A Novel Virtualization Enabled Cloud Infrastructural Framework for Enhancing Private Cloud Communication Security

Nirmalya Mukhopadhyay[1], Tushar Kanti De[2], Dipankar Chatterjee[3]

[1,2,3]Department of CSE, Faculty of Engineering & Technology, Assam Down Town University (in Collaboration with iNurture Education Solutions Pvt. Ltd.), Guwahati, India
Corresponding Author: nirmalya.njrcs@gmail.com

Research Paper Received: 31.07.2020 Revised: 15.10.2020 Accepted: 24.02.2021

Abstract—To implement security in cloud platform, the significance of virtualization can't be neglected. The hypervisors, which is also known as virtual machine monitor, offer a handful of facilities to configure the virtual machines to redesign the communication channel through which they can interact with each other or with the internet. We have studied how the facilities of virtualization can be optimally used to secure the communication between virtual machines and internet in a private cloud infrastructure. In this regard, we have tried to enhance the communication mechanism through constructing a cloud infrastructure framework. This paper deals with the isolation of virtual instances to improve the private cloud security through prevention from spreading infections between virtual machines (VMs) & also between VM & internet. Moreover, our proposal focuses on the filtering of inflow and outflow traffic to and from the outside world through a dedicated VM of a host implemented in layered switching environment.

Keywords—Cloud computing, Security, Virtualization, Hypervisor, VM, Private cloud infrastructure, Layered switching, Communication mechanism, Security framework, Inverse co-efficient index.

1. Introduction

In this information era, different data processing and data analyzing requirements have been proved essentials. Approximately, 2.5 quintillion bytes of data are produced by humans every day that need qualitative attention for business intelligence and analytical hierarchical processes (known as AHP). To complete this mammoth computation, contemporary approaches is not sufficient. Hence, the exclusive requirement of non-conventional cloud computing technology proven worthy. Similarly, today's applications are also enormously resource dependent, striving for robust computing power to satisfy increasing demand of the user's needs. While opting a titanic solution, cloud computing covers the large scale of distributed paradigm mitigating the global demands. Different cloud service providers (CSPs) have developed many services that helps in the whole analysis and processing operations. For example, AWS Kinesis, AWS Kafka, Azure Stream Analytics, Google Analysis have been developed for high-volume data analysis.

On the other hand, AWS, Azure or Google is using Hadoop and MapReduce to process those data and generate modular solutions as per requirements. The computational complexity has been optimal and general users have had least amount of operation-workflow-botheration. Virtualization, which is regarded as the primary back-bone technology of cloud computing infrastructure, enhances the computational power and storage sufficiency [6] by optimally managing the under-utilized computational resources. This fabulous solution has some major challenges in terms of security. This is more vulnerable to the malicious users because of its openness and over the internet deliverables. So, attackers have endless possibilities to compromise these systems. There can be different types of attacks such as SQLi, XSS, LDAP Injection, Session Hijacking, DoS, DDoS attacks [8][9][10] etc. to compromise a system. To protect the online systems from different attacking possibilities, researchers and developers have to repeatedly work on the security policies and enhance them to mitigate the risks.

Although virtualization provides numerous benefits through VMs but the set of challenges that are present in the system in terms of privacy and security can be so harmful. For example, data breaches, data loss, data leakage, malicious infection, traffic hijacking are having terrifying effects in the system and also in the business. Now, fortunately, according to the property of virtualization, the virtual machines are logically separated. Taking this feature as a field of opportunity, we have discussed on the policies for securing private cloud platform through a newly proposed ifrastructural framework for its inflow and outflow traffic with the help of virtualization.

2. Literature Review

In private cloud infrastructure we are looking to enhance the communication mechanism through construction of a cloud infrastructure framework which deals with the isolation of virtual instances to improve the private cloud security through prevention from spreading infections between virtual machines (VMs) & also between VM & internet. İn this regard our literature review regarding the security in private cloud stated by Prof. Shiwani Sthapak et al. [2] that DEPSKY system which mainly used in Multi-Cloud environment provides DEPSKY library which is implemented in the clients for dealing with heterogeneous interfaces of different cloud service providers. This helps to increase the service availability, massive scalability and data integrity. Alternatively, Xiaowei Yan et al. [1] implemented a cloud computing security framework focusing on increasing the security in the cloud by configuring a firewall only specific ports are being opened for important services instead of opening all ports at the same time which can lead to increase in security. Also, by the blocking ICMP messages helps unnecessary TCP services to shut down and the firewall configuration helps in non-acceptance of any traffic from the network. Although, according to Zubair Ahmad et al. [3] virtualization can be risky because in cloud virtualization on one server the virtual machines are residing and each and every virtual machine can be allocated separate roles, but can prevent attacks in different approaches. The virtual environment is having lack of control access; as a result, the hackers can easily compromise the virtual machines. And it is also very difficult to manage the virtual machines having different roles for which different persons or administrators are required. Also, by Hanqian Wu et al. [4] there can be sniffing and spoofing vulnerability in the virtual networks

by which network security between the virtual machines can be compromised. For this, a new model had been proposed in which the network security between virtual machines in the virtual network can be maintained. A firewall layer is being imposed in between the routing layer and shared network layer. Security policies are being implemented in this layer for preventing spoofing attack by dropping the corresponding ARP packets and also preventing virtual interface in the routing layer which is connected to some shared network layer to be connected to other virtual shared network.

While surveying further, we found that, Akshay Gangwani et al. [5] told that numerous security hazards can be there in VM live migration process by hypervisors like Zen, KVM, VMware. Also explained different types of attacks which can happen in live migration process. Inclusion of different methods like PALM (protection aegis of live migration in VM) which mainly deals with privacy and integrity of valuable data but the performance is compromised & TAL (trusted assurance level) which deals with the scalability and improved performance compromising the Security issues. Pooja Sharma et al. [6] came up with a three-layered architecture for security model in cloud storage system (viz authentication layer, encryption layer and recovery layer) and explains that each and every layer is responsible for securing data in the cloud. An idea about implementing some security policies between vendor and users for improving the security. Some encryption techniques for end-to-end encryption of data in transit, at rest and in process so that the attackers cannot understand the actual data. Multi factor authentication (MFA) can also be implemented for strong authentication, implementing access management like giving permission to different administrators in cloud according to their roles and also implementing SIEM and SOC system should be used for early detection and proper actions to be taken of the attacks. Kishu Gupta et al. [7] tells data is very important to every organization, therefore, security policies should be applied so that data is not been leaked to unauthorized parties. Some modules used for data leakage detection (DLD) had been proposed which explained some DLD techniques used for leakage detection like watermarking and data allocation strategy.

Amna Riaz et al. [8] made a survey on different intrusion detection systems (IDS) and made a performance analysis on the existing cloud-based IDS. In today's world despite of evolution of different IDS still some problems are existing like if a VM which is running on a hypervisor is compromised then hypervisor can have chance to get compromised and in turn can violate the other VMs hosted by the hypervisor. Also, during VM migration process attackers can gain full control of the VM and intruders can initiate or terminate the VM migration process. The insecure channels used for the VM migration helps the attacker to do active or passive attacks. N. Chandrakala et al. [11] proposed migration-based approach like deployment of discrete time markov chain (DTMC) for evaluating the security of each VM and predict the possibility of each VM to be compromised. This approach also deals with the performance overhead for better performance of cloud platform. A research by R. Velumadhava Rao et al. [12] tells that for moving into internet-based cloud model we have to mainly focus on data security and privacy as data leakage causes severe impact on business and brand and trust of an organization. Research tells that different security challenges

are there in cloud like confidentiality, integrity, availability, locality, segregation etc. To mitigate the security challenges, some solutions came up like encryption of data before it being stored in cloud and permissions to be given to some group members who are authenticated. Also, before storing the data owners have to ensure that the data is not being altered by calculating some hash values which helps in data integrity. RSA based data integrity check can be one solution. SaaS also helps in creating clear boundaries both at physical level and application level to ensure the segregation of data from different users.

Eventually, Shalu Mall et al. [13] says that nowadays use of cloud has increased rapidly by organizations and institutes. Instead of storing the data in their own local server the data owners are storing the data in the cloud. So, in this scenario the cloud security is must for protecting the data in the cloud from unauthorized access. For this, a new concept had been proposed for protecting the data using genetic algorithm. By the genetic algorithm operations (crossover and mutation) the data is encrypted and corresponding cipher text is obtained. And also, this cipher text is encrypted in blocks with the help of the private key of data owner and the encrypted blocks are stored in different locations in the cloud. As a result, the cloud service provider cannot access the data as well as cannot find the location of the data which is being stored which will help in data security in cloud. Data integrity problem and verified the client's data stored in the cloud by third party auditor (TPA) is also among the investigations by Anirudha Pratap Singh et al. [14]. To alleviate this a development had been made on an optimized public authentication protocol which will optimize the cloud storage server (CSS) size by storing homomorphic linear authenticators (HLA) and also, they use chameleon hashing and a modified chameleon authentication tree (mCAT) for performing efficient dynamic updates on client's data. The proposal comprises of three phases like setup phase used for generating key, mCAT generation and authorizing a TPA, dynamic data update phase used for block-level and fine-grained updates on client's data using mCAT and third-party auditing phase used for getting integrity-proof of client's data from the CSS.

3. Our Proposal

Virtualization is the technology through which one can create logical mimics of physical resources available. These logical instances are isolated, segregated and encapsulated files that are co-related but independent and they can behave like complete machines. By stating complete machine, we mean that they are comprises of virtual hardware, virtual operating system (known as guest operating system), virtual application. So, they can compute just like a frequently used physical machine. Below are the conceptual designs for a machine before virtualization & after virtualization.

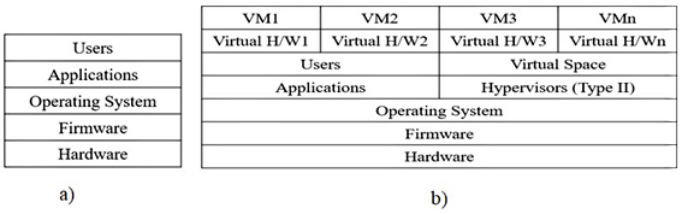

Fig. 1. Architecture of physical machine before (a) and after (b) virtualization

From the above diagrams, it is evident that we can run multiple virtual independent machines in a single physical machine. It will decrease the resource sprawl & resource spike problems

and increase degree of parallel processing. And there is an abstraction layer, which hides all the underlying details from the users. So, this is more secure than the contemporary architectures. Also, heterogeneity issues are resolved presenting multi-compatible systems. The high availability property will ensure more suitable disaster recovery. Overall, this architecture is a perfect choice for enterprises and organizations, as all of them are betting big on cloud technology. So, the researchers and developers are trying to enhance the security measures as & when possible.

In virtualization scenario, we can create a set of well-defined VMs for different purposes. Based on the roles that they will play, we can categorize them with different designations. The most primitive types can be named as Host-VM (VMH), Controller-VM (VMC), Execution-VM (VME) and Storage-VM (VMS). The VMH will work as a co-ordinator and load balancer for different VME. VMC will control all the computational requirements and will be in charge to take a decision. VME will execute whatever Instructions are placed and finally, VMS will store the relevant data or information for further processing and reference. So, the whole cluster is a four-touple. Now, we can denote it as a set of these four types of VM viz. {VMH, VMC, VME, VMS} $\in$ R that exist within the virtualization environment. Due to the fact that as the cluster can be considered as an intranet; so, it is possible to represent that as an undirected graph. All the VMs can be considered as vertex of the graph and the connections between them will be considered as edges. Mathematically, GVM $= \{N_i, \epsilon\}$ is the simplest form that we can imply to formalize the whole concept. Here, $N_i = \{N_1, N_2, N_3 \ldots\}$ is the set of all primitive VMs and $\epsilon = N_i \times N_i$, which is actually the set of all allowable connections between the VMs. Different topologies can be used to finalize the value of ϵ with in the environment.

Cloud infrastructure implementation is completely dependent on virtualization technology. One of the primitive aims of virtualization is to provide sand-boxing facility, which ensures isolation of the virtual instances. This categorically establishes one obvious benefit in case of cloud security, i.e., through this facility we can stop the spreading of infectious malicious codes from one compromised virtual machine (VM) to another. In our proposal, we have designed a framework in such a way that, every communication must be going on through such a predetermined path & VM that all other VMs will reveal no direct identity to the outer world.

In our proposal, we have taken one physical machine to create three virtual hosts within it. Each host contains two VMs. The VMs of host 1 & host 2 are running either servers or clients. Now if they are to connect with the internet, there is no direct path available through the virtual machine monitor (VMM). They have to communicate through the VMs of host 3. The VMs of host 3 are equipped with all kind of security measures. The following diagram will show the framework that we have installed & configured in the machine.

This proposed private cloud infrastructural framework uses a combination of standard and distributed switches to complete the communication between VMs and also outer world. The standard switches were used for intra-host communication and distributed switches redirects the communication of host1 and host2 into host3. Host3 has been exclusively designed as a filter for incoming and outgoing traffic. Hence any malicious content cannot surpass from one node

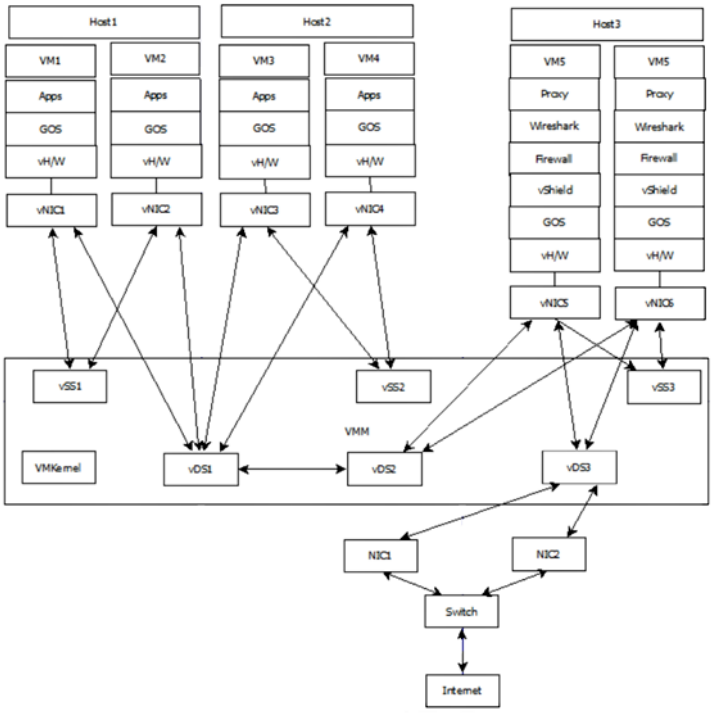

Fig. 2. The proposed private cloud infrastructural framework for secured communication

to another. All the existing models use either standard switch or distributed switch for their communication and use only the in-built firewall system as a filter. We, on the other hand have combined these two types of switches and also used a dedicated Host as a filter to accept or deny traffic after proper security check.

From Figure 2, it is evident that if VM1 wants to communicate with VM2, it has a direct connection to the virtual standard switch (vSS1). As they belong to the same host (Host1), there is a least possible amount of chance to spread infections. Although the vSS1 is configured with all the required security protocols. But, in case, that if VM1 wants to communicate with VM3, which belongs to a different host (Host 2), the chance of spreading of infection rises. So, the communication will pass through a virtual distributed switch (vDS1), which has more complex and efficient security protocols implemented & configured into it. The most vulnerable communication is when, let's say, VM1 wants to communicate to the internet. Here, VM1 has to request a connection to VM5 (or VM6, in case VM5 is not in active mode) through the layered virtual distributed switch set vDS1 & vDS2. Then, every packet will be filtered through the VM5's security tools & protocols and if found safe, will be redirected via another virtual distributed switch vDS3 to the physical switch through NIC1 (or NIC2, if required). The same will be done for any incoming traffic also. Thus, if affected, only one VM of Host 3 will be destroyed and we can repair or replace the VM easily and with minimum amount of Cost (either computational or monetary). In case of other VMs of host 1 & host 2, the same method will be followed.

While implementing this framework, we used

two VMs of Host 3 as filters, because we wanted to avoid SPOF problem. If one filter goes down then another filter will automatically be activated. The semaphore concept has been implemented for this. Semaphore offers atomic operation on signal (High) and wait (Low). This is how, we have taken care of optimal resource utilization.

Through this framework, the threat of infection on any physical or logical resources will fall down to a great extent. The overall aim to safeguard the cloud infrastructure will increase. Thus, the enterprises or organizations, running their private cloud will feel safer while connecting to the internet and communicating to the outer World.

Now, for creating a mathematical model of the entire proposed system, consider the total number of VMs in the framework as N_T. Some of these VMs are non-vulnerable or non-compromised. Let us consider that number as N_{NV}. These VMs will not be exposed to the outer-world as per our security framework. The number of VMs that are susceptible (or Open) to attack can be indicated as N_O. There will be a list of attacking VMs; we call them N_A. Now, the growth of these irresistible VMs can be shown as the differentiation value in respect to time and hence deduced as:

$$\frac{\partial N_A}{\partial T} \tag{1}$$

Whereas, the growth of the susceptible VMs can be formulated as the differentiation value in respect to time:

$$\frac{\partial N_O}{\partial T} \tag{2}$$

If we write a function ψ for denoting the total number of VMs available in the environment, then the function will look like:

$$N_T = \psi(N_{NV}, N_O, N_A) \tag{3}$$

From the equations (2) & (3), the number of attacked VMs (or compromised VMs) can be calculated as:

$$N_{COM} = \psi(N_A) \times \frac{\partial N_O}{\partial T} \tag{4}$$

So, the ratio of newly compromised VMs and susceptible VMs can be written as:

$$\frac{N_{COM}}{N_A} = \frac{\psi(N_A) \times \frac{\partial N_O}{\partial T}}{\frac{\partial N_A}{\partial T}} \tag{5}$$

If the co-efficient threat factor is T for all the susceptible VMs, then the probability of attack increases as:

$$\frac{\partial N_{COM}}{\partial N_O} = P_T(T_{NA} - N_{COM}) \tag{6}$$

Where, PT is the Probability of attack based on T and TNA is the total number of attacks done.

Considering the inverse ratio due to lesser percent of intra-VM attack and assault approach, will decrease. And hence the inverse equation will look like follows:

$$\frac{\partial T}{\partial N_A} = -C \times \frac{T}{N_A}, \tag{7}$$

where c is a constant or it can be re-written as:

$$T = \alpha \left(N_A^{-C}\right) \tag{8}$$

Where α is inverse co-efficient.

So, from the above mathematical model, it is clear that the threat coefficient factor T depends

on Inverse attack coefficient factor of N_A, which obviously will produce lesser amount of attack because of the inverse law applied on intra-VM attack and assault ratio.

4. Result and Discussion

The test-bed that we have prepared have been gone through a set of observation and based on the proposed mathematical model. N_T represents total number of VMs in the framework, N_{NV} represents invulnerable VMs (i.e., they are excluded from attack), N_O represents total number of Open VMs that are susceptible to attack and N_A are the attacking VMs. Furthermore, T_{NA} represents total number of attacks in every iteration that has been taken as one of the parameters for our simulation. Again, T is the threat factor that represents the probability of the N_O VMs that can be compromised. N_{COM} represents the total number of newly compromised VMs after every iteration.

From iteration number 1 to 5 we have observed that the numbers of newly compromised VMs are decreasing as T is increasing. This signifies that the growth rate in which the probability for a VM to get compromised decreases and hence the inverse ratio is satisfied which proves that our proposed framework and the mathematical model helps in enhancing the private cloud security. With this model, gradually, most of the VMs will be in a state where compromising is not possible.

The following table will show the number of compromised VMs have been decreased gradually based on the inverse coefficient.

The following charts display the comparisons that we got after running the simulations.

The above bar graph points out that the difference in the inverse ratio increases with proportionate to time derivation. It signifies that the number of newly compromised VMs are reducing. As a result, the risk factor for open VMs is reducing and hence the system is becoming safer.

The scatter graph depicts the peak values for each parameter along with time derivation.

Experiment for the proposed framework through the mathematical model:							
Number of Iterations	N_T	N_{NV}	N_O	N_A	T_{NA}	T	N_{COM}
1	4	0	4	1	10	0.12	4
2	4	1	3	1	10	0.36	4
3	4	2	2	2	10	0.45	3
4	4	3	1	2	10	0.75	1
5	4	4	0	2	10	0.99	0
6	8	0	8	3	10	0.12	8
7	8	2	6	3	10	0.36	7
8	8	4	4	5	10	0.45	3
9	8	6	2	5	10	0.75	1
10	8	8	0	5	10	0.99	0
11	15	0	15	5	10	0.12	11
12	15	5	10	5	10	0.36	7
13	15	9	6	5	10	0.45	4
14	15	11	4	7	10	0.75	0
15	15	15	0	7	10	0.99	0
16	20	0	16	7	10	0.12	6
17	20	6	14	7	10	0.36	3
18	20	10	10	7	10	0.45	1
19	20	16	4	10	10	0.75	0
20	20	20	0	10	10	0.99	0

TABLE 1
Simulation Results

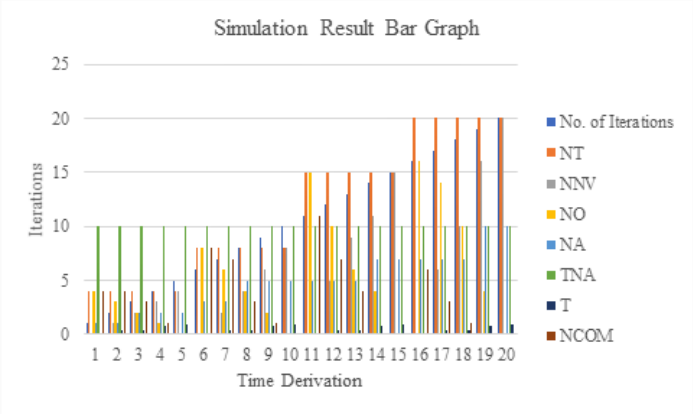

Fig. 3. Bar graph of simulation results

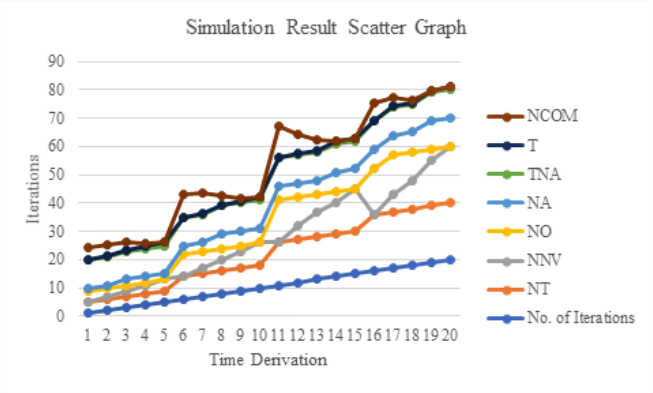

Fig. 4. Scatter graph of simulation results

Although the number of attacking VMs (N_A) are having a tangent graph here, still downfall in the graph of newly compromised VMs (N_{COM}) is clearly visible. So, betterment in the security can be ensured in our proposed model. Finally, the pivot chart also shows the aggregation values for the simulations done.

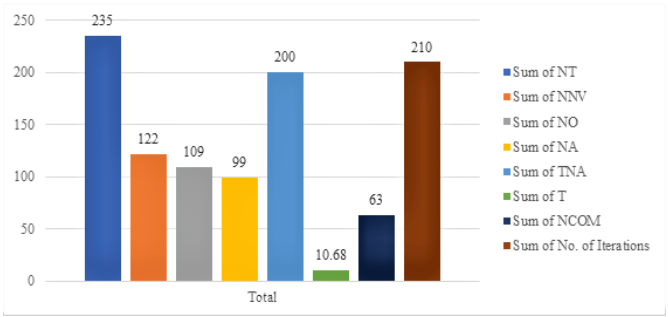

Fig. 5. Pivot Chart of simulation results

The above pivot chart shows the aggregated values for all the parameters used in our simulations. Here we can observe that the total number of newly compromised VMs (N_{COM}) is only 63 while total number of VMs (N_T) is 235 and total number of open VMs (N_O) is 109. Hence, we can say that a lot of susceptible VMs have not been compromised by following our mathematical model and infrastructural framework.

5. Conclusion

Cloud computing is a prosperous field but this should not be the opportunity for the intruder. As vulnerability issues are expanding day by day through providing increased facilities to the users. Therefore, the security measures should be developed accordingly to empower privacy & protection and to minimize the risk. Doing so, in any circumstances, is nothing but a thought to be implanted which can be a practiced law for tomorrow.

The proposed design can protect a private cloud environment from different unknown attack from the outer world. The proposed design can protect all the VMs under consideration through a master filter (VMs within a Host). This study has been carried out on VMware platform. Further enhancement in security policy can be achieved through extensive research. Secondly, implementation in various platforms will be our next objective.

Acknowledgments

We thank iNurture Education Solutions Pvt. Ltd. and Assam Down Town University to give us the chance to carry out our research works.

References

[1] Y. Xiaowei, Z. Xiaosong, C. Ting, Z. Hongtian, and L. Xiaoshan, The Research and Design of Cloud Computing Security Framework, Springer-Verlag Berlin Heidelberg, vol. 121, pp. 757-763, 2011.

[2] S. Sthapak, K. Garg and R. Shah, Enhanced Security in Cloud Computing: From Single to MultiClouds, IJERT, vol. 3, issue 2, pp. 2244-2248, February 2014.

[3] Z. Ahmad, A. M. Zeki and A Olowo, Risks with Cloud Computing Data Residency and Virtualization, Journal of Software Engineering and Intelligent System, vol. 1, issue 1, pp. 24-28, 31st August 2016.

[4] H. Wu, Y. Ding and C. Winer, Network Security for Virtual Machine in Cloud Computing, 5th International Conference on Computer Sciences and Conver-

gence Information Technology, vol. 1, pp. 18-21, 2010, doi:10.1109/iccit.2010.5711022.

[5] A. Gangwani, H. Shah and K. Shah, A Survey on Different Cloud Computing Security and Migration Techniques, IJRAR, vol. 7, issue 1, pp. 959-963, March 2020.

[6] P. Sharma and V. Jha, Boosting Security for Cloud Storage, IJRASET, vol. 8, issue V, pp. 2725-2730, May 2020, doi:10.22214/ijraset.2020.5458.

[7] K. Gupta and A. Kush, A Review on Data Leakage Detection for Secure Communication, IJEAT, vol. 7, issue 1, pp. 153-159, October 2017.

[8] A. Riaz, H. F. Ahmad, A. K. Kiani, J. Qadir, R. U. Rasool and U. Younis, Intrusion Detection Systems in Cloud Computing: A Contemporary Review of Techniques and Solutions, Journal of Information Science and Engineering, vol. 33, pp. 611-634, 2017.

[9] N. R. Tadapaneni, Cloud Computing Security Challenges, International Journal of Innovations in Engineering Research and Technology (IJIERT), vol. 7, issue 6, pp. 1-5, June 2020.

[10] M. S. Taj, S. I. Ullah, A. Salam and W. U. Khan, Enhancing Anomaly Based Intrusion Detection Techniques for Virtualization in Cloud Computing Using Machine Learning, International Journal of Computer Science and Information Security (IJCSIS), vol. 18, no. 5, pp. 68-78, May 2020.

[11] N. Chandrakala and B. Thirumala Rao, Migration of Virtual Machine to improve the Security in Cloud Computing, International Journal of Electrical and Computer Engineering, ISSN: 2088-8708, vol. 8, no. 1, pp. 210-219, February 2018.

[12] R. Velumadhava Rao and K. Selvamani, Data Security Challenges and Its Solutions in Cloud Computing, International Conference on Intelligent Computing, Communication & Convergence, vol. 48, pp. 204-209, 2015.

[13] S. Mall and S. K. Saroj, A New Security Framework for Cloud Data,8th International Conference on Advances in Computing and Communication, vol. 143, pp. 765-775, 2018.

[14] A. P. Singh and S. K. Pasupuleti, Optimized Public Auditing and Data Dynamics for Data Storage Security in Cloud Computing,6th International Conference on Advances in Computing & Communications, vol. 93, pp. 751-759, 6-8 September 2016, Cochin, India.