sentor

White Paper

Demystifying DDoS (Distributed Denial of Service) attacks

Commissioned by:

GLOBIX®

Martin Zetterlund, martin.zetterlund@sentor.se

Mathias Elväng, mathias.elvang@sentor.se

Partners and senior consultants

Sentor MSS AB

Table of Contents

1. Background

1.1. The Problem

During the last year, an increasing number of companies with critical on-line business operations and in particular on-line betting companies and banks have been victims of DDoS attacks and extortion.

1.2. Trends

Historically DDoS attacks have been aimed at high profile sites (such as Yahoo, Amazon and Microsoft) and the main purpose for the attacks have been fame, publicity or spite on behalf of the attacker.

What we now see is a trend of the DDoS attacks being used as a method of extortion, often involving some kind of organized crime entity. A DDoS attack uses multiple nodes or zombies to generate traffic against a site in order to deny service for legitimate users and clients. On one occasion Sentor uncovered a network involving more than 25,000 nodes (zombies) used for performing the attacks. This indicates increased organization behind the attacks.

The widespread uptake of Broadband /ADSL and always on Internet for home user, in combination with a low degree of protection of connected PC's, is a key factor in the spread of Zombies used in DDoS.

1.3. Objective

The aim of this paper is to explain the fast-growing problem of DDoS attacks and the mechanics and drivers behind the attacks. We will give some technical advice on how to mitigate the exposure to the attacks and how to limit the consequences. The paper is based on vast experienced gained by the authors and co-workers during 8 years of assisting our clients in responding to DoS and DDoS attacks and helping them design and implement protection for these attacks.

Some of the figures in this paper are estimates. Descriptions of techniques used to perform attacks and to protect from them are mostly conceptual rather than detailed, as attacks and attack techniques are constantly evolving.

2. History

DDoS, and its forerunner DoS, attacks are by no means a new invention and have been used for a long time in the black hat hacking Internet community. Its main purpose is, and always has been, to disrupt Internet services for organizations that have for some reason become targets of the perpetrators.

2.1. Evolution and IRC

The evolutions of the attacks have been closely linked to the IRC protocol. IRC (Internet Relay Chat) still plays a part in many of the attacks we see today. For this reason, and the fact that it describes the basic mechanisms of DoS and DDoS attacks, we will go into more detail in the subject.

IRC is a protocol for real time group communication over the Internet. Using IRC the user can communicate with others by joining named channels (sometimes referred to as "chat rooms").

On IRC, as on all more or less anonymous public online forums, users often tend to get annoyed with each other, and because of the way the IRC protocol works it is common for users to disconnect others from the IRC network when they have run out of words.

2.2. Flooding and Attack Distribution

The ways of disconnecting clients are numerous; the most interesting in our case is referred to as flooding. To "flood" another user off the IRC is achieved by sending the user more packets than the Internet connection could handle, making the server think the client is dead, and therefore disconnecting the client.

Back in those days (around 1994) Internet capacity was limited and most people used modems to connect to the Internet, so attacks didn't have to be distributed among several attacking hosts. It was sufficient to have control over a machine with a high speed Internet connection (e.g. in a university or a big company) to flood most users offline.

As the users got better connections the need for distributing the attacks grew. At first the attackers started to use more than one computer on high-speed connections to flood the victims (by simply logging into the computers and execute the commands one at the time), but as the connection speeds grew, so did the need to coordinate the attacks.

Around 1998 the first client-server distributed attack tools started to spread. Some of the more popular were stacheldraht, tfn and trinoo (see appendix A for further description). These tools were UNIX based, and the clients were spread by manually hacking the servers and installing the clients.

The tools were widely in use, but there was little publicity about it, since the victims were mostly IRC-server operators and unfortunate IRC users. Many smaller ISP's suffered from the attacks as well, even though they were not the primary targets, as the packet floods caused major outages in their networks. Even though the tools were not perfect they gave the attackers a much more powerful way to cause harm, and as Internet use grew so did the size of the DDoS networks.

2.3. High profile incidents

Early February 2000 DDoS attacks at first time made the news in a big way when a group of hackers attacked high profile sites such as Yahoo and Amazon (A Canadian hacker called Mafia-boy was later sentenced for these attacks).

The reason for the attacks seems to have been mainly to earn kudos within the hacker community. In May 2001, the infamous Code red worm hit the whitehouse.gov web-site

(or actually it missed, since the IP address of the target site was hard coded into the worm. This made it trivial to move the site and null route the hard coded IP-address).

Well worth noting is that there was no interaction in this attack. Once the worm was released, the attacker had no chance to abort the attack or change the target. The same principle was used for the attack on Windows update performed later by the Slammer worm (which also missed its target since the attacker aimed at the wrong site).

There were also rumours about extortion attempts against online bookmakers and financial institutions during 2001, however there is no evidence that it was a real problem since the tendency was not to admit to attacks.

The most interesting DDoS attack during 2002 was an attempt to knock the DNS root servers off the net, the attack was not completely successful, since some of the servers could handle the attack, most Internet users did not experience any problems. However, the attack was successful enough to raise serious questions about the core Internet infrastructures ability to handle DDoS attacks.

2.4. Current trends

In June 2003, the use of DDoS tools took a turn for the worse when spammers started to DDoS anti-spam sites such as Spamhaus.org. The notable thing about this attack is that the spammers were using DDoS for "business" purposes. This was one of the first evidence of organized crime, mafia-style extortion.

In November 2003 reports started coming in about online bookmakers being extorted with threats of DDoS attacks, and also about attacks executed. During winter/spring 2004 almost the whole of online gambling business was affected and the problem started to spread to other online business (online banks, e-merchants etc).

The widespread uptake of Broadband /ADSL and always on Internet for home user, in combination with a low degree of protection of connected PC's, is a key factor in the spread of Zombies used in DDoS.

According to the Symantec Internet Security Threat Report Volume VI September 2004, some 30,000 computers are recruited to the Zombie networks on a daily basis, compared to only 2,000 a half year ago.

There are a lot of rumours about Russian mafia connections to the DDoS extortion business. Whether this is true or not is unknown, as there is of yet no conclusive evidence. However, the allegations do not seem too far-fetched according to our experience, and what is known is that some of the attackers are from Eastern Europe (at least the group that the good guys over at National High Tech Crime Unit in the UK caught). It is important to point out that the problem is not confined to Russia or Eastern Europe.

3. Attack Tools and Types

3.1. Common attack types

3.1.1. Synflood

A synflood is a tcp based attack. In A normal TCP connection or "three way handshake", the client sends a packet (SYN) to the server, the server responds with a packet (SYN-ACK), and the client sends the third completing packet (ACK). This handshake is necessary in order to exchange the sequence numbers used by TCP to ensure session integrity.

During a synflood attack, the client only sends the first packet (SYN) ignoring any answers from the server. This results in the table of half open connections on the server being filled up, and the server will stop responding. Firewalls also keep a list over half-open connections and may be susceptible to the same form of attack.

3.1.2. UDP flood

A UDP flood aims to simply consume all the bandwidth to the attacked site, it is common to use heavily fragmented UDP packets in order to try and overload routers and firewalls. ICMP flood similar to the UDP flood, the ICMP flood aims to consume all bandwidth.

3.1.3. Spoofed attacks

The early DDoS attacks were most of the time spoofed. Spoofed means that the senders address is altered in the packets to some random address, making the attack much harder to trace. Today most attacks are not spoofed mainly for two reasons, first the clients are mostly Windows based, making the creation of arbitrary packets harder and second, the fact that many ISP's have started to filter out spoofed traffic (thus making a spoofed attack meaningless).

3.2. Tools

3.2.1. First generation

Stand alone tools, such as sync are used on single machines. The tools have to be executed individually from command line.

3.2.2. Second generation

Client - server based tools such as stacheldraht, trinoo and tfn. The client - server model enables the attacker to command large amounts of clients from a centralized server. The addition of new clients is done through hacking of machines and installing the client software.

3.2.3. Third generation

The third generation of DDoS networks has the capability to self spread and also has logistic functions to enable easy parallel use of the networks. As an example we will look at a specific Trojan found early 2004.

The Trojan was found attacking an online bookmaker among a large number of other clients during a DDoS attack. During the forensics after the attack the Trojan was disassembled and thoroughly examined in a controlled lab environment. Below is a list of some of the more interesting capabilities of the Trojan.

Control

Upon successful deployment the Trojan announces its presence on a channel on a nonpublic IRC server. Another client in the channel continually sends commands to the

channel to which the Trojan responds. The commands in this channel are for sorting the clients, for example clients with bad connections are immediately discarded. Once the Trojan has passed the first sorting round it is instructed to join one of several other channels, each channel containing about 1000 other Trojan clients. This procedure enables the attacker to split the use of the DDoS network in order to be able to simultaneously attack several sites.

Installation

The Trojan has capabilities to self-spread by utilizing known vulnerabilities in the Windows operating system. Upon a command from the controller of the IRC channel the Trojan is listening to, the Trojan may start scanning a specified subnet and infecting vulnerable machines.

DOS capabilities

The Trojan supports SYN and UDP flooding when upon command participating in a DDoS attack.

Other functionality

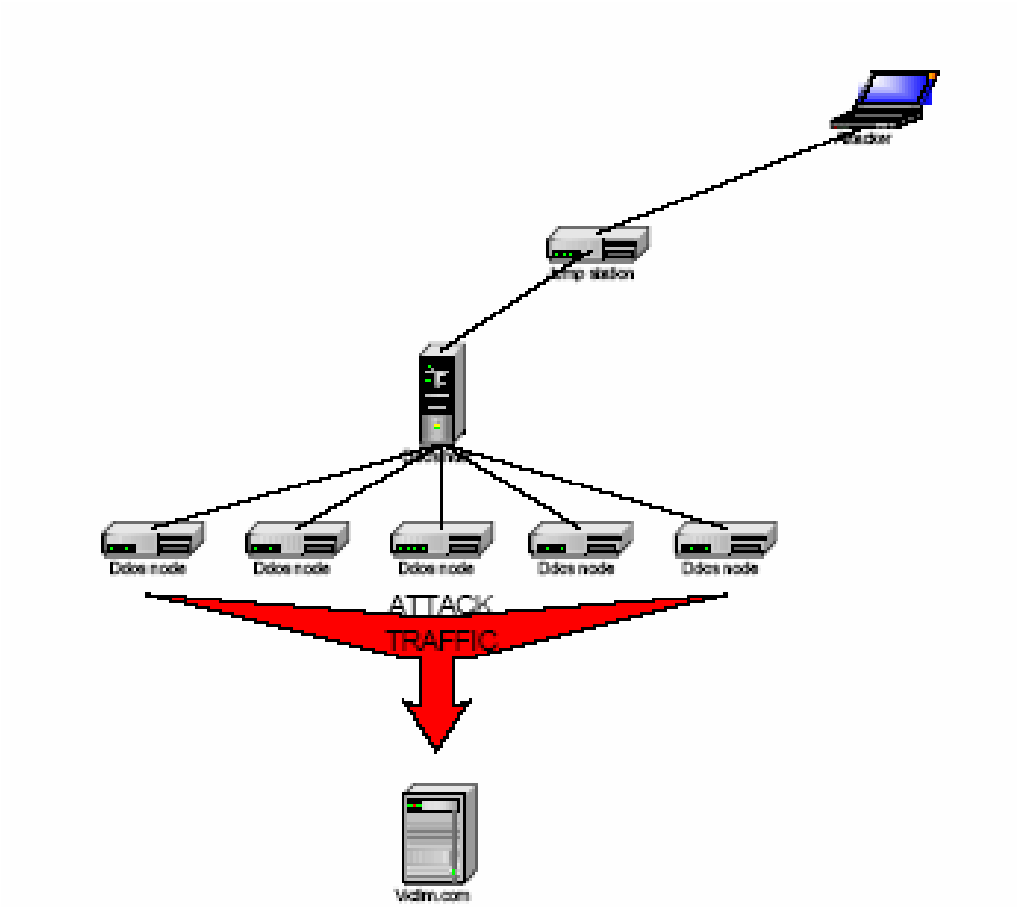
The Trojan supports several types of file/data transfer (FTP, TFTP, HTTP, IRC) as well as interactive shell access. In addition, it has various helper functions for performing data mining (keyboard logging, password extraction, product serial number extraction etc) and scanning (portscanner, user and service enumerators etc).

4. Tracing the Attack

Tracing of attacks is very important, if only to get an insight into the method of the attacks and defending against future attacks.

Tracing a DDoS attack is somewhat complicated due to the distributed nature of the attacking network. Providing the attacker does not make any apparent errors, like using his private email account to send the ransom note, the attack will have to be traced step by step.

Below is a figure of a typical attack scenario.



4.1. First Step

The first step is to find one of the DDoS nodes. This can often be done using an Intrusion Detection System or a network sniffer (providing the traffic is not spoofed). Once a node is identified the problems begin, as the owner of the machine must give up access in order to extract the DDoS software.

Providing it is possible to get hold of a node it is not that hard to find the hub-server. One way is to disassemble the binary and another is to simply infect a computer in a controlled environment and record the traffic between the node and the hub.

4.2. Next Step

In order to get any further, it is necessary to get control over the hub or at least the traffic to the hub, this can be tricky since the hub is probably placed in some country far away and people may be less than helpful.

If it is possible to get hold of the hub it is not hard to find the next step, again a network sniffer should do fine, the problem is that the next step is very likely just a jump station (a hacked computer only used to bounce traffic through). This computer is likely to be found in a country or a network where it is close to impossible to get any help.

If the jump station is found and someone can see the traffic to it, there is a slim chance that the attacker will use the same jump station again, but if he does you will probably have his IP.

4.3. Further Steps

To get any further you will probably need the help of the police. There are other ways of tracing an attack, but this is the most straightforward way.

Calling the police may be the best alternative in most cases (if you live in a country where the police has experience of this kind of attacks, e.g. in the UK or US), not that they are in any way certain to achieve any success but it is the cheapest way and they have, of late, had some success in these matters.

5. The attacks

Since there are a number of groups active in the DDoS-extortion business the scenarios often differ, but they still have much in common, described below is a typical scenario

5.1. The Ransom Note

Almost all attacks start out with a ransom note, commonly addressed to support or info e-mail addresses found on the victim's website. Below is an authentic ransom note that one of our clients received early 2004:

"Your business is targeted to attack by us. First will be attacked your site. You can increase your pipe all you want and it won't help. You have a flaw in your network that allows this to take place. You have 2 choices. You can ignore this email and try to keep your site up, > which will cost you tens of thousands of dollars, or you can send us $20.000 by Western Union to make sure that your site experiences no problems. If you send the $20.000 your site will be protected not just this weekend, but for the next 12 months. This will let you enjoy business with no worry. If you choose not to pay for our help, then you will probably not be in business much longer, as you will be under attack each weekend for the next 20 weeks, or until you close your doors. You can always choose to wait, see what happens, and then contact us for our help when you realize you can't do it yourself, however, then it will cost you more and your site will still be down. The choice is yours as we await your response".

The quality of the language varies much in the notes but the essence of the message is almost always the same:

- Pay $10 000 - $20 000
- If you don't pay we will close you down
- Resistance is futile

The ransom note is commonly received late Friday afternoon, since the attackers know it is much harder to get hold of any expertise during the weekend.

5.2. The First Attack

The first attack occurs shortly, often only a matter of minutes, after the ransom note arrives and it may or may not be preceded by a probing attack. The probe is not necessarily noticed, since it most often is only a short burst of traffic in order to determine the amount of traffic needed to knock the site off the net it may be interpreted as a network outage.

The first attack is often a SYN flood, and it may bring the site offline, generally creating panic within the victim company. After an hour or so the victim company will gather whatever technical resources still around and start thinking about a solution.

There always seems to be at least one person who thinks of the idea to change IP of the site (after all it saved the Whitehouse), the ISP is also generally contacted around this time. The ISP often generously offers to block out all traffic to the site (or has already done so since the attack often threatens their infrastructure), which of course is of no real help for the victim.

If it is possible and if they haven't gotten better advice, the victim will try to change the IP of the site. In the best case will buy them some time before the attacker changes the attack (this is an interactive human attacker, what hit the Whitehouse site was a worm). The first attack generally continues for a couple of hours to as much as a couple of days before it stops.

5.3. The calm after the first attack

The next few days nothing much will happen, the victim hopefully uses these days to figure out a solution to the problem. To implement a solution often takes more than a week.

5.4. The second attack (Larger Attack)

After the few days of calm the second attacks starts, it may or may not be preceded by a new ransom note, but it generally follows the pattern of the first attack. If no good countermeasures have been put in place the attack will again bring the site offline.

5.5. The next step

Sometimes the attackers give up after two attacks and simply go away for a while or they might administer another attack, but things generally calms down. This is a good time to make sure adequate countermeasures are put in place since it is highly likely that the attackers (or another group of attackers) will return.

6. Countermeasures

First of all it needs to be said that there are no countermeasures that will work 100% against all attacks. The bandwidth to the site will always be a limitation.

6.1. Motives

The three main motives for the attackers are:

- Fame or kudos
- Extortion
- Retaliation

The motives will very much affect the way the attack is performed and what countermeasures that are needed.

6.2. Temporary countermeasures

6.2.1. Pay the ransom

This has often been the case mostly for fear and fear of negative PR it has been kept quiet. We feel that today, this should not be the option. No good will come from this. The e-commerce community is not willing to tolerate it anymore.

6.2.2. Blacklists

If the number of attacking hosts is relatively small, the effects of the attack can be limited with the help of blacklists in the ACL (Access Control Lists) of upstream routers.

A blacklist is a list of IP-addresses that are blocked from communicating with the site, in this case the nodes of the DDoS network. In order to apply a blacklist you will have to have the full cooperation of your ISP, who often is rather reluctant to put big ACL's on their routers.

The good thing with a blacklist is that in theory you block only the offending IP-addresses and customers can access the site as usual. The bad thing is that it works mostly for smaller attacks.

6.2.3. Whitelists

If the IP-addresses for the customers are known or are within a geographically limited area it is often possible to allow only certain addresses to communicate with the site and block all other traffic.

The list of networks the customers or partners are located on is called a whitelist, and only these addresses will be able to communicate to the site. The good thing with whitelists is that they work on big attacks and that chances are good that the attacker thinks that the attack is successful (since he probably is not on the whitelist).

The bad thing is of course that few online businesses have their customers on well known networks and the ISPs are reluctant to put big ACL’s, which means that you will block many legitimate customers.

6.2.4. Moving the site

Since the nodes of the DDoS network attack a certain IP-address, it might seem like a good idea to change the IP of the site and block the IP being attacked. This has been known to work some times, but one should bear in mind that the attacker is a thinking person and that he most certainly has seen this trick before.

Since the tools used to attack the site most probably are fully automatic and reusable it is only a couple of keystrokes work to change the attack to the new IP-address.

6.3. Permanent solutions

The one most important resource in improving resistance against DDoS attacks is bandwidth. As said before there are no such thing as a 100% working DDoS mitigation solution as the potential number of clients in a DDoS network in theory is close to infinite.

Since however we live in a practical world there are good defences that actually work, at least most of the time. The main reason for this is that the number of clients of the DDoS networks is always finite.

During an attack the attacker looses nodes when the rightful owner of the computer acting as a node in the network discovers that something is wrong. When an attack is launched the users of the computers acting as nodes will experience a slowdown both in network performance and general computer performance, this increases the chances that this person realizes that something is wrong and cleans out the Trojan. This in turn means it will be more "expensive" for the attacker to mount a big attack than a small attack and that the best defence is to make sure that it is quite costly to mount an attack upon the site you are trying to protect.

As said before, the key to rise the attacking cost of a site is bandwidth but this is only true if the site itself can handle the packet-flood, which is not necessarily true. Most big firewall vendors are not able to handle SYN floods of the magnitude often seen in DDoS attacks, this means that the firewall most often will stop responding before the site runs out of bandwidth. There are a number of DDoS mitigation or Intrusion Prevention solutions that are known to handle above 1 Gb/s speeds; such as Riverhead (bought by Cisco) and TopLayer.

Since bandwidth is a limited resource and quite expensive it is important to get an estimate of how much bandwidth is needed. Below is an estimate of attacks:

- 0-50 Mbit/s a small attack, usually enough to bring a site without protection offline. Attacks of this size are quite common and may last for a long time (several days).
- 50-200 Mbit/s a mid-sized attack, commonly used against larger sites or sites with protection. Attacks of this size may last up to a day.
- 300-800 Mbit/s a large attack, this attack will probably not last very long
- 800- Attacks over 800 Mbit/s are very uncommon and there are a lot of misconceptions about this (there are a lot of more companies that think they have been hit by attacks of this magnitude than there are in reality).

The interesting thing is that the larger the attack, the more resource intensive it will be for the attacker and the more probable it is that the attack will be short. Please note that the numbers above are only estimates.

Appendix A

DOS - Denial Of Service

An attack that does not compromise the system integrity in any way, it only affects the availability of the system. DOS attacks are can either utilize a software flaw in order to crash the application or simply exhaust system resources.

DDOS - Distributed Denial Of Service

A DOS attack with several participating nodes. Usually DDOS attacks aims to exhaust system resources such as for example network bandwidth.

Blackhat

A cracker, someone bent on breaking into the system you are protecting. Oppose the less common white hat for an ally or friendly security specialist; the term grey hat is in occasional use for people with cracker skills operating within the law, e.g. in doing security evaluations.

Node

One computer in a network of computers

Zombie

A computer with an installed backdoor that lets the attacker use the machine without the owner's knowledge

Stacheldraht

A DDOS tool

Trinoo

A DDOS tool

TFN

A DDOS tool

Flooding

To overload a system with either legit requests or bogus traffic in order to fill up network bandwidth or exhaust system resources.