2019 11th International Conference on Cyber Conflict:
Silent Battle
T. Minárik, S. Alatalu, S. Biondi,
M. Signoretti, I. Tolga, G. Visky (Eds.)
2019 © NATO CCD COE Publications, Tallinn

Towards Measuring Global DDoS Attack Capacity

Artūrs Lavrenovs
NATO CCD COE
arturs.lavrenovs@ccdcoe.org

Abstract: In today's Internet, distributed denial-of-service (DDoS) attacks play an ever-increasing role and constitute a risk to any commercial, military or governmental entity that has a presence on the Internet or simply has an Internet connection. To address this threat on all levels, decision-makers have to rely on trustworthy information regarding attack capacity, sources, and the largest contributors. The lack of this information limits the ability of technicians, policymakers, and other relevant decision-makers to remediate the issue as efficiently as possible.

This research introduces a methodology for measuring the properties of individual devices participating in such attacks. These properties include rate limiting, amplification factor, and speed, which allows the calculation of each device's actual contribution to the attack capacity. This methodology was implemented as a proof of concept for the NTP protocol and the results indicate that it has promising potential. Individual measurements aggregated together provide insights into particular abused protocols: all the protocols together could provide the global DDoS attack capacity.

Keywords: *DDoS, attack capacity measurement, global DDoS attack capacity*

1. INTRODUCTION

Distributed denial of service (DDoS) attacks have been plaguing the Internet almost since its inception. Although the first large-scale network DDoS attack happened in 1999 [1], DDoS is still a serious and even growing threat to Internet-connected organizations. DDoS attacks have become almost daily news and created a large cybercrime industry offering DDoS attacks as a service as well as an immense cyber defense industry providing network filtering services, software, and hardware. Reasonable observers without any computer networking or cybersecurity background would assume that this issue has been and is currently being addressed properly to eliminate it at its root cause. The reality is that DDoS attacks have been on the rise with the increase in Internet connection speeds but mitigation efforts have only slowed down the total growth of the attacks.

DDoS attacks have become almost a household word, because when an online gaming platform or other popular resource goes offline because of continuous attacks, tens of millions of users are affected. Cybersecurity and network specialists are well aware of the attack properties, and many of the decision-makers are aware of the risks. Why, then, has this problem not been tackled in a global or at least a national manner? Due to the nature of the Internet, DDoS attacks transcend national borders and although they are illegal, there are no technical means to stop them at a national level. There must be a push for international policy from the highest-level decision-makers. These efforts cannot be made without being information-driven. The underlying causes are well-known but the question that is missing an answer is, what is the current status: total attack capacity, geographical regions and legal entities contributing the most? This information, presented in an easily digestible way together with associated risks, should be useful for non-technical decision-makers to justify taking action.

The kind of data needed and how to acquire it is investigated and the methodology for producing the missing information is proposed in this paper, resulting in the development of a proof of concept for the NTP protocol.

2. RELATED WORK

DDoS attacks are widely discussed and researched in academia. Although raw data is significantly less available to researchers than to commercial and other entities that receive DDoS attacks themselves, in some cases researchers make special agreements to access it. Attack detection and defense is a significantly explored topic; in the real world these solutions most often involve basic statistical analysis of incoming traffic. However, researchers are trying a wide range of old and new technologies like machine learning, software-defined networks, etc. to achieve better results. Less relevant research topics such as motivation, financial and criminal aspects are not reviewed here.

Protocol analysis is the default method for identifying protocols that could be abused in the future. Analysis of the protocol definition, documentation and source code of different implementations can allow researchers to identify new potentially abusable services. Some assumptions or previous research into the prevalence of the analyzed protocol must be made in advance to choose which of the many protocols to pick for analysis. If only a few devices with abusable services are found, then the overall impact for DDoS attack is negligible and malicious actors might not even bother exploiting it. Correct responsible disclosure mandates security researchers to report discovered vulnerabilities in advance to hardware and software vendors and other parties that would be responsible for issue remediation. In theory, this would preemptively mitigate the abuse of specific protocols, but the real situation is quite different. Research publications and vulnerability reports disclosed after the time period given to vendors still enable malicious actors to exploit reflection from devices that are presently not mitigated. One of the most prominent such cases was NTP DDoS. Rossow evaluated common UDP-based protocols and observed that most NTP implementations support a command to return client list that was a feature of the implementation and not defined in the protocol itself. The measured amplification factor was up to 4670, which was the largest of those measured in the research [2]. Because of the potential for abuse, the researcher conducted responsible disclosure to the security community and appropriate vendors. However, either directly due to this disclosure or inferred through released software fixes, malicious actors started exploiting it in the wild.

Attack analysis covers newly abused protocols or protocol features combining data from attack monitoring points, Internet scan data, backscatter, and other sources that present an integrated overview of the specific protocol. The success of mitigation efforts can be evaluated by notifying the system owners and continuously monitoring changes in the number of abusable devices. Czyz et al. investigated NTP DDoS in detail, additionally exploring unique protocol features that provide insight into victims [3]. This type of research does not contribute to overall DDoS attack capacity knowledge, as the produced estimates are for a fixed point in the past, possibly at the peak of the attacks, and quickly become outdated.

The ability to spoof the IP address of packets is the main cause of multiple types of attacks, including the most problematic: reflected DDoS. The Center for Applied Internet Data Analysis (CAIDA), based at the University of California's San Diego Supercomputer Center, has been conducting research into the state of IP spoofing and continuously monitoring since 2008. The CAIDA spoofer project publishes updated and historical data from their measurements. In total, 22.6% of the IPv4 AS not using NAT were spoofable in July 2018, which corresponded to 14.3% of IP address blocks [4]. In general, countries in developing regions are found to be proportionally more spoofable than those in developed countries. However, in absolute numbers, the USA has most of the spoofable IP blocks.

A. Capacity Measurement

Measuring attack capacity is not sufficiently investigated. Currently, the only methodology for measuring the overall worldwide capacity for DDoS attacks is published in the scientific literature by Leverett et al. [5]. Researchers analyzed only reflected volumetric UDP DDoS attacks, thus closely relating to this research. More specifically, four protocols were analyzed – NTP, DNS, SSDP, and SNMP. Using this methodology, it was concluded that the total estimated DDoS attack capacity is 108.49 Tb/s. As the authors acknowledged themselves, this figure is limited by factors not explored in detail; thus, in reality, it is significantly lower. This figure does not take into account the ability of the AS network to handle all the capacity at same time, device load, existing bandwidth utilization, and device computational power that might not be able to handle producing responses to fully utilize the whole available network connection.

In addition to the total capacity estimate, additional avenues to present and visualize data for easier consumption by non-technical policy-makers were explored, e.g., a map of the world with the risk posed to others attached to each individual country. This visualization allowed the important discovery that developed countries actually possess higher DDoS attack capacity than developing countries. This finding points to the lack of a policy to, at the very least, mitigate DDoS attacks or its enforcement even in developed countries. Instead of pointing the finger at developing countries, this issue should be addressed internally and at an international level.

B. Industry Research

Case studies analyzing individual attacks are occasionally published online by commercial entities receiving or mitigating DDoS attacks. This usually happens when a new protocol has started getting abused or when previous attack records are broken. The motivation behind these case studies is to advertise the ability to handle DDoS attacks to gain more clients, and the details provided in the case studies are usually very restricted so as not to reveal any commercial information or weak points in the defenses. However, these case studies have become the main point of reference when discussing DDoS attack capacity. When the question is, ‘What is the maximum realistic DDoS attack capacity?’ the answer that follows usually refers to the latest or recent published attack case study. At the time of writing this paper, the case study by Arbor reported a maximum observed DDoS attack capacity of 1.7 Tbps caused by abusing Memcache [6].

Whenever a new service gets abused for DDoS attacks, a new scanning project presenting the results publicly is usually created. The creators of these projects are organizations and individuals working in networking or cybersecurity fields who are affected by the DDoS attacks but frequently prefer to remain anonymous. The main purpose of such projects is to advise the public in general and network owners that their networks contain systems that can be abused. It can be done by either emailing a notification message to network abuse addresses, notifying only persons who have signed up their network ranges or enabled the conduct of a network-range search in their database. The goal of these projects is to minimize the number of abusable devices as much as and as quickly as possible. Sometimes, these projects cooperate with researchers from academia by providing them with raw data, so that research can concentrate on data analysis instead of technical data gathering. On its own, this research is usually limited to scanning the Internet for all the devices using specific ports and protocols, grouping the results by AS and geographical attributes and presenting them in table and graph formats. If scans are repeated, then comparisons can be made between timespans and device count decline tendency can be identified. If scans are scheduled periodically, then the current situation can be ascertained. Many open ports exposed to the Internet are being scanned by The Shadowserver Foundation, which also includes more than 10 that are most commonly used for amplified reflected DDoS attacks [7]. The Open NTP and Open Resolver projects focus on a single protocol while CyberGreen goes the furthest by calculating and assigning risks.

From the opposite side, scanning activities can be detected and presented in real time and as historical data. One of these projects is NetworkScan Mon, which aggregates data by the source and destination attributes of IP packets and presents aggregated statistics which show that in July of 2018, there was not a single protocol abusable for DDoS attacks among the top 10 ports receiving scanning activities [8]. This indicates that DDoS is a specialized niche of cybercrime and because of the required 2-pronged execution, it is less attractive to cyber criminals as opposed to most popular scanned ports which are used by services that can be directly exploited.

It is possible to monitor DDoS attacks and extract some of the attack attributes by either passively monitoring network traffic at Internet Exchange points or maintaining a distributed set of honeypots that pretend to be exploitable network services. The DDoS Mon project provides insight into worldwide DDoS attack statistics and historical trends; in July of 2018 it reported an average of about 20,000 attacked IP addresses per day [9]. Attacked IPs do not necessarily equate to a single attack or target as systems under attack can have multiple IP addresses. However, no deeper analysis into grouping separate IP addresses into a single target was provided; hence, there is potential for separate research. In the same time period, the USA and China were the most attacked countries, HTTP port 80 and HTTPS port 443 were the most targeted ports, and websites using a .com top-level domain were targeted most often. Amplification and reflection-based attacks were most common, amounting to nearly 70% of the DDoS attacks by frequency, while the most commonly abused protocols were CLDAP, NTP, and DNS. These attack statistics have drawbacks because the number of some specific abused protocol services does not reflect their overall bandwidth contribution to the attack, which is the main property of DDoS attacks.

3. MEASURING DDOS CAPACITY

Most types of DDoS attacks have effective remedies available but volumetric DDoS attacks can exhaust the resources of the whole targeted network, thus affecting all the connected services. Specifically, Reflected Amplified Volumetric DDoS attacks are the most problematic type and the proposed methodology covers only this type of attack. To mitigate these attacks, the defender must absorb and process all the received network traffic by separating legitimate packets from the attack packets. The bandwidth capacity of the attacked network is limited, not only by contractual relations between the ISP and the attacked network but also by the chosen technology and network hardware.

There are two main causes of these types of attacks – the ability to spoof IP addresses, and network services that use the UDP protocol and can produce responses significantly larger than the received requests. Volumetric attacks generate higher bandwidth than attacked networks can process. These attacks are indirect and attacking traffic is produced by unsuspecting devices running abusable network services that generate significantly larger responses than requests. To measure DDoS capacity, these devices need to be identified and their properties extracted and measured to produce the whole picture.

A. Identifying Devices

To estimate the current status of attack capacity, it is sufficient to investigate only publicly known protocols that are being abused. A whole Internet scan should provide the set of abusable devices for the attacks. Depending on the protocol and implementation choices, the scan can be either a generic protocol request or abusable functionality itself. There are differences in the information that can be extracted from this data depending on the approach, e.g., if a scan is conducted using a generic request, then a ratio of abusable to all protocol-implementing devices can be established, which might be useful. If the generic request is not the same as that abused for the attacks, an additional checking request is required before conducting further measurement. At the end of this stage, a set of only abusable devices should be produced.

B. Detecting Rate Limiting

Attackers abusing network services rely on the fact that they do not have any rate limiting. Academic and industry research usually stops at identifying the devices or estimating amplification; there is no published research regarding real-world rate limiting among the identified potentially abusable devices. Technically, rate limiting can either be explicit or implicit. The former is preset in the service's software configuration file or hardcoded in the source code, while the latter is caused by OS, hardware, or network limitations.

Technically, rate limit measurement can be implemented in two ways – sending a burst of packets and verifying the count of received packets or by analyzing every pair of response and request packet sets. Because the measurement requests are the same in most protocols, it should produce exactly the same or a very similar response packet count-wise. By using packet count from the amplification measurement step, it is possible to divide the number of received packets with a packet amplification factor to determine if the resulting value is close enough to the number of sent requests. If it is, then there is no rate limiting or it is above the selected threshold; otherwise, the resulting value approximately corresponds to the rate limit.

More precisely, rate limiting can be measured when mapping sets of response packets to each appropriate request. To some extent, it allows the differentiation of packet loss from rate limiting; as rate limiting is implemented on a per response basis, it might allow identifying exactly from which request responses stopped coming. This method is also suitable for measuring rate limiting that is not on a per second basis by detecting at what request number responses stop and at what number they restart. This type of measurement can technically be implemented in two ways. The easiest way is that every request uses a different source port number, thus every response packet set will be received by a different port. However, in DDoS attacks, all the reflected packets usually target a single port. The harder way is to use the same port and try to differentiate between responses, but depending on the tested protocol, this might be unfeasible because all the sent requests must be the same. Different protocols might possibly produce better data using different measurement methods; from a methodological perspective, it does not matter which approach is implemented as long as advantages and disadvantages are considered for every tested protocol implementation. For the proof of concept, every measurement request expects a response to different incrementing port numbers while enforcing appropriate timeouts to maintain request and response matching.

C. Estimating Network Speed

The network speed of individual devices is one of the main pieces of information lacking in attack capacity estimates. The easiest solution is to use country or specific ISP average upload data gathered by research organizations, but the issue is that abusable devices are a small part of the networks and might not be representative in terms of speed.

An important question is: can the speed of individual devices be estimated from timestamps in the current measuring methodology? If for the start and end time the minimum and maximum values are selected, then a single delayed packet skews the calculation significantly. Speed can be calculated by adding up all the received protocol payload sizes and 42 bytes as transmission overhead for each received packet and dividing it by the time difference between the last and first packet; responses with one packet cannot be processed this way and should be ignored. Although speed calculated in this manner might not necessarily correspond to the speed of the network connection for the device, it could still be a good metric. The device might not be able to fill all the bandwidth capacity available to it using a specific measured protocol. The bandwidth could have been in use in other ways at the time of measurement or the speed could have decreased over a long distance.

D. Technical Concerns

There are significant technical concerns that might affect the quality of measurement and overall viability of the proposed methodology. The measured devices might be participating at the time of measurement in real attacks, thus the measurement would not accurately reflect their capacity. If attackers are measuring themselves and selecting specific most powerful abusable devices, then the total results might get significantly skewed.

Measurement traffic looks exactly like DDoS traffic because the types of requests and responses are the same as those used by attackers. In the real world, receiving or transit networks cannot judge if the request traffic is spoofed or legitimate. Thus, the way to mitigate DDoS to an extent is to block this traffic. We have observed measurement interference from automated solutions deployed across transit networks.

The location of the measurement server both geographically and in the network affects the data. The further away the measured device is, the higher the probability of mitigation solution interference, packet loss, and delays affecting the calculation of its contribution. The same time measurement from a single point produces a view from the perspective of a single specific victim.

E. Estimating Total Attack Capacity

It might be tempting to sum up all the abused protocol measured capacity values together to produce a single value of total worldwide DDoS attack capacity. In reality, there are two major and a wide range of minor factors that limit the attack capacity.

Every network has a limited upload bandwidth capacity that is available for outgoing DDoS attack traffic. A specific network's connection capacity is directly affected by the physical technology in use, router capability, free unused capacity of the uplink and contractual agreement with the ISP or transit provider. The issue is that it is not clear where to draw the border for every network and what the capacity of every network actually is. The easiest solution would be splitting the Internet by AS and using open information from IX monitoring projects and estimating private peering capacity. However, nothing precise is possible because a single AS can contain a large number of separate networks with their own limits that decrease estimate quality as well. Even if reasonable estimates per network basis are established, then the layer of limitation could move up to the transit provider level, as their routers are often not designed to handle maximum load through all the connections at the same time.

Another major factor is that a single device could be providing multiple abusable services simultaneously. In these cases, only the protocol providing higher bandwidth should be counted towards total attack capacity. It might be easy if the protocol measurements for each IP address happen within a small time frame (seconds or minutes), but this is not the case in the designed solution. The greater the time difference between measurements per IP, the less precise it becomes. IP address reachability is affected by dynamic addressing, operating hours, network anomalies and other factors. Properly addressing these factors is crucial for future multi-protocol measurements.

F. Legal and Ethical Considerations

Cybersecurity researchers often cross into gray areas and the legal basis for cybersecurity research is still evolving around the globe. There are three main legal and ethical aspects to consider for this research: scanning the Internet to find abusable devices, measuring discovered devices, and publishing the results.

Scanning the Internet is a common occurrence, it is performed by academic and commercial researchers as well as malicious parties. Although there is no common legal framework that addresses scanning, most non-malicious researchers follow the best practices [10], [11] laid out by the developers of zmap. This allows the minimization of negative impact on the scanned networks and devices but does not negate legal liability.

The significantly higher concern is the measuring stage for every discovered device, as it requires significant interaction with the devices by sending dozens of requests and measuring replies. Scanning for TCP protocol usually involves sending 1 request and a more detailed investigation might involve multiple requests to extract the properties of the device. DDoS capacity measurement relies on the ability to detect rate limiting, thus the number of requests should exceed commonly used rate limits. A large number of requests might interfere with the measured device or cause it to hang, which opens up legal liability. At the same time, devices with abusable protocols are already abused for real-world DDoS attacks, so if they are susceptible to overload, they might be continually affected and should not be serving a critical role.

Published security research always has a risk of being abused by malicious parties. Responsible disclosure minimizes impact, but in these cases of known abused protocols, it is not effective. Furthermore, there is no easy mitigation possible for the DDoS issue. The goal of the research is to present results and encourage positive long-term changes. All three discussed ethical and legal aspects are being evaluated for further research.

4. PRELIMINARY RESULTS

A proof of concept was developed to test the proposed methodology for the NTP protocol. NTP DDoS is known to be significantly mitigated and has a small set of abusable devices to minimize the potential negative impact of the research. The abused feature is a diagnostic command monlist and not a part of the protocol itself. This feature was enabled by default for the NTP server software distribution and produced BAF up to 4670 [2]. After discovery, it was quickly abused by attackers, and at the beginning of 2014 it caused record-breaking DDoS attacks up to 400 Gbps [12]. This command returns up to 440 bytes of payload per packet and up to 100 packets containing recent client data.

A. Scanning and Measuring NTP

Since abused command is not part of the protocol, specific monlist payload must be sent as a request. Different implementations and versions of NTP servers treat this command differently. The vast majority were observed to completely disregard the request without any reply and that is the general way it is expected to detect abusable functionality. However, there are multiple other types of responses stating that the command is not supported and standard time synchronization packets were received as well; these responses are undesired.

Scanning and measuring were conducted in August 2018; as there is no known common rate limit specific to the NTP monlist command, an aggressive 100 measurements per device were used. From the full Internet scan, 943,116 UDP responses to monlist were received, the majority were deemed undesired and only 92,990 devices were actually measured. Almost 63% of the measured devices, a majority of which were located in China, did not respond at all at the measurement stage, potentially indicating network issues, some DDoS protection mechanism or aggressive one-request limits.

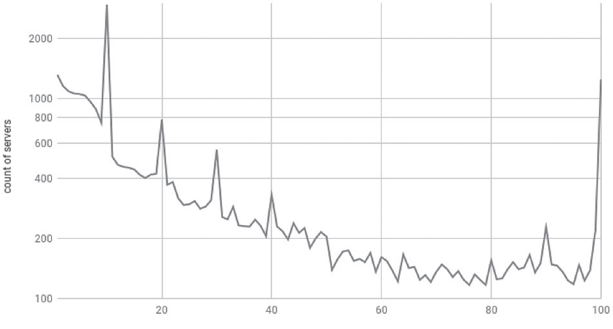
At the measurement stage, 33,325 devices responded with at least one valid monlist response containing an 80-byte payload; these devices are used for attack capacity calculation. Although some of the remaining devices provide some amplification, they are insignificant contributors to the attacks.

B. Rate Limiting

The number of responses for every request was originally expected to identify common patterns of rate limiting but the data produced just demonstrated a downward trend. This might be because it is not known in which order packets are received by the server and only monitoring the server's output on the wire would yield clear patterns. Aggregated data for the number of NTP servers per response count presented in Figure 1 portrays a much clearer picture. 1310 servers responded once (totaling 2 responses as one was received by zmap), to all requests responded only 1242 servers indicating sufficient computing power and network connection quality. However, most noteworthy are the clearly observable spikes at 10, 20, 30, 40, 80 and 90 responses.

There is nothing in the measurement system or network that relies on increments of 10 for sending or receiving packets. This indicates that some kind of rate limiting might be present, possibly set by humans. It is not necessarily explicitly defined in the configuration file of the NTP server software. It might be hardcoded as a limit inside the software or the system itself, especially for low-power embedded systems. This limit might also be present outside the devices, or it is possible that some rate limiting might be enforced by network devices in general or possibly targeting response payload known to be used mostly for DDoS attacks. It is not enforced by measurement network ISP, otherwise the full response spike would not be so significant. It is unlikely that this limit is enforced by a major IP transit provider, or that end-user networks apply these limits manually. Another possibility is that some network security solutions apply these limits automatically. Midsized ISPs are the most likely candidates that would manually create this type of limiting policy.

FIGURE 1: COUNT OF SERVERS PER COUNT OF RESPONSES

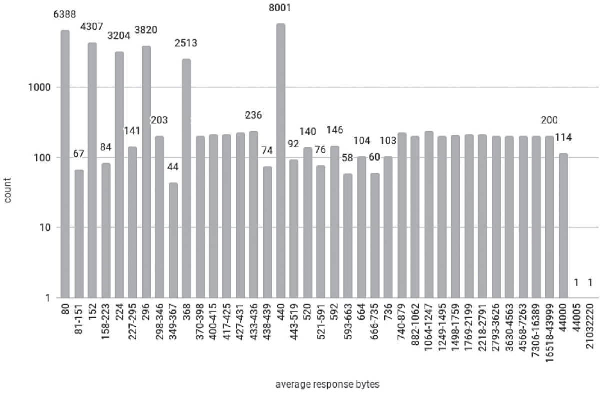


C. Response Size

Actual response size is an important metric as it allows us to calculate real-world BAF. Since the attacker's spoofable bandwidth is a limiting factor of the attack, the attacker would prefer the maximum amplification that abusable services offer. A small response is not necessarily limiting the total contribution to the attack, but it is definitely increasing network resource spending from the attacker. If no implicit or explicit rate limiting is present, then the server can utilize all the upload bandwidth available to it.

The NTP server distribution per average response size is provided in Figure 2. The average is calculated over received responses. If a single response is received, then its size will be the average. The most common values are displayed individually and uncommon values are grouped together; the highest values are the most significant ones. With an 80-byte payload, 6388 servers responded, all of which are monlist replies containing a single client entry. However, the most common response size is 440 bytes in 8001 cases, which corresponds to a single full packet monlist response. It is either an implementation issue or a mitigation effort fix for the configuration or the software itself to minimize the impact of the abuse. Only 114 servers provided maximum possible responses of 100 packets with a 440-byte payload totaling 44,000 byte responses without packet loss. Diagnostic information about a single client uses 72 bytes of response, 5 clients produce single full response packet and if there are more clients, additional response packets are generated in the same way.

FIGURE 2: NTP AVERAGE RESPONSE SIZE DISTRIBUTION



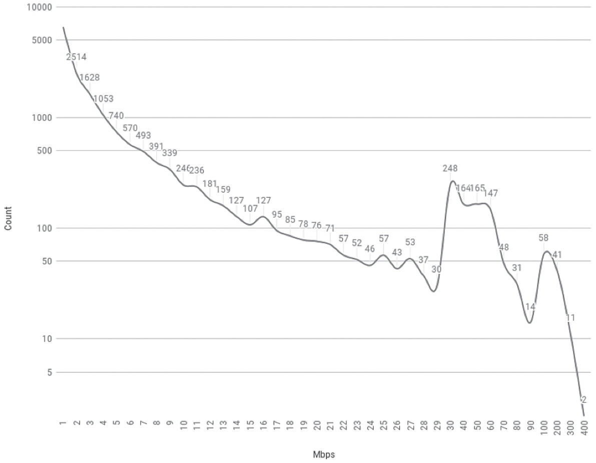
D. Response Speed

All the servers responding with an average speed above 0.5 Mbps can be considered significant contributors and are presented in Figure 3. In total 17,208 or 54% of the servers responded with speed above 0.5 Mbps, with the peak being at 0.5 – 1.5 Mbps and then decreasing; 73% of these servers had measured speeds below 5.5 Mbps.

A large portion of responses (569 servers) were received at speeds below 5 Kbps. These servers either have a low response rate and respond slowly or have a high response rate and take multiple seconds to respond from the first to last packet. Random sampling indicates that a significant portion of these devices have slow-speed wireless connections to the Internet. These servers are disregarded from further calculations. There is a spike in the number of servers responding with speed 5 – 15 Kbps; in total, 6896 servers responded with speeds between 5 and 100 Kbps. In total, 14,750 servers responded with speed below 0.5 Mbps. These are insignificant contributors to the overall DDoS attack capacity that could potentially be disregarded from the analysis set.

There are noticeable outliers with average speeds above 100 Mbps, most of which were identified as data centers and hosting providers providing virtual servers and dedicated servers for rent and supplying them with high-speed Internet connections with speeds of 100 Mbps – 1 Gbps or above. A top provider, OVH, with 15 reflectors is known for low prices and abuses. These devices are high contributors to the attack capacity.

FIGURE 3: NTP SERVERS RESPONDING WITH AVERAGE SPEEDS ABOVE 0.5 MBPS



E. Contributors to the Attack

Most NTP servers were located in the USA (8061), China (4689), Brazil (3320), Spain (2420), Turkey (1832), Indonesia (1432), Taiwan (1227), Vietnam (1226), Saudi

Arabia (1134) and Malaysia (1032). The USA is disproportionately represented in many scans, which might be surprising, but it is related to the historical availability of the Internet and a high number of legacy systems. The whole continent of Africa has very few amplifiers, about half of the countries have none. With the speed and cost of the Internet in Africa, it is expected that contribution to the total attack capacity is insignificant. Asia is a high contributor and many other network issues are caused by fast proliferation and growth of the Internet in these developing countries. A large connection count and fast speed, coupled with a lack of regulation and enforcement and general disregard for the best network management practices, all cause Asian countries to be breeding grounds for cybersecurity issues. However, as noted by existing research [5], pure count is not a good metric for estimating contribution to total attack capacity; the count needs to be balanced against upload bandwidth.

Bandwidth contribution is a significantly more important metric than overall server count. Compared to the count, significant differences can be observed – the USA and Spain contributed much more count-wise than capacity-wise. This might confirm that high-count ISPs might have low-power or low-speed embedded devices running the services without contributing significantly to the total attack capacity. China is the top contributor with about 42 Gbps total attack capacity, followed by the USA with 16 Gbps. Next are Russia and France which have low NTP server counts but very high network speeds. The rest of the top 10 are Asian countries and Brazil. Overall top contributors to the attack capacity are developed countries and developing countries with high Internet connectivity speeds.

F. NTP Attack Capacity

Summing all the calculated average speed together for the reflectors that provided more than one reply with calculated speed of at least 5 Kbps, the total speed of the NTP monlist DDoS attack was **134 Gbps**, generated by 31,389 servers. This value does not necessarily correspond to the real-world situation. There might have been competition for bandwidth with ongoing DDoS attacks. Real-world capacity could be significantly larger. Geographic distance decreases average speed as well, intermittent or permanent network quality issues would decrease measured speeds but not actual bandwidth.

There are no current estimates of NTP monlist attack capacity and no published recent attack case studies because the attack has long lost its peak capacity. It would allow the extraction of some empirical constant that potentially could be used as a multiplier for the measured capacity to produce a realistic estimate.

The real measured BAF for the 134 Gbps capacity can be calculated by dividing the total received bytes with the 100 payloads sent multiplied with payload length and server count. In this case, the real total measured **BAF** was **20.55**, which is significantly below the standard maximum of 2750. If an attacker disregards servers with large packet losses and small responses, then he can achieve attacks with multiple times larger BAFs. Whether or not attackers conduct such measurements is an open question for further research.

5. CONCLUSIONS

The proposed methodology is promising and covers aspects missing in existing ones. The implemented proof of concept produced an NTP DDoS capacity of 134 Gbps and is suitable for adaptation to different protocols. Significant technical, ethical and legal concerns were identified that require further investigation to determine if the research methodology is viable.

REFERENCES

[1] "History of DDoS Attacks," *Radware*, 13-Mar-2017. [Online]. Available: https://security.radware.com/ddos-knowledge-center/ddos-chronicles/ddos-attacks-history/. [Accessed: 04-May-2017].

[2] C. Rossow, "Amplification Hell: Revisiting Network Protocols for DDoS Abuse," in *Proceedings of the 2014 Network and Distributed System Security Symposium*, San Diego, CA, USA, 2014.

[3] J. Czyz, M. Kallitsis, M. Gharaibeh, C. Papadopoulos, M. Bailey, and M. Karir, "Taming the 800 pound gorilla: The rise and decline of NTP DDoS attacks," in *Proceedings of the 2014 Conference on Internet Measurement Conference*, 2014, pp. 435–448.

[4] Center for Applied Internet Data Analysis based at the University of California's San Diego Supercomputer Center, "State of IP Spoofing." [Online]. Available: https://spoofer.caida.org/summary.php. [Accessed: 23-Jul-2018].

[5] E. Leverett and A. Kaplan, "Towards estimating the untapped potential: a global malicious DDoS mean capacity estimate," *Journal of Cyber Policy*, vol. 2, no. 2, pp. 195–208, May 2017.

[6] C. Morales, "NETSCOUT Arbor Confirms 1.7 Tbps DDoS Attack; The Terabit Attack Era Is Upon Us," 05-Mar-2018. [Online]. Available: https://asert.arbornetworks.com/netscout-arbor-confirms-1-7-tbps-ddos-attack-terabit-attack-era-upon-us/. [Accessed: 09-Mar-2018].

[7] The Shadowserver Foundation, "The scannings will continue until the Internet improves." [Online]. Available: http://blog.shadowserver.org/2014/03/28/the-scannings-will-continue-until-the-internet-improves/. [Accessed: 30-Jun-2018].

[8] Qihoo 360 Technology Co., Ltd, "Scan volume per 10 minutes," *NetworkScan Mon*. [Online]. Available: http://scan.netlab.360.com/. [Accessed: 19-Jul-2018].

[9] Qihoo 360 Technology Co.,Ltd, "Insight into Global DDoS Threat Landscape," *DDoS Mon*. [Online]. Available: https://ddosmon.net/insight/. [Accessed: 20-Jul-2018].

[10] Z. Durumeric, E. Wustrow, and J. A. Halderman, "ZMap: Fast Internet-wide Scanning and Its Security Applications," in *Presented as part of the 22nd USENIX Security Symposium (USENIX Security 13)*, Washington, D.C., 2013, pp. 605–620.

[11] Z. Durumeric, D. Adrian, A. Mirian, M. Bailey, and J. A. Halderman, "A Search Engine Backed by Internet-Wide Scanning," in *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security*, Denver, CO, USA, 2015, pp. 542–553.

[12] M. Prince, "Technical Details Behind a 400Gbps NTP Amplification DDoS Attack," 13-Feb-2014. [Online]. Available: https://blog.cloudflare.com/technical-details-behind-a-400gbps-ntp-amplification-ddos-attack/. [Accessed: 25-Jan-2018].