Available online at www.worldscientificnews.com

World Scientific News

An International Scientific Journal

WSN 41 (2016) 222-229 EISSN 2392-2192

A Study on Denial of Service Attacks in Cluster Based Web Servers

A. Poornima, D. Maheshwari*

Rathnavel Subramaniam College of Arts & Science, Sulur, Coimbatore, India

*E-mail address: maheswari@rvsgroup.com

ABSTRACT

In today's computing world Network Security became vulnerable. As many types of attacks makes the system unsecure and reduces the performance of servers. One such attack is - Denial of Service (DoS) attacks. The drawbacks of this attacks end user application accessibility will be diminishing and the request handling capacity of the application server will drastically subsides.

***Keywords*:** Flooding, Protocol, DoS attacks

1. INTRODUCTION

According to the WWW Security a DoS attack can be described as an attack designed to render a computer or network incapable of providing normal services. A DoS attack is considered to take place only when access to a computer or network resource is intentionally blocked or degraded as a result of malicious action taken by another user. These attacks don't necessarily damage data directly or permanently, but they intentionally compromise the availability of the resources.

The most common DoS attacks target the computer network's bandwidth or connectivity. Bandwidth attacks flood the network with such a high volume of traffic that all available network resources are consumed and legitimate user requests cannot get through, resulting in degraded productivity.

(Received 01 February 2016; Accepted 10 February 2016; Date of Publication 11 February 2016)

Connectivity attacks flood a computer with such a high volume of connection requests, that all available operating system resources are consumed and the computer can no longer process legitimate user requests.

2. CATEGORIES OF DoS ATTACKS

DoS attacks can be classified into five categories based on the attacked protocol level, DoS attacks in the **Network Device Level** include attacks that might be caused either by taking advantage of bugs or weaknesses in software, or by trying to exhaust the hardware resources of network devices. One example of a network device exploit is the one that is caused by a buffer overrun error in the password checking routine. Using these exploits certain Cisco 7xx routers could be crashed by connecting to the routers via telnet and entering extremely long passwords.

In the **OS level DoS** attacks take advantage of the ways operating systems implement protocols. One example of this category of DoS attacks is the Ping of Death attack. In this attack, ICMP echo requests having total data sizes greater than the maximum IP standard size are sent to the targeted victim. This attack often has the effect of crashing the victim's machine.

Application-based attacks try to settle a machine or a service out of order either by taking advantage of specific bugs in network applications that are running on the target host or by using such applications to drain the resources of their victim. It is also possible that the attacker may have found points of high algorithmic complexity and exploits them in order to consume all available resources on a remote host. One example of an application based attack is the finger bomb. A malicious user could cause the finger routine to be recursively executed on the hostname, potentially exhausting the resources of the host.

In **data flooding attacks**, an attacker attempts to use the bandwidth available to a network, host or device at its greatest extent, by sending massive quantities of data and so causing it to process extremely large amounts of data. An attacker could attempt to use up the available bandwidth of a network by simply bombarding the targeted victim with normal, but meaningless packets with spoofed source addresses. An example is flood pinging. Simple flooding is commonly seen in the form of DDoS attacks.

DoS attacks based on protocol features take advantage of certain standard protocol features. For example several attacks exploit the fact that IP source addresses can be spoofed. Several types of DoS attacks have focused on DNS, and many of these involve attacking DNS cache on name servers. An attacker who owns a name server may coerce a victim name server into caching false records by querying the victim about the attacker's own site. A vulnerable victim name server would then refer to the rogue server and cache the answer.

According to the WWW Security on Distributed Denial of Service (DDoS) attacks: "A DDoS attack uses many computers to launch a coordinated DoS attack against one or more targets. Using client/server technology, the perpetrator is able to multiply the effectiveness of the DoS significantly by harnessing the resources of multiple unwitting accomplice computers, which serve as attack platforms".

The DDos is the most advanced form of DoS attacks. It is distinguished from other attacks by its ability to deploy its weapons in a "distributed" way over the Internet and to aggregate these forces to create lethal traffic.

A Distributed Denial of Service Attack is composed of four elements as shown in Figure 1.

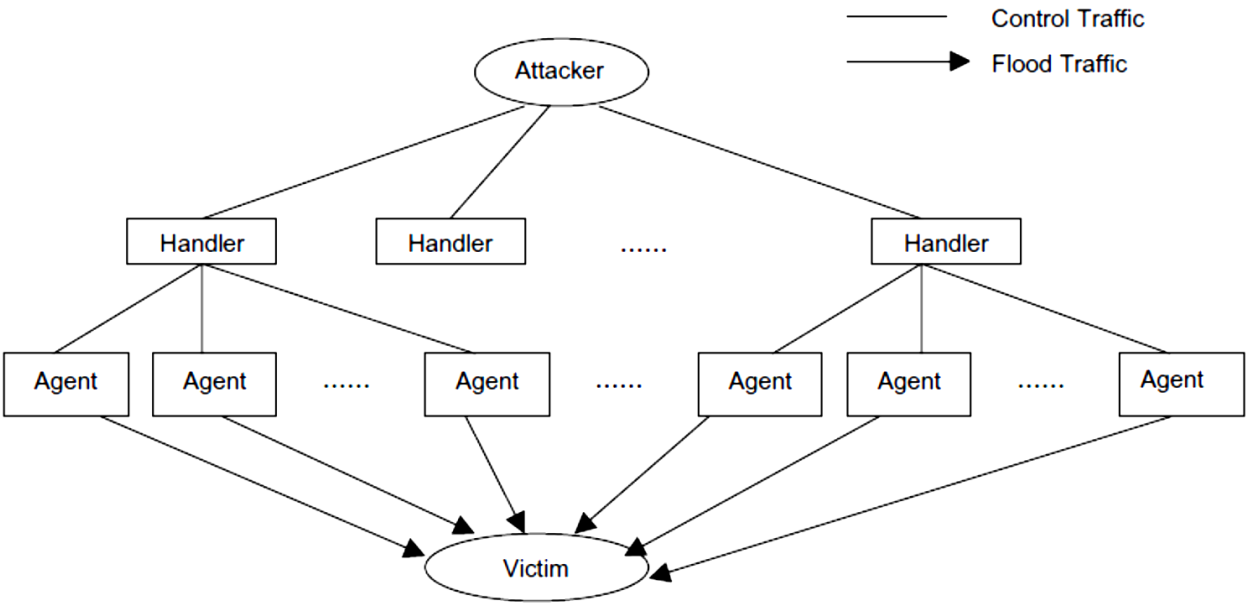

Figure 1. Architecture of DDoS attacks

- The real attacker
- The handlers or masters, which are compromised hosts with a special program running on them, capable of controlling multiple agents
- The attack daemon agents or zombie hosts, who are compromised hosts that are running a special program and are responsible for generating a stream of packets towards the intended victim. Those machines are commonly external to the victim's own network, to avoid efficient response from the victim, and external to the network of the attacker, to avoid liability if the attack is traced back.
- A victim or target host.

3. MOTIVATION

Among the many security threats in the current Internet, Distributed Denial of Service (DDoS) attacks are considered to be one of the most serious. Denial of Service (DoS) attacks aim to make the resources of the computer system of the victim unavailable or unreliable in providing their intended services. In the context of this thesis, DoS attacks try to consume and exhaust the victim's bandwidth or the server capacity. In DDoS attacks, the attacker compromises a large number of hosts in Internet and instructs them to conduct a coordinated attack. The network of the compromised hosts is called a botnet. While progress has been made in preventing or at least significantly lessening the impact of various security vulnerabilities, real progress in fighting DDoS is still missing. While automated software updates and antivirus programs can limit the number of compromised computers, there are still botnets comprising of millions of nodes. Another potential defense is to filter the packets sent by the DDoS attacker at a firewall after detecting the attack with and intrusion detection system (IDS).

These rule-based detection and filtering techniques have not been successful in filtering DDoS traffic because the DDoS attacker can send seemingly legitimate traffic.

In the case of open services, such as web servers, the DDoS attacker only needs to send large quantities of useless service requests. Thus, there might be no specific features of DDoS attack traffic that the rule-based filters can be instructed to filter. With such malicious but legitimate traffic, DDoS attackers are able to relatively easily bypass most means of DDoS defense. For the reasons explained above, much effort has been put into finding new methods for defending against DDoS attacks, also in the academic community. Among the many proposed solutions, one is to cluster network packets or service request with a learning algorithm and to use the learned clusters later to classify and filter the traffic.

The cluster-based filtering solutions do not require changes in the existing Internet model, and show promising results in DDoS defense since the problem of legitimate but malicious traffic is tackled from a different perspective. Namely, the clustering algorithms can be applied to unlabeled data, i.e., there is the unsupervised learning phase. During such a phase, the particular normal traffic distribution can be learned. Afterwards, the filter created according to the normal profile gives precedence to the traffic matching the normal classes. In such a way, the attack traffic, unless it matches the normal profile, is filtered, without the need to explicitly identify it as malicious.

4. DDoS DETECTION MECHANISMS

Detection mechanisms might not obviously seem necessary, since the DDoS attacker does not focus on hiding the attack which becomes apparent as soon as the disruption of resources causes degradation of target's services. However, from the defender's point of view, the moment of detection of the attack is important. First, earlier detection will enable faster reaction to the attack and lessening its impact. Second, detection can help in finding the sources of the attack and eventually identifying the attackers.

The two main types of detection mechanisms are determined depending on the concept that detection is based on:

- detection based on specific features of the DDoS traffic
- anomaly-based detection.

Before describing DDoS detection based on specific features of the DDoS traffic, we should mention the security platform called network intrusion detection system (NIDS). Namely, NIDS are implemented as a separate devices or software components with the function of monitoring both ingress and egress traffic in order to detect intrusions. Intrusions include different types of malicious activities and attacks, particularly DDoS attacks. The detection based on specific features of the attack traffic is often the part of signature-based network intrusion detection systems. Signatures are formed from specific features of the know attack types. Example assumptions that are made about specific features of DDoS traffic are:

- the DDoS traffic does not comply with the TCP flow control
- there is a disproportion between the packet rate at the victim and close to the sources
- the ratio of SYN packets compared to FIN and RST packets will be unbalanced in the case of SYN attacks.

However, these assumptions are not always valid. For example, an attack using a large number of bots may be conducted in such a way that each of them opens a legitimate TCP connection to the target. The attacker may send the FIN or RST packets in conjunction with the SYN.

5. DDoS SOURCE IDENTIFICATION TECHNIQUE

Source identification techniques try to identify the attack sources. In the case of a single flooding source this is still doable. In the case of DDoS, having many bots involved and with the attacker's communication to them usually being encrypted, source identification is difficult or impossible to achieve. Examples of source identification are IP traceback schemes with hash-based IP traceback scheme arguably be the most effective.

6. DDoS PREVENTION AND REACTION MECHANISMS

Prevention mechanisms are intended to prevent the attack traffic from reaching the target preferably close to the attack sources, while reaction mechanisms need to be launched as the final defense step, when prevention techniques fail and after the attack is detected. The main classes of prevention mechanisms include:

- general mechanisms which prevent host compromise
- identifying and disrupting botnets, and
- filtering techniques when deployed close to the attack sources
- filtering techniques when they are employed for the target bandwidth management.

General mechanisms include common techniques that help in improving the security of the system, such as, disabling unused services, replication of resources, installing newest security updates, and disabling IP broadcast. Since botnets are the most important resource for the DDoS attacker, it is crucial to identifying and disrupt the botnets. Nowadays, the attackers are able to compromise large botnets, having hundreds of thousands or even millions of bots. Thus many Community Emergency Response Teams (CERTs) and also the corporations such as Microsoft, have taken part in actions for disrupting some of the large botnets.

7. CLUSTERING ALGORITHMS FOR NETWORK TRAFFIC

The important factor in DDoS defense that uses cluster-based filtering is to find `the good set of clusters in the traffic. Thus many academic solutions are focused solely on the question how to cluster network traffic into qualitatively good classes. Some of the existing clustering algorithms are: Density Based Spatial Clustering of Applications with Noise (DBSCAN), K-Means and AutoClass. These algorithms are evaluated and compared from the point of accuracy of clustering traffic into known traffic classes, then from the point of algorithm speed and number of clusters that is produced.

A traffic cluster is defined by ranges of traffic feature values. All the packets in which the feature values are inside these ranges form a traffic class. The terms traffic class and traffic

cluster are often used interchangeably. Additionally, we also take as a part of the cluster definition the amount of the recorded traffic belonging to the class. The web server has exposed its services through a router on which the clustering mechanism is deployed.

The normal traffic is observed over a certain period of time. Afterwards, some off line algorithm is applied to the traffic records. The off line algorithm outputs a cluster set which represents the normal traffic profile. A filtering policy on the router is created using that profile. When the attack happens, a predefined action using the filtering policy is taken on the router as a DDoS defense reaction.

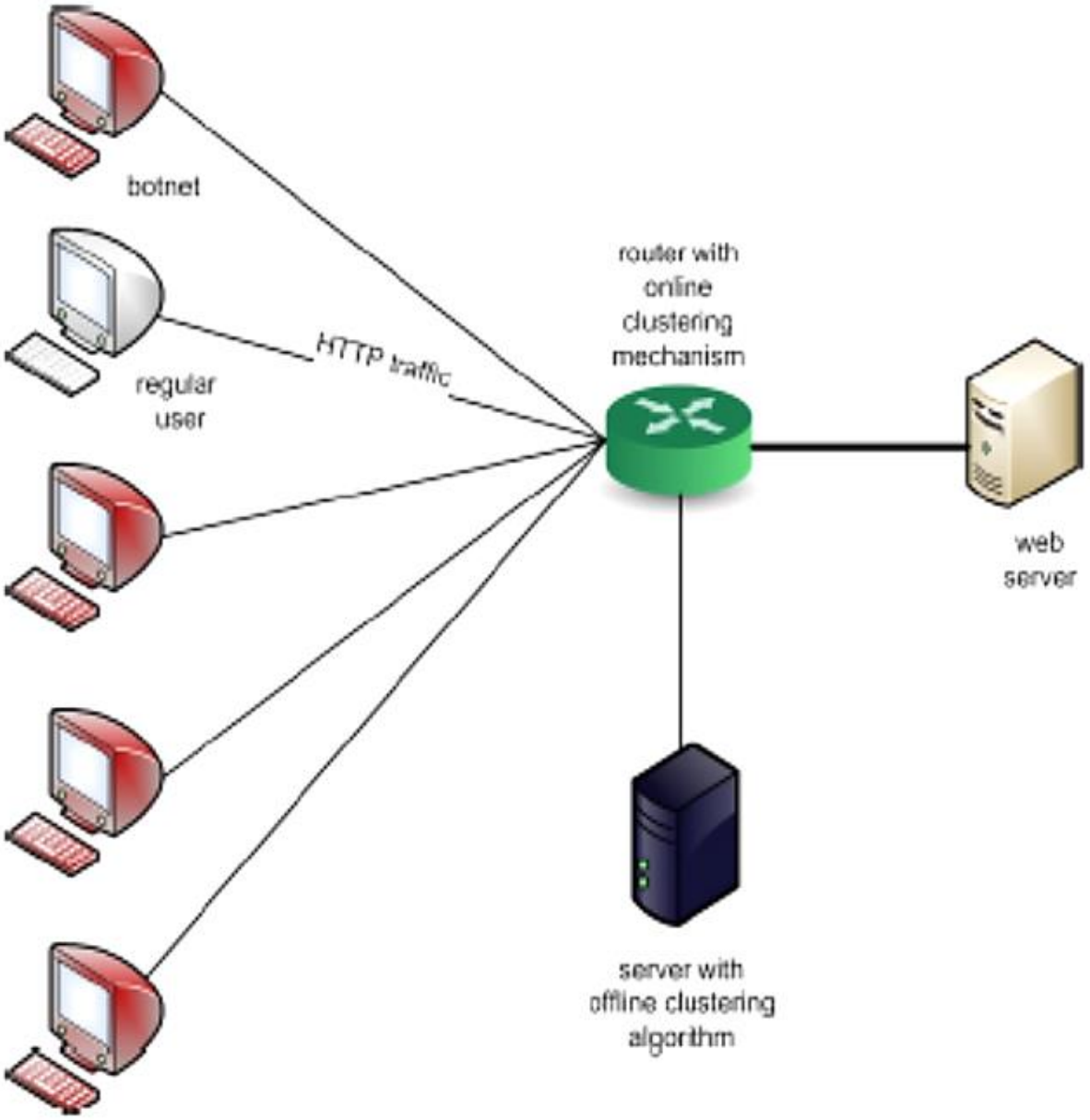

Figure 2. Scenario with clustering mechanism deployed

The heavy task of the initial cluster computation may be allocated to a separate server which executes the off line algorithm. In that case, such a server may instruct the router about the filtering policy that the router should deploy. The online algorithm might be employed on the router to cluster the traffic stream passing through the router. In the case of an attack, the online algorithm provides information about the distribution of the attack traffic. Using the output of the online algorithm, the filter reservations may be adapted so that the DDoS reaction is improved. The success of this method, however, depends on the particular clustering algorithm used. If successfully employed, this method can be also used for DDoS detection. We analyze different possibilities when only the off line algorithm is used. Depending whether the attacker or defender is more agile in changing its attack or defense strategy. Secnario with clustering mechanism is deployed in Figure 2.

8. RESULTS

The DDoS attacker might be restricted in the scope of source IP addresses or other features of traffic for conducting the attack. Such attacker capabilities depend on the available attack tools and the location in Internet and size of the botnets he controls. Also, the effectiveness of the clustering defense against such an attack depends on the traffic clustering algorithm and on the traffic features that the algorithm uses. If the algorithm finds cluster features that are specific to the normal traffic and difficult to imitate by the attacker, then the attacker cannot conduct the optimal attack and the impact of the attack will be less serious. The attacker is better able to adapt his traffic to the certain classes, and then the analysis becomes more complex. In that case, the defender could use the knowledge about the attacker to adapt the cluster reservations accordingly, or even his choice of the clustering algorithm could be influenced by that knowledge. However, the analysis in this case would require data about real DDoS attacks or a set of experiments conducted in order to find about the real capabilities of DDoS attackers.

9. CONCLUSION

DoS/DDoS is one of the main security threats in the Internet. Defending against DoS/DDoS becomes a necessary step that must be considered by the companies and ISPs. DoS/DDoS detection is regarded to be one of the main phases in overcoming the DoS/DDoS problem. In this paper, the DDoS attacker might be restricted in the scope of source IP addresses or other features of traffic for conducting the attack.

References

[1] Agrawal, R., and Srikant, R. Fast algorithms for mining association rules in large databases. In Proceedings of the *20th International Conference on Very Large Data Bases*, VLDB '94, Morgan Kaufmann Publishers Inc., pp. 487-499.

[2] Anderson, T., Roscoe, T., and Wetherall, D. Preventing internet denial-of-service with capabilities. *SIGCOMM Comput. Commun. Rev.* 34 (January 2014), 39-44

[3] Bezdek, J. C. Fuzzy Mathematics in Pattern Classification. PhD thesis, Applied Math. Center, Cornell University, Ithaca, 2013.

[4] Blake, S., Black, D., Carlson, M., Davies, E., Wang, Z., and Weiss, W. RFC 2475: An architecture for differentiated services, Dec. 2010. Status: PROPOSED STANDARD.

[5] Bretthauer, K. M., and Shetty, B. The nonlinear resource allocation problem. *Operations Research* 43(4) (2014) 670-683

[6] Bretthauer, K. M., and Shetty, B. The nonlinear knapsack problem - algorithms and applications. *European Journal of Operational Research* 138(3) (2012) 459-472

[7] Cheeseman, P., Kelly, J., Self, M., Stutz, J., Taylor, W.,and Freeman, D. Readings in knowledge acquisition and learning.Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2013,ch. AutoClass: a Bayesian classi_cation system, pp. 431-441.

[8] Cheeseman, P., and Stutz, J. Advances in knowledge discoveryand data mining. American Association for Artificial Intelligence, Menlo Park, CA, USA, 1996, ch. Bayesian classification (AutoClass): theory and results, pp. 153-180.

[9] Cooke, E., Jahanian, F., and Mcpherson, D. The zombie roundup: Understanding, detecting, and disrupting botnets. SRUTI '05: Steps to Reducing Unwanted Traffic on the Internet Workshop (2012). pp. 39-44.

[10] Cormode, G., Korn, F., Muthukrishnan, S., and Srivastava, D. Finding hierarchical heavy hitters in streaming data. *ACM Trans. Knowl. Discov. Data* 1 (February 2010), 2: 1-2: 48

[11] Cormode, G., and Muthukrishnan, S. Diamond in the rough: Finding hierarchical heavy hitters in multi-dimensional data. In Proceedings of the 23rd ACM SIGMOD International Conference on Management of Data (2012), ACM Press, pp. 155-166