A Comparative Security Evaluation for IPv4 and IPv6 Addresses

Vincent van der Eijk[*], **Erik Lamers**[†]
University of Amsterdam
{[*]vincent.vandereijk, [†]erik.lamers}@os3.nl

***Abstract*—Research into port security on IPv4 networks is widespread. This is not the case for IPv6 because, until recently scanning the IPv6 address space was considered unfeasible. With the introduction of improved enumeration methods for the IPv6 address space, scanning IPv6 addresses for open ports has become more accessible. In this research we perform a port based security evaluation of dual-stack hosts. We aggregate datasets from various sources to create a final set consisting of 4.5 million dual-stack hosts with their A and AAAA DNS records. 3.4 million hosts are reachable via ICMP and are used in this research.**

We observe that dual-stack hosts are generally more accessible over IPv4 than over IPv6. We see that 26% of hosts have a protocol exposed on IPv4 which is inaccessible over IPv6, while we only find 6% of hosts to have a protocol exposed on IPv6 which is inaccessible over IPv4. Finally, we find more than 50.000 hosts that are accessible over IPv6, but are unreachable over IPv4. Additional research should point out whether such configurations are either intentional, or the result of misconfigurations.

***Keywords*—**IPv4, IPv6, security, ports, protocols

I. Introduction

With the IPv4 address space being depleted we are forced to seek an alternative in IPv6. Since the inital draft of the IPv6 specification in 1995 [1], the standard has been struggling with it's adoption. Only from 2012 onward, a noticeable change in the usage of IPv6 can be observed. From 2015 onward the IPv6 user adoption is growing with a constant rate of approximately 5% each year, according to IPv6 usage statistics published by Google [2].

The adoption of IPv6 introduces a new attack surface because of misconfigured access control mechanisms, which is already shown by RFC 4942 that was published in 2007 [3]. Previous research performed by Czyz et al. [4] showed a noticeable difference in port based security policies between IPv4 and IPv6 networks. This research, performed in 2015, shows that the level of network security of dual-stack systems is often worse via IPv6 than via IPv4. The primary focus of the research is to study protocol or application weaknesses for IPv6 caused by lacking security policies. Both security awareness and the adoption of IPv6 have been under constant development over the past four years which makes this a suitable moment to address the discrepancy in IPv4 and IPv6 security once again [5].

The ZMap project made it possible to scan the entire IPv4 address space in under 45 minutes by developing a stateless network scanner [6]. Our research would not be possible without this tool and the IPv6 capabilities added by the research group from Technical University of Munich [7].

When comparing IPv4 and IPv6 networks we can identify various differences that might have more severe security implications if not properly mitigated [8].

- **Network address translation (NAT)**. This technique was introduced in 1994 [9] to be able to map a private address space to a network. Although this was never focused on security it acts as a secure default that makes sure public IP's cannot access devices behind the NAT. NAT is generally not used in IPv6 networks, which means that additional firewall security is required to achieve one way initialization of traffic policy that NAT achieves in IPv4 networks.
- **Firewall configuration**. The tool *iptables* which is used for firewall configuration on Linux based systems is only applicable for IPv4 rules. In order to configure a firewall for IPv6, the tool *ip6tables* is required. Unless a system administrator explicitly configures his *ip6tables*, the network is probably wide open.
- **Auto-configuration**. Most Operating Systems (OS)s these days have stateless address auto-configuration (SLAAC) for IPv6 turned on by default [10]. This means that if auto-configuration is available on the network they will get an IPv6 address assigned by default, and

try to connect to the network. Depending on the type of networking equipment used it might be as easy as enabling IPv6 on the router and all connected devices will get IPv6 connectivity. This is very user friendly, but can easly lead to misconfiguration, as shown by Borgolte et al. [11].

A. Research question:

Our main research question is defined as follows:
Has the state of IPv6 port based security compared to IPv4 port based security shifted over the last four years?
To answer this question we will look into the following subquestions:

- How do the IPv4 and IPv6 port based security rules currently compare to each other?
- Is there a trend to be seen in the port based security policies compared to four years ago?
- What portion of the IPv6 addresses used is actually reachable compared to IPv4 when it comes to dual stack hosts?

In this paper we make the following contributions:

- The dataset we use is this research consists of 4.5 million dual-stack hosts and is therefore 9 times larger than previously used datasets [4].
- We scan dual-stack hosts for a larger set of 21 protocols (23 ports), all of which might have security implications if misconfigured. With our methodology we have the potential to find more misconfigurations than any other approach.
- We define a host definition that ensures ethical issues of scanning the internet are mitigated as much as possible.

We have various data sources at our disposal that will be used to build a target list as a basis for this research. These separate datasets and the process required to build a single set is described in more detail in Section II. Section III describes the ethical implications of this research and the various measures we take to make this research as less intrusive as possible. The results are presented in Section IV were we visualize the discrepancies in IPv4 and IPv6 configurations, and present additional remarkable observations that we made during the course of this research. Section V introduces related research that we use as a basis four our discussion in Section VI. This leads to our conclusion as described in Section VII. We conclude with recommendations for future work in Section VIII.

II. Methodology

We divide this research into five main parts. First, we create a definition for a host and it's reachability. With this definition, which is described in the following section, we can develop a large universal datatset of hosts out of sanitizing multiple smaller datasets. This universal dataset is used throughout our security evaluation for IPv4 and IPv6 networks. The hosts in this universal dataset will be scanned for reachability. Subsequently, the reachable hosts will be scanned for specific protocol connectivity. Finally, all responding TCP ports will also be probed for banners. In the end all results between IPv4 and IPv6 open ports will be compared against the universal dataset.

A. Host definition

We define a host as follows: A host must have a reachable IPv4 or IPv6 address. We define reachability as a host responding to an ICMP echo request. We recognize that not all System Administrators (SA) want their network scanned and will have some firewall blocking measures in place. If the SA filters ICMP echo requests, we will not include their IP ranges in our dataset. Because we want to analyze the differences in port security, we say that a host can be reachable via either IPv4 or IPv6. This is so we do not exclude any hosts that are for example reachable via IPv6 but blocked on IPv4. A host can be any type of machine. We do not make any distinction between servers and networking devices as done by Czyz et al. [4]. With the introduction of software defined networking virtually any hardware can be used as networking equipment, making the distinction between routers and servers much more vague [12].

B. Definition of reachability

We consider a TCP port reachable if it responds with a TCP SYN+ACK to a TCP SYN as used by Czyz et al. [4]. For UDP we expect the correct UDP response to a UDP request, as this is protocol specific.

C. Building a universal dataset

We create a dataset consisting of dual-stack hosts from three separate sources. These sources are the Alexa top 1 million [13], the Rapid7 FDNS ANY [14] and the IPv6 ICMP hitlist [15]. All of those are used as input for the universal dataset. Table I outlines the dates on which the datasets were gathered. Both the Rapid7 and the Alexa datasets are used by the research of Czyz et al [4]. These datasets are primarily based on the IPv4 addresses of publicly available web servers, and therefore might introduce bias into the final dataset to contain relatively many web servers. The IPv6 ICMP hitlist [15] is generated using an algorithm to predict existing IPv6 addresses based on previously known address spaces. With the addition of this dataset we intend to reduce the bias in the dataset and to reflect a more accurate cross-section of the internet.

In order to make a comparison between IPv4 and IPv6 networks, we need to create pairs of addresses corresponding to the same host. These pairs are created based on the hostnames from the datasets mentioned in Table I. If the dataset does not

TABLE I
DATASETS USED TO CREATE UNIVERSAL DATASET

Dataset	Date	Portion of final set
Rapid7 FDNS ANY	28-12-2018	96.8%
Alexa top 1 million	07-01-2019	2.0%
IPv6 ICMP hitlist	08-01-2019	1.2%

include a hostname we use reserve DNS (rDNS) to perform a lookup on the hostname, or discard the address if no PTR record is found. Because a host must have a unique address pair according to our host definition all addresses are included in the final dataset. All the results in Section IV are based on this dataset, which contains 4.5 million unique address pairs.

D. Scanning method

To perform the network scan on a global level the tool ZMap [16] is used. In order to perform this scan in a reasonable amount of time an uplink of 1 Gbps is used. Caution is required to assure that no Denial-of-Service (DoS) is caused due to a misconfiguration or other kind of error. To make sure this does not occur during our research we use the ZMap load distribution algorithm [6]. An IPv6 capable ZMap fork is used for our probes [7].

E. Protocols

For the protocols to be scanned we will extend on the protocols scanned by Czyz et al. [4]. This research scanned for 11 distinct protocols. We intend to increase this amount by including more ports that are known for common and vulnerable services when exposed to the internet. We base these protocols on the protocols that are used the most as defined by NMAP [17] and protocols that are often misconfigured as shown by Fiebig et al. [18, 19]. Examples of these services are SMTP (25), VNC (5900), MySQL (3306), MSSQL (1433), and RDP (3389). Table II shows the protocols that have been selected for this research. The dates of probing can be found in Appendix B.

We also considered more router specific protocols such as RIP. However as these protocols often have a different implementation for IPv6, and this would not provide comparable results.

F. Banner grabbing

For the TCP protocols mentioned in Table II we perform a banner grab to confirm that the service running on that particular port is actually responding and not just a firewall responding with a TCP SYN+ACK to every request. Performing the banner grab will reduce any false positives we might encouter to a minimum. The banner grabbing is not necessarily required for the UDP probing due to the nature of the UDP protocol. In the event that a response is received from an UDP scan this is a indicator that a service is running on the targeted port. Therefore, false positives are less likely to happen for UDP scanning.

TABLE II
PROTOCOLS AND PORTS TO BE PROBED

Protocol	TCP/UDP	Port(s)
FTP	TCP	20, 21
SSH	TCP	22
Telnet	TCP	23
SMTP	TCP	25
Netbios	TCP	139
BGP	TCP	179
HTTP	TCP	80, 8080
HTTPS	TCP	443, 8443
SMB	TCP	445
IPP	TCP	631
MSSQL	TCP	1433
MySQL	TCP	3306
RDP	TCP	3389
VNC	TCP	5800, 5900-5901
Redis	TCP	6379
Elasticsearch	TCP	9200, 9300
MongoDB	TCP	27017-27019
NTP	UDP	123
SNMP(v1/2)	UDP	161
DNS	TCP/UDP	53
Memcached	TCP/UDP	11211

For the banner grabbing we will make use of the applications ZGrab and ZGrab2 [16]. Although the original ZGrab is officially deprecated, it is still required to perform banner grabbing for the TCP protocols that are not supported by ZGrab2. We will perform a generic banner grab for the remaining protocols with the original ZGrab. Appendix C provides an overview of which application is used for each protocol.

III. ETHICS

This research involves active probing of hosts on the public Internet. To minimize our impact on the various stakeholders we will adhere to the guidelines of the Menlo Report [20] while conducting our research:

- We will only use public available information to identify hosts and corresponding email addresses. (e.g., WHOIS, DNS information).
- The probing of hosts will consist of establishing a connection with the host and performing a banner grab and protocol information (SYN + service discovery), without further exploration into the correctness of authentication of the given protocol.
- We will utilize the proven ZMap host distribution algorithm to disperse load across networks [6] as it is a accepted best practice in the security community [20].
- An opt-out feature will be made available so SAs can contact us, and we will exclude their IP range from the target list. The opt-out feature is described in § III-A.

- A host is defined as reachable by ICMP echo request. This is intentional, as there is a good chance that a SA who does not want their network scanned will block ICMP echo requests. Thus by this host definition those hosts will automatically be excluded from any scans.

A. Informed consent

In order to adhere to the guidelines of the Menlo Report [20] informed consent needs to be obtained from the research subjects. Because of the large scale of this research it is not feasible to receive written consent of all subjects. Instead, we will provide an opt-out feature for our subjects while the experiment is ongoing. System administrators can use this opt-out feature to exclude their domains from our scans. This will be done with rDNS. The IP addresses used for scanning will have a PTR record pointing to a web page hosted by us. This web page will describe the purpose of the research and the impact it has on the subjects and provides detailed instructions for subjects to withdraw from the research. Furthermore, all IP addresses that we use for scanning will have a webpage running on port 80, which will state the nature of the research and a link to the webpage mentioned above. This method is also used by the ZMap research to adhere to "good internet citizenship" [6]. If the opt-out is received after the target has already been scanned, all the corresponding hosts will be discarded from all datasets. This is in line with the guidelines of the Menlo Report which states that "subjects must be free to withdraw from research participation without negative consequences." [20]

B. Data gathering and storage

As described by the Menlo Report [20] our results could potentially be used by malicious actors. All results presented in this paper are anonymized and aggregated into larger sets to prevent potential abuse. We do recognize that our dataset could be of great value to future research and intend to hand over the dataset to the SNE research group of the University of Amsterdam and Tobias Fiebig of the Delft University of Technology. The dataset is to be used for further research with a written statement that this dataset will not be distributed outside of both universities.

IV. Results

From the three datasets mentioned in the methodology we were able to extract 22.5 million unique hostnames. For the result plots we will use the unique address pairs that can be made from this set based on the hostnames. This comes down to a set of 4.5 million hosts. This address mapping is available in the raw results of the researchers. All probing dates can be found in Appendix B.

A. ICMP reachability

For the IPv4 we found that 66% of the 2.4 million addresses were reachable via ICMP. For the IPv6 67% of the 3.9 million addresses were reachable. Both protocols have similar reachability via the ICMP echo request.

When looking at the address pairs we can see that 76% were reachable on both IPv4 and IPv6. 20% is reachable on IPv4 but not on IPv6. Only 4% is reachable on IPv6 but not on IPv4. The summary of results can be seen in Table III

TABLE III
Dataset Summary

Protocol	Adresses	ASNs	Prefixes	ICMP reachable
IPv4	2.4M	13264	57030	1.6M
IPv6	3.9M	8208	15684	2.8M

B. TCP SYN/UDP scans

Overall, the majority of the hosts in our dataset has services enabled on both IPv4 and IPv6. Figure 1 shows that 72% of the hosts are reachable for at least one protocol over both IPv4 and IPv6. Roughly 10% of the hosts in our dataset only allow connections over IPv4 and do not appear to have services that respond over IPv6. In contrast, only 1.6% of the hosts only allow connections over IPv6 and do not respond to services over IPv4. The remaining part of the hosts do not have active services on either IP protocol.

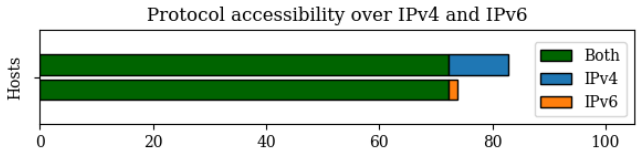

Fig. 1. **Protocol accessibility over IPv4 and IPv6, with at least one protocol reachable for the relevant IP protocol.**

The results in Figure 2 show that 6.1% of hosts have at least one service that is reachable over IPv6 while the same service is not reachable over IPv4. On the other hand, 26% of the hosts had at least one service reachable on IPv4 that was not accessible over IPv6. Furthermore, we found that 1.67% of hosts have at least one application which is reachable over IPv4 and not via IPv6 *and* at least one application which is reachable via IPv6 and not via IPv4.

The TCP SYN and UDP scans resulted in a large amount of responses from both IPv4 and IPv6 addresses. Although the host dataset contains more distinct IPv6 addresses overall, IPv4 provided more responses in absolute terms, as can be seen in Figure 3 and Figure 4. The full list of results can be seen in Table IV.

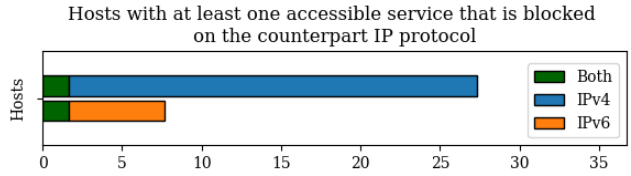

Fig. 2. **Hosts with at least one service accessible that is blocked on the counterpart IP protocol. The green bar labeled 'both' means that there is a protocol which is accessible on IPv4 which is not accessible on IPv6 *and* a protocol which is accessible on IPv6 which is not accessible on IPv4.**

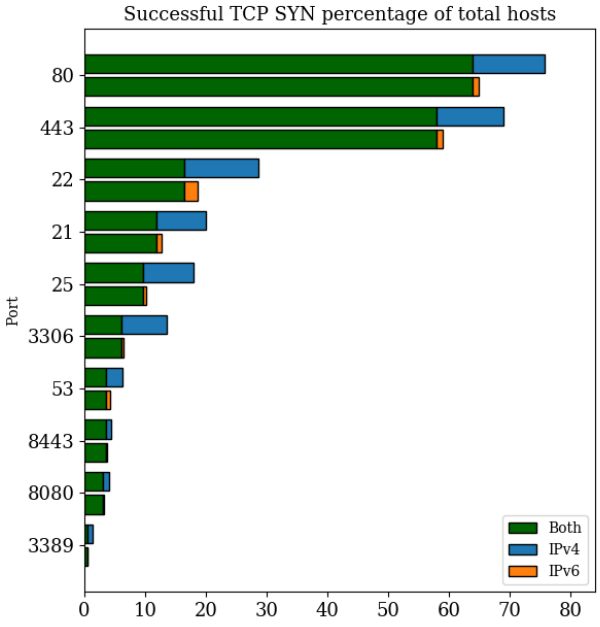

Fig. 3. **Percentage of 3.4M dual-stack hosts that were open for IPv4 and/or IPv6 for the 10 most responsive TCP applications. None of the most responsive TCP applications are more open in IPv6.**

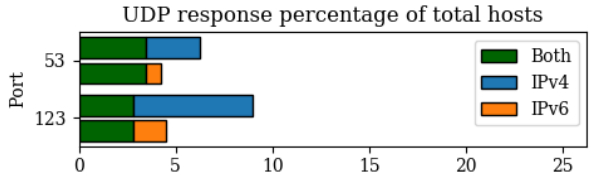

Fig. 4. **Percentage of 3.4M dual-stack hosts that were open for IPv4 and/or IPv6 for the most responsive UDP applications. None of these UDP applications are more open in IPv6.**

C. Banner grabbing

For all hosts were the TCP SYN probe was successful a banner grab was performed. Some protocols yielded no results, because either they did not respond with any banner to a connection request or some for of initial authentication was required[1]. These protocols have been omitted in the results. For all protocols that could successfully be banner grabbed, we saw an average success rate of 85%. The results of the banner grabs can be seen in Figure 5 and Table IV.

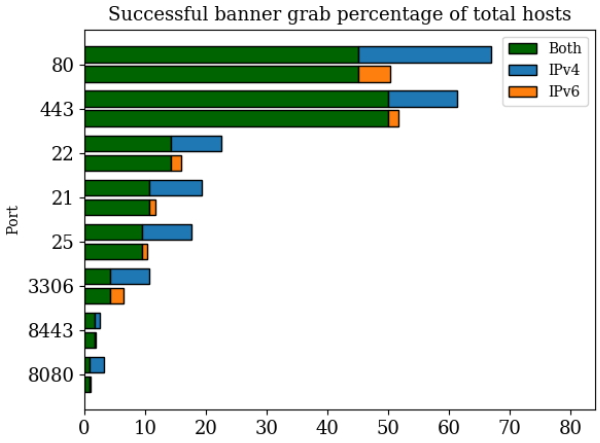

Fig. 5. **Percentage of 3.4M dual-stack hosts that were open for IPv4 and/or IPv6 on which we could successfully perform a banner-grab on for the set of hosts from Figure 3. None of the most responsive TCP applications are more open in IPv6.**

D. Local addresses

In our results we encountered a percentage of hosts with a private [21] or link-local [22, 23] address with a public address on the other IP protocol. We found that 7769 hosts have a private IPv4 address with a public IPv6 address listed in the DNS information and where the public IPv6 address is reachable over ICMP. 2331 of these hosts have a responding port on IPv6 which is unresponsive over IPv4. For IPv6, 4151 hosts have a link-local or private address with a public IPv4 address that is reachable over ICMP. 3819 of these hosts have at least one responsive IPv4 port which is unresponsive over IPv6.

For localhost addresses we found that 354 hosts have a localhost address on IPv4 with a public address for IPv6. 148 of these hosts have a responsive port on that IPv6 address. For IPv6, 4951 hosts have a localhost address on IPv6 with a public address on IPv4. 4302 of these hosts have a responsive port on that IPv4 address. This is a strong indication of a misconfiguration for both situations.

E. Received abuse reports

As described in Section III we took extensive precautions to make sure that our scans did not interfere with normal

[1]The protocols were banner grabbing was not successful include: DNS, Memcached, MongoDB, RDP and Redis.

TABLE IV
IPv4 (left), IPv6 (right) result numbers for TCP SYN and UDP scans and successful banner grabs. Percentages are from the total number of pairs. Were banner grabbing was not successful or yielded no results, the data is omitted.

IPv4			
Protocol	**Type/Port**	**Connections**	**Banners**
BGP	tcp/179	24.33K (0.71%)	1.59K (0.05%)
DNS	tcp/53	225.83K (6.59%)	-
	udp/53	217.41K (6.35%)	-
Elastics.	tcp/9200	5.53K (0.16%)	-
	tcp/9300	4.83K (0.14%)	-
FTP	tcp/20	3.5K (0.10%)	306 (0.01%)
	tcp/21	687.07K (20.06%)	663.18K (19.37%)
HTTP	tcp/80	2.59M (75.74%)	2.29M (67.05%)
	tcp/8080	137.63K (4.02%)	111.25K (3.25%)
HTTPS	tcp/443	2.36M (69.03%)	2.10M (61.32%)
	tcp/8443	151.55K (4.43%)	86.20K (2.52%)
IPP	tcp/631	3.56K (0.10%)	394 (0.01%)
Memcached	tcp/11211	4.87K (0.14%)	-
	udp/11211	45 (0.00%)	-
MongoDB	tcp/27017	5.61K (0.16%)	-
	tcp/27018	3.08K (0.09%)	-
	tcp/27019	3.00K (0.09%)	-
MSSQL	tcp/1433	32.73K (0.96%)	8.21K (0.24%)
MySQL	tcp/3306	462.64K (13.51%)	362.86K (10.59%)
Netbios	tcp/139	17.48K (0.51%)	520 (0.02%)
NTP	udp/123	311.84K (9.11%)	-
RDP	tcp/3389	43.59K (1.27%)	-
Redis	tcp/6379	5.88K (0.17%)	-
SMB	tcp/445	19.50K (0.57%)	15.34K (0.45%)
SMTP	tcp/25	615.83K (17.98%)	602.42K (17.59%)
SNMP	udp/161	1.87K (0.05%)	-
SSH	tcp/22	978.80K (28.58%)	772.07K (22.55%)
Telnet	tcp/23	17.08K (0.49%)	2.29K (0.07%)
VNC	tcp/5800	3.58K (0.10%)	-
	tcp/5900	6.66K (0.19%)	2.47K (0.07%)
	tcp/5901	5.46K (0.16%)	1.41K (0.04%)

IPv6			
Protocol	**Type/Port**	**Connections**	**Banners**
BGP	tcp/179	24.43K (0.71%)	3.16K (0.09%)
DNS	tcp/53	147.24K (4.30%)	-
	udp/53	147.71K (4.31%)	-
Elastics.	tcp/9200	1.84K (0.05%)	-
	tcp/9300	1.67K (0.05%)	-
FTP	tcp/20	301 (0.01%)	105 (0.00%)
	tcp/21	433.16K (12.65%)	399.53K (11.67%)
HTTP	tcp/80	2.22M (64.85%)	1.72M (50.23%)
	tcp/8080	108.55K (3.16%)	36.13K (1.06%)
HTTPS	tcp/443	2.02M (59.10%)	1.77M (51.69%)
	tcp/8443	127.85K (3.73%)	64.31K (1.88%)
IPP	tcp/631	1.00K (0.03%)	-
Memcached	tcp/11211	3.05K (0.09%)	-
	udp/11211	2 (0.00%)	-
MongoDB	tcp/27017	550 (0.02%)	-
	tcp/27018	278 (0.01%)	-
	tcp/27019	218 (0.01%)	-
MSSQL	tcp/1433	2.49K (0.07%)	2.07K (0.06%)
MySQL	tcp/3306	221.83K (6.48%)	159.80K (4.67%)
Netbios	tcp/139	5.26K (0.15%)	390 (0.01%)
NTP	udp/123	156.44K (4.57%)	-
RDP	tcp/3389	19.08K (0.56%)	-
Redis	tcp/6379	2.33K (0.7%)	-
SMB	tcp/445	15.10K (0.44%)	11.73K (0.34%)
SMTP	tcp/25	350.93K (10.25%)	350.00K (10.22%)
SNMP	udp/161	1.93K (0.06%)	-
SSH	tcp/22	636.09K (18.58%)	547.53K (15.99%)
Telnet	tcp/23	14.31K (0.42%)	3.22K (0.09%)
VNC	tcp/5800	259 (0.01%)	-
	tcp/5900	2.32K (0.07%)	1.19K (0.03%)
	tcp/5901	1.44K (0.04%)	513 (0.01%)

operations of the network being scanned. During our scans we did encounter automatic network monitoring that detected our scans. From the amount of responses received from the different parties we can clearly see that IPv4 networks are more heavily monitored than IPv6 networks are. The summary of the amount of emails received can be seen below:

- 25 abuse responses only listing IPv4 addresses.
- 6 abuse responses only listing IPv6 addresses.
- 1 abuse response listing both IPv4 and IPv6 addresses.

Seven of those abuse reports were also sent to the abuse email address listed on the webpage of the scanning nodes. The other abuse reports were sent to the abuse email listed in the WHOIS information of the IP addresses. Furthermore, we received three opt-out requests from SAs. Two of which were received more then five days after the scans had been completed. An explanation for the relatively few abuse responses and opt-out requests is due to the fact that our dataset is primarily based on the Rapid7 FDNS ANY dataset as shown in Table III. This dataset already excludes IP ranges from which an opt-out request has been received by Project Sonar [24]. The three opt-out request that we received during the scans were only accountable for less than 0.02% of our original dataset.

V. Related work

This research extends upon previous work conducted by Czyz et al. [4] that was published in 2016, which outlined the issue that there is a noticeable difference in port based security policies between IPv4 and IPv6 networks. We aim to observe if this difference has shifted over the last four years as both security awareness and IPv6 adoption have been constantly developing.

In two recent studies performed by Gasser et al. [7, 15] researchers investigated the feasibility of enumerating IPv6 addresses. This research made use of the Rapid7 FDNS ANY dataset (among others), which could also be valuable for our research. Additionally, Fiebig et al. [25] developed an algorithm to enumerate IPv6 addressees using DNS denial of existence (NXDOMAIN) records. What all of these papers

have in common is that they use some form of (r)DNS to gather their results, which implies that (r)DNS is an important tool to be able to gather a usable IPv6 dataset. Fiebig et al. [26] show that rDNS can be leveraged for building reliable Internet measurement datasets, that state that "rDNS can be relied on for Internet-wide studies" [26].

In early 2018, Borgolte et al. [11] showed that enumerating IPv6 hosts via DNSSEC using NSEC3 records is possible. Furthermore they also found that a lot of dual-stack hosts expose vulnerable and critical services to the Internet: "Based on the enumerated address set, we evaluated the state of security of IPv6 hosts and we have shown that many are insufficiently secured. Specifically, IPv6-enabled systems often expose critical infrastructure or sensitive and privacy-concerning information to the outside" [11].

The article *'IPv666 - Address of the Beast'*, describes the differences in security posture between IPv4 and IPv6 [8]. The researchers highlight the additional attack surfaces that IPv6 networks have, but also the challenge involved in the scanning IPv6 networks due to the enormous amount of possible addresses. The following statement made by the authors of the article was also the starting point for our research: "Unless you've explicitly set up your IPv6 firewall rules you're probably wide open" [8]. A non-conventional method is required to scan the IPv6 address space efficiently, which is confirmed with the statement that "attempting to scan across the IPv6 address space using standard high-throughput scanning tools like ZMap and MassScan will not do much of anything" [8]. Instead, they developed the tool *ipv666* that applies a statistical model to predict new IPv6 addresses based on known a known address space [27]. Although the researchers state that they intend to measure the difference in security posture between IPv4 and IPv6 networks we feel their main focus is on the weaknesses in IPv6 addressing. We aim to use the results from this article [8] and as a basis for a true comparison in security posture for dual-stack hosts.

Also, there is a human factor to this research as network misconfigurations are often a overlooked security issue as already shown by Dietrich et al. [28] and more generally by Fiebig et al. [18].

As shown in this paper we used ZMap and ZGrab to gather our results. Durumeric et al. [6] developed ZMap in order to make it possible to scan the entire IPv4 address space in under 45 minutes. Their stateless probing method allows for many parallel probes at a time. The distribution algorithm ZMap uses provides an assurance that networks are not overloaded while performing scans. Datasets such as the Rapid7 ANY DNS uses ZMap to keep it up to date [24].

VI. Discussion

Our findings on port security for IPv4 an IPv6 networks conflict with the findings of Czyz et al. [4]. They found relatively more IPv6 ports that were exposed to the Internet compared to our findings. First of all, this could be explained if the port security policies might have shifted over the past four years. Additionally, SAs might have become more aware of the need to secure and protect their IPv6 network. We used a much larger dataset which is largely based on the Rapid7 FDNS ANY dataset [14]. This dataset started out with IPv4 data sources from various earlier projects, most of which scraped the IPv4 space for usable addresses. This could be a contributing factor to the fact that we observed more reachable IPv4 addresses. Furthermore, the Rapid7 FDNS ANY dataset was larger then the other two datasets we used. As can be seen in Table I, this dataset has a majority in the final set.

Because our dataset is different then the dataset used by Czyz et al. [4], we cannot compare these results without introducing bias. This means we have to base our conclusions solely on the results we found and the general consensus that has been shown by previous research in Section V.

VII. Conclusion

When comparing IPv4 and IPv6 results, we can see that IPv4 is the more dominant protocol. In both relative and absolute numbers there are more hosts exposed over IPv4 than that hosts are exposed over IPv6, as shown in Figure 2. From this figure we observed that only 74% of the hosts in our dataset have at least one port exposed on IPv6, compared to 83% of hosts that have at least one port exposed on IPv4. Additionally, the scan results in Table IV show that the banner grabbing resulted in a higher overall amount of responses for services on IPv4 enabled hosts. As our dataset only consists of dual-stack hosts with an A and an AAAA DNS record present and in an ideal situation one would expect these numbers would match.

Figure 1 shows that four times as much hosts have at least one port reachable over IPv4 that is not reachable over IPv6 than vice-versa.

Where 7 out of 11 applications were more open for IPv6 than for IPv4 in 2015 as found by Czyz et al. [4], our results from Table IV state that 21 out of 23 ports are more open on IPv4, and only 2 ports are more open on IPv6. Both protocols that are more open on IPv6 represent only a small fraction of the entire dataset. Although we cannot compare the two studies as a whole, we did observe notable differences.

Additionally, Czyz et al [4]. found that 26% of the probed hosts had at least one application reachable over IPv6 that was not reachable over IPv4, while this was only 6.1% for our results. Instead, we found that 26% of hosts had at least one application reachable over IPv4 that was not reachable on IPv6, in contrast to the research of Czyz et al. [4] where this was only the case for 17% of the hosts in their dataset.

Our findings show that IPv4 is still the dominant protocol on the public Internet. We observed that dual-stack hosts are generally more accessible over IPv4 networks than over IPv6 networks. Our methodology did not cover a way to determine whether this observed behaviour is intentional or not, which

is an opening for future work as discussed in the following section. Besides from this unknown factor whether the found configurations are intentional or not, we observe a substantial smaller amount of protocols that are reachable over IPv6 compared to the work of Czyz et al. [4] four years ago.

VIII. Future work

Together with the Autonomous System (AS) information included in the dataset, a specific questionnaire or interview could be set out to learn about the differences in port security across organizations, while also achieving a disclosure of our findings to the responsible SAs. This interview can give an insight into the reasoning behind the differences in port security and might reveal interesting statistics.

Because our observations conflict with the findings of previous papers like Czyz et al. [4], it is important that these results are verified. Using different datasets to compare to might yield different results than the results that we found.

Looking into other pairing methods to define a host might give different results. We used the hostnames corresponding to the IP addresses to map addresses together to create unique pairs. This still leaves room for duplicate addresses existing in the mapping. Looking into other ways of pairing might provide different results on the the same data, such as mapping by IP prefix instead of by hostname.

We have a lot of results in the dataset that we did not analyze. Future research could be performed to find out if there is a correlation between the differences in port security and the type of business behind the corresponding IP-ranges.

Finally, our local addressing observations as described in § IV-D can also be leveraged for future research. Our results and dataset can be used to spot issues with operators abusing certain IP space, such as the address 1.1.1.1 being used as a local address, as experienced by Cloudflare [29].

Acknowledgements

We would like to thank Tobias Fiebig of the Delft University of Technology for the supervision and guidance during this research. Furthermore, we would like to thank Oliver Gasser of the Chair of Network Architectures and Services at the Technical University of Munich for providing us with access to their IPv6 datasets. Lastly we would like to thank Niels Sijm of the OS3 team at the University of Amsterdam for providing us with the hardware required to conduct our experiments.

References

[1] S. Deering and R. Hinden. *Internet Protocol, Version 6 (IPv6) Specification*. RFC 1883. IETF, Dec. 1995. URL: http://tools.ietf.org/rfc/rfc1883.txt.

[2] Inc. Google. *Google IPv6 Statistics*. URL: https://www.google.com/intl/en/ipv6/statistics.html (visited on 10/01/2019).

[3] E. Davies, S. Krishnan and P. Savola. *IPv6 Transition/Co-existence Security Considerations*. RFC 4942. IETF, Sept. 2007. URL: http://tools.ietf.org/rfc/rfc4942.txt.

[4] J. Czyz et al. "Don't Forget to Lock the Back Door! A Characterization of IPv6 Network Security Policy". In: *Proceedings of the 23rd Network and Distributed System Security Symposium (NDSS)*. San Diego, CA, USA: Internet Society (ISOC), Feb. 2016. ISBN: 189156241X.

[5] A. Shiranzaei and Rafiqul Z. Khan. "IPv6 Security Issues - A Systematic Review". In: *Next-Generation Networks*. Singapore: Springer, 2018, pp. 41–49. ISBN: 978-981-10-6005-2.

[6] Z. Durumeric, E. Wustrow and J. A. Halderman. "ZMap: Fast Internet-wide Scanning and Its Security Applications." In: *Proceedings of the 22rd USENIX Security Symposium*. Washington, D.C.: USENIX Association, Aug. 2013, pp. 47–53. ISBN: 978-1-931971-03-4.

[7] O. Gasser et al. "Scanning the IPv6 Internet: Towards a Comprehensive Hitlist". In: *Proceedings of the 2016 International Workshop on Traffic Monitoring and Analysis (TMA)*. arXiv: 1607.05179v1. Louvain La Neuve, Belgium: IFIP, Apr. 2016.

[8] C. Grayson and M. Newlin. *IPv666 - Address of the Beast*. URL: https://1.avala.mp/?p=285 (visited on 04/12/2018).

[9] K. Egevang and P. Francis. *The IP Network Address Translator (NAT)*. RFC 1631. IETF, May 1994. URL: http://tools.ietf.org/rfc/rfc1631.txt.

[10] S. Thomson, T. Narten and T. Jinmei. *IPv6 Stateless Address Autoconfiguration*. RFC 4862. IETF, Sept. 2007. URL: http://tools.ietf.org/rfc/rfc4862.txt.

[11] K. Borgolte et al. "Enumerating Active IPv6 Hosts for Large-scale Security Scans via DNSSEC-signed Reverse Zones". In: *Proceedings of the 39th IEEE Symposium on Security & Privacy (S&P)*. San Francisco, CA, USA: IEEE, May 2018, pp. 438–452. ISBN: 978-1-5386-4353-2.

[12] D. Kreutz et al. "Software-Defined Networking: A Comprehensive Survey". In: *Proceedings of the IEEE* 103.1 (Jan. 2015), pp. 14–76.

[13] Inc. Alexa Internet. *Alexa top 1 million dataset*. URL: http://s3.amazonaws.com/alexa-static/top-1m.csv.zip (visited on 07/01/2019).

[14] Rapid 7 open data labs. *Forward DNS (FDNS)*. URL: https://opendata.rapid7.com/sonar.fdns_v2/ (visited on 28/12/2018).

[15] O. Gasser et al. "Clusters in the Expanse: Understanding and Unbiasing IPv6 Hitlists". In: *Proceedings of the 2018 Internet Measurement Conference*. Boston, MA, USA: ACM, Nov. 2018, pp. 364–378. ISBN: 978-1-4503-5619-0.

[16] The ZMap Team. *The ZMap Project*. URL: https://zmap.io/ (visited on 01/12/2018).

[17] G. F. Lyon. *Nmap Network Scanning*. Nmap Project, 2009. ISBN: 0979958717.

[18] T. Fiebig et al. "SoK: An Analysis of Protocol Design: Avoiding Traps for Implementation and Deployment". In: *CoRR* (Aug. 2016). arXiv: 1610.05531v1.

[19] T. Fiebig et al. "Learning from the Past: Designing Secure Network Protocols". In: *Cybersecurity Best Practices: Lösungen zur Erhöhung der Cyberresilienz für Unternehmen und Behörden*. Springer, 2018, pp. 585–613. ISBN: 978-3-658-21655-9.

[20] M. Bailey et al. "The Menlo Report". In: *IEEE Security & Privacy* 10.2 (Mar. 2012), pp. 71–75.

[21] Y. Rekhter et al. *Address Allocation for Private Internets*. RFC 1918. IETF, Feb. 1996. URL: http://tools.ietf.org/rfc/rfc1918.txt.

[22] R. Hinden and S. Deering. *IP Version 6 Addressing Architecture*. RFC 4291. IETF, Feb. 2006. URL: http://tools.ietf.org/rfc/rfc4291.txt.

[23] M. Cotton and L. Vegoda. *Special Use IPv4 Addresses*. RFC 5735. IETF, Jan. 2010. URL: http://tools.ietf.org/rfc/rfc5735.txt.

[24] Rapid 7 open data labs. *About Open Data*. URL: https://opendata.rapid7.com/about/ (visited on 28/12/2018).

[25] T. Fiebig et al. "Something from nothing (There): collecting global IPv6 datasets from DNS". In: *Proceedings of the 12th Passive and Active Measurement (PAM)*. Vol. 10176. Lecture Notes in Computer Science. Sydney, Australia: Springer, Mar. 2017, pp. 30–43. ISBN: 978-3-319-54328-4.

[26] T. Fiebig et al. "In rDNS We Trust: Revisiting a Common Data-Source's Reliability". In: *Proceedings of the 13th Passive and Active Measurement (PAM)*. Vol. 10771. Lecture Notes in Computer Science. Berlin, Germany: Springer, Mar. 2018, pp. 131–145. ISBN: 978-3-319-54327-7.

[27] M. Newlin C. Grayson. *ipv666 Github*. URL: https://github.com/lavalamp-/ipv666 (visited on 04/12/2018).

[28] C. Dietrich et al. "Investigating System Operators' Perspective on Security Misconfigurations". In: *Proceedings of the 25th ACM SIGSAC Conference on Computer and Communications Security (CCS)*. Toronto, ON, Canada: ACM, Oct. 2018, pp. 1272–1289.

[29] S. Marty. *Fixing reachability to 1.1.1.1, GLOBALLY!* URL: https://blog.cloudflare.com/fixing-reachability-to-1-1-1-1-globally (visited on 04/02/2018).

Appendix A
Software versions

Software	Version
MySQL	5.7.24
Python	3.6.7
ZGrab	96cfb9f[2]
ZGrab2	65a2154[3]
ZMap	28e9dfe[4]

Appendix B
Dates of probing

Type	Date
ICMP	30-01-2019
TCP SYN/UDP	30-01-2019
Banner grabbing	31-01-2019

Appendix C
Supported banner grabing protocols

Protocol	ZGrab	ZGrab2
FTP		x
SSH		x
Telnet		x
SMTP		x
Netbios	x	
BGP	x	
HTTP		x
HTTPS		x
SMB		x
IPP		x
MSSQL		x
MySQL		x
RDP	x	
VNC	x	
Redis		
Elasticsearch	x	
MongoDB	x	
DNS		
Memcached	x	

[2] github.com/zmap/zgrab
[3] github.com/zmap/zgrab2
[4] github.com/tumi8/zmap