Distributed Intrusion Detection System using Blockchain and Cloud Computing Infrastructure

Dr. Manish Kumar[1]
Assistant Professor
Department of Computer Applications
M S Ramaiah Institute of Technology
Bangalore, India
manishkumarjsr@yahoo.com

Ashish Kumar Singh[2]
Research Scholar
Department of Computer Applications
M S Ramaiah Institute of Technology
Bangalore, India
ashish.msrit10@gmail.com

Abstract— **Intrusion Detection System is a well-known term in the domain of Network and Information Security. It's one of the important components of the Network and Information Security infrastructure. Host Intrusion Detection System (HIDS) helps to detect unauthorized use, abnormal and malicious activities on the host, whereas Network Intrusion Detection System (NIDS) helps to detect attacks and intrusion on networks. Various researchers are actively working on different approaches to improving the IDS performance and many improvements have been achieved. However, development in many other technologies and newly emerging techniques always opens the doors of opportunity to add a sharp edge to IDS and to make it more robust and reliable. This paper proposes the development of Distributed Intrusion Detection System (DIDS) using emerging and promising technologies like Blockchain upon a stable platform like cloud infrastructure.**

Keywords— Network Intrusion Detection System (NIDS), Blockchain, Cloud Computing, Distributed Intrusion Detection System (DIDS), Host Intrusion Detection System (HIDS).

I. INTRODUCTION

Intrusion Detection System (IDS) is one of the most common security systems. It is used for the protection of network infrastructure and computers from malicious activities and unauthorized usages. It detects the different types of threats. It helps users and network administrators to take preventive measures. IDS plays an important role in securing IT infrastructure. Intrusion detection systems capture and analyze the network traffic to detect suspicious activity [3][5][6]. IDS mainly works on two different approaches:

- **Anomaly detection: -** In this technique, network traffic or host OS behavior is analyzed based on various parameters and compared with the normal behavior. If the system detects any deviation from normal behavior, it raises an alarm.

- **Misuse/Signature detection:** - This technique looks for a specific pattern of behavior which is already known as an attack. All the malicious patterns and behaviors which are identified as attacks are stored in the IDS signature database. These signature databases are continuously updated and used for attack detection. The limitation of this technique is that it will not be able to detect a novel attack, as a signature will not be available.

Intrusion Detection Systems are broadly classified into two categories:

- **Network Intrusion Detection System (NIDS):-** It captures the packets from network traffic. The header of the captured packets is analyzed based on various parameters to detect malicious activities. It can be set up in the network backbone, server, switches, and gateways.

- **Host Intrusion Detection System (HIDS):-** It is installed on the individual system to detect the intrusion or misuse. HIDS analyzes the key system files, process behaviors, unusual resource utilization, unauthorized access, etc.

Based on the needs of the organization, the type of IDS can be decided. For large organizations, NIDS will be a cheaper solution. However, it is important to understand that both NIDS and HIDS use different techniques and one cannot be considered as the substitution for others.

II. DISTRIBUTED INTRUSION DETECTION SYSTEM

As discussed in the previous section, NIDS and HIDS are based on different approaches. To get the overall protection, sometimes we need to use both types of systems.

In a large network, multiple Network Intrusion Detection Systems are deployed across the network. Theses distributed IDS share the logs and alert information with each other. Such an arrangement of multiple IDS is called a Distributed Intrusion Detection System (DIDS). The type and volume of information shared among the distributed IDS is configured by the administrator and need to be fine-tuned from time to time. It facilitates advanced persistent threat analysis, network monitoring, and instant attack analysis of the whole network. It helps the administrators to get a broader view of the network attack [11].

Network monitoring and IDS alert analysis are some of the most crucial tasks. Though the security devices and software can provide adequate security one cannot rely on that completely. Attack patterns keep changing frequently. Attackers always device new and novel techniques to evade the detection and hence the security devices need to be patched and tuned accordingly.

Threat and alert analysis of the entire network helps the administrators to understand the complex attack patterns. Based on the analysis, a signature or set of rules can be generated. These can be distributed to individual IDS to protect the segment of the network from a similar attack in the future [1][4][9].

A. Advantages of DIDS

With the growing size of network infrastructure, scalability and performance is always the major concern for single-mode IDS. It's difficult for the single-mode IDS to detect the attack pattern scattered across different geographical locations of an enterprise network. DIDS has an advantage over single-mode IDS to collect and corroborate data among the peer IDS and detect the stealthy attack pattern. Many times in the case of Advanced Persistent Attack, it is observed that the attack may be initiated in some specific region of the network and then slowly spread to the entire network. In DIDS since all the IDS are connected, any attack detected in a specific region or segment of the network can help the other IDS to learn and update their rule-base. The administrator can take preventive measures and protect the rest of the network from attack.

Nowadays, attack patterns are becoming more and more complex. With the increased complexity of attacks, it is possible that the administrator of one network segment may not take a small incident seriously which may be a part of a bigger coordinated attack. However, when the attack pattern of the entire network is analyzed together, it may represent a serious threat.

The DIDS system gives the administrators the fastest and easiest way to identify the attacks coordinated across the multiple network segments. The centralized log analysis of distributed IDS allows the analyst to discover complex attack patterns and take preventive measures easily.

B. Incident Analysis with DIDS

With the growing size of networks, the number of attacks is also increasing. Identifying the intruders and malicious activities at the right time is the most important and crucial task. One needs to know how the attack initiated, where it initiated, what the attacker did, what the level of threat was, and how to prevent it.

DIDS provides a centralized platform where the threat can be detected instantly no matter in whatever network segment it occurs. As it gives an advantage to the administrator for centralized analysis, it also requires proper planning and implementation. Since the whole network infrastructure depends on DIDS, it should have potential power, flexibility, and strength to detect the threat as quickly as possible. More delay in detection will be more devastating. Keeping in view the resource requirements, scalability, and the need for computational capability, we have proposed a system that is based on cloud, which can satisfy these requirements on-demand basis. As the DIDS is based on the internetworking of individual IDS, trust, authenticity, and reliability of the alert received from the individual IDS is a matter of concern. In some situations, it is possible that the attacker has compromised a particular IDS or it is weakly configured, which can fuzz the server with the wrong alert and misguide the administrator. Keeping given trust, authenticity, and reliability issues, we have proposed the use of blockchain technology, which is one of the most promising solutions to solve these issues. The next section briefly describes how blockchain can be used in DIDS, followed by the discussion on the integration of DIDS with cloud infrastructure. Blockchain is a new and emerging technology. A detailed discussion about the blockchain and cloud technology is beyond the scope of our work and the topic of this paper. Hence only specific things related to our work have been discussed in this paper.

III. Blockchain-Based Intrusion Detection

It is observed in recent days that attackers are applying more complex and advanced techniques to attack the system and avoid detection. It is also possible that if any segment of the network is misconfigured or compromised by the attackers, then it may divert the attention of the network administrator. Since multiple IDS are connected with the centralized server, the authenticity of the logs and alerts received from the individual IDS raises a major concern [2].

To make the overall system robust and trustworthy, blockchain seems to be the most reliable solution. In the section, we will focus on the challenging issues of DIDS and how blockchain can help to resolve these [8][10][14].

A. Challenges in Distributed Intrusion Detection

There are many issues in collaborating the multiple IDS and sharing the logs with a centralized system. Trust and authenticity are major issues that need to be resolved.

Following are the requirements which Distributed Intrusion Detection System should satisfy:

- Integrity: - The integrity of the alert generated by individual IDS is very important. Logs and alerts should be tamper-proof and in no circumstances, should they be accessible and modifiable by an attacker or any individual.

- Consensus: - All the participating IDS should have a common consensus on the type and quality of alert generation.

- Scalability: - As the network size may grow and shrink, the computational demand for the centralized server should also be able to adjust accordingly.

- Privacy: - Participating individual IDS should have rights and control over selective disclose and accessibility of alert data.

DIDS is vulnerable to attacks from inside of the network, where the intruder has somehow got the authorized access to the network. Security and reliability of the central server are most crucial. The objective of central server deployment in DIDS is to collect the logs and alerts from individual IDS for centralizing analysis and easy decision making. The above-mentioned requirements become one of the most important criteria and challenge to be satisfied to build a robust DIDS [13][12].

B. *Blockchain-Based Solutions*

Blockchain provides the best promising solution to satisfy the above-mentioned requirements and challenges. The implementation of a secure distributed ledger is proposed for sharing the logs and alerts generated by the individual IDS.

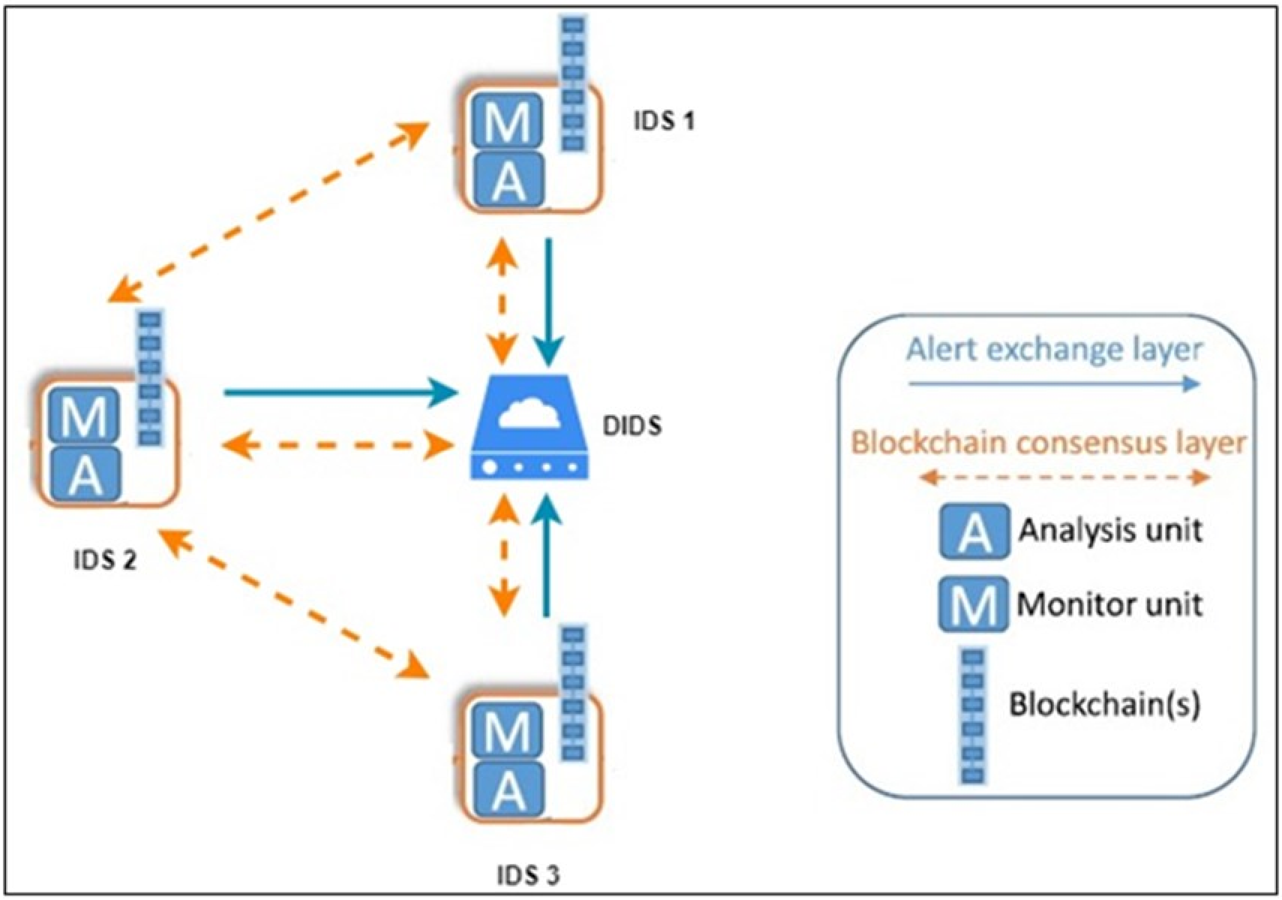

Fig 1:- DIDS Architecture Using Blockchain and Cloud Infrastructure

Alerts generated by individual IDS are stored as a transaction in the blockchain. All the individual IDS connected with the centralized server run a consensus protocol to validate the transaction before adding it into the blockchain. Such an arrangement guarantees that only validated, authenticated alerts are updated in the centralized server. These alerts are tamper-proof and accessible to all the participating IDS [15][1].

IV. Integration with Cloud Infrastructure

Though DIDS is a promising solution for protecting the entire network, it poses a lot more challenging issues like its performance and scalability. When a network is very big and individual IDS dumps the logs to a centralized server, it becomes big data. It's a challenge for the system to analyze the large size of data in real-time. Huge infrastructure and investment are required to build such real-time DIDS.

A cloud-based centralized server provides a promising solution to handle these challenges. Its a cost-effective and scalable solution. Based on the need, the capacity of the cloud based centralize server can be scaled-up and down.

The cloud-based platform not only provides a cost-effective and scalable system for handling large data size but also supports big data analytical tools. It can be used to analyze the logs and alerts data collected on centralized sever and help to identify different types of attack patterns and threats in real-time [6][7].

V. Experimental Results

For experimental analysis, we used Amazon Cloud Services. We configured the nodes on the Amazon EC2 cloud. The master nodes automatically create the computational slave nodes. The advantage of such architecture is that the user can any time increase and decrease the computational power based on the requirements and budget.

In our experiment, we tested the performance of the overall system using Auto Scaling. Amazon EC2 Auto Scaling helps to ensure that resources are available on-demand to handle the load of applications. The user can specify the maximum and the minimum number of instances in each Auto Scaling Group as per the requirements. Amazon EC2 auto-scaling ensures that resources are optimally utilized between the minimum and the maximum number of instances.

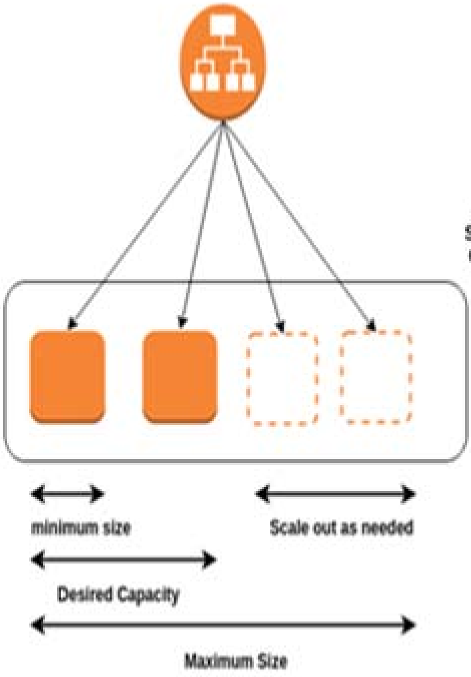

Fig 2:- Amazon EC2 Scaling

For example, the Auto Scaling group is shown in Fig:-2 has a maximum size of four instances and minimum size of one instance. It has the desired capacity of two instances. The system adjusts the number of instances within the minimum and a maximum number of instances based on the policies and criteria defined.

TABLE I. COMPARISON OF EXECUTION TIME AND SPEED-UP

Number of Instances	Log/Alert Size	Execution Time (sec)
1	500 MB	50
2	1 GB	56.5
3	2 GB	53.2
4	3 GB	54.4
5	4 GB	51.6
6	5 GB	52.6

For our experiment, the minimum number of instances was set to 1 and the maximum number of instances was set to 6. The centralized server process was configured on the master node and slave nodes were configured as independent IDS. Since the whole system was deployed on the cloud, communication time was reduced. However, in real-time experiments, communication delay can be a major factor to be analyzed which was out of the scope for this experiment.

In our experiment, the analysis of DIDS performance is carried out by performing the analysis of logs and alerts of varying size between 500MB to 5 GB. The computational power was continuously increased with the increasing size of data as shown in Table 1. The graph in Fig 3 shows that with the flexible scalability of the system computing power, the performance of the system was consistent.

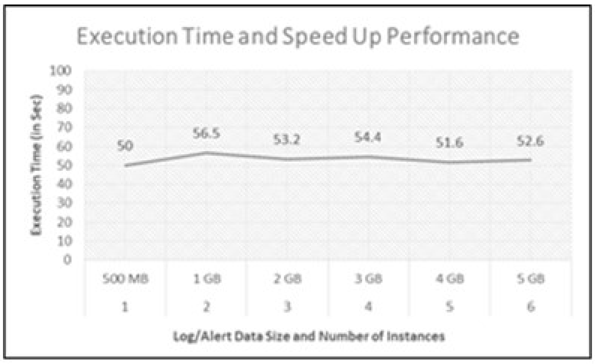

Fig. 3:- Load Vs. Speed-up Performance

VI. CONCLUSION

In this paper, we have presented the architecture of the Distributed Intrusion Detection System using Cloud Computing Infrastructure and Blockchain. We have shown the performance of the DIDS server with a varying load of data. There are many other issues like communication delay, the overhead of blockchain, cost of implementation, etc. which has to be discussed and analyzed.

REFERENCES

[1] A. A. Titorenko and A. A. Frolov, "Analysis of modern intrusion detection system," *2018 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus)*, Moscow, 2018, pp. 142-143.

[2] Alexopoulos, Nikolaos & Vasilomanolakis, Emmanouil & Réka Ivánkó, Natália & Mühlhäuser, Max. (2018). Towards Blockchain-Based Collaborative Intrusion Detection Systems: 12th International Conference, CRITIS 2017, Lucca, Italy, October 8-13, 2017.

[3] Axelsson, Stefan. Intrusion detection systems: A survey and taxonomy. Vol. 99. Technical report, 2000.

[4] H. M. Anwer, M. Farouk and A. Abdel-Hamid, "A framework for efficient network anomaly intrusion detection with features selection," *2018 9th International Conference on Information and Communication Systems (ICICS)*, Irbid, 2018, pp. 157-162.

[5] H.Debar, M. Dacier, and A. Wespi, “Towards a taxonomy of Intrusion Detection Systems”, Computer Networks, vol 31, n0. 8, pp. 805-822, 1999.

[6] Holtz, Marcelo D. ; Bernardo David ; Sousa Jr., R. T. . Building Scalable Distributed Intrusion Detection Systems Based on the MapReduce Framework. Telecomunicacoes (Santa Rita do Sapucai), v. 13, p. 22-31, 2011.

[7] J. Dean and S. Ghemawat, MapReduce: Simplified Data Processing on Large Cluster, USENIX OSDI,2004.

[8] J. Yang, C. Shen, Y. Chi, P. Xu and W. Sun, "An extensible Hadoop framework for monitoring performance metrics and events of OpenStack cloud," *2018 IEEE 3rd International Conference on Big Data Analysis (ICBDA)*, Shanghai, 2018, pp. 222-226.

[9] K. Kato and V. Klyuev, "Development of a network intrusion detection system using Apache Hadoop and Spark," *2017 IEEE Conference on Dependable and Secure Computing*, Taipei, 2017, pp. 416-423.

[10] Konstantin Shvachko, Hairong Kuang, Sanjay Radia, and Robert Chansler, "The Hadoop Distributed File System," IEEE 26th Symposium on Mass Storage Systems and Technologies (MSST), pp.1-10, 2010.

[11] S. Ghribi, A. M. Makhlouf and F. Zarai, "C-DIDS: A Cooperative and Distributed Intrusion Detection System in Cloud environment," *2018 14th International Wireless Communications & Mobile Computing Conference (IWCMC)*, Limassol, 2018, pp. 267-272.

[12] Suah Kim, Beomjoong Kim, and Hyoung Joong Kim. 2018. Intrusion Detection and Mitigation System Using Blockchain Analysis for Bitcoin Exchange. In Proceedings of the 2018 International Conference on Cloud Computing and Internet of Things (CCIOT 2018). ACM, New York, NY, USA, 40-44.

[13] W. Meng, E. W. Tischhauser, Q. Wang, Y. Wang and J. Han, "When Intrusion Detection Meets Blockchain Technology: A Review," in *IEEE Access*, vol. 6, pp. 10179-10188, 2018.

[14] Yeonhee Lee and Youngseok Lee. 2012. Toward scalable internet traffic measurement and analysis with Hadoop. *SIGCOMM Comput. Commun. Rev.* 43, 1 (January 2012), 5-13.

[15] Zohreh Abtahi Foroushani and Yue Li. 2018. Intrusion Detection System by Using Hybrid Algorithm of Data Mining Technique. In Proceedings of the 2018 7th International Conference on Software and Computer Applications (ICSCA 2018). ACM, New York, NY, USA, 119-123.