The Dark Menace: Characterizing Network-based Attacks in the Cloud

Rui Miao [*] Rahul Potharaju [‡] Minlan Yu[*] Navendu Jain[†]
[*] University of Southern California [‡] Microsoft [†] Microsoft Research

ABSTRACT

As the cloud computing market continues to grow, the cloud platform is becoming an attractive target for attackers to disrupt services and steal data, and to compromise resources to launch attacks. In this paper, using three months of NetFlow data in 2013 from a large cloud provider, we present the first large-scale characterization of *inbound* attacks towards the cloud and *outbound* attacks from the cloud. We investigate nine types of attacks ranging from network-level attacks such as DDoS to application-level attacks such as SQL injection and spam. Our analysis covers the complexity, intensity, duration, and distribution of these attacks, highlighting the key challenges in defending against attacks in the cloud. By characterizing the diversity of cloud attacks, we aim to motivate the research community towards developing future security solutions for cloud systems.

Categories and Subject Descriptors

C.2.0 [**Computer-Communication Networks**]: Security and protection; C.2.3 [**Network Operations**]: Network management

General Terms

Measurement, Security

Keywords

Attack Characterization; Network-based Attacks; DDoS

1. INTRODUCTION

The cloud computing market reached $40 billion in 2014 with a rapid growth of 23%-27% per year [1]. Hosting tens of thousands of online services, the cloud platform is increasingly becoming both the target and source of attacks. A recent survey of data center operators indicates that half of them experienced DDoS attacks, with 94% of those experiencing regular attacks [17]. Moreover, attackers can abuse hosted services or compromise VMs [30] in the cloud to target external sites via deploying botnets [29], sending spam [33, 44], selling VMs in the underground economy [23, 49], or launching DDoS attacks [21]. In April 2011, an attack on the Sony Playstation network compromising more than 100 million customer accounts was carried out by a malicious service hosted on Amazon EC2 [20]. While there have been some reports of individual attacks on enterprise and cloud networks [10, 29], to the best of our knowledge, there have not been any systematic measurement studies of attacks *on* and *off* the cloud which can guide the design of attack detection and mitigation systems. In fact, little has been published about the prevalence, diversity, and characteristics of these cloud-based attacks.

In this paper we investigate over 200 TB of NetFlow records collected from dozens of edge routers spread across multiple geographically distributed data centers of a major cloud provider. We group traffic based on public virtual IPs (VIPs) assigned to each cloud hosted service. We identify network-based attacks from the NetFlow data using four key features as also used in prior work [19, 32, 38, 48]: (1) significant traffic volume (e.g., packets per second), (2) abnormal fan-in or fan-out (e.g., number of unique clients or number of connections), (3) abnormal packet header signatures (e.g., TCP flags), and (4) communication with Internet malicious hosts [37]. Using these features, we identified nine types of attacks, ranging from various DDoS attacks to application-level attacks such as SQL injection and spam.

Due to sampling in the NetFlow data used in our study and the fact that NetFlow lacks application-level information, we do not aim at identifying *all* the attacks in the cloud. Instead, our goal is to understand the characteristics of attacks using low overhead network-level information typically collected in many data center networks. Thus, we take a conservative approach of setting the attack detection thresholds to ensure that most of the attacks we detect are real attacks.[1]

We validate the detected attacks against alerts from deployed security appliances and incident reports written by operators. Our detected attacks cover 78.5% of the inbound attack alerts from DDoS protection appliances and 83.7% of the incident reports on outbound attacks, due to the NetFlow sampling used in our study and our conservative approach. Note that the cloud provider we studied deploys a combination of software and hardware appliances to protect the infrastructure against such attacks.

Our broader goal is to (a) understand the key characteristics of these attacks to evaluate the effectiveness of existing DDoS mitigation approaches and (b) analyze their implications on building cloud-scale attack detection and mitigation solutions.

Although there have been many studies on Internet attacks, this paper presents one of the first analysis of the key characteristics of attacks to and from the cloud based on a three-month dataset.

IMC'15, October 28–30, 2015, Tokyo, Japan.
© 2015 ACM. ISBN 978-1-4503-3848-6/15/10 ...$15.00.
DOI: http://dx.doi.org/10.1145/2815675.2815707.

[1] These attacks may also include some traffic anomalies caused by flash crowds or misconfigurations. We do not distinguish them because they all impact cloud services and it is an open problem to accurately distinguish them from benign traffic.

We make the following main observations:

- We identify nine types of attacks and quantify their frequencies for inbound and outbound attacks (Section 3).
- We find that most VIPs experiencing attacks only incur one attack incident in a day. There is a very small fraction of VIPs that experience or generate many attacks (Section 4).
- We find multi-vector attacks and combinations of inbound and outbound attacks on the same VIP. While most attacks target only one VIP, there are a few cases of multiple attacks that target 20-60 VIPs simultaneously (Section 4).
- We observe high variations in attack throughput across time and VIPs, requiring cloud security solutions to have dynamic resource allocation over time and multiplexing of resources across VIPs. Attacks often have short duration (within 10 minutes), which require fast attack detection and mitigation (Section 5).
- We investigate the origins and targets of inbound and outbound attacks and identified the major types of Internet ASes that are involved in cloud-related attacks (Section 6).

Scope and Limitations. Our study analyzed traffic data from a single cloud provider and thus it may not generalize to other providers. However, the scale and diversity of our dataset, and our conversation with security operators (having a broader industry view and some having worked at other cloud networks) suggests that similar security challenges are likely faced by other providers. We collected NetFlow records from data center edge routers *before* they are filtered by the security appliances. Thus, the attacks we detected should not be interpreted as impacting the cloud infrastructure or services. Finally, since the traffic measurement is at one-minute granularity, it is likely to smooth the effect of short-lived attack spikes. Overall, our study highlights the need for developing programmable (to handle attack diversity), scalable (to handle varying intensity), and flexible approaches (for individual tenants) to protect against attacks.

2. DATASETS AND METHODOLOGY

We first present the basic setup in a major cloud provider we studied, and then describe the datasets we collected and the methodology for characterizing attacks.

2.1 Cloud provider overview

The cloud network we study comprises 10+ geographically distributed data centers across America, Europe, Asia, and Oceania, which are connected to each other and to the Internet via edge routers. Each data center hosts tens to hundreds of thousands of servers. The cloud provider hosts more than 10,000 services including web services, mobile application services, database and storage services, and data analytics. Each service is assigned a public virtual IP (VIP). The traffic to the VIP is load balanced across a group of virtual machines hosting the service; sometimes these VMs are located across multiple data centers.

Such scale of services makes the cloud an attractive target for *inbound* attacks. Incoming traffic to different services first traverses the edge routers and then the commercial security appliances (e.g., Arbor [17]). These security appliances, typically designed for enterprise-scale deployments, analyze inbound traffic to protect against a variety of well-known attacks such as TCP SYN flood, UDP flood, ICMP flood, and TCP NULL attacks; these appliances use NetFlow records for traffic monitoring. However, the detection logic is often limited to known high-volume attacks. Thus they risk missing other low-volume attack types which aim to probe vulnerabilities but that do not impact the cloud infrastructure such as stealth port scans and application-level attacks e.g., spam, SQL injection. To reduce false positives (noisy alerts), traffic thresholds for alerting can be set either on a per-tenant basis or across tenant groups on these devices.

Attackers can also abuse the cloud resources to launch *outbound* attacks. For instance, they can first launch brute-force attacks (e.g., password guessing) to compromise vulnerable VMs in the cloud. These compromised VMs may then be used for YouTube click fraud, BitTorrent hosting, Bitcoin mining, spamming, malware propagation, or launching DDoS attacks. To mitigate outbound attacks, the cloud provider we studied enforces several security mechanisms including limiting the outbound bandwidth per VM, preventing IP spoofing of egress traffic, shutting down the misbehaving VMs and isolating anomalous traffic. To our knowledge, no prior work has characterized the prevalence of outbound attacks from the cloud.

2.2 Dataset and attack detection methodology

We obtained more than 200TB NetFlow logs from a major cloud provider over three months (May, Nov, and Dec 2013). The NetFlow logs collected for our study had a 1 in 4096 packet sampling rate for both inbound and outbound traffic at the edge routers of the data centers, and aggregated over one-minute windows.[2] Since the edge routers (where we collect the logs) are located upstream of the security appliances, the attacks we detect are likely mitigated before they reach VMs hosting services in the cloud. We analyze the NetFlow data on Cosmos, a large scalable data storage system using SCOPE [26], a programming framework similar to Map-Reduce. Our SCOPE scripts use C# and SQL-like queries to perform the analysis described below.

We aggregate the NetFlow data by VIP in each one-minute window, and study the traffic to a VIP (inbound traffic) and from the same VIP (outbound traffic). For each VIP in each time window, we first filter the traffic based on the protocol number (e.g., UDP), TCP flags (e.g., TCP SYN), or port numbers (e.g., SQL traffic is filtered by TCP traffic with destination port 1433 or 3306). We then identify nine types of attacks listed in Table 1. Our attack detection is based on the following four network-level features:

Volume-based: Many volume-based attacks try to exhaust server or infrastructure resources (e.g., memory, bandwidth) by sending a large volume of traffic via a specific protocol such as TCP SYN and UDP floods, and DNS reflection attacks. We capture volume-based attacks by identifying traffic with large relative spikes. We use sequential change point detection [19, 32] by comparing the traffic volume at the current time window with the Exponentially Weighted Moving Average (EWMA) of the past 10 time windows. We then compare the difference with a change threshold of 100 packets per minute in NetFlow (1:4096 sampling rate), corresponding to an estimated value of about 7K pps in the original traffic. The threshold is suggested by the cloud security team based on the network capacity and prior attack incidents. As shown in Section 3.2, using such threshold settings, we can verify many of the attacks reported in the attack alerts and the incident reports.

Spread-based: For many services (e.g., mail, SQL, SSH), a single VIP typically connects to only a few Internet hosts in normal operation. Thus, if a VIP communicates with a large number of Internet hosts, it is likely an anomaly. To identify such anomalies, we use the NetFlow data to compute the spread of a VIP (i.e., the number of distinct Internet IPs communicating with a VIP during a time window) for inbound and outbound traffic. We then capture the relative spikes of the spread using sequential change point de-

[2] All the traffic volume numbers we show in the paper are estimated volumes calculated based on the number in the NetFlow data and the sampling rate.

Attacks	Description	Net/App	Target	Network features	Detection method	Inactive timeout
TCP SYN flood	Send many TCP SYN, UDP, ICMP packets to random or fixed ports on a server	Net	Server resources	#pkts/min	Volume-based	1 min
UDP flood		Net	Network bandwidth	#pkts/min	Volume-based	1 min
ICMP flood		Net	Server resources	#pkts/min	Volume-based	120 min
DNS reflection	A large number of DNS responses sent to a target from DNS servers (triggered by DNS requests sent by attackers with spoofed source addresses)	App	Network bandwidth	#pkts/min	Volume-based	60 min
Spam	Launch email spam to multiple SMTP servers	App	Users	fan-in/out ratio	Spread-based	60 min
Brute-force	Scan weak passwords or administrative control (using RDP, SSH, VNC)	App	Server vulnerability	fan-in/out ratio, #conn/min	Spread-based	60 min
SQL injection	Send different SQL queries to exploit software vulnerabilities	App	SQL server vulnerability	#conn/min	Spread-based	30 min
Port scan	Scan for open ports (using NULL, Xmas packets)	Net	Server vulnerability	#conn/min	Signature-based, Spread-based	60 min
Malicious web activity (TDS)	Communicate with hosts on malicious web infrastructure	App	Users	src IP/dst IP	Communication pattern-based	120 min

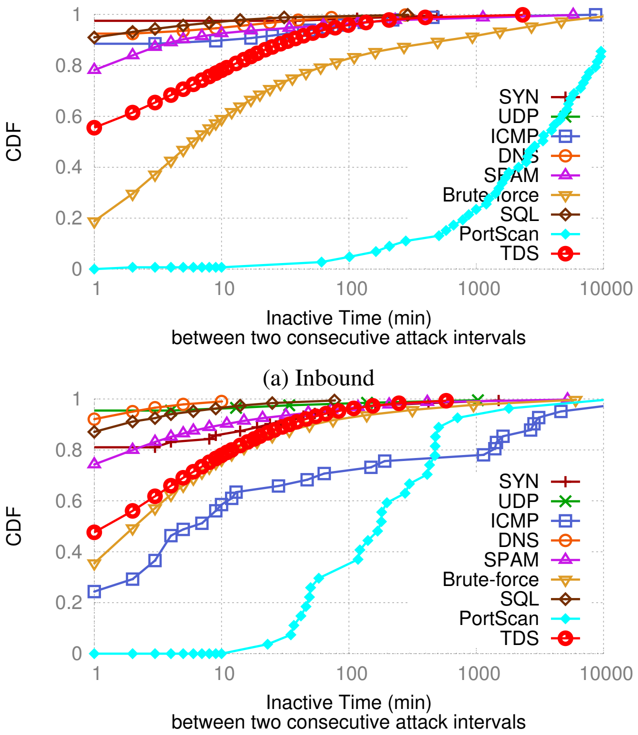
Table 1: Summary of the network-based attacks in the cloud we studied.

tection. Such spread-based detection of brute-force attacks has also been used in prior work [31]. We choose 10 and 20 Internet IPs as the threshold for brute-force and spam, respectively, and 30 connections for SQL in the sampled NetFlow, as recommended by the cloud security team.

Signature-based: Although packet payloads are not logged in our NetFlow data, we can still detect some attacks by examining the TCP flag signatures. Port scanning and stack fingerprinting tools use TCP flag settings that violate protocol specifications (and as such, they are not used by normal traffic) [6, 40]. For instance, the *TCP NULL port scan* sends TCP packets without any TCP flags, and the *TCP Xmas port scan* sends TCP packets with FIN, PSH, and URG flags (Table 1). If a VIP receives a packet with an illegal TCP flag setting during a time window, we mark the time window as under an attack. Since the NetFlow data is sampled, even a single logged packet may represent a significant number of packets with illegal TCP flag settings in the original traffic.

Communication pattern based: Previous security studies have identified blacklists of IPs in the Internet. We can identify attacks by filtering VIP traffic communicating with such blacklisted IPs. For example, the Traffic Distribution System (TDS) [37] includes a list of dedicated hosts that deliver malicious web content on the Internet. Since these hosts are hardly reachable via web links from legitimate sources, it is likely that cloud VIPs communicating with these hosts are involved in malicious web activities. In particular, these VIPs are either a victim of inbound attacks (e.g., spam, malicious advertising) or that they have been compromised to launch outbound attacks (e.g., drive-by downloads, scams, and phishing). Note that it is not always possible to infer the direction of an attack involving TDS nodes because some SYN packets may not get sampled in the NetFlow data. Thus, we distinguish the inbound from outbound communication pattern based attacks based on the destination IP in the flow records.

Counting the number of unique attacks. Given the attacks in each one minute time window, we identify the attack incidents that last multiple time windows for the same VIP. Due to NetFlow's low sampling rate, we may not be able to detect an attack over its entire duration. Therefore, similar to previous work [38, 40, 48], we *group* multiple attack windows as a single attack where the last attack interval is followed by T inactive windows (i.e., no attacks).



(a) Inbound

(b) Outbound

Figure 1: The distribution of inactive time for each attack type; the x-axis is on log-scale.

Instead of selecting a fixed T, we choose to select different T for different attacks based on analyzing the distributions of inactive times between two consecutive attack minutes of each type for both

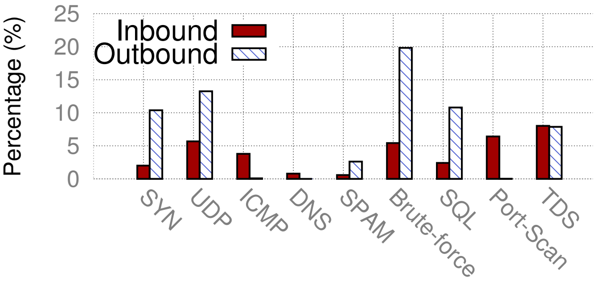


Figure 2: Percentage of total inbound and outbound attacks.

inbound and outbound attacks, as shown in Figure 1. We select the T value by generating a linear regression line between each point and the 99 percentile of each attack distribution curve and checking that the average R-squared [28] value for regression models of inbound and outbound curves is above 85%. We summarize the inactive timeout values we use for different attacks in Table 1.

3. ATTACK OVERVIEW AND VALIDATION

In this section we first give an overview of each type of inbound and outbound attack observed in our study. For validation, we compare these detected attacks using inbound attack alerts from DDoS security appliances and the attack incident reports.

3.1 Attack Overview

Figure 2 shows the distribution of inbound and outbound attacks; absolute counts omitted due to confidentiality and privacy concerns.

Flood attacks: Flood attacks (TCP, UDP, ICMP floods) in the Internet domain have been widely studied [39], and they can be launched in both inbound and outbound directions. Our analysis identified a significant increase of inbound flood attacks during Nov and Dec compared to May (breakdown by month not shown), possibly to disrupt the e-commerce sites hosted in the cloud during the busy holiday shopping season. UDP floods are often against media services hosted in the cloud and on HTTP ports. We also observe that there are about 5 times more outbound TCP SYN and about 2 times more UDP floods than inbound. This is likely because it is easier for attackers to leverage cloud resources to attack Internet users, while it is harder to attack the cloud where operators have high level of security expertise and many attacks floods can be filtered by commercial security appliances.

DNS reflection attacks: The DNS reflection attack is one of the most common amplification attacks in the Internet. It has received increasing attention from DDoS protection services [10, 17]. In DNS reflection attacks, attackers send DNS requests toward multiple open DNS servers with spoofed source address of the target, which results in a large number of DNS responses to the target from DNS servers. Since the cloud has its own DNS servers to answer DNS queries from hosted tenants, there should not be any DNS responses from the Internet to the cloud. Therefore, any activity of inbound DNS responses may signify a potential DNS reflection attack. Inbound DNS reflection attacks often come from up to 6K distinct sources (with 1500 byte full-size packets). We only observed outbound DNS responses from a single VIP hosting a DNS server at 5666 packets per second for a couple of days repeatedly.

Spam: Email services often communicate with a stable number of clients at any given time. If we see a large deviation in the number of email flows, they are likely to be spam. For instance, we observed an outbound spam eruption on a single day, which accounted for 40% of the total outbound spam instances in May. The spam traffic came from hundreds of VIPs towards thousands of external mail servers from email providers such as Yahoo and Lycos, enterprises like CenturyLink, and small clouds like SoftLayer. We observed prevalent on-off traffic pattern from the spamming VIPs. Specifically, each VIP generated slow rate spam traffic with a median of 2266 packets per second over a median of one hour period. It then subsided completely over a median of 5 hours, and launched new attacks again. We investigated these VIPs with the security team and found that most of these VIPs are free trial accounts which were quickly shut down. About 98% of these VIPs were new with no spam traffic recorded before, and the remaining ones were slow spammers lasting up to a month.

Brute-force attacks: Remote connection protocols like SSH, RDP (Remote Desktop Protocol), and VNC (Virtual Network Computing) often have just a few connections to a single VIP. If we observe many connections in the sampled NetFlow, they are likely caused by malicious behaviors such as password guessing (i.e., brute-force attacks). We observed that inbound brute-force attacks have a median of 24 distinct hosts communicating with a single VIP just in the sampled NetFlow data (i.e., there are likely other Internet hosts communicating with the VIP that are not in our sampled data). At the tail, a VIP can receive SSH traffic from up to 10K distinct hosts from the sampled NetFlow data. This can be caused by an attacker controlling multiple Internet hosts to try out different passwords in parallel. Outbound brute-force attacks have a median of 60 distinct hosts targeted by the same VIP in the sampled NetFlow data. This may be because the VIP is scanning a set of Internet servers with the same set of passwords. There are about 4 times more outbound brute-force attacks than inbound and more SSH-based brute-force attacks compared to the RDP ones, likely because the servers running in the cloud often use random ports (e.g., for RDP), and thus they are less likely to experience brute-force attacks compared to Internet hosts.

SQL injection attacks: Some attackers send a variety of malware SQL queries to exploit the vulnerability of SQL servers [16, 18]. Although these attacks are in the application layer, we can still observe such attacks in the network layer when there is a large number of connections issued towards SQL database servers. It is likely that they are exploiting all possible malformed user inputs to gain unauthorized access [16]. There are about 5 times more outbound SQL attacks compared to the inbound attacks.

Port scan: We observe many inbound port scan attacks such as TCP NULL and Xmas attacks. For example, we observed inbound traffic of 125k TCP NULL packets per second lasting for 4 minutes. Attackers usually leverage these packets to sneak through firewalls and packet filters that only handles normal TCP packets [6]. Moreover, there is a significant number of inbound TCP RST packets, which are likely caused by Internet hosts spoofing the IPs in the cloud, leading to TCP RST signals to be directed to VIPs inside the cloud. There are much fewer outbound port scans.

Malicious web activities (TDS): There are 0.039% of VIPs involved in communicating with TDS hosts in the Internet. These TDS hosts often use source ports uniformly distributed between 1024 and 5000. There is one attack incident with 89 unique Internet TDS IPs communicating with a single VIP with 31K packets per second lasting for 98 minutes.

Summary: There are more outbound attacks than inbound attacks (64.9% vs 35.1%). This implies that it is relatively easier for attackers to abuse cloud resources to launch outbound attacks than to attack the cloud-hosted tenants due to improved security over the years. At the same time, new security mechanisms need to

Attack	Inbound #detected/#alerts	Outbound #detected/#reports
TCP SYN flood	98/197	8/8
UDP flood	403/442	4/4
ICMP flood	0/0	0/0
DNS reflection	-	10/10
Spam	-	55/55
Brute-force	-	27/34
SQL injection	-	4/4
port scan	3/3	0/0
TDS	-	-
Others(Malware hosting/phishing)	-	0/14
Total	504/642=78.5%	108/129=83.7%

Table 2: Detected inbound alerts and outbound incident reports ("-" means that the alerts or reports do not support the attack type).

be developed to reduce the outbound attacks. The inbound attacks are dominated by flood attacks, brute-force attacks, port scan, and TDS attacks, while the outbound attacks are dominated by flood attacks, brute-force attacks, SQL attacks, and TDS attacks. While our study focuses on characterizing the diversity of cloud attacks, comparing across attack categories (e.g., by impact, traffic thresholds) may also reveal interesting insights. However, defining a universal metric to compare attack types is difficult because it requires normalizing the attack data across diverse metrics e.g., quantifying the impact of an attack in terms of the service downtime, privacy compromise, and the number of users impacted. Further, some of these measures may not be known till long after the attack happened. In Section 4, we study one aspect of how VIPs are affected by different attacks and leave the broader analysis to future work.

3.2 Validation

In a large heterogeneous network, it is difficult to verify whether all the detected attacks are real because it requires a snapshot of the actual traffic and the runtime application and system state before and during the attack. This problem becomes even harder given the coarsely sampled NetFlow data available for our analysis. We collect the security records including alerts from the DDoS protection hardware appliances for inbound attacks and the incident reports for outbound attacks. We compare our detected attacks with the alerts and incident reports to identify the attacks we miss. Note that the cloud provider deploys software and hardware security appliances to safeguard against these attacks so they should not be interpreted as impacting the infrastructure or tenants.

Inbound attacks: The cloud provider detects and mitigates inbound attacks using a combination of software and hardware DDoS protection appliances. These appliances generate alerts about TCP SYN floods, UDP floods, ICMP floods, and TCP NULL scan. Note that on hardware security appliances, the traffic thresholds are typically set to handle only the high-volume attacks (low-volume attacks don't cause any impact to the cloud infrastructure due to high network capacity) over a large time window, and these appliances aggregate multiple incidents together that occur close in time. Therefore, to do a side-by-side comparison with alerts from these devices, we also first group the attacks we detected based on the VIP, attack type and time window. We found that 73.2% of these attack instances were correlated. This is due to the fact that we set the traffic thresholds to (a) cover a broad range of inbound attacks, and (b) detect these attacks in their early stages. To check the latter hypothesis, we measured the detection latency of hardware security appliances by randomly sampling a few attack instances over a week and observed that these appliances detected them after an order of tens of seconds delays on average. In comparison, our detection approach (in offline mode) signaled the attack based on the NetFlow data for these instances within a minute.

Table 2 shows the number of alerts in each type and those alerts that we also detected. Overall, we successfully identified 78.5% of the alerts from hardware security appliances in our detected attacks. The remaining alerts are missed by our detection approach because of the low sampling in the NetFlow data and the false positives of these alerts. For other types of attacks, the cloud relies on individual tenants to detect and report them. However, we did not have the ground truth data to validate them.

Outbound attacks. The cloud security team detects every potential outbound attack, but they do not necessarily log all of them as incident reports to avoid false positives. Specifically, only the cases of anomalous activity reported by external sites are logged as incidents. Similar to inbound attacks, the cloud provider uses security appliances to mitigate the outbound attacks. The cloud security team may receive a complaint from an ISP when they notice malicious traffic originating from the cloud provider. Given such a complaint, the security team checks the logs of security appliances and investigates the corresponding tenant profile and payment information, and generates an incident report. Moreover, when the security team receives a complaint, the team may do traffic analysis for deeper investigation; they may also perform the VHD (Virtual Hard Disk) forensic analysis on behalf of the customer if the customer (who owns the VHD) requested it. Based on these investigations, the security team creates incident reports logging their findings such as the "attack" or "no attack found" label. We use NetSieve [45] to extract the attack information from these incident reports, and then compare it with our detected outbound attacks.

Since these incident reports come from Internet users' complaints, there is a large number of short-term transient attacks that are not covered by these reports. Therefore, we focus on the false negatives of those attacks that we missed in our approach when we compare with incident reports, rather than those attacks that we detect but are missed in these reports. For those attacks that are not covered by the incident reports, we randomly picked a few attacks for each type and investigated them. The attacks for which the packet traces got logged were verified as being mitigated by the security team.

Table 2 shows the number of incident reports (labeled as "attacks") and those that are also detected by our approach. We detect most of the attacks reported in the incident reports (83.7%). There are only two exceptions: (1) There are some incident reports about application-level attacks such as phishing and malware hosting that we cannot detect with network signatures. (2) We only investigate brute-force attacks on three remote communication protocols (SSH, RDP, VNC). Therefore we miss brute-force attacks on other protocols such as FTP.[3] There are four incident reports labeled as "no attacks" that are also covered by our detected attacks. We investigated these attacks manually and confirmed with the security team that they are real attacks (on TCP SYN floods, SSH and RDP brute-force attacks) but mislabeled. Our analysis has been leveraged by the cloud security team towards improving attack detection, reducing time to detect, and identifying correlations across attack types.

Limitations of NetFlow. Due to coarse-grained sampling in the NetFlow data collected for our study and the fact that NetFlow lacks application-level information, we do not aim at identifying all the attacks in the cloud. We may miss application-level attacks without network-level signatures and those attacks that do not appear

[3] Some incident reports do not describe the protocols that are involved in the brute-force attacks.

in sampled NetFlow (e.g., HTTP Slowloris [24]). Instead, our goal is to understand what we can learn about cloud attacks with low overhead network-level information.

Although we just detect a subset of attacks due to the conservative approach, it is still useful to understand the key characteristics of these attacks to shed light on the effectiveness of commercial attack-protection appliances, and the implications on designing future attack detection and protection solutions. Studies [12, 22, 34] have shown that sampled NetFlow does not affect the detection accuracy of flood attacks but it may underestimate the number of flows. Therefore, the number of flows we report should be viewed as a lower bound on the number in the original traffic.

4. ANALYSIS OF ATTACKS BY VIP

In our three-month trace data, there are on average 0.08% of VIPs per day under inbound attacks and 0.11% of VIPs per day generating outbound attacks. In this section we investigate these VIPs to understand the attack frequency per VIP, multi-vector attacks on the same VIP, inbound and outbound attacks on the same VIP, and attacks that involve multiple VIPs.

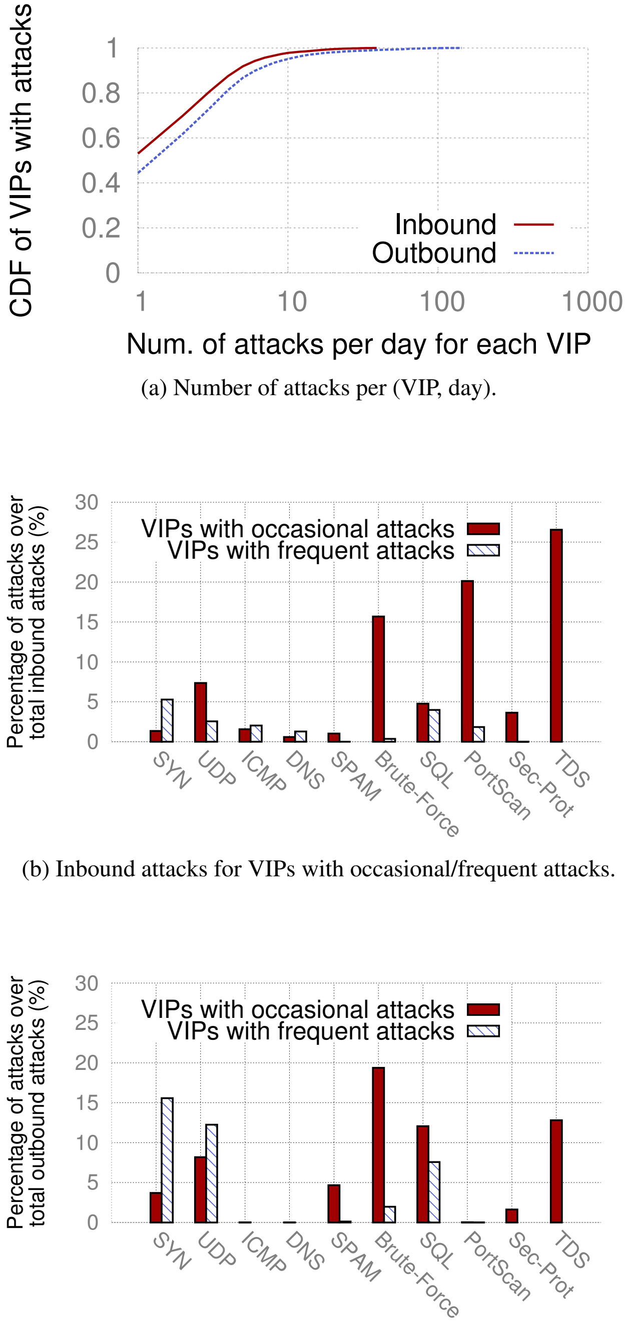
4.1 Attack frequency per VIP

Attack frequency per VIP: We count the number of attacks per VIP per day (Figure 3a). Most VIPs experiencing attacks only incur one attack incident during a day. Out of the 13K (VIP, day) pairs experiencing inbound attacks, 53% of pairs experience only one attack in a day. Out of 18K (VIP, day) pairs experiencing outbound attacks, 44% generate only one attack in a day because the misbehaving instances are aggressively shut down by the cloud security team.

At the tail, a VIP can get 39 inbound attacks in a day. This is a VIP hosting Media and HTTP services receiving frequent flood attacks (i.e., SYN, UDP, ICMP) with a median duration of 6 minutes and a median inter-arrival time of 64 minutes. For outbound attacks, there are 0.05% of outbound (VIP, day) pairs generating more than 100 attacks. We observed one VIP that generated more than 144 outbound TCP SYN flood attacks in a day to many web servers in the Internet with a median duration of 1 minute and a median inter-arrival time of 10 minutes. This VIP did not receive any inbound traffic during a whole month in the NetFlow data indicating that this VIP does not likely host any legitimate cloud service but it is only being used for malicious behavior.

VIPs with frequent and occasional attacks: We observed that there are only a few VIPs getting more than 10 attacks (2% of the inbound pairs and 5% of the outbound pairs). Therefore, we classify the VIPs into two classes: those VIPs with no more than 10 attacks per day and those with more than 10 attacks per day. Understanding the VIPs under frequent attacks is important for operators to extract the right attack signatures (e.g., popular attack sources) to protect these VIPs from future attacks.

Figure 3 shows that for inbound attacks, there are more TDS, port scan, and brute-force attacks for those VIPs with occasional attacks than those with frequent attacks (26.6% vs. 0% for TDS, 20.1% vs. 1.84% for port scan, and 15.7% vs. 0.359% for brute-force). It is natural for port scan and brute-force attacks to target VIPs with occasional attacks because these attacks search widely for vulnerabilities (e.g., open ports, weak passwords). TDS attacks also interact more with VIPs with occasional attacks, which makes TDS attacks harder to detect. Our further investigation shows that these occasional attacks mainly target applications running protocols like HTTP, HTTPS, DNS, SMTP, and SSH.



(a) Number of attacks per (VIP, day).

(b) Inbound attacks for VIPs with occasional/frequent attacks.

(c) Outbound attacks for VIPs with occasional/frequent attacks.

Figure 3: Attack characterization for VIPs with inbound and outbound attacks; the x-axis is on log-scale in the top figure.

VIPs under frequent attacks often experience relatively more TCP SYN flood attacks than those VIPs under occasional attacks (5.3% vs. 1.4%). Our investigation shows that these frequent flood attacks often target several popular cloud services on these VIPs including streaming applications, HTTP, HTTPS, and SSH.

Similarly, for outbound attacks, the VIPs with occasional attacks experience more brute-force, TDS, and spam attacks than the VIPs with frequent attacks (19.4% vs. 1.97% for brute-force, 12.8% vs. 0% for TDS, and 4.7% vs. 0.119% for spam). While attackers

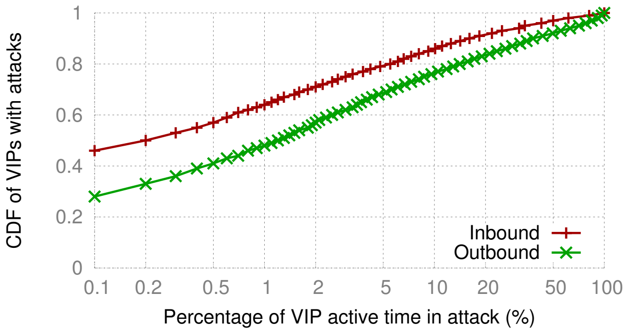


Figure 4: CDF of the percentage of VIP active time in attack.

may try to use free-trials or create fake accounts to launch them, the attack activity is only short-lived because the anomalous VMs are aggressively shut down by the cloud operators. It is challenging to detect these attacks because they come from multiple VIPs in the case of occasional attacks (e.g., spam) and they typically last only a short time. In contrast, those VIPs with frequent attacks are often the sources for TCP SYN and UDP flood attacks. For a few cases, we manually verified that these VIPs have compromised VMs, which may be sold in the underground economy [23, 49].

Fraction of VIP's lifetime involved in inbound attacks: We investigated the fraction of the time a VIP is under inbound attacks or generating outbound attacks compared to its total active time (i.e., the time that the VIP has active traffic). Figure 4 shows that 50% of VIPs experience inbound attacks for 0.2% of their active times. These are occasional attacks that do not likely affect much of their service. However, 3% of the VIPs receive inbound attack more than 50% of their operating time. Further investigation reveals that these VIPs run media, web, mail, and database services. Cloud operators need to effectively block these attacks to eliminate their impact on cloud services.

Compromised VIPs vs. malicious VIPs for outbound attacks: We also study the fraction of time a VIP generates outbound attacks compared to its active time. Note that most of these compromised VMs had weak passwords highlighting the need to enforce security best practices such as configuring cryptographically strong passwords. Figure 4 shows that 50% of VIPs generate outbound attacks for 1.2% of their active times. These VIPs are likely legitimate tenants that may have been compromised by attackers to generate outbound attacks occasionally (see Section 4.2 for one such example). In contrast, 8% of attack VIPs generate outbound attack for more than 50% of their active times. These VIPs are likely to be recruited mainly for attacks e.g., attackers may buy compromised VMs in the cloud or leverage free trial accounts.

4.2 Attacks on the same VIP

Multi-vector attacks: We observe multiple types of attacks attacking the same VIP or coming from the same VIP at the same time. This is likely because a single malicious program tries to launch multiple types of attacks to exploit the vulnerabilities of targets or to exhaust target resources in different ways. We identify these attacks if their start times to/from the same VIP differ less than five minutes. We find that 106 VIPs experience more than one type of inbound attacks simultaneously, which accounts for 6.1% of the total inbound attacks. There are 74 VIPs that experience more than

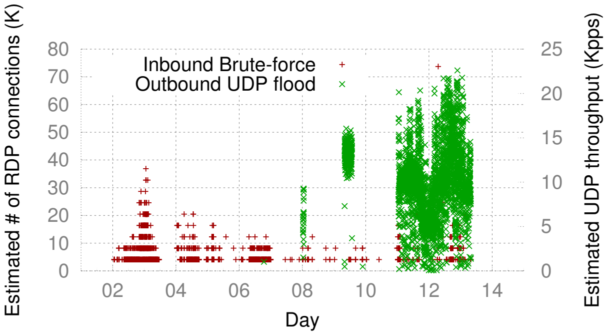


Figure 5: Inbound and outbound attacks on the same VIP. We estimate the UDP throughput and the upper bound of the number of RDP connections based on the 1 in 4096 sampling rate.

one type of outbound attacks simultaneously, which accounts for 0.83% of the total outbound attacks. Among these VIPs, 46 VIPs are targets of multi-vector volume-based attacks (i.e., TCP SYN, UDP, ICMP floods, and DNS reflection). There are 11 VIPs that launch multi-vector outbound volume-based attacks.

A new observation we make about outbound attacks is that there are 35 VIPs which launch brute-force attacks together with TCP SYN and ICMP flood attacks (which account for 22.3% of the outbound multi-vector attacks). This is likely a new attack pattern that attackers find effective in breaking Internet hosts.

Inbound and outbound attacks on the same VIP. There are also several cases of simultaneous inbound and outbound attacks. Figure 5 shows one such case. A VIP from a partner subscription was inactive (i.e., no traffic) for a long time. Starting the second day, the VIP started to receive inbound RDP brute-force attack for more than a week. These brute-force attacks originated from 85 Internet hosts, where 70.3% of attack packets are from three IP addresses within a single resident AS in Asia. These brute-force attack had a peak of estimated around 70 K flows per minute with a few packets sampled in each flow. On the eighth day, the VIP started to generate outbound UDP floods against 491 external sites. The outbound UDP attack had a peak volume at 23 Kpps, lasting for more than two days. Detecting such attacks requires first jointly analyzing the inbound and outbound traffic to identify the attack patterns of compromised VIPs, and then blocking their outbound traffic.

4.3 Attacks on multiple VIPs

If attacks of the same type start on multiple VIPs simultaneously, it is likely that these attacks are controlled by the same attacker. We identify attacks on multiple VIPs if the difference of their start times on different VIPs is less than five minutes.[4]

Figure 6 shows that most types of attacks are targeted by fewer than 10 VIPs in the 99th percentile. We also observe that most types of attacks are targeted at only one VIP in the median (not shown in the figure). Inbound brute-force attacks have simultaneous attacks on 53 VIPs in the 99th percentile and 66 VIPs in the peak. We investigated the attacks at the tail and find that there are two Internet hosts from small cloud providers that start attacking 66 VIPs at the same time, and move to other VIPs. During a single day, these two

[4] We choose five minutes because the ramp-up time is 1-3 minutes for flood attacks and the inactive time T_I (defined in Section 2) for other attacks is larger than 10 minutes.

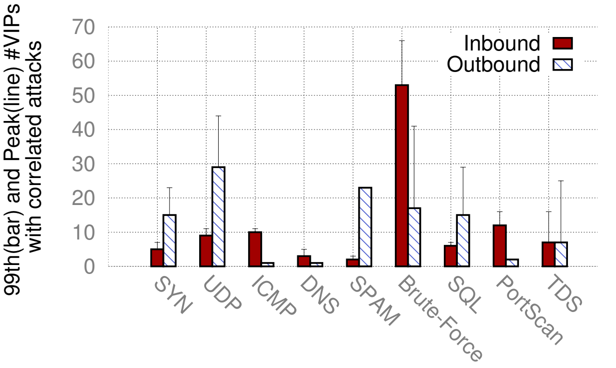


Figure 6: The 99th percentile and the peak number of VIPs simultaneously involved in the same type of attacks.

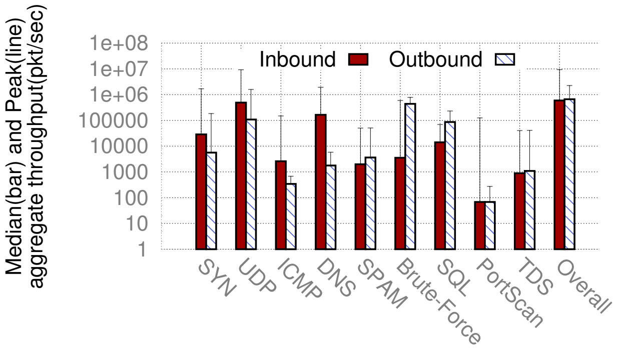


Figure 7: Median and maximum aggregate throughput by attack type; the y-axis is on log-scale.

Internet hosts attack more than 500 VIPs. These VIPs are located in five data centers in the cloud, and they belong to 8 IP subnets with different sizes (/17 to /21). The attacker scans through the entire IP subnet with up to 114.5 Kpps attack traffic per VIP. To prevent such attacks, we need to correlate the traffic to different VIPs and coordinate their attack detection and mitigation.

For outbound attacks, UDP flood, spam, brute-force, and SQL attacks involve around 20 VIPs simultaneously in the 99th percentile. In the peak, UDP flood and brute-force attacks involve more than 40 VIPs simultaneously.

4.4 Cloud services under inbound attack

We now investigate the major types of cloud services under inbound attacks. We capture the NetFlow records for VIPs receiving inbound attacks. We then filter all the attack traffic from the traffic on the VIPs, and the remaining traffic on the VIPs is mostly legitimate traffic. We use the destination port of inbound traffic to infer what type of applications and services are hosted on the VIPs. We count the application type if the traffic on the application port exceeds ten percent of its total traffic. Table 3 shows the percentage of VIPs with different types of cloud services that experience different types of inbound attacks.

Web services (HTTP/HTTPS) are major services in the cloud with 99% of the total traffic. VIPs hosting these web services receive a wide range of attacks such as SYN floods, ICMP floods, brute-force attacks, port scan, and TDS attacks. Web services receive the largest number of SYN attacks which aim to consume all the available connections of application servers. 1.2% of the SYN floods use source port 1024 and 3072, which are likely caused by a bug from an old SYN flood tool juno [9]. Blacklisting or rate-limiting these ports can help mitigate SYN floods.

We also observe other non-flood attacks targeting specific types of services. For instance, there are 35.06% VIPs hosting RDP servers with standard RDP port. The attackers often detect active RDP ports and then generate brute-force attacks against the server. TDS attacks mostly target VIPs running web services and mail services for spam, malware spreading, and malicious advertising. There are 6.94% of VIPs under attack running web services and 1.75% of VIPs running mail services.

5. ATTACK CHARACTERIZATION

We next investigate the characteristics of attacks to derive implications for the design of attack detection and prevention systems. First, to quantify the cloud bandwidth capacity needed to defend against attacks, we study the throughput of different types of attacks. Second, to understand how fast the attack detection system needs to react to attacks, we study the duration, ramp-up rate, and inter-arrival times of different types of attacks.

5.1 Attack throughput

Throughput by attack type: Figure 7 shows the median and peak aggregate throughput over the entire cloud for each type of attack and all the attacks overall. We measure the attack throughput using packets per second (pps) because the resources (CPU and memory) used to prevent these attacks are often correlated to the traffic rate. The overall inbound attack throughput has a median of 595 Kpps and a peak of 9.4 Mpps. The overall outbound attack throughput is lower with a median of 662 Kpps and a peak of 2.25 Mpps. Compared to the average throughput of legitimate traffic (54.3 Mpps for inbound and 49.7 Mpps for outbound), the median attack throughput is about 1% of the total traffic. These attacks can have a significant impact on both the cloud infrastructure (firewalls, load balancers) and services hosted in the cloud if we do not filter them at the network edge.

We now study the peak volumes of individual attacks to understand the resources we need to protect against them. TCP SYN floods have a peak throughput of 1.7 Mpps for inbound and 184 Kpps for outbound. It is important to prevent these attacks timely (e.g., using SYN cookies) before they exhaust many resources in the cloud infrastructure such as load balancers and firewalls.

The peak throughput for inbound UDP floods is 9.2 Mpps while that of outbound UDP floods is 1.6 Mpps. While these flood attacks aim to consume the cloud bandwidth or cause congestion to degrade service performance, the cloud networks are provisioned with a high network capacity to defend against them [7, 14]. Given that a software load balancer (SLB) can handle 300 Kpps per core [42] for simple Layer 4 packet processing, in the worst case handling inbound UDP floods may waste 31 extra cores in the data center infrastructure. If we fail to do in-network filtering of these UDP floods, they would cost even more resources per packet in the VMs which have more complex packet processing. However, detecting some application-level attacks (e.g., brute-force, spam), endpoint based approaches can leverage application semantics to better handle them compared to in-network defense approaches.

We also observe large variations in attack volumes over time. For inbound attacks, port scan has 1000x difference between the peak

Service(port)	Total	SYN	UDP	ICMP	DNS	SPAM	Brute-force	SQL	Portscan	TDS
RDP (3389)	35.06	0.11	0.21	0.54	0.11	0.07	33.88	0.11	0.32	0
HTTP (80,8080)	33.20	3.40	1.50	1.97	0.79	0.32	9.34	0.11	13.63	6.94
HTTPS (443)	13.27	1.22	0.29	1.40	0.21	0.07	4.44	0.04	8.05	0.14
SSH (22)	8.69	0	0.11	0	0.04	0	8.52	0	0.18	0
IP Encap (0)	6.55	0.54	1.57	0.79	1.07	0.04	0.29	0	0.39	0.04
SQL (1433, 3306)	3.11	0	0	0.07	0	0.04	1.29	1.79	0.11	0
SMTP (25)	2.75	0.04	0.04	0.04	0	0.86	0.04	0	0.04	1.75

Table 3: The percentage of total victim VIPs hosting different services involved with different inbound attacks; all numbers are in %.

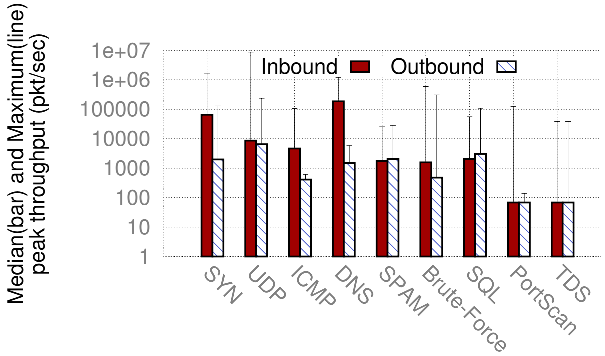


Figure 8: Median and maximum attack throughput across VIPs; the y-axis is on log-scale.

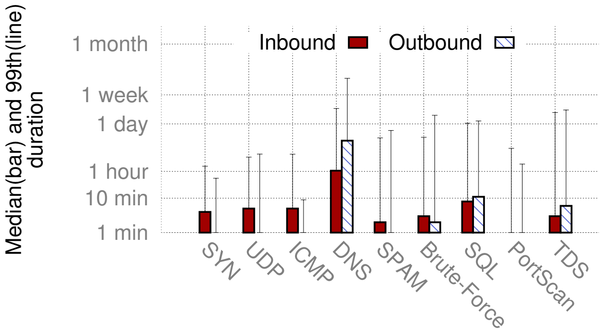


Figure 9: Median and 99th percentile of attack duration by attack type; the y-axis is on log-scale.

and median volumes. This implies that it would incur high costs and waste resources if we overprovision attack detection and mitigation solutions in hardware boxes. In comparison, elastic approaches that dynamically adjust resource usage over time may be a more cost-effective and efficient solution. The outbound attack throughput variations are relatively smaller, but for TCP SYN floods and TDS attacks, we still see a 20x-30x difference between the peak and median volumes.

Throughput per VIP: Today, cloud operators mostly focus on preventing large-volume attacks that may affect the cloud infrastructure, but they rely on tenants to secure their own services [47]. However, many of the attacks we investigated are smaller attacks targeting individual VIPs. Therefore, we study the peak attack throughput for individual VIPs and characterize the throughput differences across VIPs (Figure 8) to understand the resources individual VIPs need to defend against attacks.

We observe that some VIPs experience a high peak volume of attacks at a certain time (ranging from 100 pps to 8.7 Mpps). At times, the per-VIP peak volume is even higher than the median throughput for the entire cloud. For example, a single VIP can experience up to 8.7 Mpps inbound UDP floods. The per VIP inbound TCP SYN flood has a peak of 1.7 Mpps. We found one inbound TCP SYN flood that caused a CPU spike at the software load balancer (SLB) appliance and resulted in a restart of that appliance. However, the traffic from that device was quickly and automatically shifted to other SLBs.

There are large differences in the throughput volumes across VIPs. For example, for inbound brute-force attacks, the VIP having the peak throughput has 361 times larger volume than the VIP with the median value; for outbound brute-force attacks, this ratio is 75. Therefore, it may become too expensive to over-provision hardware security appliances for individual VIPs based on their maximum attack volumes. In comparison, resource management mechanisms that multiplex the attack protection resources across VIPs are likely to be more cost-effective.

Finally, we observe that for volume-based attacks (TCP SYN, UDP, ICMP, and DNS reflections), the peak volume of inbound attacks is 13 to 238 times higher than that of outbound. This is caused by the differences in attack resources and targets between the Internet and the cloud. For inbound attacks, there are more resources to leverage in the Internet (e.g., botnets, easily compromised personal machines) to launch high-volume attacks. These attacks also need to have high volumes to break the VIPs in the cloud, which have plenty of bandwidth and CPU resources. In contrast, outbound attacks can only leverage a few VMs in the cloud, because it is hard to compromise a large number of VMs or create multiple fake accounts to get many free VIPs.

5.2 Attack duration and inter-arrival time

Attack duration: Figure 9 shows that both inbound and outbound attacks have short duration with a median value within 10 minutes. This is consistent with other studies of Internet attacks [38, 40]. Interestingly, port scan attacks have a median duration within one minute for both inbound and outbound (but they can last for 100 minutes at the 99th percentile). There are several reasons that an attacker may launch short-duration attacks towards or from the cloud: (1) Shorter attacks are harder to be detected by cloud operators; (2) an attacker may attack one target for a short time and if not successful move quickly to another target. As a result, it is important to detect such short duration attacks in a timely fashion.

At the 99th percentile, most attacks have a duration longer than 80 minutes or even for days. For example, the TCP SYN flood attack lasts for 85 minutes at the 99th percentile. This is shorter than the previous study of Internet attacks [40], which observed

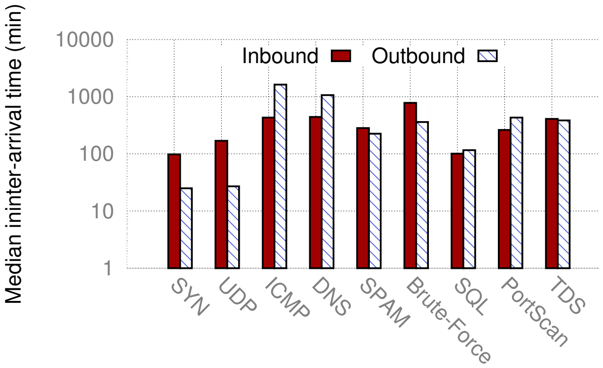


Figure 10: Median and 99th percentile of attack inter-arrival time by attack type; the y-axis is on log-scale.

that 2.4% of the attacks take more than five hours. This could be due to better security support in the cloud.

DNS reflection attacks last longer than others in both inbound and outbound directions. These attacks can sustain for a long time before being detected because they leverage many DNS resolvers simultaneously and each resolver receives relatively low query rate. Thus, it is hard to detect these attacks.

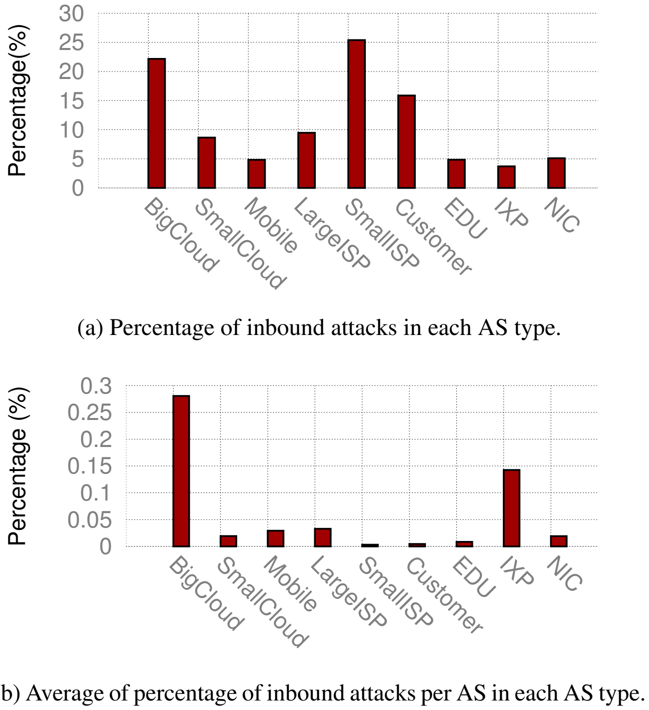
Ramp-up time: For volume-based attacks, we calculate the ramp-up time of an attack from its start time to the time when the packet rate grows to 90% of its peak. We observe a median ramp up time of 2-3 minutes for inbound attacks and 1 minute for outbound attacks. Today's flood detection solutions take about 5 minutes to detect the attacks [17, 48], and thus they are not fast enough to fully eliminate the flood attacks before they ramp up to affect the target with their peak strengths.

Inter-arrival time: We measure the inter-arrival time as the interval between the start times of two consecutive attacks to/from the same VIP. Figure 10 shows that most types of attacks have a median inter-arrival time of hundreds of minutes. The outbound TCP SYN and UDP flood attacks are shorter than inbound (about 25 minutes vs. 100 minutes). This indicates that malicious VIPs launch periodic attacks frequently. Attack protection systems can leverage such repeated attacks to identify and tune the right signatures for filtering attack traffic.

We identify two types of UDP flood attacks based on the correlations of inter-arrival time and peak attack size. 81% of the attacks have a median peak size with 8 Kpps but with large inter-arrival time (a median of 226 min). The rest 19% of the attacks have a median peak size with 457 Kpps, but with short inter-arrival time (a median of 95 min). The first type of small-scale, occasional attacks are relatively hard to distinguish from normal traffic and thus they are hard to mitigate without significant collateral damage. In contrast, the large-scale, frequent attacks require the cloud security operators to provision more resources to detect their traffic signatures and mitigate them.

6. INTERNET AS ANALYSIS

In this section we investigate the types of Internet ASes that are commonly involved in attacks to the cloud and that are under attacks from the cloud.



(a) Percentage of inbound attacks in each AS type.

(b) Average of percentage of inbound attacks per AS in each AS type.

Figure 11: Different types of ASes generating inbound attacks.

6.1 Inbound attacks

Are source IPs spoofed? We investigate whether the Internet IPs of inbound attacks are spoofed to understand the effectiveness of blacklisting on preventing different inbound attacks. Similar to prior work [40], we leverage the Anderson-Darling test (A2) [43] to determine if the IP addresses of an attack are uniformly distributed (i.e., an attack has spoofed IPs if A2 value is above 0.05). We observe that 67.1% of the TCP SYN floods have spoofed IPs. This is contrary to the study in 2006 [38] which observed that most flood attacks are not spoofed.

AS classification: We first remove those spoofed IPs and then map the IP addresses of inbound attack sources and outbound attack targets to AS numbers using Quova [11]. We use AS taxonomy repository [27] to identify AS types, which include large ISPs, small ISPs, customer networks, universities (EDU), Internet exchange points (IXP), and network information centers (NIC). We further classify big cloud (i.e., Google, Microsoft, Amazon), small cloud (i.e., web hosting services), and mobile ASes based on the AS descriptions. We count the number of attack incidents of different types for each AS class if any of its IP is involved in the attack.

Figure 11a shows the distribution of attacks across different types of ISPs. We observe that small ISPs and customer networks generate 25.4% and 15.9% of the attacks, respectively. For instance, an ISP in Asia contributed to 3.53% of the total attack packets. This is probably because these local or regional ISPs have relatively less security expertise and weak defense systems, and thus they are more likely to be compromised and leveraged by attackers.

When we calculate the average of percentage of attacks per AS (Figure 11b), we observe that there are more attacks per AS from big cloud and IXP. Individual ASes in small ISPs and customer networks do not generate many attacks on average.

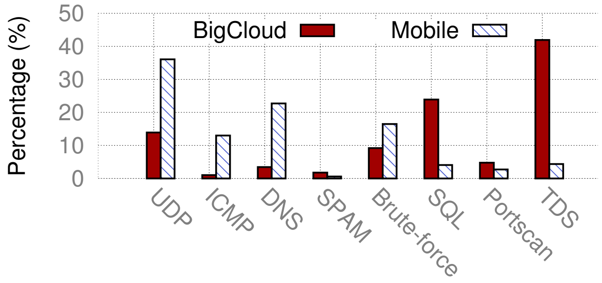


Figure 12: Percentage of inbound attacks from big clouds and mobile ASes in each attack type.

Attacks from big clouds: Figure 12 shows the distribution of attack types that originated from big clouds. UDP floods, SQL injections, and TDS attacks are the majority types. This is probably due to the availability of a large set of resources in big clouds to generate a high traffic volume and a large number of connections. In fact, big clouds contribute to 35% of TDS attacks with just 0.21% of TDS IPs.

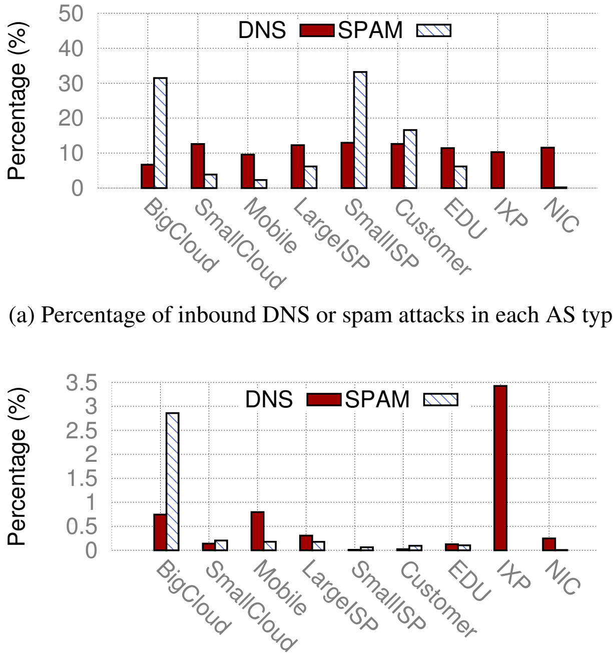
Attacks from mobile and wireless ASes: With the growth of mobile devices, attackers can try to compromise and exploit their resources for malicious activities. Given the relatively weaker software model in mobile devices compared to desktop PCs and the wide deployment of third-party apps on them, they are more likely to be compromised by malware for launching attacks. Users may also jailbreak the security restrictions and install tools (e.g., AnDOSid or mobile LOIC) to participate in botnet activities [10]. In fact, there are 2.1% of the inbound attack traffic from mobile networks.

Figure 12 shows that mobile networks mainly generate UDP floods, DNS reflections, and brute-force attacks. These attacks are harder to mitigate because simple source-based blacklisting does not work well for mobile devices. This is because most mobile devices are often located behind a NAT. While NAT may become less common with IPv6 adoption, there would be more ephemeral addresses.

Origins of DNS attacks: Figure 13a shows that the cloud we studied received a similar number of DNS attacks from all types of ASes. Figure 13b shows that if we count per AS attacks, there are more DNS attacks from IXPs. Our further investigation shows that each DNS attack involved a median value of only 17 unique DNS resolvers in the NetFlow records.

Origins of spam: Figure 13a shows that spam attacks are mainly from large cloud, small ISPs, and customer networks. For example, 81.0% of the spam packets are from Amazon Web Services (AWS) [2] in Singapore.[5] However, each individual small ISP or customer network does not generate many attacks as indicated by the number of per AS attacks shown in Figure 13b. This indicates that it is easier for attackers to leverage the free trial accounts in large clouds, the end hosts in small ISPs, and customer networks to generate spams. Prior study of spams in the Internet [46] shows many spams come from network information centers (NIC), but we observed only a single attack from NICs in our data.

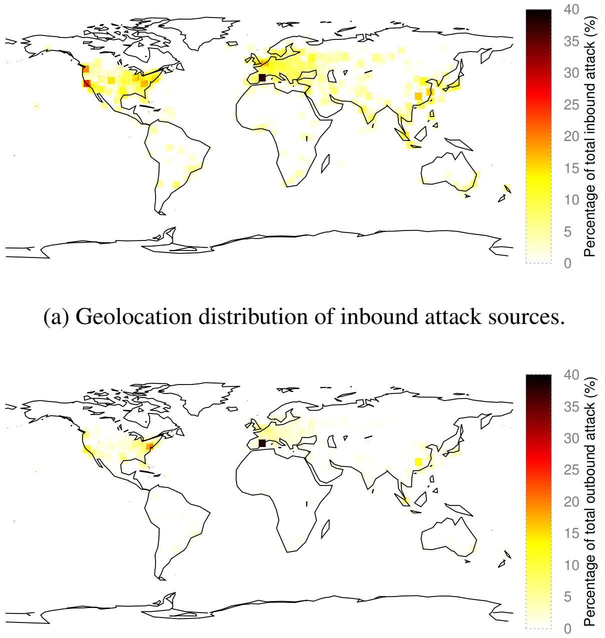
Geolocation distribution of inbound attacks: Figure 14a shows the geographical distribution of inbound attack sources. The inbound attack sources are spread mainly across places in Europe, Eastern Asia, and North America. Specifically, there is one AS in



(a) Percentage of inbound DNS or spam attacks in each AS type.

(b) Average percentage of inbound DNS or spam attacks per AS in each AS type.

Figure 13: Different types of ASes generating inbound DNS and SPAM attacks.

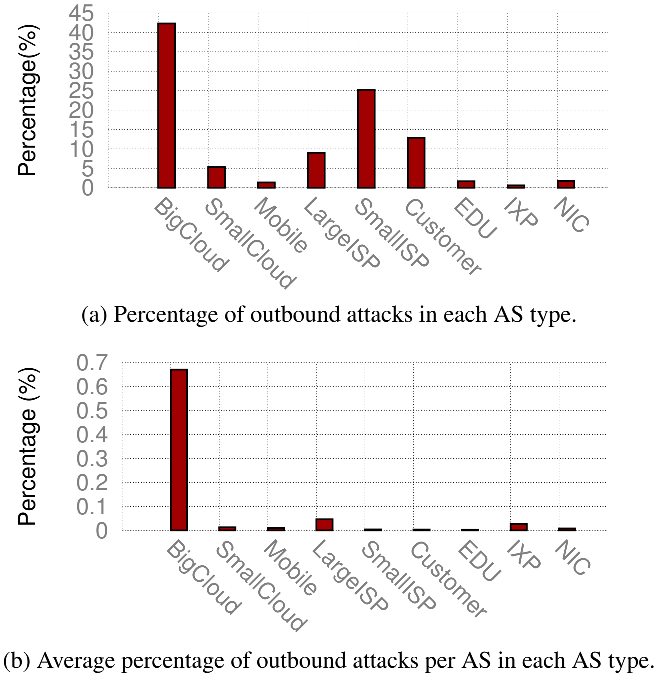

0
5
10
15
20
25
30
35
40
Percentage of total inbound attack (%)
0
5
10
15
20
25
30
35
40
Percentage of total outbound attack (%)

(a) Geolocation distribution of inbound attack sources.

(b) Geolocation distribution of outbound attack targets.

Figure 14: Attack geolocation distribution.

[5] We did not validate these spam attacks with AWS.



(a) Percentage of outbound attacks in each AS type.

(b) Average percentage of outbound attacks per AS in each AS type.

Figure 15: Different types of ASes targeted by outbound attacks.

Spain involved with more than 35% of the total inbound attacks. It mainly generated UDP flood, TDS and SQL attacks. There are ASes from the west coast of North America that are involved with more than 20% of the total inbound attacks.

6.2 Outbound attacks

Are outbound attacks clustered? Unlike Internet floods which often target a single host [48], we observe outbound UDP flood target a median of 8 hosts, while TCP SYN floods often target a median of 25 Internet hosts just in our sampled NetFlow data. This means attackers often use cloud resources to attack a group of hosts instead of individual IPs. We count the number of unique victim IPs of outbound attacks in each AS to understand if the victims are clustered on particular ASes. We find that 80% of the attacks target hosts in a single AS.

While prior work has shown that a small number of ASes are involved in a significant fraction of attacks in ISP networks [38] and distributed intrusion detection system [50], we show that cloud-related attack incidents are widely spread across many ASes. Top 10 ASes are targets of 8.9% of the attacks; top 100 ASes are targets of 16.3% of the attacks. However, there is a small portion of attacks responsible for the major attack traffic. For instance, 40% of the outbound attack packets were directed from three VIPs towards a small cloud AS in Romania, which offers web hosting, online gaming, and VPN services.

AS classes of outbound attacks: Figure 15 shows that 42% of outbound attacks are against services in big clouds. Most of these attacks are SQL injection and TDS attacks. Small ISPs and customer networks face 25% and 13% of the outbound attacks (Figure 15a), but individual ASes do not generate many attacks (less than 0.01% of the total outbound attacks per AS) (Figure 15b). Small ISPs and customer networks are the major target for brute-force attacks and spam. This is probably because these networks often lack strong security support. For example, 23.6% of the outbound DNS reflection attack packets are sent to an ISP in France. It is therefore important to coordinate security measures across the cloud infrastructure and these networks to protect against these attacks.

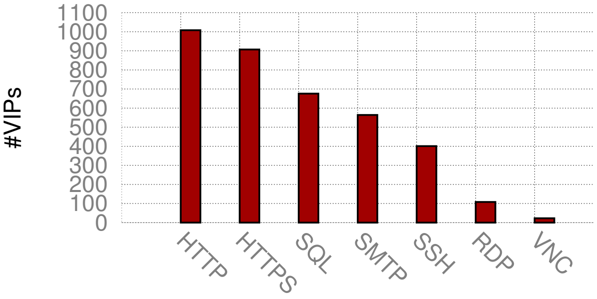


Figure 16: Top Internet applications under outbound attacks.

There are only a few brute-force attacks against mobile networks (1.4%). This may be because mobile devices are often behind a NAT preventing them from unsolicited connections, which makes it harder for attacks to get through.

Internet applications under attack: To understand the Internet applications under attacks, we investigate the destination port of outbound traffic coming from the VIPs generating outbound attacks (Figure 16). We find that most of the outbound attacks target web services (HTTP/HTTPS together form 64.5% of the attack VIPs involved with outbound attacks). For example, 69% of the outbound UDP floods use port 80 as the destination port targeting HTTP services. The other popular target services are SQL, SMTP, and SSH.

Geolocation distribution of outbound attacks: Outbound attack targets are mainly spread across places in Europe and North America (Figure 14b). There are less outbound attack targets than inbound attack sources in Eastern Asia. The same AS in Spain we discussed in inbound attacks also receives more than 35% of outbound attacks (mainly with brute-force, TDS, and SQL attacks).

7. EXISTING SECURITY PRACTICES

In this section we discuss how existing cloud security solutions handle the types of attacks observed in our study.

Inbound TCP SYN, UDP, ICMP floods, DNS reflection attacks: Cloud VMs are only accessible through virtual IPs (VIPs). Traffic towards VIPs is routed through a load-balancing infrastructure [42]. At the infrastructure, the cloud can monitor, detect and mitigate flood attacks (e.g., using SYN cookies, rate limiting, and connection limits) to help ensure that such attacks do not impact customer VM instances [7]. The tenants can also leverage scale-out (e.g., adding more VMs on demand) or scale-up (e.g., deploying resource-rich VMs) techniques to raise the bar for attacks [4].

There are multiple in-built mechanisms in cloud systems to safeguard tenants. For example, the hypervisor that hosts guest VMs in the cloud is neither directly addressable internally by other tenants nor it is externally addressable. Additional filters are put in place to block broadcast and multicast traffic, with the exception of what is needed to maintain DHCP leases. Inside the VMs, tenants can further enable web server add-ons that protect against certain DoS attacks [41]. For example, for TCP SYN flooding attacks, a security rule can be specified to track significant deviations from the norm in terms of the number of half-open TCP sessions, and then drop any further TCP SYN packets from specified sources [5].

Port scan: Unauthorized port scans are often viewed as a violation of cloud use policy and they are blocked by the cloud provider [3]. Port scans are likely to have limited effectiveness because, by default, all the inbound ports on VMs are closed except the ones opened by tenants; a tenant can author a service definition file that contains the internal endpoints that should be opened in service VMs and what roles can communicate with them [8]. Tenants can also use security groups or firewalls to further block unauthorized traffic [41].

Other inbound attacks: Some cloud providers choose not to actively block network traffic affecting individual tenants because the infrastructure does not interpret the expected behavior of customer applications. Instead, these cloud systems allow tenants to use firewall proxy devices such as Web Application Firewalls (WAFs) that terminate and forward traffic to endpoints running in a VM. Tenants can also use network ACLs or VLANs to prevent packets from certain IP addresses from reaching VMs [7].

Cloud providers can also leverage high-level signals emitted from workloads running on the cloud from customer accounts. One such signal is the number of open ports. Legitimate customers aim to minimize the susceptibility of their applications running in the cloud to any external attacks, and hence deploy services usually with a limited number of open ports (e.g., HTTP port 80). However, compromised accounts or VMs may perform a variety of anomalous activities such as running a Botnet controller or torrent services on multiple open ports. The cloud security team monitors such activities and aggressively shuts down any misbehaving tenant VMs.

Outbound attacks: To mitigate outbound attacks, the most important step is to identify fraudulent VMs when a tenant sets up a subscription. In the cloud, several anti-fraud techniques are used such as credit card validation, computing the estimated geographical distance from the IP address used for login to the billing address and ensuring it is within reasonable bounds, and determining whether the email address of the purchaser is from a free email provider.

Note that attackers can also exploit vulnerabilities in VMs of legitimate customers. If an attack is successful, the compromised VMs can then be used by attackers for malicious activity e.g., to send large amounts of unsolicited mail (spam). The cloud provider can enforce limits on how many emails a VM can send as well as prevent SMTP open relay, which can be used to spread spam [5].

8. RELATED WORK

This paper presented one of the first large-scale studies to investigate the prevalence of network-based attacks in the cloud. We compare our work with related work on detecting and understanding Internet-based attacks.

Attack detection methods: Previous works have used NetFlow logs to understand DDoS attacks and traffic anomalies [38, 22, 36, 35] in ISP networks. Our work takes a similar approach to understand a broader set of attacks in the cloud. Most previous studies on application-level attacks leverage analysis of the application content (e.g., spam [15, 25], SQL injection [18], SSH [32]). Our work shows that it is possible to detect some of these attacks by leveraging network-level signatures such as volumes, spread, TCP flags, and communication patterns. In fact, previous works have also shown that application-level attacks (e.g., spam) have strong network-level signatures [46]. We validate our network-based detection by comparing the detected attacks against the security appliance alerts and incident reports. Although network-based detection may not capture all types of application-level attacks (e.g., malware), they are more pragmatic to implement in today's cloud monitoring infrastructure.

Attack characterization: There is a large body of work on characterizing Internet-based attacks. Most prior efforts (e.g., [10, 13, 17, 40, 46, 51]) focus on one or a few types of attacks in the Internet. Given the importance of cloud services in today's Internet, understanding the attacks from/to the cloud is critical. Our study investigates a wide diversity of inbound and outbound attacks in the cloud. We differentiate DDoS attacks based on their protocols (TCP SYN, UDP, DNS reflection), and show that other types of attacks (e.g., brute-force, port scans, and TDS) also need cloud operator's attention. Moreover, we show the detailed characteristics of attacks in the cloud such as the cloud services affected by the attacks, the Internet origins and targets, and the intensity and frequency of these attacks. These results can provide guidelines for future design of attack detection and mitigation systems for the cloud.

9. CONCLUSION

We investigated the prevalence of network-based attacks both on and off the cloud. We performed the first measurement study of the characteristics of a wide range of cloud attacks that vary in complexity, intensity, duration and distribution. Our study shows a strong evidence of increasing scale, attack volume, and sophistication of these attacks. Our results have been leveraged by the cloud security team towards identifying correlations and improving mitigations for different attack types. We hope that this study motivates future research towards designing attack detection and mitigation systems for the cloud. In future work, we plan to extend our measurement study to analyze application level attacks, compare across attack categories and leverage packet traces for deeper analysis.

Acknowledgment

We thank Ramesh Govindan, Cormac Herley, Jelena Mirkovic, Vivek Narassaya, Jitu Padhye, Jay Stokes, Alec Wolman, and the anonymous reviewers for their insightful feedback and suggestions. This work is supported in part by the NSF grants CNS-1423505, CNS-1413972, and DHS-D15PC00184.

10. REFERENCES

[1] http://www.everestgrp.com/2015-04-40-billion-global-cloud-services-market-expected-to-grow-27-percent-per-annum-for-next-3-years-press-release-17218.html.

[2] Amazon web services. http://aws.amazon.com/.

[3] Amazon Web Services: Overview of Security Processes. https://d0.awsstatic.com/whitepapers/Security/AWS%20Security%20Whitepaper.pdf.

[4] AWS Best Practices for DDoS Resiliency. https://d0.awsstatic.com/whitepapers/DDoS_White_Paper_June2015.pdf.

[5] AWS Security Best Practices. https://d0.awsstatic.com/whitepapers/aws-security-best-practices.pdf.

[6] http://nmap.org/book/man-port-scanning-techniques.html.

[7] Microsoft Azure Network Security Whitepaper. http://blogs.msdn.com/b/azuresecurity/archive/2015/03/03/microsoft-azure-network-security-whitepaper-version-3-is-now-available.aspx.

[8] https://msdn.microsoft.com/en-us/library/azure/ee758711.aspx.

[9] juno.c. http://goo.gl/i1Qodc, 2013.

[10] Q4 2013 global ddos attack report. http://goo.gl/lIyRmK, 2013.

[11] Quova. http://www.quova.com, 2013.

[12] S. Ali, I. U. Haq, S. Rizvi, N. Rasheed, U. Sarfraz, S. A. Khayam, and F. Mirza. On Mitigating Sampling-induced Accuracy Loss in Traffic Anomaly Detection Systems. *ACM SIGCOMM Computer Communication Review 2010*.

[13] M. Allman, V. Paxson, and J. Terrell. A Brief History of Scanning. In *IMC*, 2007.

[14] Andrew Marshall, Michael Howard, Grant Bugher, Brian Harden. Security Best Practices For Developing Windows Azure Applications. http://download.microsoft.com/documents/uk/enterprise/88_security_best_practices_for_developing_windows_azure_applicat.pdf.

[15] I. Androutsopoulos, J. Koutsias, K. V. Chandrinos, G. Paliouras, and C. D. Spyropoulos. An Evaluation of Naive Bayesian Anti-Spam Filtering. *Proceedings of the workshop on Machine Learning in the New Information Age, 11th European Conference on Machine Learning*, 2000.

[16] C. Anley. Advanced SQL Injection in SQL Server Applications. In *Next Generation Security Software Ltd*, 2002.

[17] Arbor Networks. Insight into the global threat landscape. http://goo.gl/15oOx3, February 2013.

[18] S. Bandhakavi, P. Bisht, P. Madhusudan, and V. N. Venkatakrishnan. CANDID: Preventing SQL injection attacks using dynamic candidate evaluations. ACM CCS, 2007.

[19] M. Basseville and I. Nikiforov. Detection of Abrupt Changes: Theory and Application. Prentice Hall Englewood Cliffs, 1993.

[20] Bloomberg. Sony Network Breach Shows Amazon Clould's Appeal for Hackers. http://goo.gl/3WiAaj, 2011.

[21] G. Booth, A. Soknacki, and A. Somayaji. Cloud Security: Attacks and Current Defenses. In *ASIA*, 2013.

[22] D. Brauckhoff, B. Tellenbach, A. Wagner, M. May, and A. Lakhina. Impact of packet sampling on anomaly detection metrics. In *IMC*, 2006.

[23] J. Caballero, C. Grier, C. Kreibich, and V. Paxson. Measuring pay-per-install: The commoditization of malware distribution. In *USENIX Conference on Security*, 2011.

[24] E. Cambiaso, G. Papaleo, and M. Aiello. Taxonomy of Slow DoS Attacks to Web Applications. In *TCNDSS*. Springer, 2012.

[25] X. Carreras and L. Marquez. Boosting Trees for Anti-Spam Email Filtering. *Proceedings of RANLP*, 2001.

[26] R. Chaiken, B. Jenkins, P.-A. Larson, B. Ramsey, D. Shakib, S. Weaver, and J. Zhou. SCOPE: Easy and Efficient Parallel Processing of Massive Data Sets. *VLDB'08*.

[27] X. Dimitropoulos, D. Krioukov, G. Riley, and k. claffy. Revealing the Autonomous System Taxonomy: The Machine Learning Approach. In *Passive and Active Network Measurement Workshop (PAM)*, 2006.

[28] N. R. Draper, H. Smith, and E. Pownell. *Applied regression analysis*. Wiley New York, 1966.

[29] Google. Malware Distribution by Autonomous System. http://goo.gl/mZQeG4, 2013.

[30] C. Grier, L. Ballard, et al. Manufacturing Compromise: The Emergence of Exploit-As-A-Service. In *CCS*, 2012.

[31] L. Hellemons. Flow-based detection of ssh intrusion attempts. *Scanning*, 2012.

[32] M. Javed and V. Paxson. Detecting Stealthy, Distributed SSH Brute-forcing. CCS, 2013.

[33] C. Kanich, N. Weaver, D. McCoy, T. Halvorson, C. Kreibich, K. Levchenko, V. Paxson, G. M. Voelker, and S. Savage. Show Me the Money: Characterizing Spam-Advertised Revenue. In *USENIX SEC*, 2011.

[34] R. Kawahara, K. Ishibashi, T. Mori, N. Kamiyama, S. Harada, and S. Asano. Detection accuracy of network anomalies using sampled flow statistics. In *GLOBECOM '07. IEEE*.

[35] A. Lakhina, M. Crovella, and C. Diot. Diagnosing network-wide traffic anomalies. In *SIGCOMM*, 2004.

[36] A. Lakhina, M. Crovella, and C. Diot. Mining anomalies using traffic feature distributions. In *SIGCOMM*, 2005.

[37] Z. Li, S. Alrwais, Y. Xie, F. Yu, and X. Wang. Finding the Linchpins of the Dark Web: A Study on Topologically Dedicated Hosts on Malicious Web Infrastructures. In *Security and Privacy (SP), IEEE Symposium on*, 2013.

[38] Z. M. Mao, V. Sekar, O. Spatscheck, J. van der Merwe, and R. Vasudevan. Analyzing large DDoS attacks using multiple data sources. In *SIGCOMM Workshop on Large-scale Attack Defense*, 2006.

[39] J. Mirkovic and P. Reiher. A Taxonomy of DDoS Attack and DDoS Defense Mechanisms. *SIGCOMM CCR*, 2004.

[40] D. Moore, C. Shannon, D. J. Brown, G. M. Voelker, and S. Savage. Inferring Internet Denial-Of-Service Activity. *ACM Transactions on Computer Systems*, 2006.

[41] C. Nelson. Best practices to protect your azure deployment against "cloud drive-by" attacks. http://blogs.msdn.com/b/azuresecurity/archive/2015/07/05/best-practices-to-protect-your-azure-deployment-against-cloud-drive-by-attacks.aspx.

[42] P. Patel, D. Bansal, L. Yuan, A. Murthy, A. Greenberg, D. A. Maltz, R. Kern, H. Kumar, M. Zikos, H. Wu, et al. Ananta: Cloud Scale Load Balancing. In *ACM SIGCOMM*, 2013.

[43] V. Paxson, G. Almes, J. Mahdavi, and M. Mathis. Framework for IP performance metrics. In *RFC 2330*, 1998.

[44] A. Pitsillidis, C. Kanich, et al. Taster's Choice: A Comparative Analysis of Spam Feeds. In *IMC*, 2012.

[45] R. Potharaju, N. Jain, and C. Nita-Rotaru. Juggling the jigsaw: Towards automated problem inference from network trouble tickets. In *NSDI'13*.

[46] A. Ramachandran and N. Feamster. Understanding the network-level behavior of spammers. In *SIGCOMM*, 2006.

[47] B. Ridgway. Security best practices for windows azure solutions. *Azure Manual*, 2014.

[48] V. Sekar, N. G. Duffield, O. Spatscheck, J. E. van der Merwe, and H. Zhang. LADS: Large-scale automated DDoS detection system. In *USENIX ATC*, 2006.

[49] B. Stone-Gross, T. Holz, G. Stringhini, and G. Vigna. The underground economy of spam: A botmaster's perspective of coordinating large-scale spam campaigns. *USENIX Workshop on Large-Scale Exploits and Emergent Threats (LEET)*, 2011.

[50] V. Yegneswaran, P. Barford, and S. Jha. Global intrusion detection in the domino overlay system. In *NDSS*, 2004.

[51] V. Yegneswaran, P. Barford, and J. Ullrich. Internet intrusions: Global characteristics and prevalence. In *ACM SIGMETRICS 2003*.