Detecting, Capturing & Resolving of DDoS Attacks with Hadoop

Nakul Chorey[1] Rujuta Kate[2] Prajakta Khatavkar[3] Ms. Renuka. R. Kajale[4]
[4]Guide
[1,2,3,4]Department of Computer Engineering
[1,2,3,4]Savitribai Phule University, Pune, Maharastra, India

Abstract— Many systems use servers to manage and store their data, sometimes the servers are slowed down because of multiple user requests. Most of which are attackers or unauthorized users and some are genuine users. In computing, a denial-of- service (DoS) attack is an attempt to make a machine or network resource unavailable to its intended users. A distributed denial-of-service (DDoS) is where the attack source is more than one, often thousands of unique IP addresses. Flooding is one of the typical DDoS attacks that exploit normal TCP connections between a client and a target web server. In this project we are trying to devise a DDoS anomaly detection method on Hadoop that implements a MapReduce-based detection algorithm against the Flooding attacks.

Key words: DDoS, DoS, Map-Reduce, Hadoop

I. INTRODUCTION

[1]The cyber threat landscape is evolving faster than vendors can create mitigations for attacks. In the past, a 1 to10 GB attack would have been highly unusual and require a botnet of 100,000+ victim systems to participate. Today, 300+ GB attacks have become a norm. Botnets are increasing in number and size. Hybridization of the enterprise and cloud solutions increases the surface area of exposure to DDoS threats. To effectively mitigate an attack the scale and complexity seen today requires automated attack identification and response.

Detection of DDoS Attack is a basic measure towards defence. DDoS attacks may result in system performance degradation of the targeted network. Which can cause the services intended to the genuine users may not function or may produce delayed results.

In computing, a denial-of-service (DoS) attack is a kind of attack where the attacker attempts to stop the legitimate user from accessing the resources by disrupting the source of services. Denial of service is typically accomplished by flooding the target machine or resource with continuous requests in an attempt to overload systems and prevent some or all legitimate requests from being fulfilled.

[2]A distributed denial-of-service (DDoS) is a kind of attack where source of attack is more than one, often thousands of, unique IP addresses are used to perform the attack. Its main motive is to disrupt the regular access of resources. In the recent years the scale of DDoS attacks is increased, even reaching over 400Gbit/s.

Criminal or attackers of DoS and DDoS attacks often target sites or services which are hosted on high-profile web servers such as banks, credit card payment gateways.

The United States Computer Emergency Readiness Team (US-CERT) defines symptoms of denial-of-service attacks to include:

1) Unusually slow network responses (opening files or accessing web sites gets slower)
2) Unavailability of a particular web site or webpage.
3) Inability to access any web site.
4) Dynamic increase in the number of spam emails received (This type of DoS attack is considered an e-mail bomb)

Additional symptoms may include:

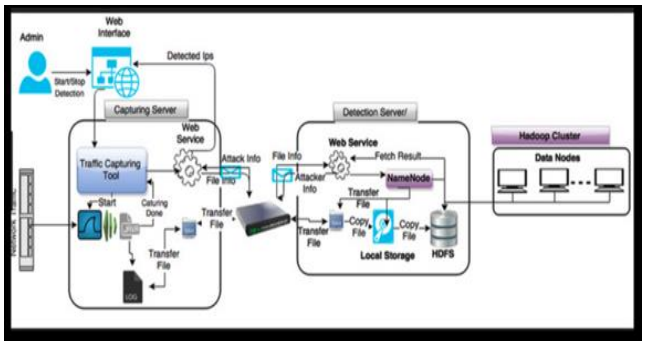
1) Disconnection of a wireless or wired internet connection.
2) Long-term denial of access to the web or any internet services.

[3]If the attack is conducted on a sufficiently large scale, entire geographical regions of Internet connectivity can be compromised without the attacker's knowledge or intent by incorrectly configured or flimsy network infrastructure equipment. HTTP GET flooding is one of the typical DDoS attacks that exploit normal TCP connections between a client and a target web server. As the volume of Internet traffic is increasing exponentially year after year, the Intrusion Detection Systems (IDSes) have faced the issue on how to assure both scalability and accuracy of analyzing the DDoS attack from these huge volumes of data.

In recent years, lighter-weight virtualization solutions have begun to emerge as an alternative to virtual machines. Because these solutions are still in their infancy, however, several research questions remain open in terms of how to effectively manage computing resources. One important problem is the management of resources in the event of virtualization. For some applications, overutilization can severely affect performance.

[6]Hadoop is an open-source distributed cluster platform that includes a distributed file system, HDFS and the programming model, MapReduce. In this project we are trying to devise a DDoS anomaly detection method on Hadoop that implements a MapReduce-based detection algorithm against the HTTP GET flooding attack.

II. PROPOSED SYSTEM



[7]Through traffic capturing and detecting the legitimate and usual user on the basis of number of requests sent from the ip addresses. Further, blocking stage takes place where the attacker ips are blocked and usual user is allowed to proceed further. The outcome of the project will be in real time in order to ensure the uninterrupted access to service. The process starts with capturing the traffic of packets using tcpdump, after that the traffic is monitored and this is done using a java based code which separates the ip addresses from

the tcpdump packets, the next step here is to check if the ip addresses are valid or not using regular expressions, the valid ip addresses are then separated and the others are discarded. Then the valid ip addresses are considered and the no of request they have sent to the server are checked and a threshold value is set, if the count of request is more than the threshold value, the ip address is considered as an attacker, else it is considered as a genuine user and allowed to access the server, and the attacker is blocked for a day as ip addresses are dynamic and they keep on changing so they get unblocked after a day.

III. CONCLUSION

In this paper we presented an efficient mechanism for detection and blocking of DDOS attack on a real time system. The proposed work is able to observe requests from genuine and malicious user using the runtime logs. The mechanism is capable of blocking the attacker's ip at an early stage thus preventing the harm to the system and saving the resources. Our method also uses Hadoop for parallel data processing which increases the overall performance.

REFERENCES

[1] Bhushan Lakhe, Practical Hadoop Security, Apress, 2014.
[2] Dacier M, Research in Attacks Intrusions and Defenses, Springer, 2017.
[3] Jayachander Surbiryala, "A Framework for Improving Security in Cloud Computing," IEEE
[4] Transactions on Computers, vol.24, pp.260-264, 2017.
[5] Jeremy Sims, "Securing Cloud SDN and Large Data Network Environments from Emerging,DDos Attacks,"IEEE Transactionon Computers, vol.11, pp.466-469,2017.
[6] Dr.Ammar Almomani, "Hadoop Based Defense Solution to Handle Distributed Denial of Service (DDoS) Attacks," in Second International conference,2013, pp.61-65
[7] Sufian Hameed, "The Efficacy of Live DDoS Detection with Hadoop," presented at 4th interna-tional conference, 2015.

