A Survey Over the Various Malware Detection Techniques used in Cloud Computing

Nancy
Student M.tech
Department of CSE
(UIT) RGPV
Bhopal, India

Dr. Sanjay Silakari
HOD Department of CSE
University Institute of Technology
(UIT) RGPV
Bhopal, India

Uday Chourasia
Assistant Professor
Department of CSE
University Institute of Technology
(UIT) RGPV
Bhopal, India

***Abstract* – Cloud computing provides various on demand services to the user. These services provide a flexible and easy mechanism to access various web applications. But in cloud computing services like infrastructure as a service, platform as a service and many more are provided to the user. But various malicious attacks can be occurs in these process, a review over the various techniques like malware detection in virtualization, system call hashing, external host monitoring etc., which used for malware detection is presented in this paper, which provides a brief overview on the techniques used for malware detection in cloud computing.**

Keywords: - Cloud Computing, Malware, malware detection, Virtualization, IDS Signature.

I INTRODUCTION

Cloud computing is a recent technological development in the computing field in which mainly focused on designing of services which can be provided to the users in same way as the basic utilities like food, water, gas, electricity and telephony. In this technology services are developed and hosted on the cloud (a network designed for storing data called datacenter) and then these services are offered to users always whenever they want to use. The cloud hosted services are delivered to users in pay-per-use, multi-tenancy, scalability, self-operability, on-demand and cost effective manner. Cloud computing is become popular because of above mention services offered to users. All the services offered by servers to users are provided by cloud service provider (CSP) which is working same as the ISP (Internet service provider) in the internet computing. In the internet technology some innovative development in virtualization and distributed computing and accessing of high speed network with low cost attract focus of users toward this technology. This technology is designed with the new concept of services provisioning to users without purchasing of these services and stored on their local memory.

Cloud computing is the process in which various type of services , resources and data is provided to the user on on-demand basis. In this an on-demand service to access various configurable resources is provided to the user. A description of the services provided by the cloud is presented in below Figure 1.1. there are services like

TYPES OF SERVICE IN CLOUD

(Software As A service) SaaS:- In this various software for user are provided on, on-demand basis. In this vendors like Google, amazon, Microsoft etc. are providing software for the user on on-demand basis.

Platform as a service (PaaS): a development platform to the user is provided to develop their applications, in that process there is no overhead for the clients system is generated and all the work is performed at cloud service provider’s server. Thus it provides a cost effective mean to develop program.

Infrastructure as a service (IaaS):- infrastructure to perform tasks are provided to the user in that virtual machines are provided by the different vendors, these vendors offers on-demand infrastructure to the user to execute their applications and tasks.

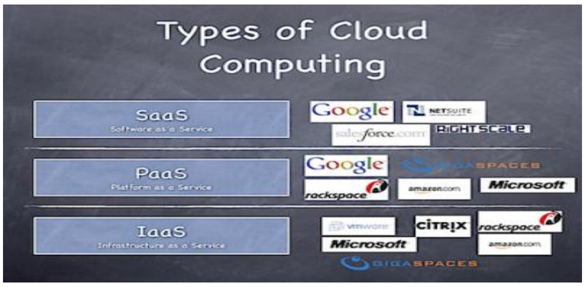

Figure 1.1: A diagram for various types of cloud services

A description for the services and service providers is presented in the above Figure 1.1, that shows vendors like Google, Microsoft, Rackspace etc. are provide services to the user just like an online office software is provided by Google, Microsoft that allow user to perform all the editing task and a cloud storage also provided to save all the updated files in the user accounts. That it offers various on-demand services to the user to perform all their tasks

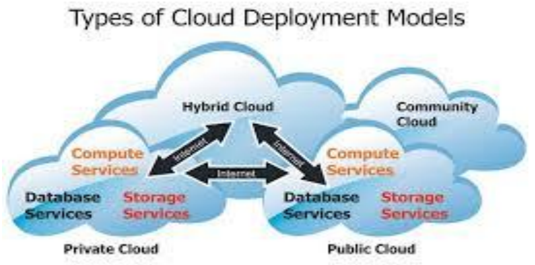

Figure 1.2: Cloud deployment model

Types Of Cloud Deployment Models
There are various type of cloud models are present in nature, which defines accessibility to the cloud.

Public cloud:- in this type of cloud architecture an open access to all the user in the world is provided to access that cloud and use applications and perform tasks. For that user need to create an account to access that cloud services. Like Google provides various cloud services to the user.

Private cloud:- that type of cloud model scope is restricted to the only for the some users, either with in a campus or organization. That provides services to the user only in that restricted area.

Community cloud:- in that model cloud services are provided with in a small region and not to allow any outsider from the community to access these services.

Hybrid cloud:- In this a combination of all the cloud models is presented which provides various on-demand services to the user.

A description for the various cloud deployment models is presented above which shows a over view of various type of cloud models.

BENEFITS OF CLOUD COMPUTING

Cloud computing provided so many services to their users in which some of the very popular services are listed below:-
On-demand self-service
A consumer can unilaterally provision computing capabilities, such as server time and network storage, as needed automatically without requiring human interaction with each service provider.
Broad network access
Capabilities are available over the network and accessed through standard mechanisms that promote use by heterogeneous thin or thick client platforms (e.g., mobile phones, tablets, laptops, and workstations).
Resource pooling
The provider's computing resources are pooled to serve multiple consumers using a multi-tenant model, with different physical and virtual resources dynamically assigned and reassigned according to consumer demand.

Rapid elasticity (Scalability)
Capabilities can be elastically provisioned and released, in some cases automatically, to scale rapidly outward and inward commensurate with demand. To the consumer, the capabilities available for provisioning often appear unlimited and can be appropriated in any quantity at any time.

Measured service
Cloud systems automatically control and optimize resource use by leveraging a metering capability at some level of abstraction appropriate to the type of service (e.g., storage, processing, bandwidth, and active user accounts). Resource usage can be monitored, controlled, and reported, providing transparency for both the provider and consumer of the utilized service.

Agility
Improves the capability of the resource provisioning to the users in such a way so that they just feel like to work with a separate resource using the concept of virtualization.

Pay-Per-Use Cost
Reductions claimed by cloud providers. A public-cloud delivery model converts capital expenditure to operational expenditure. This purportedly lowers barriers to entry, as infrastructure is typically provided by a third party and does not need to be purchased for one-time or infrequent intensive computing tasks. Pricing on a utility computing basis is fine-grained, with usage-based options and fewer IT skills are required for implementation (in-house). The e-FISCAL project's state-of-the-art repository contains several articles looking into cost aspects in more detail, most of them concluding that costs savings depend on the type of activities supported and the type of infrastructure available in-house.

Device and location independence
Enable users to access systems using a web browser regardless of their location or what device they use (e.g., PC, mobile phone). As infrastructure is off-site (typically provided by a third-party) and accessed via the Internet, users can connect from anywhere.

Easy Maintenance
Cloud computing applications is easier, because they do not need to be installed on each user's computer and can be accessed from different places.

i) Multitenancy enables sharing of resources and costs across a large pool of users thus allowing for

ii) Centralization of infrastructure in locations with lower costs (such as real estate, electricity, etc.)

iii) Peak-load capacity increases (users need not engineer for highest possible load-levels)

Utilization and Efficiency:- improvements for systems that are often only 10–20% utilized.

Reliability
Improves with the use of multiple redundant sites, which makes well-designed cloud computing suitable for business continuity and disaster recovery.

Application programming interface (API)
Accessibility to software that enables machines to interact with cloud software in the same way that a traditional user interface (e.g., a computer desktop) facilitates interaction between humans and computers. Cloud computing systems typically use Representational State Transfer (REST)-based APIs.
Productivity

May be increased when multiple users can work on the same data simultaneously, rather than waiting for it to be saved and emailed. Time may be saved as information does not need to be re-entered when fields are matched, nor do users need to install application software upgrades to their computer

MALWARE THREATS IN CLOUD COMPUTING

Over the last decade, our society has become technology dependent. People rely on computer networks to receive news, stock prices, email and online shopping. The integrity and availability of all these systems need to be defended against a number of threats. Amateur hackers, rival corporations, terrorists and even foreign governments have the motive and capability to carry out sophisticated attacks against computer systems. Therefore, the field of information security has become vitally important to the safety and economic well-being of society as a whole.

In this struggle to secure our stored data and the systems, IDS can prove to be an invaluable tool, where its goal is to perform early detection of malicious activity and possibly prevent more serious damage to the protected systems. By using IDS, one can potentially identify an attack and notify appropriate personnel immediately or prevent it from succeeding, so that the threat can be contained. IDS can also be a very useful tool for recording forensic evidence that may be used in legal proceedings if the perpetrator of a criminal breach is prosecuted.

A demonstration for the various type of malware is shown in Figure 1.3, which shows a description over the different types of malware in the system. In that case malware of widows, Linux and some others systems is presented.

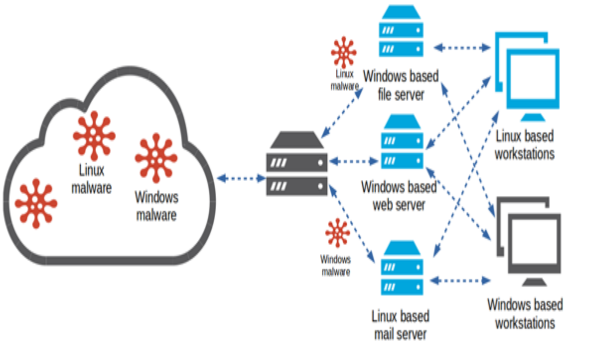

Figure 1.1: A Demonstration of the Malware in Cloud Computing

A description of the process of malware detection in cloud computing is presented in Figure1.4, which a network where various suspicious files are presents. Then a host agent based mechanism is used to detect these suspicious files and analysis by the forensics archive is presented which analyze the content and provide a threat report for these files.

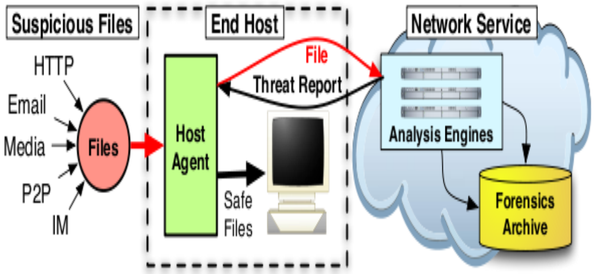

Figure 1.2: Intrusion Detection in Cloud Computing

Malware is the combination of two words called malicious and software, thus malwares are the software which puts malicious and harmful effect on the software, operating system or other components. A survey over various malwares and malware detection techniques [13] is presented which provides a description of the various types of attacks and classes of malwares, like network based malware attacks, ordinary malware attacks etc. in network malware, malwares like spywares are used to put harmful effect over users machine, in ordinary malware. malwares like autorun.inf system.inf etc. are used to put harmful effect over users machine. There are various techniques presents by the user to detect malwares in the system. Like in [8] a hybrid signature call graph scheme is used to provide a malware detection for the various types of malware attacks, in [11] a semantic aware malware detection technique is presented which uses the semantics of the files to detect malware. In [9] metamorphic malware detection technique to detect various type of malwares is presented. In [10] a file content and relation of these files is used to provide a malware detection the cloud computing which provides a secure mechanism to transfer data in cloud. In [12] a flow for the malware detection technique is presented, and security mechanism for the [7] which takes a multi agents architecture to provide security for the various type of data is presented.

FURTHER THIS PAPER ORGANIZES AS FOLLOWS

II Literature Review, in this section a description over the various malware detection techniques is presented. III Conclusion.

II LITERATURE REVIEW

A review over the various techniques which used for malware detection in cloud computing is presented in this section.

In [1] a malware detection technique to detect malware and rootkit is presented. That Takes a system call monitoring and system call hashing together and a support vector machine based external host monitoring system is also used. In monitoring system call all the system calls triggered by the users, are monitored over the parameter before execution. In system call hashing, all the stored monitored system called copies are checked before installation. Then a support vector machine based system is used which classify all the malware and rootkit attacks in virtualize cloud system. But that technique suffers in accuracy thus a new technique to provide accurate results for intrusion detection.

In [2] a support vector machine based technique is presented to detect intrusion in the cloud computing architecture. In this a support vector machine based monitoring system is used at hypervisor level to detect malware in cloud computing system. In cloud virtualization there is various type of threats can be found which requires an enhanced functionality to deal with such cloud computing threats. Thus an enhanced mechanism is required to provide an accurate result for this problem.

In [3] a description over the various malware detection techniques which used for the intrusion detection is presented. In that various machine learning techniques can be used to provide a detection mechanism for cloud computing. In that way it requires an enhanced technique to pride accurate intrusion detection for such techniques.

In [4] a malware detection technique for virtualize cloud environment is presented. There are various system resilience related risks are occurred in these virtualizations techniques. New technique is required which can deal with the issues related to the risk in the programming. In that technique a NAE (Network Analysis engine) and system analysis engine is used to deal with such issues. But these techniques are not efficient to deal with such issues

In [5] a new malware detection software technique is presented, which provides an enhanced functionality for detections of malware and enhanced forensics capability and improved deploy ability for the various software. But that technique this technique is still require an improvement in malware detection to provide an enhanced functionality for the better intrusion detection is required.

In [6] a review over the various malware detection techniques is presented, which provides a brief overview over the malware detection and malware detection techniques. There are various techniques like host based technique, malware detection in virtualized cloud scenario; malware detection for guest user is generally used to provide a malware detection system.

In [14] an ontology based malware detection technique for the cloud data is presented. In that a K-mean clustering technique is used to classify various malware attacks into different classes. On the basis of these classes a malware detection to detect malicious software is performed. In cloud user uses internet to access various applications, which can be induce any malicious software into user's machine. Thus a malware detection using these definitions is performed to detect threats in that application.

In [15] an analysis system for the IDS signatures which contains malware in a live operational scenario is presented. In that process 200 such security incidents are detected and an analysis for the detection of such events and check the vulnerability these events contains. In that a single data source is not efficient to provide better performance for the detection of such events a multiple data source based system is used to provide proper detection for these sources. Then a list of various snort signatures is presented which shows that 165 signatures provide better detection for the malware and provide better detection the malware. A blacklist and vulnerability report for that is presented.

In [16] an evaluation for the various classification techniques which used for the malware detection is presented. In these technique an OpCode N-gram based pattern are used to detect the malicious codes in the given program or applications. In that technique byte n-gram patterns are used to inspect suspicious files and provide detection for malicious codes and that provides user a way to analyze and detect malicious software in that data. In that evaluation for 1-gram, 2-gram, and n-gram based technique is provided.

Comparison for the various techniques which used for the malware detection in cloud computing

Technique	Advantage	Disadvantage
IDS Signature Based Technique [15]	Multiple sources are used to detect malware in the data	It not provide proper malware detection for the single source
Ontology based malware detection technique [14]	A definition based malware detection is provided. Prior information about the malwares is provided.	There is no updating mechanism is provided to update virus definition.
Malware detection in virtualize cloud environment [4]	Network analysis engine and system analysis engine based technique is provided to malware detection in virtualize cloud environment.	But it not able converges all the defects in the virtualize cloud environment.
Support vector machine based malware detection technique	Support vector machine a machine learning classification algorithm which provide a proper classification for the various malware attacks.	Support vector machine suffers from various inherent defects like unclassifiable regions and etc.
System call monitoring and system call hashing based technique	It provides enhanced functionality to provide security for the data in cloud storage.	It uses support vector machine for external monitoring. That suffers defects of the support vector machine.

TABLE 1 Comparison for the various techniques

III CONCLUSION

In current scenario there are various technique are used for intrusion detection in cloud computing. There are various on-demand services are provided to the user, these techniques requires an enhanced functionality to deal with such issues. A brief description over the techniques which used for malware detection in cloud computing is presented in this paper. There are techniques like n-gram based pattern detection, IDS signature, malware detection in virtualization etc. are used. An enhanced technique for the future work is proposed to provide better malware detection in cloud computing scenario.

IV REFERENCES

[1] Thu Yein Win, Huaglory Tianfield, Quentin Mair “Detection of Malware and Kernel-level Rootkits in Cloud Computing Environments” IEEE, 2015.

[2] Michael R. Watson, Noor-ul-hassan Shirazi, Angelos K. Marnerides, Andreas Mauthe and David Hutchison “Malware Detection in Cloud Computing Infrastructures” IEEE, 2015.

[3] Xiaoguang Han, Jigang Sun, Wu Qu3, Xuanxia Yao “Distributed Malware Detection based on Binary File Features in Cloud Computing Environment” IEEE, 2014

[4] Angelos K. Marnerides, Michael R. Watson, Noorulhassan Shirazi, Andreas Mauthe, and David Hutchison “Malware Analysis in Cloud Computing: Network and System Characteristics” 2013.

[5] Safaa Salam Hatem, Maged H. wafy, Mahmoud M. El-Khouly “Malware Detection in Cloud Computing” IJACSA, 2014.

[6] Imtithal A. Saeed, Ali Selamat, Ali M. A. Abuagoub “A Survey on Malware and Malware Detection Systems” IJCA, 2013.

[7] Pinz, C.I., et al., Improving the security level of the FUSION@ multi-agent architecture. Expert Syst. Appl., 2012.

[8] Ammar Ahmed E. Elhadi, M.A. Maarof, and A.H. Osman, Malware Detection Based on Hybrid Signature Behaviour Application Programming Interface Call Graph. American Journal of Applied Sciences, 2012.

[9] Kevadia Kaushal, P.S., Nilesh Prajapati, Metamorphic Malware Detection Using Statistical Analysis International Journal of Soft Computing and Engineering (IJSCE), 2012.

[10] Yanfang Ye, T.L., Shenghuo Zhu,Weiwei Zhuang,Egemen Tas,Umesh Gupta,Melih Abdulhayoglu, Combining file content and file relations for cloud based malware detection, in Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining. 2011, ACM: San Diego, California, USA. p. 222-230.

[11] Christodorescu, M., et al., Semantics-Aware Malware Detection, in Proceedings of the 2005 IEEE Symposium on Security and Privacy. 2005, IEEE Computer Society.

[12] Yin, H., et al., Panorama: capturing system-wide information flow for malware detection and analysis, in Proceedings of the 14th ACM conference on Computer and communications security. 2007, ACM: Alexandria, Virginia, USA. p. 116-127.

[13] Vinod, P., et al., Survey on Malware Detection Methods. 2009.

[14] Cristian Adrián Martínez, Gustavo Isaza Echeverri, Andrés G. Castillo Sanz “Malware Detection based on Cloud Computing integrating Intrusion Ontology representation” IEEE, 2010.

[15] Elias Raftopoulos and Xenofontas Dimitropoulos “Aquality metric for IDS signatures: in the wild the size matters” EURASIP Journal on Information Security 2013.

[16] Asaf Shabtai, Robert Moskovitch, Clint Feher, Shlomi Dolev, and Yuval Elovici “Detecting unknown malicious code by applying classification techniques on Op-Code patterns” Shabtai et al. Security Informatics 2012