International Journal on Advances in ICT for Emerging Regions 2011 04 (02) : 12 - 23

Cloud Computing Reliability, Availability and Serviceability (RAS): Issues and Challenges

Farzad Sabahi

***Abstract*—Cloud computing is one of today's most exciting technologies because of its capacity to lessen costs associated with computing while increasing flexibility and scalability for computer processes. During the past few years, cloud computing has grown from being a promising business idea to one of the fastest growing sectors of the IT industry. But on the other hand, IT organizations have expressed concerns about critical issues such as security that accompany the widespread implementation of cloud computing. Security, in particular, is one of the most debated issues in the field of cloud computing and several enterprises look at cloud computing warily due to projected security risks. Also, there are two other issues. They are the reliability and availability of the cloud which are as important as security. Although each of those three issues is associated with usage of the cloud, they will have different degrees of importance. Examination of the benefits and risks of cloud computing is necessary for a full evaluation of the viability of cloud computing.**

This article reviews issues and challenges of cloud computing's reliability, availability and security (RAS). Beginning with a brief discussion on virtualization technology, a key element of cloud infrastructure, it examines issues facing in cloud RAS fields. Then, it addresses the challenges and problems in cloud computing RAS. It also examines intrusion detection methods and outlines counter measures to improve cloud RAS.

***Index Terms*— Cloud computing, Virtualization, Reliability, Availability, Security, Threat, Intrusion, Countermeasure.**

I. INTRODUCTION

Cloud computing is based on virtualization technology, in which each user uses a virtual machine. Virtualization technology includes two levels of virtual machines, which are VMs (virtual machine) and hypervisors. The hypervisor has administrative rights to control VMs. But virtualization has some issues that could endanger system performance. From a cloud viewpoint, there are many important dimensions of virtualization technology to consider, but the hypervisor's Reliability, Availability and Serviceability (RAS) is an important aspect of virtualization technology and requires special attention. For example, from security viewpoint, if someone gets control of the hypervisor, he will gain full control of all VM that are under the hypervisor control. Consequently, cloud technology has some problems in RAS that it has inherited from virtualization technology. One such problem involves overflows of system due to excessive combination of VM to a physical server that affects availability and reliability. Because of these issues, cloud systems are vulnerable to traditional attacks as well as new attacks that some of them have migrated from virtualization.

Privacy is another issue which can decrease virtualization and cloud's overall performance, because the VMs are located practically in a multitenant environment, thus making it possible for a user to access a past tenant's information in the same space. Although the use of encryption algorithms could be a good solution for the user or cloud provider by making the appropriate arrangements, such as using advanced algorithms to wipe the user's data for avoidance from information leaks. But the use of encryption algorithms has problems as well, such as the inability of owners to recover their data when they lose the decoding key.

As we know in the world of network computing, there is a variety of attacks that can cause serious problems for Internet-based technologies such as cloud. This can make the cloud vulnerable to some attacks, like the DoS family (Denial of Service) which aims to make the target server inaccessible to legitimate users. The cloud can be a victim of DoS attacks, but it can also be part of the solution by allocating more resources to a user under a DoS attack in order to prevent the user from crashing. Therefore, applying countermeasures to deal with security problems in the cloud is critical, whereas one of the main countermeasures is controlling access control in the cloud. Generally, it seems the security countermeasures in the access control part of the cloud often involve prevention—for example, management of permissions for the account to determine access to different levels of virtualization in the cloud.

Besides security, cloud providers are also responsible for reliability and availability, because all users expect the highest level of QoS (Quality of Service). The cloud providers use some solutions such as partitioning to achieve maximum performance. But according to whether the cloud is based on public, private, or hybrid, the management and control of these performance parameters from RAS viewpoint will vary.

Manuscript received on March 14, 2011. Accepted July 16, 2011.
Farzad Sabahi is with the Azad University, School of Computer and Electrical engineering, Iran. (e-mail: f.sabahi@ieee.org).

This paper is organized as follows:

- Section 2 provides a general overview of cloud computing.
- Section 3 describes the virtualization technology that is the basis of cloud computing.
- Section 4 overviews information security policies in cloud computing.
- Section 5 comprehensively reviews the RAS factor in virtualization.
- Section 6 elaborates on the RAS factor with particular attention to cloud computing.
- Section 7 covers intrusion detection systems in cloud computing.
- Section 8 describes security management and countermeasures to take against intrusions.
- Finally, section 9 concludes the paper.

II. Cloud Computing: An Overview

Cloud computing is a network-based environment that focuses on sharing computations and resources. Clouds are Internet-based and try to reduce complexity for clients by allowing them to virtually store data, applications and technologies at a remote site rather than keeping voluminous amounts of information on personal computers or on local servers. This is accomplished using virtualization technologies in combination with self-service abilities for computing resources via network infrastructure, especially the Internet. In cloud environments, multiple virtual machines are hosted on the same physical server as infrastructure. Customers only pay for what they use and avoid having to pay for local resources such as storage and infrastructure. Cloud computing, then, ultimately refers to both applications delivered as services over the Internet, and the hardware and systems software in the datacenters that provide those services. Currently, three types of cloud environments exist: public, private, and hybrid.

A public cloud is a standard model in which providers make several resources such as applications and storage available to the public. Public cloud services may be free, or may come with an associated fee. In public cloud environments, applications are run externally by large service providers, offering some benefits over private cloud environments.

For a private cloud, a business has internal services that are not available to other people. Essentially, the term "private clouds" is a marketing term for an architecture that provides hosted services for a particular group of people behind a firewall.

A hybrid cloud is an environment in which a company provides and controls some resources internally and provides other services for public use. In this type, the cloud provider has a service whereby a private cloud can be created (and is only accessible by internal staff; it would be protected by firewalls from outside access), and a public cloud environment for access by external users is also created.

Cloud is a style of computing where massively scalable and flexible IT-related abilities are provided "as services" to external customers using Internet technologies. Cloud providers offer various services in a XaaS collection that can offer the following main services.

A. SaaS

SaaS (Software as a Service): Software as a Service is a well-known service that offers network-hosted applications. SaaS is a software application delivery model by which cloud providers develop web-based software applications and then host and operate those applications over the infrastructure (usually the Internet) for use by their customers. As a result, cloud customers do not need to buy software licenses or additional equipment and they typically only pay fees (also referred to as annuity payments) periodically to use the cloud provider's web-based software [3]. There are two major kinds of SaaS: business applications that offer software which helps various businesses perform their tasks quickly and accurately. Other type of SaaS is development tools which consist of software that is used mainly for product development and management.

B. PaaS

PaaS (Platform as a Service): In this category of service, cloud users are given a platform [4]. They can use it as their application platform independent of using their own local machine for installing those platforms.

C. IaaS

IaaS (Infrastructure as a Service): Infrastructure as a Service is a provision model in which an organization outsources the equipment used to support operations, including storage, hardware, servers and networking components. The service provider owns the equipment and is responsible for housing, running, and maintaining it. The client typically pays on a per-use basis [5]. IaaS is sometimes referred to as Hardware as a Service (HaaS).

D. Other Types of Services

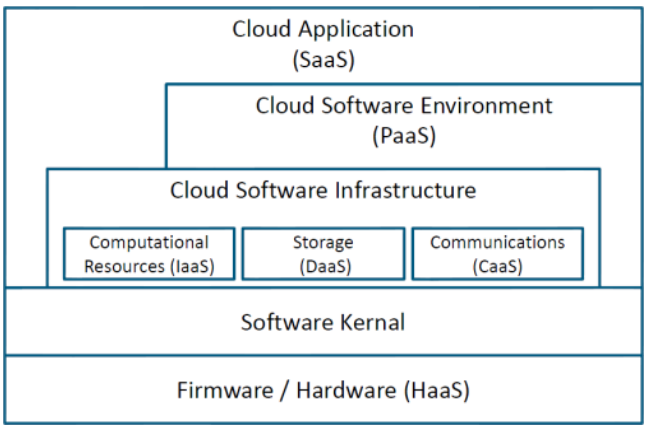

Fig. 1: Unified ontology of cloud computing.

It is important to mention that IaaS, PaaS, and SaaS are the three main categories of cloud computing services and that

the other types of cloud services are subsidiary branches of these three major categories. A typical cloud computing ontology for some of these categories is illustrated in Figure 1.

Other cloud services include the following:

- DaaS (Database as a Service): Database systems provide a user friendly interface for accessing and managing data. This type of service is very useful like many financial, business, and Internet-based applications [6].
- NaaS (Network as a Service): With NaaS, providers offer customers a virtualized network [7].
- IPMaaS (Identity and Policy Management as a Service): With this service, providers deliver identity and policy management to customers [8].

E. Cloud customers

Nowadays, many IT-related clients decide to use cloud computing for their own purposes. These can be divided into three main groups: regular customers, academics, and enterprises.

1) Regular customers

This group of users merely uses the services from the cloud [1]. They are not concerned with high performance; rather, they concentrate on the service and the privacy of their data on the cloud. SaaS is the most appropriate service for this group [9].

2) Academics

Academics usually have good networks and they often prefer to use the infrastructure that they already have to improve the performance of computations and resolve grid limits. For this group, cloud computing provides convenient access to a high-performance cluster or grid-based computation infrastructure and eliminates the need to buy new hardware.

3) Enterprises

The IT industry reaps the most considerable benefits of cloud computing [10]. Many companies have decided to enter cloud-related industries or use cloud services to reduce costs and improve performance in their own (IT-related or non-IT-related) businesses.

a) Small and mid-size enterprises

Lower costs are attractive, particularly for small enterprises that simply cannot afford the cost of solutions [4]. With distributed processing, small enterprises can afford industry-standard PCs and network servers but not expensive supercomputers. In addition, they can use cloud software instead of local software or abstruse infrastructure, which can reduce the cost of purchasing and maintaining the required software. For mid-size businesses that are growing, cloud computing can also provide a cost-effective and efficient path to enterprise-grade software and infrastructure [11].

b) Large-scale enterprises

For these enterprises, lower costs are not as important as privacy. Thus, large companies often create their own clouds or are skeptical about moving to the cloud. However, privacy of information is the most important issue, and most large companies have already spent significant amounts of money on their local systems [1]. Nowadays, large-scale enterprises often collect and analyze large amounts of data to derive business insights. However, there are at least two challenges to meet the increasing demand. First, the growth in the amount of data far surpasses the growth in the computation power of uniprocessors [12]. The growing gap between the supply and demand of computation power forces enterprises to parallelize their application codes. Unfortunately, parallel programming is both time-consuming and error-prone. Second, the emerging cloud computing pattern imposes constraints on the underlying infrastructure, which forces enterprises to rethink their application architectures.

III. VIRTUALIZATION

Virtualization is one of the most important elements of cloud computing. It is a technology that helps IT organizations optimizes their application performance in a cost-effective manner, but it can also present application delivery challenges that cause security difficulties. Most of the current interest in virtualization revolves around virtual servers, in part because virtualizing servers can result in significant cost savings. The phrase virtual machine refers to a software computer that, like a physical computer, runs an operating system and applications. An operating system on a virtual machine is called a guest operating system. A layer called a VMM (virtual machine monitor), or hypervisor, creates and controls the virtual machine's other virtual systems. Figure 2 illustrates a typical virtual machine architecture foundation in a cloud environment.

A. Hypervisor

A hypervisor (see Figure 2) is one of many virtualization techniques allowing multiple operating systems, termed guests, to run concurrently on a host computer using a feature called hardware virtualization. It is so named because it is conceptually one level higher than a supervisor.

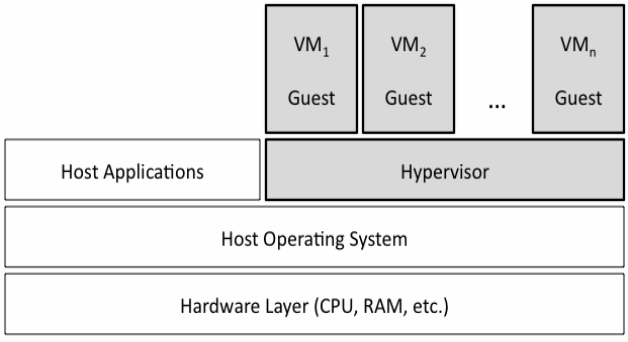

Fig. 2. Typical Virtual Machine architecture [1].

The hypervisor presents a virtual operating platform to the guest operating systems and also monitors the execution of them. Multiple instances of a variety of operating systems may share the virtualized hardware resources. Hypervisors are installed on server hardware dedicated to run guest operating systems [13].

IV. Information Security Policies

Cloud computing raises a range of important policy issues which include issues of privacy, security, anonymity, telecommunications capacity, government surveillance, reliability, and liability, among others [1]. But the most important of these issues is; security and how it is assured by the cloud provider. In addition, according to this fact that security effect on computing performance, cloud providers have to find a way to combine security and performance. For example for enterprises, the most important problem is security and privacy because they may store their sensitive data in cloud. For them, high performance processing may not be as critical as for academia users. To satisfy enterprise needs, the cloud provider has to ensure robust security and privacy more than other needs.

In cloud there are several security and privacy issues but in [14] there are the Gartner's seven well-known security issues which cloud clients should advert are listed below:

- **Privileged user access:** Sensitive data processed outside the enterprise brings with it an inherent level of risk because outsourced services bypass the "physical, logical and personnel controls" IT shops exert over in-house programs.
- **Regulatory compliance**: Customers are ultimately responsible for the security and integrity of their own data, even when it is held by a service provider. Traditional service providers are subjected to external audits and security certifications. Cloud computing providers who refuse to undergo this scrutiny are "signaling that customers can only use them for the most trivial functions," according to Gartner.
- **Data location**: When clients use the cloud, they probably will not know exactly where their data is hosted. Distributed data storage is usually used by cloud providers, but this can cause lack of control and is not good for customers who have their data in a local machine before moving to the cloud.
- **Data segregation:** Data in the cloud typically exists in a shared environment alongside data from other customers. Encryption is effective but is not a cure-all. Encryption and decryption is a classic way to cover security issues, but heretofore it could not ensure a perfect solution. While it is difficult to assure data segregation, customers must review the selected cloud's architecture to ensure data segregation is properly designed and available but without data leakage. Although data leakage has solution technology that named DLP.
- **Recovery:** If a failure occurs with the cloud, it is critical to completely restore client data. As clients prefer not to let a third-party control their data, this will cause an impasse in security policy in these challenging situations.
- **Investigative support:** Cloud services are especially difficult to investigate because logging and data for multiple customers may be co-located and spread across an ever-changing set of hosts and data centers.
- **Long-term viability:** Ideally, a cloud computing provider will never go bankrupt or be acquired by a larger company with new policies. However, clients must be sure that their data will remain available even after such an event.

V. Virtualization RAS Issues

In a traditional environment consisting of physical servers connected by a physical switch, IT organizations can get detailed management information about the traffic between the servers and the physical switch. Unfortunately, that level of information management is not typically provided by a virtual switch. In such a scenario the virtual switch has links from the physical switch via the physical NIC (Network Interface Card) attached to virtual machines. The resultant is lack of visibility into the traffic flows between and among the Virtual Machines on the same physical level affects security and performance surveying.

A potential problem also exists for virtualization when a provider combines too many Virtual Machines onto a physical server. This can result in performance problems caused by impact factors such as limited CPU cycles or I/O bottlenecks [15]. These problems can occur in a traditional physical server, but they are more likely to occur in a virtualized server because a single physical server is connected to multiple Virtual Machines all competing for critical resources.

Therefore, management tasks such as performance management and capacity planning management are more critical in a virtualized environment than in a similar physical environment. This means that IT organizations must be able to continuously monitor the real-time utilization of both physical servers and Virtual Machines. This capability allows users to avoid both over and underutilization of server resources. In addition, they will able to reallocate resources based on changing business requirements. This capability also enables IT organizations to implement policy-based remediation that helps them to ensure that their desired service levels are being met [16].

Another challenge with virtualization is cloud organization management of virtual machines sprawl [17]. In virtualized environment with virtual machine Sprawl, the number of virtual Machines running in it increases because of unnecessary new virtual Machines created rather than business necessity. Virtual machine sprawl concerns include the overuse of infrastructure. To prevent Virtual machine sprawl, a Virtual machine manager should carefully analyze the need for all new Virtual Machines and ensure that unnecessary Virtual machines migrate to other physical

servers. In addition, by migration, an unnecessary virtual machine will be able to move from one physical server to another with high availability and energy efficiency. Determination of the virtual machine destination can be challenging; it is necessary to ensure that a migrated Virtual machine keeps the same security, QoS configurations and needed privacy policies. On the other hand, the destination must assure that all the required configurations of the migrated virtual machine are kept.

A. Virtual machine security and threats

As illustrated in Figure 2, there are at least two levels of virtualization which are virtual machines and the hypervisor. Virtualization technique which is used in the virtual machines is not as new technology. Unfortunately, it has several security issues which are now migrated to cloud technology and they are not good heritage for cloud. There are also other vulnerabilities and security issues which are exclusive or may have a more critical role in the cloud environment.

As mentioned before, in the hypervisor, all users see their systems as self-contained computers isolated from other users, even though every user is served by the same machine. In this context, a virtual machine is an operating system that is managed by an underlying control program.

Hence there are various threats and attacks in this level, but some of them are important than others that mentioned below:

- **Virtual machine-level attacks:** The hypervisor or virtual machine technology used by cloud vendors are potential problems in multi-tenant architectures [18]. These technologies involve "virtual machines," remote versions of traditional on-site computer systems, including the hardware and operating system. The number of these virtual machines can be expanded or contracted on the fly to meet demand, creating tremendous efficiencies [19].
- **Cloud-provider vulnerabilities:** These could be platform-level vulnerabilities, such as SQL-injections, or cross-site scripting vulnerabilities that exist in the cloud service layer and cause insecure environments.
- **Expanded network-attack surface:** The cloud user must protect the infrastructure used to connect and interact with the cloud, a task complicated by the cloud being outside the firewall in many cases [4].
- **Authentication and authorization:** The enterprise authentication and authorization framework does not naturally extend into the cloud. Enterprises have to merge cloud security policies with their own security metrics and policies.
- **Availability of the cloud provider:** Cloud providers guarantee that their servers' uptime compares well with cloud users' own data centers and cloud providers ensure the clients which providers can handle their applications. An enterprise must be assured that a cloud provider is faithfully running a hosted application and delivering valid results [4]. Scheduled and unscheduled maintenance is another availability factor that exists and it can harm the availability ratio of the cloud provider. Although regularly scheduled maintenance does not count as downtime, unscheduled maintenance increases downtime and affects availability [20].
- **Lock-in:** There seems to be a great deal of anxiety regarding lock-in in cloud computing. The cloud provider can encrypt user data in a particular format if a user decides to migrate to another vendor or a similar situation arises [21].
- **Data control in cloud:** For mid-size businesses used to having complete visibility and control over their entire IT portfolio, moving even some components into the cloud can create operational "blind spots," with little advance warning of degraded or interrupted service [11].

B. Hypervisor Security

In a virtualization environment, there are several virtual machines that have independent security zones that are not accessible from other virtual machines that have their own zones. In a virtualization environment, a hypervisor has its own security zone and is the controlling agent for everything within the virtualization host. A hypervisor can touch and affect all of the virtual machine's actions running within the virtualization host [22]. There are multiple security zones, but these security zones exist within the same physical infrastructure that, in a more traditional sense, generally only exists within a single security zone. This can cause security issues, as if an attacker is able to take control of a hypervisor, then the attacker has full control of all the works within the territory of the hypervisor. Another major virtualization security concern is "escaping the virtual machine" or being able to reach the hypervisor at the virtual-machine level. This will become an even greater concern in the future as more APIs (Application Program Interface) are created for virtualization platforms [23]. Thereupon, so undamaged controls are to disable the functionality within a virtual machine, and this can reduce performance and availability.

1) Confronting against hypervisor security problems

As mentioned before, hypervisors are management tools, and the main goal of creating this security zone is building a trust zone. Other available virtual machines are under the approval of the hypervisor, and they can rely on it, as users are trusting that administrators of system will do what they can to do tasks properly. As for security characteristics, there are three major levels in the security management of hypervisors:

- **Authentication:** Users have to authenticate their account properly using the appropriate standard and available mechanisms.
- **Authorization:** Users must receive authorization, and they must have permission to do what they are trying to do.
- **Networking:** Using mechanisms that assure a secure connection to communicate by using available

administration applications that most likely launch and work in a different security zone than that of users.

Authentication and authorization are some of the most interesting auditing aspects of management because there are so many methods available to manage a virtual host auditing purpose [24]. The general belief is that networking is the most important issue in transactions between users and the hypervisor, but there is much more to virtualization security than just networking. Networking plays a critical role in security, but it is not solely significant for ensuring security. It is just as important to understand the APIs and basic concepts of available hypervisors and virtual machines, and how those management tools work [22]. If a security manager can address authentication, authorization, virtual hardware, and hypervisor security as well as networking security, cloud clients are well on the way to a comprehensive security policy [1, 22]. If a cloud provider at the virtualization level does not, or just depends on network security to do the tasks, then the implemented virtual environment is at risk and has poor security capability. It is a waste of money if a cloud provider spends too much money on creating a robust secure network and neglects communication among virtual machines and the hypervisor, as this can cause several problems for the provider as well as for the users.

VI. CLOUD RAS ISSUES

Using cloud means, that applications and data will move under a third-party control. The cloud services delivery model will create clouds with virtual perimeters as well as a security model with responsibilities shared between the customer and the cloud service provider. This shared-responsibility model will bring new security management challenges to the organization's IT operations staff [25]. Basically, the first question an information security officer must answer is whether he/she has adequate transparency with cloud services to manage the governance (shared responsibilities) and implementation of security management processes (preventive and detective controls) to ensure the business that the data in the cloud is appropriately protected. The answer to this question consists of two parts: what security controls must the customer provide over and above the controls inherent in the cloud platform and how should an enterprise's security management tools and processes adapt to manage security in the cloud. Both answers must be continually reevaluated based on the sensitivity of the data and the service-level changes over time [25].

A. Data Leakage

Basically, when moving to a cloud, there are two changes for customers' data. First, the data will be stored away from the customer's locale machine. Second, the data is moved from a single-tenant to a multitenant environment. These changes may raise an important concern called, *data leakage*. This has become one of the greatest organizational risks from the security standpoint [26]. Virtually every government worldwide has regulations that mandate protections for certain data types [26]. The cloud provider should have the ability to map its policy to the security mandate the user must comply with and discuss the issues.

1) DLP

Nowadays, there is an interest in the use of data leakage prevention (DLP) applications to protect sensitive data with the appearance of cloud computing. To prevent data leakage, some companies have thought of DLP products. DLP products existed before cloud computing. These products aim to ensure data confidentiality and detect unauthorized access to data, but they are not intended to be used for ensuring the integrity or availability of data. As a result, experts don't expect from DLP products to address data's integrity or availability in any cloud model.

If data is stored in a public cloud, because of its nature, using DLP products is worthless to protect the confidentiality of that data in all types of clouds. Generally in SaaS and PaaS, because cloud clients do not have control over security management used by the cloud provider, discovery of the client's data with DLP agents is not possible except when the provider puts this capacity into its service. However, it is possible by embedding DLP agents into virtual machines in IaaS to achieve some control over associated data.

In private clouds, the customer has direct control over the whole infrastructure; it is not a policy issue whether DLP agents are deployed in connection with SaaS, PaaS, or IaaS services. However, it may very well be a technical issue whether DLP agents interoperate with SaaS or PaaS services as architected [27]. In a hybrid cloud, if service is IaaS, the client could embed DLP agents for some control over data.

B. Privacy

Cloud clients' data stores in data centers that cloud providers diffuse all over the globe within hundreds of servers that communicate through the Internet have several well-known potential risks within them. Because cloud services are using the Internet as their communication infrastructure, cloud computing involves several kinds of security risks [26]. Cloud providers, especially IaaS providers, offer their customers the illusion of unlimited computer, network, and storage capacity, often coupled with a frictionless registration process that allows anyone to begin using cloud services [28]. The relative anonymity of these usage models encourages spammers, malicious users and other hackers, who have been able to conduct their activities with relative impunity [29]. PaaS providers have traditionally suffered most from such attacks; however, recent evidence shows the hackers have begun to target IaaS vendors as well [28].

As is clear in cloud-based services, a user's data is stored on the third-party's storage location [1]. A service provider must implement sufficient security measures to ensure data privacy. Generally, data encryption is a solution to ensure the privacy of the data in the databases against malicious attacks. Therefore, encryption methods have significant performance

implications on query processing in clouds. Integration of data encryption with data is useful to protect the user's data against outside malicious attacks and to restrict the liability of the service provider.

It seems protection from malicious users who might access the service provider's system is the final goal, but this is not enough when clients also prefer privacy protection from a accessing to their data by provider. Any data privacy solution will have to use particular encryption, but this causes another availability issue: data recovery [30]. Assume a user's data is encrypted with a user-known key, and the user loses his/her key. How can the provider recover his/her data when it doesn't know what the key is? If the user gives the provider authority to know the key, then this makes privacy by using a user-known encryption key useless. The simple way to solve this problem is to find a cloud provider which users can trust. This way is acceptable when data stored in the cloud is not very important. This method seems useful for enterprises with the maximum size of a small company which may decide to find trustable providers rather than finding a solution for the data recovery problem. For medium-sized companies to large-sized companies, it is more critical to develop techniques and methods that enable query processing directly over encrypted data to ensure privacy from cloud providers [30]. If the service providers themselves are not trusted, protecting the privacy of users' data is a much more challenging issue. However, for those companies it seems using a private cloud is a wise solution.

If data encryption is used as a wise solution for the data privacy problem, there are other issues in this context. One of the most important issues is ensuring the integrity of the data. Both malicious and non-malicious users can cause compromise of the integrity of the users' data when this happens and the client does not have any mechanism to analyze the integrity of the original data. Hence, new techniques have to be applied to provide methods to check the integrity of users' data hosted at the service provider side [8].

All encryption methods rely on secure and impressive key management architectures. One of the problems that can occur in an encrypted environment is encryption key management in the cloud. In the cloud environment several users may use their own encryption method, and managing these keys is another issue to address in the context of encrypted data. For example, if the cloud provides database service (DaaS), the cloud provider faces more challenges in key management architectures, such as generation, registration, storage, and update of encryption keys.

1) RAS issues in Database-based service: An example

Cloud systems provide an extremely attractive interface for managing and accessing data and have proven to be widely successful in many financial, business and Internet applications. However, they have several serious limitations in database-based service such as the following which are mentioned in [6]:

- **Database systems are difficult to scale:** Most database systems have hard limits beyond which they do not easily scale. Once users reach these scalability limits, time-consuming and expensive manual partitioning, data migration and load balancing are the only recourse.
- **Database systems are difficult to configure and maintain:** Administrative costs can easily account to a significant fraction of the total cost of ownership of a database system. Furthermore, it is extremely difficult for untrained professionals to get good performance out of most commercial systems.
- **Diversification in available systems complicates selection:** The rise of specialized database systems for specific markets complicates system selection, especially for customers whose workloads do not neatly fall into one category.
- **Peak provisioning leads to unnecessary costs:** Database workloads are often tandem in nature hence they provision for the peak often results in an excess of resources during off-peak phases and thus causes unnecessary costs.

C. Data Remanance

Data remanence is the residual physical representation of data that has been in some way erased. After storage media is erased, there may be some physical characteristics that allow data to be reconstructed [31]. As a result, any critical data must not only be protected against unauthorized access, but also it is very important that it is securely erased at the end of the data life cycle. Basically, IT organizations that have full control of their own servers use various available tools that give them the ability to destroy unwanted and important data for privacy and safety purposes. But when data is migrated to a cloud environment, they now have virtual servers that are controlled by a third party.

As a solution, IT organizations must choose cloud providers that can guarantee that all customer erased data is erased immediately and securely. A traditional solution to deleting data securely is overwriting, but this technique does not work without the collaboration of the cloud provider [4, 30]. In a cloud environment, customers can't access the physical device or the data level. Thus, there is only one solution: those customers encrypt their data with a confidential key that prevents reconstruction of the erased data from residual data.

D. Cloud Security Issues

As mentioned before, the Internet is the communication

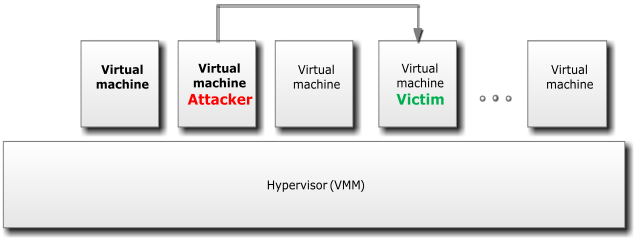

Fig. 3. Attack scenario within cloud.

infrastructure for cloud providers that use the well-known TCP/IP protocol, which uses IP addresses to identify Internet users. Similar to a physical computer in the Internet which has an IP address, a virtual machine in the Internet also has an IP address. A malicious user, whether internal or external, like a legal user who exists in network, can find these IP addresses as well. In this case, a malicious user can find out which physical servers the victim is using, and implant a malicious virtual machine at that location from which to launch an attack [28]. Because all users use the same infrastructure as the virtual machine, if a hacker steals a virtual machine or takes control of it, he also inherits the data within it. The hacker can then copy the data into his/her local machine before the cloud provider detects that the virtual machine is out of control; then the hacker can analyze the data, and may find valuable data afterward.

1) Attacks in cloud

Nowadays, there are several kinds of attacks in the IT world. Basically, the cloud can give service to legal users, but it can also give service to users who have malicious purposes. A hacker can use a cloud to host a malicious application to achieve a task, which may be a DDoS (Distributed Denial of Service) attack against the cloud itself, or arranging an attack against another user in the cloud. For example, an attacker knows that his victim [30]. This situation is similar to this scenario in that both the attacker and the victim are in the same network, but with the difference that they use virtual machines instead of a physical network (Figure 3).

a) DDoS Attacks Against Cloud

DDoS attacks typically focus a high quantity of IP packets at specific network entry elements; usually any form of hardware that operates on a Blacklist pattern is quickly overrun. In cloud computing, where the infrastructure is shared by a large number of clients, DDoS attacks have the potential of much greater impact than they do against single-tenant architectures. If the cloud does not have sufficient resources to provide services to its customers, the cause may be undesirable DDoS attacks [30]. The traditional solution for this event is to increase the number of such critical resources. But a serious problem occurs when a malicious user deliberately performs a DDoS attack using bot-nets.

Most network countermeasures cannot protect against DDoS attacks, because they cannot stop the deluge of traffic, and typically cannot distinguish good traffic from bad traffic. IPS (Intrusion Prevention Systems) are effective if the attacks are identified and have pre-existing signatures, but are ineffective if there is legitimate content with bad intentions [27]. Unfortunately, similar to IPS solutions, firewalls are vulnerable and inefficient against DDoS attacks because an attacker can easily bypass firewalls and IPSs, because they are designed to transmit legitimate traffic, and attacks generate so much legitimate like traffic from so many distinct hosts that a server, or a cloud's Internet connection, cannot handle the traffic [27].

It may be more accurate to say that DDoS protection is part of the Network Virtualization layer rather than Server Virtualization. For example, cloud systems that use virtual machines can be overcome by ARP spoofing at the network layer; DDoS protection is really about how to layer security across multivendor networks, firewalls, and load balances [32].

b) Cloud against DDoS attacks

DDoS attacks are one of the most powerful threats available in the world, especially when launched from a botnet with huge numbers of zombie machines. When a DDoS attack is launched, it sends a heavy flood of packets at a Web server from multiple sources. The cloud may be part of the solution; it's interesting to consider that websites experiencing DDoS attacks, which have limitations in server resources, can take advantage of using a cloud that provides more resources to tolerate such attacks [30]. Cloud technology also offers the benefit of flexibility, with the ability to provide resources almost real-time as necessary and almost instantaneously to avoid site shutdown.

VII. Intrusion Detection in Cloud

As we know, IDS have been used widely to detect malicious behaviors in several types of network. IDS management is an important capability for distributed IDS solutions, which makes it possible to integrate and handle different types of sensors or collect and synthesize alerts generated from multiple hosts located in the distributed environment. Facing new application scenarios in Cloud Computing, the IDS approaches yield several problems since the operator of the IDS should be the user, not the administrator of the Cloud infrastructure. Extensibility, efficient management and compatibility with virtualization-based contexts need to be introduced into many existing IDS implementations. Additionally, the Cloud providers need to enable possibilities to deploy and configure IDS for the user. Within this paper, we summarize several requirements for deploying IDS in the Cloud and propose an extensible IDS architecture that is easily used in a distributed cloud infrastructure [33, 34].

A. Intrusion detection at service level

1) IDS in SaaS

Attacks on networks are a reality in the world. Detecting and responding to those attacks is considered due diligence. The reality is that in SaaS, users will have no choice except to trust their providers to perform Intrusion Detection properly. Some providers give their users the option of getting some system logs and users can use custom application for monitoring those data, but in reality, most Intrusion Detection activities must be done by the provider and the user can only report suspicious behavior for analysis.

2) IDS in PaaS

In PaaS, similarly to SaaS, most of the Intrusion Detection activities must be done by the Cloud provider but with a little

difference. If Intrusion Detection systems are outside of the users' application, they have no choice and must rely on the provider to implement IDS. But PaaS configuration is more flexible than SaaS and users may have the choice to configure the security parameters of platforms that log on to a centralized place and users can incorporate Intrusion Detection performance [34, 35].

3) IDS in IaaS

IaaS is the most flexible service for Intrusion Detection implementation. But the most important challenge in constructing a secure cloud-computing infrastructure is Transparency. Without it, the user cannot know if the cloud provider meets significant security requirements or not. Moreover, the user cannot properly design application architecture to mitigate any risks that may exist.

B. Intrusion Detection Placement

For operating Intrusion Detection in the Cloud properly, the user must identify the possible and also proper places for hosting IDS. In the traditional network, using Intrusion Detection allows the user to monitor, detect and alert about traffic that passes over the traditional network infrastructure [9, 36]. Generally, there are some places in the network with more traffic than in other place (hotspots). Placement of IDS in the physical network part of the Cloud is similar to a traditional network, because hotspots in both of them are the same.

1) In the virtual machine and network layer

Using Intrusion Detection in the virtual machine layer allows the user to monitor the system and detect and alert about issues that may arise. In addition, using Intrusion Detection to monitor the virtual network allows the user to monitor the network traffic between the virtual machines on the host, as well as the traffic between the virtual machines and the host. It should be noted that this network is different from traditional networks and that traffic never hits it [35].

2) In the Hypervisor layer

As said before, the hypervisor presents to the guest operating systems a virtual operating platform and monitors how the guest operating systems are running [8]. Deploying Intrusion Detection in the hypervisor allows the user to monitor everything that passes between the virtual machines.

As illustrated in Figure 4, the HyIDS runs inside the hypervisor. Because the hypervisor interposes on all accesses between the guest kernel and the hardware, ISIS can monitor all operating system events and data structures for intrusions.

C. Intrusion Detection Techniques Performance

As is well known, Intrusion Detection has three well-known main groups: Host-based, Network-based, and hybrid. This section discusses performance issues of Intrusion Detection techniques in the Cloud, some of which are traditional solutions and some of which are special rectification solutions for use within the Cloud.

1) Traditional IDS solutions in cloud

a) Host-based intrusion detection

The first choice in Intrusion Detection is the traditional HIDS (Host-based Intrusion Detection), which examines events and transmissions such as what file was accessed and what application was executed. This type of IDS can be used on virtual machines as well as in the hypervisor level of Cloud environments. Using Intrusion Detection in the virtual machine layer allows the user to monitor the activity of the system and detect and alert about issues that may arise. At this level, the user can use an HIDS and have control over it. This type of IDS can detect intrusion against his/her Virtual machine. The provider may also deploy an HIDS in the hypervisor layer but only the provider is authorized to manage and configure it. In the hypervisor level, HIDS can also monitor traffic between virtual machines.

The HIDS on the virtual machine would be used by the user of the Cloud but the HIDS on the hypervisor level is for provider control; if the user wants to use the hypervisor Intrusion Detection data in his independent IDS, he would have to coordinate with the provider. This issue is likely to pose difficulties because most Cloud providers prefer not to share such data with customers due to privacy policies [26, 34]. While HIDS on the hypervisor level would be under the responsibility of the Cloud provider, deploying and managing an HIDS on the virtual machine would be the user's responsibility.

b) Network-based intrusion detection

A Network Intrusion Detection System (NIDS) is another traditional solution for performing security policies in computer networks. NIDSs work by examining network traffic but with this characteristic, only the cloud provider can deploy it. Unfortunately, in cloud, because of the nature of NIDS, this type of IDS has limitations. For example, it is unable to detect attacks within a virtual network that runs completely within the hypervisor. Also, NIDS is useless in encrypted environments. This type of placement of IDS is useful in detecting some attacks on the VMs and hypervisor but it does have three important constraints. The first is that it is not useful when it comes to malicious activities within a VM, which is fulfilled completely in the hypervisor level. Secondly, it has limited visibility into the host itself. Thirdly, if the network traffic is encrypted by users, NIDS cannot decrypt the traffic for analysis [20, 34]. Even if NIDS has all encryption keys used in the Cloud, NIDS needs more computation resources to perform the decrypting. Moreover, analyzing these data results in an increased cost of detection.

2) Performance of traditional IDS

It seems that NIDS works better than HIDS but it must be considered that HIDSs are easy to implement while NIDS are difficult or at times impossible to fulfill in the Cloud environment. In addition, in the Cloud, NIDS falls completely into the area of the provider to operate and control. This paper

has shown that Cloud users need to think more about moving toward the Cloud and also that Cloud providers should give more attention to security matters.

3) Hypervisor-based intrusion detection system

Another Intrusion Detection method is to use IDS, which launches at the hypervisor layer but is not strictly a HIDS for the hypervisor, which is called Hypervisor-based IDS (HyIDS) [34] or ISIS IDS (Intrusion Sensing and Introspection System) [36]. One of the promising technologies in this method is the use of VM introspection. This type of IDS allows users to monitor and analyze communications between VMs, between the hypervisor and VM and within the hypervisor-based virtual network. The advantage of the hypervisor-based ID is the availability of information, as it can see everything. The disadvantage is that the technology is new and users really need to know what they are looking for [21, 34]. There is a special type of Intrusion Detection in the hypervisor because of the level of accessing it contains and it has a good potential for improving the performance of intrusion detecting. As illustrated in Figure 4, the HyIDS runs inside the hypervisor. The hypervisor can interpose on all accesses between the guest VM kernel and the hardware, while HyIDS can monitor all operating system events and data structures for intrusions. Like NIDS, control and implementation of HyIDS is done entirely by the Cloud provider [34, 37].

VIII. COUNTERMEASURES

There are several traditional solutions to mitigate security problems that exist in the Internet environment and the cloud infrastructure, but the nature of clouds causes some security problems that exist especially in cloud environments. On the other hand, there are traditional countermeasures against popular Internet security problems that may be usable in clouds, but some of them must be improved or changed to use in cloud environments.

A. Access Control

To ensure the accessibility of authorized users the prevention of unauthorized access to information systems, formal procedures should be in place to control the allocation of access rights to services. The procedures should cover all stages in the lifecycle of user access, from the initial registration of new users to the final de-registration of users who no longer require access to information systems and services. Special attention should be paid, where appropriate, to the necessity to control the allocation of privileged access rights, which allow users to override system controls [38]. The following are the six control statements [38]:

- Control access to information.
- Manage user access rights.
- Encourage good access practices.
- Control access to network services.
- Control access to operating systems.
- Control access to applications and systems.

In the SaaS model, the cloud provider is responsible for managing all aspects of the network, server, and application infrastructure. In that model, since the application is delivered as a service to end users, usually via a web browser, network-based controls are becoming less relevant and are augmented or superseded by user access controls, e.g., authentication using a one-time password [27, 38]. Hence, customers should focus on user access controls (authentication, federation, privilege management, provisioning, etc.) to protect the information hosted by SaaS [39].

In the PaaS delivery model, the Cloud provider is responsible for managing access control to the network, servers, and application platform infrastructure. However, the customer is responsible for access control to the applications deployed on a PaaS platform. Access control to applications manifests as end user access management, which includes provisioning and authentication of users [28].

IaaS customers are entirely responsible for managing all aspects of access control to their resources in the cloud. Access to the virtual servers, virtual network, virtual storage, and applications hosted on an IaaS platform will have to be designed and managed by the customer. In an IaaS delivery model, access control management falls into one of the following two categories. Access control management to the host, network, and management applications that are owned and managed by the Cloud provider and user must manage access control to his/her virtual server, virtual storage, virtual networks, and applications hosted on virtual servers [30, 38].

B. Incident Countermeasure and response

One of the important issues in cloud security, similar to other IT fields, is finding problems and vulnerabilities that exist, but a more important issue is that the cloud provider has appropriate responses against all problems that it finds. Basically, the cloud systems are built on a collection of storage and process engines, driven by a configurable distributed transaction coordinator. To achieve some important parameters such as flexibility, scalability and efficient usage of resources, cloud providers must face major challenges in the area of adaptability and workload.

One of the main requirements of the cloud is the ability to be flexible; in the context of a cloud service, flexibility means dedicating resources where they are most needed [6]. This is particularly challenging in a database environment where there are large amounts of data that may need to be moved in order to reconcile data [6].

To allow high performance workloads to scale across multiple computing nodes, it is important for cloud provider to divide their data into partitions that maximize service performance. The main idea behind partitioning is to lessen the probability that a typical transaction has to access multiple nodes in cloud to compute its query.

In migration, available methods must be able to predict adaptation time and try to avoid cloud node overload by some

procedure, such as partitioning, fragmenting, breaking big data packets in smaller pieces, and maintaining the ability to execute transactions while movement occurs [36].

To balance workloads on virtual machines properly, it is necessary to analyze and classify cloud providers resource requirements to decide how these can be allocated to virtual machines.

C. Security Management in the Cloud

The relevance of various security management functions available for each cloud delivery model is dependent on the context of deployment models. As mentioned before in the introduction, there are several important parameters in cloud security management: availability management, access control management, vulnerability and problem management, patch and configuration management, countermeasure response, and cloud system use and access monitoring. Thus, according to the type of service provided, the customer or the provider must manage some or all of them independently, or perhaps partially [30]. Thus, if a cloud is a private cloud, then the cloud provider generally manages all mentioned functions. But if a cloud is a public or hybrid cloud, then who manages which aspect depends on the type of cloud and the service provided. For example, if a cloud is SaaS, then the customer must partially manage access control and monitor system use and access, and also must manage incident response, and the cloud provider must manage the other functions. In other types of clouds (PaaS and IaaS), the functions are limited to customer applications deployed in PaaS or IaaS.

IX. Conclusion

As outlined in the article, cloud computing helps IT enterprises to optimize and secure application performance in a cost effective manner. Cloud-based applications are based on network software running on a virtual machine in a virtualized environment. In view of the vital role of the hypervisor in a virtualization system, security at this level of virtualization needs special consideration. Generally, a virtual application relieves some of the management issues in enterprises because most of the maintenance, software updates, configuration and other management tasks are automated and centralized at the cloud provider's datacenter. But this way for decentralized application and access creates its own set of challenges and security problems. There are, however, risks and hidden costs in managing cloud compliance. Cloud providers often have several powerful servers and resources that provide appropriate services for their users, but the cloud is at risk to a degree similar to that of other Internet-based technologies. Unfortunately, there are some attacks for which no perfect defense exists such as a powerful DoS attack. But as paper discussed in occurrence of DoS attacks, cloud may be a good solution or mitigation because cloud providers can use mirrors or devote more resources to protecting against attacks. However this solution's performance depends on provider facilities.

Issues introduced by this paper are the main reasons for the precaution exercised by many enterprises and even some ordinary users to the adaptation of cloud computing. But benefits of using cloud have caused some of enterprises to have a plan in which cloud computing is used for less-sensitive data, but they may have local machines to store data which are of greater sensitivity. Should cloud providers wish for clients to store greater amounts of sensitive data in the cloud computing environment, improving security (and also, of course, client perception of that security) is paramount.

Whilst cloud computing is an important trend that keeps transforming and will continue to transform the IT industry, it doesn't mean that all business IT needs should move to the cloud computing model. The key to successful cloud computing initiatives is achieving a balance between the business benefits and the hidden potential risks lurking on the path to implementation.

References

[1] "Securing Virtualization in Real-World Environments," White paper2009.

[2] P. Coffee, "Cloud Computing: More Than a Virtual Stack," ed: salesforce.com.

[3] *Software as a Service*. Available: http://www.wikinvest.com/concept/Software_as_a_Service

[4] G. Reese, *Cloud Application Architectures*: O'Reilly Media, Inc.,, 2009.

[5] N. Mirzaei, "Cloud Computing," 2008.

[6] "Database as a Service," MIT-CSAIL-TR-2010-014.

[7] M. Riccuiti. Stallman: Cloud computing is stupidity. Available: http://news.cnet.com/8301-1001_3-10054253-92.html

[8] N. Antonopoulos and L. Gillam, *Cloud Computing*: Springer-Verlag London Limited, 2010.

[9] K. JACKSON, "Secure Cloud Computing: An Architecture Ontology Approach," Defense Information Systems Agency2009.

[10] R. Raja and V. Verma, "Cloud computing: An overview," Research Consultant, IIIT Hyderabad.

[11] D. Rowe. (2011, The Impact of Cloud on Mid-size Businesses. Available: http://www.macquarietelecom.com/hosting/blog/cloud-computing/impact-cloudcomputing-midsize-businesses

[12] S. Hanna, "Cloud Computing: Finding the Silver Lining," Juniper Networks2009.

[13] *Cloud Computing*. Available: http://en.wikipedia.org/wiki/Cloud_computing

[14] J. Brodkin. (2008). *Gartner: Seven cloud-computing security risks*. Available: http://www.networkworld.com/news/2008/070208-cloud.html

[15] J. Metzler. (2009, Virtualisation can make application delivery much, much harder - but you can fight back! Available: http://searchnetworking.techtarget.com.au/articles/33471-Virtualisation-can-make-application-delivery-much-much-harder-but-you-can-fight-back-

[16] "Virtualization: The next generation of application delivery challenges."

[17] (2011). *What is Cloud Sprawl and Why should I Worry About It?* Available: http://www.cloudbusinessreview.com/2011/06/08/what-is-cloud-sprawl-and-why-should-i-worry-about-it.html

[18] R. Chow, P. Golle, M. Jakobsson, E. Shi, J. Staddon, R. Masouka, and J. Molina, "Controlling Data in the Cloud: Outsourcing Computation without Outsourcing Control," presented at the CCSW'09, Chicago, Illinois, USA., 2009.

[19] D. Talbot. (2009). *Vulnerability Seen in Amazon's Cloud-Computing*. Available:

http://www.technologyreview.com/printer_friendly_article.aspx?id=23792

[20] J. W. Rittinghouse and J. F. Ransome, *Cloud Computing Implementation, Management, and Security*: Taylor and Francis Group, LLC, 2010.

[21] P. Sefton, "Privacy and data control in the era of cloud computing."

[22] Texiwill. (2009). *Is Network Security the Major Component of Virtualization Security?* Available: http://www.virtualizationpractice.com/blog/?p=350

[23] D. E. Y. Sarna, *Implementing and Developing Cloud Computing Applications*: Taylor and Francis Group, LLC, 2011.

[24] t. Ristenpart and e. al, "Hey, you, get off of my cloud: exploring information leakage in third-party compute clouds," 2009.

[25] S. K. Tim Mather, and Shahed Latif, *Cloud Security and Privacy*: O'Reilly Media, Inc., 2009.

[26] C. Almond, "A Practical Guide to Cloud Computing Security," 2009.

[27] *Cloud Security*. Available: http://cloudsecurity.trendmicro.com/

[28] N. Mead, E. Hough, and T. Sehny, "Security quality requirements engineering (SQUARE) methodolgy," Carnegie Mellon Software Engineering Institute.

[29] K. K. Fletcher, "Cloud Security requirements analysis and security policy development using a high-order object-oriented modeling," Master of science, Computer Science, Missouri University of Science and Technology, 2010.

[30] F. Sabahi, "Analysis of Security in Cloud Environments," presented at the International Conference on Computer Science and Information Technology, Chengdu, China, 2011.

[31] P. R. Gallagher, *A Guide to Understanding Data Remanence in Automated Information Systems*: The Rainbow Books, 1991.

[32] (2009). *Cloud Computing*. Available: http://groups.google.com/group/cloud-computing/browse_thread/thread/21e585b137125554

[33] S. Roschke, F. Cheng, and C. Meinel, "Intrusion Detection in the Cloud," presented at the Eighth IEEE International Conference on Dependable, Autonomic and Secure Computing, Chengdu, China, 2009.

[34] F. Sabahi, "Intrusion Detection Techniques performance in Cloud Environments," presented at the International Conference on Computer Design and Engineering, Kuala Lumpur, Malaysia, 2011.

[35] P. Cox, "Intrusion detection in a cloud computing environment," 2010.

[36] L. Litty, "Hypervisor-based Intrusion Detection," Master of Science, 2005.

[37] L. Ponemon, "Security of Cloud Computing Users," 2010.

[38] (2010). *Security Management in the Cloud*. Available: http://mscerts.net/programming/Security%20Management%20in%20the%20Cloud.aspx

[39] (2010). *Security Management in the Cloud - Access Control*. Available: http://mscerts.net/programming/Security%20Management%20in%20the%20Cloud%20-%20Access%20Control.aspx