Denial of Service Prevention and Detection

Edward Spetka, Dan Stalloch, Kyle Yager
State University of New York Institute of Technology
{spetka@cs.sunyit.edu, d.stalloch@yahoo.com, kyager@twcny.rr.com

Professor Sam Sengupta . 12, 2012
Abstract

Attack monitoring and protection is critical for an internet service provider or any host of a web server to be able to provide to its clients. It is the goal of this paper to cover some of the current and most devastating attack methods and discusses techniques to prevent against attacks including some good security policies. You will learn our approach to protecting a system from a denial of service attack. We will propose a three-step approach. The first step will be to use a special kind of acknowledgement packet to ensure the data is from a viable source. Second, we incorporate the use of a token to ensure the data is authentic and from a healthy un-compromised host. In the third step in our approach we will discuss enabling a traffic buffering and pre-analysis system to bring a second added level of security to the denial of service defense system.

1 Introduction

Denial of Service (DoS) attacks generally work either by using weaknesses in the target machines software or simply flooding the target with messages such that it wastes its resources so that it cannot respond to requests thereby rendering them unusable. Some of these attacks target the bandwidth of a network; some go for the resources of a computer on the network. The basic idea is still the same, to make a machine or a network resource unavailable to its intended users. DoS attacks make a service incapable of responding to requests in a timely manner. This can be by temporarily or indefinitely interrupting their service or suspending the service of a host. The main targets of the attacks are usually web servers, root name servers and other high profile servers. There are many different types of DoS attack. The various attacks use different techniques to cause problems on a server. Some attacks tend to focus more on attempting to tie up the bandwidth of a network. Others try to waste the resources locally on the server itself. With a vast network of computers available many look to harness many computers on the network in a distributed denial of service attack in order to do their damage. ICMP flooding is a common type of attack used to attack systems. This includes the smurf attack, ping flooding and ping of death attacks. Smurf attacks utilize the ICMP packets by broadcasting a message to hosts on the network. They would send the packets to these hosts with the return address maliciously changed to be the address of the host being targeted. Then when the broadcasted hosts received the messages they would answer back to the return address given which are all set to the same host, the target of the attack. This way that host would be overloaded with messages, clogging the network and using the bandwidth and stopping them from achieving normal functionality. Another powerful attack used often in the real world is the TCP SYN attack. This is an attack that sends numerous connection requests to a computer to waste its memory. When a user sends TCP SYN packets with fake senders to a host then the host attempts to open a connection with the host it has just received the message from. If that host does not exist this can cause many problems and be an effective DoS attack method. The problem with the TCP SYN attack is once it has been launched on the server these half open connections waiting for a host to respond that does not exist are left. These can
effectively block the legitimate users from accessing the server. Distributed denial of service attacks (DDoS) are another of the most commonly used attack techniques. Usually in a DDoS attack multiple systems flood the bandwidth or resources of a targeted system. They work together in an attempt to harm the target system. The benefits of using multiple machines to attack are the more machines you have in a coordinated attack the more traffic can be generated on the network. Some of the most powerful DDoS attacks are those that take control of computers whose security defenses have been breached. These groups of computers are then pooled together and usually controlled by some outside source. These types of attacks are becoming more and more popular as poorly secured machines give the perpetrators of the attacks more malicious machines to use in their attacks. In addition, the tools for carrying out such an attack grow in complexity in automation, making it possible for more people to be able to attempt them. One preventative measure that can be taken in an attempt to try and stop this type of attack is to use a method called Remote Monitoring Data or RMON to detect the attacks. RMON works with Simple Network Management Protocol. Unlike other attacks on a network, a DoS attack does not contain anything that the usual Intrusion Detection System would flag as a threat. Thus it is vital to monitor the network for less obvious tells such as unusual bandwidth usage. Also, another variant of RMON, RMON2 was used, and can sometimes provide better protection. The downside is that RMON2 is incompatible with most modern networks. Other methods require restructuring the network architecture in order for them to be used. These other methods are also unreliable and showed a tendency to produce false positives. These were mostly cause by the bandwidth thresholds being set too low to account for the usual changes in Internet traffic. Researchers took normal traffic captured with RMON1 over a 4 month period, then injected traffic to simulate a DoS attack, rather than capture data during a genuine attack. The first conclusion they reached from the captured data was that, as mentioned earlier, variability in the bandwidth being used is large. This was especially true during normal business hours, and thus the time of day had to be taken into account. They then tested both a standard traffic modeling technique, and multilayer perceptron, to try and detect the (simulated) attacks, noting that both had good results. However, the perceptron had both a higher detection rate, and a lower rate of false positives, which are time-consuming to deal with from the users perspective. They also tried varying the size of the DoS attack packets, and noted that the perceptron had both a lower success rate on the larger packets, and also a noticeably high rate of false positives (3% vs 0.6% on the smaller packet size). They explain that this is because large packet traffic is the norm on their data set. The perceptron also improves in performance, in terms of both detection rate, and false positive rate, relative to the length of the training period. The last test showed that most of the misses and false positives occur during working hours. The researchers in these experiments were able to show that although network monitoring can be a good indicator of what is going on in the network, it is hard to detect and prevent denial of service attacks. Since the packets often used in attacks are not abnormal they are able to bypass intrusion detection systems. This is why a more specific system is needed to prevent and detect denial of service attacks.
2 Background

2.1 Considerations

ARP (Address Resolution Protocol) Cache Poisoning although a very relevant topic for DoS attacks is something we choose not to focus on as it is easily circumvented in unix systems where one can just turn off promiscuous mode in the kernel. However even when not running on a unix system there are programs like arpwatch to monitor any alterations to the cache that would allow an attacker to act as any other computer in the network. DNS (Domain Name Service) Cache Poisoning is also a topic we choose to ignore for similar reasons of just being able to update to the most recent version for the operating system to stop the problem. It might also be helpful to notice that BIND (Berkley Internet Name Daemons) are the ones which have had this problem occur the most. SNMP attacks in this very same strain of issues are equally easy to defeat by updating to the SNMP 3 version. One could also turn off the feature for those devices that dont need to be using it but this may take a long time as lots of devices have it on by default. [1] We also may focus more on the extend detection of DoS attacks such as on a network instead of a singular computer because there are scanners that could be free or bought such as The CyberCopZombie Scan. These scans can be used to find out if there is any potential for your computer to become a DDoS bot. [3] WebSOS which is an adapted form of SOS for allowing clients to access a site or sites that is/are under attack by a DoS or DDoS. What it does is allow for clients to gain a certificate through the use of a couple of One Time Coins which act like one time passwords. Once the certificate is obtained then a Graphical Turning Test or CAPTCHA as it is best known can be issued to verify that the user is truly a human. However technology is starting to reach a point where CAPTCHAAs cant stop all forms of automated computer systems from gaining entry. It is partly due to this that the SOS involved makes several hop points but only one secret servlet and one beacon randomly chosen upon set up. These two are also set up in such a manner so that the source is close to the beacon and the beacon close to the servlet to insure quick access. Another thing that can defeat this set up is the use of very large groups of bot networks but they must be on the order of 70 thousand to 150 thousand or more. This can be considered a minor concern due to the majority of bot networks not nearing this size. One thing that could be slightly annoying is the fact that when the system is used it adds a little bit of time to access the web site(s). The nice part though is that it can be turned off when the site(s) are not under an attack. Due to the possible defeats to this system we only take the idea of the One Time Coins in the form of our tokens with a counter or Time to Live. [5] Adaptive Intrusion Detection and Prevention combines the chi-square goodness fit test with check control and multiple iterations in order to increase positive DoS attack detections and decrease false positives. When compared to just control chart alone it does this job very well since the testing showed that it had an 86

2.2 Reasons

Certain groups of people have recently used DoS and DDoS attacks to show a lack in security on major sites but there are good ways
to do this versus bad ways of doing it. Sadly a majority of them have used the bad ways which is to post information obtained during the attack(s) publicly. Even more recently some of these groups have also used these attacks to bring down sites as a form of protests leading to the coining of the term hacktivism which one then could begin to call the people doing it hacktivits. [2]

2.3 Special Concerns

Wireless data networks contain five major issues when it comes to DoS attacks which are Signaling, Peer-to-Peer applications, Excessive port scanning, Battery drain and Malfunctioning air card. In Signaling the attacker simply reopens a connection sending out a small amount of data in an attempt to bring down the radio network controller. Peer-to-peer applications can be used so much that it just brings down a network so this type could be intentional or completely incidental. Excessive port scanning is where a few specific ports were targeted by worms however this has been taken care of through a tool that can inspect packets on a wireless network. Battery drain is also a low volume attack that stops a mobile device from using sleep mode to prevent less battery power loss. Malfunctioning air card was an incident of a singular mobile devices air card not acting properly but took quite some time to isolate and find. We can safely ignore all but Signaling and Battery drain because one could simply limit the number of connections or bandwidth for the Peer-to-peer applications, Excessive port scanning has been dealt with already and Malfunctioning air cards can be stopped by a little extra quality insurance. The last one is assuming it wasnt done on purpose but if it was done intentionally then isolating the source and cutting off the connection seems to be the best option. [4] RREQ flooding that cant be detected normally in MANETs are ones that use IP addresses in the domain but not present currently in the network and issuing more than one with increased TTLs before the ring traversal time. [6] Voice over Internet Protocol is very easily defeated through flooding which can be very cheap for the attacker through the use of thousands of invitations. [7]

3 Related Work

We found multiple other suggested methods of defending against DoS attacks. It has been observed that the traffic analysis method commonly triggers on legitimate traffic, preventing actual users from accessing the services, essentially, causing the very problem they’re trying to prevent. Additionally, because what separates legitimate and non-legitimate traffic in the eyes of the system is rarely disclosed, it can have a negative effect on the development of new protocols, that would inevitable have a different signature than the traffic of more established protocols. So, one group suggested new approach, basically giving permission for someone to send, this permission being reflected in their resulting packets. This way, non-legitimate traffic can be dropped. Also examined were other methods to defend against DoS attacks, each having notable weaknesses. For example, the Source Filtering method, is only effective if it is fully implemented by all machines in the network. Additionally, tracing back to the source of the packets only helps to identify those involved in the attack, and does nothing to stop the
attack itself. On top of that, the attacks are often coming from legitimate users who have infected machines which are part of a botnet. Overlay Filtering is another defense mechanism that relies on moving all traffic through an intermediate note that can mark legitimate packets, so the destination can drop all others, but this system can be bypassed by the attacker finding out what that mark on the legitimate packets is, since the same mark is used for all legitimate traffic. A new method was devised that uses a variation of Overlay filtering, by using tokens to indicate that the source has obtained permission to transmit to the destination, however, each token is used for a 1-time agreement to accept transmitted packets, so a compromised token doesn't give an attacker free reign to flood a network. Similar approaches tried to keep out unwanted packets with existing routers and such, whereas this method attempts to redesign the architecture itself to be better at handling DoS attacks. In this system, each destination creates its own tokens that count as permission for someone to send to them, which the source would ask for when trying to communicate with them. Routers along the route to the destination can simply drop packets that do not have valid tokens. The data for the tokens is encrypted enough that, by the time a brute-force attack could crack them, they would no longer be valid. In the event of an attacker somehow finding out what the token is, such as snooping the links along the path the packets take, it will only be able to disrupt that particular connection, instead of being able to disrupt all connections to that destination.

4 Prerequisite Knowledge

Black hole filtering was devised to drop unwanted packets before they entered a network. They are placed at the edge of an ISP network, and all traffic believed to be malicious can be sent there to be dropped. The source or destination of the addresses is used to determine which packets are malicious. Similar to other DoS prevention strategies, this means it is vulnerable to source address spoofing. The black hole system is generally triggered by internal Border Gateway Protocol updates sent out from the network operations center. Once the situation is resolved, another iBGP update allows the traffic to resume along the usual route. In order for this to work, the triggering device must be an iBGP peer with all of the edge routers, or the route reflectors in each cluster of the network, if those are being used. Additionally, the routers are the edge of the system require a static route, which will correspond to an unused IP address space, which is where the packets will be dropped. The trigger also gets a static route, which it will send out via a BGP update to its peers, and next hop for the address that is under attack can be set to the address previously mentioned. The edge routers will then receive this update, look at the static routing entry for that address, and set the route to that address as null0. All traffic bound for this destination will then be dropped at these edge routers. Once the attack has passed, the static route is removed from the trigger. As before, the resulting BGP update will inform the network. It is important to point out the adding and removing of the static route must be done by a network administrator. Any system which requires human intervention relies on the reaction time and the judgment of the human
Secure Overlay Service (SOS) is an overlay architecture that stops attacks by having it stop popular communication jamming points. It uses SOAPs (Secure Overlay Access Points) to tunnel the traffic from the source to the beacons which in turn go through the secret servlets and finally to the destination. These nodes are selected in secret and have a cryptographic or Graphic Turning Test to ensure that a human is using them rather than a bot. It issues to the legitimate users of the system a one time certificate so that even if a user becomes infected they may be stopped. It also uses the principle of a distributed firewall to prevent spoofing. [5]

Session Initiation Protocol (SIP) is a protocol that has come to be seen as the best choice for Voice over Internet Protocol (VoIP). It is made up of Registrars which strictly associate a users SIP address to a specific IP address, Proxies that send along retrieval and uploading of information to SIP networks and User Agents (UA) that choose when retrieval and uploading of information happens. It has a text based format similar to HTTP with sender in the header along with any routing information and destination on the first line. The most susceptible part of the SIP to a DoS attack is the registrars also known sometimes as redirectors and the proxies. Both of these can be referred to as proxies under certain circumstances and when given a properly formatted text with an irresolvable URI or domain name can become unresponsive for quite some time. It would also be a good idea to know about synchronous versus asynchronous SIP proxies. [7]

5 Application

The challenge is to build a denial of service prevention system that is both scalable and secure. One of the most important challenges is to fight of malicious hosts while still allowing the requests from legitimate traffic to not be dropped. We propose a hybrid approach to solve this problem. The proposed method uses special packets before a request is sent which basically functions as an authentication system. When any host would like to request information from a server it must first send a packet requesting authentication to the server. Then the user would receive a token if it was thought to be legitimate. We suggest this with overlay filtering, the overlay check performs the check on the legitimate packet by checking the token for authenticity. The problem with overlay is the user finding out the legitimacy marking on the packet because they can mark their own packets as legitimate and then bypass this system. By incorporating this method the server now has 2 layers of protection against malicious hosts trying to DoS attack the system. Besides checking to make sure the packet has a legitimate token, we could also have a check that the users IP address is not already listed in our servers cache. We use a small set domain name cache list size in order to achieve quick access to the DNS listing. The reason this check is essential is it can save a server from wasting resources for trying to resolve a domain name from the DNS listing.
This can either be due to unresolvable domain names or names which are last to reach in the listing therefore causing the server to wait. By doing this check those unresolved domain names cannot cause a DoS. When a user first makes a request to the server for a token, the server responds with a request authorization token. This is distributed back upon request based on the previously mentioned cache check, time, number of requests from that IP in given time. Once authenticated the host can send the server the original request packet along with the token. This token has time to live built into it for further verification of authenticity of the host. It also has a counter to keep track of the number of access attempts. This is different from the standard protocol that consists of sending one request to the server and having a response sent back. This method of interaction between server and host disregards any information about the user leaving them susceptible to DoS attacks. Once the authentication token has been granted by the server then the host will have permission to send its request. Before the request can be served there is another layer of security through which it must pass. A filter system which uses the token to validate the server requests authenticity.

6 Critique

Although a hybrid method may be able to prevent a larger number denial of service attacks it still has weak points. One of the biggest problems with the idea of using a token in the authentication system is that if the token is compromised, the system is then vulnerable to a variety of DoS attacks. Another problem with this system is the number of additional extra handshakes between the hosts and server which add unnecessary congestion to the network. In our study we looked at WebSOS. This system uses captcha or cryptography as part of its authentication system which makes this a much more secure system to protect against DoS attacks.

7 Future Work

8 Conclusion

Some companies try to deal with DoS by overprovisioning, spending 75% more to increase their bandwidth a corresponding amount, to mitigate a DoS attack. However, recent DDoS attacks had rates of over a million packets per second, with some reaching almost five times that. Such an attack would occupy roughly 40 Gbps of bandwidth. This is not something that overprovisioning can hope to adequately counter. Additionally, the losses companies could suffer for unexpected downtime could be severe, in a 2009 survey, a number of companies that responded estimated hundred-thousand dollar losses in revenue alone for even an hour of such downtime. Of these companies, the majority of them relied on overprovisioning as a defense method, but even the most extensive among them only reached 7.5 Gbps bandwidth vs the 40 Gbps attack bandwidth mentioned previously. Thus, these companies are spending large amounts of money for a measure that will not ultimately protect them from a determined DDoS attacker.

Our hybrid approach could work well and improve the ability to prevent and detect de-
nial of service on a server. More work is necessary in order to prove the increased efficiency of the hybrid technique as opposed to only using any of the other techniques available today. Perhaps if its performance could be proven then it could in the future become part of what people believe to be Quality of Service. Quality of Service (QoS) is the ability to provide different priority to different applications, users, or data flows, or to guarantee a certain level of performance to a data flow. Sometimes QoS guarantees a required bit rate, delay, jitter or packet dropping probability and/or bit error rate may be guaranteed.
References


http://delivery.acm.org/10.1145/1590000/1582581/p926-nadeem.pdf?ip=149.15.108.152&acc=ACTIVE
SERVICECFID=223348477CFTOKEN=51692084aCM=135499057c,f11031018fcnd27c=182c6405848e7>.