Less is More? Exploring the Impact of Scaled-Down Network Telescopes on Security and Research

Arthur V. C. Camargo, Leandro M. Bertholdo, Lisandro Zambenedetti Granville

[1]Instituto de Informática – Universidade Federal do Rio Grande do Sul (UFRGS)
Caixa Postal 15.064 – 91.501-970 – Porto Alegre, RS – Brazil

avccamargo@inf.ufrgs.br, berthold@penta.ufrgs.br, granville@inf.ufrgs.br

Abstract. *Cyber threat intelligence relies on network telescopes for detecting attack, and emerging threats, traditionally utilizing a substantial portion of the IPv4 address space. However, the escalating scarcity and value of this resource force universities and companies to grapple with the challenge of re-purposing their address spaces, potentially impacting cybersecurity effectiveness and hindering research efforts. In this paper we investigate the historical usage of IPv4 addressing space in network telescopes and explores the impact of reducing this space on their ability to identify attackers and collect valuable research data. Our findings reveal that even halving the allocated space for a network telescope may still permits the detection of 80% of unique cyber attack sources, and the address allocation schema have low influence in this detection.*

1. Introduction

Network telescopes, also known as "darknets" [Moore et al. 2004], capture and record unsolicited Internet traffic directed towards globally routed but unused IP address space. While network telescopes have been utilized for years, they remain essential tools for detecting and studying cyber threats and global events.

The main application of network telescopes includes monitoring and analyzing Internet traffic, helping cybersecurity experts identify new threats, attack patterns, and understand the behavior of potential attackers. They have been used for years to observe cyberattacks on an Internet scale, such as botnets [Antonakakis et al. 2017], distributed denial of service (DDoS) [Moore et al. 2004, Jonker et al. 2017], and network scan campaigns [Richter and Berger 2019, Cabana et al. 2021], providing a myriad of insights on malicious, unwanted, and unexpected behavior of cyberattacks.

Network telescopes typically collect large volumes of data. For example, a /16 sensor can generate between 100GB and 1TB of data daily. Analyzing this massive amount of data is challenging, but the use of artificial intelligence (AI) techniques is transforming this scenario [d'Andréa et al. 2023]. Advanced IA methods significantly improve one's ability to understand and interpret the data collected by those sensors, thereby providing deeper insights into emerging cyber threats.

Despite their many benefits, network telescopes also come with an inherent challenge: IPv4 address space is now a scarce and expensive resource. For example, an address space /19 (8,192 IP addresses) is rated between US$ 357,990 to US$ 395,673 [IPv4 Global 2023]. In this context, companies, research networks, and universities face

a growing pressure to release their IPv4 address space used in those sensors in favor of other uses, or even to sell or rent these addresses.

Recent studies, such as the exploration of dynamic darknets in cloud environments [Pauley et al. 2023], encounter financial challenges due to the rising costs of IPv4 address space. Beginning January 1, 2024, major cloud providers like AWS and Google Cloud will impose new charges for IPv4 usage [Huides et al. 2023] [Google Cloud 2023]. These costs present obstacles for ongoing and future telescope research in cloud platforms.

In this paper, we present a dual-fold investigation aimed at understanding the current state of network telescopes. Our first contribution involves a meticulous literature review to discern the presence and magnitude of a potential decline in the size of network telescopes. Our second contribution is an initial exploration of our network telescope to investigate the consequences of reducing its address space. Our primary objective is to identify an optimal approach for allocating the limited IPv4 address resource while preserving some of its effectiveness in threat detection and intelligence. This exploration takes into account various sampling techniques that network telescopes can employ.

Our analysis involves two network telescopes, exploring the effects of reduced IPv4 address space on traffic volume and unique source detection. The primary goal is to identify optimal strategies for allocating the finite IPv4 resource, contributing to a better understanding of network telescope dynamics and offering insights for addressing challenges posed by IPv4 scarcity in future network telescopes developments.

This paper is organized as follows: In Section 2, we provide a literature review addressing key concepts and new approaches for network telescopes. In Section 3, we detail how we assess the impact of reducing its IPv4 address space. In Section 4, we presents our findings, and in Section 5 we summarize our conclusions.

2. Background, Definitions and Related Work

A network telescope or "darknet" is a technique that allows us to capture traffic destined to unused IP address space on the Internet. Due to the unadvertised nature of those "dark" spaces, all the traffic received in this infrastructure is unsolicited and very likely malicious, with a very low number of false positives. The term "network telescope" is known by various alternative terms such as darkspace, darknet, or blackhole [Fachkha and Debbabi 2016]. Throughout this paper, we consistently use the term "network telescope" to maintain terminology harmony.

Recent advancements in data analysis, automation techniques, and artificial intelligence have leveraged the usability of large datasets generated by network telescopes. This enhancement aims to boost their effectiveness in cyber threat analysis, resulting in increased interest in network telescopes in recent years. Some research showcasing this evolution can be found in [Soro et al. 2020] and [Cabana et al. 2021].

[Soro et al. 2020] introduced community detection algorithms applied to represent network telescope activity as a graph, grouping hosts infected by a botnet that is actively scanning the network in search of vulnerable services. [Cabana et al. 2021] utilized a combination of network telescope traffic analysis and artificial intelligence to analyze reconnaissance attack campaigns against industrial control systems, allowing an automatic determination of the threat level associated with each campaign.

Network telescope sensors have various applications, including the analysis of Internet scan campaigns [Richter and Berger 2019], locating botnets [Le Malécot and Inoue 2014], observing the proliferation of Internet worms [Harder et al. 2006], and the analysis of Internet Backscattering Radiation (IBR) [Balkanli and Zincir-Heywood 2014]. These sensors are valuable for detecting and studying such threats, how they spread across the Internet and how attackers select their targets. However, [Cooke et al. 2004b] shows that the size of blocks allocated for this purpose affects its capabilities.

An alternative for expanding network telescopes size, or begin new initiatives, is IPv6. However, current challenges in effectively scanning the IPv6 address space often lead attackers to prefer IPv4. The perceive IPv6 scans from the viewpoint of IPv6 network telescopes show that scans of the IPv6 address space are infrequent and exhibit significantly different characteristics compared to the more common IPv4 scans[Richter et al. 2022]. There has been some research made by RIPE [Strowes et al. 2020] using a very large address space, although, as expected it only received very little scans compared to IPv4. These findings suggest that IPv4 network telescopes will continue to play a significant role in the foreseeable future.

Additionally, despite the existence of large, well-known global network telescopes like CAIDA [CAIDA 2024] and Merit [Merit Network 2024], many organizations face constraints in allocating addresses for threat intelligence. Deploying and maintain a network telescope is often hindered by the scarcity of unused IP address space, making it a significant entry-level barrier. Our research has noted a limited examination of the resources invested by companies and researchers in network telescopes. Thus, in Section 4.1, we review existing literature to identify utilized network telescopes in research.

Several telescopes, including CAIDA and Merit, have reduced their size over the years. One of the earliest approach to employing a smaller address space for a network telescope while preserving the benefits for cyber threat intelligence was proposed by [Harrop and Armitage 2005]. The authors suggested sparsely populating the sensors between actively used address spaces coining their approach as “greynets”,–a mixture of unused address space within specific subnets. However, this is not an option when you relinquish a part of the address space, as you lose ownership.

Following the subject of comparing or reducing of address space usage of network telescopes, there are other authors who explores that topic.

[Pemberton et al. 2007] explores sampling network telescopes and focus on the arrival density of backscatter radiation using a /16 network. The paper uses four schemes to slice its address space, horizontal, a contiguous /24 block, vertical, 1 address of each /24 and, Random-30 and Random-256, that respectively, chooses random addresses from 30 and 256 /24s respectively. Their work concludes that deploying a number of random /32 networks across the telescope is the best way to predict backscatter radiation activity.

[Chindipha et al. 2018] compares how different subnets behave on the matter of collecting IBR. The article computes the overlap of unique sources IPs across the whole network telescope. It concludes that lower /24 subnets do receive more unique origin hosts that the others, noting that 67% of the sources does only scan one IP of the sensor.

[Soro et al. 2019] compare 3 network telescopes with different sizes, /19, /15, and

three /24 to assess how the size of each sensor influence the efficiency and detection of different types of events. The objective was to assess how the size of each telescope influences the efficiency and detection of different types of events. The study presented evidence that the sources of traffic significantly vary based on the IP range and the size of the network telescope. The authors analyzed one week of data, aggregating it around autonomous systems (ASes). They demonstrated that reducing the network telescope by half minimally affects the visibility of network scans but results in different behavior in backscatter analysis when considering ASes.

Recent developments addressing the scarcity of IPv4 address space for constructing new network telescopes propose other approaches while upholding their primary goals for cyber intelligence.

The "cloud-native Internet telescope" [Pauley et al. 2023] suggests deploying short-lived telescopes on virtual machines within a cloud provider, leasing IPv4 address space, and releasing it after use. Their results indicate that optimal price performance per IPv4 is achieved in 8 minutes, and 90% of the steady-state traffic to a given IP address, compared with a regular network telescope, can be observed after only 72 minutes.

Another approach, called "meta-telescope" [Wagner et al. 2023], proposes identifying "unlikely to be used" address space in central points of the Internet (a.k.a Internet Exchanges) and capturing unsolicited traffic to this address space. In their research, they were able to capture unsolicited traffic for more than 350k /24 blocks in over 7k ASes.

Our contribution, aims to investigate the normally used range of addresses utilized in telescopes across different organizations. To the best of our knowledge, we have not found any work related to identifying how much IP resources companies and researchers have spent on network telescopes Additionally, we explored how reducing the number of addresses impacts its threat intelligence. Differently, from [Pemberton et al. 2007] and [Chindipha et al. 2018], we do not focus on collecting backscatter radiation data (TCP-ACK coming from the victims possible DDoS attacks) and are mainly worried about the number of requests and unique sources that can be gathered, which has not been done previously.

3. Methodologies

In this section, we initially review the literature to identify all known network telescopes, along with their deployment characteristics, address space usage, and relevant research results based on each infrastructure. In the second part, we investigate the impact of reducing the IPv4 address space in an existing network telescope. This section delves into the methodology employed in both cases.

3.1. Network Telescopes over time

To understand the current landscape of IP address utilization in network telescopes, we conducted a literature review. Specifically, we gathered information on the types of sensors being used for research or production purposes, aiming to present a fresh perspective on their deployment trends over time. Additionally, we intend to show the status and information gathered by the projects and its approaches.

To achieve this objective, we selected data from research papers and significant projects related to network telescopes. Given our primary focus on studying address space

usage, certain key characteristics are deemed essential for each subject under review. These include the number of IPv4 addresses utilized, the date of deployment, and the primary objectives pursued by the authors.

To make the data more accurate, we only select survey, reviews, essays, databases and papers related to network telescopes that are at least from year 2000. Non-relevant works were not selected (e.g., white paper, experimental studies, and reports lacking the information we are intending to collect) . To be included, the documents must prove to be informative and descriptive and meet one of the following criteria: (i) published by a respected organization with strong scientific endeavor or (ii) published in influential journals and conferences. In addition to those requirements, it is important that the source expresses the steps used to deploy and maintain the network telescope, as well as the results that were achieved during its activity.

To collect high-quality research and studies, we conducted a manual search for keywords related to each topic on Google Scholar. We used keywords such as "darknets," "network telescope," or "network blackhole" in our search. After reading the abstract of each article, we excluded those that were not eligible. Then, were selected relevant studies for further reading while discarding off-topic articles and papers.

3.2. Network Telescope reduction impact

To explore the implications of reducing the address space utilized in a Network Telescope, we conducted a comprehensive analysis with one month of data collected from two different sources at different times: the Japanese NICTER Darknet in 2018 [Han et al. 2022] and the Darknet-BR, a new capture of a telescope utilized by [Soro et al. 2019] in 2019 and which now gathered information from December 2023 to January 2024.

So as to investigate if there are more voluminous addresses, we inspected and compared the number of requests each address received in both telescopes, which are explained in Section 3.2.2. Our main focus was the capacity of the telescope to capture unique sources. That way, we calculated the expected value of unique sources for reduced versions of both NICTER and Darknet-BR in Section 3.2.3. Finally, in Section 3.2.4 , sampling strategies based on [Pemberton et al. 2007] approach was implemented in order to understand how different allocations affect the sensors potential.

3.2.1. Network telescope Datasets

The Network Incident Analysis Center for Tactical Emergency Response (NICTER) darknet is an integration of large-scale network monitoring for the analysis of cyber threats, such as botnets or DDoS attacks. Its dataset encompasses information from eight sensors distributed worldwide, covering networks ranging from /20 to /17. The dataset archives one month of data from October 2018 and includes only TCP SYN packets.

The Darknet-BR is a Brazilian network telescope operating on a /19 IPv4 prefix over which we have full control, enabling us to perform more in-depth analysis on the captured packets. We used a dataset from December 14, 2023, to January 14, 2024, in our analysis. Table 1 shows the period during which each dataset was collected, the daily volume of collected data, and the address space size of each sensor.

Sensor	Size	Volume	Sensor	Size	Volume
NICTER-A	/17	10GB per day	NICTER-F	/18	6GB per day
NICTER-B	/18	6GB per day	NICTER-G	/21	800MB per day
NICTER-C	/20	2GB per day	NICTER-H	/21	800MB per day
NICTER-D	/20	2GB per day	Darknet-BR	/19	3GB per day
NICTER-E	/19	3GB per day			

Table 1. Network Telescope Datasets from NICT and Darknet-BR.

3.2.2. Distribution per IP

For the purpose of further analysing the relation of address space and threat detection abilities, was made an examination of the number of requests received by each IP address in the network telescopes.

For validation purpose we apply our method in another network telescope. We select from a dataset provided by NICTER, a telescope with several IP spaces. After analysing all of its sensors and notice they presents similar behaviour, we select just the one for further comparison (sensor E). This sensor have the same size as Darknet-BR telescope. It worth to mention we just analyzed the number of scan events (TCP-SYN) and unique attacker sources in this study. We do not consider UDP or ICMP data to keep the comparison possibility–NICTER-E just provided TCP-SYN data.

3.2.3. Expected value of unique sources

In order to estimate the number of unique sources expected for allocating different sizes of network telescopes, we utilized a probabilistic approach based on the number of different destinations each source has. Considering N the set of addresses in the network telescope and S a subset of N in a way that $S \subseteq N$. Naming K as the set of source IPs, captured by the telescope and T_k the number of times that same origin address $k \in K$ appears. The probability of k being observed by the smaller version S considering an uniform distribution is given by P_k in Equation (1).

In addition, we grouped the number of IPs k that target the same number of T_k together as $G_t k$, as they provide roughly the same information for our model. That way, it is possible to deduct the expected number of unique sources that the reduced version of the sensor will capture utilizing the formula of expected value—see $E[S]$ in Equation (2).

$$P_k = 1 - \frac{\binom{|N|-T_k}{|S|}}{\binom{|N|}{|S|}} \tag{1}$$

$$E[S] = \sum_{k}^{k \in K} P_k \times G_t k \tag{2}$$

As our work revolves around gathering unique sources, the complement probability of not perceiving any attack is utilized. That way, the only parameter related to the reduced network telescope is its size $|S|$, meaning that the formula is independent on the individual addresses being picked in the smaller version.

3.2.4. Sampling strategies

One way that organizations have to mitigate the problems of IPv4 exhaustion and scarcity is to reduce their telescopes. To do that, it faces the dilemma of how to divide the blocks. Said that, using the Darknet-BR, we considered two possible alternatives: (1) a reduction from a /19 (8,192 IPs) to a contiguous /20 (4,096 IPs), evaluating if there is differences between the first and second /20, or; (2) reducing it to several /24s by adopting a sample strategy proposed by [Pemberton et al. 2007]. While the first solution (1) is the simpler one, we consider to study the sampling solution (2) as a possibility to minimize the network telescope lost potential.

For the first study, we consider to analyze just the horizontal sampling from [Pemberton et al. 2007], since vertical sampling we consider operationally unfeasible– when delegating a prefix we lose control over the that range. In horizontal sampling, we select half of the /24 blocks for the telescope, while the others will be assigned for users. As for the second study, we considered selecting alternating /24 blocks, in a way that they are equally spaced, trying to cover a larger area of addresses. Complementary, we observed individual blocks at specific locations within the network telescope, such as the beginning, end, and middle, exploring potential correlations with the findings of [Harrop and Armitage 2005], and [Chindipha et al. 2018].

4. Results

In this section, we present our results from research on identifying network telescope initiatives over the years and our investigation into the impact of reducing the address space of a currently operational network telescope.

4.1. Address space utilization on Network Telescopes over time

After conducting our paper review on network telescopes (see Section 3.1), we summarize the main telescopes initiatives we identified in Table 2. It is important to note that network telescopes in the cybersecurity industry typically do not publish their data, address space size, or even their existence to safeguard the secrecy of their initiatives. Consequently, we could not gather much information about those environments.

Additionally, most network telescope initiatives referred in research also does not disclose the address space they use. They justify this approach to avoid *adversary traffic*—when an attacker avoids scanning or using the network telescope address to avoid being identified. We have listed the address spaces that we could verify.

Table 2. Summary of Network Telescope projects referred from 2000 to present

IPv4 Addr.	Year	Name	Comments
50,331,648	2010	APNIC/ARIN	APNIC and ARIN collaborated on IBR research utilizing unallocated addresses 1/8, 50/8, and 107/8. This telescope had a lifespan of 1 week in 2006.
17,048,576	2001	Internet Motion Sensor	Arbor Networks and the University of Michigan project deploys sensors in diverse locations to enhance the diversity, sparsity, and size of a Network Telescope. The IMS initiative seems ending in 2004 and spanning into Merit Telescope.

Continued on next page

Table 2 continued from previous page

IPv4 Addr.	Year	Name	Comments
16,777,216	2005	MERIT	Merit Network Telescope used the 35/8 address from 2005 to 2018. After this date the Michigan University formalized the Orion telescope with a smallest address space.
16,777,216	2001	UCSD-CAIDA	The UCSD Network Telescope, a project from the University of San Diego/US was built on the globally routed 44/8 prefix (former AMPRNet) from 2001 to 2019.
12,582,912	2019	UCSD-CAIDA	The UCSD Network Telescope reduced it size from a /8 to a /9 and /10 network.
∼2,000,000	2012	SWITCH	Collect data from the address space from multiple networks across Switzerland .
626,944	2004	Team Cymru	Multiple sensors deployed by the company Team Cymru.
524,288	2014	Farsight	Farsight's Network Telescope, now part of DomainTools, offers data through subscription.
475,136	2018	ORION-MERIT	Michigan State University's project, known as the Observatory for Cyber-Risk Insights and Outages of Networks, focuses on Internet backscatter radiation. Designed and engineered with support from the US-NSF, it consists of 1,856 /24 subnets.
270,000	2005	NICTER	The Japanese organization NICTER (Network Incident Analysis Center for Tactical Emergency Response) integrates its network telescope for large-scale monitoring and analysis of cyber threats, including botnets and DDoS.
178,000	2018	MIT-Akamai	The first network telescope built over a Content Delivery Network (CDN) infrastructure. It is composed of two IPs on each of the 89,000 Akamai servers across the globe.
131,072	2018	NL-Darknet	Network telescope maintained by SurfNET in the Netherlands.
65,636	2019	HEAnet	Ireland's National Education and Research Network Telescope.
65,536	2004	IUCC/IDC Telescope	The Israel InterUniversity Computation Center (IUCC) Network Telescope.
65,536	2006	Anonymous	The University of Wellington, NZ, utilizes an undisclosed /16 network telescope to test various address sampling strategies for measuring arrival density.
65,536	2021	Anonymous	An undisclosed enterprise network telescope identifies a specific stateless-scanning malware, and a response is forged to slow down the malware's propagation, deceiving botnet scanners. The research is being conducted in Germany.
8,192	2018	BR-Darknet	A /19 network telescope in Brazil (used in this paper).
4,096	2017	JP-Darknet	Another /20 network telescope hosted in Japan.
4,096	2010	INRIA	French Telescope at INRIA's High Security Laboratory.
765	2017	IT-Darknet	Italian network telescope
512	2009	Rhodes University	The first known network telescope in the Afrinic Region.
512	2013	KISTI	Science and Technology Security Center South Korea (KISTI), does provide 2 sensors with a /24 mask.
256	2006	–	After 2006, numerous minor initiatives deployed temporary network telescopes with short lifespans (1-2 years), ranging from /28 to /24, for specific research.

From the paper review we observed that the majority of significant network telescopes emerged between 2000 and 2007, a period characterized by fewer issues related to IPv4 allocation. The onset of IPv4 address exhaustion was first announced by the Regional Internet Registry (RIR) in Asia in 2011, followed by announcements from other RIRs in subsequent years [APNIC 2024, RIPE 2024, LACNIC 2024, AFRINIC 2024].

The IPv4 exhaustion resulted in the absence of new relevant network telescope initiatives and even a reduction in existing network telescopes in recent years. For instance, UCSD-NT/CAIDA, which is part of the US Protected Repository for the Defense of Infrastructure against Cyber Threats (PREDICT) program and its successor, the Information Marketplace for Policy and Analysis of Cyber-risk & Trust (IMPACT) [CAIDA 2018], saw a reduction in its IPv4 address space from a /8 to a /9 plus /10 in 2019.

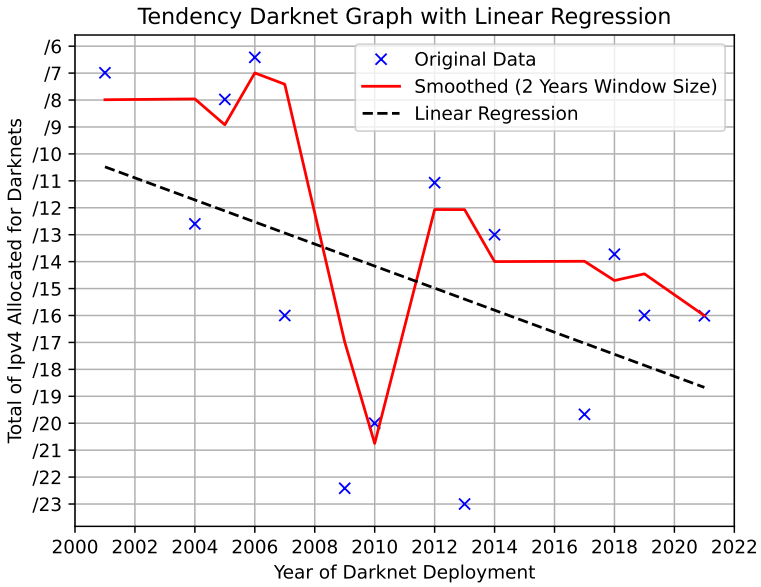


Figure 1. Tendency graph of IPv4 usage on Network Telescopes in each year, considering the first deployment of each one. Here, we omitted data from initiatives who reduced the size of their telescope (i.e., UCSD and MERIT).

In Fig. 1, we can better visualize the decline in the utilization of public IPv4 addresses for network telescopes over the years. The figure is built from a more comprehensive list of publications we investigated, encompassing 28 network telescopes (blue crosses). From this graph, is reasonable to infer an initial reduction around 2010, correlated with the depletion of IPv4 address in RIRs. The second reduction, around 2021 may be linked to the escalating prices of IPv4 in the market. Notably, during this period, portions of addresses from large network telescopes shifted to major companies such as Google and Amazon.

The reduction trend become more evident when we visualize the smoothed tendency over a two-year time window (red line). A linear regression (dashed back line) for this tendency also point to a possible deallocation of IPv4 address space for Network Telescopes after 2018– as observed in the case of UCSD-NT/CAIDA.

Given the difficulty of maintaining this address space active, new research initiatives such as *meta-telescope* [Wagner et al. 2023] and *DScope*[Pauley et al. 2023] aim to explore new ways to deploy temporary telescopes.

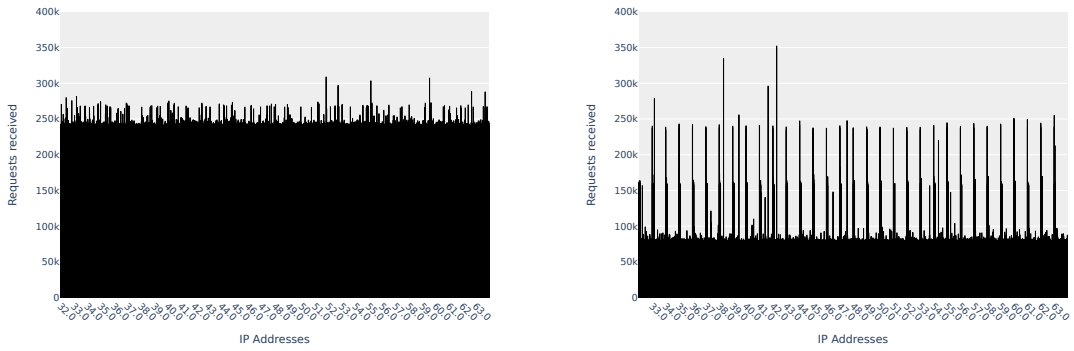
4.2. Analysis of IPv4 Address Space Reduction in Network Telescopes

As described in Section 3.2, we then explored the datasets and analyzed possible strategies to overcome problems related to the scarcity of IPv4 address space. Initially, we explored

the distribution of the requests per destination IP in all the datasets we listed.

In Fig. 2 we can visualize the distribution of requests received per IP address in both telescopes in a period of 31 days (one month). Here we depict Darknet-BR for the period of Dec/2023 and NICTER-E for Oct/2018. It's noteworthy that other NICTER sensors have shown a similar behaviour when compared with NICTER-E. In all datasets, the telnet scans (TCP/23) were the primary target; however, they were more prominent in 2018.

In DarknetBR (Fig. 2a), we received way more requests overall. It's probable that the chronological factor is the main reason–more scans and malicious activities in 2023 than in 2018. Other point that is worth to mention is that NICTER-E telescope show a discrepancy in the first 25 addresses of each /24 block, indicating that most telnet scans target the first IPs in each /24 prefix, for example the IPs ending on .1 or .2. There were some cases in the literature [Cooke et al. 2004a], [Chindipha et al. 2018] showing this increase of requests in lower IP addresses at the beginning of the /24 block, commonly allocated for gateways (*e.g.,* routers).



(a) Number of requests per IP, Darknet-BR. **(b) Number of requests per IP, NICTER-E.**

Figure 2. Distribution of received scans (TCP-SYN) per IP address in two different network telescopes. Both using a /19 block in one month.

Applying the method described in Section 3.2.3, we calculated the expected value of unique sources for all sizes of S for both sensors (Darknet-BR and NICTER-E). Fig. 3 and Fig. 4 show the percentage of the unique sources that would be visible when we project a reduction in the number of addresses being utilized by the telescope. The figure illustrate the projection of the percentage from the original telescope that we can reach (Y-axis) by the number of hosts needed (X-axis). This estimation allows us to assess how the reduction will impact the capture of unique sources by providing information on the origins already captured and the desired size of the new telescope.

When we compare the projected results for both telescopes, some similarities become evident. Both telescopes expect to capture more than 80% of unique sources when the address space is reduced by half, and more than 60% when reduced to a quarter. Our main observation here is that most of the identified scanned sources tend to scan several addresses in both telescopes.

Another observation is related to the incline of the curve in the two graphs. In this regard, NICTER-E shows slower growth in the beginning. This result is attributed

to different attack methods, with NICTER-E registering most scans at lower addresses in each /24 (scans for routers). The takeaway here is that reducing the address space too much may impact in our ability to detect certain types or methods of scans.

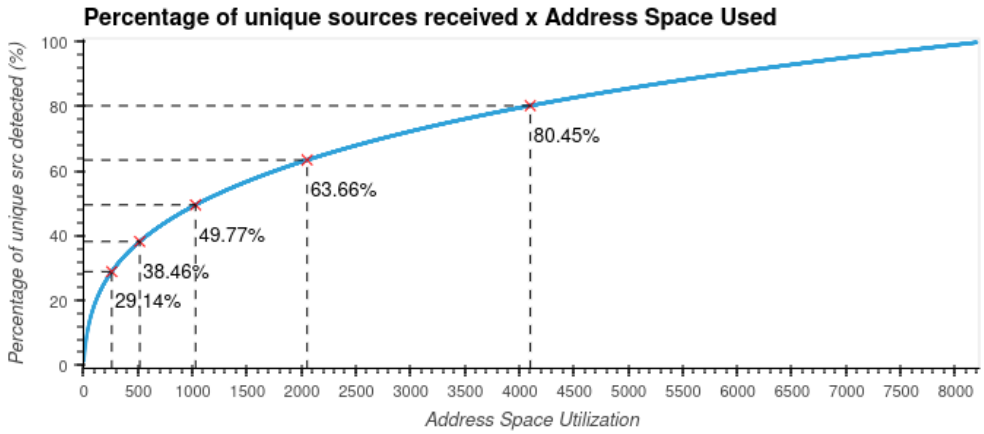


Figure 3. Darknet-BR /19 Percentage of unique sources received per IP used

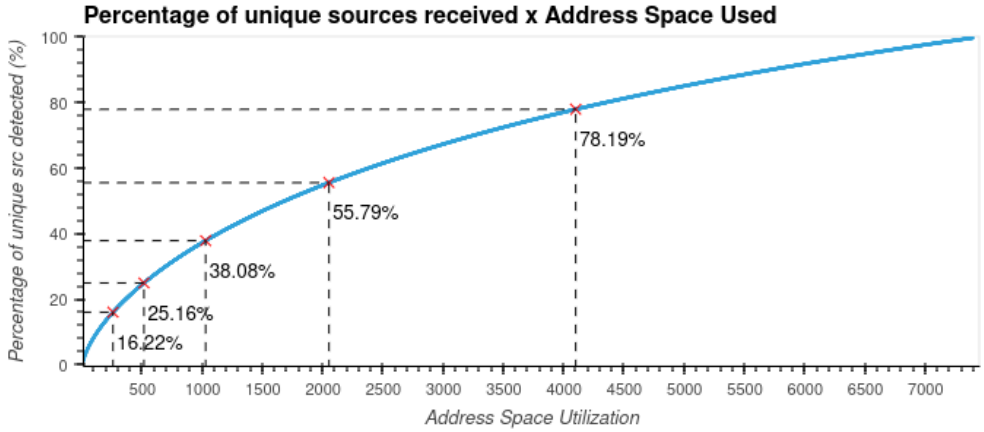


Figure 4. NICTER-E /19 Percentage of unique sources received per IP used

Additionally to the computation of the estimations used in Section 3.2.3, we also relied on different sampling approaches to make those results more concrete. Here, we arranged the Darknet-BR dataset in 4 different subnets allocation schemas (e.g., one /20 or sets of /24), and gathered the real number of unique sources that would be perceived in each context, as described in Section 3.2.4.

In Table 3, we present the results of reducing the size of Darknet-BR by half, i.e., from 8k addresses to 4k addresses. The figure illustrates the impact on unique sources and the number of requests observed for each applied sampling method.

Our results from testing four different allocation methods over our Darknet-BR dataset do not show a significant difference, less than 0.1% for identifying unique scan sources or requests. This occurs because of the uniform distribution present in the requests per IP addresses, meaning we did not find any particular set of IPs being more targeted than others.

Table 3. Number of unique sources and requests seem by different methods.

Method	Unique Sources (%)	Number requests (%)
Total	100.00	100.00
Low /20	80.30	50.03
High /20	80.26	49.97
Even /24 allocation	80.26	50.01
Odd /24 allocation	80.39	49.99

As to understand more about this behaviour, we also examined some individual /24 blocks in the beginning, end and in the middle of the /20 with the objective to assert that there are not a specific block that is significantly different from the others. Although the first block does show more unique sources, it just sees 28.35% while the worst /24 observed 28.11% of the total number of sources, a minimal difference of only 0.24%.

5. Conclusions

In this work, we review the literature to gather information about all the networks telescopes that has been deployed in the last 23 years. We identified just 28 distinct initiatives, where we believe 18 are still active today. The reduction in the IPv4 address space for network telescopes is a trend, with the larger telescopes yielding part of their space for other organizations.

A second contribution of our work is the analysis of two network telescopes with the objective of better understanding the effects of reducing their address space. We achieve this by estimating the number of unique sources each one would receive in the case of reduction using an expected value formula and applying sampling techniques. Our findings indicate that reducing a /19 telescope in half will still maintain more than 80% of its total number of unique sources detected, although it would perceive only 50% of the number of requests. Additionally, the addressing schema adopted in the reduction, such as splitting into /24 or just selecting one /20, has a low influence of less than 0.1%.

References

AFRINIC (2024). Afrinic ipv4 exhaustion statistics. `https://stats.afrinic.net/ipv4/exhaustion/ipv4_available`. (Accessed on 2024/04/16).

Antonakakis, M., April, T., Bailey, M., Bernhard, M., Bursztein, E., Cochran, J., Durumeric, Z., Halderman, J. A., Invernizzi, L., Kallitsis, M., et al. (2017). Understanding the mirai botnet. In *26th USENIX security symposium (USENIX Security 17)*, pages 1093–1110.

APNIC (2024). Apic ipv4 exhaustion. `https://www.apnic.net/manage-ip/ipv4-exhaustion/`. (Accessed on 2024/04/16).

Balkanli, E. and Zincir-Heywood, A. N. (2014). On the analysis of backscatter traffic. In *39th Annual IEEE Conference on Local Computer Networks Workshops*, pages 671–678. IEEE.

Cabana, O., Youssef, A. M., Debbabi, M., Lebel, B., Kassouf, M., Atallah, R., and Agba, B. L. (2021). Threat intelligence generation using network telescope data for industrial control systems. *IEEE Transactions on Information Forensics and Security*.

CAIDA (2018). Supporting research and development of security technologies through network and security data collection. `https://apps.dtic.mil/sti/trecms/pdf/AD1054333.pdf`. (Accessed on 01/21/2024).

CAIDA (2024). Historical and Near-Real-Time UCSD Network Telescope Traffic Dataset. `https://www.caida.org/catalog/datasets/telescope-near-real-time_dataset`. Accessed on 2024/04/16.

Chindipha, S. D., Irwin, B., and Herbert, A. (2018). Effectiveness of sampling a small sized network telescope in internet background radiation data collection. In *Southern Africa Telecommunication Networks and Applications Conference (SATNAC)*.

Cooke, E., Bailey, M., Mao, Z. M., Watson, D., Jahanian, F., and McPherson, D. (2004a). Toward understanding distributed blackhole placement. In *Proceedings of the 2004 ACM workshop on Rapid malcode*, pages 54–64.

Cooke, E., Bailey, M., Watson, D., Jahanian, F., and Nazario, J. (2004b). The internet motion sensor: A distributed global scoped internet threat monitoring system. *Technical Report CSE-TR-491-04*.

d'Andréa, E., François, J., Festor, O., and Zakroum, M. (2023). Multi-label classification of hosts observed through a darknet. In *NOMS 2023-2023 IEEE/IFIP Network Operations and Management Symposium*, pages 1–6. IEEE.

Fachkha, C. and Debbabi, M. (2016). Darknet as a source of cyber intelligence: Survey, taxonomy, and characterization. *IEEE Communications Surveys & Tutorials*.

Google Cloud (2023). Google cloud virtual private cloud (vpc) pricing. `https://cloud.google.com/vpc/pricing-announce-external-ips?hl=pt-br`. Accessed: 2024/04/16.

Han, C., Takeuchi, J., Takahashi, T., and Inoue, D. (2022). Dark-tracer: Early detection framework for malware activity based on anomalous spatiotemporal patterns. *IEEE Access*, 10:13038–13058.

Harder, U., Johnson, M. W., Bradley, J. T., and Knottenbelt, W. J. (2006). Observing internet worm and virus attacks with a small network telescope. *Electronic Notes in Theoretical Computer Science*, 151(3):47–59.

Harrop, W. and Armitage, G. (2005). Defining and evaluating greynets (sparse darknets). In *The IEEE Conference on Local Computer Networks 30th Anniversary (LCN'05) l*, pages 344–350. IEEE.

Huides, A., Santhanam, A., and Lehwess, M. (2023). Identify and optimize public ipv4 address usage on aws — networking & content delivery. `https://aws.amazon.com/blogs/networking-and-content-delivery/identify-and-optimize-public-ipv4-address-usage-on-aws/`. Accessed: 2024/04/16.

IPv4 Global (2023). November 2023 sales report. `https://ipv4.global/reports/november-2023-sales-report/`. Accessed: 2024/04/16.

Jonker, M., King, A., Krupp, J., Rossow, C., Sperotto, A., and Dainotti, A. (2017). Millions of targets under attack: a macroscopic characterization of the dos ecosystem. In *Proceedings of the 2017 Internet Measurement Conference*, IMC '17, page 100–113, New York, NY, USA. Association for Computing Machinery.

LACNIC (2024). Estadísticas de asignación de lacnic. `https://www.lacnic.net/999/1/lacnic/`. (Accessed on 2024/04/16).

Le Malécot, E. and Inoue, D. (2014). The carna botnet through the lens of a network telescope. In *Foundations and Practice of Security: 6th International Symposium, FPS 2013, La Rochelle, France, October 21-22, 2013, Revised Selected Papers*, pages 426–441. Springer.

Merit Network (2024). Orion Network Telescope – Merit. `https://www.merit.edu/initiatives/orion-network-telescope/`. Accessed: 2024/04/16.

Moore, D., Shannon, C., Voelker, G. M., and Savage, S. (2004). Network telescopes: Technical report.

Pauley, E., Barford, P., and McDaniel, P. (2023). DScope: A Cloud-Native internet telescope. In *32nd USENIX Security Symposium (USENIX Security 23)*, pages 5989–6006, Anaheim, CA. USENIX Association.

Pemberton, D., Komisarczuk, P., and Welch, I. (2007). Internet background radiation arrival density and network telescope sampling strategies. In *2007 Australasian Telecommunication Networks and Applications Conference*, pages 246–252.

Richter, P. and Berger, A. (2019). Scanning the scanners: Sensing the internet from a massively distributed network telescope. In *Proceedings of the Internet Measurement Conference*, pages 144–157.

Richter, P., Gasser, O., and Berger, A. (2022). Illuminating large-scale ipv6 scanning in the internet. In *Proceedings of the 22nd ACM Internet Measurement Conference*, IMC '22, page 410–418, New York, NY, USA. Association for Computing Machinery.

RIPE (2024). What is ipv4 run out? `https://www.ripe.net/manage-ips-and-asns/ipv4/ipv4-run-out/`. (Accessed on 2024/04/16).

Soro, F., Allegretta, M., Mellia, M., Drago, I., and Bertholdo, L. M. (2020). Sensing the noise: Uncovering communities in darknet traffic. In *2020 Mediterranean Communication and Computer Networking Conference (MedComNet)*, pages 1–8.

Soro, F., Drago, I., Trevisan, M., Mellia, M., Ceron, J., and J. Santanna, J. (2019). Are darknets all the same? on darknet visibility for security monitoring. In *2019 IEEE International Symposium on Local and Metropolitan Area Networks (LANMAN)*, pages 1–6. ISSN: 1944-0375.

Strowes, S. D., Aben, E., Wilhelm, R., Obser, F., Stagni, R., and Formoso, A. (2020). Debogonising 2a10::/12: Analysis of one week's visibility of a new/12. In *TMA*.

Wagner, D., Ranadive, S. A., Griffioen, H., Kallitsis, M., Dainotti, A., Smaragdakis, G., and Feldmann, A. (2023). How to operate a meta-telescope in your spare time. In *Proceedings of the 2023 ACM on Internet Measurement Conference*, IMC '23, page 328–343, New York, NY, USA. Association for Computing Machinery.