

An Integrated Intrusion Detection and Diagnosis for Clouds

Junaid Arshad
School of Computing
University of Leeds, LS2 9JT
ja@comp.leeds.ac.uk

Abstract

Improving dependability of Cloud computing is critical to the realization of its potential as a technology to facilitate the development of large-scale, flexible computing infrastructures, available on-demand to meet computation requirements of compute intensive workloads. However, unique characteristics such as; Diversity, Flexibility and Mobility inherited by such systems require dedicated efforts to address emanating challenges. The emphasis of our research is to investigate these novel challenges with respect to security in general and intrusion detection and diagnosis in particular. In this paper, we describe our overall research objectives and summarize our efforts to quantify security for compute intensive workloads in clouds. We also demonstrate application of our security metrics to resource acquisition process for clouds which represents initiatory step for our research.

1. Introduction

The comprehensive adoption of *Virtualization* [13], has elicited the emergence of Cloud Computing as a technology with promising prospects to facilitate the development of large scale, flexible computing infrastructures, available on-demand to meet the computational requirements of compute intensive workloads. With compelling characteristics such as; Live Migration, Isolation, Customization and Portability, Cloud Computing has witnessed widespread acceptance thereby increasing the value attached with such infrastructures. The virtual machine technology [6] has profound role in it. Amazon [2], Google [5] and GoGrid [4] represent some of commercial Cloud computing initiatives whereas Nimbus [10] and OpenNebula [11] represent academic efforts to establish a Cloud.

However, in order to stimulate extensive adoption of Clouds, there is need to develop mechanisms focused at improving the dependability of such infrastructures. Related to this, Clouds inherit unique characteristics such as Diversity, Mobility and Flexibility from virtual machines which require dedicated efforts to address the

challenges emanating due to these characteristics. The emphasis of our research is to investigate these challenges with respect to security in general and intrusion detection and diagnosis in particular. This paper describes the overall objectives of our research along with the description of methodology we envisage to adopt to achieve those objectives. In order to facilitate provision of customized services, we propose to incorporate application specific security requirements using service level agreements. In this regard, we summarize our efforts for the quantification of security for Clouds from the perspective of compute intensive workloads. Finally, we demonstrate the usefulness of our security metrics by presenting their application to resource acquisition process for clouds.

2. Cloud Computing

Cloud Computing has been defined in different ways by different sources as described by [12]. However, for the purpose of this research, we define Clouds as under:

A Cloud is a high performance computing infrastructure based on system virtual machines to provide on-demand resource provision according to the service level agreements established between a consumer and a resource provider.

A system virtual machine, as described in the above definition, serves as the fundamental unit for the realization of a Cloud infrastructure and emulates a complete and independent operating environment. As described in the above definition, Cloud Computing involves on-demand provision of virtualized resources based on service level agreements (SLA) thereby facilitating the user to acquire resources at runtime by defining the specifications of the resource required. The user and the resource provider are expected to negotiate the terms and conditions of the resource usage through service level agreements so as to protect the quality of service being committed at resource acquisition stage. These agreements usually revolve around the quality of service (QoS) that a service provider provides to a service requestor and usually include only resource specific requirements such as CPU, Memory and Bandwidth. However, for the purpose of this research,

we envisage to add security requirements as part of a SLA along with the resource specific requirements.

3. Intrusion Detection and Diagnosis

Intrusion detection and diagnosis is a well established research domain focused at improving the security of a system. Historically, an intrusion detection system (IDS) strives to facilitate a system administrator by raising an alert whenever it detects an intrusion. Contemporary IDSs can be broadly classified as host or network based with respect to their location. As suggested by their names, a host based intrusion detection system is located on the host being monitored and therefore has the benefit of maximum visibility of the monitored system. However, it has the disadvantage of being prone to get compromised in the event of a successful intrusion taking control of the monitored system. A network based IDS on the other hand, is usually installed at the edge of a network and has the advantage of being isolated from the monitored system. However, it has the disadvantage of reduced visibility of the monitored system. Leveraging the isolation provided by virtual machines, Hypervisor based IDS have been proposed [8,9] which combines the benefits of both host and network based IDS thereby improving the security of the monitored system.

However, virtual machines introduce new challenges due to characteristics such as; Mobility, Flexibility and Diversity of the monitored systems. We describe each of these as under;

Mobility: With respect to Mobility, virtual machines allow migration of an entire virtual machine from one physical machine to another with minimal downtime i.e. *Live Migration*. This introduces challenges for intrusion detection such as; How does an IDS comprehend when a monitored virtual machine is migrated to another physical domain. How does an IDS respond to a virtual machine being added to the physical node monitored by it. How does an IDS comprehend with the reason for migration etc.

Flexibility: Virtualization facilitates on-demand creation and deletion of virtual machines to existing physical resources thereby augmenting the flexibility of the infrastructure. An intrusion detection system, therefore, will have to take into the dynamic nature of monitored system including issues such as on-demand creation and deletion of virtual machines.

Diversity: Virtualization supports the ability to host multiple different computing environments on a single physical resource whereby a detection and diagnosis system will have to monitor multiple virtual machines with possibly varying security requirements. This requires identification of the fact that a particular malicious attempt can have a different degree of impact

for different applications. We define this to be the *Level of Severity (LoS)* of a security breach for a particular application. There can be different implications of this concept such as invoking an appropriate recovery mechanism based on the LoS of a particular attack.

From the perspective of intrusion diagnosis, it is traditionally defined as the process to investigate the cause of a successful intrusion. However, for the purpose of this research, we define diagnosis to be the process of evaluating the severity of an intrusion for a monitored host, a virtual machine in our case. In contemporary intrusion detection systems, this function is often rendered responsibility of a system administrator who is supposed to use his technical expertise and offline analysis to decide the severity of an alert raised by the IDS. The severity of an intrusion has significant implications as described in the previous section.

4. An Integrated Approach for Intrusion Detection and Diagnosis

From the discussion in previous section, we hold that virtualized computing systems such as Clouds introduce novel challenges to develop efficient intrusion detection and diagnosis systems. For the purpose of our research, we propose to address the challenges raised due to flexible and diverse nature of virtualized systems. In particular, we envisage developing an abstract model for intrusion detection and diagnosis system for virtualized systems in general and Clouds in particular. As part of this research we intend to use dynamic reconfiguration to address the issues due to flexibility of the system and use mathematical modelling to carry out the severity analysis for intrusions detected by the detection system. Figure 1 describes a basic architecture of the proposed system.

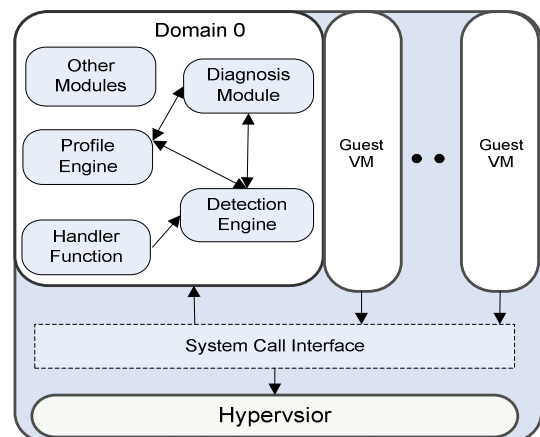


Figure 1: The basic architecture of proposed system

Related to this, intrusion detection systems usually require human intervention for effective intrusion

analysis and response, isolating it from other security systems. We believe that an integrated approach for intrusion detection and diagnosis will facilitate provision of dependable virtualized computing systems. As compared to the existing, isolated detection and diagnosis systems, an integrated approach has the benefit of reduced response times and assurance for end-to-end operation by virtue of automation. This improves ability of the system to minimize the time required to efficiently respond to an intrusion. Consequently, reducing the disruption in the service being delivered which is an important metric to evaluate the dependability of a system.

As described in figure 1, the system is envisaged to be developed as part of the domain 0, the most privileged domain, of a virtualized resource. This choice enables us to leverage the strong isolation property of virtualization. And, consequently improves the dependability of the system itself in the event of a monitored virtual machine getting compromised. Another motivation behind this choice is to benefit from the principle of recursive construction for fault tolerant components [3] which improves the dependability of the computing infrastructure as a whole. From figure 1, the guest virtual machines use system call interface to interact with the hypervisor which triggers a handler function in the domain 0 to activate the proposed system. Furthermore, both detection and diagnosis modules are envisaged to use a profile engine which is supposed to generate VM-specific profiles to assist the detection and diagnosis processes. We aim to develop the policy and detection engines to comprehend with the flexibility of the computing infrastructure as a whole by applying dynamic reconfiguration. Dynamic reconfiguration is envisaged to facilitate the policy and detection engine to comprehend with the addition of new VM by instating new security profile and associating a detection engine at runtime. With regards to evaluating the severity of intrusions, we envisage to develop a mathematical model which can take into account the security requirements of a workload along with other parameters.

5. Quantification of Security

In order to facilitate our objective of evaluating the severity of an intrusion for a monitored virtual machine, we propose to gather security requirements for virtual machines using SLA during resource acquisition process. This required quantifying security into a set of requirements so as to facilitate negotiation and facilitate assurances to fulfil those requirements. As part of this exercise, we have formulated seven security requirements summarized in table 1 which are in accordance with the three attributes of security i.e. Integrity, Availability and Confidentiality as described

by [1]. These are envisaged to be specified as part of a resource request along with their priorities by the consumer of the service as explained in the next section.

Security Attributes	Requirements
Integrity	Workload State Integrity
	Guest OS Integrity
Availability	Zombie Protection
	Denial of Service Attacks
	Malicious Resource Exhaustion
	Platform Attacks
Confidentiality	Backdoor Protection

Table 1: Proposed Security Requirements

In order to demonstrate the applicability of our research, we have implemented our security requirements to be incorporated in the resource acquisition process. For this purpose, we have used Haizea [7], an open source resource manager for Clouds, mainly due to its ability to work with two of the established open source Cloud infrastructures i.e. Nimbus and OpenNebula. The fundamental resource provisioning abstraction in Haizea is a *lease* which describes the type and quantity of resources required such as; 10 nodes with 1GB memory and 1GB bandwidth. An important parameter of a lease is the start time which defines the type of lease i.e. Best Effort (BE) or Advanced Reservation (AR) Lease.

Figure 2 describes two of the experiments we conducted to demonstrate the applicability of our security metrics. In order to explain these conflicts, we take an example of a scenario with three leases where lease1 and lease2 represent workloads from particle physics applications. These workloads are assumed to have highest priority for DoS protection as part of the security requirements. However, lease3 represents an e-social science workload with highest priority for backdoor protection to ensure confidentiality of data being used in the computation. We also assume that the resource pool for this simulation has resources to accommodate only one of the two competing leases.

This, therefore, generates a conflict between lease1 and lease2. The execution pattern in this case depends on the type of lease and the request time for each lease. In figure 2 (a), we describe the case which simulates a conflict between *lease1* (BE) and *lease2* (AR) i.e. the two leases with similar security requirements. As an

AR lease has priority over BE lease in our simulations, lease1 is suspended to execute lease2 on the same resources. An interesting observation here is that even though the request for lease2 arrives after lease1 has started its execution, lease1 is pre-empted to facilitate execution of lease2. The reason for this arrangement is the configuration of our simulation environment which prioritizes AR leases over BE leases. Therefore, during the execution of lease2, lease1 is suspended which causes a delay in the completion time of lease1. However, lease3 which represents a BE lease with different security requirements is executed uninterrupted.

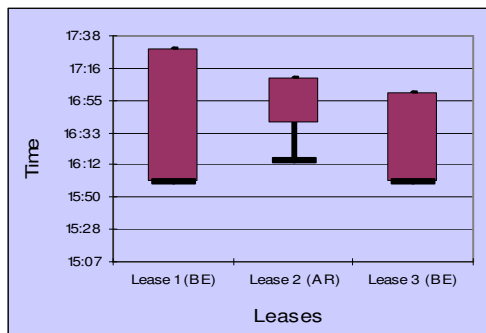


Figure 2(a): Conflict between AR and BE

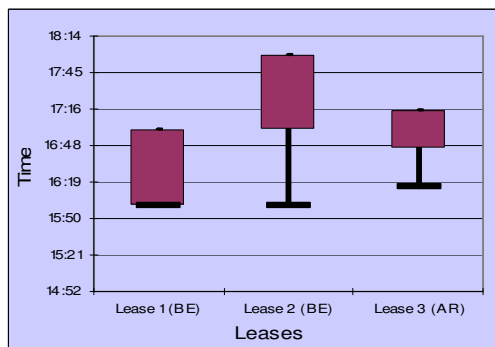


Figure 2(b): Conflict between two BE leases

In the second experiment, we simulate a scenario where the leases with conflicting security requirements are requested at the same time and share same type i.e. both are BE leases. The result of this experiment is shown in figure 2(b). As both the leases have same priority being BE leases, the resource manager applies First Come First Server (FCFS) rule to resolve this conflict. Consequently, lease2 is made to wait until lease1 finishes its execution. The AR lease i.e. lease3 in this case runs smoothly as it is not involved in the conflict.

6. Conclusions and Future Directions

We hold that unique characteristics of virtual machines raise novel challenges which need dedicated efforts whilst developing security solutions for virtualized

computing systems. The overall objective of our research is to improve the dependability of Cloud computing infrastructures with special emphasis on security aspect of dependability. In this paper, we have presented our research objectives to achieve integrated intrusion detection and diagnosis system for virtualized computing systems in general and Clouds in particular. In accordance with our research objectives, we envisage developing a comprehensive abstract model for intrusion detection and diagnosis which is able to address the challenges identified in this paper. To the best of our knowledge, we are the first to carryout such research for virtualized computing systems. This will be followed by an evaluation of the model to evaluate the effectiveness of the approach. We also presented our initial efforts to quantify security and demonstrated their application in resource acquisition process. We aim to utilize this research for effective intrusion diagnosis whereby security requirements of a workload can contribute to calculate level of severity of a particular intrusion. To the best of our knowledge, we also pioneer research efforts in this regard.

References

- [1] Algirdas Avi zienis, Jean-Claude Laprie, Brian Randell and Carl Landwehr: Basic Concepts and Taxonomy of Dependable and Secure Computing, IEEE Transaction on Dependable And Secure Computing, Vol. 1, No. 1, January-March 2004
- [2] Amazon Elastic Computing Cloud Available at: aws.amazon.com/ec2
- [3] B. Randell; Recursively Structured Distributed Computing Systems. In Software System Design Methods, Skwirzynski, J.K. (ed), pp 35-52 NATO ASI Series: Series F, Computer and Systems Sciences, 22 Springer-Verlag, 1986, ISBN 0-387-16765-X
- [4] GoGrid: Scalable Load-Balanced Windows and Linux Cloud-Server Hosting. Available at: <http://www.gogrid.com/>
- [5] Google Cloud. Available at: www.googlecloud.com
- [6] Goldberg, R.P.: A survey of virtual machine research. IEEE Computer. 7, 34-45 (1974)
- [7] Haizea: An open source VM-based lease manager. Available at: <http://haizea.cs.uchicago.edu/index.html>
- [8] Marcos Laureano, Carlos Maziero and Edgard Jamhour: Intrusion Detection in Virtual Machine Environments in the Proceedings of the 30th EUROMICRO Conference (EUROMICRO'04)
- [9] Lionel Litty: Hypervisor-based Intrusion Detection MS Thesis, University of Toronto 2005. Available at: www.cs.toronto.edu/~llitty/papers/MS.pdf
- [10] Nimbus. Available at: www.workspace.globus.org
- [11] OpenNebula Project. <http://www.opennebula.org>
- [12] A. Weiss: Computing in the Clouds, netWorker, Volume 11, Issue 4, Pg 16-25, Dec. 2007
- [13] Ravi Subramaniam; [The Siamese Twins of IT Infrastructure: Grid and Virtualization](#), Open Grid Forum 2007.