Network Situational Awareness

Applying Concepts to a Specific Data Source

Chris Canning
Joan Downing
Chris King
Bob Weiland

12/7/2009

Executive Summary

We were tasked with performing an analysis using the MAWI Sample Point Data F data set provided through the WIDE Project. Our analysis was based on data collected over 24 hours of network traffic on the 144.44.0.0/16 class B network, which will be referred to as "our network". This network was chosen through analyzing the network traffic provided. Our network had the third highest flows and network traffic, so we figured this would provide us with enough data to perform a thorough analysis.

Our network was comprised of a number of Web Servers, DNS Servers, and Mail Servers, which were determined by analyzing flow traffic corresponding to services that would be hosted by the respective types of servers. Servers that processed a large amount of data on a known port would be more likely to host that service. In addition to servers, there were a number of client machines hosted on the network that had traffic similar to that of a normal user.

Reviewing the network flow traffic revealed a number of details about our network. First, there was a peak in traffic between the hours of 02:00 UTC and 09:00 UTC, which would indicate that the traffic analyzed could likely be for a large corporate network or a number of corporate networks. Also, our network was open and both clients and servers had an external face on the network.

In order to improve our network, we would provide a number or recommendations. We recommend the use of DHCP, bandwidth monitoring, virus scanning, verifying firewall configurations and using SSL for email transmissions. DHCP would limit the external view of the network. Bandwidth monitoring would help in load balancing the network. Our network contained a number of flows that were consistent with malware-infected machines; implementing network-wide virus scanning would help mitigate these issues. Additionally, the infected machines were sending traffic through known IRC ports, which should be closed, as well as other known vulnerable ports, on the firewall. Lastly, there were severs that still sent POP email ports commonly known to send text in the clear over port 110. Changing these POP servers to use SSL through port 995 would provide a higher level of privacy for these emails.

Finally, we also looked at the impact of a DDoS attack would have on our network. In order to do this, we analyzed the previous DDoS attack on Estonia and reviewed the technique used for this attack. We took this profile and placed it against our network. If a similar DDoS attack was made against our network in its current state, it would effectively impact the legitimate traffic on the network and would be likely to take down a number of routers on our network. Due to the externality of web servers, they would be the likely target. Routers would likely fail first during a DDoS attack against our network since they would be required to handle the increased load being sent to our network.

It is our recommendation that our network be aware of the risks and mitigate issues that we have found and also perform regular reviews of network flow traffic using SiLK Tools and other netflow analyzers. Understanding the current network traffic is beneficial for finding and eliminating anomalies and planning for disaster recovery.

Table of Contents

Introduction

In our report, we will first take a look at the profile of our network (144.44.0.0/16). This will discuss the hosts that make up our network, the key resources produced and consumed by our network, significant choke points, dependencies and the other features present in the network flows. Our profile will provide a view of how we determined the network was constructed as well as provide a recommendation for how to improve this structure. Lastly, we focus on the Internet-wide DDoS attack against Estonia and how a similar attack would affect our network.

Network Profile

Assumptions

During our review of the netflow traffic, it was necessary to make a number of assumptions when forming our view of the network structure. First, we assumed that every IP address was dedicated to a separate host. In most network settings, some servers are configured to use multiple IP addresses which helps load balance network traffic to different services. We also assumed that traffic sent through lesser used ports were malicious. This would include traffic sent through to known IRC ports. We also assumed that the flow traffic increase between 02:00 UTC and 09:00 UTC indicates an influx in normal traffic and is not indicative of an attack or anomalous behavior.

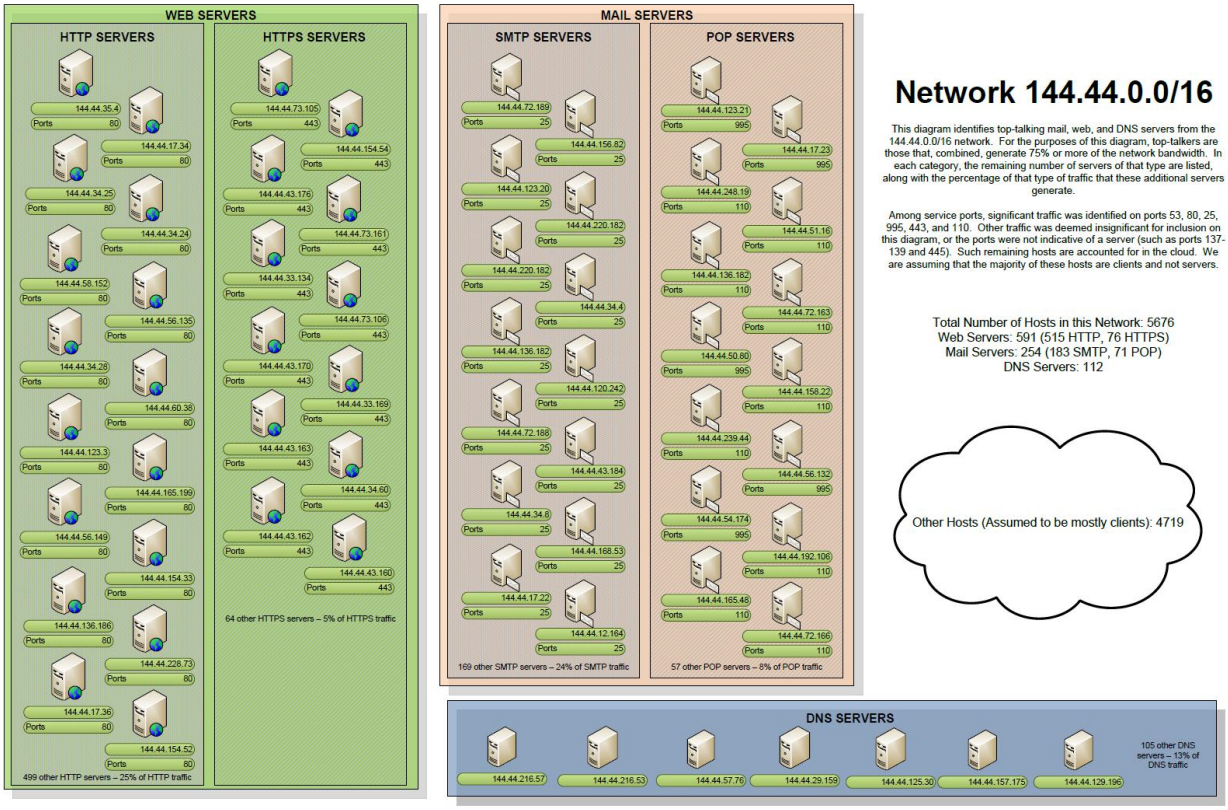
Network Diagram

Designing a complete diagram of a Class B network would be impractical, as our analysis identified flows from 5,676 unique hosts in the 144.44.0.0/16 network. Instead, we identified all of the service ports in use on our network. Then, we identified how many hosts were communicating from each of these service ports. Finally, a list of the top talkers on each of these ports was created so that we could plot them on our diagram.

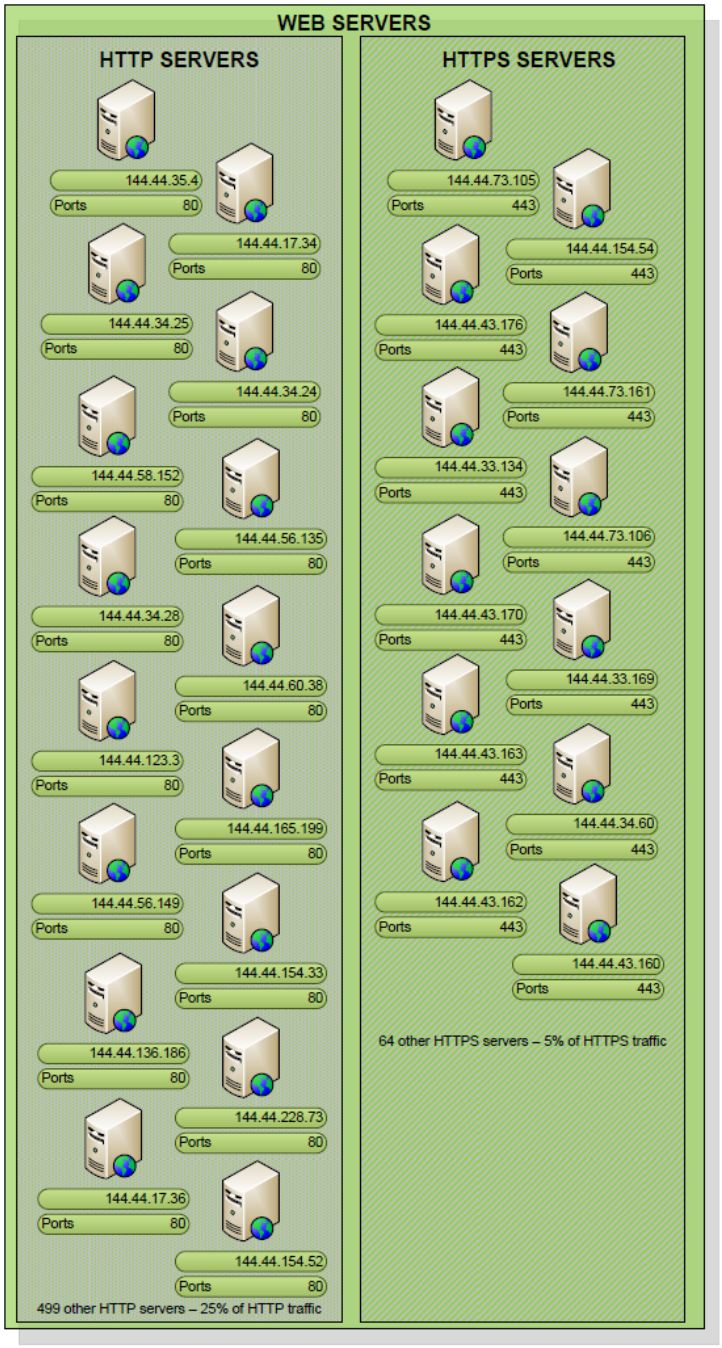
Our list of top service ports used on our network segment was created with the following SiLK query: `rwfilter --pass=stdout --saddress=144.44.0.0/16 --sport=1-1024 <<all traffic files>> | rwstats --count=25 --bytes -sport`
This query filters based on a source address in the network range and a source port in the service port range, and then outputs a list of the top 25 ports based on number of bytes sent. From this list, we determined that ports 80, 995, 110, 25, 443, and 53 were hosting services that we were interested in mapping and at a significant level, though Port 80 (HTTP) accounted for 90% of bytes within this filter group. We excluded ports such as Telnet and SSH from our analysis because they were not used significantly and are not necessarily the primary public service being offered by a host.

Each of the following segments of our overall diagram follows the same format. The top talkers which cumulatively account for 75% or more of traffic by bytes are plotted, while others are omitted for brevity. The icons are sorted from top to bottom and left to right by number of bytes sent from that port. The data for each generating diagram was produced by a SiLK command of the following format: `rwfilter --pass=stdout --saddress=144.44.0.0/16 --sport=<<80,443,25,995,110,53>> <<all traffic files>> | rwstats --count=25 --bytes --sip`
This command generates, based on the source port selected, a list of the 25 top-talkers by bytes sent for that particular protocol.

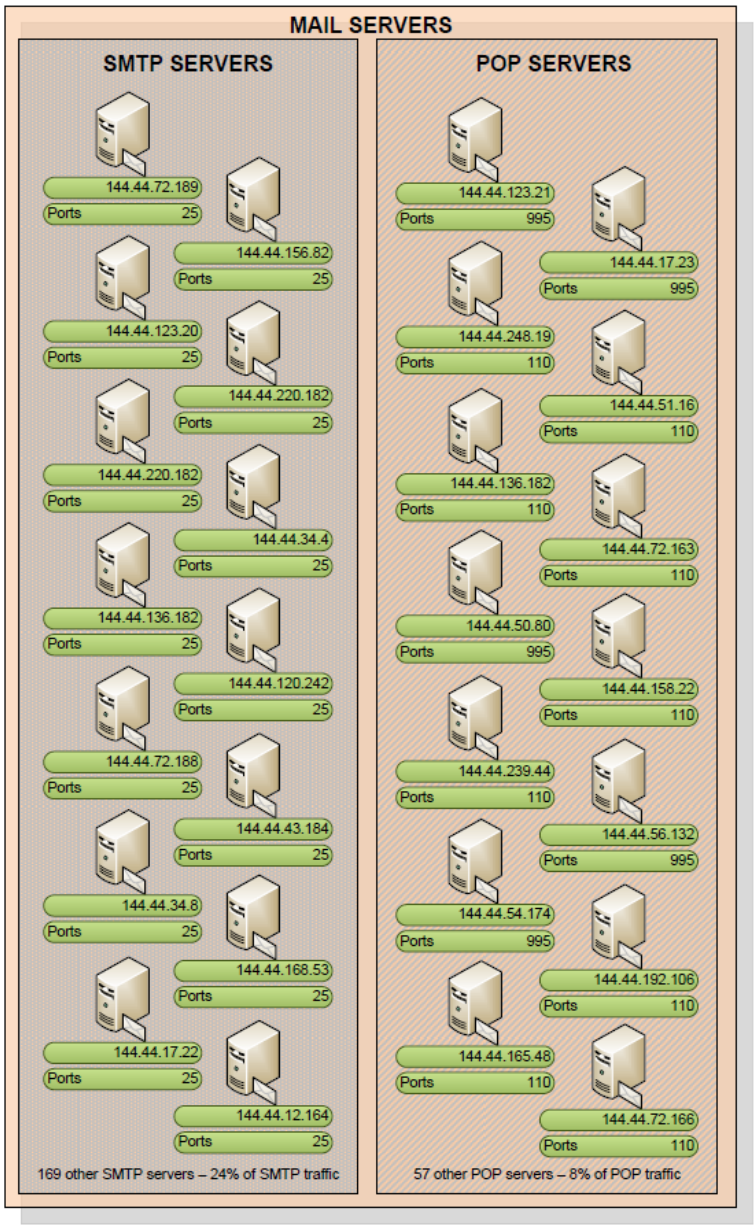
If we were unable to identify a significant service being offered from a host, it was assumed that this host was a likely a client instead of a server. These are accounted for in the cloud portion of our diagram.



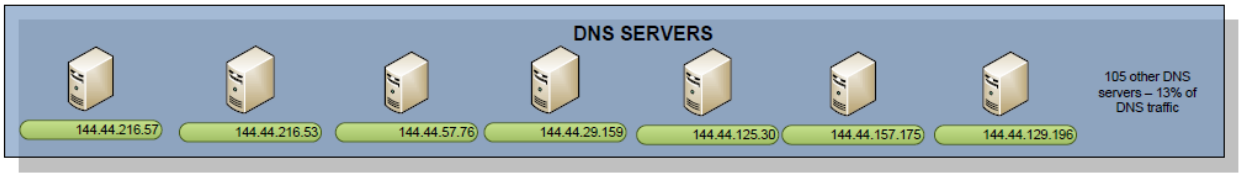
First, since web servers made up a large portion of traffic, we analyzed hosts which sent data from Port 80 or 443. Port 80 is used for standard HTTP communication, while port 443 is used for HTTPS, or HTTP with SSL encryption. Among the 515 servers using port 80, 5 of them accounted for 50% of all port 80 traffic by byte count, while only 1 HTTPS server out of 76 accounts for 63% of port 443 traffic. Combining the two types, there were a total of 591 web servers.



The next set of servers we diagrammed was mail servers. In total, 254 mail servers were observed. Again, we broke this larger category into two sets: SMTP servers and POP servers. SMTP (Simple Mail Transfer Protocol) is the primary mail transfer protocol which is used for sending mail between servers and by clients sending mail, while the POP (Post Office Protocol) is used by clients for receiving mail. Among the SMTP servers, 6 hosts account for 52% of SMTP traffic, out of a total of 183 hosts. However, among POP servers, only 3 hosts account for 67% of POP traffic, out of a total of 71 hosts. Another interesting observation about the POP servers is that the top two utilize SSL, but many others do not.



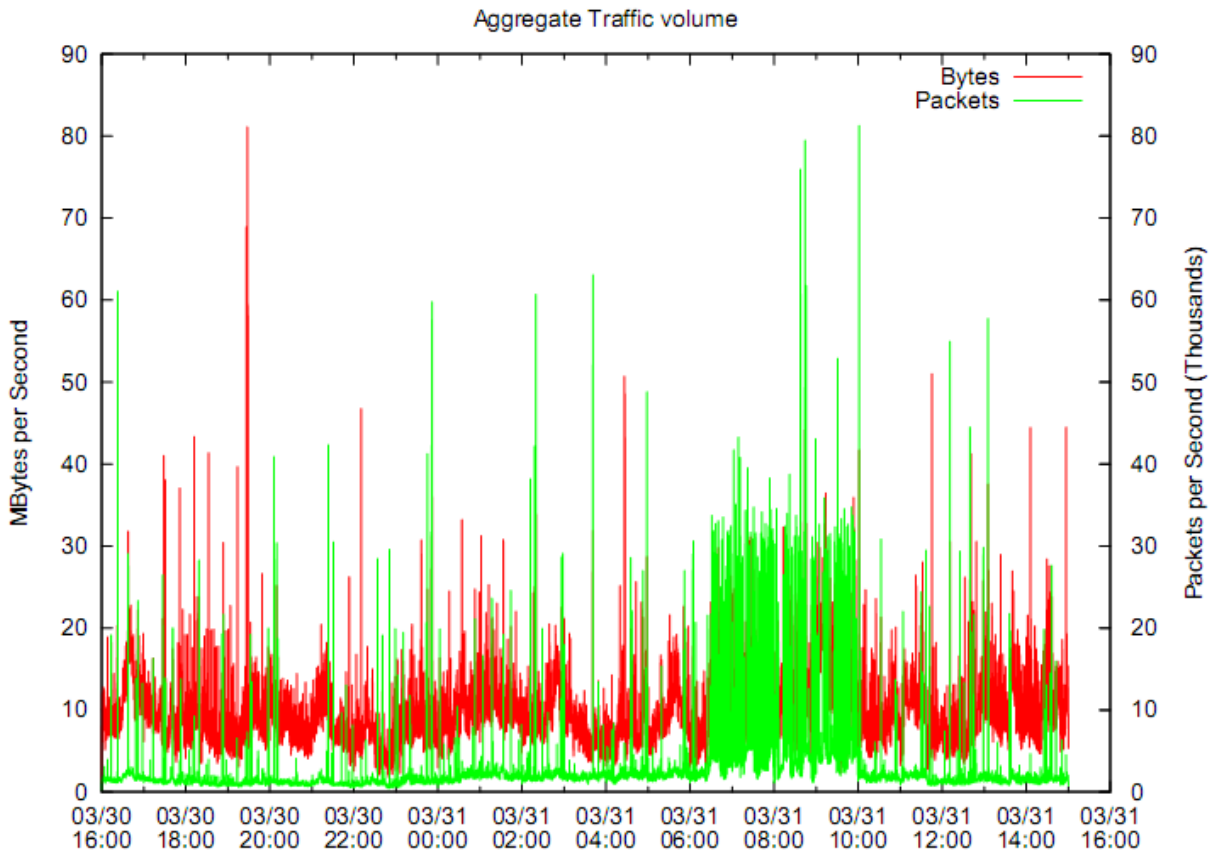
Finally, we analyzed the DNS servers identified on our network. While DNS did not produce nearly as much bandwidth as either web servers or mail servers, it was the most active protocol in terms of number of flows. This is not particularly surprising because many network transmissions first require a brief DNS session. On our network, a single DNS server accounts for 57% of DNS traffic, out of a total of 112 servers. The seven hosts plotted on our diagram combined consume 87% of DNS bandwidth.



Overall, our network appears to be fairly normal, offering both web and email services from a variety of servers. Few unusual protocols were identified as being used to any large extent. Thus, the network seems to resemble a large corporate network, possibly with several subsidiaries and decent control over the services running on the network. This would also explain the large number of web servers which account for almost 10% of all the hosts we detected traffic from on the 144.44.0.0/16 network. Additionally, we detected no traffic both originating and terminating in the 144.44.0.0/16 network, which indicates that it lies entirely on one side of the MAWI detection point.

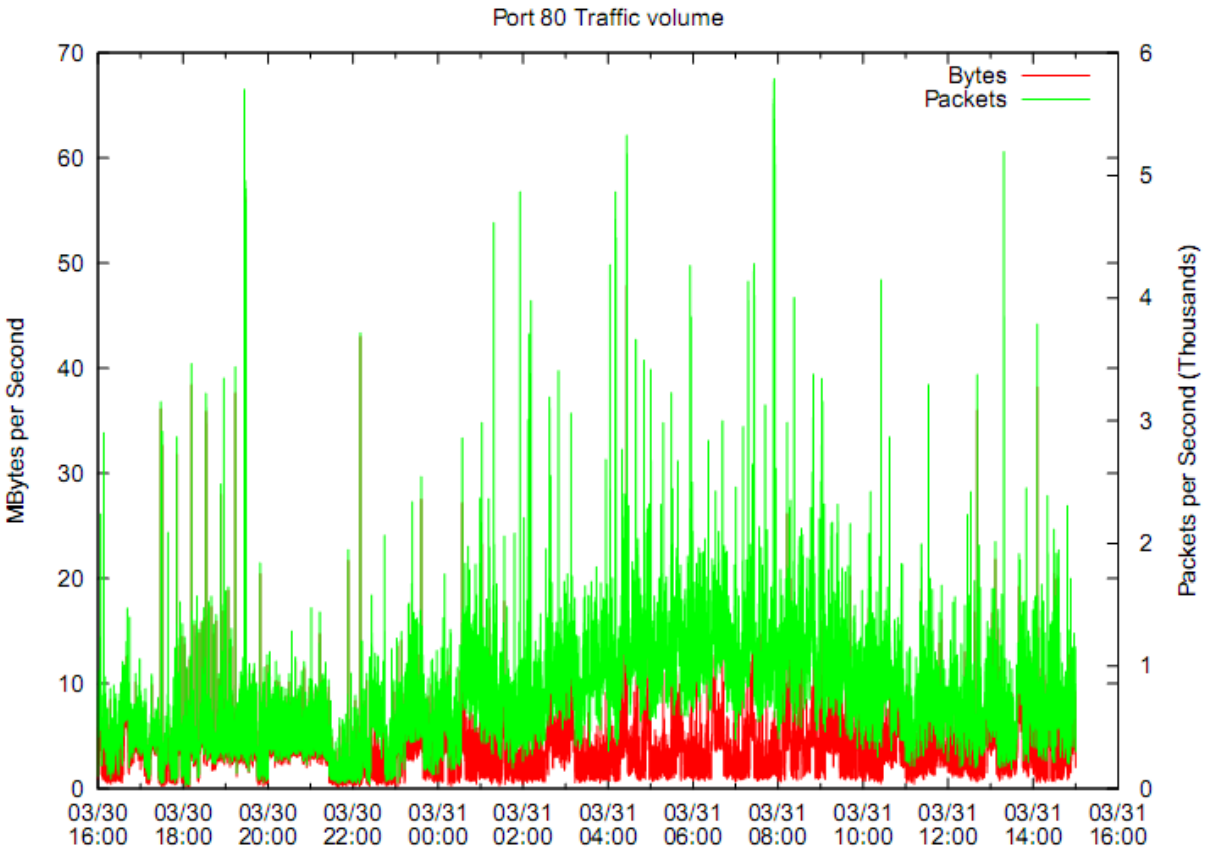
Network Flow Profile

In order to determine our network structure, we have analyzed both the aggregated flow volumes as well as specific flow traffic through the network. In viewing the volume of traffic, we were able to determine an eight hour period that had an escalated amount of traffic than the normal flows for the remaining sixteen hours of netflow traffic that was analyzed. The figure below is based on the aggregated traffic sent over the 24 hours period. The red lines indicate the amount of data sent per second and the green lines indicate the packets per second. This figure shows an increase in packets between 06:00 UTC and 10:00 UTC. Our flow analysis indicated that this was caused by an increase in ICMP traffic.



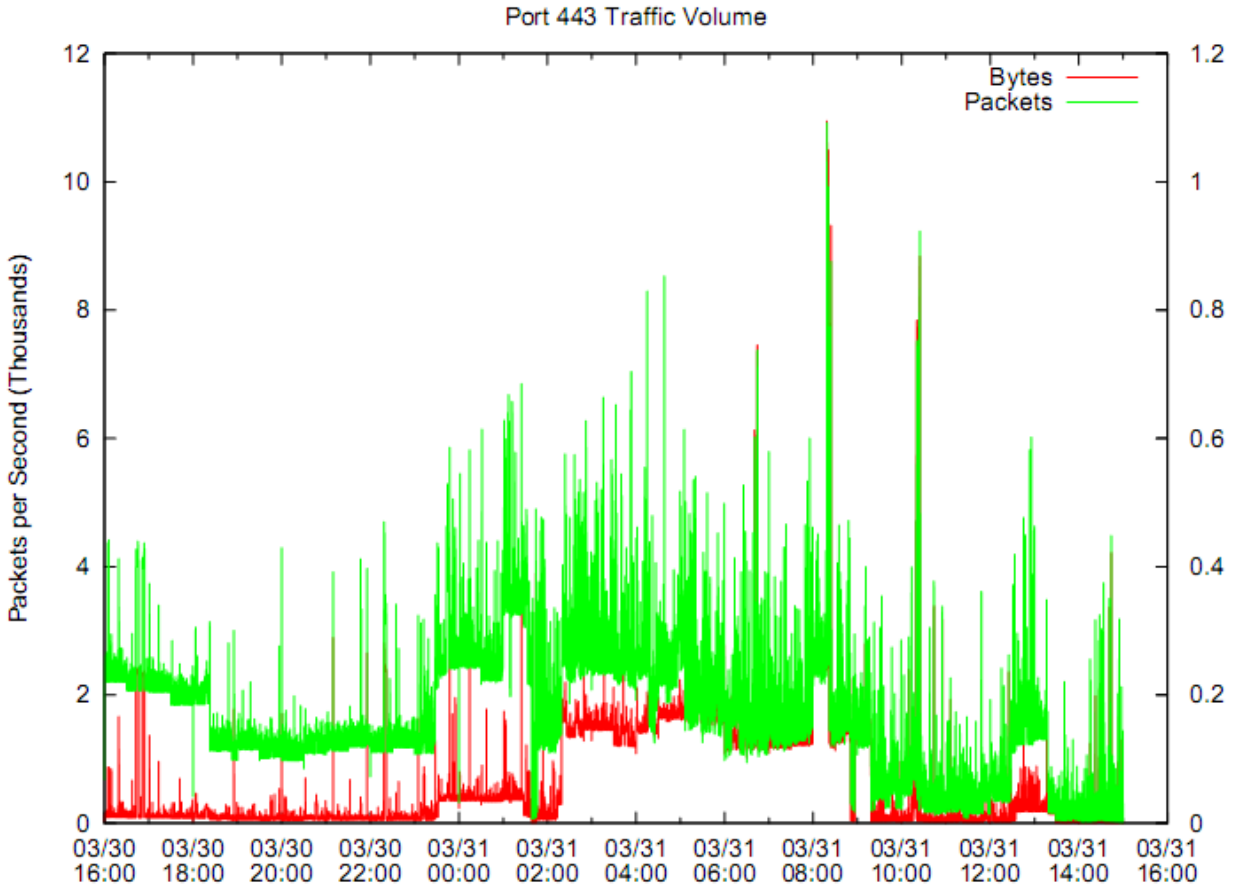
1 - Aggregate Traffic Flow Volume

Our next figure shows the traffic sent to port 80. As the figure reveals, there is an escalation in traffic that begins roughly around 02:00 UTC and lasts until 10:00 UTC or 11:00 UTC. This increased traffic may be a signature sign of a corporate network.



2 - Port 80 Traffic Flow Volume

The third figure shows the traffic sent via port 443. There was a lower volume of this traffic than that through port 80, but this figure shows the jump in traffic between 02:00 UTC and 09:00 UTC. It is clear that SSL traffic was transmitted in higher volume between these hours. This likely indicates that the users of this network perform a lot of Secure Web Traffic via SSL during the work day or a business service that is used by a number of corporations is accessed via port 443 throughout a set 8-hour work day. In either case, this provides strong evidence that the flow traffic through our network is related to that of a corporation.



3 - Port 443 Traffic Flow Volume

Suggestions for Improvement

Based on our assumptions about this network and the analyzed traffic, we offer several suggestions for improvement.

- Utilize DHCP in the network.
    - DHCP limits outsider views of the network and forces attackers to take more difficult actions to scan the network. It also provides an easier way to bring machines onto the network without custom IP configurations.
    - Servers should be placed in a DMZ, with clients masked through DHCP and Network Address Translation (NAT). This allows the servers to be open to the world (behind a firewall of course) while clients are given a second level of protection since they are not in the DMZ.
- Have a bandwidth monitoring program in place.
    - Since our network has high traffic loads, it should be monitored in case of failure. The organization should track bandwidth usage through flow sensors or other devices and install backup routers, switches, and firewalls to prevent downtime.
- Implement network virus scanning.
    - We noticed several hosts with port 113 (IRC) open, which is not used by many corporate services. It is most likely a virus beaconing to a command and control server, so we recommend a more robust virus scanning capability to prevent further infection.
- Properly configure firewalls
    - Firewalls should be configured to filter unwanted traffic but should also have ingress and egress filtering. Ingress filtering is necessary to prevent unwanted malware and bad traffic from entering the network. Since it is obvious that malware is on the network, better firewall rules and configuration is warranted. Egress filtering prevents improper traffic from leaving the network. This would prevent beaconing by the port 113 malware as well as increased protections against data ex-filtration.
- Change POP3 servers to Port 995 (POP + SSL)
    - Several POP servers are currently running unencrypted on port 110, so this should be configured with SSL to prevent possible man in the middle attacks.

Event Analysis

Distributed Denial of Service Attacks against Estonia

On the 27th of April 2007 the "most wired country in Europe" was at the receiving end of a cyber attack during Estonia's plight to move a Soviet World War II statue, the Bronze Soldier of Tallin, from its original site to the Defence Forces Cemetery of Tallinn, located within the country [4][5]. Although the attacks were not noticed until the morning of the 27th, when government officials tried to sign into their accounts, it is believed that the first attack was launched around 10 p.m. on the 26th. The first wave of the distributed denial of service (DDoS) attacks lasted about a week and focused on the Reform Party's website, the Estonian parliament website, as well as other political and government websites. After the first week of attacks all targeted websites were knocked offline. The second wave of DDoS attacks started on the 3rd of May and the list of targets expanded to include major news publications within the country; the attacks systematically knocked the news sites offline making it hard for them to tell the rest of the world what was happening in Estonia. On May 8th, the third and final wave of attacks targeted Estonia's banking system and ultimately forced the largest bank, Hansabank, to shut down all Internet-based operations. Many other organizations were targeted throughout the attacks; such targets included news, media, and universities websites [3][4].

Analysis of the Estonia DDoS Attack Data

ATLAS (Active Threat Level Analysis System) is able to provide Arbor Networks with DDoS attack data from around the world, allowing those with access to the data repository the ability to analyze and report findings pertaining to the events. Joe Nazario, a blogger for Arbor Networks, received access to the Estonia DDoS data, analyzed it, and reported his findings. The following is a breakdown of what Nazario was able to find from his two week analysis of the data using internal tools and reporting systems [2].

In the two weeks there were a total of 128 attacks; of those most were ICMP flood attacks, approximately 89.8%, while the others were TCP SYN and general traffic flood attacks, 3% and 7% respectively (see Figure 1). As for the targeted websites, there was an uneven distribution of attacks amongst the targeted websites: the bulk of the attacks were received by The State

Portal and other government websites as well as the Ministry of Finance, 28% and 27% respectively (see Figure 2) [2].

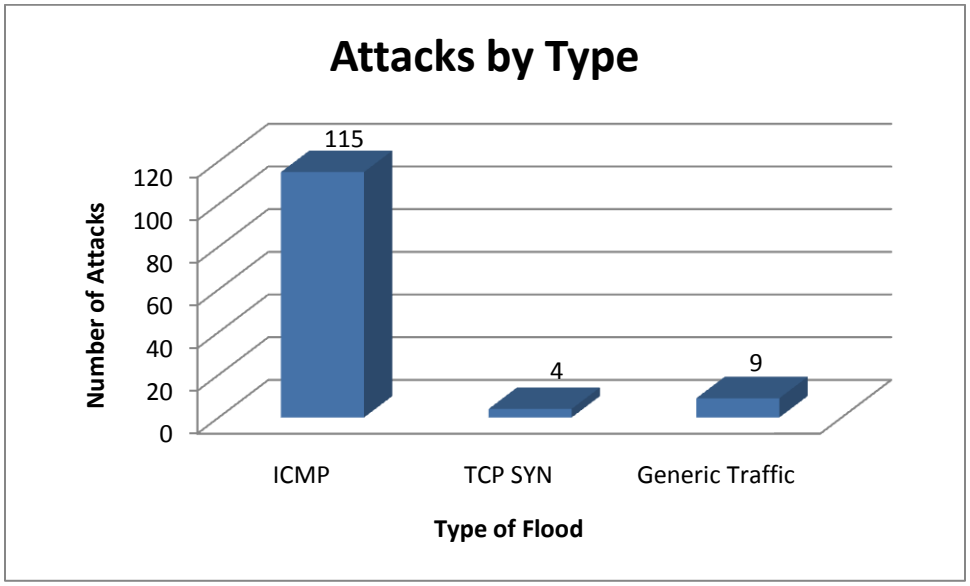


Figure 1

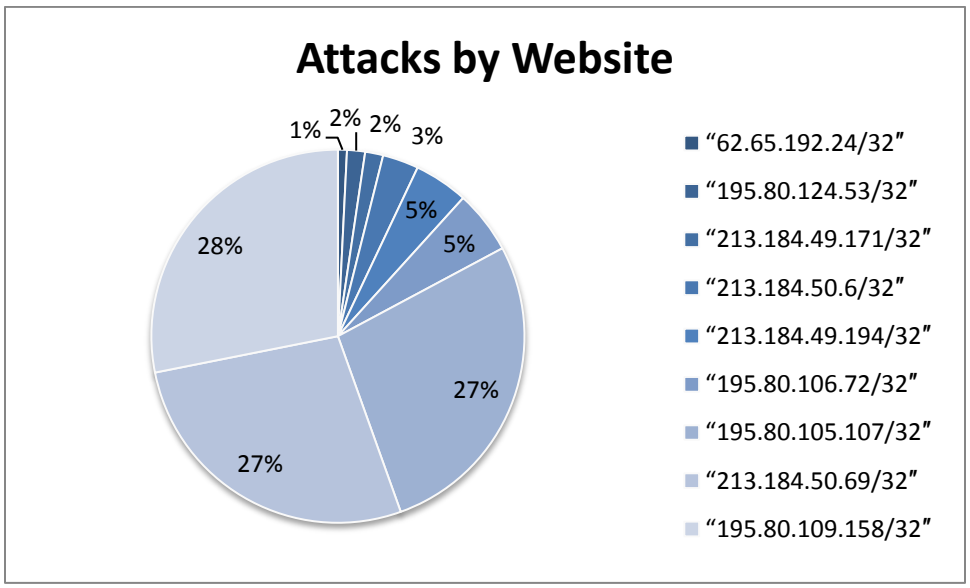


Figure 2

During the two weeks of data collected and analyzed the attacks occurred on the 3rd, 4th, 8th, 9th, and 11th of May. Although the number of attacks started out relatively high they peaked on the 9th, with a total of 58, then abruptly decreased to almost nothing (see Figure 3). The number of packets sent correlates closely with the number of attacks; the first day of attacks

averaged around 1,000 packets and the number rapidly increased until the 9th when it peaked at over four million packets of information per second [3]. The attacks also varied in length and bandwidth. Although the attacks ranged from under a minute to 10 hours or more approximately 62% of them were between one minute and one hour (see Figure 4). The bandwidth used in the attacks also different immensely with a range from less than 10 Mbps to 95 Mbps; however most of the attacks used between 10 and 30 Mbps, approximately 40.6% (see Figure 5) [2].

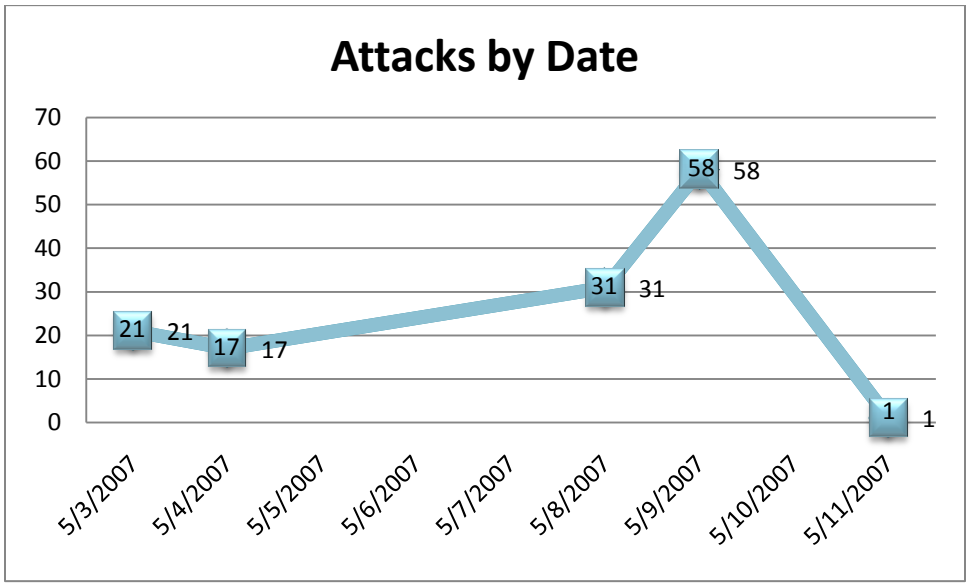


Figure 3

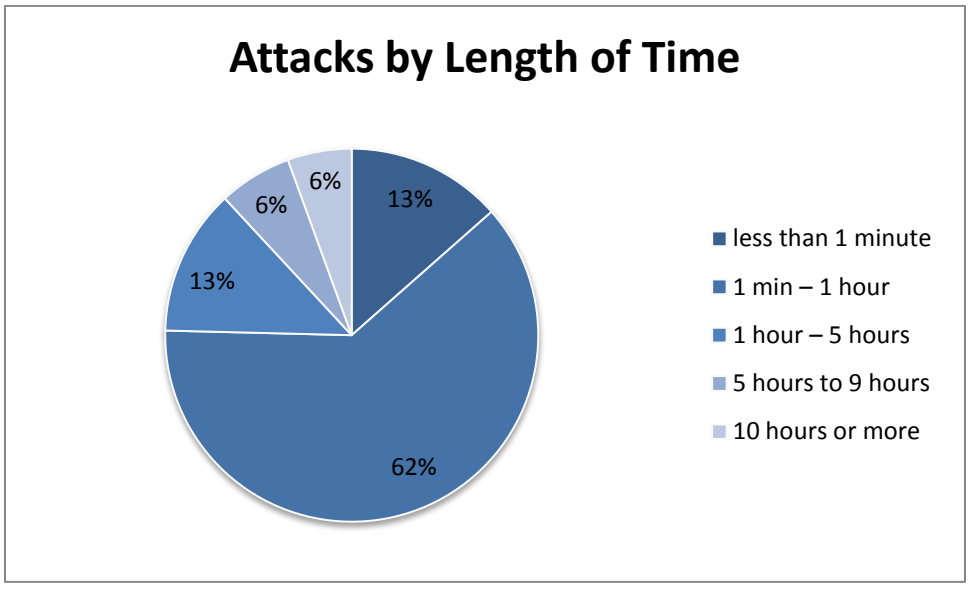


Figure 4

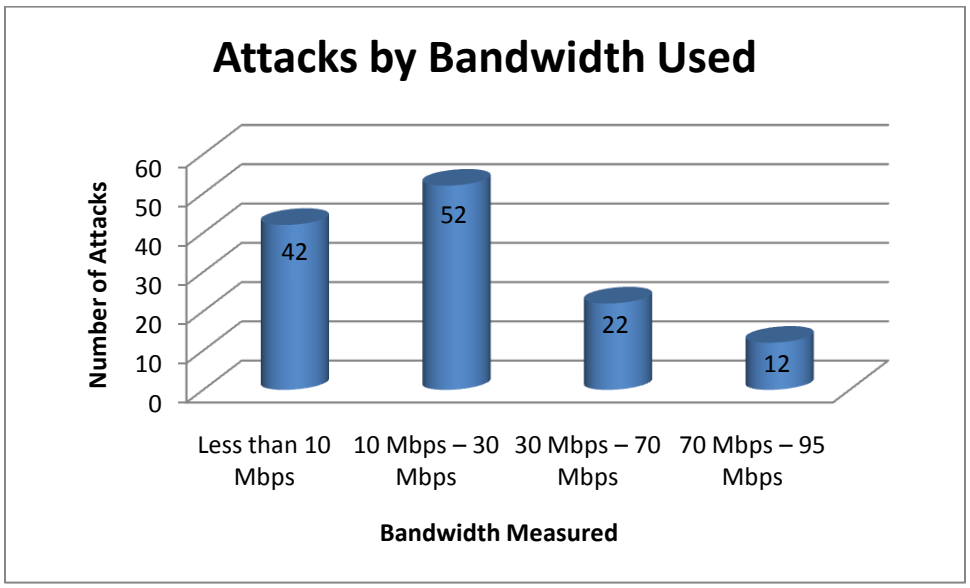


Figure 5

Execution of Estonia DDoS Attack

Estonia was known as the "most wired country in Europe" because their government and business infrastructure relied heavily on e-systems, technology, and the Internet; this, along with no defense protocols in place, made them very vulnerable to cyber attacks and exploits. Thus when the DDoS attacks started on the 26th, crippling all of Estonia, it was not much of a surprise; however discovering how the attacks were executed was. It was thought that the attacks were orchestrated and carried out by the Russians and it may very well have been; but the data showed that Estonia was receiving attacks from over fifty different countries (including the United States). The main difference in this particular DDoS attack was many computers and users involved were "volunteers" or script kiddies (trouble makers who execute code written by hackers). Basically, the hackers (or Russians) organized the attacks by posting on weblogs, web journals, Russian websites and chat rooms detailed lists of dates, times, and targets as well as instructions on how to carry out the cyber attacks that would bring down Estonia's infrastructure [3]. The instructions had the script kiddies executing a simple ping attack which sends a single response to a web server over and over slowing it down until it becomes so overwhelmed it shuts down. Along with those who volunteered many other computers were seized and used to help carry out the attacks. These hijacked computers were part of botnets

located around the world and their main mission was to flood specific Internet addresses with any type of data that would clog up the networks. The final, more sophisticated and intelligent aspect of the attacks were hackers that gained access to and defaced many individual websites; they used this tactic to make their message and propaganda public [1].

DDoS Attack Effects on Our Network

A DDoS of this nature would shutdown our network very effectively. Aggregate traffic on our network peaks at roughly 90Mbps, with averages of 30Mbps. A DDoS attack, like the one in Estonia, had traffic peaks as high as 95Mbps. This is well over our network peaks, and sustained traffic at that speed would cause our network to be extremely slow or inaccessible.

Since our network is most likely an e-commerce site, the web servers would be the most likely target. Web servers are often targets because the attacker has some sort of issue with the owner of the site and the website is the most public part of the company. However, mail and DNS may also be flooded as side effects of the attack. Depending on the type of attack and its scope, the attacker may only knock out a single server, or the entire network. Load balancing is very important in this case.

With these high amounts of traffic, the routers may actually fail before the servers do. Having multiple routers is important to prevent this failure from occurring. Ideally a "graceful failure" is the best failure condition since it allows the site to slow down without a sudden outage and to come back up as traffic decreases again.

DDoS Attack Effect on Netflows

A DDoS attack will manifest itself in flows through several indicators:

- Large increases in flows to a specific port (most likely 80 or 443)
- High traffic flows from certain IP addresses.
- Overall flows may *decrease* due to reduced network performance.
- ICMP flows may increase since many DDoS attacks are simply ping requests.

Conclusion

It is important for organizations to understand how their network works. This includes having a comprehensive understanding of the traffic that is transmitted through the network. It is important that legitimate traffic is profiled so that anomalies can be easier to detect and to eliminate through security. Since a network is required to operate and function differently through a day or week, depending on users, services hosted, etc., network analysts should have a clear understanding of the timeline patterns on the network.

In understanding a comprehensive view on the flow of traffic and the type of traffic transmitted via their network, organizations would be better equipped with managing bandwidth, detecting (potentially malicious) anomalous network traffic and aid in planning against incidents and disasters and would help in formulating incident response guides and expectations for the organization. Network flow analysis is a powerful tool for providing situational awareness to an organization. Implementing this analysis and monitoring provides another layer in defense of preventing and quickly detecting incidents.

Bibliography

[1] Davis, Joshua. "Hackers Take Down the Most Wired Country in Europe." *Wired News*. 21 Aug. 2007. Web. 22 Nov. 2009. <http://www.wired.com/politics/security/magazine/15-09/ff_estonia?currentPage=1>.

[2] Nazario, Jose. "Estonian DDoS Attacks; A summary to date." *Security to the Core | Arbor Networks Security*. 17 May 2007. Web. 22 Nov. 2009. <http://asert.arbornetworks.com/2007/05/estonian-ddos-attacks-a-summary-to-date/>.

[3] Richards, Jason. "Denial-of-Service: The Estonian Cyberwar and Its Implications for U.S. National Security." *International Affairs Review -- George Washington's Elliott School of International Relations*. Web. 22 Nov. 2009. <http://www.iar-gwu.org/node/65>.

[4] Traynor, Ian. "Russia accused of unleashing cyberwar to disable Estonia." *Guardian.co.uk*. 17 May 2007. Web. 22 Nov. 2009. <http://www.guardian.co.uk/world/2007/may/17/topstories3.russia>.

[5] Zarakhovich, Yuri. "Estonians Under Siege in Moscow." TIME. 02 May 2007. Web. 22 Nov. 2009. <http://www.time.com/time/world/article/0,8599,1616943,00.html>.