348 → 49002 [SYN] Seq=0 Win=64240 Len=0 MSS=1460 SACK_PERM=1 TSva
12.0.0.1	11.0.0.1	TCP	49001 → 55188 [SYN, ACK] Seq=0 Ack=1 Win=64896 Len=0 MSS=1260 SACK_
11.0.0.1	12.0.0.1	TCP	55188 → 49001 [ACK] Seq=1 Ack=1 Win=64256 Len=0 TSval=630219423 TSe
12.0.0.1	11.0.0.1	TCP	49003 → 19774 [SYN, ACK] Seq=0 Ack=1 Win=64896 Len=0 MSS=1260 SACK_
11.0.0.1	12.0.0.1	TCP	19774 → 49003 [ACK] Seq=1 Ack=1 Win=64256 Len=0 TSval=630219423 TSe
12.0.0.1	11.0.0.1	TCP	49002 → 22348 [SYN, ACK] Seq=0 Ack=1 Win=64896 Len=0 MSS=1260 SACK_
11.0.0.1	12.0.0.1	TCP	22348 → 49002 [ACK] Seq=1 Ack=1 Win=64256 Len=0 TSval=630219423 TSe
11.0.0.1	12.0.0.1	TLSv1.2	Client Hello
11.0.0.1	12.0.0.1	TLSv1.2	Client Hello
11.0.0.1	12.0.0.1	TLSv1.2	Client Hello
12.0.0.1	11.0.0.1	TCP	49001 → 55188 [ACK] Seq=1 Ack=518 Win=64384 Len=0 TSval=3698352536
12.0.0.1	11.0.0.1	TCP	49003 → 19774 [ACK] Seq=1 Ack=518 Win=64384 Len=0 TSval=3698352536
12.0.0.1	11.0.0.1	TCP	49002 → 22348 [ACK] Seq=1 Ack=518 Win=64384 Len=0 TSval=3698352536

Figure 42: Packet capture at SIG A.

7.3 CETP Connection Establishment

7.3.1 CETP connection delay

The CETP connection establishment occurs between two CES nodes. Hence, the hosts connected to them are unaware of whether the traffic is routed over IP or SCION, and calculating Round Trip Time (RTT) between them does not provide any valuable information. The switch-over takes place before the data packets are sent out. Hence, the time taken to perform this switch-over is of more interest.

To measure the delay introduced by the CES-to-SCION switch-over, the time difference between the DNS query and the DNS response of the host is calculated. As the complete flow of switch-over is performed within the DNS query and the response of the host, this provides a reliable way to find the delay. This delay is calculated for both the scenarios: CES-to-CES and CES Signaling over SIG. Table 6 shows the time difference between the DNS query and the DNS response on *hosta1*.

In the CES-to-CES case, the traffic travels within the container. Hence, the delay between the DNS query and the response is small. Whereas, in the case of switch-over, the traffic travels from one container to the SCION attachment point and back to the container. As part of the CETP negotiation there are three TCP connections initiated by the outbound CES node. A simple ping between the two SIGs produces a rtt of 110ms, and this would that the rtt for the TCP connection would be around 330ms. However, due to re-transmissions the rtt value doubles for each connection i.e. 600ms. This increase is caused by using OpenVPN for connecting SIG with its parent node and a lower MTU value set on SIG. The SCION AP in the solution above is located in Switzerland, thereby increasing the delay. Configuring an AP closer to the solution would reduce this delay. Also, the Switzerland AP is more powerful than the ASes that are ASes available in Finland that run on VMs. There have been a range of performance tests performed on these SCION ASes available in Finalnd presented in the report [39].

Case	**Delay**
CES-to-CES	0.04sec
CES Signaling over SIG	0.67sec

Table 6: Delay betwen DNS query and response on *hosta1*.

7.3.2 CETP policy

CETP connection establishment consists of two layers of message exchange: CETP-C2C layer and CETP-H2H. The CETP-C2C layer is concerned with negotiating network-related policies between the two CES nodes. During this negotiation, the two CES nodes exchange rules governing the data plane traffic. Also, this layer establishes trust between the two nodes.

SIG-to-SIG communication has all the benefits a SCION network provides. The use of non-routable IP addresses hides the source identity rendering the packets useless in case sniffed. It is not possible to spoof a SIG-to-SIG connection since they are bound to their SCION hosts. The SIG-to-SIG connection between two nodes confirms the authenticity of the hosts running them (Only if both CES and SIG are run on the same host). Considering the above points can lead to the optimization of the CETP policies:

Group	Code	Description	Relevant in SCION
CES	*cesid*	FQDN-based ID of the CES node	required
	caces	The CA address for CES validation	not required as SIG-to-SIG connection already establishes the trust between the hosts
	headersignature	Signature of the CETP packet	not required as SIG provides encapsulation
	ratelimit	The rate limit for session	required
Control	*dstep*	FQDN-based destination endpoint ID	required
	caep	The CA address for endpoint validation	not required as SIG-to-SIG connection already establishes the trust between the hosts
	terminate	Contains session terminating information	required
	ack	The acknowledgement number	required
	hard_ttl	the hard-time out for the session state	required
	idle_ttl	the idle-time out for the session state	required

Table 7: CETP policy elements and its relevance in SCION

- CES validation can be skipped when SIG is available. Due to this, CA parameters from the policies can be eliminated.
- SIG itself provides encapsulation, thereby making it possible to avoid encrypting CETP messages.
- CETP can include payload with the SIG option, providing the host an option to send data plane traffic over SIG. (Current implementation does not support this)

Table 7, shows the list of the CETP policy elements and its relevance when the communication is over SCION.

7.4 SIG Performance

Although the SIG encapsulates the IP packet with the SCION header and traverses the SCION network, it is interesting to know the performance metrics such as round-

trip-time and bandwidth of the SIG. The SIG appears to perform poorly (abnormally high rate of reordered and dropped packets and consequently low throughput with TCP) when running in a SCIONLab User AS with a provider link using OpenVPN [36]. However, performance tests were run over SIG, and the results tabulated.

7.4.1 Round-trip-time

In the RTT calculation, packets are sent from one node to another, and the response time is tabulated for varying numbers of packets. Table 8 shows the RTT tests run between SIG A and SIG B (11.0.0.1 to 12.0.0.1). Every packet from SIG A to SIG B or vice versa would pass through the SCION AP located at Switzerland. Therefore, for an average of 100ms RTT one way delay from the VM in Finland to the SCION AP would be 100/4 = 25ms.

Packet count	**Avg RTT**	**Loss %**	**Total time**
1	112.76ms	0	0ms
5	100.78ms	0	4s
10	113.48ms	0	9s
15	112.46ms	0	14s
20	110.72ms	0	19s
50	117.05ms	0	49s

Table 8: RTT measurements from SIG A to SIG B.

Throughout the tests, the SIG tunnel was stable and provided uniform results. The packets here travel from SIG A to SCION AP and back to SIG B.

7.4.2 Bandwidth

In the bandwidth calculation, both SIG A and SIG B are configured as server and client simultaneously. Bandwidth tests from 1Mbps upto 10Gbps were performed and the results are tabulated in Table 9. Achieved bandwidth is calculated for both

direction: Client to Server (C->S) and Server to Client (S->C).

Throughout the tests, the SIG tunnel was stable and provided uniform results. SIG performed well in terms of the bandwidth and no loss were observed.

Attempted BW(bps)	**Achieved BW(bps)**	
	C ->S	**S ->C**
1M	1M	0.99M
10M	9.83M	9.91M
50M	49M	49.5M
100M	98M	98.9M
200M	196M	198M
500M	490M	495M
1G	981M	990M
10G	9.92G	9.96G

Table 9: Bandwidth measurements of the SIG tunnel.

7.4.3 Data plane traffic

In the case of CES signaling over SIG, only the TCP traffic is switched from routed IP to SCION. This implementation can be updated to allow all the traffic to switch-over SCION's SIG. However, it is critical to understand that only the traffic(TCP) between the two CES nodes is switched from IP to SCION. This can include data traffic over TCP between the nodes. But, the TCP data traffic between the hosts (hosta1 and hostb1) will not switch from IP to SCION, and this is because each host is allocated a proxy address, and the data plane traffic is carried over any of the available options (vxlan, ipsec, gre, and geneve).

8 Conclusion

In this thesis, the main objective was to integrate SCION and CES, thus providing an option to the CES firewall to switch signaling traffic from routed IP to SCION, using SCION-IP-Gateway. This switch-over option would effectively allow CES to benefit from the features provided by the SCION, such as defense against DDoS, multipath communication, BGP free communication, and many more. SCION can also benefit from the CES firewall which provides defenses against application layer DoS attacks for the end-host domain, missing in SCION.

Customer Edge Switching (CES) is a firewall solution intended to replace the traditional NAT along with some extensions. CES works on the principle of trust-based communication by enforcing cooperative behavior between hosts within a network served by the CES node. While CES provides added security by providing features such as masking the source IP address with proxy addresses, application-layer packet filtering, host authenticity verification, it is still plagued by some of the common attacks present on the current Internet. BGP-route hijacking, man-in-the-middle attacks, and DDoS attacks are some examples to name.

SCION is proposed as a new Internet architecture, which provides defenses against some of the commonly seen attacks on the Internet by design. Along with the defenses, it also provides features such as high availability, transparency, scalability, and support for heterogeneous trust. Realizing the SCION network would require utilizing SCION network infrastructure and changes in the protocol stack. However, SCION provides a feature for the end-hosts in an IP network to connect to SCION using SCION-IP-Gateway (SIG). SIG encapsulates the IP packet with the SCION header and traverses the SCION network, and upon reaching the destination, the SCION header is stripped and forwarded to the host.

End-domains can benefit from the integration of CES and SCION, where CES provides host-level authenticity by cooperative behavior concept, and SCION can provide network-level security by design. Control plane traffic is the most critical transaction of the solution. Hence, the proposed solution focuses on switching only the control plane/Signaling traffic between the two CES nodes from routed IP to SCION.

The main objective of the thesis has multiple sub-objectives under it, and achiev-

ing each sub-objective would reach the final implementation. The sub-objectives are: ensuring CES and SCION can co-exist and function independently within the same orchestration, ensuring SIG connection between the two SCION AS, the interaction between CES and SIG, proactive and reactive phases of the switch-over, monitoring of the switch-over when started and fall back to default when the switch-over fails.

The sub-objectives are implemented in the said order, and the final implementation is achieved. The proactive, reactive, and monitoring phases of the implementation are crucial as they involve adding software code to the CES code base. In the proactive phase, the CES node identifies SIG running on the host and configures itself to accept traffic from them. In the reactive phase, upon receiving the NAPTR response with the remote SIG, CES verifies the route to the remote SIG and configures itself to send traffic over SIG.

Evaluation of the implemented solution includes aspects such as design verification, packet visualization of end-user and the switch-over from IP to SCION, delay calculation of CETP connection establishment and its optimization suggestions, and SIG performance. Design verification demonstrates the implemented design can handle the various scenarios that can occur during the switch-over. It also verifies if the solution can fall back to its default behavior whenever inconsistencies occur. Packet visualization provides evidence of the working solution at various nodes, especially the end-user. Evidence suggests the solution has no change from an end-user perspective. CETP connection establishment has an increased delay, and the reason for this is the SCION AP between the connected SIGs. Choosing a closer SCION AP and optimizing the CETP establishment can reduce the delay. Finally, performance-related tests are run against the connected SIGs, and the results look promising even though the SIG running over OpenVPN has drawbacks.

8.1 Future Work

Future work can include: providing an option to CES, allowing it to switch not only TCP traffic but any traffic of choice. CES can make use of the SIG-to-SIG trust and optimize the connection establishment procedures. In the solution implemented, only traffic between the CES nodes is switched over to SCION. However, this can be extended to the hosts as well. A host can mention in its policy to inform CES to use SIG for all traffic originating from that host. Typically, CES would be supporting SIG for end hosts.

The SCION ASes configured on the Linux containers use OpenVPN to connect to the parent AS, and this causes performance issues. OpenVPN can be eliminated if the SCION AS can have a publicly reachable IP address. Also, the SCION AP used in the solution was located in Switzerland, making the comparison between CES and SIG unfair. It would be interesting to compare CES and SCION in a fair environment.

The solution mentions the benefits provided by the SCION architecture, such as defense against attacks. Future work can include performing different types of attacks against the SIG and verifying its credibility.

References

[1] Santos, Jesús Llorente and Kantola, Raimo and Beijar, Nicklas and Leppäaho, Petri. Implementing NAT traversal with Private Realm Gateway *2013 IEEE International Conference on Communications, pages=3581-3586, 2013*

[2] B. Trammell, J. Smith and A. Perrig *Adding Path Awareness to the Internet Architecture, IEEE Internet Computing, vol. 22, no. 2, pp. 96-102, Mar./Apr. 2018*

[3] S. Deering and R. Hinden. Internet Protocol, Version6 (IPv6) Specification *RFC1883 (Proposed Standard) 1995. Obsoleted by RFC2460.*

[4] K. Egevang, and P. Francis. *The IP Network Address Translator (NAT) RFC1631, 1994.*

[5] J. Rosenberg, R. Mahy, and P. Matthews. *Session Traversal Utilities for NAT (STUN), RFC5389, 2008.*

[6] J. Rosenberg, R. Mahy, and P. Matthews. *Traversal Using Relays around NAT (TURN): Relay Extensions to Session Traversal Utilities for NAT (STUN), RFC5766, 2010.*

[7] J. Rosenberg. *Interactive Connectivity Establishment (ICE): A Protocol for Network Address Translator (NAT) Traversal for Offer/Answer Protocols, RFC5245, 2010.*

[8] R. Kantola. 2010. *Implementing Trust-to-Trust with Customer Edge Switching. in press for AMCA in connection with AINA.*

[9] E. K. Çetinkaya and J. P. G. Sterbenz *"A taxonomy of network challenges," 2013 9th International Conference on the Design of Reliable Communication Networks (DRCN), 2013, pp. 322-330.*

[10] Alex A. Stewart and Marta F. Antoszkiewicz *BGP Route Analysis and Management Systems, The University of Northern Iowa*

[11] Van Jacobson, Diana K. Smetters, James D. Thornton, Michael F. Plass, Nicholas H. Briggs, and Rebecca L. Braynard. *Networking named content. In Proceedings of the International Conference on Emerging Networking Experiments and Technologies, 2009. Page 13*

[12] Dipankar Raychaudhuri, Kiran Nagaraja, and Arun Venkataramani. *MobilityFirst: A robust and trustworthy mobility-centric architecture for the future Internet. ACM SIGMOBILE Mobile Computing and Communications Review, July 2012. Page 14.*

[13] Maryam Pahlevan *Signaling and Policy Enforcement for Co-operative Firewalls, Master Thesis. Department of Communication and Networking, Alto University, 2013.*

[14] Forouzan, Behrouz A., and Sophia Chung. Fegan. *Data Communications and Networking . 4. ed. Boston: McGraw-Hill, 2007. Print.*

[15] Llorente Santos, Jesús *Private realm gateway, Master Thesis. Department of Communication and Networking, Alto University, 2012.*

[16] Timothy Rooney *IP Address Management, John Wiley and Sons, Ltd, pages = 141-142, year 2010*

[17] Kurose, James F. and Ross, Keith W. *Computer Networking: A Top-Down Approach (6th Edition), 2012, Pearson*

[18] Stewart, James Michael. *Fundamentals of Network Security Firewalls and Vpns, 1st edition, Jones and Bartlett Publishers Incorporated, 2010.*

[19] Kantola, Raimo, Llorente Santos, Jesus, Beijar, Nicklas *Policy-based communications for 5G mobile with customer edge switching, Security and Communication Networks, Security Comm. Networks, 2016.*

[20] Muhammad Hassaan Bin Mohsin *Security Policy Management for a Cooperative Firewall, Master Thesis. Department of Communication and Networking, Alto University, 2018.*

[21] A. P. R. M. R. DAVID BARRERA, LAURENT CHUAT and P. SZALACHOWSKI *The scion internet architecture, vol. 60, 2017.*

[22] Stephanos Matsumoto, Raphael M. Reischuk, Pawel Szalachowski, Tiffany Hyun-Jin Kim, and Adrian Perrig *Authentication Challenges in a Global Environment. ACM Trans, 2017.*

[23] R. Hakimi, Y. M. Saputra and B. Nugraha *Case study analysis on BGP: Prefix hijacking and transit AS, 2016 10th International Conference on Telecommunication Systems Services and Applications (TSSA), 2016.*

[24] Taeho Lee, Christos Pappas, Adrian Perrig, Virgil Gligor, and Yih-Chun Hu. *The case for in-network replay suppression. In Proceedings of the ACM Asia Conference on Computer and Communications Security (AsiaCCS), April 2017.*

[25] Georgios Mantas, Natalia Stakhanova, Hugo Gonzalez, Hossein Hadian Jazi, and Ali A. Ghorbani. *Application-layer denial of service attacks: taxonomy and survey. International Journal of Information and Computer Security, 7(2-4):216–239, 2015.*

[26] The RIPE NCC has run out of IPv4 Addresses *Accessed on: AUG. 17, 2021. [Online]. Available: https://www.ripe.net/publications/news/about-ripe-ncc-and-ripe/the-ripe-ncc-has-run-out-of-ipv4-addresses*

[27] R. M. R. L. C. Adrian Perrig, Pawel Szalachowski *"SCION: A Secure Inter-net Architecture,", 2017 [Online]. Available: https://pszal.github.io/papers/SCION-book.pdf*

[28] SCALABILITY, CONTROL, AND ISOLATION ON NEXTGENERATION NETWORKS *Accessed on: AUG. 17, 2021. [Online]. Available: https://www.scion-architecture.net/*

[29] ITU Internet Usage Statistics *Accessed on: AUG. 21, 2021. [Online]. Available: https://www.itu.int/en/ITU-D/Statistics/Pages/stat/default.aspx*

[30] High-Availability Communication Infrastructure *Accessed on: AUG. 21, 2021. [Online]. Available: https://version.aalto.fi/gitlab/scion-research-comnet-aalto/documents/-/blob/master/SCION-Aalto-Sep-2020.pdf*

[31] E. Z. Network Security Group *Accessed on: AUG. 17, 2021. [Online]. Available: https://docs.scionlab.org/*

[32] TRAFICOM *Accessed on: NOV. 25, 2021. [Online]. Available: https://www.kyberturvallisuuskeskus.fi/en/tcp-implementations-vulnerable-denial-service*

[33] Prefix hijack example *Accessed on: NOV. 25, 2021. [Online]. Available: https://blog.apnic.net/2019/06/07/large-european-routing-leak-sends-traffic-through-china-telecom/*

[34] R. Kantola, "Principles of Customer Edge Traversal Protocol," 2012. *Accessed on: NOV. 25, 2021. [Online]. Available: http://www.re2ee.org/*

[35] Create SCION User AS *Accessed on: NOV. 25, 2021. [Online]. Available: https://www.scionlab.org/*

[36] SCION-IP-Gateway *Accessed on: NOV. 25, 2021. [Online]. Available: https://docs.scionlab.org/content/apps/remote_sig.html*

[37] CES orchestration *Accessed on: NOV. 25, 2021. [Online]. Available: https://github.com/Aalto5G/CustomerEdgeSwitching/tree/master/orchestration/lxc*

[38] CES Code *Accessed on: NOV. 25, 2021. [Online]. Available: https://github.com/Aalto5G/CustomerEdgeSwitching*

[39] Performance Comparison of SCION with Routed IP on Virtual Machines *Accessed on: NOV. 25, 2021. [Online]. Available: http://www.re2ee.org/*