Stellar: Network Attack Mitigation using Advanced Blackholing

Christoph Dietzel
TU Berlin/DE-CIX
christoph@inet.tu-berlin.de

Matthias Wichtlhuber
DE-CIX
matthias.wichtlhuber@rnd.de-cix.net

Georgios Smaragdakis
TU Berlin
georgios@inet.tu-berlin.de

Anja Feldmann
Max Planck Institute for Informatics
anja@mpi-inf.mpg.de

ABSTRACT

Network attacks, including Distributed Denial-of-Service (DDoS), continuously increase in terms of bandwidth along with damage (recent attacks exceed 1.7 Tbps) and have a devastating impact on the targeted companies/governments. Over the years, mitigation techniques, ranging from blackholing to policy-based filtering at routers, and on to traffic scrubbing, have been added to the network operator's toolbox. Even though these mitigation techniques provide some protection, they either yield severe collateral damage, e.g., dropping legitimate traffic (blackholing), are cost-intensive, or do not scale well for Tbps level attacks (ACL filtering, traffic scrubbing), or require cooperation and sharing of resources (Flowspec).

In this paper, we propose Advanced Blackholing and its system realization Stellar. Advanced blackholing builds upon the scalability of blackholing while limiting collateral damage by increasing its granularity. Moreover, Stellar reduces the required level of cooperation to enhance mitigation effectiveness. We show that fine-grained blackholing can be realized, e.g., at a major IXP, by combining available hardware filters with novel signaling mechanisms. We evaluate the scalability and performance of Stellar at a large IXP that interconnects more than 800 networks, exchanges more than 6 Tbps traffic, and witnesses many network attacks every day. Our results show that network attacks, e.g., DDoS amplification attacks, can be successfully mitigated while the networks and services under attack continue to operate untroubled.

CCS CONCEPTS

• **Networks → Denial-of-service attacks**; *Network components*; *Network measurement*; *Network services*;

KEYWORDS

BGP; IXP; Blackholing; DDoS Mitigation.

1 INTRODUCTION

The revolution of the digital age fueled by the Internet has attracted the good but the evil alike. While the threats executed over the Internet are multifaceted from a criminalistics perspective, e.g., fraud, data and identity theft, espionage, or cyber terrorism, the dominant network threat is Denial-of-Service (DoS) attacks [2]. The goal of DoS attacks is to force a service or system to become unavailable by consuming crucial resources. These resources can be computing power at the servers or exploitation of application-layer vulnerabilities, i.e., semantic attacks, or network bandwidth, i.e., volumetric attacks. To conduct such volumetric attacks, adversaries often use Distributed DoS (DDoS). Traffic from numerous distributed sources is generated and steered towards a target service to make it unavailable. Once the network links to the target are congested due to the DDoS attack, legitimate traffic that traverses the same links is also affected.

DDoS threats are continuously increasing in terms of volume, frequency, and complexity. While the largest observed and publicly reported attacks were between 50 to 200 Gbps before 2015 [59, 60, 70], current peaks are an order of magnitude higher and exceeded 1 Tbps [9, 48] in 2016, and 1.7 Tbps [57] in early 2018. We also observe a massive rise in the number of DDoS attacks. Jonker et al. [41] report that a third of all active /24 networks were targeted by DDoS attacks between 2016 and 2017. Similar observations are reported by the security industry [3, 19]. A particularly prominent DDoS attack type is amplification attacks [64, 65]. They take advantage of protocol design flaws, whereby a relatively small request triggers a significantly larger response. With a spoofed source IP address [49] the response traffic is amplified and reflected to the target. Vulnerable protocols include classical protocols such as NTP, DNS, and/or SNMP [20, 64], as well as relatively new protocols, e.g., DNSSEC [74] and memcached [5, 57]. Amplification factors of up to $50,000\times$ have been witnessed in the wild [73]. To exemplify, a request of 15 bytes can trigger a 750 Kbytes response.

1.1 DDoS Mitigation: State of the Art

This alarming increase in DDoS attacks and their sophistication and severity, e.g., see [56, 77], demands scalable yet cost-effective countermeasures. However, at this point, we are left with various mitigation techniques and tools that can *partially* counteract the impact of the attacks. These include: (i) *Traffic Scrubbing Services (TSS)*, (ii) *Router Access Control List Filters (ACL)*, (iii) *Remotely Triggered Black Hole (RTBH)*, and (iv) *BGP Flowspec*.

Traffic Scrubbing Services (TSS): offer all-round carefree services to their subscribers. They redirect the traffic of a service to specialized hardware either via DNS redirection or BGP delegation [43]. There they classify traffic as unwanted or benign and send the benign "scrubbed" traffic to its original destination or move the destination to their network [4, 30, 43, 75]. The convenience and fine-grained filtering of TSS comes at significant recurring costs and requires in-time subscription and setup. Moreover, it currently has inherent limitations, e.g., per packet or per flow processing for deep packet inspection, which can reduce effectiveness [75] and does not cope with Tbps-level attacks [48]. Moreover, it may reroute traffic and, thus, impose performance penalties, and is vulnerable to evasion tactics [42].

ACL Filters: are often used by Internet Service Providers (ISPs) and Internet Exchange Points (IXPs) to overcome specific network problems. They deploy policy-based filters that drop unwanted traffic at their AS border routers. The implementations and capabilities depend on the vendor-specific hardware, e.g., ACL rules or QoS classifiers. Such filters can work well if the hardware is homogeneous, the network engineers have sufficient expertise, and the network management system supports the automated deployment of filters. However, such systems typically do not scale well and, given that the filtering location is beyond the ingress points of the network, the bandwidth to a neighbor AS can still be exhausted.

Remotely Triggered Black Hole (RTBH): also referred to as BGP Blackholing, is an operational DDoS mitigation technique [16]. ASes under attack can signal upstream ISPs [24, 40] or IXPs [22, 50] to drop traffic to specific IP prefixes. Using BGP to trigger blackholing is simple to realize and lowers the entry barrier for ASes, but limits the level of granularity of the blackhole (to IP prefixes) and the acceptance of neighboring ASes. Despite substantial growth of blackholing usage (it quadrupled between 2015 and 2017 [33]) that is evidence of its effectiveness to drop large volumes of attack traffic [26], unfortunately, it is coarse-grained. BGP blackholing also drops legitimate traffic to the prefix under attack and thereby causes collateral damage. Essentially, this makes the IP prefix partially unreachable. For RTBH to be effective, cooperation between network operators to act upon receiving a blackhole signal (typically, a BGP community) is required, see Section 2. Namely, it requires that BGP messages for prefixes more specific than /24 in IPv4 are propagated, thus, networks operators have to set up exceptions for blackholing to accept BGP messages such prefixes, e.g., /32 in IPv4.

BGP Flowspec: the BGP flow specification feature, also referred to as Flowspec, allows the deployment and propagation of more fine-grained filters (compared to RTBH) across AS domain borders, e.g., to mitigate DDoS attacks [18]. Flow specifications can match a particular flow with a source, destination, layer-4 (L4) parameters, packet characteristics such as length and fragment, and allow to specify a drop rate limit. Flowspec has received some adoption in *intra-domain* environments and has been shown to have good reaction time and performance [11, 66]. However, in the *inter-domain* environment, Flowspec has received little attention. Among the reasons is that in this environment it relies on trust and cooperation among different networks, as well as on the sharing of computational and network resources. In other words, providing one's resources to solve someone else's problem. This also raises liability questions, thus, it is challenging to implement in a highly competitive environment.

In summary, all of the above techniques can help mitigate DDoS. However, each of the above solutions has its own significant shortcomings. Among the main limitations of TSS is its cost and resource hunger. Among the drawbacks of ACL filters are the limited scalability and the demand for customization. RTBH suffers from its coarseness and the fact that it requires cooperation between ISPs and/or IXPs to make the service effective. Thus, compliance with RTBH signaling is complicated. Flowspec is not a popular choice in the inter-domain environment as it relies on trust and a high degree of cooperation among, potentially, competitive networks. Moreover, only TSS provides networks under attack feedback (telemetry) regarding the state and volume of the attack, i.e., if it is still ongoing or is over. Flowspec standardization [52] only recommends telemetry but ultimately it is a vendor-specific implementation. For a summary of the pros and cons see Table 1.

	TSS	ACL filters	RTBH	Flowspec	Advanced Blackholing
Granularity	✓	✓	✗	✓	✓
Signaling complexity	✗	✗	✗	✗	✓
Cooperation	•	•	✗	✗	✓
Resource sharing	✓	✓	✓	✗	✓
Telemetry	✓	✗	✗	•	✓
Scalability	✗	•	✓	✓	✓
Resources	✗	✗	✓	✗	✓
Performance	✗	✓	✓	✓	✓
Reaction time	✗	✗	✓	✓	✓
Costs	✗	•	✓	✓	✓

Table 1: Advanced Blackholing vs. DDoS mitigation solutions. ✓: advantage, ✗: disadvantage, •: neutral.

1.2 Advanced Blackholing in a Nutshell

In this paper, we propose another approach for attack mitigation, called Advanced Blackholing (Advanced BH). Advanced Blackholing does not require trust, cooperation, and sharing of resources among networks. It builds upon the excellent scalability of RTBH (to aggressively drop volumetric attack traffic) while incorporating the good properties of ACLs, Flowspec, and TSS (fine-granular filtering) in a lightweight fashion. Thus, Advanced Blackholing offers a new service in between RTBH and TSS and, as we will show, it can be deployed at scale, e.g., at IXPs.

IXPs offer an ideal deployment location for DDoS traffic mitigation as many ISPs use them to exchange traffic, e.g., more than 800 networks and more than 6 Tbps at DE-CIX in Frankfurt or AMS-IX in Amsterdam. Notice that by enabling such a service in one of these large IXPs, hundreds of member networks (as well as their customers and peer networks) will immediately benefit without the need to change anything else in the Internet protocols and the operation of the member networks, and without cooperation and coordination between two member networks. Previous examples of such innovation includes SDX [14, 15, 36, 37]. IXPs can also easily absorb the largest attacks seen to date[1][7, 21, 57], as they have Tbps of capacity before the attack reaches the ports (Gbps of capacity) of their members. Moreover, IXPs have existing routing infrastructure via the route servers [63], they have experienced network management teams, and they are increasingly hosting critical infrastructure, such as root DNS and NTP servers [71].

On the data plane, Advanced Blackholing combines on demand fine-grained filtering based on layer $2-4$ header information with rate limiting. This can be done via vendor specific filters or SDN OpenFlow rules. IXP members can trigger Advanced Blackholing filters either via BGP attributes or SDN on their ports to drop or shape attack traffic. Thus, Advanced Blackholing achieves scalable scrubbing while giving feedback about the state and volume of the attack (telemetry) to the Advanced Blackholing users.

[1]Examples: DE-CIX has 25 Tbps connected capacity [21]; AMS-IX has 26 Tbps connected capacity [7]; to the best of the knowledge of the authors, the largest DDoS attack reported to date did not exceed 1.7 Tbps [57].

Our prototype, Stellar[2], relies on filtering and rate limiting of traffic and on BGP communities for signaling. We focus on the latter for the prototype to enable fast adoption in practice. Moreover, we show that Stellar requires no configuration by the IXP members, light configuration by the IXP operator, and the attack traffic is dropped at the IXP.

Thus, in summary, our major contributions are:

- We underline the need for Advanced Blackholing via a measurement study of a current RTBH service and highlight shortcomings including collateral damage, lack of granularity, and lack of honoring of RTBH signals. For example, for many attacks most of the attack traffic could have been removed by more fine-grained filters, e.g., application port specific filters, without any collateral damage. Moreover, due to the complex signaling of RTBH, to our surprise, more than two thirds of the IXP members do not react to RTBH signals.
- We introduce and present Advanced Blackholing, a new mitigation technique which keeps the advantages of RTBH while incorporating the advantages of ACL filters and the power of fine-grained filtering of Flowspec and traffic scrubbing.
- We describe the design and implementation of the Stellar system, which realizes Advanced Blackholing at a major European IXP. We report on our initial experience with it to mitigate attacks.

Overall, our novel mitigation technique Advanced Blackholing tackles DDoS mitigation while not suffering from the disadvantages of limited scalability, need for cooperation, and high costs. Thus, it can cope with increasingly popular Tbps-level attacks while the networks and services under attack continue to operate as usual.

2 RTBH LIMITATIONS

Before delving into the details of blackholing, we provide the necessary IXP background and insights on how RTBH is deployed at IXPs. Then we collect and analyze traffic flows at a large IXP during attacks. To our surprise, we notice that the majority of the IXP members do not honor RTBH signals, i.e., they do not drop attack traffic for prefixes that are tagged with blackholing communities.

2.1 IXPs 101

An IXP, see Figure 1, consists of two major components, namely a switching fabric and a route server. The switching fabric – the IXP data plane – is a layer-2 (L2) infrastructure for exchanging traffic between IXP members (see solid lines). On the control plane IXPs offer two options: (a) direct bi-lateral peering between IXP members [1, 14] enabled by the data plane, or (b) via a route server (multi-lateral peering) [34, 63]. With an increasing number of members, many medium to large IXPs offer route servers as free value-added services to their members. Route servers enable peering at scale, i.e., with a single BGP session (see dashed arrows) a member can establish peerings (exchange routes and interconnect) with all other route server users that would require a separate BGP session with each one of them.

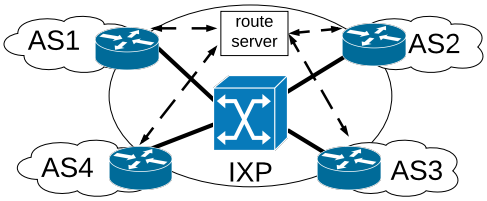


Figure 1: An IXP overview.

2.2 DDoS Mitigation at IXPs

IXP members can use the route server to steer their announcements by annotating them with BGP communities. Examples include selective advertisements to certain peers or advertisements to all/none. IXP members can also use the IXP's blackholing service to drop traffic destined to a prefix that they announce. Therefore, the prefix owner advertises the prefix to the route server along with an IXP BGP blackholing community [26]. Such communities are known to the IXP members and in many cases, this information is publicly available on the IXP's website. The number of ASes that use blackholing globally has quadrupled in the last three years, 60% of which rely on the IXP-based variant [33].

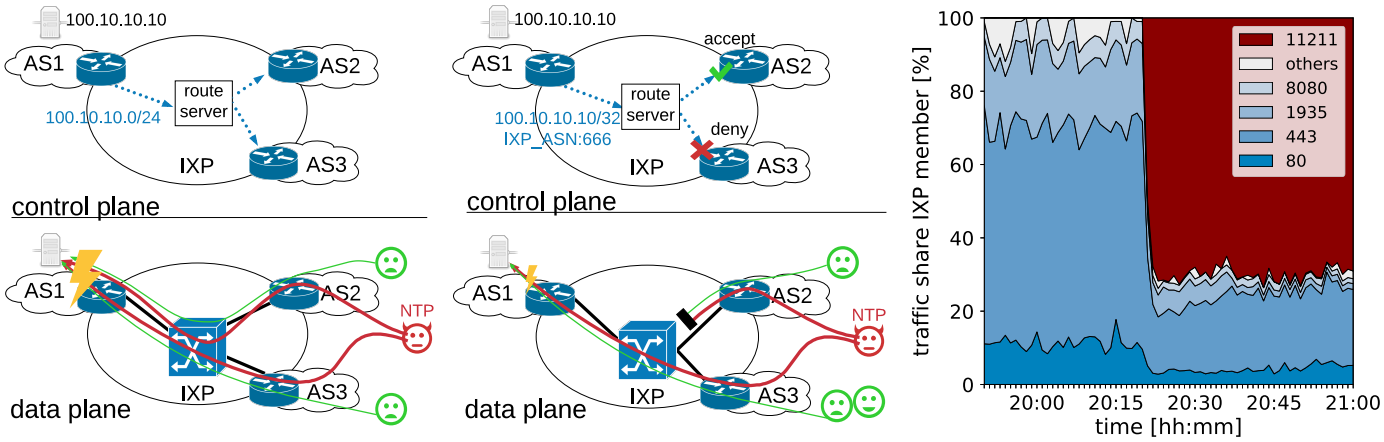
Consider the scenario in Figure 2(a). AS1 advertises the prefix `100.10.10.0/24`, see Figure 2(a) (top). Now an adversary attacks a web service that is running on IP address `100.10.10.10`, which is part of the advertised prefix. The consequence is that users from neither AS2 nor AS3 can use the service due to the overloaded network port of AS1, see Figure 2(a) (bottom). Now, AS1 uses the IXP's blackhole service by sending an update for `100.10.10.10/32` annotated with the standardized blackhole community (`IXP_ASN:666`). The route server propagates this update to the other peers, namely AS2 and AS3 as shown in Figure 2(b) (top). If an AS (e.g., AS2) accepts the announcement, the `next hop` IP is changed to the IXP's blackholing IP. This ensures that traffic to this prefix via AS2 is dropped at the IXP's null interface as shown in Figure 2(b) (bottom). Notice that this causes collateral damage because all the legitimate traffic via AS2 is also dropped. If an AS (e.g., AS3) does not honor the update, neither attack traffic nor legitimate traffic is dropped. This implies that if the majority of the attack is via AS3, the attempted mitigation fails. Recall, IXPs are typically *carrier* and *policy neutral*. They will only blackhole traffic if the owner[3] of the prefix instructs them to do so.

2.3 RTBH: Collateral Damage

RTBH is heavily used [33], but suffers from severe shortcomings [26]. To underline the extent of RTBH collateral damage, we rely on IPFIX data from one of the largest European IXPs and refer to it as L-IXP. We focus on a memcached amplification attack [64, 73] from the 29th April 2018. It lasted for several hours with traffic levels of up to 40 Gbps. Figure 2(c) shows the normalized traffic towards

[2]In astronomy, a stellar object is an object of great mass but not large enough to be characterized as a black hole. Thus, some objects escape from its attraction. This analogy describes how Advanced Blackholing differs from RTBH.

[3]Typically, the IXPs require the members to register the ownership of their prefixes in Internet Routing Registries (IRR), and check before they accept announcements of prefixes at the route server [63].



(a) Attack scenario. (b) Mitigation via RTBH. (c) Collateral damage of RTBH.

Figure 2: IXP: Attacks and their mitigation using RTBH.

the target IXP member (mainly an IPv4 /32) before and during the attack. The IP appears to host a Web service as is evident from the port usage of 443 (HTTPS) and 80/8080 (HTTP) before the attack. At 20:21 CET, we observe a sudden and huge increase in the relative volume for UDP port 11211. This is a classical signature of an amplification attack. With RTBH, *all* traffic to this IP is dropped. This includes the remaining *legitimate* Web traffic. Ideally, one would blackhole only traffic from port 11211 so that the fraction of Web traffic would return to its typical level. In this case there would be no or minimal collateral damage. Indeed, the potential of collateral damage is even worse if an IP is shared among multiple co-location services and/or across tenants, e.g., at a cloud provider.

So far, we focus on the overall traffic to an IP under attack via the IXP. Next, we compute the relative port distribution for all blackholing traffic during the two last weeks of April 2018[4], see Figure 3(a). We compare its port distribution to that of non-blackholed traffic. Significant differences are identified by using a one-tailed Welch's unequal variances t-test with significance level 0.02, i.e., a non-existing difference may only be measured with a probability of 2%. All differences are significant.

We find that L4 ports 0, 123, 389, 11211, 53, 19 are the most prominent ones to show increased traffic volume. These ports and their associated applications (NTP, LDAP, memcached, domain, and chargen) are known to be highly susceptible to (amplification) DDoS attacks [5, 20, 39, 64, 73]. The source port distribution of all other—not blackholed—traffic is very different. We notice that the UDP accounts for 99.94% of the blackholed traffic while the share of TCP is as low as 0.03%. This has to be expected, as TCP is a connection-oriented protocol. If one direction is broken, e.g., via RTBH, no TCP connection can be successful. Thus, the small fraction of TCP control packets in the blackhole can relate to a much larger potential traffic share that is likely collateral damage. After all, TCP in non-blackholed traffic accounts for 86.81%. In summary, we find that UDP and amplification attack prone ports are dominant when analyzing attack traffic and that there is a significant risk of blocking legitimate traffic.

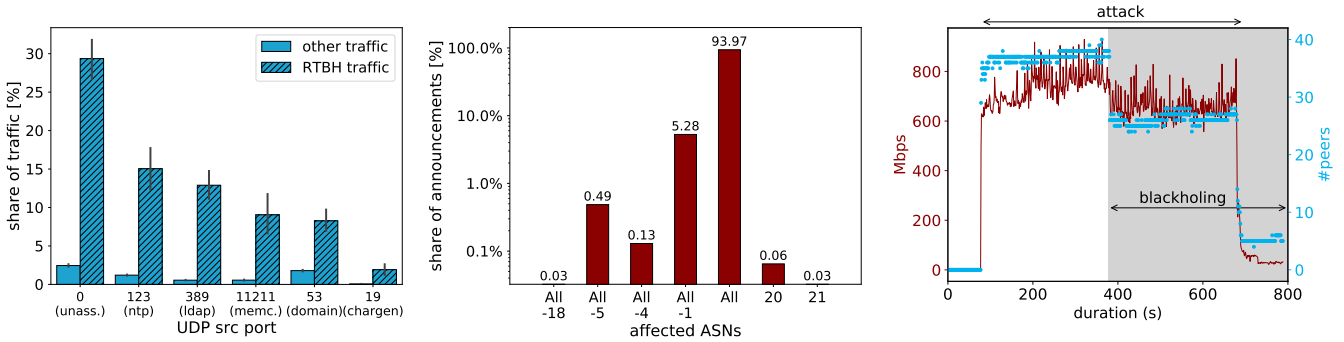
2.4 RTBH: Compliance Check

From Figure 2(c) we also notice that the attack is persistent and upon further investigation, we find that a large number of peers are involved. This motivates us to explore to what extent IXP members honor the signal to blackhole a specific prefix. We find, see Figure 3(b), that for more than 93% of the blackholing events, the prefix owner asks all route server participants (using an IXP BGP community without exceptions) to blackhole the traffic. However, almost 70% of these IXP members do not honor the blackholing community. Among the possible reasons are: (a) they choose to not participate in RTBH, (b) they do not accept updates for more specific prefixes than /24 because this requires some changes to the default configurations, or (c) they made a mistake in their configuration (fat-finger error).

To assess the effectiveness of current RTBH we perform a controlled experiment towards a test web server we operate. More specifically, we run an active experiment that attacks a single IPv4 (/32) address that we operate in our AS using a booter service [68]. To comply with measurements ethics we carefully design the experiment and take a number of measures: (a) inform and synchronize with the IXP operator about the attack, (b) take precaution that sufficient network bandwidth is available so that the likelihood of members being harmed by the targeted attack is minimized, (c) use an experimental AS with no customer traffic that we operate and is a member of the IXP, (d) utilize an unused /24 prefix that was allocated and announced only for the purpose of the experiment, (e) are prepared to shut down the experimental AS and stop the traffic by withdrawing the /24, and (f) run the experiment for a short duration (10 minutes).

The experimental AS receives routes from more than 650 IXP members (ASes) at the IXP route server via multi-lateral peering. It does not have any bilateral peerings. When launching the attack via the booter service, we request a short-duration attack (less than 10 minutes) of peak traffic of about 1 Gbps, whereby 10 Gbps is the

[4]We focus on IPv4 traffic as IPv6 blackholing traffic is less than 1% of the overall blackholing traffic. More than 98% of the blackholed prefixes are IPv4 /32 addresses.



(a) UDP source ports of blackholed traffic across RTBH events with 95% confidence intervals.

(b) Usage of policy control for RTBH at L-IXP.

(c) Active DDoS attack exposing RTBH ineffectiveness.

Figure 3: Properties and shortcomings of RTBH.

port capacity of the experimental AS at the IXP. 280 seconds after the start of the attack the experimental AS signals an RTBH route update for the /32 prefix with the IXP blackholing community to the route server. Figure 3(c) shows the result of the attack as well as the impact of the blackhole mitigation.

We observe that the DDoS attack increases the traffic levels to slightly less than 1 Gbps at its peak. Traffic is being received via quite diverse network paths from almost 40 different peers. Interestingly, RTBH has little effect on the traffic volume. It still amounts to roughly 600 to 800 Mbps. We note that the number of peers that the traffic is received from has decreased by only 25%. The reason that it does not decrease further is that the other ASes did not accept the /32 blackholing announcement. This underlines why RTBH by itself is not a sufficient DDoS mitigation technique.

3 ADVANCED BLACKHOLING CONCEPT

Given the major limitations of existing DDoS mitigation techniques, we propose Advanced Blackholing which combines the advantages of RTBH, i.e., easily dropping volumetric attack traffic, with some of the granularity of ACL filters and traffic scrubbing services. To do so, we first present the corresponding design requirements before presenting our concept.

3.1 DDoS Mitigation: Requirements

To tackle the challenges imposed by the ever increasing DDoS attacks we need to (i) improve our mitigation techniques to keep up the arms race with ever more frequent attacks and their raising traffic levels [41, 57], and (ii) remain cost-effective and with as little overhead as possible. In detail, we should meet the following requirements.

Granularity: Coarse-grained filtering is cost-effective but imposes significant collateral damage, see Section 2.3. We argue for a mechanism that can offer fine-grained filtering and operates on arbitrary, customizable source and destination packet header fields, i.e., MAC, IP, transport protocol, transport layer port, and meaningful combinations of them. While sacrificing some granularity, we can gain efficiency and cost-effectiveness. Thus, this is a desirable sweet spot.

Signaling complexity: For mitigation to be effective, it needs to be responsive, effective for any given attack, and easy to use. Current mitigation solutions may involve many stakeholders, e.g., for RTBH at IXPs all members should participate. As we also showed in Section 2.4, the level of compliance to RTBH signaling is very low. For traffic scrubbing services, contracts are necessary which may require a setup time and traffic may have to be rerouted [43]. Thus, the required signaling complexity can be significant, as many parties may have to honor more specific prefixes than used in regular routing.

Cooperation: RTBH requires cooperation among network operators to act upon receiving a prefix with a blackhole community. Flowspec inherits this requirement, and it also relies on trust and resource sharing when it is applied in the inter-domain environment [66]. Unfortunately, these requirements are hard to satisfy when networks with diverse resources as well as different or even conflicting policies and business strategies form the Internet. Thus, Advanced Blackholing should lower the necessary levels of cooperation among the involved networks and ideally not require changes in the operation of these networks.

Telemetry: A major challenge in DDoS mitigation is the victims' ability to determine when an attack is over. They have to either get direct feedback from the scrubbing service or terminate the RTBH, ACL filter, or the scrubbing which may lead to immediate congestion if the attack is still ongoing. Such probing behavior has been observed in practice [26, 33]. Thus, we argue that a well-designed DDoS mitigation system should enable the network under attack to still receive telemetry information about the status of the attack. This can, e.g., be a well-defined sub-sample of the attack traffic, which does not exceed a fixed reserved bandwidth share. Moreover, traffic statistics about the discarded traffic should be made available to inform operational decisions including terminating or triggering further mitigation actions.

Scalability: DDoS mitigation has to scale along many dimensions including (a) performance (number of attacks and volume of attack traffic), (b) filtering resources (on both the data and control plane), (c) reaction time (until mitigation takes effect), (d) configuration complexity, and (e) number of users of the mitigation technique. Hereby, understanding the hardware limitations of network devices can be a major challenge due to the closed source mentality of many network hardware vendors.

Cost: The challenge here is to meet the above design requirements while keeping costs low. Costs include but are not limited to hardware, software, configuration, management, and human resources.

3.2 Our Proposal: Advanced Blackholing

The focus of our proposal are IXPs. They are ideal locations for deploying DDoS mitigation services as they are open to innovation, have experience with DDoS services already (RTBH), are carrier and policy neutral, carry large volumes of traffic, and have a large member base already. Moreover, the IXP hardware is in principle capable of processing extra capacity and performing filtering at scale and on demand.

We propose **Advanced Blackholing**, a network attack mitigation concept for protecting IXP members (e.g., their IXP port, internal network, or Web service) from DDoS attacks. Advanced Blackholing builds upon RTBH but allows configuring fine-grained filtering and shaping rules directly in the IXP's hardware. By doing so, it reduces signaling complexity from one-to-all communication to one-to-IXP communication. Thereby, Advanced Blackholing enables better traffic filtering, a simple signaling interface at the IXP that does not need cooperation among peers, traffic telemetry for discarded packets, and a scalable architecture to handle the volume and frequency of future large-scale DDoS attacks with minimal additional cost.

To filter unwanted traffic with Advanced Blackholing, fine-grained filter rules are instantiated by the IXP on behalf of a member who owns the IP address under attack. The filtering rules, in the remainder of the paper referred to as *blackholing rules*, can be a combination of L2-L4 header information, including MAC and IP address (IPv4 and IPv6), transport protocol, or TCP/UDP port. To request a blackholing rule the member signals its request to the IXP. For this purpose, Advanced Blackholing can leverage protocols such as BGP, which enables in-band network specific inter-domain signaling, or remote customer facing APIs. Since the victim AS and IXP are the only involved parties, no further cooperation of the other members is required.

Advanced Blackholing, in contrast to RTBH, is not an all-or-nothing approach. It supports rate-limiting traffic matching a blackholing rule. Thus, valuable telemetry about the state of the attack or forensics after the incident can be collected and analyzed directly by the IXP member. We show that realizing Advanced Blackholing is possible with existing hardware deployed at IXPs. It can also be realized within an SDX [8, 37]. Both realizations ensure that the needed data and control plane resources are within the IXP's operational boundaries while offering good filter update rates and reaction times. Indeed, Advanced Blackholing scales up to the point where attacks exceed the IXP's connected capacity, e.g., 25 Tbps connected member capacity at DE-CIX Frankfurt in Summer of 2017 [21].

3.3 Advanced Blackholing: An Example

To highlight the benefits of Advanced Blackholing, Figure 4 revisits the previous IXP toy example (introduced in Section 2.2), now with an Advanced Blackholing system. On the control plane (top) Advanced Blackholing is triggered by signaling an IP (`100.10.10.10/32`) with the `ADV_BH signal`, e.g., the NTP port (`IXP:123`), via a single BGP announcement (if BGP communities are used for signaling). This signal is received by the route server which forwards the signal to the IXP's Advanced Blackholing tool-chain. Next, the desired filter/shaper is deployed within the IXP's data plane. Thus, the unwanted attack traffic, e.g., an NTP amplification attack, is filtered but benign web traffic, e.g., HTTP/HTTPS is still forwarded. Consequently, congestion at AS1's IXP port and interior network are resolved and the Web service is restored. The DDoS amplification attack is mitigated.

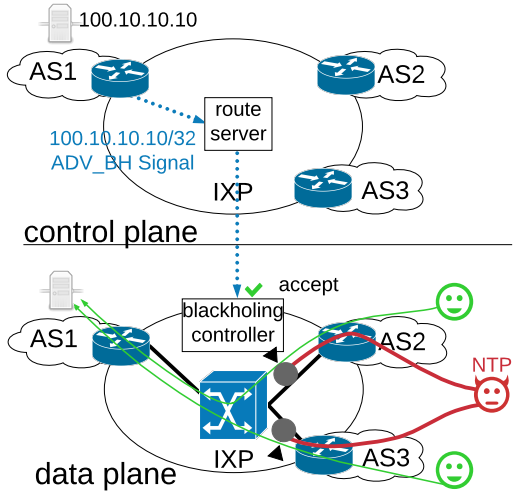


Figure 4: Example scenario of Advanced Blackholing.

Notice that with Advanced Blackholing the risk of collateral damage is minimized. Moreover, Advanced Blackholing does not require the cooperation of all IXP members (in this case AS2 and AS3). It suffices that the IXP's blackholing controller honors the /32 announcement.

4 STELLAR: ADVANCED BLACKHOLING SYSTEM

We build Stellar to realize the benefits of Advanced Blackholing. To encapsulate functionality, we discuss alternative design options and highlight the advantages of the chosen implementations. Stellar's design consists of three layers: *signaling*, *management*, and *filtering*, see Figure 5. *Signaling* handles the IXP members' signal to discard traffic according to well-defined filter rules (the blackholing rules) and forwards them to the blackholing manager. In addition, signaling assures resilience (fall back, member state), consistency (authentication), and security (authorization). The blackholing manager—as part of *management*—holds the state of all signaled blackholing rules from all IXP members and resolves existing conflicts. At the same time, the blackholing manager compiles the blackholing rules to IXP hardware specific configurations. Those are deployed on the distributed switching fabric to drop or shape traffic (*filtering*). In this way, we encapsulate hardware and vendor specific aspects including configuration, parameters, and hardware capabilities.

Stellar's architecture respects the fact that IXPs per se offer L2 connectivity only. While member ASes communicate reachability

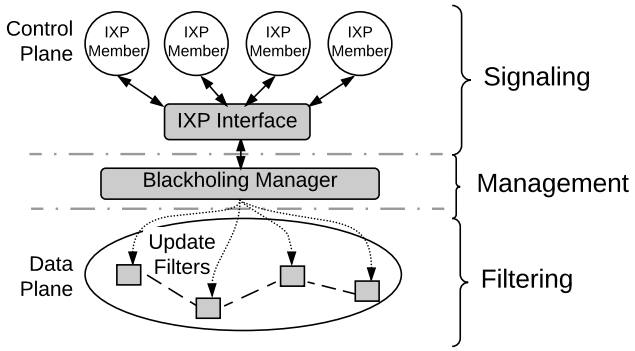


Figure 5: Stellar architecture for Advanced Blackholing.

via BGP (L3) they cannot directly interact with the IXP's forwarding infrastructure (IXP data plane) via L3. Such interactions are only feasible via other member's routers or the route server (control plane). Our proposed system architecture respects this design and augments it with a well-defined cross-layer interface to enable members to use the IXP's hardware for the purpose of network attack mitigation.

4.1 Implementation Constraints

Stellar's implementation is subject to constraints imposed by both IXP members as well as IXP operators.

4.1.1 IXP Member Constraints. To facilitate fast and seamless adoption of Stellar it is imperative to keep the entrance barrier low. Thus, as Advanced Blackholing extends the capabilities of RTBH, the signaling layer should be implemented in a way that allows IXP members to continue using their existing blackholing toolchains. Moreover, no major configuration updates or system upgrades should be required. The current state-of-the-art signaling for RTBH is to annotate BGP announcements with the standardized blackholing community [47]. Consequently, a signaling mechanism on top of BGP is favorable but challenging given BGP's limited expressiveness and extensibility. Yet another constraint is BGP's security design, which comes with a number of shortcomings, namely, prefixes are hijacked [69], limited adoption of prefix authentication [35], lack of cryptographic integrity and authenticity over BGP attributes [72], AS paths (or any other attribute) are modified for traffic engineering [61], and routing tables are flooded with more specifics [51].

4.1.2 IXP Operator Constraints. To be economically affordable, new services should be seamlessly integrated in the existing IXP system landscape. This includes interoperability with existing configuration management, monitoring, and even provisioning processes to keep additional operational costs (OPEX) low. For performance reasons filtering implementation in hardware is obligatory. Yet most hardware has limits, e.g., number of filters per port or line card, and kind of filters. Still, it is economically critical to exploit existing hardware resources and to rely only on new or additional hardware if inevitable, to keep capital expenditure (CAPEX) low. Even adding an edge switch at a major IXP already easily exceeds several millions of euros. Thus, filtering should be realized using the capabilities of existing hardware. Moreover, as IXPs are critical infrastructure [23, 27] their availability is indispensable. Traffic forwarding has to be guaranteed at all times. This constraint implies that: (i) filtering has to be highly resilient and include fallbacks towards simple forwarding of all traffic; (ii) management has to do "admission control" (limit the number of blackholing rules) to ensure that the hardware resource limitations of the IXP's forwarding hardware are respected.

4.2 Implementation Choices

While there are many implementation options, the choices for the signaling interface and the filter implementation based on the above constraints are critical for Stellar.

4.2.1 Signaling Interface. The member ASes communicate with Stellar through a signaling interface of the IXP. It must enable a member to express blackholing rules, signal new ones, update existing ones, withdraw them, and implicitly withdraw them (in case of system failures). Thus, the choice of how to interact with the IXP is essential. Among the candidates are: remote (i) *Application Programming Interface (API)*, (ii) *BGP-4*, (iii) *BGP with Flowspec*, or (iv) *BGP with extended communities*.

APIs: are a well-established way to interact across system boundaries and can be designed to scale, perform, be highly available, and secure. However, in inter-domain settings public remote APIs are rarely used since they introduce potential points of failures and introduce complexity. Moreover, APIs are alien to network engineers and their default toolbox.

BGP-4: Network engineers favor in-band protocols, here BGP. However, BGP-4 [62] per se does not offer the capabilities to signal anything more specific than a /32 prefix (single IPv4 address) or /128 for IPv6 respectively.

Flowspec: To tackle the above limitations, numerous extensions were introduced including Flowspec – "Dissemination of Flow Specification Rules" [52]. Flowspec defines the BGP Network Layer Reachability Information (NLRI) encoding format to distribute fine-grained L2-L4 traffic flow specifications. At first glance, Flowspec appears to be a perfect candidate as signaling interface, in particular, as inter-domain cooperation and coordination for traffic filtering is one of the highlighted use cases [52]. Flowspec is a useful tool for bilateral peerings among IXP members, but it also requires cooperation and inherits the signaling complexity of blackholing which can render RTBH ineffective. Furthermore, in practice, Flowspec announcements consume scarce hardware resources of routers that are not under the control of the owner and require non-trivial resource monitoring. Thus, the adoption of Flowspec in the inter-domain environment is questionable [11, 66] even though most hardware vendors support it since 2014. For the IXP specific setting, additional limitations are: (a) lack of Flowspec implementation for the route server software stack, and (b) no IPv6 standardization [54]. Therefore, we decided against Flowspec for Stellar.

BGP extended communities: An alternative in-band BGP interface that allows communication of meta information to peers is BGP communities. The "BGP Communities Attribute" [13] introduced in 1996 allows the tagging of routes with numeric values and is heavily used by ASes, e.g., for traffic engineering [61], network troubleshooting [29], location information [32], and RTBH [47]. BGP communities are heavily used in inter-domain routing and

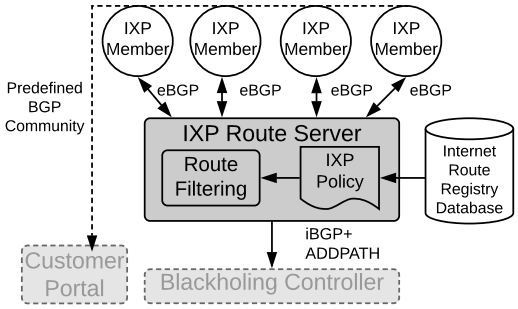


Figure 6: Signaling layer implementation.

IXP settings, and are supported by all route server implementations as of today. The communities themselves are only tags without any semantics. Some predefined values have been agreed upon—the well-known communities—while others are AS or IXP specific. We choose BGP extended communities [67] for signaling since extended communities provide a sufficiently large numbering space and allow us to define a distinct community namespace for blackholing rules. This choice satisfies all constraints imposed by IXP members and IXP operators.

4.2.2 Filtering. The filtering layer must ensure that *all* port specific filter rules including blackholing rules are applied to the traffic. Among the possible choices are: (i) *traffic diversion systems*, *filtering with SDN hardware*, and *filtering with vendor specific ACLs*.

Traffic diversion systems: either powerful computational clusters [78] or specialized hardware [58], require a substantial investment into equipment typically not yet integrated in today's IXP infrastructure.

SDN hardware: in principle, offers both the ability to configure via OpenFlow [53] or P4 [10], and realize filters with the match-action abstraction efficiently [12]. Moreover, with per flow counters it is possible to gather statistics about filtered as well as forwarded traffic and, thus, provide telemetry feedback. In principle, SDN based solutions are a good option. However, at this point most IXPs are not yet realized as SDX and, accordingly, additional hardware investments would be mandatory. Even if an IXP decided to make such an investment, there is a lack of hardware support by vendors currently.

Filtering with vendor specific ACLs: To achieve a cost-effective deployment, we need to rely on vendor specific router ACL features which are commonly used to realize Quality of Service (QoS) policies. Such features are offered by almost all switching/routing hardware vendors under a different name, e.g., Cisco offers Extended ACLs [17], Juniper offers Firewall Filters [44], and Alcatel Lucent offers QoS Policies [6]. These ACL features enable dropping, shaping for telemetry feedback, and forwarding on the data plane.

4.3 Signaling Implementation

In the following, we describe the implementation of Stellar's signaling layer, i.e., how IXP members signal Advanced Blackholing and express the concrete blackholing rule by means of BGP, and how rules are validated and forwarded to the blackholing controller, see Figure 6.

The central element of Stellar's signaling is the IXP's route server, i.e., the member facing IXP interface which Stellar uses to collect signals from the members. The route server is placed in the IXP peering LAN and maintains BGP sessions with all IXP members (demanded by terms and conditions). The route server collects the signals from all members and performs *route filtering* on import to assure routing hygiene based on an *IXP policy*. The IXP policy ensures that each member can only announce prefixes that are not in conflict with Internet Route Registry databases (IRRs) [55], BOGONS [28], and RPKI validation [38]. This does not interfere with prefix delegations which can be represented by all validation databases. The benefit of integrating Stellar with the route server is that we immediately inherit routing hygiene as well as its high availability design (IXPs multilateral peerings rely on them), member state monitoring (offered by BGP), and debugging interfaces. For instance, members can rely on looking glasses for debugging [31, 46].

Notably, as opposed to RTBH, the route server does not reflect signals back to the other members, i.e., signaling only involves the IXP member under attack and the IXP. Instead, the southbound interface of the signaling layer connects the route server with the *blackholing controller* and consumes the signals. From the route server perspective the blackholing controller is just an IXP member router. Nevertheless, there are three main differences. First, the blackholing controller uses internal BGP (iBGP) instead of external BGP (eBGP) to ensure forwarding of updates. Consequently, no own AS number is required. Second, the blackholing controller is passive, i.e., it only collects announcements but never announces any routes. Third, the blackholing controller uses BGP's recently standardized ADD-PATH capability [76] to bypass BGP best path selection at the route server. This is essential for a number of corner cases, e.g., to be able to honor the same prefix from different member ASes with diverging blackholing rules.

For a member to trigger Stellar to filter traffic destined to a prefix it needs to annotate the BGP route announcement with a specific BGP extended community. This community encodes a reference to a specific blackholing rule, e.g., drop traffic from UDP source port 123 (NTP), and can be predefined by the IXP or by the IXP member via a *customer portal* (self-service portal). Currently, the IXP offers a shared set of predefined blackholing rules for common attack patterns but custom blackholing rules can be defined as well.

4.4 Blackholing Management Implementation

Stellar's management layer consists of two components: A *blackholing controller* and a *network manager*, see Figure 7. The blackholing controller is responsible for tracking blackholing rules and their changes as signaled by the route server on behalf of the members. The network manager realizes the blackholing rule changes by computing the hardware specific configuration changes.

The blackholing controller implements a BGP parser and a BGP processor. The first maintains the iBGP session with the route server and parses the incoming message stream. The latter processes the semantics of BGP messages and stores the announced routes in a Routing Information Base (RIB). Next, the controller calculates differences between RIB snapshots. Essentially, these differences represent a set of abstract, i.e., still hardware-independent, configuration changes that must be applied to the network to reflect all

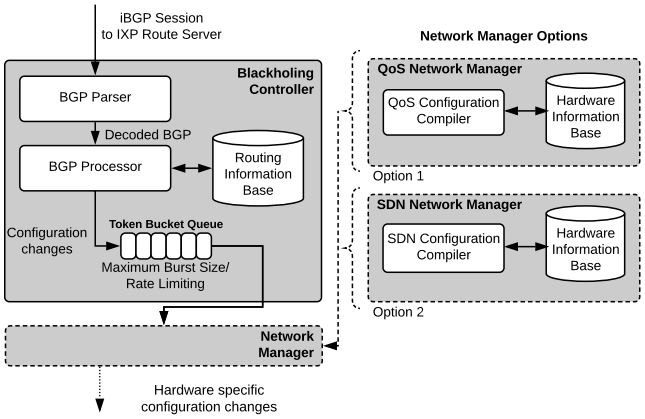


Figure 7: Management layer implementation.

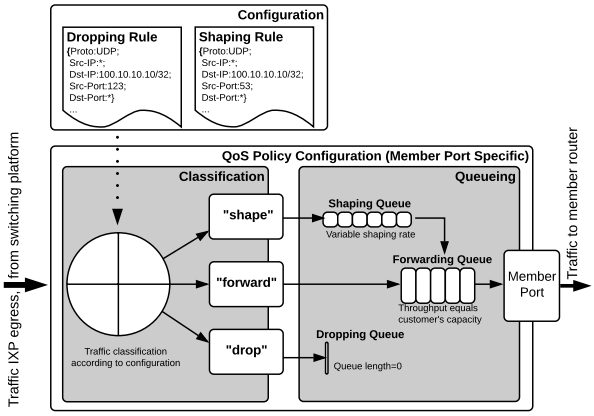


Figure 8: Filtering layer impl. via QoS policies.

requested blackholing rules. These configuration changes are then forwarded via a software queue to the network manager.

The network manager compiles the abstract configuration changes in hardware specific ones. Notice that a high-level configuration change may require multiple low-level configuration changes on several switches. To limit the number of configuration changes within any time interval to a rate that is manageable by the switch hardware, the queue uses a Token Bucket algorithm. This ensures that the configurable Maximum Burst Size (MBS) and a reasonable long-term rate limit is never exceeded.

The requested configuration changes are dequeued by the network manager. On an abstract level, the network manager compiles abstract configuration changes as provided by the blackholing controller into hardware-specific ones. Then, the changes are deployed on the IXP's forwarding hardware (data plane) while respecting the hardware limitations of the actual IXP infrastructure. As mentioned above, there are multiple different feasible implementations for the network manager. Indeed, the feasible set depends on the capabilities of the existing IXP hardware. We realized two solutions, one using vendor-specific ACL filters to realize QoS policies and an SDN-based solution. Below, we focus on the former. For details on the latter, we refer the reader to our recent demo [25] based on the SDX platform [37].

Each network manager has access to a description of the hardware limitations via a *hardware information base*. This includes, e.g., the number of allowed QoS policies per port or the number of OpenFlow rules that may be installed on a switch. Using this information, the *configuration compiler* can ensure that the limitations are respected.

4.5 Filtering Implementation (QoS)

Stellar's filtering component depends primarily on the specific IXP hardware deployed. Our implementation is targeting deployment at L-IXP, a major European IXP, with more than 800 members, more than 6 Tbps peak traffic, and distributed across more than 20 IXP PoPs in the metropolitan area it operates in. The IXP platform is not SDN-based and we thus use the implementation option that is available, namely, QoS policies.

Recall that QoS policies can be applied either on *all* IXP member ingress ports or on the *destination* IXP egress port. In the former case the traffic is dropped or shaped upon entering the IXP's platform. In the latter case, it is dropped or shaped upon leaving the IXP's platform. We opt for the latter, mainly to minimize the complexity of the configuration procedure. Configuration changes are minimized, as only the destination port has to be updated rather than the configurations of all other $(n-1)$ ports. Moreover, an update from one IXP member only causes changes to the port configuration of exactly this IXP member. Thus, causality is maintained. This implementation option is currently possible at L-IXP as it has (at least today) enough capacity to carry the network attack traffic across the IXP platform to the bottleneck IXP member ports. Shaping traffic on egress also has the advantage that telemetry information is made available to the member IXP. Nevertheless, moving egress filters to ingress filters may be a good choice for future work and to enable deployment at smaller IXPs where the platform capacity is a bottleneck.

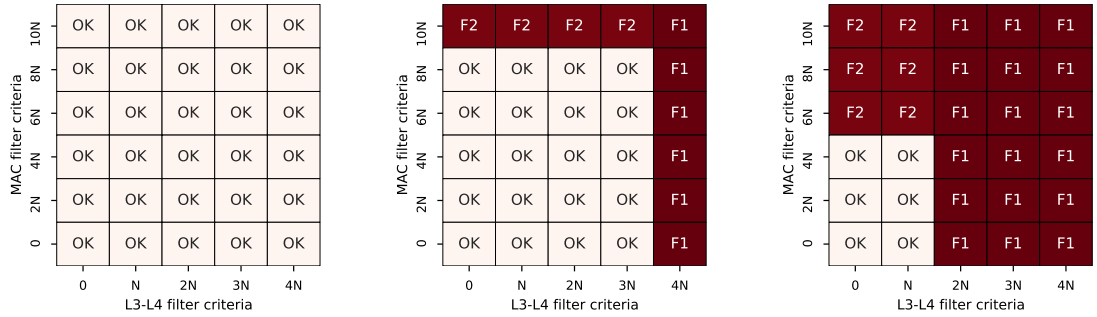
An overview of how the QoS policies are realized is given in Figure 8. The policies are IXP port specific QoS dropping and shaping rules coming from the network manager. These are used to configure the ports, which classify the packet streams. Our configuration distinguishes between three options, "shape", "forward", and "drop". Packets tagged with drop are redirected to a zero-length queue for immediate dropping. Packets tagged with forward are put in the egress queue of the member port. Packets tagged with shape are put in a queue which is shaped according to the specification. Packets passing the shaping queue enter the forward queue.

5 EVALUATION

In this section we first explore how well Stellar scales in the L-IXP test lab setting and then present how Stellar, when deployed at the L-IXP, is able to handle network attacks.

5.1 Scaling and Performance

To ensure that Stellar can be deployed at the IXP, it is essential that the limits of the IXP's hardware are respected. Thus, we check in a lab setup: (a) how Stellar scales with an increasing number of filters and ports, and (b) if the configuration update frequency limits are sufficient to support Stellar. Our approach mainly aims to measure the effect of Stellar's properties on Ternary Content-Addressable



(a) 20% of IXP member ASes. (b) 60% of IXP members ASes. (c) 100% of IXP members ASes.

Figure 9: Stellar scaling limits by IXP member adoption rate: (a) 20% (2× of RTBH users today), (b) 60%, and (c) 100%; N is the 95^{th} percentile of the currently observed parallel RTBHs by any member on any port; OK: no hardware resource depletion, $F1/F2$: hardware limit of L3-L4/MAC filters exceeded; an increased adoption rate leads to less available filters per port.

Memory (TCAM) [45]. The TCAM is used to implement matching header information in hardware. Its size and update behavior constitute the main resource bottleneck of Stellar. We do not measure traffic bottlenecks as it is difficult to generate enough traffic and the IXP's vendor explicitly guarantees support of the relevant features in hardware at line rate (verified by observing no impact on hardware's CPU utilization).

Our test setup consists of an IXP Edge Router[5] (ER), a hardware accelerated packet generator, a resource monitor, and Stellar. The ER is configured with a production configuration of more than 350 member ports, which corresponds to the ER with the largest port density. The evaluation of the scaling limits of the hardware aims to verify whether there are both sufficient resources per member port (by increasing the blackholing rules per port) as well as sufficient system wide hardware resources (by increasing the total number of QoS policies), as Stellar can exhaust either one. We generate a number of blackholing rules per port. Each rule uses MAC filters to filter traffic from a specific AS (as is needed for RTBH policy control) and a random subset of L3-L4 rules to filter traffic (the Advanced Blackholing capabilities). We then increase the fraction of ASes (ports) using Stellar as well as the number of blackholing rules per port.

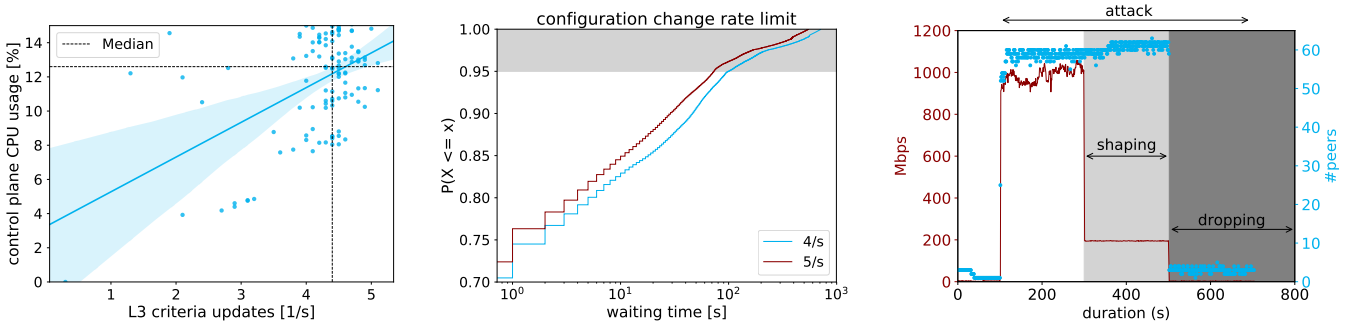
Figure 9 shows the limits for three different percentages of IXP member ASes actively using Stellar. We start at 20%, twice the percentage of IXP member ASes that currently use RTBH daily, and go via 60%, a high adoption rate, to 100% IXP member adoption rate. For each adoption rate we show whether there are sufficient hardware resources for increasing numbers of QoS filters along both dimensions, namely, MAC filters (shown on y-axis) and L3-L4 filters (shown on x-axis). OK corresponds to sufficient resources; F1 and F2 correspond to insufficient resources meaning respectively that the total number of filter criteria for QoS policies (L3-L4) is exceeded or the maximum number of MAC filters per port (L2) is exceeded. For each plot, we increase the number of MAC filters from 0 to $10N$ and the number of L2-L3 filters from 0 to $4N$. We choose N to be the 95^{th} percentile of the number of currently observed parallel RTBHs on any port by any IXP member. For a more detailed analysis of current RTBH usage, see [26].

[5]IXPs often deploy routers but configure them to act as switches.

Not surprisingly, the feasible region —e.g., the region marked OK— decreases when the IXP members are using more blackholing rules or if more IXP member ASes are using Stellar. Let us consider the starting point, Figure 9(a): While the number of Stellar users corresponds to twice the current (June 2018) users of RTBH, they each use more rules to take advantage of novel Advanced Blackholing capabilities. We do not see any scalability limits. Next, see Figure 9(b), we presume an increased adoption rate to 60% of the IXP members. There still is a huge headroom for the number of MAC filters (8N) and the number of L2-L4 filters (3N). With an adoption rate of 100%, see Figure 9(c), the safety margin decreases but still is substantial. Note, the experiments presume that every single IXP member increases their number of parallel blackholing rules at the same time (a stretch test). Thus, we conclude that Stellar can be deployed without exhausting the IXP platform filtering resources.

Next, we ask if the hardware can sustain the update frequency imposed by Advanced Blackholing. A review of Stellar and the IXP's ER hardware shows that the limiting factor is the ER's CPU resources. Notably, the ER's control plane runs a real-time OS and the current configuration imposes a hard CPU limit of 15% for configuration tasks. To check if this suffices, we measure the ER's CPU resources while increasing the number of blackholing rule addition and removal operations. Figure 10(a) shows how many rule updates were processed during a five-second interval (scaled to 1 second) vs. the corresponding CPU usage. With a 15% CPU usage, the ER can handle a median of 4.33 rule updates per second.

To predict how long it takes until a blackholing rule takes effect, we perform a controlled experiment on top of the Token Bucket Queue of the blackholing controller: We enqueue configuration changes generated from the traces of L-IXP's RTBH service. These configuration changes are dequeued with a rate of 4/5 per second, respectively (to mimic the median rule update rate of 4.33). Figure 10(b) shows how long each update stays in the queue, i.e., the time from blackholing signal to configuration. 70% of all configuration changes are well below 1 second and the 95^{th} percentile is below 100 seconds. Given the setup overhead of alternative solutions, e.g., TSS and ACLs, we consider these results to be acceptable from an IXP member perspective.



(a) Control plane CPU usage vs. L3 criteria update rate (linear regression, 95% confidence interval).

(b) Required queuing for different announcement frequencies.

(c) Active DDoS attack to assess Advanced Blackholing effectiveness.

Figure 10: Evaluation: (a) scalability, (b) blackholing manager queuing rate, and (c) DDoS experiment with Stellar.

5.2 Functionality

To validate the functionality of Stellar we conduct extensive lab experiments where we use a hardware accelerated traffic generator to simulate legitimate as well as attack traffic and a high-speed measurement system on the member's port to confirm Advanced Blackholing's effectiveness. The traffic generator allows us to construct traffic flows with certain properties, for instance a common UDP source port, at high bandwidth, e.g., to generate NTP, DNS, and benign traffic with a bandwidth of 10 Gbps. This traffic is then directed to different IP addresses within an IXP member AS via the monitored member port—which has a capacity of 1 Gbps and is immediately congested. Our experiments confirm that the ER with Stellar behaves as expected: Flows redirected to a *dropping* queue are not forwarded to the IXP member, while flows redirected to a *shaping* queue share the shaping queue's rate limit. Flows that are forwarded share the *forwarding* queue's rate limit. By redirecting the NTP or DNS flows to the dropping or the shaping queue, the benign traffic flow passes the port untouched, for each targeted IP address.

5.3 Internet Experiments

To demonstrate the capabilities of Stellar at L-IXP we repeat the experiment from Section 2.4 (Figure 3(c)). We use the same booter service as before with the same precautions to launch an active attack against a single IP address in our experimental AS which peers at the IXP. However, this time we use Advanced Blackholing via Stellar. The mitigation—started 200 seconds after the attack—is highly successful, see Figure 10(c). In more detail, the DDoS attack starts at 100 seconds and causes an immediate increase to about 1 Gbps of attack traffic from about 60 peers. The attack is an NTP reflection attack and, thus, we mitigate it by blocking or shaping traffic from UDP source port 123. Hence, we send a BGP update for the IP (`/32` prefix) tagged with BGP community `IXP:2:123`. Hereby, **2** refers to UDP source traffic and **123** to port 123. 200 seconds into the attack we trigger an update, which shapes the traffic to a rate limit of 200 Mbps for telemetry purposes. We see that the traffic level indeed quickly drops to 200 Mbps. Note, the number of peers remains constant. 200 seconds later, we signal Stellar to drop all UDP and NTP traffic. This reduces traffic close to zero and reduces the number of peers. The remaining minimal traffic deviation are mainly ARP traffic from some peers. Thus, Stellar is convincingly able to mitigate the attack which RTBH (see Figure 3(c)) in contrast was unable to achieve. Stellar reduces the attack traffic and the number of peers that send attack traffic, to a shaped minimum or close to zero.

6 DISCUSSION

Applicability in other network scenarios: We demonstrate the efficiency of Advanced Blackholing in an IXP setting by implementing and evaluating Stellar at one of the largest IXPs. Nonetheless, Advanced Blackholing can also be applied to other settings, e.g., within an Internet Service Provider (ISP) or any network platform that handles large numbers of flows. To realize Stellar, the main ingredients are a filtering mechanism to efficiently discard traffic and a signaling interface. In an ISP context this can be the top-level route reflector or the SDN controller in a data center. Thus, we argue that Stellar (by using alternative options) is deployable in other settings as well, and as part of our future work we will explore these.

Improved utilization: Advanced Blackholing allows IXP member ASes to regain control of their hardware resources, in particular, their port capacity. Since attack traffic is dropped before using the member ports' capacity at the IXP egress, IXP members do not need to over-provision to cope with volumetric attacks and the imposed congestion decreases. Rather, the attack traffic is absorbed by the IXP with its Tbps spare capacity. This is feasible, as IXPs often deploy large edge routers but configure them to act as switches, thereby under-utilizing available hardware options. With Stellar these resources are used to offer a compelling new service. Similar observations apply to ISPs and data center networks, where it is again possible to use available resources to drop traffic early.

Combining Advanced Blackholing with other solutions: Advanced Blackholing can be combined very successfully with other mitigation techniques, in particular, traffic scrubbing services. These services can be used to signal Advanced Blackholing to drop traffic with specific features efficiently. Thus, attacks with known patterns can be dropped at no cost. This option frees resources for expensive deep packet inspection and machine learning processing used by traffic scrubbing services to extract patterns of yet unknown attacks. Moreover, Advanced Blackholing can be used to only send

a limited volume of traffic to the scrubbing service. Thus, Advanced Blackholing can drastically reduce the cost of scrubbing services without sacrificing their efficiency. The former is possible as the high cost of traffic scrubbing services is caused by first carrying all traffic to their premises and, then, applying the filters within their scrubbing clusters—a very expensive constellation. With Stellar, we can dynamically determine which and how much (sample attack traffic) reaches a monitor facility, e.g., a scrubbing center. Stellar together with deep packet inspection of attack traffic can be used to, e.g., infer attack signatures or an attack start/end.

Limitations: Advanced Blackholing combines the scalability of blackholing with fine-grained filtering of ACL and scrubbing solutions to combat many but not all attacks. Thus, Advanced Blackholing can often mitigate volumetric *network* attacks saturating the target's network resources, as well as attacks on specific protocols that consequently deny specific protocols or services. Still, it is not designed to handle semantic attacks, i.e., low bandwidth attacks exploiting specific operating system or application vulnerabilities, e.g., a kernel or Web server bug. To tackle these, other techniques, including deep packet inspection and machine learning approaches are promising candidates. However, if there is a specific L2-L4 signature of the attack, Advanced Blackholing can even mitigate these.

7 CONCLUSION

Network attacks are more frequent and voluminous than ever. Recent attacks have taken everybody by surprise as attack traffic volume is in the order of Tbps, and still growing to set new records. At this scale, attack mitigation techniques such as traffic scrubbing and ACL are either inefficient (high set up time, high resource requirements) or are prohibitively expensive. Flowspec, a popular intra-domain attack mitigation technique, relies on trust, cooperation, and sharing of resources among different networks when deployed in the inter-domain environment. Unfortunately, these requirements are hard to satisfy when networks with diverse resources as well as different or even conflicting policies and business strategies form the Internet. A cheap and highly scalable mitigation technique is blackholing. Unfortunately, our measurement-driven analysis shows that blackholing suffers from severe shortcomings, namely, collateral damage as it drops all traffic to an attack destination (including legitimate traffic), and has high signaling complexity that limits effectiveness.

In this paper we argue that a new attack mitigation technique is needed which inherits the scalability and low cost of blackholing, does not require cooperation of network operators, while providing fine-grained filtering, simple signaling, and telemetry capabilities of expensive techniques such as traffic scrubbing. We propose Advanced Blackholing, which satisfies all these desired properties. We design and implement Stellar, and operate it at a large IXP to make the benefits of Advanced Blackholing available to the hundreds of the IXP member ASes. IXP member ASes can utilize Stellar with minimal configuration changes to their setup, contrary to traditional blackholing (RTBH) and other mitigation techniques.

Our evaluation shows that Stellar scales well even if Advanced Blackholing requests and attack traffic increases at very high levels. Stellar also allows fast responses and is highly configurable, e.g., it provides traffic shaping to give telemetry feedback on the attack to its users. We are currently deploying Stellar as a service at a large IXP and we plan to install it at other IXPs in the near future.

ACKNOWLEDGMENTS

We would like to thank the CoNEXT 2018 anonymous reviewers and our shepherd, Ítalo Cunha (Universidade Federal de Minas Gerais), for their constructive comments, as well as our colleagues Daniel Kopp, Daniel Spierling, and the entire development team for their ongoing implementation and deployment efforts.

This work and its dissemination efforts were supported in part by the European Research Council (ERC) grant ResolutioNet (ERC-StG-679158) and by the Leibniz Prize project funds of DFG–German Research Foundation: Gottfried Wilhelm Leibniz-Preis 2011 (FKZ FE570/4-1).

REFERENCES

[1] B. Ager, N. Chatzis, A. Feldmann, N. Sarrar, S. Uhlig, and W. Willinger. Anatomy of a Large European IXP. *ACM SIGCOMM* (2012).
[2] Akamai. State of the Internet Security Report (Q3 2016). https://www.akamai.com/us/en/multimedia/documents/state-of-the-internet/q3-2016-state-of-the-internet-security-report.pdf. (2016).
[3] Akamai. State of the Internet Security Report (Q4 2017). https://www.akamai.com/us/en/multimedia/documents/state-of-the-internet/q4-2017-state-of-the-internet-security-report.pdf. (2017).
[4] Akamai. Prolexic Technologies by Akamai. https://www.akamai.com/us/en/cloud-security.jsp. (2018).
[5] Akamai. State of the Internet Security Report (Attack Spotlight: Memcached). https://www.akamai.com/us/en/multimedia/documents/state-of-the-internet/soti-summer-2018-attack-spotlight.pdf. (2018).
[6] Alcatel Lucent. QoS Policies. https://infoproducts.alcatel-lucent.com/html/0_add-h-f/93-0077-HTML/7750_SR_OS_QoS_Guide/QoS.html. (2018).
[7] AMS-IX. Annual Report. https://ams-ix.net/annual_report/2017. (2017).
[8] G. Antichi, I. Castro, M. Chiesa, E. L. Fernandes, R. Lapeyrade, D. Kopp, J. H. Han, M. Bruyere, C. Dietzel, M. Gusat, A. W. Moore, P. Owezarski, S. Uhlig, and M. Canini. ENDEAVOUR: A Scalable SDN Architecture for Real-World IXPs. *IEEE Journal on Selected Areas in Communications* 35, 11 (2017).
[9] M. Antonakakis, T. April, M. Bailey, M. Bernhard, E. Bursztein, J. Cochran, Z. Durumeric, J. A. Halderman, L. Invernizzi, M. Kallitsis, D. Kumar, C. Lever, Z. Ma, J. Mason, D. Menscher, C. Seaman, N. Sullivan, K. Thomas, and Y. Zhou. Understanding the Mirai Botnet. *USENIX Security Symposium* (2017).
[10] P. Bosshart, D. Daly, G. Gibb, M. Izzard, N. McKeown, J. Rexford, C. Schlesinger, D. Talayco, A. Vahdat, G. Varghese, and D. Walker. P4: Programming Protocol-independent Packet Processors. *ACM CCR* 44, 3 (2014).
[11] K. Carriello. Arm Yourself Against DDoS Attacks: Using BGP Flow Specification for Advanced Mitigation Architectures. http://forum.ix.br/files/apresentacao/. (2017).
[12] P. Chaignon, K. Lazri, J. François, T. Delmas, and O. Festor. Oko: Extending Open vSwitch with Stateful Filters. *ACM SOSR* (2018).
[13] R. Chandra, P. Traina, and T. Li. BGP Communities Attribute. IETF RFC 1997. (1996).
[14] N. Chatzis, G. Smaragdakis, A. Feldmann, and W. Willinger. There is More to IXPs than Meets the Eye. *ACM CCR* 45, 5 (2013).
[15] M. Chiesa, C. Dietzel, G. Antichi, M. Bruyere, I. Castro, M. Gusat, T. King, A. Moore, T. Nguyen, P. Owezarski, S. Uhlig, and M. Canini. Inter-domain Networking Innovation on Steroids: Empowering IXPs with SDN Capabilities. *IEEE Communications Magazine* 54, 10 (2016), 102–108.
[16] Cisco. Remotely Triggered Black Hole Filtering - Destination Based and Source Based. http://www.cisco.com/c/dam/en_us/about/security/intelligence/blackhole.pdf. (2005).
[17] Cisco. Configure Commonly Used IP ACLs. https://www.cisco.com/c/en/us/support/docs/ip/access-lists/26448-ACLsamples.html. (2018).
[18] Cisco. Implementing BGP Flowspec. https://www.cisco.com/c/en/us/td/docs/routers/asr9000/software/asr9k_r5-2/routing/configuration/guide/b_routing_cg52xasr9k/b_routing_cg52xasr9k_chapter_011.html. (2018).
[19] Corero Network Security. Corero DDoS Trends Report (Q2-Q3 2017). https://www.corero.com/resources/reports/2017-ddos-trends-report. (2017).
[20] 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. *ACM IMC* (2014).
[21] DE-CIX. Connected capacity in Frankfurt exceeds 25 Terabits. https://www.de-cix.net/en/news-events/news/connected-capacity-exceeds-25-terabits.

(2017).
[22] DE-CIX. DE-CIX Blackholing – Fight DDoS Attacks. http://www.de-cix.net/products-services/de-cix-frankfurt/blackholing/. (2018).
[23] Department of Homeland Security. Communications Sector-Specific Plan: An Annex to the National Infrastructure Protection Plan. https://www.dhs.gov/sites/default/files/publications/nipp-ssp-communications-2010-508.pdf. (2010).
[24] Deutsche Telekom. AS3320 BGP Communities. https://onestep.net/documents/AS3320_BGP_Communities_v1.1.pdf. (2005).
[25] C. Dietzel, G. Antichi, I. Castro, E. Fernandes, M. Chiesa, and D. Kopp. SDN-enabled Traffic Engineering and Advanced Blackholing at IXPs. *ACM SOSR* (2017).
[26] C. Dietzel, A. Feldmann, and T. King. Blackholing at IXPs: On the Effectiveness of DDoS Mitigation in the Wild. *PAM* (2016).
[27] European Union Agency for Network and Information Security. Critical Infrastructures and Services, Internet Infrastructure: Internet Interconnections. http://enisa.europa.eu/activities/Resilience-and-CIIP/critical-infrastructure-and-services/inter-x/. (2010).
[28] N. Feamster, J. Jung, and H. Balakrishnan. An Empirical Study of "Bogon" Route Advertisements. *ACM CCR* 35, 1 (2005).
[29] K. Foster. Application of BGP Communities. *The Internet Protocol Journal* 6, 2 (2003).
[30] D. Gillman, Y. Lin, B. Maggs, and R. K. Sitaraman. Protecting Websites from Attack with Secure Delivery Networks. *IEEE Computer Magazine* 48, 4 (2015).
[31] V. Giotsas, A. Dhamdhere, and kc claffy. Periscope: Unifying Looking Glass Querying. *PAM* (2016).
[32] V. Giotsas, C. Dietzel, G. Smaragdakis, A. Feldmann, A. Berger, and E. Aben. Detecting Peering Infrastructure Outages in the Wild. *ACM SIGCOMM* (2017).
[33] V. Giotsas, G. Smaragdakis, C. Dietzel, P. Richter, A. Feldmann, and A. Berger. Inferring BGP Blackholing Activity in the Internet. *ACM IMC* (2017).
[34] V. Giotsas, S. Zhou, M. Luckie, and kc claffy. Inferring Multilateral Peering. *ACM CoNEXT* (2013).
[35] S. Goldberg. Why is It Taking So Long to Secure Internet Routing? *Communications of the ACM* 57, 10 (2014).
[36] A. Gupta, R. MacDavid, R. Birkner, M. Canini, N. Feamster, J. Rexford, and L. Vanbever. An Industrial-Scale Software Defined Internet Exchange Point. *USENIX NSDI* (2016).
[37] A. Gupta, L. Vanbever, M. Shahbaz, S. P. Donovan, B. Schlinker, N. Feamster, J. Rexford, S. Shenker, R. Clark, and E. Katz-Bassett. SDX: A Software Defined Internet Exchange. *ACM SIGCOMM* (2014).
[38] E. Heilman, D. Cooper, L. Reyzin, and S. Goldberg. From the Consent of the Routed: Improving the Transparency of the RPKI. *ACM SIGCOMM* (2014).
[39] N. Hinze, M. Nawrocki, M. Jonker, A. Dainotti, T. C. Schmidt, and M. Wählisch. On the Potential of BGP Flowspec for DDoS Mitigation at Two Sources: ISP and IXP. *ACM SIGCOMM (Posters)* (2018).
[40] Hurricane Electric. Customer Blackhole Community. (2006). http://www.he.net/adm/blackhole.html.
[41] M. Jonker, A. King, J. Krupp, C. Rossow, A. Sperotto, and A. Dainotti. Millions of Targets under Attack: a Macroscopic Characterization of the DoS Ecosystem. *ACM IMC* (2017).
[42] M. Jonker and A. Sperotto. Measuring Exposure in DDoS Protection Services. *IFIP/IEEE CNSM* (2017).
[43] M. Jonker, A. Sperotto, R. van Rijswijk-Deij, R. Sadre, and A. Pras. Measuring the Adoption of DDoS Protection Services. *ACM IMC* (2016).
[44] Juniper. Configuring Firewall Filters. https://www.juniper.net/documentation/en_US/junos/topics/task/configuration/firewall-filter-ex-series-cli.html. (2018).
[45] R. Karam, R. Puri, S. Ghosh, and S. Bhunia. Emerging Trends in Design and Applications of Memory-based Computing and Content-addressable Memories. *Proceedings of the IEEE* 103, 8 (2015).
[46] A. Khan, T. Kwon, H. C. Kim, and Y. Choi. AS-level Topology Collection through Looking Glass Servers. *ACM IMC* (2013).
[47] T. King, C. Dietzel, J. Snijders, G. Doering, and G. Hankins. BLACKHOLE Community. IETF RFC 7999. (2016).
[48] B. Krebs. KrebsOnSecurity Hit With Record DDoS. https://krebsonsecurity.com/2016/09/krebsonsecurity-hit-with-record-ddos. (2016).
[49] F. Lichtblau, F. Streibelt, T. Krüger, P. Richter, and A. Feldmann. Detection, Classification, and Analysis of Inter-domain Traffic with Spoofed Source IP Addresses. *ACM IMC* (2017).
[50] LINX. LINX: Black Holing Support for DDoS Attack. https://www.linx.net/files/hotlinx/hotlinx-34.pdf. (2013).
[51] A. Lutu, M. Bagnulo, and O. Maennel. The BGP Visibility Scanner. *IEEE INFOCOM Workshops* (2013).
[52] P. Marques, N. Sheth, R. Raszuk, B. Greene, J. Mauch, and D. McPherson. Dissemination of Flow Specification Rules. IETF RFC 5575. (2009).
[53] N. McKeown, T. Anderson, H. Balakrishnan, G. Parulkar, L. Peterson, J. Rexford, S. Shenker, and J. Turner. OpenFlow: Enabling Innovation in Campus Networks. *ACM CCR* 38, 2 (2008).
[54] D. McPherson, R. Raszuk, B. Pithawala, A. Karch, and S. Hares. Dissemination of Flow Specification Rules for IPv6. IETF draft. https://tools.ietf.org/html/draft-ietf-idr-flow-spec-v6-09. (2017).
[55] Merit Network, Inc. IRR - Internet Routing Registry. http://www.irr.net. (2018).
[56] J. Mirkovic and P. Reiher. A Taxonomy of DDoS Attack and DDoS Defense Mechanisms. *ACM CCR* 34, 2 (2004), 39–53.
[57] C. Morales. NETSCOUT Arbor Confirms 1.7 Tbps DDoS Attack; The Terabit Attack Era Is Upon Us. https://asert.arbornetworks.com/netscout-arbor-confirms-1-7-tbps-ddos-attack-terabit-attack-era-upon-us/. (2018).
[58] Netscout. Arbor Availability Protection System. https://www.netscout.com/product/arbor-availability-protection-system. (2018).
[59] C. Osborne. 2014 DDoS Attacks: Heavier and in higher Volume. https://www.zdnet.com/article/2014-ddos-attacks-heavier-and-in-higher-volume/. (2014).
[60] M. Prince. The DDoS That Almost Broke the Internet. https://blog.cloudflare.com/the-ddos-that-almost-broke-the-internet/. (2013).
[61] B. Quoitin, C. Pelsser, L. Swinnen, O. Bonaventure, and S. Uhlig. Interdomain Traffic Engineering with BGP. *IEEE Communications Magazine* 41, 5 (2003), 122–128.
[62] Y. Rekhter, T. Li, and S. Hares. A Border Gateway Protocol 4 (BGP-4). IETF RFC 4271. (2006).
[63] P. Richter, G. Smaragdakis, A. Feldmann, N. Chatzis, J. Boettger, and W. Willinger. Peering at Peerings: On the Role of IXP Route Servers. *ACM IMC* (2014).
[64] C. Rossow. Amplification Hell: Revisiting Network Protocols for DDoS Abuse. *NDSS* (2014).
[65] F. J. Ryba, M. Orlinski, M. Wählisch, C. Rossow, and T. C. Schmidt. Amplification and DRDoS Attack Defense–A Survey and New Perspectives. *arXiv preprint arXiv:1505.07892* (2015).
[66] J. Ryburn. DDoS Mitigation Using BGP Flowspec. NANOG 63. (2015).
[67] S. Sangli, D. Tappan, and Y. Rekhter. BGP Extended Communities Attribute. IETF RFC 4360. (2006).
[68] J. J. Santanna, R. van Rijswijk-Deij, R. Hofstede, A. Sperotto, M. Wierbosch, L. Z. Granville, and A. Pras. Booters – An analysis of DDoS-as-a-service Attacks. *IFIP/IEEE IM* (2015).
[69] P. Sermpezis, V. Kotronis, A. Dainotti, and X. Dimitropoulos. A Survey Among Network Operators on BGP Prefix Hijacking. *ACM CCR* 48, 1 (2018).
[70] D. Smith. Turkey DNS Servers Under Attack. https://blog.radware.com/security/2015/12/turkey-dns-servers-under-attack/. (2015).
[71] R. Stapleton-Gray and W. Woodcock. National Internet Defense – Small States on the Skirmish Line. *Communications of the ACM* 54, 3 (2011).
[72] F. Streibelt, F. Lichtblau, R. Beverly, A. Feldmann, C. Pelsser, G. Smaragdakis, and R. Bush. BGP Communities: Even more Worms in the Routing Can. *ACM IMC* (2018).
[73] US-CERT. UDP-Based Amplification Attacks. https://www.us-cert.gov/ncas/alerts/TA14-017A. (2018).
[74] R. van Rijswijk-Deij, A. Sperotto, and A. Pras. DNSSEC and its Potential for DDoS Attacks: A comprehensive Measurement Study. *ACM IMC* (2014).
[75] T. Vissers, T. Van Goethem, W. Joosen, and N. Nikiforakis. Maneuvering around Clouds: Bypassing cloud-based Security Providers. *ACM CCS* (2015).
[76] D. Walton, A. Retana, E. Chen, and J. Scudder. Advertisement of Multiple Paths in BGP. IETF RFC 7911. (2016).
[77] S. T. Zargar, J. Joshi, and D. Tipper. A Survey of Defense Mechanisms against Distributed Denial of Service (DDoS) Flooding Attacks. *IEEE Communications Surveys & Tutorials* 15, 4 (2013).
[78] P. Zilberman, R. Puzis, and Y. Elovici. On Network Footprint of Traffic Inspection and Filtering at Global Scrubbing Centers. *IEEE Transactions on Dependable and Secure Computing* 14, 5 (2017).