A Secure Access Control Architecture for Multi-Tenancy Cloud Environments

Ronald Beaubrun
Department of Computer Science and Software Engineering
Laval University
Quebec, Canada
e-mail: ronald.beaubrun@ift.ulaval.ca

Alejandro Quintero
Department of Computer and Software Engineering
Polytechnique Montreal
Montreal, Canada
e-mail: alejandro.quintero@polymtl.ca

***Abstract*— In multi-tenancy cloud environments, physical resources are transparently shared by multiple Virtual Machines (VMs) belonging to multiple users. Implementing an efficient access control mechanism in such environments can prevent unauthorized access to the Cloud resources. In this paper, we propose an access control mechanism that provides scalable and secure access control to the Cloud in the context of multi-tenancy cloud environments. Such a mechanism will prevent malicious tenants from generating and sending unauthorized traffic to the Cloud network.**

Keywords-access control; cloud computing; hypervisor; multi-tenancy; security.

I. INTRODUCTION

Cloud computing is a flexible and cost-effective platform for providing business and consumer services over the Internet [1][8]. Such a platform is utilized by multiple customers who share computing resources, including CPU time, network bandwidth, data storage space, with other users, which refers to multi-tenancy [2]. By multi-tenancy, Clouds provide simultaneous, secure hosting of services for various customers utilizing the same infrastructure resources [3][9]. However, in multi-tenancy cloud environments, one customer can gain unauthorized access to the information of other customers. In this context, it is important to control the access of network entities to such information.

Access control is a security feature that controls how users and systems communicate and interact with other systems and resources. In general, there are three types of access control: physical access control, technical access control and administrative access control [5][11]. Physical access control refers to the implementation of security measures in a defined structure in order to prevent unauthorized access to sensitive materials. Examples of such control include: security guards, picture IDs, locked and dead-bolted steel doors, biometrics, closed-circuit surveillance cameras and motion or thermal alarm systems. Technical access control employs the technology as a basis for controlling the access to sensitive information throughout a physical structure and over a network. Examples of technical access control are: encryption, smart cards, network authentication, Access Control Lists (ACLs) and file integrity auditing software. Administrative access control defines the human factors of security. All levels of the personnel within an organization are involved in such control. Administrative access control also determines which users have access to which resources and information.

The above types of access control can be integrated into security architectures in order to preserve the integrity, confidentiality and availability of resources that are collocated in multi-tenancy Cloud environment. In this paper, we investigate the use of technical access control for proposing a secure access control mechanism in the context of multi-tenancy cloud environments. Such a mechanism will prevent malicious insiders from generating and sending unauthorized traffic to the cloud network.

The rest of the paper is organized as follows. Section II introduces the context and background related to access control in multi-tenancy cloud environments. Section III presents the main assumptions and principles of the proposed architecture. Section IV explains a use case scenario, whereas Section V gives some concluding remarks.

II. CONTEXT AND BACKGROUND

As illustrated in Figure 1, a multi-tenant Cloud service provider has three essential elements: the Cloud manager, the hypervisor and the Virtual Machines (VMs) [6]. The Cloud manager is a console of management provided for clients in order to manage their Cloud infrastructure, which means creating, shutting down, or starting the instances. The hypervisor, also called Virtual Machine Manager (VMM), allows multiple operating systems (guests or virtual machines) to run concurrently on a host server. Its main responsibility is to manage the application's operating systems (OSs) and their use of the system resources (e.g., CPU, memory and storage). Its role is to control the host processor and resources, and also to allocate what is needed to each operating system.

A VM is an isolated guest operating system installation within a normal host operating system. In this context, each client may have one or more VMs, as one physical server can host several VMs. In such an environment, one client can send unlimited amount of traffic to another client. Accordingly, a malicious agent can rent a VM on the same host where the target VM resides. This malicious agent can send unauthorized traffic to the target VM and violate the security of the target VM [10].

 ISBN: 978-1-61208-845-7

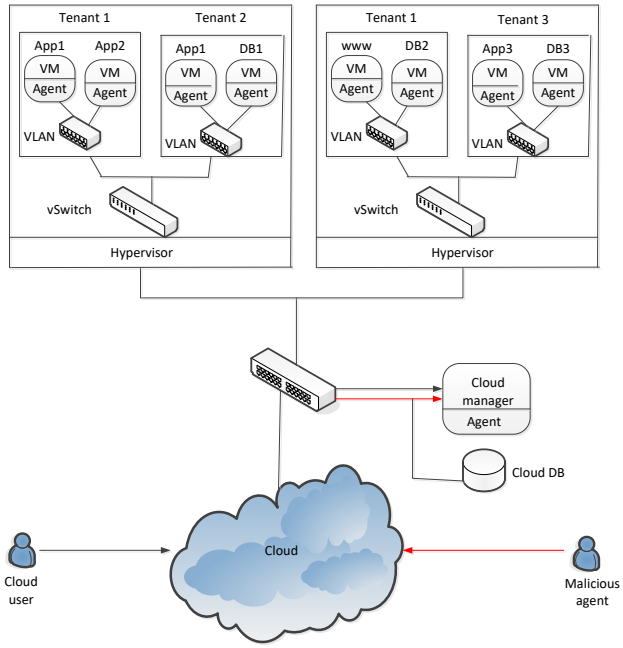

Figure 1. A model for a multi-tenant cloud service provider [6].

The unauthorized traffic may contain some script or malware which violates the confidentiality or the integrity of the target VM data. Sending such traffic to another VM makes it possible to perform other sorts of attacks. For instance, a malicious agent who owns a VM can perform VM Hopping over another user who is co-located at the same host. With VM hopping, an attacker has the control of one VM and tries to gain the control of another VM. VM hopping allows an attacker to move from one virtual server to the next one, or even to gain the root access to the physical hardware. VM hopping is a considerable threat because several VMs can run on the same host, which makes them the targets for the attacker. By performing this attack, a malicious user can violate the security and steal the data of other users who are located at the same server while compromising the hypervisor file system [4].

In addition, the malicious insider can perform Denial of Service (DoS) attacks. These kinds of attacks exhaust the resources of the Cloud network, such as bandwidth and computing power, by sending large amount of unauthorized traffic to other VMs.

III. EXISTING METHODS AND MODELS

In this section, we discuss the main existing methods and models for controlling access in the context of multi-tenancy cloud environments.

A. *Distributed access control*

The Distributed Access Control (DAC) architecture was proposed by Thomas et al. [12]. Such an architecture has three main components: the Cloud Service Provider (CSP), the Cloud Service Consumer (CSC) and the Identity Provider (IdP). The CSC requests the resources or services hosted by the CSPs. In this stage, the CSC should be first authenticated to ensure that unauthorized users do not access the services from the CSP. The main responsibility of the CSP is to host and to provide various services or resources to the CSCs. As a result, for avoiding illegal and unauthorized access by CSCs, proper authorization and authentication of CSCs are required.

Moreover, in DAC architecture, the IdP plays a great role since it generates identity tokens to the users. By using this identity token, a user can request the access to the cloud. Such a user may subscribe to services from multiple CSPs to meet the resource requirements. In this case, a federated identity management approach is required. The CSCs can use the identity tokens generated by the IdPs and these cloud users can exchange such tokens with various CSPs in the federation [12].

Analysis and results of DAC architecture reveal that using such an architecture is important in the domain of distributed applications or service computing. However, this model has some limitations. In particular, there is no effective mechanism which meets all access control requirements.

B. *Adaptive access algorithm*

Wang et al. [13] added trust management to the Role-Based Access Control (RBAC) in order to propose an adaptive access algorithm for cloud environments. This model is based on loyalty, i. e., a user is restricted only when its behavior contains malicious behavior. More specifically, the user request is first analyzed, and based on trust evaluation, the user becomes dynamically authorized. Here, user's trust is calculated according to user's behavior. In other words, the user access to the resource is dynamically based on calculation. As a result, by establishing dynamic mapping between roles and trust values, this model is able to determine the security level and control the user's access to the resources.

The trust-role-based-access control model claims that it can efficiently control user's malicious behavior. However, this model depends on the trust values, as the trust evaluation process needs to be improved in order to become widely used.

C. *Multi-tenancy access control model*

Multi-Tenancy Access Control Model (MTACM) is a security architecture which embeds the security duty separation principle in multi-tenancy cloud environments [14]. The main idea of MTACM is based on limiting the management privilege of CSP and letting the customers manage the security of their own business. In this model, the duty separation mechanism between cloud service provider and cloud customer is handled by a management module. However, the management module is not user-friendly for customers, as the cloud customer has to take care of the data security.

 ISBN: 978-1-61208-845-7

D. *Role-based multi-tenancy access control*

Role-Based Multi-Tenancy Access Control (RB-MTAC) applies identity management to determine user's identity and applicable roles [15]. Such a model combines two important concepts in access control under multi-tenancy access environment: identity management and role-based access control. In this context, Yang et al. [15] believe that this combination makes it easier to manage privileges that protect the security of application systems and data privacy. Providing a set of privileges and identity management schemes for corporations in cloud computing environment is the main contribution of this security model.

This scheme can be used to easily change employee privileges when a personnel member leaves an organization or when we want to grant employees more access without the need to modify all employee privileges one by one. However, RB-MTAC is not independent, and for implementing it in a cloud computing system, a directory service is needed.

E. *CloudPolice*

Popa et al. [7] proposed CloudPolice, a system that implements a hypervisor-based access control mechanism for multi-tenancy cloud environments. Since hypervisors are generally trusted, network-independent, close to VMs and fully software programmable, CloudPolice seems to be effective to prevent denial of service (DoS) attacks from malicious agents who send unauthorized traffic to their targets. As a result, CloudPolice acts as stateful firewalls and creates a state for each flow.

However, there are several major concerns for the feasibility of CloudPolice. The first concern is the ability for the hypervisor to act on per flow state, as the hypervisor should be ready to act on every single flow. The second concern is the ability to install new state with low enough latencies for new traffic flows, as we should make sure that the hypervisor is able to create a state for each new incoming flow very fast. As a result, the hypervisor should be able to create states for all new flows without latency (or at least with acceptable latency) and also act on the states that already exist in the buffer. Also, CloudPolice imposes overheads in the system, as the destination hypervisor receives all the traffic and decides to pass or drop the traffic based on the security attributes of the target virtual machines.

IV. THE PROPOSED ARCHITECTURE

This section defines the main assumptions, as well as the design and principles of the proposed architecture.

A. *Main assumptions*

The proposed architecture deals with the concept of Inter-VM traffic, which is the transmission of any data packet to and from one virtual machine. In other words, when the hypervisor encounters inter-VM traffic, the traffic does not pass through the physical switch or router, as the virtual switch that is located at the hypervisor forwards the packet to the destination VM. At this point, the following assumptions need to be done:

- The virtual machines and physical servers are co-located at the same cloud provider. If the entire system is not part of the Cloud, then for sending traffic to another Cloud, the traffic should pass through a real router or firewall. In this case, the policies that are implemented in the firewall should be enforced.
- Each physical server has only one hypervisor. In this case, the security attributes and access control lists of all virtual machines that belong to a physical server are located at one hypervisor. If we have multiple hypervisors on a physical server, we should apply an extra process for realizing which hypervisor contains the access control lists of certain virtual machines.
- Each physical server is hosting at least one tenant, and each tenant has at least one virtual machine. Since each virtual machine should be registered as a tenant, if a tenant is registered in the Cloud, a virtual machine should be assigned to that tenant.
- All access control lists are defined and stored in the hypervisor.
- In its startup process, a hypervisor sends an update message to the other hypervisors that are located at the same Cloud. This update message contains the IP address and the ID of virtual machines that are located at that hypervisor.

B. *Architecture principles*

The principles of the proposed architecture are based on control packets, which is the core element for verifying security permissions of virtual machines in multi-tenancy Cloud environments. In this section, we explain the elements of the proposed access control architecture, which is illustrated in Figure 2.

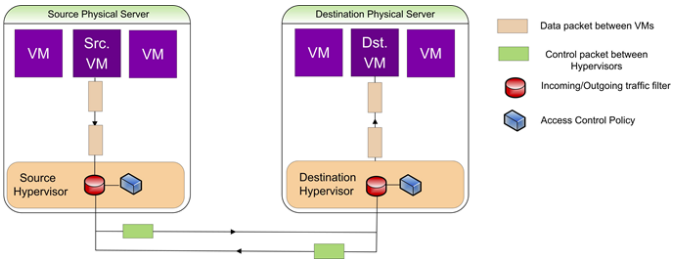

Figure 2. Principles of the proposed architecture.

- Source (Src.) VM is a virtual machine that is installed on the source hypervisor, as the latter is located at the physical source server. The source VM is then sending traffic packets to a virtual machine in the same Cloud called Dst. VM.
- Destination (Dst.) VM is installed at the destination hypervisor, and this hypervisor is located at the destination physical server.

- A data packet is a packet that the source VM wants to send to the destination VM.
- A control packet is a special packet that is generated by the source hypervisor. Its content represents the specifications of the source and destination VMs.
- Incoming/outgoing traffic filter is a lightweight IDS that is integrated in the hypervisor. It compares the control packet with the access control lists of destination VM.
- An access control list is a set of security permission that defines the level of security of each virtual machine.

C. Architecture design

The main goal of the proposed architecture is to block and drop undesired packets as close as possible of the source hypervisor. As illustrated in Figure 2, when the source VM sends traffic to the destination VM, such traffic has to pass through the source hypervisor. As soon as a data packet reaches the hypervisor, it generates a control packet which consists of the necessary information for access control checking, such as the source IP address, the destination IP address, the port numbers, as well as the protocol type. Such a control packet has to be sent to the destination hypervisor which checks its content and decides whether the traffic can be delivered to the destination hypervisor. If the source VM is permitted to send the so-called traffic to the destination VM, the destination hypervisor adds a pass or drop value to the control packet payload, and sends it back to the source hypervisor. According to this value, the source hypervisor threats the awaiting traffic.

As illustrated in Figure 3, the process starts when a VM initiates to send some traffic to another VM. As soon as such traffic is received by the source hypervisor, it checks the packet and looks for the destination address that is located at the inserted IP packet header. If the destination address belongs to a virtual machine in the same cloud, we will have two possibilities. The first case considers that the destination address is located at the same physical server. In this case, the architecture checks the access control policy of the destination VM, and can decide whether to pass or drop the traffic. The second case occurs when the destination address is located at a different physical server. In this case, the source hypervisor generates and sends the control packet to the destination hypervisor. Then, it waits for the response control packet.

Beside such possibilities, there may be an exception, when the destination address does not belong to any VM in this cloud, which means that the source and destination addresses belong to two devices that are not co-located at the same Cloud. In this case, the architecture only has to pass the traffic to the default gateway of the source hypervisor (router, switch or firewall).

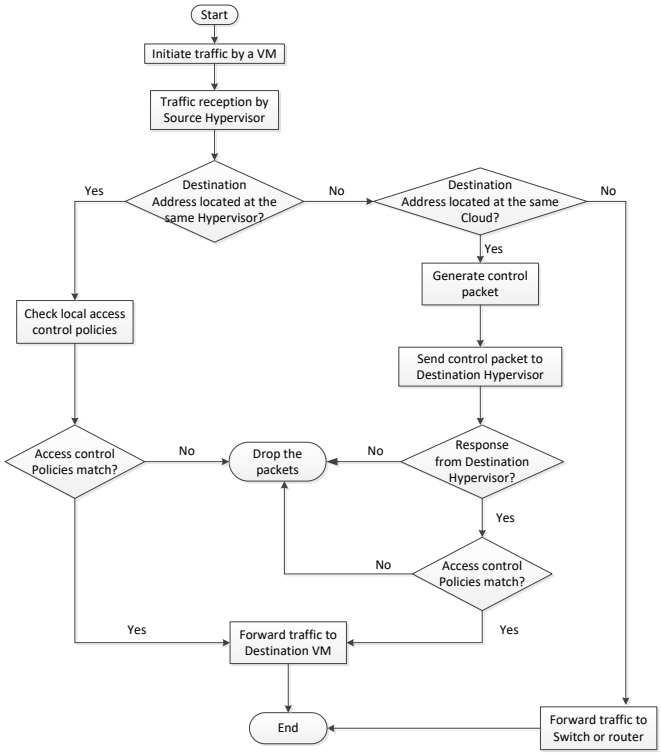

Figure 3. General mechanism flowchart.

The main part of the mechanism starts if the destination address belongs to a VM that is located at a destination hypervisor. In this case, the whole traffic should wait until the source hypervisor generates and sends a control packet to the destination hypervisor. Hence, the decision will be made based on the response control packet. Figure 4 shows the main tasks of the destination hypervisor when it receives the control packet from the source hypervisor. More precisely, the destination hypervisor selects one of the following actions:

- Insert a pass value to the control packet if the access control policy of the destination VM matches, and accept the traffic from the source VM.
- Insert a drop value to the control packet if the access control policy of the destination VM does not match, as the source VM is not authorized to send the traffic to the destination VM.
- Insert a null value to the control packet if the destination address is not found in the destination hypervisor. This may happen if the control packet is sent to the hypervisor by mistake, or if the VM destination is migrated to another hypervisor, whereas the source hypervisor is not informed about such migration.

 ISBN: 978-1-61208-845-7

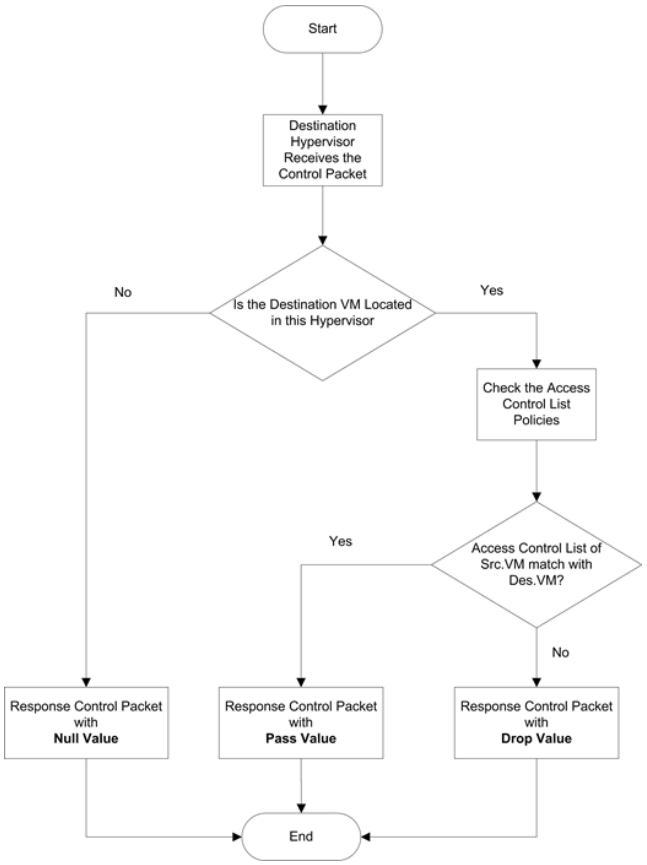

Figure 4. Destination hypervisor's tasks after control packet reception.

After inserting the proper value to the control packet, the destination hypervisor returns the edited control packet to the source hypervisor. The response control packet contains the decision and the action to be taken for the traffic. In the case of a drop value, the source hypervisor drops the traffic right away, as such traffic will not even exit the hypervisor, which means no wasted and unnecessary traffic in the network. Consequently, the network bandwidth does not suffer from extra and unwanted traffic. Finally, the pass value indicates that the access control policy matches between the source and destination, whereas the source VM and the traffic will pass throughout the destination hypervisor.

V. A USE CASE SCENARIO

In this section, we analyze a use case scenario which enables to tackle the problem of sending unauthorized traffic to a VM in the context of multi-tenancy Cloud environments. This scenario is illustrated in Figure 5, where a public Cloud is connected to the Internet, using a router and three physical servers that are connected to a layer-2 switch. In this scenario, the function of the router is to route the internal traffic of the Cloud to the Internet. Apparently, the router serves as a controller, enabling the networked devices to talk to each other efficiently.

In this scenario, there are 3 physical servers, as well as 10 virtual machines. These virtual machines belong to 4 tenants. The multi-tenancy topology of this Cloud is as follows:

- Server 1: Tenant 1 (VM1, VM2) and Tenant 2 (VM3)
- Server 2: Tenant 1 (VM4, VM5) and Tenant 3 (VM6, VM7)
- Server 3: Tenant 4 (VM8) and Tenant 3 (VM9, VM10)

It is important to mention that the process of controlling the access is executed in the hypervisors. In this context, the scenario has two phases: the first phase consists of generating control packets, whereas in the second phase, the destination hypervisor investigates the information and decides about the destiny of the packet. More specifically, in phase one of the scenario, the VM Source sends a traffic flow to the hypervisor source, as illustrated in stage 1 of Figure 6. Then, the source hypervisor generates a control packet. The content of this control packet is based on the traffic to be sent from source VM3 to destination VM8.

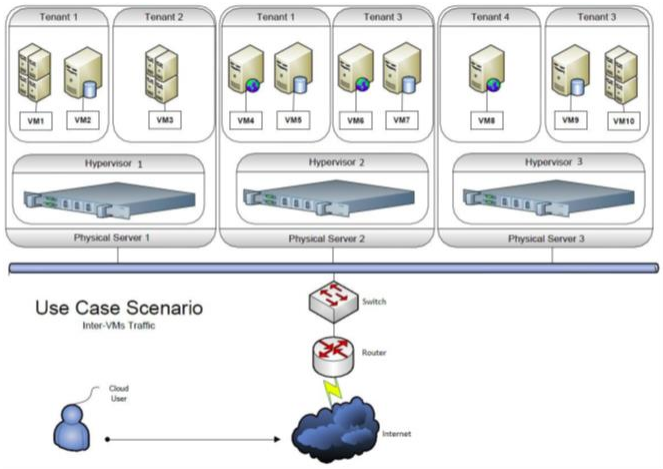

Figure 5. A use case scenario for multi-tenancy cloud access control.

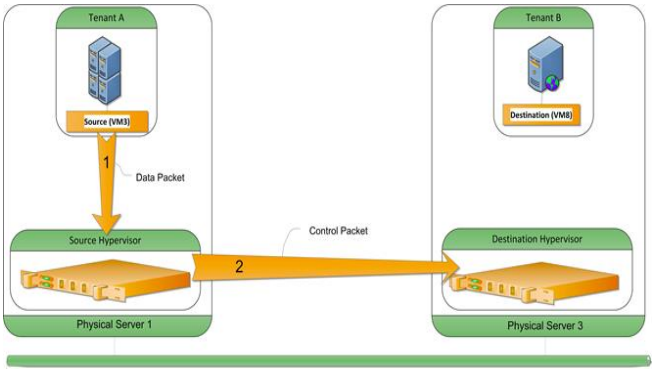

Figure 6. Illustration of phase one of the scenario.

As illustrated in stage 2 of Figure 6, the source hypervisor sends the control packet to the destination hypervisor in order to check the access control policy of the VM destination.

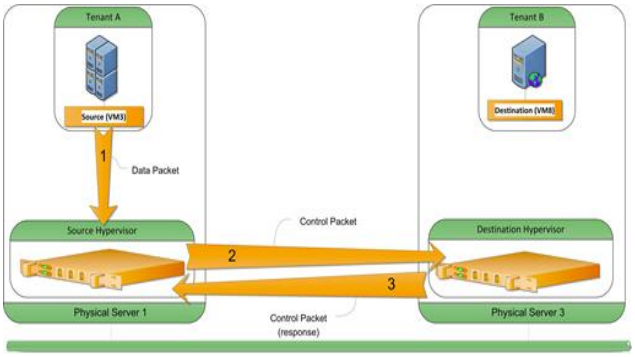

Figure 7. Illustration of phase two of the scenario.

In phase two, the control packet arrives at the destination hypervisor which checks the access control lists (ACLs) to verify if VM3 is authorized to send traffic to VM8. If the ACLs related to VM8 match, the destination hypervisor sends back a pass value within the control packet (called response control packet) to the source hypervisor, as illustrated in stage 3 of Figure 7. The response control packet enables the hypervisor source to decide what to do with the traffic that is waiting in the source hypervisor. Hence, if the security attributes of VM8 do not match the data packet, then the destination hypervisor sends a drop signal to the source hypervisor.

VI. CONCLUSION

The access control architecture proposed in this paper for multi-tenancy Cloud environments satisfies a number of the requirements, such as scalability and security. This architecture is scalable in the sense that, if the number of VMs grows, we only need to implement this architecture in the hypervisor of each physical server without any extra changes in the system. Besides that, the architecture enables to maintain the security of information in the Cloud system by controlling the traffic sent from one hypervisor to another hypervisor and by enforcing the security policies in the hypervisor. Using such an architecture leads to better performance by avoiding unnecessary traffic and dedicating the Cloud resources to necessary traffic. Future works will focus on implementing a prototype of the proposed architecture on a real Cloud environment.

REFERENCES

[1] K. Hashizume, D. G. Rosado, E. Fernández-Medina, and E. B. Fernandez, "An analysis of security issues for cloud computing," Journal of Internet Services and Applications 4:5, vol. 4, 2013, pp. 5-18.

[2] Z. Minqi, Z. Rong, X. Wei, Q. Weining, and Z. Aoying, "Security and Privacy in Cloud Computing: A Survey," in 6th International Conference Semantics Knowledge and Grid (SKG), , Beijing, 2010, pp. 105-112.

[3] C. J. Guo, W. Sun, Y. Huang, Z. H. Wang, and B. Gao, "A framework for native multitenancy application development and management," in 9th International Conference on E-Commerce Technology/4th International Conference on Enterprise Computing, Ecommerce and E-Services, Tokyo, 2007, pp. 551-558.

[4] A. Jasti, P. Shah, R. Nagaraj, and R. Pendse, "Security in Multi-Tenancy Cloud," in International Carnahan Conference on Security Technology (ICCST), , San Jose, CA, 2010, pp. 35-41.

[5] S. Harris, CISSP All-in-One Exam Guide, Sixth ed. New York: McGraw-Hill, 2013.

[6] K. Benzidane, S. Khoudali, and A. Sekkaki, "Autonomous Agent-based Inspection for inter-VM Traffic in a Cloud Environment," in 7th International Conference for Internet Technology and Secured Transactions (ICITST-2012), London, 2012, pp. 656-661.

[7] L. Popa, M. Yu, S. Y. Ko, S. Ratnasamy, and I. Stoica, "CloudPolice: Taking Access Control out of the Network," in ACM Workshop on Hot Topics in Networks. HotNets, Monterey, CA, USA, 2010, pp. 1-6.

[8] M. Auxilia and K. Raja, “Dynamic Access Control Model for Cloud Computing,” Sixth International Conference on Advanced Computing (lCoAC), pp. 47-56, 2014.

[9] G. Kappes, A. Hatzieleftheriou, and S. V. Anastasiadis., “Multitenant Access Control for Cloud-Aware Distributed Filesystems,” IEEE Transactions on Dependable and Secure Computing, Vol. 16, No. 6, pp. 1070-1085, 2019.

[10] S. J. De and S. Ruj, “Efficient Decentralized Attribute Based Access Control for Mobile Clouds,” IEEE Transactions on Cloud Computing, Vol. 8, No. 1, pp. 124-137, 2020.

[11] K. Albulayhi, A. Abuhussein, F. Alsubaei, and F.T. Sheldon, “Fine-Grained Access Control in the Era of Cloud Computing: An Analytical Review,” 10th Annual Computing and Communication Workshop and Conference (CCWC), pp. 748 – 755, 2020.

[12] M. V. Thomas and K. C. Sekaran, "An Access Control Model for Cloud Computing Environments," in 2nd International Conference on Advanced Computing, Networking and Security (ADCONS), Mangalore, pp. 226-231, 2013.

[13] W. Wenhui, H. Jing, S. Meina, and W. Xiaohui, "The design of a trust and role based access control model in cloud computing," in 6th International Conference on Pervasive Computing and Applications (ICPCA), Port Elizabeth, pp. 330-334, 2011.

[14] X.-Y. Li, Y. Shi, Y. Guo, and W. Ma, "Multi-Tenancy Based Access Control in Cloud," in International Conference on Computational Intelligence and Software Engineering (CiSE), Wuhan, pp. 1-4, 2010.

[15] S.-J. Yang, P.-C. Lai, and J. Lin, "Design Role-Based Multi-tenancy Access Control Scheme for Cloud Services," in International Symposium on Biometrics and Security Technologies (ISBAST), Chengdu, pp. 273-279, 2013.

 ISBN: 978-1-61208-845-7