An Empirical Study into Factors that Create Configuration Inconsistencies between IPv4 and IPv6 Systems

Identifying the Managerial and Scientific Implications

D. Rieffe

Management of Technology
TU Delft

Cyber Security & Forensic Technology
PwC Nederland

An Empirical Study into Factors that Create Configuration Inconsistencies between IPv4 and IPv6 Systems

Identifying the Managerial and Scientific Implications

by

D. Rieffe

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Tuesday September 24, 2019 at 12:30.

Student number:	4654862	
Project duration:	Februari 1, 2019 – September 1, 2019	
Thesis committee:	Prof. dr. ir. M.F.W.H.A. Janssen	*TU Delft - ICT and Governance*, Chairman
	Dr.-Ing. T. Fiebig	*TU Delft - ICT and Governance*, Supervisor
	Prof. dr. ing. A.J Klievink	*TU Delft - Organization and Governance*, Critical Observer
	D. Switzer	*PwC Netherlands*, External Member

An electronic version of this thesis is available at `http://repository.tudelft.nl/`.

TU Delft Delft University of Technology

pwc

Executive Summery

This study performed a qualitative analysis of the rising problem of IPv6 misconfiguration. From scans of the public IP ranges of the internet is visible that currently 4,8% of all dual-stack hosts have a different configuration on IPv4 than IPv6. However, from an IT security perspective, the reasons for a difference in IP configuration are limited. In this thesis, the IP configuration process is studied. Furthermore, we reached out to the owners of identified misconfigured hosts and discussed with them the reason behind the misconfiguration.

This goal of this study is to discovering why there are differences in port configuration between IPv4 and IPv6. This thesis aims to increase the literature, knowledge, and awareness levels of IPv6 port configuration. For this, the following research question is drawn:

What is the reason for the existence of inconsistencies in open port policies on dual-stack hosts?

For answering the main research question, four corresponding sub-questions are drafted. If answered they will aid in answering the main research question.

SQ1: *What actors play a part in the IPv6 configuration?*
SQ2: *Which guidelines are considered to be most critical for port-based firewalling policies?*
SQ3: *What processes lead to IPv4/IPv6 misconfigurations and inconsistencies that can be found in the wild?*
SQ4: *Is there evidence of active exploitation of IPv6 misconfigurations?*

Methodology

Most literature based on IPv6 port configuration has a technical approach to the problem. However, reviewing the IPv6 configuration from a managerial perspective is lacking. The identified knowledge gap was filled by an exploratory qualitative study. For this, we performed thirteen interviews with IPv6 security experts and owners of identified misconfigured hosts. From these interviews, we got insights on how IP configuration is performed and why IPv6 configuration is often poorly configured.

Results

The thirteen interviews were transcribed and analyzed with ATLAS.ti. From the interviews, we found that the larger the company, the more formal the port configuration becomes. This process increases the number of reviews, but also time to implement a change. Next, we discovered that there are no real use cases for implementing IPv6 and only IT networking enthusiasts prefer to use IPv6. There is IPv4 scarcity but IPv4 lifetime extension technologies (Carrier-Grade NAT and NAT) are currently preferred over implementing IPv6.

Discussion and Conclusion

In the discussion, we compared our results with current literature. We discuss two reasons for IPv6 implementation, Europol and Rabobank. Secondly, we discuss two methods for organizations to migrate from IPv4 to IPv6. In one case IPv4 is used internally and IPv6 externally, and in the other case the entire network is running IPv6. Thirdly, a comparison is made between existing technology adoption theories and the adoption of IPv6. Also, a similar technology migration on the internet is added in this section. Finally, management awareness and the different actors in the IPv6 landscape are discussed.

IPv6 configuration has proven to be a difficult subject. Not all parties are ready for IPv6, and the general awareness and knowledge of IPv6 lacks behind. Thus far, there are little use cases for IPv6 implementation, which results in no priorities on IPv6. Resulting in improper and half configurations of firewalls and differences in open ports on dual-stack hosts. It is expected that in the future, with an increase of cost of IPv4

extension technologies and IPv4 scarcity, the world will migrate to IPv6. If this increases the knowledge and awareness of IPv6, it is expected that the number of misconfigured devices will decrease.

Furthermore, it is essential that besides technicians also managers become aware of this IPv6 problem. Growth in awareness and knowledge concerning IPv6 will be two factors that will greatly reduce the number of misconfigurations. Currently, there is no clear actor who is responsible for addressing this problem to managers. In this study, we discuss that Internet governance agencies, e.g., RIPE, could take this role.

Contents

List of Abbreviations

CGN	Carrier-Grade NAT
CIA	Confidentialy Integrity Availability
CLI	Command Line Interface
CSV	Comma Seperated File
CVE	Common Vulnerabilties and Exposures
DDoS	Distributed Denial of Service
DNS	Domain Name Server
DoS	Denial of Service
GDPR	General Data Protection Regulation
GUI	Graphical User Interface
HTTP	HyperText Transfer Protocol
HTTPS	HyperText Transfer Protocol Secure
IANA	Internet Assigned Numbers Authority
ICMP	Internet Control Message Protocol
IETF	Internet Engineering Task Force
IP	Internet Protocol
IPv4	Internet Protocol version 4
IPv6	Internet Protocol version 6
ISP	Internet Service Provider
IT	Information Technology
LAN	Local Area Network
NAT	Network Address Translation
NIST	National Institute of Standards and Technology
OSI	Open Systems Interconnection
RDP	Remote Deskto Protocol
RFC	Request for Comments
RIR	Regional Internet registry
SME	Small Medium Enterprises
SQL	Structured Query Language
SSH	Secure Shell
TCP	Transmission Control Protocol
TLD	Top-Level Domain
UDP	User Datagram Protocol
URL	Uniform Resource Locator
VPN	Virtual Private Network
VPS	Virtual Private Server
VM	Virtual Machine
WWW	World Wide Web

List of Figures

List of Tables

1 Introduction

The Internet is rapidly changing. An increasing number of systems is being connected to the Internet to improve availability, performance, and connectivity every day. All services need to be up and running, and more companies try to achieve high availability. Systems are attempted to follow the five 9 rule where a maximum downtime of 5.39 minutes per year is allowed [88]. This increased connectivity and up-time provide substantial benefits to companies and customers. The constant access to knowledge increases work performance, efficiency, and reduces costs which can all lead to increased profit margins [37, 53].

However, these benefits are not without their costs. When systems are being connected to the Internet, they become vulnerable to attacks over the Internet. An attack is an attempt to damage or steal a system or to access and extract information. This attack could limit functionality, increase cost, or break privacy laws. In 2019, these attacks are considered as one of the major threats against markets. Unfortunately, most companies are insufficiently aware of these threats, and their digital infrastructure is where their weaknesses lie. Even when aware of their weakness, some system administrators turn a blind eye and hope that it is not discovered and exploited. A famous quote from Stephane Nappo, Global Chief Information Security Officer at Société Générale International Banking, says, "It takes 20 years to build a reputation and few minutes of cyber-incident to ruin it" [66]. It demonstrates the importance of computer security and the risks involved when a person does not perform computer security appropriately. The history of data breaches and hacks shows that most did not target any sophisticated state-sponsored methods but ones with simple configuration mistakes or weak patch management. One of the most known and shocking examples is Equifax, a consumer and credit card reporting agency. In the end, the breach had been lead back to a single employee who neglected to apply a security patch to a weakness that had been known for months [59].

A well-known model used in computer security to guide policymakers is the CIA triad. Sometimes, it is also referred to as the AIC model to avoid confusion with the intelligence agency of the United States. The three elements of the triad are availability, integrity, and confidentiality and are considered the three most important elements of computer security [104]. Confidentiality refers to only authorized users being able to access the select information. Integrity is the guarantee that data or information remains unchanged by unauthorized parties and availability entails that all authorized users can access information. For each company or governmental agency, choices have to be made between the three triads and the amount of financial investment in them. For example, backup servers could be required when high availability is necessary, leading to extra hardware and maintenance and thus extra costs.

Fortunately, there is already an increase in security awareness and a growing number of companies considering these threats seriously. Computer security incidents receive more attention from higher management, and they represent a major topic in boardroom discussions. Accordingly, larger resources are allocated to keep the digital infrastructure secure [111]. These increasing investments lead to a larger security community, that aims to help each other to fight back malicious actors. Frameworks are being developed by security professionals on how to set up security controls. Personnel is being educated on how to avoid a security incident or data leakages.

However, an occasionally overlooked aspect is the deployment of IPv6, the successor of IPv4. Both are essential IP protocols that provide each machine with an address. IPv6 is being developed since 1998, and while similar to IPv4, there is an immensely larger number of IPv6 addresses available compared to IPv4. The IP protocol uses a 2^{32} bit IP address with a reduced space because of policy restrictions while the IPv6 protocol uses a 2^{128} bit address.

Based on Google's information, in 2019, approximately 24% of their users connect to Googles servers with the use of IPv6 [39]. Nowadays, there is a migration going on from IPv4 to IPv6.The migration started to get attention in 2012 when on June 6, the Internet Society, in partnership with several companies held World IPv6 Launch Day. Since 2015 the IPv6 adoption has been gently growing approximately 5% per year. Nevertheless, Internet experts do not dare to give predictions on the future developed and adoption of IPv6

Currently, both IPv4 and IPv6 are supported on the Internet, and they are being used alongside each other. An IPv4 address cannot connect to an IPv6 address and vice versa, meaning that hosts and servers need to support both protocols during the migration phase from IPv4 to IPv6. To set up an IPv6 connection, host, client, and the route between them need to support IPv6. However, not all system administrators are aware that their systems are compatible with IPv6 or that they are using it, leading to security implications since not all security measures applied on IPv4 are implemented on IPv6. Servers are running services that can be accessed over the Internet. A computer uses ports to differentiate the different services running on one host. Though not all Internet users should be able to access all services that are running on a server. With the aid of port blocking policies, certain services can have restricted access. An open port means a method to connect to a specific service on a server. However, open ports and misconfigured services can decrease the security of a system. Simply stated, if all ports are closed on an Internet-facing device, it is almost impossible for a malicious actor to get unauthorized access to this system over the Internet. Open ports form a weakness within the security link, and additional steps are required to block ports on IPv6. However, open ports, running services, and no firewall are needed for a server to have Internet connectivity. The system administrators who are responsible for the configuration occasionally erroneously misconfigure IPv6 [19].

1.1. Focus and Goal

This research focuses on the process of deployment, awareness, and security configuration concerning IPv6. This study was not only done from a technical but also from a managerial perspective. If the problem is to be looked upon from a technical perspective, it is clear that the solution is a reasonably straightforward fix. However, the mistakes made within these configurations possibly originate from other sources, raising questions surrounding the entire process of IPv6 configuration. The following topics are focused on to answer these questions:

1. **General state of IPv6 awareness** A thorough investigation of IPv6 support within companies was performed. This support can be both intentional and unintentional. Not all companies are aware of whether their hardware and ISP are IPv6 compatible. The goal in this section is to determine what their reasons are for supporting IPv6. This insight could be used to scope better the research phases to come. By analyzing these previously mentioned phases, it may be possible to determine the core root of this specific problem. Additionally, the differences and similarities between different companies can be identified. From a managerial perspective, it could be helpful to know which branches or sizes of companies are more aware of the IPv6 migration and what their implications could be by adopting or rejecting IPv6.

2. **Discover and map the configuration process on determining open port policies.** This topic attempts to gain insights into the entire process of IPv6 configuration. Despite that the technical solution is normally easy to implement, mistakes were still found too often. By identifying the different stakeholders and actors and their respective roles, a more clear picture of the configuration process can be created.

3. **Analyze the security posture of IPv6** Research will be performed to the current state of IPv6 security. In what way are organizations at the moment adding IPv6 security to their global IT landscape. Furthermore, research will be performed into the security advantages of IPv6 over IPv4.

4. **Future or current migration to IPv6.** Shortly, companies should migrate to IPv6. This topic addresses how companies are planning their migration or how far they are in the migration.

Based on these topics, the following research question was drafted:

> What is the reason for the existence of inconsistencies in open port policies on dual-stack hosts?

The first step towards answering this research question is literature research on IPv6, the background, current adoption, and applications [96, 113]. Next, data from potentially vulnerable companies were gathered about the process of IPv6 configuration, and its setup was identified. Last but not least, recommendations were drafted on how to best approach the adoption of IPv6 both, from a technical and a managerial perspective.

A series of sub-questions were drafted to allow a structured and thorough process that will lead back to a satisfactory answer to the main research question. The questions are listed below:

1. What actors play a part in the IPv6 configuration?
2. Which guidelines are considered to be most critical for port-based firewalling policies?
3. What processes lead to IPv4/IPv6 misconfigurations and inconsistencies that can be found in the wild?
4. Is there evidence of active exploitation of IPv6 misconfigurations?

1.2. Problem Description

In 2016, Czyz et al. published 'Don't forget to lock the back door' which described differences between open ports on IPv4 and IPv6 [19]. Simply stated, there are different security measurements implemented in both protocols. If other services are not properly configured, these could have large security implications. In 2019, another paper 'A Comparative Security Evaluation for IPv4 and IPv6 addresses' showed there are still important differences between the two protocols on how they are currently deployed on the Internet. Most attacks on the Internet are not sophisticated cyber attacks but scripts that use well-known misconfigurations and blindly try to use these. Almost 40% are bots scouring the Internet for different reasons. In 20% of these cases, it is known that these bots have malicious intent. A simple misconfiguration could be an easy target for such bots [46].

These differences between the two protocols give rise to questions concerning why there are configuration differences in the first place, even with the knowledge of the current number of malicious traffic. Thus far, it is unknown whether these differences are intentional or if they are the result of a misconfiguration. In other words, this thesis will try to uncover if the differences are erroneous.

There are multiple stakeholders involved in this subject. As the Internet is freely accessible for everybody, anyone with an Internet connection can add content. Without proper knowledge, this could result in false/improper security configuration. When keeping in mind that we want to keep the Internet freely accessible, it is hard to uphold limitations on adding content or connecting hardware to the Internet. However, servers that are compromised by hackers could be used for malicious intent. This intent could be harmful to people.

Another roadblock is that the Internet is a vast and anonymous space. Multiple parties are responsible for different elements. Furthermore, the Internet has no country boundaries and is accessible from anywhere. The number of differences between the configurations is enormous, and it is almost impossible to find companies and hosting parties responsible for the servers on this scale.

1.3. Intended Audience

Because this thesis has both, a technical and a managerial perspective, it is potentially relevant for system administrators as well as their managers, meaning that this thesis has a rather broad target audience that is

summarized in the two paragraphs below. This report is written with these different audiences in mind, and not all sections of this thesis are relevant for all targets. Only readers with a clear interest in both the technical as managerial perspective are invited to read the entire report.

IT Administrators & Security Personnel This thesis aims to give a clear overview of the technical background on IT and connectivity over the Internet. It addresses the possibilities, changes, and threats IPv6 brings or will bring to the Internet. Based on this thesis, administrators should have a better understanding of how to use IPv6 in their networks.

IT Management For IT governance, this thesis should be valuable because it is not only insightful into the basic knowledge of the Internet. It also elucidates how different companies in different branches set up their management controls around IT audit and IT security. This thesis gives a touche on how companies could implement and which parties are responsible for the implementation and execution of the migration from IPv4 to IPv6.

1.3.1. Actors

With the two mentioned target groups, we can briefly introduce the actors involved. First, are the technicians who are performing the configuration process. Technicians are at the core because thoroughness from them is directly linked to the safeness of a configuration. Next, are their managers. They play a role because they prioritize on what the engineers work. Engineers and managers are actors part of companies which are connected to the Internet. Another group of actors are ISPs, Internet Service Parties, and hosting parties. They supply connectivity and infrastructure to companies who are connected. Finally, there are internet governance agencies who are responsible for maintaining, spreading knowledge, and improving the Internet

1.4. Societal Relevance

To keep the Internet accessible to everyone, it is crucial that it undergoes the migration from IPv4 to IPv6. If the migration is withheld too long, the chances are that more security implications will occur and cost will increase in the future. Current measures that increase the lifespan of IPv4 could withhold innovation and the overall quality of the Internet. The initiative for this endeavor should first and foremost come from companies, universities, and other large institutions. The normal user will soon follow automatically, as they will not be able to notice any difference. Likely, the transition is not advancing as hoped because network engineers and IT administrators may be afraid to change towards a system that they perceive as a dangerous unknown. A better understanding of these security issues could hopefully lead to a learning process that should allow for a more smooth transition into IPv6.

1.5. Management of Technology

This thesis is written for the Management of Technology program at Delft University of Technology. On the MOT official website is MOT described as *"In the Management of Technology programme you learn to explore and understand technology as a corporate resource - a resource that allows a firm to keep many different balls in the air."* For this thesis, the chosen technology is IPv6. IPv6 is a technology currently being adopted on the Internet. With the aid of skills acquired in the MOT program, the IPv6 configuration and adoption process is thoroughly analyzed.

Furthermore, the problem is examined from both a technical as managerial perspective as is the aim of the Management of Technology curriculum. One goal of this study is to have a broader perspective on a technical problem and see if root causes can be linked to managerial decisions. The skills required for performing such a study affiliate with the skills thought in the MOT program.

2

Background Knowledge

The following chapter will discuss the required knowledge for further reading of this thesis. If topics discussed in the final chapters are not understood, then this chapter can give an idea about the fundamental theory behind it. It is split up into multiple sections: first, a general explanation will be given about computer security in general, with the recent developments in the field. Next, the required technical knowledge will be explained around the TCP/IP model, the internet in general, tech, and how information is transferred from one computer to another.

2.1. What is Cyber Security

This first section should improve knowledge of the general field of Computer Security. It is used to help the reader understand the position of security in the global digital landscape. Computer Security or cybersecurity is the protection of items that have value for a person or an organization [6]. These items are called assets. Within the concept of cybersecurity, assets can be split up into three different categories: software, hardware, and data. Each of these items can have a different value.

By determining the value of an asset, a more accurate risk assessment can be made to analyze which assets should be protected, and what methods should be applied to each asset. A security event could occur if method, opportunity, and motive intertwine [73]. A security event is an event that influences daily operations. It can impact the security of the systems or the data. There are multiple sources and methods on how a security incident can be discovered or reported. This section will further elaborate on malicious actors, attack methods, and consequences of a security incident.

2.1.1. Vulnerability-Threat-Control Paradigm

The Vulnerability-Threat-Control framework is well known within the field of cybersecurity. It prevents any assets from harm and allows security personnel to mitigate the situation. Vulnerability-Threat-Control stands for

Vulnerability A vulnerability is a weakness in a system, e.g., in a procedure, design, or implementation that might be exploited to cause loss or harm to a person or organization [74].

Threat A threat to a computing system is a set of circumstances that has the potential to cause loss or harm. Threats can come from multiple sources both from human or computer origin. When the former is true, the threat can be described as an attack. In case of a computer origin, a malicious actor is not involved. In this case, it can thus not be considered an attack by humans [74].

Control A control or countermeasure that is used to mitigate vulnerabilities and threats [74].

By implementing checks, or by following guidelines, vulnerabilities and threats can be kept to a minimum [73]. Vulnerability will be further elaborated upon in the corresponding section below. An example of these

guidelines is the NIST cybersecurity framework with multiple standards on how to set up controls. Such steps are known as risk assessment and risk mitigation procedures [91].

2.2. Malicious Actor

Various actors can be defined within the field of cybersecurity. A malicious actor is a person or entity who purposefully misuses a vulnerability within a system. This abuse of a vulnerability is referred to as an exploit, e.g., a predefined script or program that attacks a vulnerable machine executes code, and delivers a payload the malicious actor has prepared. Through this, a hacker can, for example, acquire remote access to a specific machine by bypassing the normally required authentication method [109]. However, there are multiple other methods to misuse a vulnerability. Different actors and their capabilities were listed below [44].

Amateur An amateur, or *script kiddie*, is defined as "One who relies on pre-made exploit programs and files, 'scripts', to conduct his hacking, and refuses to bother to learn how they work. The script kiddie flies in the face of all that the hacker subculture stands for - the pursuit of knowledge, respect for skills, and motivation to self-teach are just three of the hacker ideals that the script kiddie ignores" by the urban dictionary. They do not have hacking skills and use other developed methods to perform an attack without having a fundamental understanding of the tools [107].

Hackers A hacker, also known as a cracker, is a more professional actor who does not rely on pre-made scripts but attempts to find vulnerabilities by themselves and develop exploits for this. They then attempt to disable or gain access to a system. They are more knowledgeable about computer systems when compared to a script kiddie. Not all hackers are criminals; most of them do it for sports, curiosity, or the challenge. There are Black Hat, Grey Hat, and White Hat hackers. Black Hat hackers are criminals who use their skill for criminal intent and do not always have permission to break into the systems. On the other side, White Hat hackers are often referred to as pen testers and are mostly *hackers for hire*. They test the security of an organization and give recommendations on how to improve it. A pen tester or penetration tester often work as consultants for an IT security company. A Grey Hat hacker is between those mentioned above. They often perform white hat activities to hide their black hat work [45].

Professional Criminals The numbers on organized cybercrime has rapidly increased over the past years. Professional criminals have discovered how lucrative computer crime can be. They mostly operate from a monetary perspective or status. Examples of professional crimes are identity theft, large scale stealing of credit card information, or the release of ransomware. Ransomware is software that encrypts a hard drive for a certain down payment. It is possible to retrieve the personal data when paying a hefty fee. Their skills are often highly professional, and they try to operate within complete anonymity [73].

Nation States For nations, because of the increase of connectivity, the cyberspace has also become a vital field to operate. Multiple nations have started to increase their cyber operations to prepare for cyber warfare, both openly and behind closed doors. They operate in the highest spectrum of violence and are believed to have a broad knowledge and skill. They mainly focus on intelligence gathering and large-scale sabotage of industry systems. Thus far, most cyber operations by nation-states remain unknown to the general public. However, Stuxnet is a suspected example of modern cyber warfare by the United States of America and Israel [29]. Stuxnet was a highly sophisticated piece of malware that was discovered in 2010. Its purpose was to attack Iranian SCADA systems to slow down Iran's nuclear program. However, the USA an Israel never confirmed that they were responsible for Stuxnet.

2.3. Cyber Security Roles

Besides offering a picture of malicious actors in the cybersecurity landscape, also some cybersecurity roles and job titles will be discussed. Their key role is to take responsibility for protecting networks, infrastructure, and computer systems. There are more roles then described in this chapter, but a few suitable for this study are briefly elaborated here.

Chief Information Technology Officer The CISO is C level executive that is responsible for the vision of the protection of all assets of a company. They direct staff to minimize risk to minimize IT risk.

Security Officer A security officer is a position filled by a person who forms the bridge between management and the technicians. A security officer implements information security into a company.

Ethical Hacker An ethical hacker or white-hat hacker is a security specialist who attempts to penetrate networks or computer systems. They can be seen as hackers for hire who actively search for weak systems in a network.

Computer Security Response Responder A team member of the Computer Emergency Response Team (CERT) that responds in case a cyber attack is launched to a company.

Security Analyst An security analyst has a more hands-on job that investigates networks, software, or hardware with available tools. They try to remedy vulnerabilities and search for security incidents.

Security Architect A security architect is responsible for designing or partially designing a security system.

Security Auditor A security auditor is a person responsible for auditing the systems. An auditor will provide a report of the current state of the infrastructure.

2.4. Common Attack Methods

Multiple attack methods are used by a malicious actor to compromise a system successfully. Two different attack goals are: to compromise a system or to disable access to a system.

2.4.1. Password Attacks

There are multiple methods for a malicious actor to obtain unauthorized access to a system. The first and probably the most famous one is a password attack. This attack relies on the fact that a password is a weak link in the security chain. It is seen as a weak method of authentication. People recycle passwords across platforms, use easy-to-remember passwords, or fail to change default passwords [104]. If a person reuses passwords on multiple systems or websites and one of them is compromised by an attacker, in theory, all other places the user has the same credentials (username + password) are also vulnerable. Even with extensive security and high-quality encryption, systems could easily be compromised if the same credentials are leaked via another platform. Hackers use these human mistakes to gain unwanted access. They seek weaker authentication methods to recover credentials and then reuse them on vital services (banks, e-mail). Conversely, even if a password never leaked, there are other methods for malicious actors to try and break the encryption by, e.g., brute-force, dictionary attack, and password spraying [73]:

Brute force In a brute-force attack, a malicious actor systematically tries all possible combinations of letters, numbers, and symbols. These attacks can be automated with scripts. To minimize the risk of a brute force, websites or companies often have password policies to use at least one capital letter, number, symbol, and a minimal amount of characters, increasing the breaking time exponentially [104].
There are methods were system administrators use password policies that after a fixed amount of tries, a password resets or an account becomes disabled. This method is a workable defense against a brute force attack. However, if a malicious actor can capture a hash which, for ease of this example, is the encrypted result of a password, they can try and brute force it on a local system and are not dependent on the feedback from a system that returns that a password is incorrect.

Dictionary attack A dictionary attack is a more sophisticated brute-force attack where instead of trying all combinations, only well-known passwords are tried. It uses large databases with previously leaked, common or expected passwords. The password list RockYou.txt is a well-known example of a collection of 32 million passwords that have been collected throughout the years. It is advisable to check if a password is not in the RockYou password list [54]. With a dictionary attack, attacks are more targeted and faster brute force attacks which decreases the breaking time [104]. Have I been Pwned created a website with 550 million breached passwords and is updated it regularly. It is possible to enter a password on this website and see if it is being used on a password list [43].

Password spraying Password spraying is, in some ways, the opposite of a normal brute force. Instead of having one username and try all possible passwords with that username, it is reversed. An attacker tries a considerable amount of usernames with a small specific amount of passwords. This attack is efficient because usually, usernames are stored in plain text, not encrypted, while passwords are if security is properly implemented, stored in ciphertext. Ciphertext is the encrypted version of the plaintext. Generally, this is achieved by performing mathematical functions. If a database is compromised and an attacker can extract all usernames and passwords, the attacker will have all usernames in plaintext while the passwords are not usable. With a password spraying attack, a malicious actor has a method to, in a reasonable amount of time, gain access to weakly protected accounts [104].

2.4.2. Social Engineering

In some cases, it is easiest to acquire someone's password just by asking for it. This method is commonly known as social engineering. With social engineering, an attacker attempts to gain the trust of the victim and make the victim give them the required information. In this case, an attacker can impersonate another person, flatter, or deceive to achieve the goal. Furthermore, with spoofing techniques, a malicious actor can easily hide his or her identity and pretend to be someone else. Examples of spoofing techniques are e-mail of phone spoofing where they can send e-mails or make phone calls under the identity of another person. The identity that malicious attacker impersonates can range from an acquaintance of the victim or a person with authority. Stating that you are from a governmental agency or bank can have a massive impact on the cooperation of victims. Security personnel always states that people are the weakest link in cybersecurity, and hackers often take advantage of this. It does not matter how extensive security is when employees fall for social engineering techniques and give the attacks access to the systems [42].

Phishing is a form of cybercrime where attackers try to trick people into giving up their information. It is mostly known as e-mail phishing were people are directed to misleading websites to fill in their information. An attacker can, for instance, recreate the PayPal website and send an e-mail to the victim that his or her account has an issue and if the person can log in to confirm some information. If the victim falls for the scam, he or she will provide the credential to the malicious actor. Attackers can also use phishing to trick a victim into downloading malicious files or malware, which can infect and give access to the victim's computer. According to security experts, there are multiple categories of phishing emails [42].

Large scale phishing A typical phishing attack where an attacker sends out multiple not personalized e-mails to a large group of people [104].

Spear phishing With spear phishing, an attacker first gathers personal information about his or her victim to create a more authentic e-mail. It will focus on a smaller specific group and can seem to be written by a colleague from within a firm [42].

Whaling Whaling is a more extensive form of spear-phishing where attackers target high-placed employees, such as the CEO or CFO [104].

2.4.3. Denial of Service

With a Denial of Service or a DoS, a system is temporarily disabled from service, disabling normal and legitimate traffic. This downtime happens when a server is flooded with traffic. An attacker can thus compromise the availability of an IT system by voluntarily overloading a server. A solution from security personnel is to block that specific device so it cannot access the attacked server. However, when this happens from different sources, it becomes harder to defend against. In this case, the attack is referred to as a Distributed Denial of Service, or DDoS. In that case, it becomes harder to update the firewall with the correct configuration to keep all systems safe. Confidentiality and integrity are not compromised with these attacks [58]. In figure 2.1 is an example of a DDoS displayed. Here is visible how multiple actors connect at the same time to one server and thus overloading this system. A DDoS can happen by multiple actors but often is performed by one malicious party that controls multiple systems via a command and control server. From this command and control server, it is possible to control multiple systems and direct the traffic to the victim's machine. A DoS is on average a relatively easy attack, but with a rather low impact because the confidentially and integrity is not harmed [71].

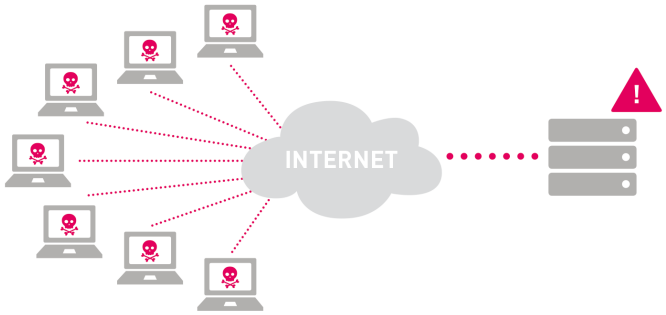


Figure 2.1: Graphical display of a DDoS attack performed by multiple clients. A DDoS is an attack to (temporary) disable a service.[1]

2.4.4. Vulnarabilities

In the process of writing and configuring computer software, mistakes can occur that created weaknesses in a software package. Vulnerabilities can be defined as "The degree to which a system, or part of it, may react adversely during the occurrence of a hazardous event" [80]. A malicious actor can use these vulnerabilities to access a system or network. Well-known weaknesses are weakly configured systems, default passwords, or mistakes in the source code of software packages. Hackers can use these mistakes to gain access to systems or increase their privilege level to obtain more rights. Normally in computer systems, there are different levels of access. An average user, if configured properly, does not have Administrator or sudo access to a system, and only certain people or certain accounts have privileged access. Administrator or sudo access is required to make modifications to systems that could impact the security or performance. It is called privilege escalation if a malicious actor got access to a user account and attempts to increase his rights to the level of administrator. These increased rights would allow a hacker to perform lateral movement through a network [64]. Lateral movement means that a hacker has access to one or multiple clients, which do not necessarily have access to vital data but uses this to get access to more vital servers, e.g., a file server or e-mail server. These vital systems are normally not directly accessible if a network is configured adequate, but they are accessible via weaker or more often used computers. Software packages and Linux distributions are available that are equipped with pre-defined scripts to attack well-known vulnerabilities [11, 27, 50]. Generally, these tools are developed for penetration testers, white hat hackers, who use it for testing purposes. However, the same software packages can also be applicable for malicious activities by black hat hackers.

Fortunately, the field of cybersecurity is fighting back to reduce the number of weaknesses. Currently, there is a global community of security specialist addressing these threats together to identify, locate, report, and distribute knowledge on this subject. Any system administrator and security expert who discovers a new deficiency should report it to a central agency named MITRE Corporation (an American non-profit organization). They operate the national Cybersecurity FFRDC (Federally Funded Research and Development Centers), collect all found deficiencies, and tag those with a unique CVE (Common Vulnerability Exposure) [63]. A CVE is a list of entries of a publicly known weakness. All vulnerabilities and CVE's are placed in a database and shared to the public [79]. With the aid of specialized responsible disclosure protocols, vendors receive a notification beforehand so they can develop and release security patches before the vulnerability is made known to the public. An example of security experts that search for unknown vulnerabilities is Google project, *project zero*, a team of Google's top security researchers that have the goal to search and find security flaws in software [40]. In the end, companies monitor the published CVE's and identify if they are of relevance to their products. If they are, security patches released by the vendors should be applied. If they do not follow these procedures, possibilities are their systems can become vulnerable [56, 91]. Within the Dutch government, the NCSC (National Cyber Security Centre) is responsible for monitoring CVE's, performing an impact analysis, and distributing background information and solutions to the public [67].

2.4.5. Vulnerabilties and Misconfiguration

The previous section discusses vulnerabilities in software or hardware. Mostly discussing mistakes in source code of the software. The general perspective of incidents on the Internet is often believed to be caused by

[1] https://blog.paessler.com/types-of-ddos-attacks

programming errors. Programming errors that allow software to crash or circumvent authentication methods. Often the public thinks that zero days (to the public unknown programming errors) are used for attacks. However, this is not always the case. Attacks are also due to simple configuration mistakes or lacking to timely patch vulnerable systems. Attacks that are performed with zero days are only executed by state actors who do not leave a trace. State of the Art software can still be vulnerable if it is not correctly implemented. A Next-Generation Firewall is only as good as the rules that are implemented by the firewall administrator. If it is not configured correctly, weaknesses and security holes will exist. Moreover, software needs regular updates because of mistakes are being found. Too use Equifax again as an example. An update for the vulnerability used in the Equifax attack was released months before the attack occurred. The reason why this company was breached was that they did not correctly patch the vulnerable system [22]. So, it is crucial to see the difference between a vulnerability by the system/network/firewall administrator and the software/hardware vendor who supplies the software/hardware. The relation to IPv6 configuration, it is expected that most differences between the IPv4 and IPv6 are because of the configurations made the system administrators.

2.5. Consequences of a Cyber Attack

Too frequently, people are not aware of the negative consequences of a cyber event for a business. IT security experts have created a list with multiple consequences and identified five elements that could have an impact of an attack. This list gives guidelines and guidance on the possible consequences. Knowing the negative consequences of an attack could help in better mitigating the problems. Sometimes a solution to risk must be accepted because the solution is too expensive or not achievable. Knowing the consequences is a requirement for correctly knowing how to handle the risk [1, 104].

1. **Physical/Digital harm** Physical/Digital harm is a negative effect on something or someone. Damaged or destroyed assets are examples of this effect

2. **Economic harm** Economical harm can be described as the adverse financial effect or economic consequences, hence, loss of profit or stock price.

3. **Psychological harm** Psychological harm is the loss of well-being of individuals.

4. **Reputational harm** The general opinion of a company or entity can be seen as reputational harm.

5. **Social/Societal harm** With societal harm, there is a more significant risk at stake. There is a change of disruption of society if these systems are attacked. This can occur for companies or organizations that have a crucial impact on society, and the population relies on their continuance. An article released by NCTV talks about the Dutch dependence of vital processes and systems. Systems discussed in this article are, for example, water- or power-plant. In case one of those systems would disrupt, it could have a tremendous impact on the people who rely on those systems [68].

2.6. History of the Internet

The Internet is developed in the '60s by the United States Department of Defence. Back then, it was known as the ARPANET initially funded by the Advanced Research Project Agency (ARPA). The reason why this was a revolutionary new concept is that it made use of packet-switching. Packet Switching is a communication method where data is split up into multiple smaller blocks and streamed over a shared network individually. Before packet switching was used, all communication was circuit-switched. When communication was set up, a dedicated channel was created between the two nodes. Analog telephone networks are an example of this. The first message that was retrieved over the Internet was from Professor Leonard Kleinrock's laboratory at University of California, Los Angeles (UCLA) to another node at Stanford Research Institute (SRI) in 1969. At the beginning of the Internet, researchers were the primary users of the Internet to share information between different research institutes [52].

The Internet is often confused with the World Wide Web, Tim Berners-Lee developed that at CERN in 1980. The invention of the World Wide Web made it possible to display documents and information over the Internet with a specific software application, now called web browsers. This invention made the Internet more accessible for the general public, and halfway through the '90s, the Internet became usable for a broader audience. A web browser helped in the adoption of the Internet in the general public. With more people

starting to see the benefit of the Internet, more people adopted this technology which increased fast communication like e-mail, IM, and VoIP. With all the new possibilities, social networks and online shopping came into existence to form the Internet as we know it today [52].

2.7. Internet Governance

This section will give an overview of Internet governance and parties involved in maintaining, updating, and improving the Internet. First, some general organization will be discussed. Then, a brief overview will be given about multiple non-profit organizations responsible for Internet governance.

1. **Internet hosting service** An Internet hosting service is responsible for running applications accessible via the Internet. It thus allows people or companies to access these services online, the supply computing resources or storage. Examples of hosting services are websites or e-mail servers. In other words, they allow storing information, data, and applications on their servers. Hosting services are ideal for companies who temporally have to scale up their services or do not have the hardware to host some services themselves.

2. **Domain name registrar** A domain name registrar is a company that is responsible for managing domain names. Domain names must be coupled to IP addresses. In section, DNS will be further elaborated on what domains are and their purpose on the Internet. Domain name registrars are responsible for adding, removing, and updating this information.

3. **ISP** An ISP, or Internet Service Provider, is an organization that provides Internet access to companies and private use. Examples of ISP's in the Netherlands are KPN or Ziggo. They can be non-profit or privately owned.

Through the years, multiple companies started to perform governance and improve the Internet. Some noteworthy companies are the following.

1. **IANA** IANA is a department of ICANN, which is an American non-profit organization that supervises the total IP allocations. Furthermore, They govern other Internet-related tasks, like giving out top-level domains. Top-level domains are the highest zones that are being used on the Internet. Examples of top-level domains are .com, .nl, and .info.

2. **RIR** RIR stands for the Regional Internet Register, which consists of five organizations which are responsible for registration, managing, and allocating IP addresses between the different continents. IANA delegates these processes (registration, managing, and allocating) to the following five RIRs [86], see figure 2.2:

   (a) The African Network Information Center (AFRINIC) serves Africa

   (b) The American Registry for Internet Numbers (ARIN) serves Antarctica, Canada, parts of the Caribbean, and the United States

   (c) The Asia-Pacific Network Information Centre (APNIC) serves East Asia, Oceania, South Asia, and Southeast Asia

   (d) The Latin America and Caribbean Network Information Centre (LACNIC) serves most of the Caribbean and all of Latin America

   (e) The Réseaux IP Européens Network Coordination Centre (RIPE NCC) serves Europe, Central Asia, Russia, and West Asia

3. **SIDN** In the Netherlands, SIDN is responsible for managing the .nl domain. They regulate the registration of the .nl domain names and provide knowledge about protocols and support initiatives for improvement of the Internet for users [102].

[2] https://upload.wikimedia.org/wikipedia/commons/9/95/Regional_internet_Registries_world_map.svg

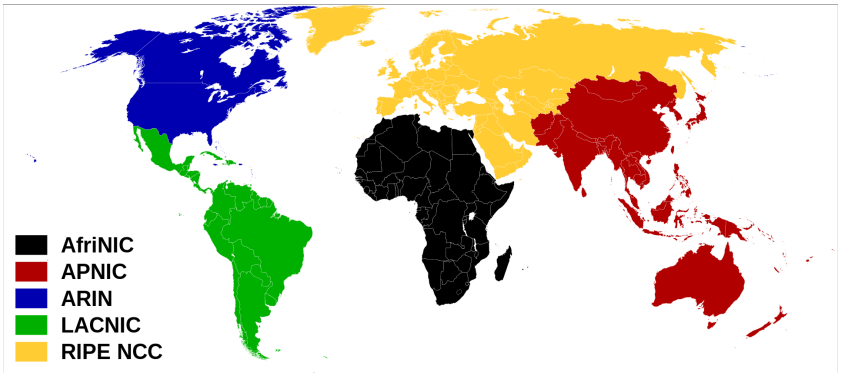


Figure 2.2: The world' five RIR's resposible for different regions. Each RIR manages the registration and allocation of IP addresses in their region. [2]

2.8. OSI Model

The OSI model is a seven-layer conceptual network model that consists of a series of protocols with specific functions set up in different layers. The OSI model is shown in Figure 2.3. OSI stands for Open System Interconnection and is a conceptual framework that shows how a network operates and how digital communication between different systems is performed. The model is used by IT personnel to give a more visual indication of how computers communicate with each other. By identifying the different layers, it is easier to pin down to potential issues. It is a universal standard for managers, software developers, and physical vendors. A piece of information travels through the seven layers from top to bottom [99], where layer seven is the closest to the user and layer one is the result of the input by the user translated into bits and bytes.

Layer 7 – Application This protocol is closest to the user. With this protocol, the application layer, users interact with a computer. An example of an application is a web browser [4, 104].

Layer 6 - Presentation The presentation layer is responsible for transforming data that was received by the application layer into a format that is required for layer 5. Examples of this are encryption or compression of the data [4, 104].

Layer 5 – Session This layer establishes, maintains, and terminates sessions between two computers. An example of a protocol in this layer is NetBIOS, NetBIOS is a deprecated file sharing and name protocol, which was used by Windows to exchange data over a LAN [4, 104].

Layer 4 – Transport The information received by the previous layer is transported into this layer. There are two essential protocols within the transport layer, TCP and UDP. TCP stands for Transmission Control Protocol and is used for a reliable connection. Furthermore, it is responsible for segmentation, acknowledgment, and traffic control. An acknowledgment means that the receiver acknowledges that it successfully received the package, meaning that in this protocol have error detection and correction inbuilt. This built-in mechanism ensures that packets that get lost in transfer are resent. The second protocol, UDP or User Datagram Protocol does not have those controls, and when a packet is lost in the transfer, the receiver will fail to receive the packet. There are advantages and disadvantages for both protocols, UDP being the fastest and TCP the most reliable. A use case for TCP is the transmission of a document. The receiver would like to receive the entire document because missing pieces could lead to incomplete data. On the other hand, UDP is preferred in video calling where the speed of data transfer is more relevant than the completeness of data [60].

Layer 3 – Network Routing protocols are defined here. A packet created by one of the previous layers can take multiple routes over the Internet to reach its destination. The network layer is a protocol that is responsible for finding an efficient route based on its original and destination IP address [4, 104].

Layer 2 – Data Link The data link layer is responsible for transforming the packets from the network layer into frames for the final transmission in the physical layer. Different formats are available, and the data link layer uses information from the hardware of the network to decide which protocol to use. However, Ethernet

(IEEE 802.3) is mostly used in modern technology [4, 104].

Layer 1 – Physical The physical is the final layer where the frames are converted into bits. Bits can then be transported over the medium that supplies the physical connection. A bit is the smallest representation of data available and can only have two options, on or off. Mostly known as 1 and 0. The physical layer is also responsible for receiving the incoming bits and the transmission to the data link layer. Then the entire process is performed the other way around to get the bits into a presentable form for the end-user [4, 104].

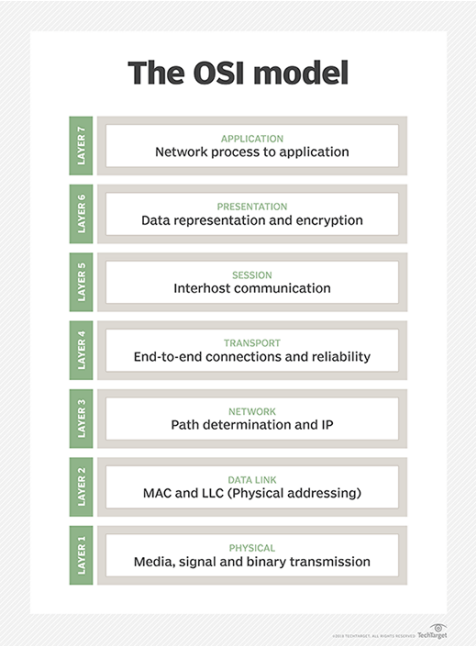


Figure 2.3: Illustration of the seven layers of the OSI model. The OSI model is a conceptual model that consists of a series of protocols with specific functions. [3]

2.9. TCP/IP

The Internet is based on a protocol developed by Robert E. Kahn and Vint Cerf in 1970. During their research, they developed the Internet protocol suite, commonly known as TCP/IP. The TCP/IP model is a protocol that offers a communication channel on how packets should be sent, routed, transmitted, and received at different nodes. It is partly comparable with the OSI model. The difference is that, while the OSI is a conceptual model used for better understanding, is TCP/IP a protocol for establishing connections. It is as mentioned before packet-switched and information is thus split up into different packages and transferred over the Internet [76].

Within the TCP/IP protocol, mechanisms for error correction are in place, meaning it corrects flaws in the connection by correcting the missing packages and making sure all packages are delivered. When a packet is sent from transmitter to receiver, the receiver must acknowledge the packet. If the receiver does not acknowledge a packet, the sender will resend the packet to make sure that all information was delivered. Currently, the Internet Engineering Task Force is responsible for maintaining the TCP/IP protocol [13].

2.9.1. IP Addresses

One important part of the TCP/IP protocol is an IP Address. An IP address is a numerical label attached to every device connected to a computer network. Simply stated, without an IP address is it impossible to participate in the communication on the Internet. Devices connected over the Internet need to "know" where to send their information or packets. For this, they use IP addresses, which are comparable to zip codes. Currently, there are two main versions of the IP protocol, IPv4 and IPv6. At the moment, most traffic is routed

[3] https://searchnetworking.techtarget.com/definition/OSI

over the IPv4 protocol. IP addresses are usually written in a human-readable notation visible below [104]. Both IP addresses route to google.com.

```
IPv4:  172.217.164.174
IPv6:  2607:f8b0:4004:815::200e
```

IPv4 uses a 32-bit address which results in 4.29 billion unique IP addresses. The Internet Assigned Numbers Authority (IANA) globally manages the IP addresses space, and there are five Regional Internet Registries (RIR) that manages the allocation of IP address. These institutions are responsible for registering the owners behind an IP address. They allocate for example ranges of IP addresses for ISPs who then allocate these addresses to private end-users. Mostly are these ranges allocated by subnet. A subnet is a cluster of IP addresses close to each other for efficient routing.

When the IPv4 protocol was developed, it was never expected that the Internet would grow to its current size. Because of this explosive growth, all IPv4 addresses are now allocated, meaning that IANA no longer has any IP blocks available for larger organization [25]. However, there are still IPv4 addresses available in smaller ranges. For example, it is still possible to acquire or buy 8.192 IP addresses (/19) in one range. However, 65.536 (/16) becomes more difficult. Those ranges are becoming more scarce by the day. Also, organizations who own IP ranges, for example, ISPs, can still give out IP addresses to end-users. In contrast, larger organizations who need entire subnets cannot get one anymore on IPv4 as it is no longer available. The paragraph IP Depletion will give a more extensive explanation on this subject.

With the increase in world population, there is currently almost one IPv4 address per two people, which could lead to IP shortage. Fortunately, this problem was already foreseen in 1998, and the IETF (Internet Engineering Task Force) developed IPv6 as an alternative resolving the shortage [21]. It became the Internet standard in 2017 [101]. IPv6 uses 128-bit addresses that, in theory, allow for 3.4×10^{38} unique addresses. The actual number is slightly inferior to the total amount because some ranges of IP addresses are reserved for special purposes. A few examples are addresses reserved for local networks (Link-Local Unicast) and special addresses that are being used in the IPv4-IPv6 Translation [16]. In comparison to IPv4, there are 7.9×10^{28} times more IPv6. Therefore, it is doubtful that there will ever be a need for a new IP system [75]. Currently, both IPv4 and IPv6 are used on the Internet and based on Google's information, approximately 24% of Google users connected to Google with IPv6 in 2019 [39]. Currently, ISPs and hosts need to be able to support IPv6 to make a reliable IPv6 connection on the Internet. A reliable IPv6 connection can only be set up when the client, host, and route in between all support IPv6.

At this moment, both IPv4 and IPv6 are supported on the Internet, but they are being used alongside each other, meaning that an IPv4 address cannot connect to an IPv6 address. Additionally, IPv6 was not developed to be compatible with IPv4 but to replace it eventually completely.

2.9.2. IP Depletion

As mentioned before, this number of IP addresses is insufficient with the current growth of the Internet. A more permanent solution is IPv6, which has an immensely larger number of IP addresses. Especially with the newly connected devices where and multiple users having multiple systems is it not expected that the demand will decrease in the future. In theory, the transfer from IPv4 to IPv6 should have gone considerably more effortlessly; however, only 25% of the Internet now uses IPv6. Figure 2.4 is a graph by Google that quantifies all IPv6 connections over the past years. It can be seen that the IP adoption has been very moderate over the past years and was only initiated in 2013 despite IPv6 being developed around 1998.

Due to differences between the two protocols, most network administrators are reluctant to switch to IPv6, as they do not feel knowledgeable enough. Older organizations lack motives to undergo the transition to IPv6 as their IPv4 services function, while new organizations, especially ones in Asia, are confronted with the depletion of IPv4.

[4] https://www.google.com/intl/en/ipv6/statistics.html

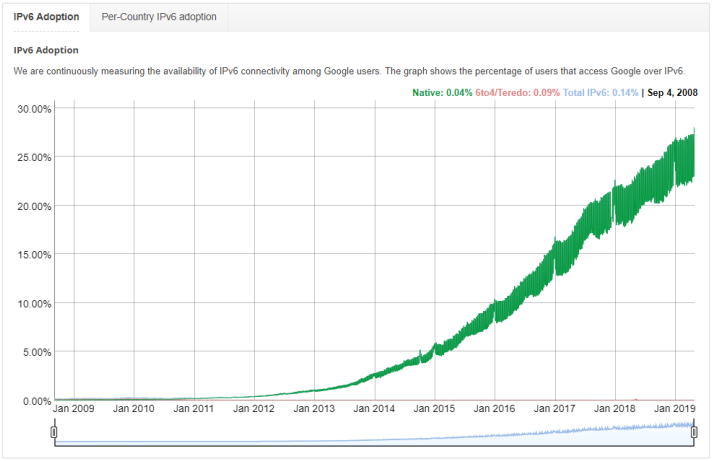


Figure 2.4: This graph shows a percentage of the number of connections to google over IPv6 in the past 10 years.[4]

2.9.3. Ports

Another part of the TCP/IP and UDP protocol is computer ports, where ports are the endpoints of communication. Ports are the gates on the server where the information is being retrieved or sent. They are not the route in between. Ports are based on a 16-bit number, resulting in 65535 different ports. Ports are similar between IPv4 or IPv6 protocols. Normally, a server supports multiple applications, where applications can be seen as software programs operating different processes. A server is capable of performing different tasks, and these different tasks require different ports to communicate. Examples of this are the hosting of a database or website. It requires different ports because different applications require different protocols. Ports can thus be seen as gates to enter services on the computer, which can be open or closed. The first 1024 ports are known as the well-known ports and are reserved for often-used applications. Examples are 443 (HTTPS), 80 (HTTP), and 22 (SSH) [104]. The reason why different applications require different ports is that as soon as one application is running on a port, it is reserved by that application. For two devices to communicate via a protocol, they need to speak a similar *language*, so they understand each other. HTTP requires a different *language* then HTTPS and thus need a different port to run on.

From a security perspective, it is best practice to have as few as possible ports open, as each open port increases the risk for a security breach. This strategy is often referred to as the principle of least privilege [93]. A firewall on a system is a mechanism that manages the opening and closing of ports. A firewall can be locally on a computer or centralized to create segmentation between the Internet and a local subnet. Furthermore, firewalls monitor network traffic and are responsible for blocking and filtering network traffic on ports that are not required. A firewall knows with traffic is allowed with the aid of rules. System or network administrators build rules that state which communication is allowed and which traffic should be blocked. For simplicity, we can state that there is inbound and outbound traffic, where inbound traffic is a connection that is being initiated from another device to the host. For example, clients connecting to a server too receive the content of a website. Outbound traffic is where the hosts set up the connection himself to another device. This is, for example, could be a server that checks if updates are available. Normally the rules for inbound traffic are more strict than for outbound traffic. [5, 45]. The important ports for this research are summarized in Table 2.1.
The table shows which ports are scanned on the hosts. In theory, every device can decide which protocol should run on which port. Nonetheless trough years some ports became claimed by often used protocols. The first 1024 ports are known as well-known ports and reserved for privileged services. Although not all ports in the table are below 1024, they are the default ports these applications run on. Every protocol has it's own weaknesses and risk to be accessible via the internet [104].

2.9.4. Network Address Translation

There are several reasons why the roll-out of IPv6 has stagnated, and the most well-known one is NAT or Network Address Translation. NAT was developed mid-1990 and is a method for remapping IP addresses from one address space to another by changing the network address. Researchers had already expected that the

Port Number	Protocol	Explenation
21	FTP	File Transfer Protocol, Used for the exchange of files between computers
22	SSH	A protocol for secure connections over a unsecure network
23	Telnet	A protocol for remote access to machine
25	SMTP	One of the protocols used in sending e-mail over the internet
139	NetBIOS	NetBIOS is a protocol for computers on a local network
179	BGP	Border Gateway Protocol is the most important routing protocol on the internet
80, 8080	HTTP	Hyper Text Transfer Protocol used for communication in a app or web-browser
443, 8443	HTTPS	HTTPS is a secure extension of the HTTP protocol
445	SMB	A windows protocol used for file sharing
631	IPP	A protocol used for communication between clients and printers
3306	MySQL	A language to create and maintain databases
389, 3389	RDP	A Microsoft protocol to give a GUI interface to another computer
5800, 5900, 5901	VNC	VNC is a universal desktop sharing protocol to remotely access another device
9200, 9300	Elasticsearch	An effective search engine to search to large data sets
11211	Memcached	Memcached is caching protocol to speeed up delivery of a website
27017, 27018, 27019	MongoDB	MongoDB is a protocol for accessing databases
1433	MsSQL	A similair protocol as MySQL
6379	Redis	A protocol for in-memory structure storage

Table 2.1: All ports that were scanned for this study. Each port has a port number and corresponding protocol. A brief explanation is added where the port is used for.

RFC1918 name	IP address range	Number of addresses
24-bit block	10.0.0.0 – 10.255.255.255	16.777.216
20-bit block	172.16.0.0 – 172.31.255.255	1.048.576
16-bit block	192.168.0.0 – 192.168.255.255	65.536

Table 2.2: The following IP addresses are reserved by IANA for private networks. These addresses are not directly routable and are used behind a NAT.

number of available IP addresses would not suffice and in RFC2663 the principles of NAT were elaborated [103]. An RFC is a Request for Comments which describes protocols or other aspects used on the Internet [17]. Because of NAT, multiple end-users or clients can access the Internet via one public IP address. Researchers reserved the private IP ranges 192.168.0.0/16, 10.0.0.0/8, 172.16.0.0/12 for internal use [83]. Table 2.2 shows these IP ranges. When a device is behind a NAT, it receives a private IP address from the router. Normally, is a DCHP server is responsible for handing out IP addresses on the private side of the network. Those are leases and temporarily assigned. It is possible that when a device reconnects to the network the next day, it will receive a different IP address. A connected device can, for example, receive an IP address like 192.168.1.11. Traffic is then routed via the router, and when it accesses the Internet, it uses the IP address (a public address) of the router. This way, when a reply is sent to that specific device, it is first transferred to the router, and then the router forwards the information to that specific device (192.168.1.1) [83]. With NAT it is thus possible to connect multiple devices with an internal IP address to the Internet via one public IP address. See Figure 2.5 for a graphical illustration.

Besides the ability to have multiple devices connected to the Internet via one public IP address, NAT has an additional security benefit: IP masquerading. It hides the local subnet, consisting of private IP addresses to the outside world. This increases the privacy users experience when browsing the Internet. Conversely, there are downsides to NAT, because the Internet was developed to work with a peer to peer connection. The sender fills in the destination IP address, and the packet knows its final destination. However, with NAT, the IP address given to the packet by the sender is not the final destination but the public IP address (router's IP) known by the sender. There could still be a thousand different clients behind that one IP address. Currently, routers are responsible for changing that destination IP from public to private, but they were not developed for this purpose. This method could lead to the latency of a network connection [84]. IPv6 was amongst others developed to restore the original way the Internet was meant to operate, as such avoiding this problem. The IPv6 space will be so ample that every device again can receive its own publicly routable address.

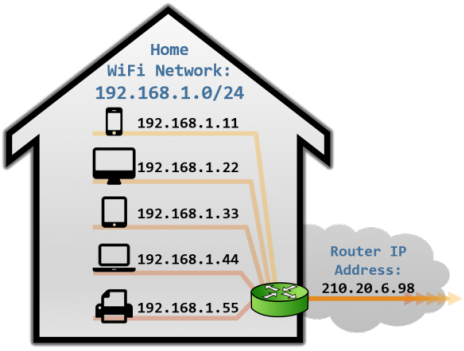


Figure 2.5: Graphical illustration on how NAT works. The public Address 210.20.6.98 is translated to the different private IP addresses like 192.168.1.11.[5]

2.9.5. Carrier-grade NAT

Another method to increase the lifetime of IPv4 is Carrier-grade NAT (CGN). Carrier-grade NAT is a similar method to NAT in private networks, wherewith the aid of network translation IP address can be used multiple times in different *zones*. The RFC for CGN was released in 2012 with as the main reason to continually support IPv4 until IPv6 is fully deployed. The difference is while with normal NAT, the translation is made at the end-users home or office, in CGN the translation is performed in the private IP range of an ISP. The IP range used by ISP's is 100.64.0.0/10 [85]. This range of IP addresses can thus exist at every ISP, which again greatly increases the amount of available IPv4 addresses. However, there are downsides to CGN; it again breaks the end-to-end idea of the IP protocol. Moreover, hosts in a GCN will not be directly routable via the internet and end-users cannot port-forward in edge routers. For the average user, CGN could be a solution, but for network enthusiast and companies, it is not desirable. As mentioned in the RFC, it could increase the lifespan of IPv4 but will not be a solution to the need for public IP addresses and the IP exhaustion [36]. A visual presentation on CGN is shown in Figure 2.6.

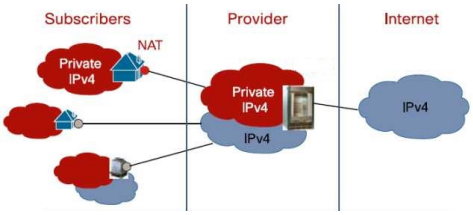


Figure 2.6: Carrier-grade NAT. A similar process as normal NAT but then at ISP level rather than at a home location.[6]

2.10. DNS

DNS or Domain Name Server is a decentralized naming system for computers and services on the Internet. It is developed and implemented around 1985. It is comparable to a phone book for the Internet as it links domain names to the corresponding IP addresses. The downside of IP addresses is that they are hard to remember for people and domain names are not. For this DNS was developed, as a DNS server links a memorable domain name to an IP address through a DNS lookup. Organizations and people can register a domain name (example.com) at a DNS registrar which are responsible for maintaining the phone book. In this database, they link an IP address to a chosen domain name. These organizations can also modify or delete entries into the central registry [60].

There are multiple DNS lookups, but two are of interest for this research: an 'A record' lookup that returns a 32-bit (IPv4) IP address and an 'AAAA record' lookup that returns a 128-bit (IPv6) IP address [98]. A schematic DNS lookup can be seen in Figure 2.7. The reverse process is also possible, in this case, the IP address is

[5]https://www.practicalnetworking.net/series/nat/why-nat/
[6]https://networkingnerd.net/2012/01/05/double-nat-nat/

known, and the corresponding hostname is requested from the central database. This action is often referred to as rDNS [34].

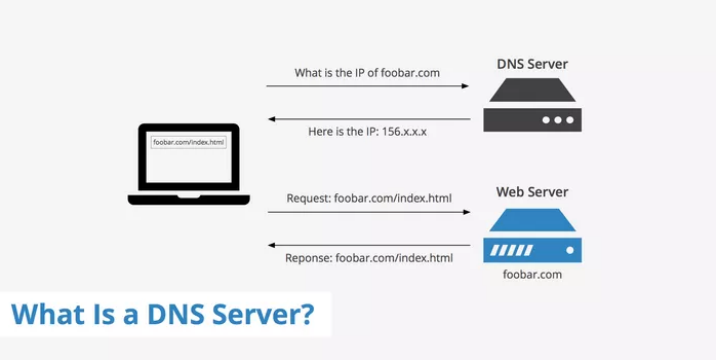


Figure 2.7: Schematic overview on how DNS works. A client requests the IP address of a domain name, and the DNS server responds with the corresponding IP. The client uses that IP address to request the required data from the web server[7]

2.11. IPv6 transition protocols

There are other protocols developed for a successful migration from IPv4 to IPv6. Thus far only dual-stack has been discussed but, next to dual-stack tunneling protocols have been developed. Tunneling protocols are a mechanism that aid network engineers to migrate from IPv4 to IPv6. Tunneling is a method that provides the existing infrastructure to support IPv6 traffic. At the moment is it of importance for successful transmission to IPv6 the existing IPv4 infrastructure remains functional while IPv6 is slowly deployed in a network. Before some tunneling protocols are addressed, some information around encapsulation will be elaborated.

Encapsulation is a process that happens within the TCP/IP stack where headers and tails are being added to some data. An application starts with some data, and through every layer in the TCP/IP stack, information is added. For example, sending an e-mail over a network. It starts that an email is sent via the application layer to the transport layer. The transport layer adds then its data like which port is being used and further pushes the packet to the internet layer. In the internet layer, more information is added to the data like the source address and the destination address. It is then further pushed through to the network layer, which also adds information to the packet. It is then finally sent through to the physical network link to its destination. In every layer has the packet a specific name:

1. **Frame** Encapsulated data created in the Network Access layer
2. **Packet** Encapsulated data created in the Network Layer.
3. **Segment** Encapsulated data created by the Transport layer

Every protocol uses its own language and has its own way to build up a packet. The reason why IPv4 and IPv6 are not compatible is that the packet headers are different between the two protocols. However, there are protocols that translate packets from IPv4 to IPv6 and vice versa.

6to4 6to4 is a protocol that is being used to transfer IPv6 packets over an IPv4 only network, usually the internet. 6to4 is essential for the initial phases of IPv6 support but was only developed as a transition method, and not to be used all the time on the internet. The advantage of 6to4 is that it does not require explicit tunnels [65].

Teredo Teredo is a protocol that gives IPv6 connectivity on an IPv6 enabled host in an IPv4 network only. It can work from behind a NAT on IPv4 [97].

ISATAP ISATAP is a transition mechanism for transmitting IPv6 packets between dual-stack devices on an IPv4 running network. It creates an IPv6 address based on the acquired IPv4 address [106].

[7] https://www.keycdn.com/support/what-is-a-dns-server

2.12. Disabeling IPv6

Often, system administrators disable IPv6. They tend to do this because of their lack of knowledge. However, there are both security and business implications as to why disabling IPv6 is a bad policy. As a technology, IPv4 was developed almost 40 years ago for another purpose than what we use it for today. It was not intended for the current size of the Internet, and there was no need for security. IPv4 was developed to operate in a network of researchers who knew each other. There was no distrust between the communicating parties, which is not valid with today's size of the internet. IPv6 resolves some of the IPv4-issues related to the Internet's size. First, some advantages of IPv6 are summarized.

1. The number of IPv6 addresses is much larger than the number of IPv4 addresses. There is no need for NAT anymore, increasing the speed and robustness of a connection.

2. Within the IPv6 protocol is IPv6 itself responsible for handling the fragmentation of the packets instead of the router. The protocols used in IPv6 are better suited for this than a router.

3. By default, IPv6 is encrypted using the IPSec algorithm. The information in the header and packet have improved security when transferred. A header of an IP packet consists of a fixed amount of fields that are being used to add the required information for reliable transport about the packet. For example, the source and destination address are part of the header. That data from the sender is added after the header in a packet

At this moment, researchers are looking into security implications and migration problems concerning IPv6. Researcher at the Cooperative Cyber Defence Centre of Excellence (CCDCOE), showed that it is possible to set up so-called covert channels on IPv6. A covert channel is a process where a communication bridge is set up between two systems that is not supposed to be allowed by the security policies. In this case, communication which is normally blocked is still possible via this covert channel. Normally the network traffic is routed via a firewall which inspects the incoming and outgoing packages. With a covert channel, packages between two systems directly communicate with each other, whereas this should only be allowed via a firewall. Often security engineers develop tools that make use of these vulnerabilities to show what the possibilities are. In case of exploiting covert channels is *v00d00N3t* developed. With this tool, they were able to set up a covert channel over ICMPv6 [7]. ICMPv6 is a protocol that is used for error reporting and not communication. Software called Network Intrusion Detection Systems (NIDS) monitors network traffic. Normally, NIDS software can detect these breaches. However, such tools are generally not compatible with IPv6 and thus cannot detect the malicious session over IPv6. Snort is a well-known software package and often used NIDS that supports both IPv4 and IPv6. However, a paper by Kumar et al. concluded that Snort is not fully configured to detect all potential IPv6 security events. Such configurations are referred to as detection rules and are lacking on IPv6 [51].

Thus far security implications around IPv6 enabled networks and devices have been discussed. However, there are also security implications when IPv6 is disabled. When IPv6 is half configured or not properly disabled, it becomes possible to set up a malicious DCHP server by a hacker, using a tool that gives clients within an IPv4 network an IPv6 address. A hacker has then the ability to change DNS settings over IPv6 to redirect users to potentially malicious websites. Most client devices prefer to use IPv6 instead of IPv4. A hacker can exploit this setting to enable IPv6 on a network from a malicious device and start making connections over IPv6. If this is not properly blocked/disabled users could be communicating over IPv6 without even knowing. Thus, disabling IPv6 within a network can also lead to security implications [35].

3

Methodology

This chapter elaborates on the methodology that is used for this research. It will discuss the setting and the different approaches used in this study. Furthermore, arguments are provided about the decisions made about the collection and processing of data. This qualitative research focuses on non-numerical data. The focus lies on the *why* rather than the *what* of social phenomena. Examples of qualitative research are interviews, case studies, observations, and action research [96]. Its counterpart, quantitative research, focuses on logical and statistical procedures. In that instance, the focus lies on numbers and is mostly theory proving.

For this thesis, qualitative research has been executed by performing interviews as the main source of data collection. The phenomenon that will be addressed will be the IPv6 configuration. Previous research executed by van der Eijk et al. [24] found that there are configuration differences between open ports on both IP protocols. This research elaborates on this by asking the question *why* these differences occur and what the route cause is of these differences.

In research are two types of studies. Researchers often categorize studies in observational and experimental studies. The first one, observational, is a study were researchers draw samples from a population and where treatment is not under experimental control. In experimental research, an intervention is introduced by the researchers, and they study the effect of that intervention [87]. This thesis performs an observational study because no treatment has been performed on the subjects. In-depth can be said that this is a cross-sectional study. In comparison to a longitudinal study, data is collected at one specific point in time. In a longitudinal study, data is collected over time, mostly referred to as before and after [81]. The goal of this research is to find differences in configuration between open ports. It is not to test the effect of increasing awareness concerning IPv6.

Certain steps have to be followed to perform adequate qualitative research and to answer the *why* as mentioned earlier. Making this research an exploratory study. Therefore, it will not be possible to give conclusive evidence. Rather this research will search through the topic at different levels [100]. Nor is it possible to do thorough quantitative research because no conclusive literature has been published yet about the configuration differences between IPv4 and IPv6. Thus far, only Czyz et al. [19] published an article about IP configuration. However, they focused more on the technical aspect and not the managerial processes behind port configuration. For this study, interviews were conducted with security experts and system administrators. The interviews can be further split up into the sampling, data collection, and data analysis. In the following subsections, first, these methodologies will be defined, and finally, ethical considerations are discussed.

For this study, a semi grounded theory approach was used. Because of its nature of a theory-building and exploratory research. Grounded theory is a research approach where theory building and theory verification go hand in hand [38]. Thus far there are little theories concerning IP configuration on dual-stack hosts. An advantage of collecting data and directly analyzing, it becomes possible to generate theories close to reality. Furthermore, it gives a better understanding of the process with regards to personal experiences of system administrators. For this, research methods will not be carved in stone beforehand but adjusted to the finding during the progress of this study.

3.1. Interviews

Interviews are a common way of data collection in qualitative research. An interview is a method to collect data about the decision-making process around IPv4, IPv6 configuration at companies. By performing interviews, information can be gathered about the context. Furthermore, interviews are a method to analyze a person's behavior in their context [95]. An interview supplies up to date information directly from the source and professionals in the field. In an interview, primary data is gathered by the interviewer. Primary data is data that is collected by the researchers first-hand. This data is in contrast with secondary data, where the data is collected from already available data sources (e.g., literature) [2]. It was decided to perform interviews because no conclusive facts about IP configuration could be found in current literature. Furthermore, this research is similar to research performed by Dietrich [22] about security misconfigurations. That research focused on investigating the engineers perspective on security misconfiguration. Here they also approached the human component around misconfiguration. This research is similar because a difference between IPv4 and IPv6 configuration could be seen as a misconfiguration. In Dietrich et al. the focus was on misconfiguration in general while for this research, the focus to port configuration. For this, a similar approach was taken in this study.

During an interview, there are two individuals, the interviewer and interviewee or respondent, as he or she is frequently named. The interviewer is responsible for logically extracting the data from the interviewee. There are different styles of interviews to do this, but in research, three different styles are frequently used; the structured interview, the semi-structured interview, and unstructured interview [95]. For this study, it was decided to perform semi-structured interviews because this exploratory study has a clear goal of what data must be gathered from the interviewee. This research uses a script with sample questions that are asked during an interview. However, the question will be phrased open-ended, so there is room to deviate from the script. Furthermore, because of the semi-structured nature, it is possible to ask additional questions when relevant topics were touched. This methodology broadens the interviewer's possibilities to go into more depth during an interview.

There are multiple methods to perform an interview, while in this study is focused on telephone and physical interviews. Both of these methods have advantages and disadvantages. An in-person interview can often lead to better result because it is easier to communicate verbally and even possible to communicate non-verbally. However, they are often more expensive and time-consuming than telephone interviews [95].

An interview should always start with an explanation of the respondent's role in an interview. In this introduction, the purpose of the study should be clearly explained. Furthermore, interviewees rights must be clearly stated by the researcher and accepted by the respondent. A researcher does not need to be biased during an interview. Only after the evaluation of an interview can it be determined what is important and what is not for general research. Therefore, the researcher should always be careful in steering an interview. This why the interview should be minuted and if possible, recorded. Then, it should be transcribed for processing and so that no minor detail is overlooked. A downside is that not all potential respondents are willing to participate in a recorded interview. As a consequence, it becomes harder to find suitable respondents. Furthermore, these respondents will generally be only willing to participate with interest in research and IT configuration results in a bias in the sample [41].

Finally, the sampling procedure is elaborated. To understand which respondents are approached to participate in this study, the unit of analysis and unit of observation need to be discussed. The unit of analysis identifies and localizes the data to be collected. Also, there is a strong correlation between the unit of analysis, the collection methods, and sample size [96]. Depending on the used unit of analysis and collection method, the sample size is determined. In this case, the unit of analysis is individuals who perform IP configuration on internet-connected hosts. From this point, the unit of observation can be determined. The unit of observation will be based on a data set enclosing configuration differences gathered by the researchers of the University of Amsterdam [24]. Here, the unit of observation encloses the owners or responsible IT engineers of the servers that are listed in the data set.

A final concept that should be discussed is saturation. Saturation is used in qualitative research as a criterion for ending the research. Saturation means that no additional data is being collected. For instance, saturation can be reached when a researcher sees similar instances over and over again [90].

3.2. Structured Approach

The previous paragraph describes the method used in this study. This section will discuss steps performed in this study and elaborates on methods applied in this study. This study is based on the work of articles published by Czyz et al. [19] and van der Eijk et al. [24]. These studies mapped the current landscape of IP configuration on dual-stack hosts. However, both parties approached this from a technical perspective but did not look at the human part of IP configuration. However, there have been studies performed to the human side of security misconfiguration in the past [22]. In this study, we tried to combine the human perspective of misconfigurations with the previously performed research of the technical side of IP configuration. We did this by first search for literature on both topics. However, we found that there is no conclusive literature yet that combines both topics. So, the data gathered about dual-stack host configurations by van der Eijk was used to identify and target potentially misconfigured devices. With this data, it was possible to recover owners or responsible parties of misconfigured hosts. By executing interviews with these owners or parties, we were able to determine if the difference between open ports on IPv4 and IPv6 were made unintentionally. Finally, if the configuration was made unintentional, were in the process they thought the misconfiguration slipped in. In the current landscape, interviews were the best applicable method for data collection. The reason why we chose this will be further elaborated in the upcoming paragraph.

3.2.1. Qualitative Research

For this study, a qualitative approach with interviews was chosen. There are several reasons why the decision was made to use interviews as the main source of data. First, this study is exploratory. Before this study, there were no conclusive theories about the configuration of dual-stack hosts. First theories had to be drafted based on gathered qualitative data before these theories could be tested with quantitative data. However, it has been proven difficult to gather data concerning vulnerabilities. An IP misconfiguration could be considered as a vulnerability in a system. For example, a study by Li et al. [55] shows that a majority of contacts did not respond to their notification of a vulnerability. Furthermore, even if they did remediation was often only partial. Next, they found that direct contact is the best approach for notifications. However, this is an extremely time-consuming process because every host has to be handled uniquely, and the process cannot be easily automated with scripts. A second study by Jhaveri et al. [48] discusses a model for abuse reporting and state that the Internet has a highly fragmented ownership structure. At the beginning of this study, it was not clear what parts of the Internet infrastructure would be most applicable for data collection. The different parts of the Internet infrastructure are in this case considered the ISPs, hosting parties, companies connected to the Internet, governmental institutions, or Internet governance agencies. Within mind that there was no clear theory drafted yet, the low response rate on vulnerability notification, and unclear target group it was not seen practical to collect the needed required of quantitative data in the given period. First, a clear focus group had to be determined, and theories had to be drafted before qualitative data could be collected.

3.3. Ethical Considerations

Finally, the ethical consideration concerning this study will be elaborated. Ethics are important in technology and research because it is important to think of the consequences a technology or research can have. During this research, emotionally driven people are being investigated, and this is why the research should follow a set of guidelines. This statement is especially true in qualitative research [108]. For this study, the ethical considerations are split up into three sections; confidentiality, informed consent, and harm [72].

3.3.1. Confidentiality

Confidentially is important in most studies. The interviewee's personal information can be required, but it is not ethical to neglect to be cautious with that information. Privacy is linked to this subject. The information that is made publicly available should not be sufficient to recover the data sources used in the research [72]. Next, great care must be taken with the gathered data set with regards to the privacy and confidentiality of respondents. Although the content of the data set with IP configuration is publicly available, it is gathered in one database creates an easily accessible source for malicious intent. No targets which can be exploited are identifiable based on the data set, but it shows multiple entries that are potentially vulnerable. Normally, in case of malicious intent, a port scan is the first step to see which hosts are vulnerable and what protocols could potentially be misused. Moreover, port scans can be detected and leave traces, which could identify a malicious actor is trying to break in.

There are discussions about the legality of port scans, and it is in a grey area if port scans are allowed without proper consent [12]. Nevertheless, the data set must be handled with care. It should not be made publicly accessible or shared with third parties, even though the data set contains only publicly available and no private information.

3.3.2. Informed Consent

Informed consent focuses itself on giving the required information to the interviewee before the interview started. It is important to make the purpose of the study clear to the respondent, and they agree with the study. Furthermore, the respondent must agree to their information being used in this study [72].

3.3.3. Harm

The final ethical consideration is harm. During the research, it is vital to make sure no harm will be done to the respondent or researcher. In this study, changes are minimal that there is any physical or mental harm. However, one subject should not be overlooked, the fault identification process. Fault identification process entails that faults are identified and subjects are confronted with these potential mistakes. During this study, potential, configuration mistakes could be identified. Generally, people are not willing to cooperate if they are identified and correct them on their mistakes. Thus, there should be a balance between pointing out errors and gathering research data [72].

4

Collecting Qualitative Data

This chapter will elaborate on all the steps that are executed in this study. This study is part of a more extensive study, and different steps were performed by others or in collaboration with other researchers. Simply stated are there three phases in this study. The collection of quantitative data, the collection of qualitative data, and the analysis of the qualitative data. This chapter will discuss all these different steps and which the researcher was responsible for each part. Before the quantitative data collection phase, there was knowledge about a possible difference in IP configuration on dual-stack hosts. Only after this phase numbers were available about the current dual-stack configuration landscape. However, this not yet gave the answer *why* there was a difference. To answer this why, qualitative data had to be collected, the second phase. This method was based on identified hosts from the first phase. Finally, this data was transcribed and analyzed in the final phase. These phases will chronologically be discussed in this chapter. Moreover, it will describe what work had to be executed in every step and by whom.

4.1. Vulnerable Host Identification

This research is based on a data set gathered by two researchers from the University of Amsterdam. This data set consists of all hostnames, corresponding IP addresses, and results of their scans. Their method will be briefly described in this chapter. Additional information can be found in their paper [24]. Their first step was the collection of hostnames from dual-stack hosts. For this, they used three different data sources; Alexa top 1 million [3], Rapid7 FDNS ANY [82], and IPv6 ICMP [36]. In the end, most data was gathered from the Rapid7 project. Finally, an entry was only useful if it comprised each element: a hostname, IPv4, and IPv6-address. Missing records were added with reverse DNS queries to find the corresponding hostname.

The next step by van der Eijk et al. was to scan the found IP addresses. The correct term is a network scan or port scan. It determines if there are any open ports on a system. During a port scan, a specific network package is sent to a port. If the protocol running behind that port responds, the port is considered open. Traditional port scanners like Nmap take a long time to scan larger subnets on the Internet. However, with the development of ZMap in 2013, it became possible to reduce this time greatly [23]. With ZMap, it is possible to scan both IP protocols (IPv4 and IPv6). The following well-known ports were scanned: FTP, SSH, Telnet, SMTP, NetBIOS, BGP, HTTP, HTTPS, SMB, IPP, MySQL, RDP, VNC, Redis, Elasticsearch, MongoDB, NTP, SNMP, DNS, Memcached. All the scan data, together with their corresponding hostnames, were added in an SQL data. SQL stands for structured query language and is a method to communicate with a database. An example of a scan overview can be seen in Tables & and &. This method made it possible to search and analyze the data within a reasonable period. Altogether, this resulted in a database that included tables for Hostnames, IPv4, and IPv6 addresses. This database was the final contribution by the previous researchers and the data was handed over for this study.

In our case, a Virtual Machine in an ESXi environment was used for analyzing the data. It was running on a Ubuntu 18.04 LTS on which MySQL was installed. This machine was not accessible via the Internet because of the potential hazardous data that was in the data. Initial imports of the data failed because the data was corrupted. After multiple attempts, the import of new data was successful. Most work on the dataset was

executed via CLI (command-line interface) from a local network. Table 4.1 gives some facts and figures about the data. It will give a brief illustration of the consistency of the dataset.

TLD	Total	Correct Config	Incorrect Config	Percentage Incorrect
Total	22.514.910	21.432.137	1.082.773	4,8%
Alexa	93.497	91.236	2.261	2,4%
.arpa	73	73	0	0,0%
.com	7.311.192	7.010.569	300.623	4,1%
.int	37	37	0	0,0%
.edu	69.939	68.710	1.229	1,8%
.eu	361.155	338.905	22.250	6,2%
.gov	8.194	8.163	31	0,4%
.mil	10.781	10.779	2	0,0%
.net	1.512.595	1.448.107	64.488	4,3%
.org	676.716	642.992	33.724	5,0%
.ac.uk	8.194	8.094	100	1,2%
.co.uk	436.747	427.093	9.654	2,2%
.gov.uk	114	108	6	5,3%

Table 4.1: An overview of popular top level domains (TLD) and the percentages of proper configured and misconfigured servers.

Before the data was analyzed, different configuration scenarios where determined. Table 4.2 shows the possible scenarios the configuration could occur. We looked at if it was an assumable configuration from a security perspective. It is possible that from an availability perspective, other assumptions are made, but from a security perspective should their not be a difference between both protocols. However, there are millions of configuration which would allow for an exception. The large size of the data, as seen in Table 4.1 meant that a selection had to be made. For this purpose, it was decided to first focus on .nl and .com domains. It was assumed that this would increase the chances to find English or Dutch language subjects. Next, the selected data were extracted to CSV format, meaning the data was more easily searched through and shared between researchers. The data were selected based on the following criteria:

1. Hostname must end on .com or .nl
2. Pingable on both IPv4 and IPv6
3. At least one port configuration was different between the two IP protocols. For example, port 22 was closed on IPv4 and open on IPv6.

The extract from the SQL data consisted of the hostname, IPv4, IPv6, and open/closed ports on IPv4 and IPv6. The next step was to analyze further the data and tag ports that differ between both protocols. These differences are finally summarized together with the total open ports on IPv4 and IPv6. The files were converted to .xlsx format.

IPv4		IPv6	Logic
Open	=	Open	Normal
Open	>	Open	Normal
Open	<	Open	Unusual
Open		Closed	Normal
Closed		Open	Highly Unusual

Table 4.2: A logical, from a security perspective, explanation about possible scenarios of port configuration.

4.2. Ownership Identification

The result of the previous phases were a .com and a .nl xlsx format files with a top-level domain, whether it is reachable over ICMP (pingable), and the results of the scans. Only hostnames with a complete of information were used. Thus, the .nl data was reduced from 1.600.891 to 57.613 entries and the .com from 7.311.192 to 293.884 entries. These numbers of entries were more manageable than the 22 million entries in the initial data. A primary analysis of the data revealed that multiple hostnames resolved to the same IP address. Showing that a web server can run multiple websites on one web server. HTTP request form a browser to a

web server include the hostname, and with this hostname, the web server can find the correct corresponding data. For this research, the interesting part was the layer 4, 5 configurations, and not in the websites running on a server. Hence, we targeted unique IP addresses and did not look at the hostnames that resolved to that corresponding IP address. The 'remove duplicates' function of Excel allowed to reduce the final number of entries in the .nl data to 2.964 and the .com data to 32.655. Finally, the data was sorted on the number of total ports open on IPv6. It was assumed that the more ports are open, the higher the risk of exposure, although this assumption could be considered to be open for debate. Another method to analyze these data would be to sort the database based on the number of port differences. In any case, the final step was to subtract any entry where no open IPv6 port was found. This measure does not indicate that there are no single IPv6 ports are open as only a small selection of ports was scanned. In theory, there could still be large security implications through one of these unscanned ports on those servers.

During the next phase, hosting information had to be recovered from the potentially misconfigured servers. There were two methods possible to find information about these hosts. Most registered domains in the data resolved to websites. When ports 80 or 443 were open, chances were that there was a web server responsible for hosting a website. This setup could easily be checked by just entering the hostname into a web browser and see if the hostname resolved and a website loaded. Another possible entry was, e.g., mail.example.com. In this case, the server was likely a mail server based on the registered URL. If port 25 was open on this device, it was almost certainly an email server. When the complete hostname would be used, the URL would resolve directly to the mail server. This server would not have any HTTP/HTTPS port open. Thus, it would not give any feedback when inserted into a browser. However, when only the second-level domain was used (mail.**example.com**), it would often be possible to resolve, and the company's owners website would be returned. In this case, it was assumed that the same owner held the second-level domain as of the subdomain. The website that was then loaded would often provide the owner's contact information which could be added to the spreadsheet.

Another method to find information about websites is with the aid of WHOIS queries. Whois is a protocol to find information about the domain name or IP address from a global database. With the aid of a query, the database will respond with the contact information like the owner, ISP, or DNS registrar of the entered IP address/hostname. This lookup can be done via the command *whois* on de command line, or with the aid of websites. In most cases, `https://www.ip-tracker.org/` was a helpful website to do these queries. Here the ISP was added as contact information to the Excel sheet.

This gave the research team two different methods for gathering contact information. A human interpretation was needed to determine who was responsible for the host and the configuration of that specific machine. Furthermore, there are differences if the server was maintained at a home location or in a large server park. Depending on the host and location, different parties can be responsible for the server or website. Before the analyzes, five different scenario's were determined.

1. A server at a home location maintained and configured by a computer enthusiast.
2. A server at an office location maintained and configured by an individual or a small team.
3. A server at a vast office location maintained and configured by large professional teams.
4. A server located at a hosting party, dedicated for web traffic configured by the hosting party.
5. A virtual private server (VPS) at a hosting party.

In this fourth case, the hosting party is generally responsible for maintaining the server and blocking traffic to the hosts. In other cases, the person him- or herself is responsible for the server. Hence, in cases 1, 2, and 3, the DNS registrar only links the hostname to the IP address. ISP only manage the internet connection but have no configuration rights on the server itself. It is even more problematic to determine the nature of a hostname associated with a hosting provider. In those cases, the hosting party is responsible for the total infrastructure, but the person who rented the system (e.g., VPS) has full access to install applications and open or closed ports.

The researchers needed to determine which of the five scenarios applied to every case. This evaluation was

up to their interpretation, but some guidelines were drawn to improve the decision-making process. For instance, if the ISP was, e.g., Ziggo, KPN, XS4All, it was likely that the server fell under the first case enumerated above. These ISP's focus on providing Internet to personal users and do not have many corporate clients. Another guideline was if multiple distinct websites were running on the same server, it was probably hosted at an external party. These guidelines simplified the analysis of the .nl data because of the experience and knowledge the researchers had from Dutch internet providers. The final goal was to find a corresponding email address or phone number from a person who was responsible for the port configuration process. This contact information was added to the excel sheet.

4.3. Internal IPv4/External IPv6 Configuration

This section will continue on another slice of the data analysis part. During the initial analysis of the data, other remarkable findings were discovered in the public DNS record. The initial data consists of hostnames and their corresponding IP addresses. A hostname resolves according to a correct configuration to a public IPv4 and a Public IPv6 address. However, in the public DNS record, private IP addresses are registered. The Chapter' Background Knowledge' specified that public IP addresses are accessible via the Internet and that private IP addresses are behind a NAT solution, entailing that public IP addresses are unique while, in different NAT solutions, the same IP address can be present. Normally, a DNS record consists of an A record (IPv4) and AAAA record (IPv6) that resolve to IP addresses of the format *93.184.216.34* in the case of an A record and to the format *2606:2800:220:1:248:1893:25c8:1946* in the case of AAAA. These are both public IP addresses. However, in the data examples were found that the A record resolved to a private IP like *192.168.1.1* while the public IP address was an accessible IPv6 address over the Internet. These IP addresses were active over ICMP (ping) and could even be scanned by the scanners. These differences are possible because of different characteristics between the IPv4 and IPv6 protocol. NAT was a solution to increase the lifetime of IPv4 that is not implemented in IPv6. By default is there thus no NAT in IPv6 and in the IPv6 protocol networks do not have a private IPv6 address [14]. There are scenarios that it is possible, but not by the default. There is a possibility in IPv6 to use a private address (Link-Local), but this is only used for routing within a network. It cannot be used to access the Internet. It is automatically configured by the device and replaced as soon as it is connected to the Internet.

However, from a security perspective is combining a private A record and public AAAA record an illogical up. If a computer would try to access a domain name with an internal IPv4 configuration, it could be directed to unwanted locations. Because it is private, it routes to an internal IP, so it directs to different devices depending on which private network a person is. In larger networks with many private IP addresses, it could be possible that the same private IP address is used as registered in the public DNS A record. This machine would then be directed to an unwanted location what could have undesirable consequences, for example, accessing information that was not intended. In a smaller environment, the chances are small that the same internal IP address would live twice and the connection would time out. This setup does not have security implications however this up is not desirable

The previous section discusses the security implications of a person who tries to resolve from an external network, which does not necessarily have large security implications. However, from a security perspective, this is presumed not to be the case for the IPv6 address. A device with an internal IPv4 IP address is part of an internal system and is not supposed to be connected directly to the Internet. In-home networks, a NAT solution is often used as a firewall [112]. With a NAT, it is possible to implement a *default deny* policy. This policy blocks all unsolicited incoming packages before it can enter the internal network. Internet-connected machines cannot address devices that are connected in the private IP range behind a router. These devices are only accessible from the Internet when the appropriate ports are open, and the packages are forwarded to a specific device behind a NAT. A use case for this could be storage that is required to be accessed from multiple locations. In that instance, a public port on the router (e.g., 5500) will be forwarded to a samba share (445). Port forwarding is used to make services that are behind a NAT publicly accessible. Figure 4.1 gives an example of port forwarding. It may look like that a firewall and NAT are the same, but there are differences between the two. A firewall has more options to analyze network traffic. While a NAT only blocks packages, a firewall will also look inside packets for analysis. Furthermore blocking custom ports or IP addresses is easier to implement in a firewall.

So why is it different in IPv6? In the case of IPv4, a person deliberately configured the router to access the samba share from outside the network. If no port forwarding were applied, the samba share would not be accessible from the Internet. This scenario is likely in home and small company network storage. However, because of the difference in protocols (IPv4, IPv6), IPv6 allows the samba share to be directly connected to the Internet and to be accessible from anywhere in the world since a normal configured IPv6 network (no NAT) is directly linked to the Internet. See the graphical display of the difference in IPv4 and IPv6 routing in Figure 4.2. In IPv4, all devices have the IP address 123.123.123.123 when accessing the Internet while in IPv6 (red), they all have a unique address with the same prefix (first 16 digits). However, the IPv6 space is immensely larger than the IPv4 space and is thus harder to scan for vulnerable hosts by malicious actors. Chances to identify vulnerable hosts on IPv6 with port scans are abundantly lower than on IPv4. This case is not true when system administrators register their IPv6 address in the public DNS record. According to Netcraft [69], in May 2019 there were *235,011,143* unique domain names. Tools make it possible for malicious actors and researchers to retrieve the IPv6 address behind these domain names easily. This number is much smaller (3.4×10^{38}) than the total number of IPv6 addresses, making it more feasible to scan them all.

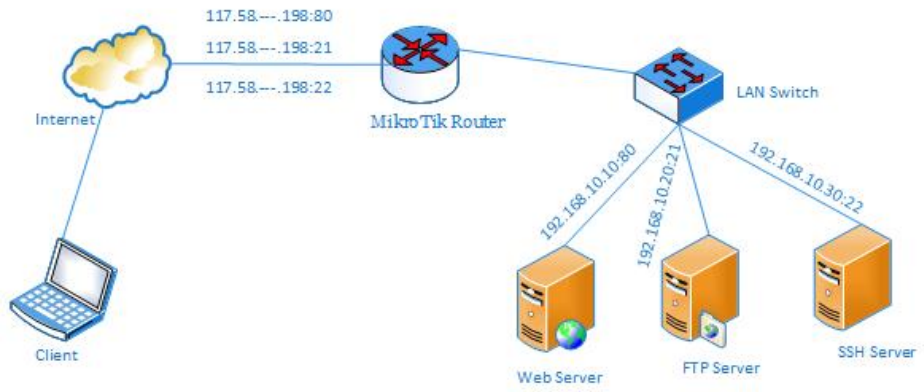


Figure 4.1: The port forwarding principle. A client connects to the router on, for example, port 80. The router then knows to forward the port from port 80 to the web server. An equivalent happens for port 22 on the SSH server.[1]

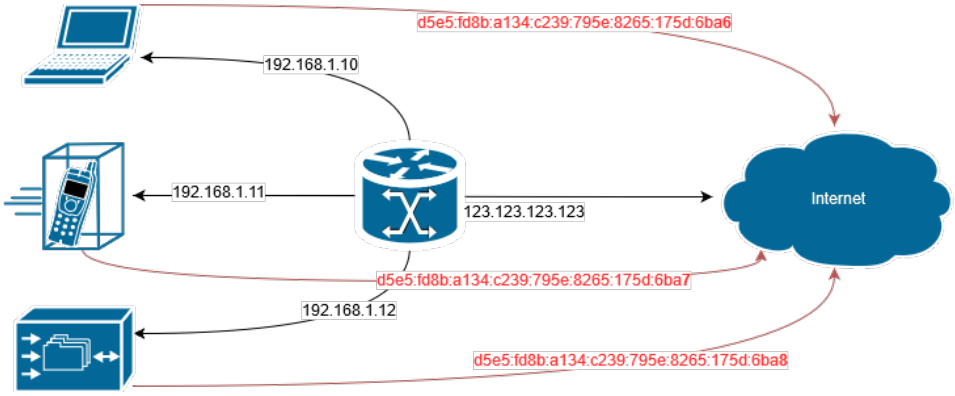


Figure 4.2: Graphical representation of the differences between IPv4 and IPv6. In IPv4 (black) all devices connected to the router access the Internet via the same IP address (123.123.123.123) While in the IPv6 case (red), all three devices access the Internet with their unique IPv6 address.

As a summary of the previous section, certain security implications can occur when an internal configured IPv4 address is added to the public DNS. For this, the total data was searched for registered internal IP addresses (the three private IP ranges are 192.168.0.0/16, 10.0.0.0/8, 172.16.0.0/12). To summarize, the SQL had the following requirements:

[1] https://systemzone.net/mikrotik-port-forwarding-using-winbox

1. ICMP alive on IPv6

2. IPv4 address 192.168.xxx.xxx and 10.xxx.xxx.xxx.

For selecting these private IP addresses, it was chosen to exclude 172.16.0.0/12 private addresses in the data extract because it was impossible to write an SQL query that would exclude the public IP addresses. However, visual inspection from the 17.16.xxx.xxx range did not identify any .nl domain names.

The extract resulted in 6792 unique domain names. This extract comprised the hostname, IPv4, IPv6, and open ports on IPv6. IPv4 ports were not extracted because no internal addresses are accessible via the Internet. In Excel, the top-level domain was extracted in a new column to make searching per country more straightforward and easy sum functions were added to calculate the total number of open ports.

The next phase for this extract was to determine the ownership of the hosts. In contrast to the previously described ownership identification phase, this one is more complicated. A quick analysis of the data revealed that most of these connections fell into scenario one from the previous section, A server at a home location maintained and configured by a computer enthusiast, making ip-tracker.org an unreliable source for ownership identification. Furthermore, all IPv6 traffic was blocked at the physical location where the research was executed. Thus, it was impossible to resolve the hostnames and see if any services were running on port 80, 8080, 443, and 8443. Two different methods were employed to overcome this situtation. First, the domains were stripped of their sub-domains, and only the second-level domain names were resolved via a web browser. In some cases, the second-level domain resolved over IPv4 to an external address, and information could be gathered from that website. Here, it was presumed that the owner stayed consistent between the second-level domain and sub-domains. Secondly, the domain was looked up via sidn.nl. This organization is responsible for the management of all domains in the .nl top-level domain. Behind every registered domain name, some information is publicly accessible about the organization that/person who registered the domain. The registered email address was used for further analysis.

4.4. Recruitment

This paragraph will discuss the recruitment process performed. At this stage, there is data with parties that could be contacted to ask if they wanted to participate in this research by taking part in an interview. It was decided to contact parties over the telephone from the gathered information in the data. The article written by Stock et al. was used as a reference to up the recruitment and interview part. They provide information about how to successfully notify vulnerabilities on the web [105]. Stock states that trust and the ability to reach the correct contact is the hardest challenge.

The data was sorted on open IPv6 ports in the previous phases. Furthermore, was it decided to start contacting parties from the top because they were likely to be more interesting targets. Assumptions were made that the higher the difference in open ports, the higher the risk. However, this did not take into account that some open ports have more implications on security than others. In any case, the data was too large to find the most relevant cases by hand. Moreover, a spreadsheet listed all parties that had already been contacted.

Three researchers performed were involved in establishing the initial contact. After the initial contact, they logged the progress for each of these companies into different categories. The different categories were e.g., not interested, interested, call back later, or email sent. For consistency purposes, each call was performed according to a previously established script about, the research because not all researchers had a full understanding of the technologies used (TCP/IP, DNS). This script can be found in the appendix of this report: E. This script also served another purpose. It helps to avoid any similarity to a phishing call like the well-known one by 'Microsoft Tech Support' about a problem with a computer [77]. In that spirit, the script accentuates that students from the TU Delft performed the research. To perform these calls, Cisco Jabber was used in combination with a Skype premium account. The researchers were trained by participating in each other's calls. Furthermore, with two different software packages work could be executed alongside each other.

After the training period, the time came to start reaching out to the parties in the data. Each call followed the structure of first briefly introducing the research and addressing that a of port-configurations drew the attention to their server. In practice, it appeared that many interviewees had only limited knowledge about

IP protocols, a problem similar to that which Stock et al. faced in their study. Most people admitted that they had previously heard of an IP address, but did not know about the existence of IPv4 and IPv6. This setback made the research more difficult because some understanding of the IP protocol was required to participate in the interview. In that case, we asked to be put through to somebody from the technical department. They responded to this request in three ways. First, they said that their tech support was not interested in calls from individuals who are not clients. The conversation would then be ended, and the person was thanked for their time. That company was labeled as 'not interested.' In the second scenario, the call was redirected to a tech support worker who was asked the questions stated in the script. Most of the time, people were interested in and requested our scan results alongside the goal of this study over email. An example email was added in the appendix, see Appendix A. For privacy reasons, these people were asked if they were the person responsible for the server before sending any information about the potentially vulnerable server. The final scenario was that the tech support was only available via email. This email would not contain any information concerning the ports. This information was only shared when an email confirmed that the tech support team was responsible for the specific server. When these individuals confirmed their willingness to participate in this research, an appointment was to conduct the interview.

Besides recruiting interviews from the data, experts from the field were contacted if they want to participate in this research. These experts were found via personal/professional relationships and the relationships from PwC. In this case, the focus was more directed on their experiences in port policies and what they had seen at companies they worked at. People contacted via this route were on average more willing to participate because of reference.

4.4.1. Limitations and Roadblocks

During the recruitment process, some problems occurred why a party was not able to or interested in helping in this study. There were multiple reasons which can be summarized in knowledge, time, and privacy.

The first problem that occurred was the difference of responsibility between hosting provider and user. As mentioned earlier, the researchers had to guess if the user or provider was maintaining the server. It seemed that most users were responsible for layer seven configuration but the rest laid at the hosting provider. Most end-user did not know the firewall configuration of their machines. Although the subject was not required to have a broad technical background, some knowledge was required to give clear answers to our questions. After some time, it was decided to focus on the hosting parties, because most end users had no experience with firewall configuration whatsoever. It was assumed that employees of hosting parties would be more knowledge. This assumption seemed to be accurate. However, this led to additional problems. Many providers were interested and thankful for the information, however not willing to participate in an interview because of privacy reasons. Despite the fact we were not interested in client data, they understandably said that they were not able to participate in our research because of client secrecy.

Another reason was time and interest. For hosting providers, after they had received the information about the vulnerable system, there was no reason to take part in the study. Luckily, not all were of that opinion. Furthermore, on Fridays, it was more difficult to come in contact companies. Many employees may have their day off on Friday.

Next, another limitation we faced was reaching the correct person. This limitation is something that had also been mentioned by Stock's article. Normally, we first came in contact with a non-technical person, and it took multiple redirects to find the right person. Another closely related roadblock was that this person was sometimes not willing to cooperate. Altogether this long procedure was not always fruitful.

4.4.2. Facts and Figures

To show the number of parties we contacted some small figures are elaborated. From the data, we identified approximately 200 different parties. In this list are ISPs, companies, hosting providers, individuals. Next, we were able to find approximately 90 telephone numbers, 120 email addresses. In the end, we had email contact with more than 60 parties, and even more, companies have been contacted. We only emailed a party if they asked for more information. Finally two tables (Table 4.3, 4.4) are added to show the amount of misconfigured devices and the corresponding ports. They are split up by the port configuration, IPv4 open only, IPv6 open only, both open, and both closed.

IPv4	IPv6	DNS		FTP		HTTP		HTTPS	
1	0	596.766	4%	444.542	3%	385.014	3%	360.932	2%
0	1	122.804	1%	109.818	1%	156.903	1%	146.619	1%
1	1	1.062.968	7%	7.518.553	52%	13.103.650	90%	11.932.924	82%
0	0	12.764.791	88%	6.474.416	45%	901.762	6%	2.106.854	14%
Total		14.547.329	100%	14.547.329	100%	14.547.329	100%	14.547.329	100%

Table 4.3: Together with the next table Table 4.4 an overview is presented of often used ports on both protocols. It presents ratios between the current state of port configuration on IPv4 and IPv6

IPv4	IPv6	MYSQL		RDP		SMB		SSH	
1	0	594.777	4%	78.847	1%	5.884	0%	612.531	4%
0	1	37.911	0%	3.032	0%	10.661	0%	458.396	3%
1	1	2.221.542	15%	5.1307	0%	50.352	0%	4.057.780	28%
0	0	11.693.099	80%	14.414.143	99%	11.693.099	99%	9.418.622	65%
Total		14.547.329	100%	14.547.329	100%	11.759.996	100%	14.547.329	100%

Table 4.4: Follow up of Table 4.3.

4.5. Interview Setup

Interviews were either done over the telephone or in person. Most interviewees that were found via the data participated in a telephone interview, and persons that had been contacted via the researcher's network were visited in person. For consistency, it was tried to create similar interview whether over the phone or in person. The interview started with a brief introduction about the researchers and the study. Most interviewees are IT specialist of which the technical background was confirmed, allowing core problems to be tackled in most of these interviews. However, in some interviews, the respondent's networking knowledge exceeded that of the interviewer.

The interview was guided based on prepared questions but could be deviated from if interesting topics were brought up. There was a small difference between the prepared questions from respondents recruited from the internal/external data and the *normal* data. The difference was a few additional questions about the configuration that was found. However, all the other questions stayed the same. The questions could be segregated into three (four in case of internal/external, third section) categories. The full question list was added to the appendix, see appendix B.

- **General information about the participant and company.** It was helpful for the researcher to know the background of the person or company. It was interesting to know the sector (private or public), the industry operated.
- **In-depth questions about the configuration procedure concerning open public ports.** Trying to reveal how companies apply their port configuration policies in practice.
- **In-depth questions about the internal/external configuration (only if applicable).** Question why there was a difference between the IPv4 and IPv6 address configuration. Mostly if it was intentional or unintentional.
- **Questions concerning security incidents or potential incidents based on IP configuration.** Broad questions concerning a, IP configuration or when they did not have a security incident; they were asked to for their speculations surrounding what would happen if it had. Finally, their opinion was asked on how to mitigate these problems best.

We based these questions on the work of Dietrich et al. [22]. We followed a similar line of questioning as the questionnaire developed for this study. There are similarities between the goals of the research, finding the reason behind a configuration. However, instead of focusing on misconfigurations in general, we targeted our questions on the IPv6 configuration process. It was decided to focus on mapping the port configuration process (both on IPv4 and IPv6) and possible security implications of misconfigured IPv6 systems. Finally, a draft was created reviewed by other researchers before the first interview was conducted. This way is how the initial interview questions were formed for this study. During the interview, notes were taken, and an audio recording was made for later analysis. The period of the interviews was different. On average, were interviews conducted in person longer, because of the more in-depth follow-up questions.

4.5.1. Interview Phases

Interviews were performed in multiple rounds. Although there are no clear borders, three different stages can be identified: Testing the interview questions, gathering of the data, verification of the data. Because this research was based on theory building, it was not clear what we were looking for in the first place. It was also not certain in which direction the answers of the respondents would lead the interview. First, to ensure that the initial interview questions would provide sufficient answers, a practice interview was up with a network administrator from the TU Delft. The participant knew that this was the first interview and was asked to give feedback on the interview questions afterward. This interview was not used in the final analysis but provided some useful starting points. Next, interviews were performed with experts from the field and subjects from the data. After every interview, the data was analyzed and coded. These were not worked out in detail yet, but if certain interesting topics started to emerge, they were added to a secondary question list. This second question list is also added in the appendix, see B. These questions were used in later stages of the research to focus the research and more direct search for findings concerning IPv6 security and configuration. Furthermore, we noticed that on multiple topics, we were starting to reach saturation, and while other concepts started to emerge from the data. For example, after multiple interviews, we had a clear picture of the configuration processes at organizations, and different interviewees gave repetitive answers. After saturation was hit on these topics, it was unnecessary to continue questioning these topics. Particularly because most interviewees were on the clock. We chose to remove questions concerning the port configuration and add questions of IPv6 adoption and awareness. In the following interviews, these topics and findings were discussed with the interviewees. To check for saturation in these topics, we evaluated all our findings in the interviews twelve and thirteen, the final two. They stated they agreed with our findings and were not giving us additional information, stating that saturation was also reached on the topics of IPv6 adoption and current state of awareness.

4.6. Data Analysis

After each interview, it was transcribed by the researchers. It was chosen to do word-for-word transcription, meaning that most filler words were removed while grammar mistakes were not. Next, the transcript was loaded into the software ATLAS.ti for further analysis. ATLAS.ti was chosen because of its ease of use. It is a powerful software to code pieces of text and for creating links between the codes. Nonetheless, the quality of the codings is still dependent on the researcher's ability to identify the key elements. Therefore, to get more accustomed to ATLAS.ti and coding principles, an appointment was made with a post-doc researcher from the TU Delft, increasing the quality of data analysis from the transcripts.

Before the coding started, each interview was re-read and listened to in full. These recording had the important added value of helping the coder to extract the most interesting elements according to the interviewee based on volume and intonation. After this process, the next step is open coding. During open coding, a researcher reads through the information and start to create temporary labels for parts of the data, allowing summarizing the general topics in the unstructured gathered data. At this time, there is no direct link to an existing theory yet. Open coding is mainly focused on what emerges from the data. The goals should be to treat all data as equals because in this phase, the researcher cannot determine the importance of codes yet, entailing that data should be structured, but no relationships should be sought yet.

After the open coding phase, groups started to emerge, and free codes were created. A free code is a code that is not directly linked to a piece of text from the transcripts. These free codes were used as main structures to analyze the data further. Furthermore, the data was cleaned during an iterative process through merging, deletion, and renaming. Next, it was attempted to establish relationships between free codes and the codes from the open coding phase. This iterative process was performed to search for words to link codes together. Finally, the entire codebook was manually inspected if the search terms missed codes. This process was done by sorting the codes on the occurrence and treating codes that occurred multiple times with more importance then codes who occurred once. If codes occurred more then five times, they were to with a relationship to a free code. Codes with a lower occurrence were overlooked and see if they were usable for this study. Finally, free codes were linked together if applicable.

4.7. Summary of Approach

This chapter discusses all the steps taken for this study. These steps can be roughly separated into three phases. The first phase responsible party was the UvA. Researchers from the TU Delft executed the second and third phase. The first part entailed the identification of dual-stack hosts and scanning on both protocols resulting in a SQL database with an organized overview of the hostnames and the result of the scans. The second phase involved the identification of owners of the misconfigured hosts. Structured steps were taken to find system administrators that were responsible for the configuration of the devices and asked if they wanted to participate in an interview. After the recruitment phase, the final phase, the research was executed. Thirteen interviews were conducted with IPv6 security experts and administrators from the database. Finally, these interviews were transcribed and analyzed with the aid of ATLAS.ti. In table 4.5, an overview is given of all the steps performed. Furthermore, it shows which party has contributed to which phase.

	Steps	UvA	Main Researcher	Research Assistants
1	Hostname Identification	✓		
2	Scanning	✓		
3	Database Creation	✓		
4	Database Anaysis		✓	
5	Data Extraction		✓	
6	Hostname Lookups		✓	✓
7	Ownership Identif		✓	✓
8	Ownership Identif (Inter Exter)		✓	
9	Interviewee Recruitment		✓	✓
10	Interview Preperation		✓	
11	Interview Execution		✓	
12	Transcribing interviews		✓	✓
13	Analyzing transcripts		✓	
14	Writing Report		✓	

Table 4.5: An overview which party was responsible and aided for every step taken in this research.

5

Results

This chapter will discuss the results and findings of the interviews. These results are based on the analyses of the interview transcripts. As a reminder, the interview question of this study can be found in Appendix B. Please refer to chapter four for an explanation of how these results were studied.

5.1. Information Sources

A total of thirteen interviews have been conducted, excluding the trial interview with the university. Table 5.1 summarizes the different respondents who participated in this research. The interviewees were separated between different groups, creating a better understanding of their role in and view on the problems surrounding port configuration and security in general. We discovered that there are considerable differences in answers based on these different perspectives. For example, people who focus on the security aspect of the problem tend to look at the problem differently than people who focus on availability. This table reflects a broad perspective on the global landscape. As such, it presents the different actors' opinions. In this chapter's text, interviewees will be referred to as I#. Where the # is the number for an interview, for example, I2 is interview two.

All thirteen interviews were coded to a total of 855 unique codes. This number takes into consideration the free codes and the merging of overlapping codes. An overview of theses free codes is visible in Table 5.2. These free codes were referred to as first-level codes or top-level codes (in red). Furthermore, the second layer of codes (yellow) was created to specify these top-level codes further. A final free code (green) was used to introduce an essential general topic that could be linked to a second layer code. To summarize, in the appendix are the word webs added that were created by ATLAS.ti (See appendix D). These are all the word webs made for this study. Segments of some of these webs from these total extracts are added as figures in this chapter.

Interview	Occupation	Company size	Job title	Security	Tech. driven	Sector	From data set
I1	Technical	Medium	Network consultant	No	Yes	Network	No
I2	Technical	Medium	Cloud engineer	No	Yes	Hosting	Yes
I3	Technical	Medium	System integrator	No	No	Data transfer	Yes
I4	Technical	Small	Network engineer	No	Yes	ISP	Yes
I5	Technical	Small	Software engineer	No	Yes	Software development	Yes
I6	Managerial	Large	Consultant	No	Yes	ISP	Yes
I7	Managerial	Medium	System architect	No	Yes	Hosting	Yes
I8	Technical	Large	Security engineer	Yes	No	Public sector	Yes
I9	Technical	Large	Network engineer	No	No	Public sector	Yes
I10	Managerial	Medium	CEO	Yes	Yes	Security	No
I11	Managerial	Large	System engineer	Yes	Yes	Network	No
I12	Technical	Large	Network engineer	No	Yes	ISP	Yes
I13	Managerial	N/A	Consultant	Yes	No	Security	No

Table 5.1: Summary to give an insight on which kind of parties and people have participated in this research

Top Level Code	Second Level Code	Number of Links
Solution	IPv6 problems	21
	Security	5
	Possible implementations	9
Company	Large Company	13
	Medium Company	4
	Small company	13
	Technology driven	7
	Not technology driven	2
Technology Push		8
Demand Pull		13
Port configuration	Misconfiguration	26
	Port configuration process	14
	Security implications	11
Perspective	Managerial perspective	27
	Technical perspective	16
NAT	NAT64	2
	Carrier-grade NAT	7
	Normal NAT	5
IPv6 Adoption	Reason for IPv6 adoption	23
	Reason for no IPv6 adoption	29
Occupation	Managerial	6
	Technical	7
Internal/External config	Opinion	8
	Unintenional	6
	Intentional	13

Table 5.2: All first-level codes and second-level codes and the number of connections to codes that have been identified in the interviews

5.2. Security in General

Before addressing the port configuration processes and IPv6, some general remarks about computer security were mentioned in this section. All actors indicated that computer security does not yet receive the attention it requires. Topics that were mentioned include the high pressure on hiring adequate personnel despite the scarcity of security professionals who can face the upcoming challenges. Moreover, I13 also mentioned that networking employees are also hard to come by. Finding an employee that performs well in both fronts is even harder. Port configuration is part of the difficult field of networking. The field lays heavy requirements on knowledge and hard skills. The reasons mentioned above make it hard to find skillful employees, especially for smaller companies. Based on the interviews, most top experts are employed at larger companies or work as an external consultant for the larger companies. Of course, finances are a factor in finding the correct personnel. Another statement from one of the interviewees (I11) is that Security Officers or C level executives focus their attention on application security. Application security entails the security and improvement of applications. For example, the finding and fixing of bugs in software. In case of improper application security, it is, for example, possible to circumvent the login screen and login without the required credentials. For them, application security is a more gasping topic, because they have an understanding of what applications are and where they are used for in the organization. Most C-level executives have little to no experience with networking. Therefore, they do not give it the necessary attention, and thus port configuration is often lacking behind.

Moreover, security personnel faces the problem that the Internet as it is today was never meant to grow to its current scale. We have changed the application of the internet to fit needs that were not envisioned initially. When it was developed, it was based on complete trust between all users, without any malicious intent. We can state that this is not the case anymore, making the job of security personnel even more difficult. Unfortunately, it is hard to move away from the current insecure protocols because they are embedded in the core infrastructure of the Internet. Therefore, modifications and workarounds have been created to ensure safe communication on the Internet. However, this also increases complexity and inefficiency.

Almost all the actors mentioned a monetary perspective. Security is always a cost, and there is seldom an immediate gain. For a company, it is hard to determine its cyber risk and which measures it should take to keep its assets safe. Furthermore, it is hard to assess the impact of a cyber-attack until one occurs. Addition-

ally, most companies choose Availability over Confidentiality and Integrity (CIA triad). I2 pointed out: "The knowledge in companies is low, and an administrator is happy when it is working and leaves running as soon as it works." Less availability means downtime, which could lead to employees being unable to perform their job. Such a thing is expensive and therefore of importance in management. Managers understand the consequences of no up-time and associated costs. However, managers struggle to envisage the consequences of a cyber-attack. They are reserved about the trade-off between security measures that protect their company and the limiting implications of these measures. For instance, a drastic measure would be to have no internet connectivity because most attacks occur from the internet. However, no Internet connectivity greatly decreases productivity.

Parties of different sizes and technology maturity have been interviewed during this study. They indicated that larger companies had a more formal approach to security in comparison to smaller ones. Here, formal refers to a situation where multiple actors are involved, and checks are in place. In those large companies, each process is evaluated by multiple actors (e.g., Security Officer). Only then is the process accepted and cleared for execution. Additionally, these large companies divide tasks among different actors when configurations are scheduled to be changed. This formality, including the logging of all changes, is advantageous because of preventing errors; however, the long and tedious process of approval, execution, and verification has been known to make some employees take unauthorized shortcuts. On the other hand, this is not the case in smaller companies due to the more modest size of their system administration team. This way causes difficulties because they are not large enough to give employees individual responsibilities, which is fundamental to properly implement security processes according to the 'separation of work' principle. Indeed, in contrast to larger companies, employees do not check each other's work due to the limited workforce. I4 eluded on this subject: "You need to have people who think with the availability cap on, and you need to have other actors who think with the security cap on."

5.3. Port Configuration Process

One aspect that was discussed with the respondents was the port configuration process, as shown in Figures 5.1, 5.2, and 5.3. This process was discussed with I1 to I5, I7, I8. As a reminder, ports grant access to running processes on a server to which a client can connect. A firewall is responsible for allowing or blocking access. In essence, there are two configuration methods for a firewall: locally or centrally. A local firewall is running on the device itself and blocks or allows connections to local ports. A centralized firewall determines whether each of the connections routed through it is authorized. As such, this type of firewall safeguards the entire network as it filters any entering connection and, like a local firewall, regulates the access to a device. If a person should not have access to a system, he or she should not be able to connect to that system, often referred to as least privilege access. The attack surface, possibilities to attack a system, increases if security personnel does not follow this guideline. For example, on almost all Linux systems, port 22 (SSH), can be accessed from anywhere, implying a risk associated with keeping that port open. However, SSH requires password authentication for a successful connection. In the case of a strong password, it could still take years before a system is breached and thus still can be considered safe.

We only approached people concerning their public IP addresses. In general, people are more strict with open port policies on public IP addresses than on private IP addresses. These are private because they are situated behind a NAT and are not publicly accessible if no port forwarding has been configured. The respondents, mostly system administrators, revealed that the basic security processes of configuring public ports have some overlap with the least privilege method. On a side note, the security of internal networks is a whole other topic. Where public ports are concerned, an example will be given on how most configuration processes occur. In large companies, a process always starts with a request for a need from the organization (I1, I3, I8, I9, and I10). For example, A new service is offered to clients, new hardware is implemented, and vendor access is required, or an employee needs (temporary) access to certain systems. The request is then translated to a task containing what is needed, by whom, why, and sometimes for how long the configuration needs to change. The task is divided into sub-tasks that are assigned to different employees. For instance, a firewall engineer could be responsible for creating a NAT rule on the centralized firewall, while a system engineer is responsible for changing the firewall rules on the local system. When the task is fully described, it is reviewed by a Security Officer, who asses the risks of configuration and approves the task and sub-tasks. They are then executed. Furthermore, these companies have predefined protocols as to what is allowed and

what not. The more mature a company, the more predefined protocols they have on what is allowed and how certain request should be handled. There are occasions that a request does not fit into a predefined protocol. If it is a unique request, it is discussed between the administrators and security personnel. Multiple respondents stated that these unique requests could lead to long discussions. Altogether, an overall example of a port configuration process cannot be given for smaller companies. There, the system operators have much more freedom in the way they want to execute their work. In those companies, there is a much larger spread in knowledge and thus operational procedures.

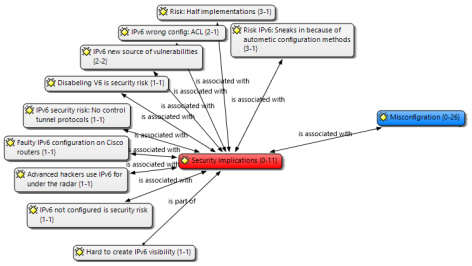


Figure 5.1: All codes concerning security implications around port configuration.

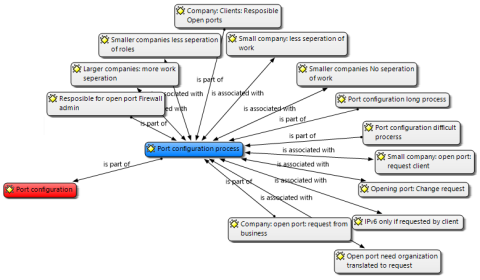


Figure 5.2: The findings on how companies set up a their port configuration process.

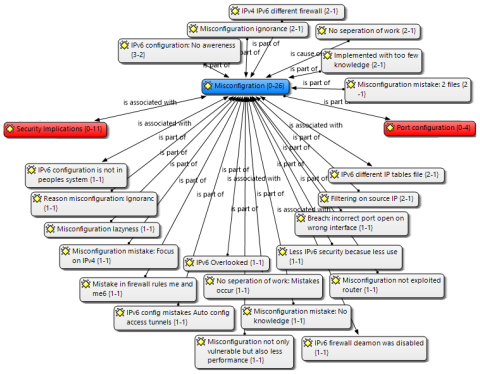


Figure 5.3: A summary of responses regarding reasons behind a misconfiguration on dual-stack hosts.

5.3.1. Opinions

Most opinions that came forward in the interviews were largely in line with the general security findings of this study. Most respondents said that large companies use a process with different checks along the way that impacts productivity. Sometimes, employees of these companies spend more time on reports than on the operation itself. I10 said that from his management perspective, it is good that there are processes, but from his engineering perspective, it greatly influences the freedom of working. I12 said that he preferred to make use of local firewalls instead of global firewalls because it can take a long time until a task is approved. As visible in Table 5.1 has I10 a managerial background while I12 has a technical background, showing that engineers will take shortcuts because of time limitations. Furthermore, they also mentioned that less sepa-

ration of work could lead to mistakes and that smaller companies face these problems. On the other hand, larger companies face problems with huge configuration files. If these are not properly maintained, it could be possible that there are firewall rules on one side of the network to allow access to a system somewhere else in the network. If that system is phased out after a certain period, all firewall rules should also be removed. However, in practice, this is not always the case. There is no security risk, as long as that specific IP address (from phased out system) is not reused, because the route leads to nothing. Nevertheless, if after a certain period that IP address is reused, it could lead to unwanted routes and possible security implications.

5.3.2. Verification

Finally, we asked our interviewees about their port verification process, see the output of ATLAS.ti in Figure 5.4. It is possible with the aid of tools (Nmap, Zmap) to see which ports are open or closed. Besides, we noticed a relationship between the verification process and company size. Smaller companies have little to no verification processes while in larger companies, Security Officers have periodic assessments of their vulnerability scans (I1). Here must be noted that there is a difference between internal and external scanning. Internal scanning is performed in a private network in search of security vulnerabilities. However, vulnerability scanning is a tool to search for vulnerabilities in a network and do not necessarily inform about open ports, especially not on external ports. We found that even most large companies did not do have any verification of their open ports. I8 and I9 are from a large technology company and stated that their company just got permission to start testing with a vulnerability assessment tool. I2 said that once in a while, a network engineer looked through the firewall rules and reviewed if they were all still applicable, but there was no clear protocol. In the case of smaller companies, verification was done after configuration changes. Because it is a smaller company, fewer changes to the configuration occurred and there they stated that only after configuration changes, changes could occur in firewall rules. Therefore, every time, a new system was adopted; some verification was performed but noting periodically. The interesting finding is that almost all respondents said that verification is an important part of the process but could not lay their finger on why there were no clear verification processes.

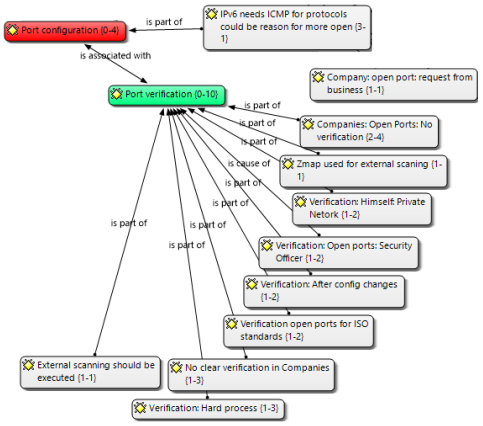


Figure 5.4: This figure displays the codes that relate to the verification process of IP configuration.

5.3.3. IPv6 Port Configuration

The previous paragraphs discussed the port configuration process in general and why misconfigurations in general occurred. However, for this study, we were interested in the difference between IPv4 and IPv6 configuration. As an external party (the researchers), is it hard to estimate if a port should be publicly accessible or not. However, we were able to identify different configurations between the two protocols (IPv4, IPv6). First, we asked our respondents if their organizations supported IPv6. All respondents confirmed that their organizations, in some way, support IPv6. We hoped that we were able to find a person that was unknowingly supporting IPv6, but we were did not encounter such a person/company. Next, we asked if their configuration process was similar on IPv4 and IPv6. They stated that it was hard to determine because the request for IPv6 is low. Some parties on default support both IPv4 and IPv6 (I1, I4, I7, I8, and I9). Other parties only support IPv6 if there is a request from a client (I2, I6, I10, I12). However, parties that only support IPv6 on request said that it did not often occur that a client requests an IPv6 address. People are not aware of the existence

of IPv6, and there is much less awareness about IPv6. They know that it is there but forget to configure it correctly. One example that most respondents gave is iptables. Iptables is a local firewall used in Linux distributions. Most network engineers have experience with the configuration of iptables and properly add the correct firewall rules on IPv4. However, what is often forgotten is that IPv6 uses ip6tables, a different daemon using a different configuration file. If that file is not configured properly, the systems still can have open ports on IPv6 and thus could still be vulnerable. Often mentioned reasons where that the lack of knowledge and awareness on IPv6 is why ip6tables is forgotten. We found a similar instance with me and me6, as explained by I12. Me/Me6 is part of IPFW, which is a firewall for FreeBSD. FreeBSD is an open-source operating system based on UNIX. A similar problem occurred with this distribution. Next, we also had an instance were a DNS record was erroneously configured. The DNS record for IPv4 was changed while the IPv6 record was left on the default. Adrift in the migration occurred, and our scans identified this drift. As a summary, almost all respondents stated that these mistakes were made because of a lack of knowledge, used to IPv4, ignorance, laziness, and awareness.

Finally, we analyzed the difference between IPv4, IPv6 configuration, and company size. We saw that here again, smaller companies have more struggles to keep the configuration between both protocols similar. We noticed that the argument that large companies could have a drift in their configuration, because of the huge files, was not valid. We were not able to identify these mistakes. We noticed that the larger the company or the more technology-oriented, the more automation was used in the process. Automation tools, like Ansible and Puppet, are adding programs that help with software provisioning and configuration maintenance, both on the operation as application level. Recommendations were given about automation by larger institutions as I6, I7, I8, I9, and I11. These companies use predefined scripts and deploy them over their entire infrastructure with one command. In this case, if a mistake is made, it will be identified earlier because multiple systems will be incorrectly configured instead of that one device. One company (I7), medium-sized but highly technology-driven, even had feedback loops which checked if their script were executed properly. If these checks detected a drift in configuration, an alarm was triggered, and the script was rerun to redeploy the correct configuration. To summarize this approach, a respondent stated that people always make mistakes, and a party should automate as much as possible to decrease the chances for people to make mistakes.

5.3.4. Logical Differences

Despite saying that all found configuration with a difference between ports is erroneously configured, we asked if in their opinion to think about a logical reason why there was a difference in port configuration. We asked this question mainly to respondents with a technical background. There were two reasons why they thought differences could exist. First ICMP (ping), a protocol that operating systems use for error messages. It is normal practice for system administrators to block ICMP on IPv4 because of an old DoS attack named ping of death. Ping of death was an attack where a too large package was sent on purpose, which would crash a system. Most systems nowadays are not vulnerable to the old ping of death, but there are similar attacks on ICMP still in use. One solution for system administrators is to block ICMP on IPv4 to prevent these attacks. However, this does not work for IPv6; IPv6 requires ICMP. ICMP gained with IPv6 a much more significant role in its operation. The protocol will not work properly if all ICMP packages are blocked. So, from a technical perspective, it could be plausible that ICMP is closed on IPv4 and IPv6 open.

Secondly, we used DNS to identify IPv4 and IPv6 addresses. Here we assumed that both records resolved to the same physical or virtual host. However, this is not necessarily the case. For example, a web server could be running on two different hosts. This can be the same website with the same appearance but running on two systems with different operating systems. The IPv4 record could direct to a Windows server while the IPv6 record directs to a Debian server, but for the user the result is similar. For us, with the current scan results, are we unable to determine what OS the server is running on. We tried to look for hosts that had one protocol SSH open and on the other RDP, but this also gave to many false positives. There were multiple hosts with both protocols open and thus still unclear what of OS it was. It could be possible to determine the OS with more in-depth scanning; however, this is a more invasive and longer process. Furthermore, some ethical consideration could come into play because it could cross the delicate line between research and “hacking.” Therefore, this question remains open, and we cannot determine how much systems are configured to run on two different hosts. During our research, we did not identify an intentional case. However, as mentioned before, we found one unintentional case where a person forgot to change its AAAA record. So, this could be

the second reason why there are differences between the two protocols. Nevertheless, it is highly unlikely that this is a significant amount.

5.3.5. Opinions About Process

We also asked our respondents what their opinion was about the current security process in their companies. Furthermore, we also went into more depth concerning their opinions of their current configuration protocols. Lastly, some recommendations by the interviewers are discussed in this section.

These next three topics originated from hosting or VPS providers. Of course, the recommendations are per company branch different, and these recommendations apply to them. The first one was from a small hosting company who said that they had protocols for port configuration, but in his opinion, some time should be spent to increase the number of protocols. Now they only had some for the basic use cases, but the number should be increased. Currently, too many topics were discussed with the entire team, a time-consuming process according to him. Secondly, we noticed that many incorrect configured devices are virtual private servers. We discussed this with two hosting providers, and they said that it occurred regularly that multiple clients rent just one VPS. Furthermore, most VPS providers supply a VPS with on default IPv6 enabled, which is regularly forgotten by those users. He said that a large part of his client portfolio are do-it-yourselves in the IT business. With the rise of the cloud and virtual hosting, it has become relatively easy to receive a system that is directly connected to the internet. Before virtualization and cloud migration, people had to invest in hardware before it was possible to have a system directly. With cloud virtualization, this threshold has become drastically lower. People who are relatively inexperienced with port configuration now can acquire a system that is directly routable from the Internet. These people are prone to make mistakes and forget to apply the correct firewall rules. Probably because of lack of knowledge and awareness. Two hosting providers addressed this problem and came with two solutions. First, to take away these responsibilities by the clients. Give clients access but fewer rights, and if users wanted to open or close ports, they have to request it via the hosting party itself. The other solution was to create a GUI (graphical user interface) around the process. On most UNIX systems, firewall configuration is done via CLI. A CLI interface does not necessarily give an ideal overview. If a more user-friendly interface is created, with an agent in between that can translate the user input to commands, it should be less prone to human error.

Another discussing worth mentioning was with an engineer working at a medium, non-technical driven company (I3). He agreed that there should be no difference between both protocols; however, from his perspective, he did not mind that it was not configured properly. The company he worked for did not have private data, and if an incident occurred, it would be an inconvenience but not the end of the world. A party should think about the balance between security and costs. In their case, security was not one of the main targets. They chose to allocate their finances on other projects and accepted the security risks. Furthermore, he stated that it was an accepted risk by management and engineers.

5.4. Security Implications

With the introduction of IPv6, a new source of security implications occurred. This paragraph will briefly talk about security implications that are introduced because of IPv6. From the interviews, there are three scenario's where security risk concerning IPv6 can occur.

1. IPv6 is fully enabled
2. IPv6 is half configured
3. IPv6 is not configured

First, IPv6 has been developed relatively long ago, but the adoption rate is still lacking behind. The pool of active users is relatively small. One security consultant (I6), specialized in IPv6, stated that there are still child diseases in IPv6. We need to start using IPv6 on a larger scale to identify these mistakes. So far, there have been more CVE concerning IPv4 than IPv6. According to him fully enabling IPv6 is the safest option for a system administrator. However, it should be configured adequately. Next are half-configured or inadequate configured networks. We talked with multiple parties who said that half configurations are a security risk. Specifically, when IPv6 is configured with not the required knowledge. Most devices (Mac, Windows 10) now on default choose IPv6 as their preferred network. If there is some Router Advertisement picked up by a

device, it will switch over to IPv6 and create a tunnel to the outside. One respondent said that it is comparable as a device that sets up an unwanted VPN connection. If IPv6 is not correctly blocked on the outside firewall changes are that a device can start, communicating over IPv6 without the correct firewall policies implemented. Then a device that is usually behind NAT directly communicating to the internet. The tool from Fox-IT discussed in the background research works with a similar approach. Finally, there is the scenario where there is no IPv6 implemented. This setup can give serious security concerns because all devices are IPv6 ready. If a system administrator disables IPv6, they have to be sure also to disable it on the clients, block all tunneling protocols, and block all traffic on IPv6. Else, because of the auto configurations abilities of IPv6, it can slowly sneak into a network.

5.4.1. Summary Port Configuration

When the previous paragraph is overlooked, some exciting findings come up. All actors agree that there should not be a difference between the two protocols; it increases the risk and the attack surface. However, it should also be noted that security is a cost and should be weighted in comparison to spending and convenience. Furthermore, we found a wide spread of different reasons why configuration mistakes were made. In the end, all respondents agreed that there is not enough knowledge and awareness is concerning IPv6 and IPv6 security.

5.5. IPv6 Adoption

So far looked at port configuration processes on both IPv4 and IPv6 and elaborated our findings why there are differences. Most respondents gave lack of knowledge, IPv6 adoption, and awareness as an argument why these mistakes occur. For this, we also talked with our respondents about the past of IPv6 adoption, its current state, and their future expectations. Furthermore, we looked at the differences between the managerial and technical perspective on how the IPv6 adoption process. We will talk about reasons why companies should adopt IPv6 and opinions will be presented why IPv6 adoption is such a heavy burden, see Figures 5.5 and 5.6. However, it must be stated that for this research, we only talked to IT personnel in Western Europe.

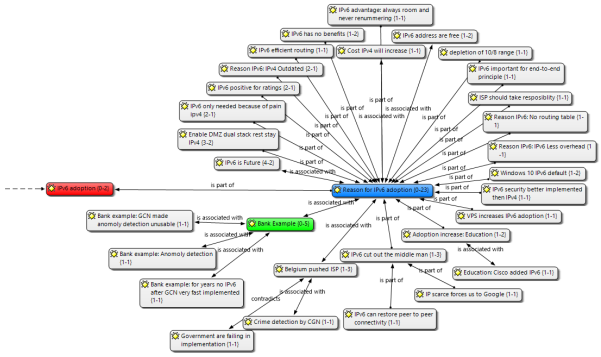


Figure 5.5: Schematic of advantages of IPv6 adoption

From a technical perspective are there several reasons why IPv6 is the preferred protocol. First, IPv6 is a more modern protocol with more options and better to build security. Furthermore, it is a more clean protocol with less overhead then IPv4. If the two headers of the protocols are compared, there is much less overhead in the IPv6 protocol then IPv4. Furthermore, IPv6 tackles the problem we have with IPv4, the IP depletion. Where we have problems with finding IPv4 addresses, we will not have this problem in the foreseeable future. The address space of IPv6 is considerably larger than IPv4. For example, if a company hires a VPS at a hosting firm, they will normally receive one IPv4 address. There are scenarios that engineers need to use multiple interfaces and thus multiple IP addresses. It is possible to increase the amount of IP addresses; however, most hosting providers have a monthly fee for this. If IPv6 is implemented on this system, every VPS on default receives a /64 to allocate. A /64 is approximately 1.8 x 10^{19} unique IP addresses. Therefore, every VPS that is hired will have more public IP addresses than the entire IPv4 address space. There is no scenario where a system administrator needs more than that amount of IP addresses. This is also the case for office networks when a company grows; it can become harder to find usable subnets to allocate new devices. System administrators have to create a new IP plan and renumber the devices. A company mentioned that IPv6 could be useful in case of a merger because in theory an every IP address needs to be unique in a network. If two

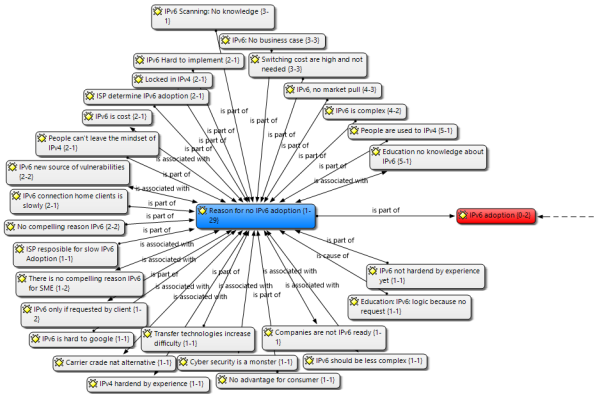


Figure 5.6: Schematic of disadvantages of IPv6 adoption.

existing infrastructures need to merged chances are that there are duplicate private IPv4 addresses. If both companies have a fully implemented IPv6 stack changes of this are zero to none.

Furthermore, there is another argument why IPv6 is the favorite protocol over IPv6. When IPv4 was developed, it was meant to have a peer-to-peer connection between two users. Both users had an IP address, and they would connect directly to each other. With the IP scarcity, various tricks have been developed to increase the lifespan of IPv4 before IPv6 is fully developed. We found that almost all respondents were not in favor of NAT. Most of them confirmed that NAT was a workaround and not something that should be used on default. It breaks the end-to-end principle of the IP protocol. One respondent knew the writers of the RFC for NAT and said that the group who developed NAT regrets developing the RFC. Furthermore, because of the NAT system, administrators are used to NAT and search for similar properties in IPv6. However, on default, IPv6 does not use NAT because there is no IPv6 scarcity. One respondent stated that we are so used to NAT that we now cramp as much as information we can in one IP address translate this with NAT rules. The stretch is out of the IPv4 protocol, and we limit ourselves in our abilities to stick to IPv4 and NAT. Thus clearly looked from a technical perspective is IPv6 the superior protocol over IPv4. However, these topics mentioned above are mostly nice to have and are not appealing for the normal user. From a technical perspective are these features interesting, but they do not appeal to the average user. The normal user or manager is interested in a safe and workable infrastructure do not mind if traffic is routed over IPv4 or IPv6. We cannot put numbers to it, but most people do not even know the difference and do not care. The performance could improve on a larger scale, but it is not noticeable for the individual. We asked multiple respondents: “Image if you want to implement IPv6 in your company, what arguments can you give your superiors why they should allocate finances for this change.” These respondents answered the same. If there is an operational infrastructure, there is no compelling reason to switch to IPv6. The switching cost are enormous; new investment in hardware should be made, it will take network administrators a long time to implement, and many funds have to be spent on training and education. There was no argument why a company should enable IPv6 in their corporate network. Furthermore, there is IP scarcity for starting companies, but already established companies who own public addresses there is no reason to change. There are exchanges where companies can acquire IPv4 addresses. At the time of writing public IP addresses ranges from 22-25 dollar [47]. One respondent (I7) told us that when he started working at the company, they acquired public IPs from RIPE every time they were running out. However, this is not possible anymore, and they continue buying. For them, the cost of buying new IP addresses are lower than only support IPv6. The cost of implementing IPv6 is just too high, and it is easier to buy IP addresses or use the transition methods (NAT). As long as the cost does not weigh the benefits of IPv6, management will not allocate funds to help IPv6 adoption. Finally, we compared the difference in IPv6 adoption between the company sizes and technology orientation. We saw that technology orientation is a large factor in IPv6 adoption. The more technical focused, the more IPv6 adoption is visible. Furthermore, we were not able to find a link between company size and adoption. The adoption rate, in general, is still too low to make a clear statement on that. From I6, we learned that Google’s and Facebook’s of this world are fully IPv6 ready; however, these are also highly technical companies. Only those companies have the funds to attract the required knowledge too fully and adequate to implement IPv6 in their infrastructure.

The previous section discusses the IPv6 adoption within companies. Thus, which protocol companies use for their internal routing. Which is not necessary traffic over the internet. Beside company traffic, there is also personal traffic from home users to Google, Facebook, Netflix. These large companies are all dual-stack and support both protocols; However, home routers are mostly still on IPv4. We were not able to contact the large ISPs, but we talked with other parties who work with the large ISPs. They told us that most of these parties have a large enough pool of public IP addresses that, for them, there is no need to change to IPv6. Furthermore, it is less costly to implement a Carrier-Grade NAT (CGN) than to switch to IPv6. Most home users do not even notice that they are behind a CGN. It is expected that the cost of CGN will rise, but thus far, the costs do not exceed IPv6 implementation. We asked multiple respondents about their expectation, and the reaction from I13 is obvious: "I have been wrong about this question for so many times, I do not dare to answer this anymore." Moreover, for most home users, they will never notice that they are behind a CGN. The downside of being behind a CGN is that a connection can not be set up directly to a server. A use case where a direct connection is required is employees working from home who have to connect with a VPN to the office. Nevertheless, most traffic from home users is outbound traffic, and there are only a few reasons why normal users would like to have inbound traffic to their home router. There are scenarios where a NAS (network storage) should be accessible, but often vendor provides cloud access via their website. Most popular smart home products, like Google's NEST or Phillips Hue, connect to the server of the vendor and via that route, a home user can access their products. All traffic is then routed via their servers. For now, that few home users that require a directly routable public IP is the pool large enough.

Furthermore, we talked with our respondents about reasons why a web server should be accessible over IPv6. There should be noted that most websites are hosted at hosting providers. Web hosting is on average, not expensive and outsourcing is easier than maintaining the hardware locally. Larger hosting providers now mostly support dual-stack. However, we asked why they supported both protocols, and for this, there is not a compelling reason why a website should be accessible over IPv4. There are very few IPv6 only users because that would greatly influence their accessibility. Besides, most users still have IPv4 access or use IPv6 to IPv4 tunneling techniques to access IPv4 only websites. It is possible to host a website on IPv4, but only IPv6 would be a problem. An example, *nu.nl*, is a popular Dutch news site that does not have an AAAA record. While that website is at the 1222 position in the Alexa 1 million. Respondent I2 mentioned that websites which are accessible over IPv6 appear higher in Google searches. The google example was the only reason most parties could think of why a website should be accessible other then it is supposed to be.

Finally, we asked our respondents if a party should be responsible for the IPv6 adoption and if so, who should take that responsibility. Most respondents thought that IPv6 should not be pushed and the market should solve the adoption by itself. The IPv6 adoption will come if the market requires it. If a responsible party should be pointed, most interviews agreed that ISPs should take that role. An ISP has the most experience with IPv6 and could aid in a faster adoption. If they increase the support of IPv6, more web server will be contacted over the new protocol, leading to increased awareness and more priorities for IPv6 security. Lastly, we asked what the role of the government should be. Almost all respondents stated that the government should not implement policies and regulation. It is the responsibility of the market to solve this. However, the government could give a boost to the IPv6 adoption to make sure that their services are IPv6 compatible.

5.5.1. Summary Adoption

To summarize the adoption discussion we had with the respondents, we noticed that IPv6 adoption is a tedious process. For us, it was interesting to find out that there is no compelling reason to switch to IPv6. From a technical perspective, there is, but no respondent was able to give us a clear argument to pitch to management why a company should be fully IPv6 enabled. Other then there could be scarcity soon. However, how soon is a question none of our respondents dared to touch.

5.6. IPv6 Implementation Examples

From the interviews, we talked with different parties who were fully IPv6 ready. In this paragraph, two companies are discussed on how they implemented IPv6 and why for them, it was possible to implement IPv6. The first one is a small-to-medium security company fully running IPv6. Their client portfolio exists mostly from other SME, where they assist with security questions and help with security solutions. There was for this party no reason to change to IPv6; however, they did it to be more acquainted with IPv6. They expected that

in the future, they would receive a question about IPv6 implementation; but thus far, none of their clients has ever asked any questions concerning IPv6. To be ahead of these questions and to train themselves with the protocol, they decided to implement IPv6. The respondents also stated at the end of his explanation, that if he looked through his management glasses, there is no reason why his company's network is now better or more secure. The other company is a larger, technology-oriented company. Most of their servers and workstations are fully running dual-stack. Clients who access their network on default receive IPv6, and they do not make use of any tunneling or proxies to access the internet. The reason why this company was able to adopt IPv6 relatively easy is that they have a large IPv4 space. Before the IPv6 adoption, all IPv4 traffic was already publicly routed, and they do not use NAT. Because of the already open network setup and being used to route traffic publicly, it was relatively easy to implement IPv6 in their networks. Furthermore, they use automation tools to keep track of configuration and try to keep the manual configuration to a minimum. Their main reason was to implement IPv6 is that they can and in their opinion, help the adoption because it is the future.

5.7. Internal/External Configuration

This paragraph will discuss the findings from the interviews performed with respondents for the Internal IPv4 External IPv6 data set. From the data set, we extracted configured servers with an internal IPv4 address (192.168.0.0 – 192.168.255.255, 10.0.0.0 – 10.255.255.255) and publicly routable and pingable over IPv6. There could be a security risk if system administrators are not aware when their system is directly routable. This section will discuss why system administrators used this setup if it was an intentional or unintentional setup, and opinions from other interviews about the neatness of this setup.

From the scan, we found 6791 devices that were configured according to the requirements mentioned above. In total, 39 of those devices had the top-level domain NL. Some hostnames occurred multiple times in the list, which reduced the total unique hostnames to approximately twenty. We interviewed parties from this list and found out that the configuration could both be intentional and unintentional. First, intentional configurations are discussed. We found out that all parties who had an intentional configuration were home labs from IT enthusiast. A home lab is a lab where IT personnel can do safe experiments. Normally, testing and playing in a production environment at work is a no-go. Therefore, they set up their own "production" environment to experiment. All respondents mentioned that these were no production environments, and in case of an incident, there would not have been any harmful consequences. In such an environment should a different number of security policies be applied. Furthermore, they all stated that they were aware that their systems were directly accessible over IPv6 and that the IPv6 address was directly routable. The lack of an internal DNS server was the reason for this setup. These parties do not have an internal DNS server but wanted the ability to call on hostnames in their internal network. Moreover, they did not want to use self-signed certificates and thus used Let's Encrypt as Certificate Authority. These were the reasons for an internal/external configuration. Next, all respondent stated that they were aware of the risks and that all systems were properly firewalled on both IPv4 and IPv6. The final reason why they enabled IPv6 in their home labs was that they were not given the opportunity at their work to be acquainted with the protocol but wanted to gain more experience with IPv6.

We also found a party who was not aware of the configuration. For them, it was an unintentional configuration. However, after evaluation, they confirmed that there were no security risks. Nevertheless, they agreed with us that it was not a formalized configuration. The found configuration came to existence as follows. This company is a large company with domain-joined laptops. Employees can work from home via a VPN. If a system is domain joined, it has a setting to push its IPv4 and IPv6 address to the active directory DNS. This active directory DNS is then pushed to a public DNS. This way, the private IPv4 address of the client home location is pushed to the public DNS of this company. For security reasons, we will not go into more depth concerning this configuration. The respondent confirmed from his perspective that there was no imminent security risk; however, he addressed it to the administrators responsible for the Active Directory.

Finally, we asked multiple parties about their opinion of adding private IP addresses in the public DNS. All respondents confirmed it was not a clean configuration and that the cause was probably ignorance. Furthermore, they expected that the system administrators were unaware of the configuration. Based on the findings talked in the previous paragraph, we can a partly state that is assumption was incorrect.

5.8. Limitations

This section will discuss the limitations concerning the results of this study. Most are concerning the respondents regarding this research. First, are the people who participated in this study. Because it was an interview, it often required at least an hour of a person's time. We noticed that people on average were willing to answer some questions via e-mail, but we noticed that they started to withdraw when we started about an interview. Next, in some way, we asked parties about client data, and multiple hosting providers and ISPs were keeping back on sharing client data, meaning that we were often not able to get to the source of the problem. Furthermore, most technical people were willing to participate in the research since managerial people were unable to understand the problem. This resulted that we mostly came in contact with engineers or people who came from a technical background and picked up management task through their career. Next, we asked about during recruitment we told people about their potential misconfiguration on IPv6. Nevertheless, we noticed that some people were unaware of the existence of IPv6. These people were not willing to participate and if they kept a door open during a call, ignored our follow-up email. In the end, most of our participants had some to extensive experience with IP protocols or IPv6.

The next limitation is the information in the data set. We found systems based on their DNS record. The IPv6 address space is so enormous that it is impossible to scan the entire space. Assumed is that there are more servers configured dual-stack that were not picked up in our scans. All these systems are connected to the internet but are not scanned. However, a malicious actor could also have problems locating these systems.

Finally, the interview size. A total of 13 interviews have been conducted with different system administrators. Only touching the tip of the iceberg of the total configured devices. Furthermore, almost every setup is unique, which makes it impossible to generalize the results. Additionally, it must be stated that all respondents were living in Western Europe. We created a bias concerning IP scarcity by doing so. Most established Western countries have a pool op IPv4 addresses they can use for the near future. The respondents stated that this is less the case for upcoming industries as China and India because their pool is sufficiently smaller. During this study, we were unable to come in contact with those actors and discover their opinion about IP depletion.

5.9. Summary Results

Our results are based on thirteen interviews. We conducted interviews with multiple actors in different fields. We had interviews that were more focused on the technical implementation of IPv6 and with a more managerial approach. We discussed multiple topics with our interviews and found that security, in general, is still lacking behind. Reasons are the priority of availability over security and the lack of knowledge of adequate security personnel. Next, we discovered how a business approaches the port configuration process and found differences between smaller and larger companies. We discovered that the larger the organization, the more formal this process becomes, resulting in more overhead and less efficient processes. The verification phase of the port configuration is often forgotten while all actors stated that it is an important step. We compared IPv4 and IPv6 port configuration and found there is less awareness of IPv6. Furthermore, we discussed technical reasons why IPv6 is forgotten or harder to implement. Finally, we discussed how automation could aid in decreasing the chances of a misconfiguration. Next, we elaborated on legitimate reasons of differences between IPv4 and IPv6 configuration from a security perspective. ICMP can be blocked on IPv4 but not IPv6, and the A record could point to a different server then the AAAA record. Next, we analyzed the current IPv6 adoption on a managerial and technical level. We went into more depth on the transition and transition technologies currently in use. Finally, we discussed the results of some internal/external configurations we discovered and drew the limitations of our research approach.

6
Discussion

This chapter will discuss our findings from the previous chapter. In this chapter, we will look at future expectations of IPv6. Furthermore, similarities are discussed between IPv6 adoption and similar protocol migrations on the Internet. Finally, we analyzed the IPv6 adoption with theories from literature.

6.1. Implementation Cases

From the interviews, we discovered that there are several compelling reasons for western countries to migrate to IPv6. There are reasons from a technical perspective, but the advantages do not way the cost yet. An impulse from another area is thus required for a company or organization to decide to implement IPv6. We will discuss two cases for IPv6 implementation, where the business requested it. Both cases had a similar reason why IPv4 is the preferred protocol over IPv6. In both cases is cyber-crime the reason for IPv6 adoption.

6.1.1. Bank Implementation

In 2014 Rabobank did an internal investigation to see if IPv6 was required for their company. Both for the internal network as the corporate websites (www.rabobank.nl). Beforehand it was expected that the answer would be no; The internal private addressing is large enough (17 million), Rabobank has 65.000 public IP addresses, and a migration cost time and money. Furthermore, migrations can cause additional business and security risks. However, they discovered that because of the IP scarcity, more of their clients made connection request behind a CGN. Moreover, they expected that this number would grow in the future. This change could cause a problem because their Security Operations Centre reacted with the following statement: "The majority of the SOC tooling for protection of traffic becomes unreliable or unusable." The reason behind this is that their tools make use of IP addresses. When multiple customers are behind one IP address, it becomes harder to flag a malicious transaction. An IP address is part of the profile banks create to identify if a hacker is trying to transfer money illegally. Additionally, they stated that more than 15% of their customers behind a NAT are not acceptable, greatly decreasing Rabobanks level of security and thus their services to their customers. For Rabobank, this is a compelling reason to switch from IPv4 to IPv6 [30, 31].

6.1.2. Europol

The following example is a similar request but from a European perspective. In 2017 Europol held a workshop with 35 EU countries concerning the rise of cyber-crime and Carrier-Grade NAT. With CGN, it becomes impossible for ISPs to comply with legal requests to identify individuals. An IP address is often the only source of information that is being used to link crimes to individuals. Furthermore, it increases the chances that individuals without malicious intent are being investigated by law enforcement. With CGN, they could share their IP with others and potential criminals. The inability to identify people with an IP address has created a difficult situation for European law and government officials. The problem is especially large in mobile access because 90% of all mobile internet providers have adopted this technology [26, 94].

6.1.3. Analysis

These two reasons are interesting because of the request for IPv6 from the business. We compared these two examples with a country with one of the highest adoption rates: Belgium. An interview by Network World

discusses the reasons why the adoption rate is high in Belgium [62]. Interestingly is that there is one that applies to both scenarios. The interviewee states that there is a *secret* understanding between the Belgium ISP association, cyber police, regulators, and the ministry of economic affairs that limits sharing IPv4 addresses. The contract encompasses that one IPv4 address can be shared with a maximum of 16 subscribers, a hefty limit on the Carrier-Grade NAT possibilities. Increases the cost of maintaining IPv4. The increased cost of IPv4 tips the scales on IPv6 implementation and this is one of the reasons why IPv6 adoption is at 45% In Belgium, in 2016. This situation is interesting because it conflicts with the finding we had from the interviews. During our interviews, almost all respondents stated that the government should not play a role in IPv6 adoption. Nevertheless, as seen in Belgium, governmental interference could aid in the adoption speed of IPv6.

6.2. Implementation Process

The previous two cases explained the reasons why a company or organization should implement IPv6. Based on the interviews, we would like to describe an approach to how a company could implement IPv6. This high-level implementation method is mainly focused on SMEs but could apply to other companies as well. This approach is based on combined advice gathered via the interviews. If a company starts implementing IPv6, there are roughly speaking two places to implement, internal routing and accessibility over the Internet. In our results, we were not able to identify any real compelling reasons to switch to internal IPv6. Thus, implementing IPv6 first on a connection to the Internet is a more logical step. For security reasons, a normal network is segmented, meaning some parts of the network cannot connect. For example, a printer in one office should not be able to connect to a vital server at another office. Traffic from the Internet is routed via a Demilitarized zone (DMZ) to an internal network, see Figure 6.1. There are security advantages to use a DMZ in a corporate network. For example, servers that should be accessible via the Internet (e.g., email server, web server, VPN) can be placed in a DMZ. By applying this method are these servers directly accessible from the Internet, but the internal workstations not directly routable. Furthermore, there are only a few use cases why a company should allow access to an internal workstation directly from the Internet. It is best practice to use a VPN for those applications. In the case of IPv6 implementation, in the nearby future, only a DMZ should be dual-stack configured. This setup would already solve the largest problem with IPv4, the scarcity. By only enabling the DMZ, the amount of devices on which IPv6 should be implemented is greatly reduced. This method should also decrease the cost of IPv6 implementation, and there are no real downsides to this approach now. Internal devices can then use proxies to translate the internal IPv4 traffic from IPv4 to IPv6. A proxy is a device (see Figure 6.2) located between client and server, most of the time placed in a DMZ. A client connects to a proxy, and the proxy connects to the outside world. This way, there is no direct connection between the client and the server, but all via the proxy. A proxy can translate the IPv4 and IPv6 addresses to each other.

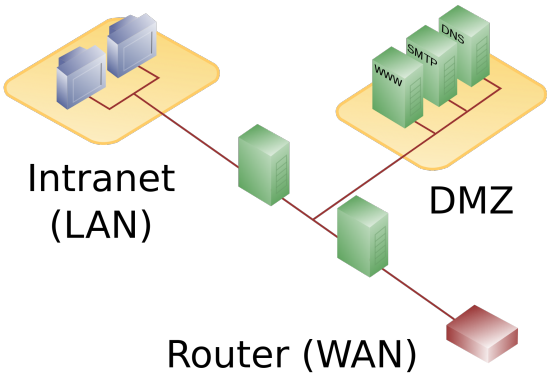


Figure 6.1: A DMZ is a method for creating network segmentation. Systems that should be easily accessible are placed more in front of the network topology. A secondary firewall can be used to protect the more vital systems. Network segmentation is a helpful method to implement security on a network level.[1]

Secondly, how to tackle internal routing. As mentioned before, there is no real reason for a company to go IPv6 internal. However, IPv6 is there and should be considered. It is advised not to implement IPv6 yet, but

[1] https://en.wikipedia.org/wiki/DMZ_(computing)
[2] https://hide-ip-proxy.com/what-is-a-proxy-server/

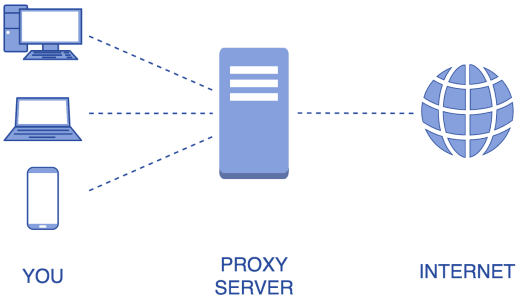


Figure 6.2: A proxy is a system that sits between a client and a server on a network. The client connects to the proxy, and the proxy connects to the Internet. This way, there will never be a direct link between the client and the Internet. [2]

to create an implementation plan. This plan can be laid aside, but yearly reviewed and has to be updated, in case necessary. Advice for IPv6 implementation is to streamline this with other technology life cycles. All technology goes through this cycle. The same goes for networking technology, and after a certain period, improved software and hardware is developed that should replace the old techniques. IPv6 can be added when there is a comprehensive update released to minimize cost. An update means downtime and downtime should be decreased to a minimum. A company may decide not to implement IPv6 this cycle. It then should analyze what the expected lifespan is of the next cycle. If it is, for example, five years, they should look if they expect that the company can rely on IPv4 for these next five years. If this company fails to make a proper analysis, it could lead to extra cost, because IPv6 is then implemented not during a technology life cycle. The implementation process is part of risk management and has to be taken into account when a company wants to support IPv6 internally.

6.3. Related Work

This section discusses the related work of this study. Other studies have been conducted on similar subjects to this study. Topics addressed in those studies are the discovery of IPv6 addresses, the allocation of IPv4 addresses, IP scanning, and IT misconfigurations.

6.3.1. IPv6 Addresses Discovery

Thus far, IPv6 was able to stay on the background, and although there were misconfigurations, researchers and malicious actors were not able to identify them because of the vast address space. The IPv6 is just too large to scan. A scan with modern tools over the entire IPv6 address space would take approximately 7.5×10^{23} year. IPv6 had the ability to make use of an incorrect method of security named security through obscurity. A method for security to keep it a secret, while not securing it properly. Because of the enormous IPv6 address space, it was next to impossible to identify misconfigured or improve systems. However, this changes when people register an AAAA record in DNS for their IPv6 address. As a reminder, the AAAA record is the link between a hostname (example.com) and an IPv6 address.

In 2017 and 2108, two articles published by Fiebig et al. [33] and Borgolte et al. [8] show methods how to recover IPv6 addresses from DNS. Fiebig et al. identified 5.8 million unique IP address, and Borgolte identified 2.2 million addresses. This number is lower then Fiebig but more difficult to detect and thus harder to mitigate. Combining the knowledge of those studies with our knowledge that a large percentage of IPv6 hosts are incorrectly configured could have large negative outcomes. To make matters worse, multiple interviewees stated that people register their IPv6 address in DNS because also for the owners of a host, an IPv6 address is difficult to remember. Next, for a valid certificate registering in DNS is required. For this, people are actively registering their IPv6 addresses in DNS, not realizing that this makes their host identifiable, putting an end to the 'security through obscurity' method.

6.3.2. IPv4 Address Allocation

Another study, by Richter et al. [84], identified fundamentals of the IPv4 exhaustion. They performed research on the current ecosystem of the IPv4 address space. Furthermore, they also mapped the blocks of IPv4 addresses is assigned to each RIR. Blocks of IP addresses are being assigned to RIRs per /8. Every /8 consist of roughly 16,7 million addresses. For a more visual representation, every IP address octet can range from 0-255. So an IP address is represented as **0-255**.0-255.0-255.0-255. The first octet (in bold) is assigned to each RIR. Table 6.1 shows an overview of the number of blocks in every RIRs responsibility. North-America, Europe, and Asia currently most ranges. While the Internet was being developed, IPv4 addresses were allocated without cost to parties who requested it. During that time, this process was not as monitored as today. Moreover, the first phase of registering an IPv4 address was performed by one man named John Postel [78]. By then, IPv4 addresses were allocated quite freely without the thought that IP could be a scarcity. By the time Africa and South America started to adopt the Internet, large blocks of IP addresses were already allocated to western countries. Countries who were part of the early adopters of the Internet feel less pain of IPv4 scarcity because they can make use of a large pool of IP addresses which were allocated to them in the past. They can rearrange this supply and use it as efficiently as possible. However, upcoming countries in Africa and South America have very few IP addresses to allocate in the first place. It is questionable if these seven blocks are enough even with IPv4 lifetime extension methods as NAT and CGN. In 2017 the adoption rate of Internet users in Africa was approximately 22% while Europe had the highest adoption rate of 80% [89]. It is expected that the adoption rate of Africa will grow in the future, and it is highly questionable if their current blocks will be sufficient for this future growth.

RIR	Number of Blocks (/8)
ARIN	100.9
APNIC	51.7
RIPE	48.6
LACNIC	11.1
AFRINIC	7.1
Other	29

Table 6.1: An overview of the assigned blocks to each RIR. Furthermore, the number of people living in each RIR are also added. Number blocks who fall into the other category are reserved by IANA, privately owned, or property of the United States Department of Defense.

Richter et al. also discusses the current routable IPv4 address space. They state that a third of all IPv4 addresses are currently not usable for routing, meaning that they can not be used for public routing. Moreover, they also discuss studies that found that even routable addresses are not being fully utilized. Thereupon is a large part of the IPv4 space still available. In theory, is it possible to renumber the IPv4 address space for more efficient routing but this will require both technical as political polices and guidelines. However, this renumbering will be problematic with the current landscape of the Internet. To summarize, they suggest three options for solving the IPv4 scarcity; IPv6, Carrier-Grade NAT, and partly reassigning address blocks. Like our findings, they expect that cost will determine which of these three will finally solve the IPv4 scarcity.

6.3.3. Scanning IPv6 space

As mentioned earlier in this thesis, there have been studies performed on the difference in port configuration on dual-stack servers. Czyz et al. published 'Don't Forget to Lock the Back Door! A Characterization of IPv6 Network Security Policy' and did a similar study as for this thesis [19]. They identified a difference in port configuration on IPv4 and IPv6. However, there are differences to this study. They scanned 520.000 dual-stack servers and 25.000 dual-stack routers while for this study, 12.7 million dual-stack hosts were scanned. Furthermore, they published in 2016, and by then the IPv6 adoption was only 10% compared to the 25% at the time of writing. The main difference between Czyz et al. and our study is that we attempted to answer the question of why there is a difference in configuration. Not only from a technical perspective but also from a people perspective.

Besides Czyz et al. another party performed scanning research on dual-stack hosts. However, Fehrenbach et al. [32] approach were also from a technical perspective. They performed a full port scan (65535 ports) on all dual-stack hosts on the Alexa one million. Furthermore, they performed version detection and product detection of the hosts, meaning that identification was made of the what kind of protocols was running on a server. It is possible to identify if a server is running a specific version or a specific distribution. If our

results could be compared with their work we could, for example, identify if specific distributions are more vulnerable for misconfiguration.

6.3.4. Misconfigurations

Finally, an article related to this study is Investigating System Operators' Perspective on Security Misconfigurations by Dietrich et al. [22]. This paper discusses a similar approach to this thesis to search for the human component in security misconfiguration. A study was performed from a system operators perspective on security misconfiguration. Next, root causes were determined on the factors that operators receive. There are similarities to this thesis; however, our study focused on misconfiguration in port configuration while Dietrich et al. looked into misconfiguration in a broader perspective. The study proofed to be helpful for a better apprehension of the human factor of misconfigurations.

6.4. Literature Adoption Theories

We analyzed the IPv6 adoption and IPv4 scarcity with adoption theories from literature. One compelling theory is the Technology Push/Market Pull framework [92]. Technology push is when R&D initiates the innovations while market pull comes from that the public requires new technology [10]. In the time that IPv6 was developed, the world was running on sufficient IPv4 addresses. IPv6 was created because it was expected that we would run out of IPv4 addresses; however, at that time, it was not required. It was expected that the market started adopting IPv6 from itself, but as we can see today, this was not the case. Techniques as NAT and Carrier-Grade NAT gained popularity over IPv6, which postponed the IPv6 adoption. Because of the techniques, there is no market pull for new technology. It can be discussed that there currently is some technology push. This could be one reason why there is a difference between the IP configurations on both protocols. Some hosting providers (VPS and web-server) on default, enable IPv6. System administrators do not ask for or are not aware that their system is now directly accessible over IPv6. Thus, the push from some ISPs could be a reason why on servers there is a difference in port configuration.

This problem could be analyzed with the technology push/market pull framework. Technology push is when R&D initiates the innovations while market pull comes from that the public requires new technology. Technology push has been evident through the past years where researchers have developed IPv6 while not required because IPv4 sufficed, visible from Google's data. However, some market pull developments are expected soon. The upcoming developments of the 5G network and the advances around the Internet of Things may break this path dependency and stimulate the migration from IPv4 to IPv6 [92].

Another principle that could apply to the development and adoption is radical and incremental innovation. Incremental innovation states that small and safe changes are made to a known product or service. This method is a safer and less risky business approach. While radical innovation is a more complex process where an entirely new process or service is developed. IPv6 can be seen as a radical innovation, and NAT and CGN are incremental innovations. These incremental innovations on IPv4 have increased the ability of IPv4 and are increasing its lifespan [92]. Without a protocol as NAT, the amount of IP addresses would not have been sufficient years ago. These suggestions could be confirmed by the diffusion patterns in new high-tech products. Throughout history, it has been difficult to predict what the adaption phase of a product would become. The chances are that adoption takes multiple years before an innovation finally takes off [70]., commonly linked to the network effect of a product. The network effect states that a good or service can increase in value if they are being used in a larger network [28]. An example of this is that it is more profitable to develop a mobile app for Android or iOS than for Windows Phone because the number of users of Windows phones is much lower than the other two. This effect can also be seen in the IPv6 adoption, the amount of users is low, so it is less profitable to create accessibility over IPv6. Some parties could influence the network effect. Large technology organizations are trying to take a leading role in IPv6 implementation. Google, Facebook, and YouTube all adopted IPv6. Furthermore, the Dutch government is trying to increase the network effect by enabling support on their services [57, 110].

Another theoretical principle that applies to IPv6 adoption is path dependency [20]. Path decency is a famous principle where historical events influence future development. One of the well-known examples is the QWERTY keyboard. It was developed during the time of the typewriters to slow down the typing speed. If typed on a typewriter to fast chances were that hammers would get stuck in each other. However, our

modern keyboards do not have these problems anymore but are still according to the QWERTY layout. Other keyboard settings (Dvorak) have proven to be faster, but it seems impossible to change to this keyboard setting. We know that Dvorak will increase productivity, but the switching costs are enormous. The adoption of IPv6 has similarities to this problem. All systems are currently running on IPv4 (QWERTY), and it seems easier to continue working with this then implement IPv6 (Dvorak). We are used to IPv4, not only the possibilities but also the limitations, and it is hard to unlearn IPv4 configuration and start over with IPv6. One respondent confirmed this and stated that we were stuck in IPv4. In an economic perspective, these are named as switch barriers [49]. The cost of unlearning IPv4, reeducation of IPv6, and the risks involved can lead to unwanted costs.

6.5. Similar Technology Adoption Cases

We were interested if we could find similar cases as the transition from IPv4 to IPv6. We found a similar instance that is currently in the end phase of the adoption. In the past years, the number of HTTPS has been growing exponentially. The old protocol HTTP has lost its popularity, and many websites migrated from HTTP to HTTPS. If looked at the adoption phase of HTTPS, there are a few distinct reasons why the adoption has picked up in such an exponential way. HTTPS is a secure connection between two parties, where all data in transit is encrypted. For setting up an HTTPS connection, the server requires a certificate that proves that he is whom he says he is. Certificate Authorities, who check an identity, review these certificates. There is a subscription fee involved for those certificates. These fees' withheld the adoption of HTTPS at non-corporate organizations. On the other hand, the reasons for adoption are clear. HTTPS websites are considered safer by the public, modern browsers mark HTTP websites as insecure, and secure websites are being ranked higher in search engines [9]. However, the real adoption came when a company, Let's Encrypt, came with the ability to have free certificates. In April 2016 they launched a method to greatly decrease the overhead in certificate verification, which made the road clear for mass adoption. In 2018, they already issued 350 million certificates leading to be responsible for 80% of all certificates on the internet [18, 114].

So, what lessons can be identified in this for IPv6 adoption? There has not been, nor is it likely that an event will kick-start the adaptation of IPv6 similar to that that Let's Encrypt had on the HTTPS adoption. However, some cases could aid in IPv6 adoption. An example is awareness creation, IPv4 is becoming a legacy protocol, and it would aid if browsers would give feedback to the users if their connection is established over IPv6. They currently warn for HTTP websites and similar methods can be used for IPv4. An organization, internet.nl, already have tools for website testing and analyze if it is future proof. If technology giants as Google or Firefox pick this up, it is believed to increase knowledge and awareness. Google already is aiding in this process by lower-ranking IPv4 websites.

6.6. Technology Disruption

We found from literature and our interviews that there is a general awareness and knowledge problem concerning IPv6, which is presumed the reason for IPv6 misconfiguration. We expect that this can be lead back to the current adoption speed. However, during our interviews and research, we found reasons that could inflict the adoption speed, IPv6 security, and security in general. In this section, we would like to discuss some expected events or initiatives that could break the IPv4 chain and speed up the IPv6 adoption and awareness.

1. First is the general state of security in companies and organizations. Multiple respondents mentioned that the General Data Protection Regulation (GDPR) affects the security awareness level of companies. Because of GDPR, companies are now responsible for handling data with care and will be fined if they fail to do so. Cyber-crime is an often-mentioned risk for a company, and GDPR gave a monetary reason for improving the security posture, improving security posture could increase the risk concerning improper IPv4 and IPv6 configuration.

2. A second reason is the rise of the Internet of Things. IoT devices have gained popularity in the past years. 58% of the companies are investing in IoT possibilities [61]. The rise of IoT devices can significantly increase the number of required IP addresses, which would put even more pressure on the IPv4 scarcity. However, this is highly questionable because it is unknown if all these IoT devices are going to use the IP protocol. On average is the IP protocol quite power-hungry while most IoT devices are developed to run years on a small button cell.

3. The third is education. Education thus far has mainly been focused on IPv4. Multiple respondents stated that education concerning IPv6 lacks behind. It is slowly starting to be adopted in educational programs. One respondent who worked a computer science guest lecturer at a university stated that the only IPv6 education students got was one lecture by him. Furthermore, Cisco only recently added IPv6 to their training modules. More education could increase the knowledge and awareness with the protocol, resulting in fewer configuration mistakes.

4. Finally, we attempted to gather information on IPv6 security incidents by asking our respondents, reaching out to a leading cybersecurity company, and by a by searching in public sources. None of the parties could give us one example of a substantial breach where IPv6 security played a role. Of course, there is some bias because companies attempt to hide security breaches in fear of damaging their public image. Even when the public is made aware of such a breach, there is generally only minimal information. An IPv6 security incident could increase awareness and show the security risks. Nonetheless, it could also lead to more fear of IPv6 and the start of blocking all IPv6 traffic.

6.7. Lessons to Actions

Finally, in this discussion, we would like to elaborate on the different actors related to the IPv6 problem. We want to discuss the actors and what their responsibilities could be. A problem we identified is that there is no responsible actor for the migration process. From the interviews, we also discovered that there is no clear answer who should take responsibility. It is not that one party does not take her responsibility; it is unclear which party should take the responsibility to address and solve this problem. Furthermore, IPv6 has proven by the slow adoption, that it is not the ideal solution for IP scarcity. Nevertheless, there are no clear alternatives. Other protocols as CGN and NAT also have its limitations but are becoming more rooted on the Internet. Both solutions have limitations, and different actors feel those different limitations. The different actors are enumerated below:

1. **ISP** ISPs are commercial institutions, where profit is their main factor. An ISP will ponder between IPv6 adoption and CGN based on the costs of implementation. Thus far has proven that CGN is a more profitable solution for them. Moreover, from the interviews, we found that ISPs are presumed to have the most knowledge on IPv6. They should try to spread this knowledge. Next, because most traffic on the Internet is Google, Facebook, and YouTube, and those services are fully IPv6 ready ISP should roll-out IPv6 at their clients faster to increase the network effect. The increased network effect improves the awareness and hopefully, the number of misconfigured devices.

2. **Hosting Providers** On average hosting providers (Web-hosting and VPS) are progressive in the IPv6 adoption. Popular websites are starting to run dual-stack, nevertheless not all websites are available on both protocols. They should continue the adoption and the roll-out of IPv6 on their servers. Furthermore, they could aid their clients with IPv6 awareness and informing that a VPS is IPv6 enabled. Possibilities are to create a *things to do after installing a VPS* and add *configure the IPv6 firewall* on it or to design an interactive GUI to aid with the configuration.

3. **Government** The government could play a role in IPv6 adoption. Most parties stated that the government should not get involved, but it could aid in the adoption with rules and regulations. Cyber-crime and security are factors important for a government, and if CGN would put too much pressure on this, they could act. From the example in Belgium, it is proven that regulations concerning CGN will increase IPv6 adoption.

4. **Education** All interviewees agreed that lack of knowledge was a reason for the misconfiguration on the hosts. Educational institutions are responsible for spreading knowledge and educating network and security professionals. Currently, the main curriculum and all the examples are based on IPv4. These institutions should revise their program and focus more on IPv6. Another example is the CISSP program [15]. CISSP is currently the certification required for practitioners, managers, and executives who are interested in security practices and principles. IPv6 is mentioned in the book but only briefly. It is stated as a future protocol with a slow adoption rate. Security implications of IPv6 are in any way mentioned in the book. Addressing this issue by CISSP could aid in the management awareness creation of IPv6.

5. **Internet Governance** Finally, internet governance could play a role in creating awareness and knowledge. Engineers will not start working on IPv6 as long as their managers will not allocate time and funds

on this. A goal should thus be to educate and teach IPv6 to management. Internet Governance should play a role in this. For example, RIPE offers training material in Internet protocols like BGP, DNS, and IPv6. Nonetheless, all the training material concerning IPv6 is focused on technical implementation and deployment. They do not offer a *executive summery* course of IPv6 were IPv6 is not explained on a technical level but more with a high-level approach. If possible, they should create a short training program that educates managers in the new possibilities that IPv6 brings and the risks involved. They should elaborate on the risks and possibles concerning, no IPv6 implementation, half configurations, and a fully running IPv6 network. RIPE focused thus far their efforts on educating engineers. However, based on our findings, they should move their focus to the management level of an IT department.

6.8. Summary Discussion

In this chapter, we discussed multiple subjects concerning the adoption and configuration of IPv6. We elaborated on two IPv6 implementation cases. Interestingly was that the reason for implementation was not the technical features of IPv6 but the IP scarcity from clients connecting to servers. In both cases, IP addresses are used in the detection of cyber-crime, and CGN impacts the usability of crime detection tools. We added a related work section that discusses articles related to this thesis, both on IP scanning as the human perspective. Next, we elaborated on two different methods of how IPv6 could be implemented in a network. There is a difference between internal and external implementation. It is advised first to make services accessible from the Internet over IPv6. If this is implemented properly, IPv6 could be implemented internally. Subsequent, we discussed some technology adoption theories and saw similarities between the adoption of IPv6 and those theories. We talked about methods of innovation and how new products are deployed in the market. Following, we discussed the HTTPS adoption case. HTTPS is a protocol for encrypted transfer, were websites moved from HTTP to HTTPS. We analyzed if we were able to see similarities between the HTTP, HTTPS case, and the transition from IPv4 to IPv6. Finally, we elaborated on the different actors in the field and what their responsibilities could be to increase the IPv6 adoption and improve the IPv4, IPv6 port configuration.

7

Conclusion

This study had the goal of discovering why there are differences in port configuration between IPv4 and IPv6. Researchers from the University of Amsterdam indicated these differences and this study elaborated upon this work with the question *why*. For this, an exploratory approach was used by identifying the ownership of the presumed misconfigured devices and questioning if the responsible actors were aware of their configuration. Furthermore, we approached different security and network specialists with experience from the field and discussed the port configuration processes and the current state of IPv6 adoption. In this final chapter, the research and sub-questions are evaluated, and we will discuss the practical implications. Finally, recommendations for future work will be given, and a personal reflection is added.

7.1. Research Questions

This section will answer the main research question and its corresponding sub-questions. The research question is formed based on the goal of this study. First, the sub-questions will be repeated and answered before continuing to the main research question.

What actors play a part in the IPv6 configuration?

We found that multiple actors participate in IPv6 configuration. There is a difference in the number of actors involved depending on the company size. In smaller companies, there is less separation of work, and the person responsible for security and availability is the same person. In smaller organizations, the system administrator is the only actor. In larger companies, a client with a request, a system administrator, a network engineer, and a security officer participate in IPv6 configuration.

Which guidelines are considered to be most critical for port-based firewalling policies?

Companies follow the least privileged method, stating that the attack surface should be as small as possible. If it is not necessary to have a port open, it should not be open. Most companies follow a similar method where a client makes a request. If it is a common request, a protocol dictates how this request should be handled. A standard protocol is to translate this request into a task. This task is then further separated into sub-tasks, which are evaluated by a security officer. If a security officer approves the tasks, the change is executed. An uncommon task is evaluated into more depth and a risk assessment is based on the consequences of the open port. If approved, a task is created, and a similar procedure is executed. On average, it can be stated that the larger the company, the more formal the process. In smaller companies, it is the system administrator that weighs the risk and executes the change on all systems.

What processes lead to IPv4/IPv6 misconfigurations and inconsistencies that can be found in the wild?

This question is answered with the aid of interviews performed with subjects from the data set of misconfigured devices. We found multiple reasons why configuration mistakes were made. We found scenarios were an IPv6 daemon was disabled, a firewall script was used that forgot to configure the IPv6 ports, the DNS entry was not changed correctly, and because of certain features of a configuration method a specific firewall. Interviewees stated that the underlying reason was the complexity, insufficient awareness, and knowledge of IPv6.

Is there evidence of active exploitation of IPv6 misconfigurations?

This question has proven difficult to answer because during this study we were unable to find security incidents with the consequences of an IPv6 misconfiguration. This question was attempted to answer with interviews, literature study, and news articles. It is not possible to state that there are no security incidents on IPv6, because companies who had an incident do not share the reason behind an incident with public sources. Most security incidents are shared at all.

What is the reason for the existence of inconsistencies in open port policies on dual-stack hosts?

Combining the knowledge gathered from the interviews and the literature study, we can state that IPv6 configuration has proven to be a difficult subject. The protocol is complicated; however, the adoption of IPv6 is starting to increase. Not all parties are ready for IPv6, and the general awareness and knowledge of IPv6 lacks behind. Thus far, there are little use cases for IPv6 implementation, which results in no priorities on IPv6 from management. At the moment, IPv4 extension technologies are more popular then IPv6, resulting in IPv6 being in no men's land with too little awareness and knowledge. This situation leads to improper and half configurations of firewalls and differences in open ports on dual-stack hosts.

7.2. Practical Implications

There are different types of insights that can be gained from this thesis. It attempted to discover reasons on both, managerial and technical perspective, behind misconfiguration on IPv6. Although there are no direct consequences of an IPv6 misconfiguration, it increases the attack surface for malicious actors. A company should always minimize the risk of exposure and thus properly configure IPv6. However, management should be aware of these risks and allocate appropriate resources for properly implementing or disabling IPv6. This study could aid in that problem as it provides intelligible handles on security implications of a misconfiguration for both parties, although it is written mainly from a technical IPv6 perspective. Altogether, there is a presumed relationship between the awareness level and IPv6 misconfigurations. This study, together with the work of the researchers from the UvA, can create this awareness. It constitutes a warning for both managers and engineers to check whether their IPv6 configuration is handled correctly. It is expected that as soon as the problem is identified, engineers know whether their systems are IPv6 enabled, a solution could be implemented relatively easily.

Concretely, people should make themselves aware of the problem. From there, a solution can be sought because technical employees, as well as management employees, can search for one. This thesis can be seen as a tool to familiarize oneself and one's management with the problem. This familiarization is fundamental because it means that resources can be made available to address the issue in several ways because the consequences of IPv6 misconfiguration are no longer seen as acceptable risks. For instance, personnel could receive training on IPv6, can spend time on the configuration process, or the team could be expanded. From there, companies can get their IP configuration in order and keep it that way.

7.2.1. Actor Action Points

This study searched for why there are inconsistencies in the port configuration on dual-stack hosts. We attempted to identify possible sources of the problem and discussed our results with the aid of similar migrations and adoption theories from literature. However, these lessons identified should be changed to lessons learned. We identified that the knowledge and awareness level is too low concerning these misconfigurations. Engineers and networking enthusiast are working and experimenting with IPv6, but management is

not aware of the additional risks concerning a half and inconsistent configurations on IPv6. In this study, we identified actors, and we suggest action points what the different actors can undertake.

1. **ISP** It is assumed that most knowledge concerning IPv6 is at ISPs. They should spread this knowledge by spreading and educating this knowledge to their clients. Furthermore, larger internet parties (Google, Facebook, Amazon) are supporting IPv6. The limiting factors are ISPs who do not supply an IPv6 address to their clients. They should continue and increase the speed of the roll-out of IPv6 and decrease the number of devices behind a CGN.

2. **Education** Educational institutions should revise their curriculum and focus more on IPv6 than IPv4. If an improved number of engineers are being educated in IPv6, more awareness will be created, and it is expected that fewer misconfigurations will be made. The main source of IPv6 misconfiguration is inexperience and unawareness. Not only universities and colleges should revise their programs but also training programs, and computer security certifications should focus their attention more on IPv6.

3. **Internet Governance** Finally, Internet Governance parties should address this problem, for example, RIPE. RIPE is an institution who should take responsibility in the education and start informing about the consequences of IPv6 adoption. Thus far, they focused their energy on training engineers who implement IPv6. However, a forgotten group is the managers who should also be trained in the consequences of the current IPv6 adoption. Managers should be made aware of the consequences for their company, and what would the effect, if they decide not to implement IPv6, half configure IPv6, or implement IPv6 in their networks. Thus far, management has not been given the possibility to get acquainted and educated with IPv6. Without this possibility, they are not able to make a correct and informed decision.

7.3. Generalization and Limitations

For this study, data was collected with the aid of interviews from network and security experts. It must be stated that all interviewees were living in European countries and working for European or American companies. One of our findings was that there is not enough IPv4 scarcity to feel the negative effect of the IPv4 scarcity. However, we can state this for western countries, and did not collect data from African/South-American countries. We know from the literature that currently, most IPv4 addresses are allocated to the western countries, and when third world countries started to be connected, large chunks of IP addresses were already allocated and reserved. Furthermore, we state that there is no real need for IPv6 because the IPv4 infrastructure is still sufficient (with the IPv4 lifetime extension methods). However, this statement is questionable for third world countries. It would be useful to perform a study focused on third world countries to look into the current deployment of IPv4 and the expectations if their current pool of IPv4 is large enough for a reliable connection to the Internet. Especially in mind with the rising number of connected people in those countries. To summarize, most of our findings apply to western countries, and additional work is required to discover the current and future state of these countries.

7.4. Future Work

During this study, we identified multiple gaps that impact the security concerning IPv6 port configuration and adoption. These gaps were identified by the researchers or by a suggestion of an interviewee.

The first is a technical recommendation suggested by a hosting provider. Currently, the data set shows misconfigured devices. But it is unknown what the underlying operating systems are. During this study, we encountered Ubuntu and BSD distributions, but this is only a small sample off all operating systems available. A future extension could be to analyze misconfigured devices and see if there are operating systems more prone to misconfiguration. With knowledge of operating systems, it would be possible to draw recommendations on how to change these operating systems or implement safety measures that alert system administrators for possible misconfigurations.

Another technical finding could be to analyze the current use cases of dual-stack hosts. This research focused on misconfigured devices and did not review what the entire landscape of dual-stack configured devices are. It would be useful to know what the use cases are for the dual-stack configured devices (Web-hosting, VPS). If it is known what devices are more vulnerable for mistakes, a more specific solution could be drafted, and

the leading target group could be approached. If the target group is known, a solution could be more specific on this. If the main target group is approached first, the attack surface would be reduced quickest.

During this study, IP misconfigurations were linked to IPv6 adoption. A higher adoption rate would lead to more awareness and presumably to fewer configuration mistakes. During this study, the initial data set was sorted, and IPv6 adoption rates were added per TLD (Appendix C). A more in-depth statistical analysis on IPv6 adoption and configurations could be a useful addition to understand the scope of IPv6 configuration better. If statistical methods could prove a clear correlation between the two, a strong argument could be formed for the urgency of IPv6 adoption.

At the moment, the Internet uses IPv4 with transition methods and IPv6. For maintaining both infrastructures, costs are involved. It would be beneficial from a managerial perspective if analysis could be performed on the cost of maintaining these infrastructures. Not only the current cost but also the expected costs of maintaining IPv4. It is expected that the cost of IPv4 will increase in the future with CGN and the auctions of IPv4 ranges. If a financial analysis could shed light on the expected cost, a more weighted decision-making process could be made about the implementation of IPv6. Currently, there are too many unknowns concerning implementation costs.

7.5. Personal Reflection

While I enjoyed performing this study, there were some roadblocks along the way. For this thesis, large parts of fundamental knowledge were required, both on the managerial as the technical perspective of IPv6. A literature study aided in getting this knowledge up to the required level. However, there is a literature gap on IPv6 configuration and adoption on a managerial level. Currently, most articles are about the technical implementation and features of IPv6, making it difficult to create a theoretical foundation for this problem.

Secondly, the amount of gathered data. The initial data set was immense, and it was difficult to identify the possible entries. This process took longer than expected to identify the correct subject. Even after identifying the interesting subjects, it proved difficult to recruit people for an interview. We noticed that people were unaware of IPv6, so could not aid us in our research and people who were aware did not want to share information because of client data, making recruitment challenging and hard to overcome. However, when evaluation the interview phase, it was enjoyable to have these conversations. The diversity of people and their different views on IPv6 gave food for thought. Furthermore, it was a pleasant experience when those thoughts became concepts, and links were made with the material taught in the MOT curriculum. The people who were willing to help were thankful for our efforts. In general, people agreed that this study was a relevant topic and thanked us for informing them of a potential misconfiguration. Most of them tried to resolve the issue within a day or informed the responsible party about the configuration.Knowing that we had a small contribution in a more secure Internet, is a positive feeling.

In the end, I mostly learned that IT security is not always challenging. The solution is often relatively easy, but awareness is the problem. I often approached computer security as a technical problem, and I enjoyed looking at this via the managerial perspective. Furthermore, the analysis of the data set opened my eyes to the scale of the Internet, how large it is, and how efficiently it works. The Internet is a beautiful piece of technology.

Bibliography

[1] Ioannis Agrafiotis, Jason R C Nurse, Michael Goldsmith, Sadie Creese, and David Upton. A taxonomy of cyber-harms: Defining the impacts of cyber-attacks and understanding how they propagate. *Journal of Cybersecurity*, 4(1), 10 2018. ISSN 2057-2085. doi: 10.1093/cybsec/tyy006. URL https://doi.org/10.1093/cybsec/tyy006.

[2] Victor Ajayi. Primary sources of data and secondary sources of data. 09 2017. doi: 10.13140/RG.2.2.24292.68481.

[3] Amazon. Alexa One Million, 2019.

[4] Madhav Bahl. OSI Model Layers — "Explained". *Medium*, 2018. URL https://medium.com/@madhavbahl10/osi-model-layers-explained-ee1d43058c1f.

[5] Kevin Beaver. Inbound vs. outbound firewall rules: Comparing the differences, 2014. URL https://searchsecurity.techtarget.com/answer/Comparing-firewalls-Differences-between-an-inbound-outbound-firewall.

[6] Matt Bishop. What is computer security? *IEEE Security and Privacy*, 1(1):67–69, January 2003. ISSN 1540-7993. doi: 10.1109/MSECP.2003.1176998. URL http://dx.doi.org/10.1109/MSECP.2003.1176998.

[7] Bernhards Blumbers. Hedgehog in the Fog : Creating and Detecting IPv6 Transition Mechanism-Based Information Exfiltration Covert Channels. *NordSec*, 2016. doi: 10.1007/978-3-319-47560-8.

[8] K. Borgolte, S. Hao, T. Fiebig, and G. Vigna. Enumerating active ipv6 hosts for large-scale security scans via dnssec-signed reverse zones. In *2018 IEEE Symposium on Security and Privacy (SP)*, pages 770–784, May 2018. doi: 10.1109/SP.2018.00027.

[9] Kevin Borgolte, Tobias Fiebig, Shuang Hao, Christopher Kruegel, and Giovanni Vigna. Cloud strife: Mitigating the security risks of domain-validated certificates. 07 2018. doi: 10.1145/3232755.3232859.

[10] Alexander Brem and Kai-Ingo Voigt. Integration of market pull and technology push in the corporate front end and innovation management insights from the german software industry. *Technovation*, 29(5):351 – 367, 2009. ISSN 0166-4972. doi: https://doi.org/10.1016/j.technovation.2008.06.003. URL http://www.sciencedirect.com/science/article/pii/S0166497208000898. Technology Management in the Service Economy.

[11] James Broad and Andrew Bindner. Hacking with Kali. *Hacking with Kali*, pages 185–199, 2014. doi: 10.1016/B978-0-12-407749-2.00022-7.

[12] Calyptix. Is Port Scanning Legal? Answers for IT Companies, 2017. URL https://www.calyptix.com/top-threats/port-scanning-legal-answers-companies/.

[13] V. Cerf and R. Kahn. A protocol for packet network intercommunication. *IEEE Transactions on Communications*, 22(5):637–648, May 1974. ISSN 0090-6778. doi: 10.1109/TCOM.1974.1092259.

[14] Deka Ganesh Chandra, Margaret Kathing, and Das Prashanta Kumar. A comparative study on ipv4 and ipv6. *2013 International Conference on Communication Systems and Network Technologies*, pages 286–289, 2013.

[15] Mike Chapple, James Michael Stewart, and Darril Gibson. *(ISC)2 CISSP Certified Information Systems Security Professional Official Study Guide*. SYBEX Inc., Alameda, CA, USA, 8th edition, 2018. ISBN 1119475937, 9781119475934.

[16] M Cotton, L Vegoda, and B Haberman. Special-Purpose IP Address Registries. Technical report, April 2013. URL https://www.rfc-editor.org/info/rfc6890.

[17] Stephen Crocker. How the Internet Got Its Rules, 2009. URL https://www.nytimes.com/2009/04/07/opinion/07crocker.html?_r=1.

[18] Dan Cvrcek. Let's Encrypt in the spotlight - Magic of Security, 2017. URL https://magicofsecurity.com/lets-encrypt-in-the-spotlight/.

[19] Jakub Czyz, Matthew Luckie, Mark Allman, and Michael Bailey. Don't Forget to Lock the Back Door! A Characterization of IPv6 Network Security Policy. *NDSS*, pages 21–24, 2016.

[20] Paul David. Path dependence, its critics and the quest for 'historical economics'. *EconWPA, Economic History*, 01 2005.

[21] S. Deering and R. Hinden. Internet Protocol, Version 6 (IPv6). *IETF*, 1998. URL https://tools.ietf.org/html/rfc2460.

[22] Constanze Dietrich, Katharina Krombholz, Kevin Borgolte, and Tobias Fiebig. Investigating system operators' perspective on security misconfigurations. pages 1272–1289, 10 2018. doi: 10.1145/3243734.3243794.

[23] Zakir Durumeric, Eric Wustrow, and J. Alex Halderman. Zmap: Fast internet-wide scanning and its security applications. In *Presented as part of the 22nd USENIX Security Symposium (USENIX Security 13)*, pages 605–620, Washington, D.C., 2013. USENIX. ISBN 978-1-931971-03-4. URL https://www.usenix.org/conference/usenixsecurity13/technical-sessions/paper/durumeric.

[24] Vincent Van Der Eijk and Erik Lamers. A Comparative Security Evaluation for IPv4 and IPv6 Addresses. pages 1–9, 2019.

[25] Tom Espiner. IANA allocates final IPv4 address blocks, 2011. URL https://www.zdnet.com/article/iana-allocates-final-ipv4-address-blocks/.

[26] Europol. Are you sharing the same IP address as a criminal? Law enforcement call for the end of Carrier Grade NAT (CGN) to increase accountability online, 2017.

[27] Exploit DB. Exploits Database by Offensive Security, 2018. URL https://www.exploit-db.com/.

[28] Joseph Farrell and Paul Klemperer. Coordination and Lock-In: Competition with Switching Costs and Network Effects. *Handbook of Industrial Organization*, 3(06):1967–2072, 2007. ISSN 1573448X. doi: 10.1016/S1573-448X(06)03031-7.

[29] James P. Farwell and Rafal Rohozinski. Stuxnet and the future of cyber war. *Survival*, 53(1):23–40, 2011. doi: 10.1080/00396338.2011.555586. URL https://doi.org/10.1080/00396338.2011.555586.

[30] Friso Feenstra. Rabobank IPv6 - What happend after, . URL https://www.ripe.net/support/training/ripe-ncc-educa/presentations/friso-feenstra-rabobank-ipv6-what-happened-after.pdf.

[31] Friso Feenstra. That is why Rabobank has IPv6. . URL https://ripe74.ripe.net/wp-content/uploads/presentations/3-That-is-why-Rabobank-has-IPv6.pdf.

[32] Patrik Fehrenbach. Security Evaluation of Dual-Stack Systems. *Troopers*, 2016. URL https://www.troopers.de/media/filer_public/81/81/81819c66-3db3-42c7-8a53-c6bfa30c08f9/tr16_ipv6_sec_summit_security_evaluation_of_dual-stack_systems_pf.pdf.

[33] Tobias Fiebig, Kevin Borgolte, Shuang Hao, Christopher Kruegel, and Giovanni Vigna. Something from nothing (there): Collecting global ipv6 datasets from dns. pages 30–43, 02 2017. ISBN 978-3-319-54327-7. doi: 10.1007/978-3-319-54328-4_3.

[34] Tobias Fiebig, Kevin Borgolte, Shuang Hao, Christopher Kruegel, Giovanni Vigna, and Anja Feldmann. In rDNS We Trust: Revisiting a Common Data-Source's Reliability. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 10771 LNCS:131–145, 2018. ISSN 16113349. doi: 10.1007/978-3-319-76481-8_10.

[35] FOX IT. mitm6 – compromising IPv4 networks via IPv6, 2018. URL https://blog.fox-it.com/2018/01/11/mitm6-compromising-ipv4-networks-via-ipv6/.

[36] Oliver Gasser, Quirin Scheitle, Pawel Foremski, Qasim Lone, Maciej Korczynski, Stephen D. Strowes, Luuk Hendriks, and Georg Carle. Clusters in the Expanse: Understanding and Unbiasing IPv6 Hitlists. *Proceedings of the 2018 Internet Measurement Conference (IMC '18)*, 2018. URL http://arxiv.org/abs/1806.01633.

[37] David Gewirtz. The astonishing hidden and personal costs of IT downtime (and how predictive analytics might help), 2017.

[38] Barney G. Glaser and Anselm L. Strauss. *The Discovery of Grounded Theory*. Routledge, jul 2017. ISBN 9780203793206. doi: 10.4324/9780203793206. URL https://www.taylorfrancis.com/books/9781351522168.

[39] Google. IPv6 Adoption, 2019. URL https://www.google.com/intl/en/ipv6/statistics.html.

[40] Andy Greenberg. Meet 'Project Zero,' Google's Secret Team of Bug-Hunting Hackers | WIRED, 2014. URL https://www.wired.com/2014/07/google-project-zero/.

[41] Jaber Gubrium, James Holstein, Amir Marvasti, and K.D. McKinney. *The SAGE Handbook of Interview Research: The Complexity of the Craft, second edition*. 01 2012. doi: 10.4135/9781452218403.

[42] Surbhi Gupta, Abhishek Singhal, and Akanksha Kapoor. A Literature Survey on Social Engineering Attacks : Phishing Attack. *International Conference on Computing, Communication and Automation (ICCCA)*, pages 537–540, 2016. doi: 10.1109/CCAA.2016.7813778.

[43] Have I been Pwned. Pwned Passwords, 2019. URL https://haveibeenpwned.com/Passwords.

[44] Thomas Holt. Know Your Enemy : The Social Dynamics of Hacking. *The Honeynet Project*, 2012.

[45] Bill Holtsnider and Brian Jaffe. *IT Manager's Handbook*. Elsevier, Burlington, 1st edition, 2010. ISBN 9780123751102.

[46] Matthew Hughes. Bots drove nearly 40% of internet traffic last year — and the naughty ones are getting smarter, 2019.

[47] IPv4.GLOBAL. IPv4 Address Auctions, 2019. URL https://auctions.ipv4.global/.

[48] Mohammad Hanif Jhaveri, Orcun Cetin, Carlos Gañán, Tyler Moore, and Michel Van Eeten. Abuse reporting and the fight against cybercrime. *ACM Comput. Surv.*, 49(4):68:1–68:27, January 2017. ISSN 0360-0300. doi: 10.1145/3003147. URL http://doi.acm.org/10.1145/3003147.

[49] Michael A Jones, David L Mothersbaugh, and Sharon E Beatty. Switching barriers and repurchase intentions in services. *Journal of Retailing*, 76(2):259 – 274, 2000. ISSN 0022-4359. doi: https://doi.org/10.1016/S0022-4359(00)00024-5.

[50] Kali Linux. Kali Linux | Penetration Testing and Ethical Hacking Linux Distribution, 2018. URL https://www.kali.org/.

[51] Sumit Kumar, Computer Science, and Guru Nanak. IPv6 Network Security using Snort. 2(8):17–22, 2013.

[52] Barry M Leiner, Vinton G Cerf, David D Clark, Robert E Kahn, Leonard Kleinrock, Daniel C Lynch, Jon Postel, Larry G Roberts, and Stephen Wolff. Brief History of the Internet. 1997.

[53] Andrew Lerner. The Cost of Downtime, 2014. URL https://blogs.gartner.com/andrew-lerner/2014/07/16/the-cost-of-downtime/.

[54] John Leyden. RockYou hack reveals easy-to-crack passwords, 2010. URL https://www.theregister.co.uk/2010/01/21/lame_passwords_exposed_by_rockyou_hack/.

[55] Frank Li, Zakir Durumeric, Jakub Czyz, Mohammad Karami, Michael Bailey, Damon McCoy, Stefan Savage, and Vern Paxson. You've got vulnerability: Exploring effective vulnerability notifications. In *25th USENIX Security Symposium (USENIX Security 16)*, pages 1033–1050, Austin, TX, August 2016. USENIX Association. ISBN 978-1-931971-32-4.

[56] Richard Lippmann, Seth Webster, and Douglas Stetson. The effect of identifying vulnerabilities and patching software on the utility of network intrusion detection. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2516 (October):307–326, 2002. ISSN 16113349. doi: 10.1007/3-540-36084-0_17.

[57] Logius. Overheidsbreed IPv6-nummerplankader. pages 1–24, 2016.

[58] Georgios Loukas. Protection against Denial of Service Attacks : A Survey. *The Computer Journal*, 53(7): 1020–1037, 2009. doi: 10.1093/comjnl/bxp078.

[59] David Lumb. Former Equifax CEO blames breach on one IT employee, 2017.

[60] Steve Manzuik, Ken Pfeil, and Andre Gold. *Network Security Assessment*. 2007. ISBN 978-1-59749-101-3.

[61] Bill Martorelli and Michele Pelino. Global IoT Services For Connected Business Operations. *Q4 2018 The Forrester Wave ™ : Global IoT Services For Connected Business*, 2018.

[62] Paul McNamara. Why Belgium leads the world in IPv6 adoption, 2016. URL `https://www.networkworld.com/article/3100968/why-belgium-leads-the-world-in-ipv6-adoption.html`.

[63] MITRE. National Cybersecurity FFRDC, 2018. URL `https://www.mitre.org/centers/national-cybersecurity-ffrdc/who-we-are`.

[64] M Monshizadeh, P Naldurg, and V N Venkatakrishnan. MACE: Detecting privilege escalation vulnerabilities in web applications. *Proceedings of the ACM Conference on Computer and Communications Security*, pages 690–701, 2014. ISSN 15437221. doi: 10.1145/2660267.2660337.

[65] Keith Moore and Brian E. Carpenter. Connection of IPv6 Domains via IPv4 Clouds. *IETF*, 2001. URL `https://tools.ietf.org/html/rfc3056`.

[66] Stéphane Nappo. CISO of the Week, Stéphane Nappo, Société Générale - Cyber Startup Observatory, 2019. URL `https://cyberstartupobservatory.com/cyber-startup-observatory-ciso-week-stephane-nappo-societe-generale/`.

[67] NCSC. Beveiligingsadviezen, 2018. URL `https://www.ncsc.nl/actueel/beveiligingsadviezen`.

[68] NCTV. Ontwrichting van de maatschappij ligt op de loer, 2019. URL `https://www.nctv.nl/actueel/nieuws/2019/csbn-2019-ontwrichting-maatschappij-ligt-op-de-loer.aspx`.

[69] Netcraft. Web Server Survey, 2019. URL `https://news.netcraft.com/archives/category/web-server-survey/`.

[70] Roland Ortt and Jan Schoormans. The pattern of development and diffusion of breakthrough communication technologies. *European Journal of Innovation Management*, 7:292–302, 12 2004. doi: 10.1108/14601060410565047.

[71] Tao Peng, Christopher Leckie, and Kotagiri Ramamohanarao. Survey of network-based defense mechanisms countering the dos and ddos problems. *ACM Comput. Surv.*, 39, 04 2007. doi: 10.1145/1216370.1216373.

[72] Peter Allmark, Jonathan Boote, Eleni Chambers, Amada Clarke, Ann McDonnel, Andrew Thompson, and Angelena Mary Tod. Ethical issues in the use of In-depth interviews: Literature review and discussion. *Research Ethics Review*, 5(2):48–54, 2009.

[73] Charles Pfleeger. *Security in Computing*. Pearson Education, Inc., 5th edition, 2015. ISBN 9780134085043.

[74] Charles Pfleeger and Shari Pfleeger. *Analyzing Computer Security*. Pearson Education, Massachusetts, second edition, 2012. ISBN 0-13-278946-9.

[75] Atik Pilihanto. Information Security Reading Room A Complete Guide on IPv6 Attack and Defense. *SANS Institute Readin Room*, 2011.

[76] David M Piscitello and A Lyman Chapin. *Open Systems Networking: TCP/IP and OSI*. 1993. ISBN 0201563347.

[77] Politie. Helpdeskfraude (Tech Support Scam of Microsoft scam) | politie.nl, 2019. URL `https://www.politie.nl/themas/microsoft-techscam.html`.

[78] Jon Postel. RFC 790. *IETF*, 1981. URL `https://tools.ietf.org/html/rfc790`.

[79] Anbalagan Prasanth. *A Study of Software Security Problem Disclosure, Correction and Patching Processes*. PhD thesis, North Carolina State University, 2011.

[80] Virendra Proag. The Concept of Vulnerability and Resilience. *Procedia Economics and Finance*, 18 (September):369–376, 2014. ISSN 22125671. doi: 10.1016/S2212-5671(14)00952-6.

[81] William R Shadish, Thomas D Cook, and Donald Thomas Campbell. Quasi-experimental designs for generalized causal inference. *Evaluation and Program Planning - EVAL PROGRAM PLANN*, 27, 01 2004.

[82] Rapid 7. Forward DNS (FDNS) | Rapid7 Open Data, 2019. URL `https://opendata.rapid7.com/sonar.fdns_v2/`.

[83] Y. Rekhter, B. Moskowitz, D. Karrenberg, G. J., and E. Lear. Address Allocation for Private Internets. Technical report, feb 1996. URL `https://www.rfc-editor.org/info/rfc1918`.

[84] Philipp Richter, Mark Allman, Randy Bush, Vern Paxson, and U C Berkeley Icsi. A Primer on IPv4 Scarcity. *ACM SIGCOMM Computer Communication*, 45, 2014. doi: 10.1145/2766330.2766335.

[85] Philipp Richter, Florian Wohlfart, Narseo Vallina-rodriguez, Mark Allman, Randy Bush, Anja Feldmann, Christian Kreibich, Nicholas Weaver, and Vern Paxson. A Multi-perspective Analysis of Carrier-Grade NAT Deployment. *Proceedings of ACM IMC*, 2016. doi: 10.1145/2987443.2987474.

[86] RIPE NCC. Regional Internet Registry — RIPE Network Coordination Centre, 2017. URL `https://www.ripe.net/about-us/what-we-do/regional-internet-registry`.

[87] Paul R. Rosenbaum. *Two Simple Models for Observational Studies*, pages 65–94. Springer New York, New York, NY, 2010. ISBN 978-1-4419-1213-8. doi: 10.1007/978-1-4419-1213-8_3. URL `https://doi.org/10.1007/978-1-4419-1213-8_3`.

[88] Margeret Rouse. What is 99.999 (Five nines or Five 9s)?, 2010. URL `https://searchcio.techtarget.com/definition/99999`.

[89] Brahima Sanou. Facts and 2017 figures. *ITU Telecommunication Development*, 2017.

[90] Benjamin Saunders, Julius Sim, Tom Kingstone, Shula Baker, Jackie Waterfield, Bernadette Bartlam, Heather Burroughs, and Clare Jinks. Saturation in qualitative research: exploring its conceptualization and operationalization. *Quality & Quantity*, 52(4):1893–1907, Jul 2018. ISSN 1573-7845. doi: 10.1007/s11135-017-0574-8. URL `https://doi.org/10.1007/s11135-017-0574-8`.

[91] Karen Scarfone and Angela Orebaugh. Technical Guide to Information Security Testing and Assessment Recommendations of the National Institute of Standards and Technology. *Nist Special Publication*, 800: 1–80, 2008. doi: 10.6028/NIST.SP.800-115.

[92] Melissa Schilling. *Strategic Management of Technological Innovation*. 01 2016. ISBN 1259539067.

[93] F. B. Schneider. Least privilege and more [computer security]. *IEEE Security Privacy*, 1(5):55–59, Sep. 2003. ISSN 1540-7993. doi: 10.1109/MSECP.2003.1236236.

[94] Security.nl. Europol wil einde aan Carrier Grade NAT bij internetproviders, 2017. URL https://www.security.nl/posting/535623/Europol+wil+einde+aan+Carrier+Grade+NAT+bij+internetproviders.

[95] Irving Seidman. *Interviewing as Qualitative Research*. Teachers College Press, London, third edition, 2006. ISBN 978-0-8077-4666-0.

[96] R. Sekaran & Bougie. *Research methods for business*. 2013. ISBN 978-1-119-94225-2. doi: 10.1017/CBO9781107415324.004.

[97] Vishal Sharma and Rajesh Kumar. Teredo tunneling-based secure transmission between uavs and ground ad hoc networks. *International Journal of Communication Systems*, 30(7):e3144, 2017. doi: 10.1002/dac.3144. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/dac.3144. e3144 dac.3144.

[98] Keith Shaw. What is DNS and how does it work? *Network World*, 2018. URL https://www.networkworld.com/article/3268449/what-is-dns-and-how-does-it-work.html.

[99] Keith Shaw. The OSI model explained: How to understand (and remember) the 7 layer network model, 2018. URL https://www.networkworld.com/article/3239677/the-osi-model-explained-how-to-understand-and-remember-the-7-layer-network-model.html.

[100] Patricia Shields and Nandhini Rangarajan. *A Playbook for Research Methods: Integrating Conceptual Frameworks and Project Management*. 07 2013. ISBN 10:1-58107-247-3.

[101] Aftab Siddiqui. RFC 8200 - IPv6 has been standardized, 2017. URL https://www.internetsociety.org/blog/2017/07/rfc-8200-ipv6-has-been-standardized/.

[102] SIDN. SIDN : Over SIDN, 2019. URL https://www.sidn.nl/t/over-sidn.

[103] Pyda Srisuresh and Matt Holdrege. IP Network Address Translator (NAT) Terminology and Considerations. 1999. URL https://tools.ietf.org/html/rfc2663.

[104] James Steward. *CISSP official Study Guide*. 7 edition, 2015. ISBN 9781119042716.

[105] Ben Stock, Giancarlo Pellegrino, Frank Li, Michael Backes, and Christian Rossow. Didn't you hear me? - towards more successful web vulnerability notifications. 01 2018. doi: 10.14722/ndss.2018.23183.

[106] Fred Templin. RFC 5214 - Intra-Site Automatic Tunnel Addressing Protocol (ISATAP), 2008. URL https://tools.ietf.org/html/rfc5214.

[107] Urban Dictionary. Script Kiddie, 2019. URL https://www.urbandictionary.com/define.php?term=scriptkiddie.

[108] Ibo van de Poel. *Ethics, Technology, and Engineering: An Introduction*. Wiley-Blackwell, 2011.

[109] Karim Vellani. *Strategic Secutiry Management - A Risk Assessment Guide for Decision Makers*. 2007. ISBN 9780123708977.

[110] VNG. Ruim 50% Nederlandse gemeenten bereikbaar via IPv6, 2019. URL https://www.vngrealisatie.nl/nieuws/ruim-50-nederlandse-gemeenten-bereikbaar-ipv6.

[111] Rossouw von Solms and Johan van Niekerk. From information security to cyber security. *Computers and Security*, 38:97–102, 2013. ISSN 01674048. doi: 10.1016/j.cose.2013.04.004. URL http://dx.doi.org/10.1016/j.cose.2013.04.004.

[112] Johannes Weber. Why NAT has nothing to do with Security! | Blog Webernetz.net, 2013. URL https://blog.webernetz.net/why-nat-has-nothing-to-do-with-security/.

[113] Jane Webster and Richard T Watson. LIT REV - Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly*, 26(2):xiii – xxiii, 2002. ISSN 02767783. doi: 10.1.1.104.6570.

[114] Zack Whittaker. Three years later, Let's Encrypt has issued over 380 million HTTPS certificates, 2018. URL https://techcrunch.com/2018/09/14/three-years-later-lets-encrypt-now-secures-75-of-the-web/?guccounter=1&guce_referrer_us=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce_referrer_cs=LmSTXZU6xRH7cu3Julb7hA.

A

Email

Email I

Beste [redacted]

Ik e-mail je omdat je contact heb gehad met een van mijn collega's. Ik begreep van hem dat je wat vragen hebt rondom ons onderzoek.

Ik zal u kort uitleggen welke stappen wij ondernomen hebben. Voor het gemak zal ik het stapsgewijs toelichten.

1. De hostnames zijn verzameld uit project Sonarr van Rapid 7.
2. Op deze hostnames zijn A & AAAA record lookups gedaan.
3. Met ZMAP zijn scans gedaan op de betreffende IPv4 en IPv6 adressen.
4. Vervolgens zijn deze resultaten naast elkaar gezet en geanalyseerd
5. Met behulp van WHOIS querys is de informatie achter de hostnames achterhaald.

Vanuit daar zijn wij contact gaan zoeken met bedrijven om te achterhalen waar deze verschillen in configuratie vandaan zijn gekomen. Ik hoop dat dit u een heldere uitleg geeft van onze stappen.

Bij deze wil ik u ook graag een overzicht geven van de scanresultaten van de systemen. U komt met 2 IP adressen voor in onze dataset. Hieronder vindt u de IP adressen, hostnames en scan resultaten.

[redacted]
tcp_ssh_22 : IPv4 = 0, IPv6 = 1
tcp_dns_53 : IPv4 = 1, IPv6 = 1
tcp_ftp_21 : IPv4 = 1, IPv6 = 1
tcp_http_80 : IPv4 = 1, IPv6 = 1
tcp_https_443 : IPv4 = 1, IPv6 = 1
tcp_https_8443 : IPv4 = 1, IPv6 = 1
tcp_smtp_25 : IPv4 = 1, IPv6 = 1
udp_dns_53 : IPv4 = 1, IPv6 = 1

[redacted]
tcp_ssh_22 : IPv4 = 1, IPv6 = 0
tcp_dns_53 : IPv4 = 0, IPv6 = 1
tcp_ftp_21 : IPv4 = 1, IPv6 = 1
tcp_http_80 : IPv4 = 1, IPv6 = 1
tcp_https_443 : IPv4 = 1, IPv6 = 1

tcp_smtp_25 : IPv4 = 0, IPv6 = 1
udp_dns_53 : IPv4 = 0, IPv6 = 1

Wij proberen nu te achterhalen waarom er verschillen tussen de 2 zijn. Niet alleen vanuit technisch vlak maar ook vanuit management perspectief, daarnaast proberen we te kijken naar potentiële oplossingen en wat voor controls dit soort problemen zou kunnen oplossen. Onze vraag aan u is of u of een van uw collega's bereid is voor een kort telefonisch gesprek zodat wij een interview kunnen afnemen om hierover data te verzamelen. Het uiteindelijke resultaat wordt gepubliceerd, maar alle data zal confidentieel zijn. De gebruikte bronnen zijn alleen bekent bij de onderzoekers en het zal onmogelijk zijn om vanuit de publicatie te achterhalen welke bronnen zijn gebruikt, namen en bedrijven zullen nooit bekend worden gemaakt.

Ik hoop dat ik u bij deze een duidelijke uitleg heb gegeven over ons onderzoek en mocht u verder nog vragen hebben hoor ik het graag. Ik hoop dat ik op uw medewerking kan regekenen.

Met vriendelijke groet,

Dennis Rieffe

TU Delft Faculty of Technology, Policy and Management
Jaffalaan 5
2628 BX Delft

Translation Email I:

Dear [redacted]

I am sending you this email because you have been in contact with one of my colleagues. He told me that you had a few questions about our research.

I would like to briefly address the steps that we have taken so far. For clarity's sake, I maintained the order in which we took them.

1. The hostnames were taken from the project Sonarr by Rapid 7.
2. A & AAAA lookups were performed on these hostnames.
3. ZMAP was used to scan the corresponding IPv4 and IPv6 addresses.
4. Then, the results were compared and analysed
5. The information corresponding to the hostnames was retrieved using WHOIS queries.

From this point, we sought contact with companies to understand where the differences in configuration stem from. I hope this provides you with a comprehensive explanation of the steps we have taken.

I would like to take this opportunity to also provide you with our scan results. In our data set, we found that you have two distinct IP addresses. Bellow, you can find the IP addresses, hostnames and scan results.

[redacted]

tcp_ssh_22 : IPv4 = 0, IPv6 = 1
tcp_dns_53 : IPv4 = 1, IPv6 = 1
tcp_ftp_21 : IPv4 = 1, IPv6 = 1
tcp_http_80 : IPv4 = 1, IPv6 = 1
tcp_https_443 : IPv4 = 1, IPv6 = 1

tcp_https_8443 : IPv4 = 1, IPv6 = 1
tcp_smtp_25 : IPv4 = 1, IPv6 = 1
udp_dns_53 : IPv4 = 1, IPv6 = 1

[redacted]
tcp_ssh_22 : IPv4 = 1, IPv6 = 0
tcp_dns_53 : IPv4 = 0, IPv6 = 1
tcp_ftp_21 : IPv4 = 1, IPv6 = 1
tcp_http_80 : IPv4 = 1, IPv6 = 1
tcp_https_443 : IPv4 = 1, IPv6 = 1
tcp_smtp_25 : IPv4 = 0, IPv6 = 1
udp_dns_53 : IPv4 = 0, IPv6 = 1

We are attempting to understand why there are differences between these two. Not only are we interested in the technical, but also in the management aspect. Moreover, we try to find potential solutions and control measures could solve this kind of issues. Our question to you is whether you or one of your colleagues would be willing to participate in a telephone interview with us so that we can collect data. The final result of this research shall be published, but all data shall be confidential. Any source that we use will only be known to the researchers and shall not be made public in the publication. Names or Companies shall never be identifiable.

I hope i have provided you with a satisfactory explanation about our research. I shall be happy to answer any question you may have. I hope for your cooperation.

Thank you for your time,

Dennis Rieffe

TU Delft Faculty of Technology, Policy and Management
Jaffalaan 5
2628 BX Delft

Email II

Geachte Meneer/Mevrouw,

Ik heb zojuist telefonisch contact gehad met uw bedrijf en heb via hun dit e-mail adres gekregen.

Ik ben bezig met een onderzoek vanuit de TU Delft naar de kennis en het gebruik van IPv6 en heb daarbij poorten gescanned van Dual-Stack hosts. Hierbij kwam een webserver voorbij die bij u in beheer is. Deze webserver draait op de IP adressen [redacted]. Mijn onderzoek focust zich op de verschillen tussen in openstaande poorten tussen IPv4 en IPv6 en ik heb gezien dat bij uw webserver verschillen zitten tussen de twee.

Zou ik hier met u een keer contact over mogen hebben wat hierachter de beweegredenen zijn.

Alvast hartelijk bedankt en ik hoop graag iets van u te horen. Mocht u meer informatie willen schroom dat niet om dit te vragen.

Met vriendelijke groet,

Dennis Rieffe

TU Delft Faculty of Technology, Policy and Management
Jaffalaan 5
2628 BX Delft

Translation Email II:

Dear Sir / Madam,

I have just called your company and they have provided me with your email address.

I am currently working on a research with TU Delft about the use of IPv6. In that context, we have scanned ports of Dual-Stack hosts. In that process, a webserver of which you are the administrator came to our attention. The webserver is accessible on IP addresses [redacted]. My research focusses on the differences in open port policy between IPv4 and IPv6, and I have noticed that this is the case in your webserver.

Would you be open to me contacting you about your motivations?

Thank you very much in advance,
I shall be happy to answer any question you may have,
Looking forward to your reply,

Dennis Rieffe

TU Delft Faculty of Technology, Policy and Management
Jaffalaan 5
2628 BX Delft

Email III

Beste Meneer/Mevrouw,

In het kader van een onderzoek naar IPv6 configuratie op Dual-Stack hosts van de Technische Universiteit Delft wil ik u graag wat vragen stellen. Wij onderzoeken de verschillen in openstaande poorten tussen IPv4 en IPv6, en uw server komt hierbij naar boven. Wij zijn aan uw contact gegevens gekomen door de hostname [redacted] op te zoeken op sidn.nl.

Aan de hand van publieke hostnames hebben wij scans uitgevoerd op openstaande poorten. Hierbij is een server van u naar boven gekomen, [redacted]. Wanneer er een DNS lookup wordt gedaan op IPv4 resolved het naar een intern adres terwijl het IPv6 adres een publiek adres is. Klopt het dat deze server bij u in beheer is? Uit een Cyber Security perspectief is dit een ongebruikelijke situatie die gevaren met zich mee kan brengen vandaar dat ik u hier graag over op de hoogte van wil stellen.

Voor ons onderzoek willen wij u graag vragen naar de mogelijke achterliggende beweegredenen. Zou ik u hiervoor daarom mogen benaderen? Ik wil u graag in ieder geval graag hartelijk bedanken voor uw tijd en ik hoop van u te horen. Mocht u meer informatie willen over ons onderzoek schroom dan niet om dit te vragen.

Met vriendelijke groet,

Dennis Rieffe

TU Delft Faculty of Technology, Policy and Management
Jaffalaan 5
2628 BX Delft

Translation Email III:

Dear Sir / Madam,

As part of a study into IPv6 configuration on Dual-Stack hosts at Delft University of Technology, I would like to ask you some questions. We investigate the differences in open ports between IPv4 and IPv6, and your server came up. We got your contact details by looking up the hostname [redacted] on sidn.nl.

Based on public hostnames, we performed scans on open ports. A server of yours has emerged, [redacted]. When a DNS lookup is done on IPv4, it resolved to an internal address while the IPv6 address is a public address. Is it true that this server is under your control? From a cyber security perspective, this is an unusual situation that can cause dangers, so I would like to inform you about this.

For our research we would like to ask you about the possible underlying reasons. May I contact you for this? In any case, I would like to thank you for your time and I hope to hear from you. If you want more information about our research, don't hesitate to ask.

Sincerely,

Dennis Rieffe

TU Delft Faculty of Technology, Policy and Management
Jaffalaan 5
2628 BX Delft

B

Interview Questions

Interview I

General information about the respondent

- What is the size of is your organization in employees (or monetary)?
- When has your organization been founded?
- How many years of experience do you have in your current field?
- What is your position within the company?
- What kind of IT systems do you mostly operate (endpoints, storage, Databases, Web, Network)?

IPv6 and port configuration

- Does your organization support IPv6?
- Who is responsible for the opening and closing of the external ports? Do you agree with this?
- Who determines which ports should be open or closed? Do you agree with this?
- Who is responsible for verifying which ports are open on internet facing machines? Do you agree with this?
- Are processes similar for IPv6 and IPv4?

Internal/External setup (only if applicable to respondent)

- What is the reason that there was an internal IPv4 external IPv6 configuration?
- Is this an intentional configuration?
- In your professional opinion, is this a neat configuration?
- In your opinion is this a save configuration?

Security configuration and incidents

- Have you encountered security incidents based on IP address misconfigurations? (both IPv4 and IPv6)?
- How and why do you think this misconfiguration occurred?
- On average, how many security misconfiguration lead to an incident and how many were mitigated in time? (quantify if possible)
- How did misconfigurations affect you and the way your company approaches and handles security?
- What do you think could be a potential solution to mitigate these problems?

Closing question

- Do you have any other remarks you wish to share with me?

Interview II

General information about the respondent

- How many years of experience do you have in your current field?
- What is your position within the company?
- What kind of IT systems do you mostly operate (endpoints, storage, Databases, Web, Network)?

Follow-up questions

- Does your organization support IPv6?
- What do you think of NAT (local and Carrier-Grade NAT)
- What is the reason why there are differences between the port configurations?
- Do you see differences in security posture between size, technology driven in companies and why?
- Do you see difference in IPv6 maturity per country?
- Do you see why management has a reason to switch too IPv6?
- Who is responsible for verifying which ports are open on internet facing machines? Do you agree with this
- What do you think could be a potential solution to mitigate these problems? Also from a managerial perspective
- Who is responsible for addressing this problem (is it a problem)?
- What could break the chain of IPv6?
- What is your opinion on how we are handling the transfer to IPv6, with dual-stack and tunneling protocols?
- Do you notice that IPv4 is picked up and moved to IPv6?
- Education around IPv6?
- Devil's advocate, why not use IPv6?
- Why would you use IPv6?

Security configuration and incidents

- Have you encountered security incidents based on IP address misconfigurations? (both IPv4 and IPv6)
- How and why do you think this misconfiguration occurred?

Closing question

- Do you have any other remarks you wish to share with me?

C

Dataset TLD extract

Top Level Domain	Country	IPv6 Adoption Rate (Google or Akami when starred)	# Hostnames	Correct Configuration	Incorrect Configuration	
Total			22514910	21432137	1082773	4.8%
Alexa			93497	91236	2261	2.4%
.arpa			73	73	0	0.0%
.com			7311192	7010569	300623	4.1%
.int			37	37	0	0.0%
.edu			69939	68710	1229	1.8%
.eu			361155	338905	22250	6.2%
.gov			8194	8163	31	0.4%
.mil			10781	10779	2	0.0%
.net			1512595	1448107	64488	4.3%
.org			676716	642992	33724	5.0%
.ac.uk			8194	8094	100	1.2%
.co.uk			436747	427093	9654	2.2%
.gov.uk			114	108	6	5.3%
.ac	Ascension Island	n/a	416	367	49	11.8%
.ad	Andorra	0*	93	83	10	10.8%
.ae	United Arab Emirates	13.71	2163	2032	131	6.1%
.af	Afghanistan	0.04	92	86	6	6.5%
.ag	Antigua and Barbuda	0	1043	937	106	10.2%
.ai	Anguilla	0	1365	1310	55	4.0%
.al	Albania	0	1067	1041	26	2.4%
.am	Armenia	1.64	2347	2291	56	2.4%
.ao	Angola	0	32	29	3	9.4%
.aq	Antartica	0*	4	3	1	25.0%
.ar	Argentina	7.71	9575	9001	574	6.0%
.as	American Samoa	0	339	325	14	4.1%
.at	Austria	10.47	102002	93091	8911	8.7%
.au	Australia	16.92	87591	65449	22142	25.3%
.aw	Aruba	0	12	12	0	0.0%
.ax	Åland	24.95	103	101	2	1.9%
.az	Azerbaijan	0	1864	1858	6	0.3%
.ba	Bosnia and Herzegovina	4.83	1161	1138	23	2.0%

.bb	Barbados	0.01	7	7	0	0.0%
.bd	Bangladesh	0.04	1040	1034	6	0.6%
.be	Belgium	52.93	286966	273162	13804	4.8%
.bf	Burkina Faso	0	34	34	0	0.0%
.bg	Bulgaria	1.45	1730	1327	403	23.3%
.bh	Bahrain	0*	11	11	0	0.0%
.bi	Burundi	0	62	48	14	22.6%
.bj	Benin	0.01	38	36	2	5.3%
.bm	Bermuda	0*	14	14	0	0.0%
.bn	Brunei	0	6	6	0	0.0%
.bo	Bolivia	13.45	160	126	34	21.3%
.br	Brazil	27.56	101645	95012	6633	6.5%
.bs	Bahamas	0	12	8	4	33.3%
.bt	Bhutan	4.32	384	340	44	11.5%
.bw	Botswana	0	85	84	1	1.2%
.by	Belarus	0.04	11227	10898	329	2.9%
.bz	Belize	3.29	1144	1094	50	4.4%
.ca	Canada	23.37	57789	54294	3495	6.0%
.cc	Cocos Islands	0*	10866	10381	485	4.5%
.cd	Democratic Republic of the Congo	0.02	74	68	6	8.1%
.cf	Central African Republic	0	1970	1828	142	7.2%
.cg	Republic of the Congo	0.61	45	31	14	31.1%
.ch	Switzerland	29.31	171924	164514	7410	4.3%
.ci	Ivory Coast	0.04	206	192	14	6.8%
.ck	Cook Islands	0*	3	3	0	0.0%
.cl	Chile	0.13	9047	6594	2453	27.1%
.cm	Cameroon	0.07	403	377	26	6.5%
.cn	People's Republic of China	5.11	12471	11991	480	3.8%
.co	Colombia	1.1	37292	36369	923	2.5%
.cr	Costa Rica	0.11	515	495	20	3.9%
.cu	Cuba	0	49	49	0	0.0%
.cv	Cape Verde	0	34	34	0	0.0%
.cw	Curaçao	0.04	24	24	0	0.0%
.cx	Christmas Islans	0*	818	781	37	4.5%
.cy	Cyprus	0.1	283	282	1	0.4%

.cz	Czech Republic	12.03	407290	384879	22411	5.5%
.de	Germany	42.69	3834730	3540598	294132	7.7%
.dj	Djibouti	0	126	118	8	6.3%
.dk	Denmark	3.55	192851	192039	812	0.4%
.dm	Dominica	0	164	164	0	0.0%
.do	Dominican Republic	1.22	333	311	22	6.6%
.dz	Algeria	0	71	68	3	4.2%
.ec	Ecuador	18.94	628	489	139	22.1%
.ee	Estonia	26.4	54737	54591	146	0.3%
.eg	Egypt	1.58	38	38	0	0.0%
.es	Spain	2.51	125519	122796	2723	2.2%
.et	Ethiopia	0	9	9	0	0.0%
.fi	Finland	26.39	37724	30538	7186	19.0%
.fj	Fiji	0	37	36	1	2.7%
.fm	Federated States of Micronesia	0*	978	849	129	13.2%
.fo	Faroe Islands	7.82	158	155	3	1.9%
.fr	France	32.08	462141	436191	25950	5.6%
.ga	Gabon	13.39	2590	2436	154	5.9%
.gd	Grenada	0.02	129	113	16	12.4%
.ge	Georgia	0	681	668	13	1.9%
.gf	French Guiana	38.04	4	4	0	0.0%
.gg	Guernsey	0*	6821	6798	23	0.3%
.gh	Ghana	0	113	112	1	0.9%
.gi	Gibraltar	0.2*	12	12	0	0.0%
.gl	Greenland	0	561	547	14	2.5%
.gm	The Gambia	0	82	77	5	6.1%
.gp	Guadeloupe	24.88	60	58	2	3.3%
.gq	Equatorial Guinea	0	895	791	104	11.6%
.gr	Greece	36.95	46624	46004	620	1.3%
.gs	South Georgia and the South Sandwich Islands	n/a	264	247	17	6.4%
.gt	Guatamala	10.16	210	209	1	0.5%
.gw	Guinea-Bissau	0	15	15	0	0.0%
.gy	Guyana	0	131	124	7	5.3%
.hk	Hong Kong	0.25	2116	2064	52	2.5%

.hm	Heard Islands and McDonald Islands	n/a	16	16	0	0.0%
.hn	Honduras	0.01	102	91	11	10.8%
.hr	Croatia	0.04	2473	2447	26	1.1%
.ht	Haiti	0	100	95	5	5.0%
.hu	Hungary	21.5	70232	59988	10244	14.6%
.id	Indonesia	0.3	5610	5043	567	10.1%
.ie	Ireland	18.95	12941	12814	127	1.0%
.il	Israel	4.44	1623	1555	68	4.2%
.im	Isle of Man	0.02	1328	1271	57	4.3%
.in	India	36.56	30468	29689	779	2.6%
.io	British Indian Ocean Territory	0*	101948	98679	3269	3.2%
.iq	Iraq	0.01	42	42	0	0.0%
.ir	Iran	2.85	2813	2761	52	1.8%
.is	Iceland	2.2	1896	1801	95	5.0%
.it	Italy	3.48	80094	77563	2531	3.2%
.je	Jersey	0.24	215	193	22	10.2%
.jm	Jamaica	0	26	26	0	0.0%
.jo	Jordan	0.4	50	50	0	0.0%
.jp	Japan	31.39	52612	50158	2454	4.7%
.ke	Kenya	4.59	1057	1038	19	1.8%
.kg	Kyrgyzstan	0	423	412	11	2.6%
.kh	Cambodia	0.05	77	60	17	22.1%
.ki	Kiribati	0*	22	22	0	0.0%
.kn	Saint Kitts and Nevis	0.02	6	6	0	0.0%
.kr	South Korea	5.2	962	915	47	4.9%
.kw	Kuwait	0	50	50	0	0.0%
.ky	Cayman Islands	0	50	44	6	12.0%
.kz	Kazakhstan	0.01	14553	14438	115	0.8%
.la	Laos	0.03	574	559	15	2.6%
.lb	Lebanon	0.35	38	36	2	5.3%
.lc	Saint Lucia	0	107	102	5	4.7%
.li	Liechtenstein	0.5*	4363	4166	197	4.5%
.lk	Sri Lanka	18.54	634	615	19	3.0%
.ls	Lesotho	0	10	10	0	0.0%
.lt	Lithuania	0.12	4277	3173	1104	25.8%

.lu	Luxemburg	29.35	4224	3885	339	8.0%
.lv	Latvia	1.88	2007	1931	76	3.8%
.ly	Libya	0	767	759	8	1.0%
.ma	Morocco	0	1550	1402	148	9.5%
.mc	Monaco	0*	64	57	7	10.9%
.md	Moldova	2.41	1117	1045	72	6.4%
.me	Montenegro	0	58618	56727	1891	3.2%
.mg	Madegascar	0	322	299	23	7.1%
.mh	Marshall Islands	0*	8	8	0	0.0%
.mk	North Macedonia	0.02	453	437	16	3.5%
.ml	Mali	0	3739	3478	261	7.0%
.mm	Myanmar	5.05	10	10	0	0.0%
.mn	Mongolia	0	203	183	20	9.9%
.mo	Macau	10.2*	44	44	0	0.0%
.mp	Northern Mariana Islands	0*	25	25	0	0.0%
.mq	Martinique	26.04	18	16	2	11.1%
.mr	Mauritania	0	17	17	0	0.0%
.ms	Montserrat	0*	278	251	27	9.7%
.mt	Malta	0.02	180	170	10	5.6%
.mu	Mauritius	0.01	267	250	17	6.4%
.mv	Maladives	2.2*	66	66	0	0.0%
.mw	Malawi	0.02	22	16	6	27.3%
.mx	Mexico	27.75	18350	18059	291	1.6%
.my	Malaysia	38.52	2344	2315	29	1.2%
.mz	Mozambique	0.1	40	38	2	5.0%
.na	Namibia	0	69	61	8	11.6%
.nc	New Caledonia	0.21	761	698	63	8.3%
.ne	Niger	0.05	6	6	0	0.0%
.nf	Norflok Islands	0*	38	38	0	0.0%
.ng	Nigeria	0	4671	4563	108	2.3%
.ni	Nicaragua	0.33	44	40	4	9.1%
.nl	Netherlands	17.08	1600891	1542821	58070	3.6%
.no	Norway	13	76428	75638	790	1.0%
.np	Nepal	0.05	580	554	26	4.5%
.nr	Nauru	0*	22	22	0	0.0%

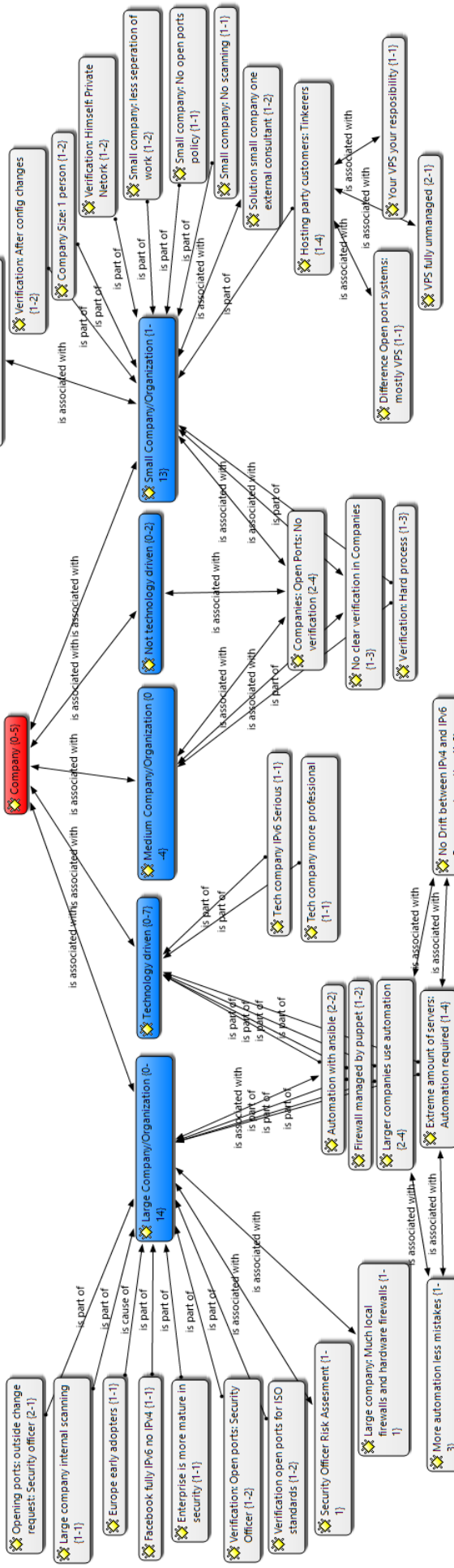
.nu	Niue	0*	40475	39716	759	1.9%
.nz	New Zeeland	20.27	21509	19532	1977	9.2%
.om	Oman	4.77	21	19	2	9.5%
.pa	Panama	0.11	50	50	0	0.0%
.pe	Peru	17.02	937	829	108	11.5%
.pf	French Polynesia	0*	9	8	1	11.1%
.pg	Papua New Guniea	0	7	6	1	14.3%
.ph	Philippines	0.49	979	865	114	11.6%
.pk	Pakistan	0.04	1397	1375	22	1.6%
.pl	Poland	10.95	138078	132572	5506	4.0%
.pm	Saint-Pierre and Miquelon	0*	1020	1001	19	1.9%
.pn	Pitacairn Islands	n/a	13	12	1	7.7%
.pr	Puerto Rico	4.01	31	27	4	12.9%
.ps	Palestine	0*	100	97	3	3.0%
.pt	Portugal	25.4	7811	7449	362	4.6%
.pw	Palau	0*	6618	6483	135	2.0%
.py	Paraguay	0.21	246	213	33	13.4%
.qa	Qatar	0.06	140	138	2	1.4%
.re	Réunion	24.67	2429	2250	179	7.4%
.ro	Romania	14.63	55476	54927	549	1.0%
.rs	Serbia	0.04	5726	5666	60	1.0%
.ru	Russia	3.48	624223	610011	14212	2.3%
.rw	Rwanda	0.08	67	58	9	13.4%
.sa	Saudi Arabia	15.05	436	362	74	17.0%
.sb	Solomon Islands	0	41	41	0	0.0%
.sc	Seychelles	0*	181	165	16	8.8%
.sd	Sudan	0.29	55	44	11	20.0%
.se	Sweden	5.65	313386	311823	1563	0.5%
.sg	Singapore	9.2*	1483	1431	52	3.5%
.sh	Saint Helena	0*	1668	1590	78	4.7%
.si	Slovania	10.41	9749	9622	127	1.3%
.sk	Slovakia	4.41	42111	38737	3374	8.0%
.sl	Sierra Leone	0	9	9	0	0.0%
.sm	San Marino	0*	26	19	7	26.9%
.sn	Senegal	0	237	168	69	29.1%

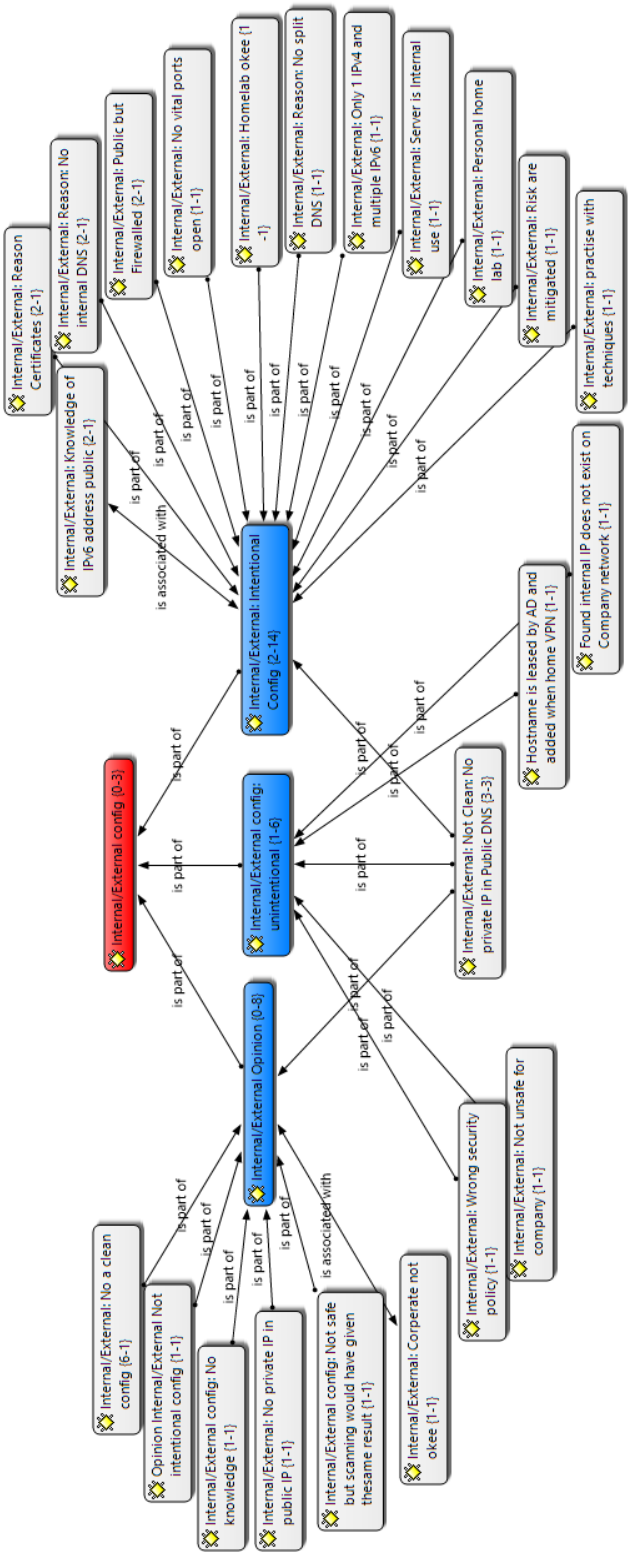
.so	Somalia	0	355	341	14	3.9%
.sr	Suriname	0.02	80	71	9	11.3%
.st	Sãn Tomé and Príncipe	0*	1198	1153	45	3.8%
.su	Soviet Union	n/a	12032	11748	284	2.4%
.sv	El Salvador	0.01	53	49	4	7.5%
.sx	Saint Martin	n/a	206	195	11	5.3%
.sy	Syria	0.01	16	15	1	6.3%
.sz	Eswatini	0	2	2	0	0.0%
.tc	Turks and Caicos Islands	0	135	120	15	11.1%
.td	Chad	0.01	13	11	2	15.4%
.tf	French Southern and Antarctic Lands	0*	423	400	23	5.4%
.tg	Togo	3.63	43	41	2	4.7%
.th	Thailand	23.12	3737	3431	306	8.2%
.tj	Tajikstan	0.04	496	488	8	1.6%
.tk	Tokelau	0*	7744	7336	408	5.3%
.tl	East Timor	0.07	78	74	4	5.1%
.tm	Turkmenistan	1.48	111	88	23	20.7%
.tn	Tunesia	0.01	1499	1175	324	21.6%
.to	Tonga	0.1*	4490	4328	162	3.6%
.tr	Turkey	0.01	3260	3209	51	1.6%
.tt	Trinidad and Tobago	20.55	77	74	3	3.9%
.tv	Tuvalu	0*	16001	15203	798	5.0%
.tw	Taiwan	32.07	33700	32937	763	2.3%
.tz	Tanzania	0.11	724	701	23	3.2%
.ua	Ukraine	1.99	322848	320682	2166	0.7%
.ug	Uganda	0.27	514	476	38	7.4%
.uk	United Kingdom	23.26	496169	485057	11112	2.2%
.us	United States of America	38.13	46705	44952	1753	3.8%
.uy	Uruguay	36.04	1732	1625	107	6.2%
.uz	Uzbekistan	0.1	643	625	18	2.8%
.va	Vatican City	0*	18	18	0	0.0%
.vc	Saint Vincent and the Grenadines	0	365	353	12	3.3%
.ve	Venezuela	0	865	846	19	2.2%
.vg	British Virgin Islands	0.75	107	92	15	14.0%
.vi	United States Virgin Islands	0.09	18	18	0	0.0%

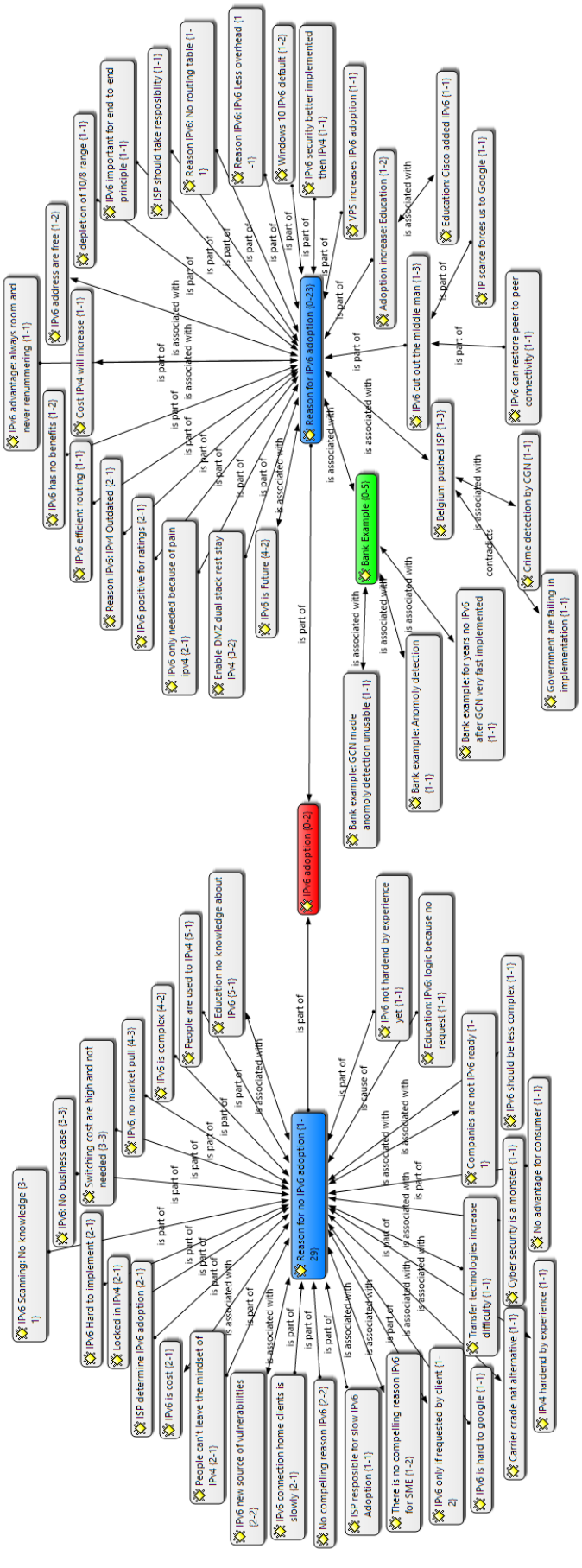
.vn	Vietnam	38.54	807	741	66	8.2%
.vu	Vanuatu	0	144	133	11	7.6%
.wf	Wallis and Futuna	0*	1424	1420	4	0.3%
.ws	Samoa	0*	3399	3297	102	3.0%
.ye	Yemen	0.02	20	15	5	25.0%
.yt	Mayotte	42.88	370	350	20	5.4%
.za	South Africa	0.5	8499	7895	604	7.1%
.zm	Zambia	0.71	67	64	3	4.5%
.zw	Zimbabwe	6.11	124	119	5	4.0%

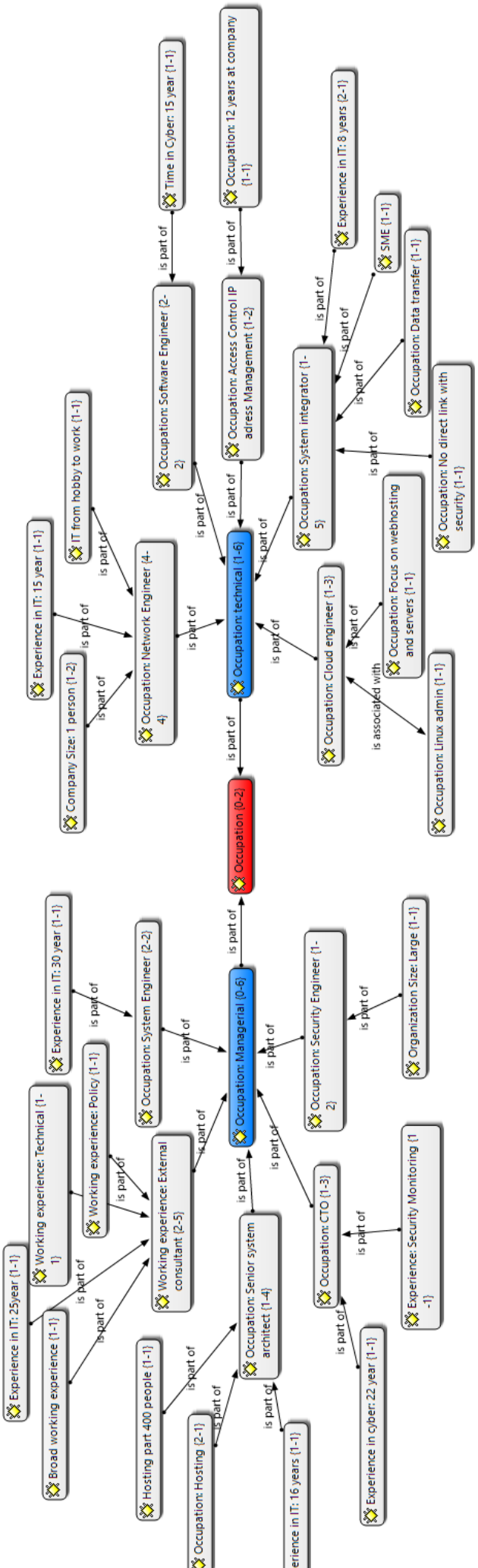
D

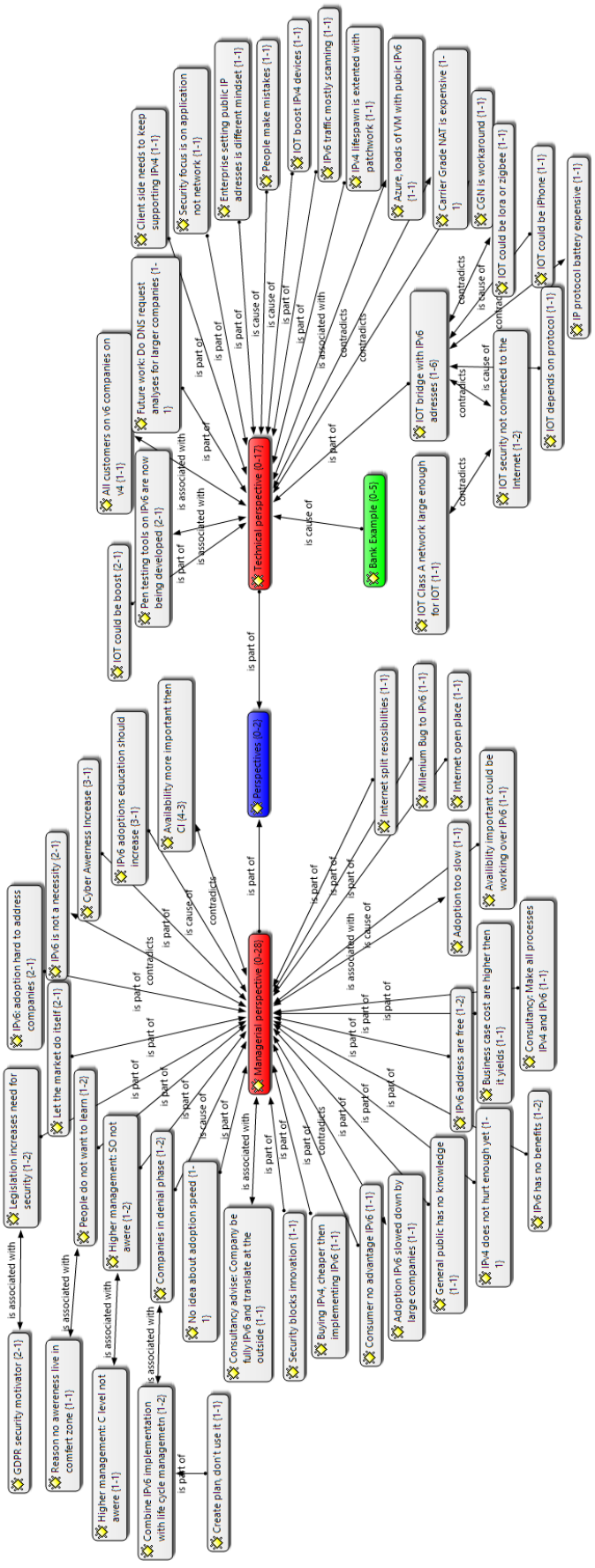
ATLAS.ti Results

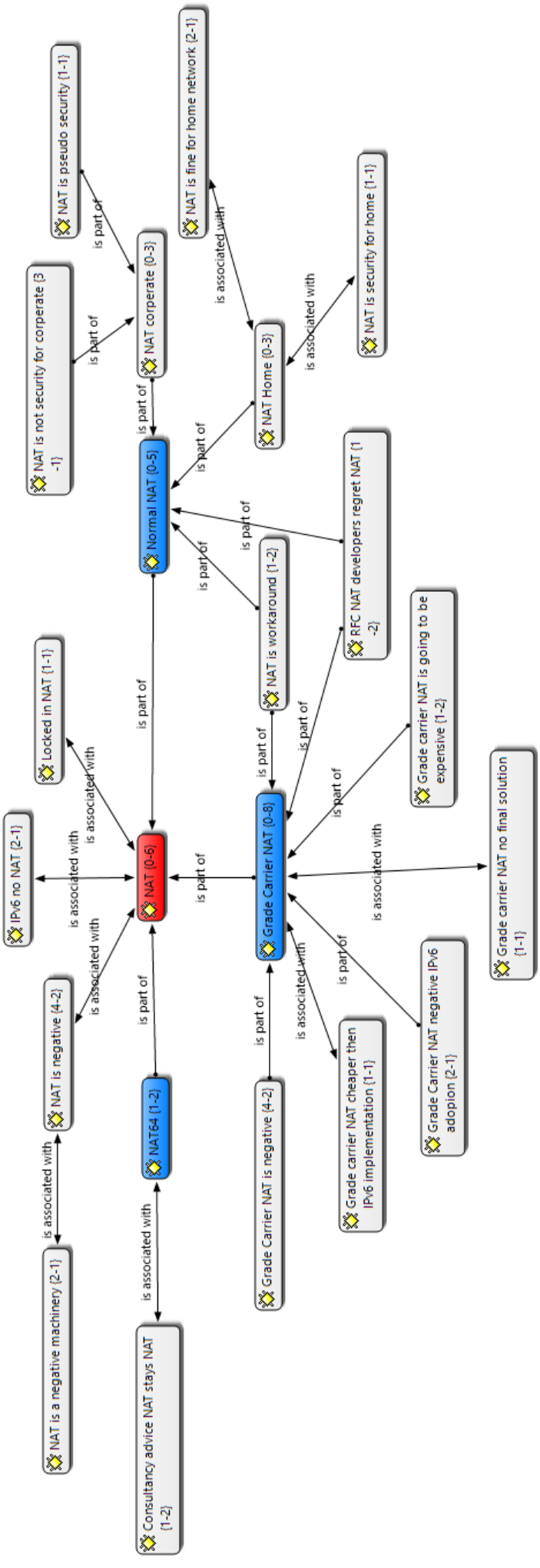

Company (0-5)
Small Company/Organization (1-13)
Medium Company/Organization (0-4)
Large Company/Organization (0-14)
Technology driven (0-7)
Not technology driven (0-2)
Smaller companies less secure (2-1)
Verification: After config changes (1-2)
Company Size: 1 person (1-2)
Verification: Himself Private Netork (1-2)
Small company: less seperation of work (1-2)
Small company: No open ports policy (1-1)
Small company: No scanning (1-1)
Solution small company one external consultant (1-2)
Hosting party customers: Tinkerers (1-4)
Your VPS your resposibility (1-1)
VPS fully unmanaged (2-1)
Difference Open port systems: mostly VPS (1-1)
Companies: Open Ports: No verification (2-4)
No clear verification in Companies (1-3)
Verification: Hard process (1-3)
Tech company IPv6 Serious (1-1)
Tech company more professional (1-1)
No Drift between IPv4 and IPv6 Because automation (1-2)
Automation with ansible (2-2)
Firewall managed by puppet (1-2)
Larger companies use automation (2-4)
Extreme amount of servers: Automation required (1-4)
Opening ports: outside change request: Security officer (2-1)
Large company internal scanning (1-1)
Europe early adopters (1-1)
Facebook fully IPv6 no IPv4 (1-1)
Enterprise is more mature in security (1-1)
Verification: Open ports: Security Officer (1-2)
Verification open ports for ISO standards (1-2)
Security Officer Risk Assessment (1-1)
Large company: Much local firewalls and hardware firewalls (1-1)
More automation less mistakes (1-3)
is associated with
is part of
is cause of

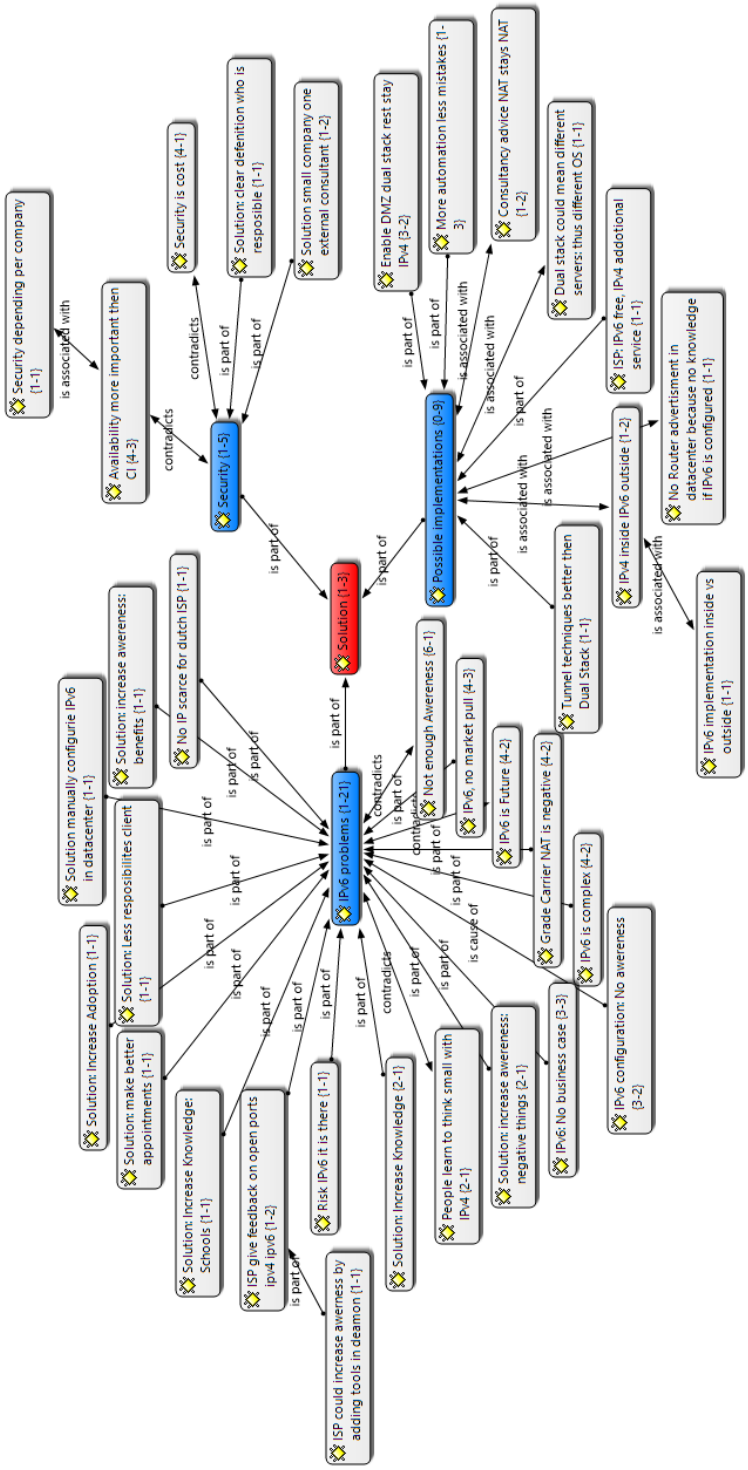

Internal/External: Reason Certificates {2-1}
Internal/External: Reason: No internal DNS {2-1}
Internal/External: Public but Firewalled {2-1}
Internal/External: No vital ports open {1-1}
Internal/External: Homelab okee {1-1}
Internal/External: Reason: No split DNS {1-1}
Internal/External: Only 1 IPv4 and multiple IPv6 {1-1}
Internal/External: Server is Internal use {1-1}
Internal/External: Personal home lab {1-1}
Internal/External: Risk are mitigated {1-1}
Internal/External: practise with techniques {1-1}
Internal/External: Knowledge of IPv6 address public {2-1}
Internal/External: Intentional Config {2-14}
Found internal IP does not exist on Company network {1-1}
Hostname is leased by AD and added when home VPN {1-1}
Internal/External config {0-3}
Internal/External config: unintentional {1-6}
Internal/External: Not Clean: No private IP in Public DNS {3-3}
Internal/External: No a clean config {6-1}
Opinion Internal/External Not intentional config {1-1}
Internal/External Opinion {0-8}
Internal/External: Wrong security policy {1-1}
Internal/External: Not unsafe for company {1-1}
Internal/External config: No knowledge {1-1}
Internal/External: No private IP in public IP {1-1}
Internal/External config: Not safe but scanning would have given thesame result {1-1}
Internal/External: Corperate not okee {1-1}
is part of
is associated with

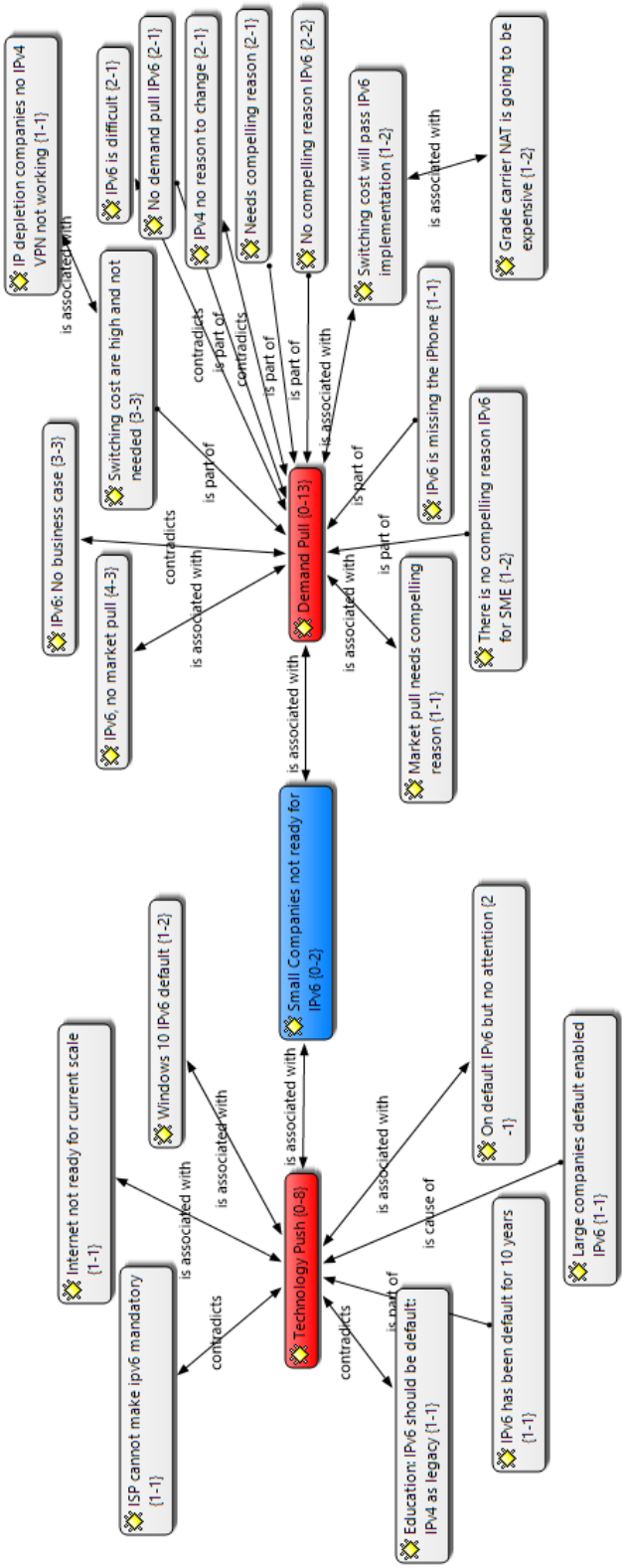

IPv6 adoption (0-2)
Reason for no IPv6 adoption (1-29)
Reason for IPv6 adoption (0-23)
Bank Example (0-5)
IPv6 Scanning: No knowledge (3-1)
IPv6: No business case (3-3)
IPv6 Hard to implement (2-1)
Locked in IPv4 (2-1)
ISP determine IPv6 adoption (2-1)
IPv6 is cost (2-1)
People can't leave the mindset of IPv4 (2-1)
IPv6 new source of vulnerabilities (2-2)
IPv6 connection home clients is slowly (2-1)
No compelling reason IPv6 (2-2)
ISP responsible for slow IPv6 Adoption (1-1)
There is no compelling reason IPv6 for SME (1-2)
IPv6 only if requested by client (1-2)
IPv6 is hard to google (1-1)
Carrier grade nat alternative (1-1)
IPv4 hardend by experience (1-1)
Switching cost are high and not needed (3-3)
IPv6, no market pull (4-3)
IPv6 is complex (4-2)
People are used to IPv4 (5-1)
Education no knowledge about IPv6 (5-1)
IPv6 not hardend by experience yet (1-1)
Education: IPv6 logic because no request (1-1)
Companies are not IPv6 ready (1-1)
Transfer technologies increase difficulty (1-1)
Cyber security is a monster (1-1)
IPv6 should be less complex (1-1)
No advantage for consumer (1-1)
IPv6 advantage: always room and never renummering (1-1)
IPv6 has no benefits (1-2)
IPv6 efficient routing (1-1)
Reason IPv6: IPv4 Outdated (2-1)
IPv6 positive for ratings (2-1)
IPv6 only needed because of pain ipv4 (2-1)
Enable DMZ dual stack rest stay IPv4 (3-2)
IPv6 is Future (4-2)
IPv6 address are free (1-2)
Cost IPv4 will increase (1-1)
depletion of 10/8 range (1-1)
IPv6 important for end-to-end principle (1-1)
ISP should take resposibility (1-1)
Reason IPv6: No routing table (1-1)
Reason IPv6: IPv6 Less overhead (1-1)
Windows 10 IPv6 default (1-2)
IPv6 security better implemented then IPv4 (1-1)
VPS increases IPv6 adoption (1-1)
Adoption increase: Education (1-2)
Education: Cisco added IPv6 (1-1)
IP scarce forces us to Google (1-1)
IPv6 cut out the middle man (1-3)
IPv6 can restore peer to peer connectivity (1-1)
Belgium pushed ISP (1-3)
Crime detection by CGN (1-1)
Government are failing in implementation (1-1)
Bank example: GCN made anomoly detection unusable (1-1)
Bank example: Anomoly detection (1-1)
Bank example: for years no IPv6 after GCN very fast implemented (1-1)
is part of
is associated with
is cause of
contradicts


Occupation [0-2]
Occupation: technical [1-6]
Occupation: Managerial [0-6]
Occupation: Network Engineer [4-4]
Occupation: Software Engineer [2-2]
Occupation: Access Control IP adress Management [1-2]
Occupation: System integrator [1-5]
Occupation: Cloud engineer [1-3]
Occupation: Linux admin [1-1]
Occupation: Focus on webhosting and servers [1-1]
Occupation: No direct link with security [1-1]
Occupation: System Engineer [2-2]
Occupation: Security Engineer [1-2]
Occupation: Senior system architect [1-4]
Occupation: CTO [1-3]
Occupation: Hosting [2-1]
Experience in IT: 15 year [1-1]
Company Size: 1 person [1-2]
IT from hobby to work [1-1]
Time in Cyber: 15 year [1-1]
Occupation: 12 years at company [1-1]
Experience in IT: 8 years [2-1]
SME [1-1]
Occupation: Data transfer [1-1]
Experience in IT: 30 year [1-1]
Working experience: Technical [1-1]
Working experience: Policy [1-1]
Working experience: External consultant [2-5]
Experience in IT: 25year [1-1]
Broad working experience [1-1]
Hosting part 400 people [1-1]
Organization Size: Large [1-1]
Experience: Security Monitoring [1-1]
Experience in cyber: 22 year [1-1]
Experience in IT: 16 years [1-1]
is part of
is associated with


Managerial perspective (0-26)
Technical perspective (0-17)
Perspectives (0-2)
Bank Example (0-5)
Legislation increases need for security (1-2)
IPv6 adoption hard to address companies (2-1)
Let the market do itself (2-1)
IPv6 is not a necessity (2-1)
People do not want to learn (1-2)
Cyber Awerness Increase (3-1)
IPv6 adoptions education should increase (3-1)
Availability more important then CI (4-3)
Higher management: SO not avere (1-2)
Companies in denial phase (1-2)
No idea about adoption speed (1-1)
Consultancy adviser: Company be fully IPv6 and translate at the outside (1-1)
Security blocks innovation (1-1)
Buying IPv4, cheaper then implementing IPv6 (1-1)
Consumer no advantage IPv6 (1-1)
Adoption IPv6 slowed down by large companies (1-1)
General public has no knowledge (1-1)
IPv4 does not hurt enough yet (1-1)
IPv6 has no benefits (1-2)
IPv6 address are free (1-2)
Business case cost are higher then it yields (1-1)
Consultancy: Make all processes IPv4 and IPv6 (1-1)
Adoption too slow (1-1)
Availibility important could be working over IPv6 (1-1)
Internet split resosibilities (1-1)
Milenium Bug to IPv6 (1-1)
Internet open place (1-1)
GDPR security motivator (2-1)
Reason no awereness live in comfert zone (1-1)
Higher management: C level not avere (1-1)
Combine IPv6 implementation with life cycle managemetn (1-2)
Create plan, don't use it (1-1)
IOT could be boost (2-1)
Pen testing tools on IPv6 are now being developed (2-1)
All customers on v6 companies on v4 (1-1)
Future work: Do DNS request analyses for larger companies (1-1)
Client side needs to keep supporting IPv4 (1-1)
Security focus is on application not network (1-1)
Enterprise setting public IP adresses is different mindset (1-1)
People make mistakes (1-1)
IOT boost IPv4 devices (1-1)
IPv6 traffic mostly scanning (1-1)
IPv4 lifespawn is extented with patchwork (1-1)
Azure. loads of VM with pubic IPv6 (1-1)
Carrier Grade NAT is expensive (1-1)
CGN is workaround (1-1)
IOT could be lora or zigbee (1-1)
IOT could be iPhone (1-1)
IOT depends on protocol (1-1)
IP protocol battery expensive (1-1)
IOT bridge with IPv6 adresses (1-6)
IOT Class A network large enough for IOT (1-1)
IOT security not connected to the internet (1-2)
is part of
is cause of
contradicts
is associated with


NAT is pseudo security {1-1}
NAT is fine for home network {2-1}
NAT is not security for corperate {3-1}
NAT corperate {0-3}
NAT Home {0-3}
NAT is security for home {1-1}
Normal NAT {0-5}
RFC NAT developers regret NAT {1-2}
NAT is workaround {1-2}
Locked in NAT {1-1}
IPv6 no NAT {2-1}
NAT {0-6}
Grade Carrier NAT {0-8}
Grade carrier NAT is going to be expensive {1-2}
Grade carrier NAT no final solution {1-1}
NAT is negative {4-2}
NAT64 {1-2}
Grade Carrier NAT is negative {4-2}
Grade carrier NAT cheaper then IPv6 implementation {1-1}
Grade Carrier NAT negative IPv6 adopion {2-1}
NAT is a negative machinery {2-1}
Consultancy advice NAT stays NAT {1-2}
is part of
is associated with


Security depending per company {1-1}
is associated with
Availability more important then CI {4-3}
contradicts
Security is cost {4-1}
contradicts
Security {1-5}
is part of
Solution clear defenition who is resposible {1-1}
is part of
Solution small company one external consultant {1-2}
is part of
Solution {1-3}
is part of
Enable DMZ dual stack rest stay IPv4 {3-2}
is part of
More automation less mistakes {1-3}
is part of
Possible implementations {0-9}
is associated with
Consultancy advice NAT stays NAT {1-2}
is associated with
Dual stack could mean different servers: thus different OS {1-1}
is associated with
ISP: IPv6 free, IPv4 addotional service {1-1}
is part of
IPv4 inside IPv6 outside {1-2}
is associated with
No Router advertisment in datacenter because no knowledge if IPv6 is configured {1-1}
is associated with
IPv6 implementation inside vs outside {1-1}
is part of
Tunnel techniques better then Dual Stack {1-1}
is part of
IPv6 problems {1-21}
contradicts
Not enough Awereness {6-1}
is part of
IPv6, no market pull {4-3}
is part of
IPv6 is Future {4-2}
is part of
Grade Carrier NAT is negative {4-2}
is part of
IPv6 is complex {4-2}
is part of
IPv6 configuration: No awereness {3-2}
is part of
IPv6: No business case {3-3}
is part of
Solution: increase awereness: negative things {2-1}
is part of
People learn to think small with IPv4 {2-1}
is part of
Solution: Increase Knowledge {2-1}
is part of
Risk IPv6 it is there {1-1}
is part of
ISP give feedback on open ports ipv4 ipv6 {1-2}
is part of
ISP could increase awerness by adding tools in deamon {1-1}
is part of
Solution: Increase Knowledge: Schools {1-1}
is part of
Solution: make better appointments {1-1}
is part of
Solution: Increase Adoption {1-1}
is part of
Solution: Less resposibilites client {1-1}
is part of
Solution manually configurie IPv6 in datacenter {1-1}
is part of
Solution: increase awereness: benefits {1-1}
is part of
No IP scarce for dutch ISP {1-1}
is part of


IP depletion companies no IPv4 VPN not working {1-1}
IPv6 is difficult {2-1}
No demand pull IPv6 {2-1}
IPv4 no reason to change {2-1}
Needs compelling reason {2-1}
No compelling reason IPv6 {2-2}
Switching cost will pass IPv6 implementation {1-2}
Grade carrier NAT is going to be expensive {1-2}
Switching cost are high and not needed {3-3}
IPv6: No business case {3-3}
IPv6, no market pull {4-3}
Demand Pull {0-13}
IPv6 is missing the iPhone {1-1}
There is no compelling reason IPv6 for SME {1-2}
Market pull needs compelling reason {1-1}
Small Companies not ready for IPv6 {0-2}
Internet not ready for current scale {1-1}
Windows 10 IPv6 default {1-2}
On default IPv6 but no attention {2-1}
Technology Push {0-8}
ISP cannot make ipv6 mandatory {1-1}
Education: IPv6 should be default: IPv4 as legacy {1-1}
IPv6 has been default for 10 years {1-1}
Large companies default enabled IPv6 {1-1}
is associated with
contradicts
is part of
is cause of

Company: Clients: Resposible Open ports [1-1]
Small company: less seperation of work [1-2]
Smaller companies: No seperation of work [1-1]
Port configuration long process [1-1]
Port configuration difficult process [1-1]
Small company: open port: request: client [1-1]
Opening port: Change request [1-1]
IPv6 only: if requested by client [1-2]
Open port need organization translated to request [1-1]
Smaller companies less seperation of roles [2-1]
Larger companies: more work seperation [2-1]
Resposible for open port Firewall admin [3-1]
Port configuration process [0-14]
IPv6 needs ICMP for protocols could be reason for more open [3-1]
Company: open port: request from business [1-1]
Companies: Open Ports: No verification [2-4]
Znap used for external scaning [1-1]
Verification: Himself Private Netork [1-2]
Verification: Open ports: Security Officer [1-2]
Verification: After config changes [1-2]
Verification open ports for ISO standards [1-2]
No clear verification in Companies [1-3]
Verification: Hard process [1-3]
Port verification [0-10]
External scanning should be executed [1-1]
Port configuration [0-4]
Misconfiguration mistake: 2 files [2-1]
No seperation of work [2-1]
Implemented with too few knowledge [2-1]
Misconfiguration ignorance [2-1]
IPv4 IPv6 different firewall [2-1]
IPv6 configuration: No awareness [3-2]
Misconfiguration [0-26]
IPv6 different IP tables file [2-1]
Filtering on source IP [2-1]
Breach: incorrect port open on wrong interface [1-1]
Less IPv6 security because less use [1-1]
Misconfiguration not exploited router [1-1]
IPv6 Overlooked [1-1]
No seperation of work: Mistakes occur [1-1]
Misconfiguration mistake: No knowledge [1-1]
IPv6 firewall deamon was disabled [1-1]
Misconfiguration not only vulnerable but also less performance [1-1]
IPv6 config mistakes Auto config access tunnels [1-1]
Mistake in firewall rules me and me6 [1-1]
Misconfiguration mistake: Focus on IPv4 [1-1]
Misconfiguration lazyness [1-1]
Reason misconfiguration: Ignoranc [1-1]
IPv6 configuration is not in peoples system [1-1]
Security implications [0-11]
Risk: Half implementations [3-1]
Risk IPv6: Sneaks in because of autometic configuration methods [3-1]
IPv6 wrong config-ACL [2-1]
IPv6 new source of vulnerabilities [2-2]
Disabeling V6 is security risk [1-1]
IPv6 security risk: No control tunnel protocols [1-1]
Faulty IPv6 configuration on Cisco routers [1-1]
Advanced hackers use IPv6 for under the radar [1-1]
IPv6 not configured is security risk [1-1]
Hard to create IPv6 visibility [1-1]

E

Recruitment Script

Script for introduction call

Introduction
Introduce yourself. Who are you and use TU Delft as research institution?

Goal
Our goal of this research is to increase awareness for IPv4 and IPv6 on dual stack hosts and to find out if the configuration differences between open port policy are deliberately or on purpose.

Explain the procedure
With Zmap a scan tool similar like Nmap, which is a ports scanner Public IP addresses on Dual stack hosts are scanned on well-known ports.

FTP, SSH, Telnet, SMTP, Netbios, BGP, HTTP, HTTPS, SMB, IPP, mySQL, RDP, VNC, Redis, Elasticsearch, MongoDB, NTP, SNMP, DNS, Memecached. (If you have questions about these protocols you can forward them to me.)

A dual stack host is the same host that is accessible via both IPv4 and IPv6. It would be logic that there would be similarities between open port policies. Unfortunately, there are differences and potential differences are.

IPv4 > IPv6	logic
IPv4 < IPv6	not logic
IPv4 and no IPv6	logic
No IPv4 and open IPv6	Very Unlogic
IPv4 resolves to internal address and IPv6 is active	Risk

What we found on your system, (we can name IP address and hostname depends on the technical knowledge of the person calling) that there is one of the above-mentioned scenario's.

DNS explanation for not technical personnel
A DNS server revolves hosts to IP addresses. This is done because people are not able to remember IP addresses really well and this is why they created a very large lookup register to connect IP addresses to hostnames, which are easier to understand.

Do you have knowledge about the configuration setup or could you bring me into contact with a person who has knowledge about these topics and are you willing to participate in a interview with one of our researchers?

Skype account: [redacted]
Password: [redacted]

Werkinstructie opzoeken
Uit de SQL database is een CSV extract gemaakt. Deze is vervolgens aangevuld en verduidelijkt met wat formules en filters. De .xlsx file heten final dataset com en final dataset nl. De .com heeft ongeveer 300.000 entries dus hij kan wat traag zijn.

Momenteel is het gesorteerd op openstaande poorten op IPv6. Daarnaast rekent het automatisch het verschil in openstaande poorten uit. Wat opvalt is dat op IPv6 vooral 22/SSH open staat. Daarnaast is het ook duidelijk in de dataset dat er meerdere websites draaien op een IP-adres.

Als je wilt kan je zelf spelen met de dataset en opzoek gaan naar interessante gevallen. De makkelijkste manier is denk ik door te sorteren op verschillen en vervolgens een filter te zetten op de hoogste waardes van IPv6. Dan komen de grootste risicogroepen omhoog.

Ik heb nog een extra extract gemaakt uit de database. Deze vond Tobias zelf behoorlijk interessant en dit zijn alle, dus geen selectie op .com en .nl, die op IPv4 resolven naar een intern adres en bij IPv6 resolven naar een extern publiek adres. Dit betekent dat ze in principe achter een NAT/Firewall zijn geconfigureerd maar dus waarschijnlijk niet de juiste security maatregelen hebben die zijn toegepast dan wanneer ze bewust internet facing waren gemaakt. Deze zijn wat lastiger te achterhalen wie de beheerders zijn, maar zijn wel erg interessant om even doorheen te kijken.

Er zit een verschil in de persoon die gebruik maakt van een server en een instantie die hem beheerd. In principe kijken we naar de 2e laag en niet naar de gebruikers die ergens een server hosten. Dit zijn hosting providers. Zij zijn verantwoordelijk voor de security van de systemen. In principe is het doel om te achterhalen wie de hosting providers zijn van de websites, want de gebruikers hebben daar niets mee te maken.

Er zijn verschillende manier om te achterhalen wie de hosting provider is. Hiervoor zijn meerdere websites beschikbaar. Het is ook mogelijk om het IP-adres in een google te gooien met lookup of info en kijken wat daarbij komt.

Mocht het niet duidelijk zijn is de verwachting dat het thuis/intern gehost wordt. In dat geval is het mogelijk om met bedrijf zelf contact op te nemen om dit te bevestigen. Vervolgens wanneer er meerdere URL resolven naar een hetzelfde IP address en de websites niets met elkaar te maken hebben kun je verwachten dat het door een externe hosting partij wordt gedaan.