21 Years of Distributed Denial-of-Service: A Call to Action

Eric Osterweil and Angelos Stavrou, George Mason University
Lixia Zhang, University of California, Los Angeles

We are falling behind in the war against distributed denial-of-service attacks. Unless we act now, the future of the Internet could be at stake.

Distributed denial-of-service (DDoS) attacks from robot networks (botnets) are now 21 years old, and during that time, they have been a growing blight on the Internet, swelling in size and complexity. Although DDoS mitigation techniques have been evolving, they seem to be outpaced by the increasing sophistication and diversification of the techniques, tactics, and procedures, in other words, "behaviors," used by attackers. In this article (our second of two, the first part ran in the July 2020 issue), we would like to sound an alarm: we have been losing ground to our adversaries in the DDoS war, and we must take corrective action now.

We posit that the research community is ideally suited to formulate and investigate fundamental questions about how we got here. For example, are there foundational issues in the existing network architecture that make DDoS attacks easy? How can our network infrastructure be enhanced to address the principles that enable the DDoS problem? To reap the gains of any suggested mitigations, how can we incentivize the development and deployment of necessary changes? To move our defensive posture forward we need to step up a level, take a fresh look at the problem, and consider approaches that take into consideration the current state of the Internet and networking as a whole.

DDOS SCRUBBING

Detecting attack traffic—as opposed to normal traffic or that which is abnormal but benign—is crucial, especially if a service provider intends to use a DDoS mitigation service to remediate attacks. Internet services, such as the

Digital Object Identifier 10.1109/MC.2020.2993330
Date of current version: 30 July 2020

 0018-9162/20©2020IEEE

EDITOR **JAMES BRET MICHAEL**
IEEE Senior Member; bmichael@nps.edu

web, email, messaging, streaming and video delivery, and so forth, are typically equipped to handle expected loads and user behavior under "peacetime" conditions. This provisioning often has a monetary component: buying transit traffic capacities, setting up server infrastructure, and staffing operations, among others. It also includes capacity plans for periodic above-average bursts of traffic, slow Transport Layer Security (TLS) handshakes, and other suboptimal but nonthreatening client behaviors. However, DDoS attacks are designed to overcome these capacities. Once an attack is detected, the common response is to apply network- or application-level filters to packets and/or packet flows to inspect and "scrub" attack traffic away from legitimate flows.

WHY SCRUBBING?

The ultimate goal of DDoS mitigation is to ensure the uninterrupted delivery of "good" traffic to its intended destination. Scrubbing uses techniques that range from measuring traffic heuristics and various vendor-specific traffic-assessment techniques to approaches that address complex behaviors at the application layer. In many cases, detecting and discerning DDoS attacks is difficult. The primary issue in differentiating attack traffic is that, from a network point of view, traffic sources (individual hosts and networks) can appear to be sending both proper application traffic and DDoS traffic. For example, a single large home-access network might have well-functioning hosts transacting with a website and separate compromised hosts (bots) sending DDoS traffic. This blending of traffic can make remediation problematic. For example, is a User Datagram Protocol packet a legitimate Domain Name System query or a genuine Network Time Protocol query? Is a Transmission Control Protocol (TCP) synchronization packet actually trying to set up a TCP connection or an authentic TLS 1.3 zero-round-trip-time resumption?

In these types of situations, it is challenging to separate attack traffic based solely on the information carried in the network/transport layer protocol headers. Many attacks start becoming distinguishable from benign traffic only after multiple round trips between a client and a service, and defenders are often very averse to acting too fast and thereby dropping legitimate traffic (false positives). It is often the case that remediation must assess application-level semantics to discern attack traffic from nonattack traffic, especially given that behaviors continue to become more complex.

In today's Internet architecture, any host can send Internet Protocol (IP) packets to any other host in the world, which precisely explains why scrubbing is the industry's last line of defense. While approaches such as FlowSpec, remotely triggered blackhole (RTBH) filtering, and the Internet Engineering Task Force's new DDoS Open Threat Signaling (DOTS) Working Group (WG)[1] all attempt to enhance the network/transport layer to aid in DDoS defense, scrubbing centers are the fallback solution that catches all DDoS attacks that get through. Scrubbing centers are sites dedicated to mitigation, and they are typically provisioned with high bandwidth and specialized hardware and/or software. Because discerning the difference between traffic bursts and attacks and then erecting mitigation defenses requires time, some Internet services engage an always-on mitigation provider. In these cases, all traffic is sent through mitigation machinery, even during peacetime, so that detection and mitigation can be simultaneously handled with high confidence and low latency.

STATE OF THE ART: SCRUBBING CENTERS

Some companies offer access to scrubbing centers as a commercial product, dubbed *mitigation as a service* (*MaaS*).[2–4] While any network can deploy mitigation solutions for self-protection, mitigation providers generally offer quantifiably better traffic filtering and higher bandwidth protection than one's own solutions. Mitigation providers invest in large transit and peering capacities, often in excess of 1 TB/s in aggregate. Moreover, they deploy their infrastructure across the Internet and have a global view of traffic volumes, which are inspected by augmented conventional mitigation appliances, and they have 24/7 security operation centers that monitor traffic and include large cybersecurity threat intelligence teams. The scale of their operations enables defenders, information security teams, incident response teams, and researchers to create profiles that can attribute attacks to specific families of malware.

Occasionally, data from scrubbing centers can point to the actor(s) responsible. This is because large mitigation providers with multiple clients can observe a broader cross section of both normal and attack traffic, which enables deeper analysis. At the same time, one may also ask whether this fact reflects the benefit of MaaS providers' positions or whether it is an indication that our defenses are in need of basic research to overcome the inherent asymmetry between attackers and victims. Would the dependency on MaaS further drive the Internet toward centralization if all Internet customers could afford the cost of the service?

Furthermore, since attack sources originate from increasingly distributed networks with ever-higher transit capacities, it is unlikely that mitigation providers' topological diversity

would be able to keep up with attackers' dispersion. Figure 1 illustrates the fact that as attack source distributions continue to grow, scrubbing centers represent a relatively centralized and inconveniently located solution to the DDoS distributed threat. That is, if any given scrubbing center intends to mitigate attack traffic originated from n different networks, it likely will need to be able to absorb $n \times BW$ attack traffic (where BW is the bandwidth of the network that attacking bots are deployed on). With the multitudes of readily compromisable devices that exist on increasingly well-provisioned networks, this means that scrubbing centers face the challenge of keeping pace with n other networks that are all attempting to maximize their own bandwidth for an optimal customer experience.

ASYMMETRIC THREAT: FUNDAMENTALS OF THE PROBLEM

Twenty-one years of fighting DDoS has taught us a lot. During that time, technology advances have greatly increased the bandwidth of our service networks, but that growth has also augmented attack capacities. The problem is that using the increased bandwidth does not in itself constitute abuse. In fact, it has actually benefited attackers more because for every remediation instance that has more bandwidth, every attacking bot (the numbers of which are growing fast) has additional bandwidth, too. This is because Moore's law has brought more abundant, cheap, powerful, and, unfortunately, easily compromisable end devices onto the Internet at a rate that not only meets but often outpaces the growth of our remediation infrastructure. This is what frames the asymmetry and ultimate impedance mismatch of DDoS: there are ever-larger numbers of attack sources that can launch at any time, while malicious traffic is being triaged by relatively fewer remediation infrastructures.

What's more, attack sources can be far from the remediation centers. As malevolent traffic heads toward scrubbing centers, it aggregates. This leads to congested transit links, which, in turn, can lead to service disruptions in the network. At this time, scrubbing centers are tactical necessities that provide critical protections to Internet services and keep the Internet operating. Strategically, however, this approach pushes us toward the losing end of the DDoS arms race, with a growing edge capacity and increasingly

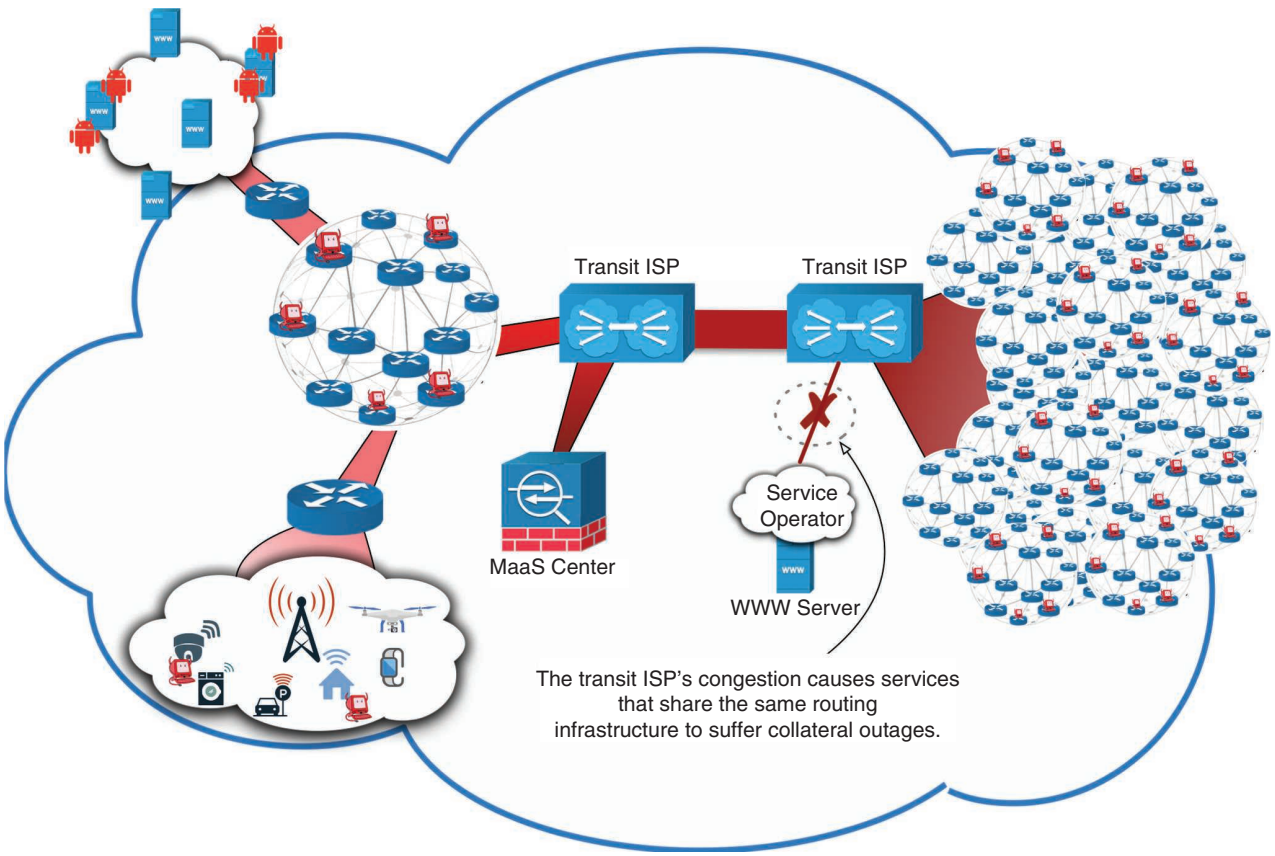

Figure 1. As DDoS traffic aggregates on its way from many sources to fewer scrubbing centers, mitigatable traffic volumes can become overwhelming (and result in collateral outages to other services). WWW: World Wide Web; ISP: Internet service provider.

well-provisioned compromised devices (bots) and with no progress toward fixing the root cause.

WHAT WOULD HELP

The problem space of DDoS can be seen as rather intricate: attackers' behaviors are highly variable, they come from a multitude of different topological locations, they are carried by numerous routing infrastructures, and they keep evolving to a higher complexity. However, if one steps up a level, one can realize that all the preceding issues are the symptoms of much deeper root causes. We believe that now is a critical time to identify and address the fundamental source of DDoS instead of falling further into more expensive and complicated tail-chasing mitigations. We should begin identifying whether there exist any design tenets and architectural features that have made DDoS the asymmetric problem we face today so that we can use this understanding to jumpstart the development of long-term effective solutions. One core observation from our current practice is that combating the distributed nature of DDoS from a relatively small number of centralized vantage points misaligns many central aspects of the nature of the threat. Enabling remediations at the edge, where attacking nodes reside, seems to realign one fundamental aspect of the overall problem by mitigating the effects of DDoS before it traverses the network toward the end target.

Given that many attacks also capitalize on application-level semantics, one potential approach might be to investigate solutions that can be effectuated in the network by providing the network layer with information about application semantics, but this requirement would conflict with today's network layered design. Alternatively, some investigations into flow control consider relatively fewer innovations in the network and push logic to the edge. In either case, in the long run, unless mitigations become effective at the edge of our network and administrative boundaries, we have a dismal chance to succeed in holding the floodgates against the ever-increasing number of vulnerable devices fueled by Moore's law and compounded with the always-growing network capacity. It is plainly clear that we must address the root cause of this ballooning volumetric DDoS threat, starting now.

EXISTING DEFENSES THAT USE THE NETWORK SUBSTRATE

Today, spoofed source addresses play a large role in the biggest DDoS attacks on the Internet (directly or via reflection). Approaches to alleviate source address spoofing, discussed at length in the first part of this article series, including two operational guidance documents [called Best Common Practices (BCPs) 38 and 84], advise operators to configure their networks to disallow outbound and inbound packets with spoofed source addresses.[5,6] Although much effort has gone into promoting the mitigation of spoofed traffic, the deployment success has been debated, and the sizes of recent DDoS reflector attacks with spoofed source addresses is undeniably still growing.[7,8]

One reason for the lagging adoption could be because there is a misalignment of costs and benefits. That is, the source network operators and their customers deploying and managing these protections, including home-access networks, open access networks, and other client-side providers, do not receive the derived protections. Indeed, this has been noted in operational communities as well: "The costs ... [are] not directly [being] borne by the potential beneficiaries of deploying the solution."[9] Another reason comes from the extremely diverse composition of the Internet: not all providers possess the same level of comprehension of best practices nor do they care equally about the well-being of the Internet as a whole at a global scale.

While the interdomain routing system that carries traffic has been undergoing security enhancements through a relatively new set of standards called the Resource Public Key Infrastructure,[10] this does not help DDoS remediation. These efforts focus on verifying IP addresses and autonomous system number resource holders as well as the propagation of routing information in the control plane, but they do not address security issues in the data plane, which is where source address spoofing occurs.

RESEARCH LITERATURE

Previous research, such as the Traffic Validation Architecture[11] and Pushback,[12] proposed forms of distributed remediations by using in-network deployment. These approaches likely did not gain traction because they did not align deployment incentives with the parties paying the costs (just as with BCPs 38 and 84). A more recent proposal, Stellar,[13] doubles down on the same fundamentals as methods including FlowSpec, RTBH, and the DOTS WG: a black-holing framework focused on deployments in large Internet exchange points. This work aims to reduce black-holing's collateral damage by pushing remediations closer to the attacking sources. These techniques innovate in the space where defenders are already falling behind attackers.

In addition to research literature that chronicles diverse approaches, techniques, and tools for investigating the symptoms of today's manifestations of DDoS, investigations also exist that examine what could be the underlying causes of DDoS and what architectural changes could help with mitigation.[14] Indeed, with a more fundamental perspective in mind, in 2010 the National Science Foundation created a program to fund research into future Internet architectures.[15] Similar DARPA efforts include Extreme DDoS Defense[16] and Open, Programmable, Secure 5G.[17] Among the proposed new architecture designs, Named Data Networking (NDN)[18] shows that the architecture of the Internet itself could provide the foundation for effective long-term solutions. A brand-new

architecture would certainly take time to be widely deployed, and it would be especially difficult inside the Internet backbone. There, a new solution would be at odds with what carriers and large operators can currently offer. However, the incremental deployment of such an architecture at the network edges seems, relatively speaking, to be more feasible and easier. Such an approach could have the potential to gradually curtail attacks because the assaults' sources largely reside at the edges.

DISCUSSION

Have we made fundamental enhancements to our DDoS defense arsenal during the past 21 years? While technical innovations are clear, the fundamentals of our remediations appear to still focus on building network-level floodgates using brute force and enhanced deep packet inspection scrubbing. The landscape of cheap and compromisable bots has only become more fertile for miscreants and more damaging to Internet safety. Increases in bandwidth have benefited both genuine Internet services and attacking bots, with the latter's gains being multiplied by the asymmetric scaling factor, and the age of the Internet of Things and 5G has the potential to make the situation far more severe. To make matters worse, our needs and applications have become more complex, and even our security and privacy protections (such as TLS, HTTPS, and so on) have become a hurdle, making DDoS more difficult to mitigate inside the network.

As a starting point for discussions, we must recognize the weakness in our existing network architecture, which is so easily abused by DDoS attackers. We must enhance the Internet in ways that undercut the properties that DDoS attack vectors rely on, and this likely requires staged planning and efforts. In the long run, the whole community needs to converge around a strategy for a future direction and an architecture with resilience to abuse. In the meantime, tactically, we also need new solutions that enable us to mount distributed defenses. Federal investment into future Internet architectures has produced new strategic designs, such as NDN, which show the promise of incremental deployability starting from the edge. However, any new architecture will inevitably face challenges, including policy hurdles, economic considerations, and possibly even legislative actions.

While the Internet waits for needed protections via architectural changes, historical lessons suggest that to gain deployment traction for short-term solutions, early adopters must be rewarded economically, qualitatively, and in other palpable ways. Aligning the deployment costs of defenses with direct benefits is going to be key to changing the tide of our war on DDoS. Debates exist about which layers' expressiveness and intelligence should be incorporated, but investigations must be conducted to understand the fundamentals that enable the DDoS problem and how network defenses will be applied. This article is a call to action: we need to turn the tables on miscreants while we still can. We do need today's tactical remediations—they are buying us time—but fundamental strategic solutions are the Internet's only hope in the long run.

REFERENCES

1. A. Mortensen, F. Andreasen, T. Reddy, N. Teague, R. Compton, and C. Gray, "Distributed-denial-of-service open threat signaling (DOTS) architecture," IETF Secretariat, Internet-Draft draft-ietf-dots-architecture-10, Mar. 6, 2020. [Online]. Available: https://tools.ietf.org/html/draft-ietf-dots-architecture
2. "Akamai DDoS protection," Akamai, Cambridge, MA, 2019. [Online]. Available: https://www.akamai.com/us/en/resources/ddos-protection.jsp
3. "CloudFlare advanced DDoS attack protection," CloudFlare, San Francisco, 2020. [Online]. Available: https://www.cloudflare.com/ddos/
4. "Neustar defense and performance," Neustar, Sterling, VA, 2020. [Online]. Available: https://www.security.neustar/digital-defense/ddos-protection
5. P. Ferguson and D. Senie, "Network ingress filtering: Defeating denial of service attacks which employ IP source address spoofing," RFCEditor, BCP38, 2000. [Online]. Available: http://www.rfc-editor.org/rfc/rfc2827.txt
6. F. Baker and P. Savola, "Ingress filtering for multihomed networks," RFC Editor, BCP 84, 2004. [Online]. Available: https://tools.ietf.org/html/bcp84
7. M. Antonakakis et al., "Understanding the mirai botnet," in *Proc. 26th USENIX Security Symp.*, 2017, pp. 1093–1110. doi: 10.5555/3241189.3241275.
8. B. Krebs, "KrebsOnSecurity hit with record DDoS," KrebsOnSecurity, Sept. 21, 2016. [Online]. Available: https://krebsonsecurity.com/2016/09/krebsonsecurity-hit-with-record-ddos/
9. G. Huston, "Why is securing the Internet so hard?" presented at the Asia-Pacific Regional Internet Conf. Operational Technologies (APRICOT'19), Daejon, South Korea, Feb. 18–28, 2019.
10. M. Lepinski and S. Kent, "An infrastructure to support secure internet routing," Internet Engineering Task Force, Tech. Rep. RFC-6480, Feb. 2012. [Online]. Available: https://tools.ietf.org/html/rfc6480
11. X. Yang, D. Wetherall, and T. Anderson, "TVA: A DoS-limiting network architecture," *IEEE/ACM Trans. Network*, vol. 16, no. 6, pp. 1267–1280, 2008. doi: 10.1109/TNET.2007.914506.
12. J. Ioannidis and S. M. Bellovin, "Implementing pushback: Router-based defense against DDoS attacks," in *Proc. Network and Distributed System Security Symp. (NDSS)*, vol. 2002, pp. 6–8. doi: 10.7916/D8R78MXV.

13. C. Dietzel, G. Smaragdakis, M. Wichtlhuber, and A. Feldmann, "Stellar: Network attack mitigation using advanced blackholing," in *Proc. ACM 14th Int. Conf. Emerging Networking Experiments and Technologies*, 2018, pp. 152–164. doi: 10.1145/3281411.3281413.
14. M. Handley and A. Greenhalgh, "Steps towards a DoS-resistant internet architecture," in *Proc. ACM SIGCOMM Workshop Future Directions Network Architecture*, 2004, pp. 49–56. doi: 10.1145/1016707.1016717.
15. "NSF announces Future Internet Architecture awards," National Science Foundation, Alexandria, VA, Aug. 2010. [Online]. Available: https://www.nsf.gov/news/news_summ.jsp?cntn_id=117611
16. J. M. Smith, "Extreme DDoS Defense (XD3)," DARPA, Arlington, VA. [Online]. Available: https://www.darpa.mil/program/extreme-ddos-defense
17. J. M. Smith, "Open, programmable, secure 5G (OPS-5G)," DARPA, Arlington, VA. [Online]. Available: https://www.darpa.mil/program/open-programmable-secure-5g
18. L. Zhang et al., "Named data networking," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 44, no. 3, pp. 66–73, 2014. doi: 10.1145/2656877.2656887.

ERIC OSTERWEIL is an assistant professor in the Department of Computer Science at George Mason University. Contact him at eoster@gmu.edu.

ANGELOS STAVROU is a professor in the Department of Computer Science at George Mason University. Contact him at astavrou@gmu.edu.

LIXIA ZHANG is the Jonathan B. Postel Professor of Computer Science at the University of California, Los Angeles. Contact her at lixia@cs.ucla.edu.

Digital Object Identifier 10.1109/MC.2020.3007050