RESEARCH ARTICLE OPEN ACCESS

Anomaly Detection in Infrastructure Service of Cloud Computing

N.Praveena[1], SK.Sofia[2], D.Srinivasulu [3]
M.Tech [1], Assistant professor [2], Professor [3]
H.O.D of Computer Science[3]
Department of Computer Science
Priyadarshini College of Engineering and Technology
Nellore
Andhra Pradesh - India

ABSTRACT
Cloud services are prominent within the private, public and commercial domains. Many of these services are expected to be always on and have a critical nature; therefore, security and resilience are increasingly important aspects. In order to remain resilient, a cloud needs to possess the ability to react not only to known threats, but also to new challenges that target cloud infrastructures. In this paper we introduce and discuss an online cloud anomaly detection approach, comprising dedicated detection components of our cloud resilience architecture. More specifically, we exhibit the applicability of novelty detection under the one-class support Vector Machine (SVM) formulation at the hypervisor level, through the utilisation of features gathered at the system and network levels of a cloud node. We demonstrate that our scheme can reach a high detection accuracy of over 90% whilst detecting various types of malware and DoS attacks. Furthermore, we evaluate the merits of considering not only system-level data, but also network-level data depending on the attack type. Finally, the paper shows that our approach to detection using dedicated monitoring components per VM is particularly applicable to cloud scenarios and leads to a flexible detection system capable of detecting new malware strains with no prior knowledge of their functionality or their underlying instructions.

Keywords:- Security, resilience, invasive software, network-level security and protection.

I. INTRODUCTION

Cloud data centers are beginning to be used for a range of always-on services across private, public and commercial domains. These need to be secure and resilient in the face of challenges that include cyber attacks as well as component failures and mis-configurations. However, clouds have characteristics and intrinsic internal operational structures that impair the use of traditional detection systems. In particular, the range of beneficial properties offered by the cloud, such as service transparency and elasticity, introduce a number of vulnerabilities which are the outcome of its underlying virtualised nature. Moreover, an indirect problem lies with the cloud's external dependency on IP networks, where their resilience and security has been extensively studied, but nevertheless remains an issue [1]

The approach taken in this paper relies on the principles and guidelines provided by an existing resilience framework [2]. The underlying assumption is that in the near future, cloud infrastructures will be increasingly subjected to novel attacks and other anomalies, for which conventional signature based detection systems will be insufficiently equipped and therefore ineffective. Moreover, the majority of current signature-based schemes employ resource intensive deep packet inspection (DPI) that relies heavily on payload information where in many cases this payload can be encrypted, thus extra decryption cost is incurred. Our proposed scheme goes beyond these limitations since its operation does not depend on a-priori attack signatures and it does not consider payload information, but rather depends on per-flow meta-statistics as derived from packet header and volumetric information (i.e. counts of packets, bytes, etc.). Nonetheless, we argue that our scheme can synergistically operate with signature-based approaches on an online basis in scenarios were decryption is feasible and cost-effective.

Overall, it is our goal to develop detection techniques that are specifically targeted at the cloud and integrate with the infrastructure itself in order to, not only detect, but also provide resilience through remediation.

At the infrastructure level we consider: the elements that make up a cloud data centre, i.e. cloud nodes, which are hardware servers that run a hypervisor in order to host a number of Virtual Machines (VMs); and network infrastructure elements that provide the connectivity within the cloud and connectivity to external service users. A cloud service is provided through one or more interconnected VMs that offer access to the outside world. Cloud services can be divided into three categories based on the amount of control retained by the cloud providers. Software as a Service (SaaS) retains the most control and allows customers to access software functionality on demand, but little else. Platform as a Service (PaaS) provides customers with a choice of execution environment, development tools, etc., but not the ability to administer their own Operating System (OS). Infrastructure as a Service (IaaS) relinquishes the most control by providing customers with the

ability to install and administer their own choice of OS and install and run anything on the provided virtualised hardware; as such, IaaS clouds present the most challenges in terms of maintaining a properly functioning system. Such a system would ideally be free from malware and from vulnerabilities that could lead to an attack. It is for this reason that we focus on this type of cloud since security measures applicable to IaaS clouds will also be relevant for other cloud types.

In order to increase the resilience of cloud infrastructures we have already defined resilience architecture in our previous works [3], [4] that comprises anomaly detection, remediation and also coordination elements. However, this paper discusses two particular components within this architecture that deal with anomaly detection at the system and network level. The elements presented here form the basis in which different detection techniques can be hosted and further allow the identification and attribution of anomalies.

In this paper we discuss the detection of anomalies using a novelty detection approach that employs the one-class Support Vector Machine (SVM) algorithm and demonstrate the effectiveness of detection under different anomaly types. More specifically, we evaluate our approach using malware and Denial of Service (DoS) attacks as emulated within a controlled experimental test bed. The malware samples used are Kelihos and multiple variants of Zeus. We have selected these particular malware samples and their variants since they have been identified as posing recent and evolving threats for a range of Windows OS flavors that have already compromised more than 3.6 million machines worldwide between 2010 and 2014; mainly due to their varying and sophisticated evasion techniques, as well as their stealthy propagation1. Our contributions are as follows:

- Experiments carried out in this work are done so in the context of an overall cloud resilience architecture under the implementation of one-class Support Vector Machines (SVMs). The resulting experimental findings show that anomalies can be effectively detected online, with minimal time cost for reasonably realistic data samples per Virtual Machine (VM), using the one-class SVM approach, with an overall accuracy of greater than 90% in most cases.
- Our work is the first to explicitly address the aspect of malware detection in pragmatic cloud-oriented scenarios as performed by cloud providers, such as VM live-migration.
- We provide an online novelty detection implementation that allows the adaptive SVM-specific parameter estimation for providing better detection accuracy benefits.
- This work assesses the VM-based feature selection spectrum (i.e. system, network-based or joint datasets) with respect to the detection performance benefits on two distinct network-wise attacks (malware and DDoS) under novelty detection.

II. RELATED WORK

A. Malware and Detection Method

One of the biggest challenges within the development of resilient and secure cloud-oriented mechanisms is related to the adequate identification and detection of malware. This is due to the fact that, in the majority of cases, malware is the first point of initiation for large-scale Distributed Denial of Service (DDoS) attacks, phishing and email spamming [3], [8], mainly through the deployment of botware. Current methods of detecting attacks on cloud infrastructures or the VMs resident within them do not sufficiently address cloud specific issues. Despite the huge efforts employed in past studies regarding the behaviour of certain types of malware in the Internet, so far little has been done to tackle malware presence in clouds. In particular, the studies in aimed to adjust the performance of traditional Intrusion Detection Systems (IDS) under signature-based techniques that employ Deep Packet Inspection (DPI) on network packets. Moreover, work in studied system-related features on monitored VMs by employing Virtual Machine Introspection (VMI) methods in order to detect threats on a given VM's Operating System (OS).

Nevertheless, despite the important lessons learned from these studies they do not develop an overall online detection strategy that considers real-time measurement samples from each VM. Further, these approaches are purely signature based, and as such are not in a position to provide a robust scheme for any future threats posed by novel malware strains due to their simplistic rule-based nature. Each solution to detection is performed in an isolated manner and neglects to consider the unique topology of the cloud, which is at its heart a network of interconnected nodes, each with their own isolated execution environments. If a detection system is to perform effectively within a cloud it is required to possess the capability of communicating detected faults and challenges across the whole infrastructure, especially if it is to perform as part of a larger, autonomous and self-organising, cloud resilience system.

III. MALWARE DETECTION TECNIQUES

Since malware has different types, behaviors and different level of risk, the same detection methods and mechanisms cannot be used in all cases. It is impractical to have just one security software to efficiently handle the malwares. Hence having different detection methods for different environments becomes unavoidable. This study had focused on the most common and powerful techniques such as malicious based

detection, anomaly based. The experiment added a great value to the field of malware detection since it was able to detect many malwares which were not detectable by normal detection methods, going forward, we can clearly see that the detection process needs more computer processing power and advance techniques to make sure that the nature and behavior of malware are clear and covered from all the angles and views.

A. Malicious Based Detection

Cloud computing is recognized as an alternative to traditional information technology due to its intrinsic resource-sharing and low-maintenance characteristics. In cloud computing, the cloud service providers (CSPs), such as Amazon, are able to deliver various services to cloud users with the help of powerful data centers. By migrating the local data management systems into cloud servers, users can enjoy high-quality services and save significant investments on their local infrastructures. One of the most fundamental services offered by cloud providers is data storage.

However, it also poses a significant risk to the confidentiality of those stored files. Specifically, the cloud servers managed by cloud providers are not fully trusted by users while the data files stored in the cloud may be sensitive and confidential, such as business plans and other data. To preserve data privacy, a basic solution is to encrypt data files, and then upload the encrypted data into the cloud. Unfortunately, while uploading the data malware files can also be uploaded. To detect the malware and sending the alert message using the malicious based detection.

According to intrusion detection system, they suggested a detection system to expose intruders and attacks in a cloud computing environment based on the malicious method. This system which is used to check the malware files which are present in the cloud infrastructure. After finding the malware files it sends the alerts to the providers.

Malware detection in cloud computing presented a model to detect malware on cloud computing integrating intrusion ontology representation using malicious based methods. This model uses multiple engine services which follows a set of defined parameters and standards for web service technologies. This model is founded on analysis with specific applications residing on the client. It can enhance their performance if they are moved to the network, where instead of running complicated software on every host, it gives each process a light to enter the system files. Then it sends them to the network to be analyzed by multiple engines and then to decide whether or not they are executed according to the report of threat delivered. This model is a multi-engine based file analysis service deployed in cloud computing, via a group of protocols and standards for web services. It is used to identify the files with malicious codes through the remote analysis by multiple engines and send the alert to the service provider.

B. *Anomaly Based Detection*

Anomaly-based detection looks for unexpected or abnormal behaviour indicators, which indicate the presence of malware. In more detail, anomaly based detection creates a baseline of expected operation. After this baseline has been created, any different form of baseline is recognized as malware. We have identified that the anomaly based detection technique uses the previous knowledge of what is known as normal to find out what is malicious. A special type of anomaly based detection techniques is specification based detection. A specification based detection uses set of rules to determine what is considered as normal, with the purpose of making a decision about the maliciousness of the program that breaches the rule set. The basic limitation of the specification based system technique is the difficulty to correctly determine the program or system behaviour.

IV. CONCLUSION

In this paper we introduce an online anomaly detection method that can be applied at the hypervisor level of the cloud infrastructure. The method is embodied by a resilience architecture that was initially defined in [4], further explored and which comprises the System Analysis Engine (SAE) and Network Analysis Engine (NAE) components. These exist as sub modules of the architecture's Cloud Resilience Managers (CRMs), which perform detection at the end-system, and in the network respectively. Our evaluation focused on detecting anomalies as produced by a variety of malware strains from the Kelihos and Zeus samples under the formulation of a novelty detector that employs the one-class Support Vector Machine (SVM) algorithm. Moreover, in order to empower the generic properties of our detection approach we also assess the detection of anomalies by the SAE and NAE during the onset of DoS attacks.

Overall, this work performs online anomaly detection under two pragmatic cloud scenarios, based on suggestions by cloud operators, which emulate "static" detection as well as detection under the scenario of VM "live" migration. The results obtained by strictly utilizing system-level data in our SAE detection, which was supported by an automatic SVM-specific parameter selection process, have shown excellent detection for all samples of malware under a variety of conditions (i.e. static and migration

analysis) with an overall detection accuracy rate of well above 95%. Hence, we have demonstrated that the extracted features for classifier training were appropriate for our purposes and aided towards the detection of the investigated anomalies under minimal time cost throughout the training and testing phase. Nonetheless, in order to further the investigation, this feature set can easily be expanded to include statistics derived from CPU usage and a deeper introspection of process handles, which could be beneficial for the detection of– highly –stealthy malware. However, the consideration of new features would naturally invoke a computational trade- off, since deeper introspection will require more system resources.

The results derived from the experiments based on network-level detection of DoS attacks have also justified that the network features used were sufficient for the detection of such challenges, since the detection accuracy rate also reached well above 90%. However, when attempting to detect the examined Zeus and Kelihos malware samples using a strictly network-based feature set the gained results were inconclusive with low detection accuracy rates and un- acceptable recall. In parallel, we have also observed minimal improvement in the evaluation metrics when considering a joint dataset, which was composed of both end-system and network level data. Hence, despite experiencing good results from the detection conducted using system-based features in the SAE we concluded that is not possible to improve the results obtained from the NAE through the combination of feature sets. Therefore, we demonstrate that by extending the feature set explicitly under the one-class SVM formulation would not necessarily lead to higher detection accuracy rates. However, as we show in our other work using the Ensemble Empirical Mode Decomposition (E-EMD) algorithm [8], a joint dataset could lead to good detection accuracy levels, thus we argue that the effectiveness of a feature set is strongly related with the exact mathematical formulation of a given detection algorithm.

In general, the detection approach presented in this paper is designed to be adaptive and respond to new threats and challenges online and in real time under minimal computational cost. Given the promising results presented through this work, we argue that our novel solution can overcome the commonly used signature-based intrusion detection solutions that are currently governing the domain of cloud security and further benefit cloud data center operations where security and resilience are of paramount importance.

ACKNOWLEDGEMENT

This work was supported by **SK. Sofia**, M.Tech. Assistant Professor, Department of Computer Science and Engineering and **D. Srinivasulu**, M.Tech.,(Ph.D). Professor, H.O.D of Computer Science and Engineering by their valuable guidance, constant encouragement, constructive criticism and keen interest.

REFERENCES

[1] A. Marnerides, C. James, A. Schaeffer, S. Sait, A. Mauthe, and H. Murthy, "Multi-level network resilience: Traffic anal- ysis, anomaly detection and simulation," ICTACT Journal on Communication Technology, Special Issue on Next Generation Wireless Networks and Applications, vol. 2, pp. 345–356, June 2011.

[2] J. P. G. Sterbenz, D. Hutchison, E. K. Çetinkaya, A. Jabbar, J. P. Rohrer, M. Schö ller, and P. Smith, "Resilience and survivability in communication networks: Strategies, principles, and survey of disciplines," Comput. Netw., vol. 54, no. 8, pp. 1245–1265, Jun. 2010. [Online]. Available:http://dx.doi.org/10.1016/j.comnet.2010.03.005

[3] A. K. Marnerides, M. R. Watson, N. Shirazi, A. Mauthe, and D. Hutchison, "Malware analysis in cloud computing: Network and system characteristics," IEEE Globecom 2013, 2013.

[4] M. R. Watson, N. Shirazi, A. K. Marnerides, A. Mauthe, and D. Hutchison, "Towards a distributed, self-organizing approach to malware detection in cloud computing," 7th IFIP/IFISC IWSOS, 2013.

[5] M. Garnaeva, "Kelihos/Hlux Botnet Returns with New Tech- niques." Securelist http://www.securelist.com/en/blog/655/ Kelihos Hlux botnet returns with new techniques.

[6] H. Binsalleeh, T. Ormerod, A. Boukhtouta, P. Sinha, A. Youssef, M. Debbabi, and L. Wang, "On the analysis of the zeus botnet crimeware toolkit," in Privacy Security and Trust (PST), 2010 Eighth Annual International Conference on, Aug 2010, pp. 31–38.

[7] T. Brewster, "GameOver Zeus returns: thieving malware rises a month after police actions," Guardian Newspaper, 11, July, 2014, http://www.theguardian.com/technology/2014l/

11/gameover-zeus-criminal-malware-police hacking.

[8] A. K. Marnerides, P. Spachos, P. Chatzimisios, and A. Mauthe, "Malware detection in the cloud under ensemble empirical model decomposition," in Proceedings of the 6th IEEE International Conference on Networking and Computing, 2015.

[9] L. Kaufman, "Data security in the world of cloud computing," Security Privacy, IEEE, vol. 7, no. 4, pp. 61–64, July 2009.

[10] M. Christodorescu, R. Sailer, D. L. Schales, D. Sgandurra, and D. Zamboni, "Cloud security is not (just) virtualization security: A short paper," in Proceedings of the 2009 ACM Workshop on Cloud Computing Security, ser. CCSW '09. New York, NY, USA: ACM, 2009, pp. 97–102. [Online]. Available: http://doi.acm.org/10.1145/1655008.1655022

[11] N. Gruschka and M. Jensen, "Attack surfaces: A taxonomy for attacks on cloud services," in Cloud Computing (CLOUD), 2010 IEEE 3rd International Conference on, July 2010, pp. 276–279.

[12] Y. Chen, V. Paxson, and R. H. Katz, "Whats new about cloud computing security?" EECS Department, University of California, Berkeley, Tech. Rep. UCB/EECS-2010-5, Jan 2010. [Online]. Available:http://www.eecs.berkeley.edu/Pubs/TchRpts/2010/ EECS-2010-5.html