ISSN 2007-9737

A Survey on Information Security in Cloud Computing

Sandra Dinora Orantes Jiménez, Eleazar Aguirre Anaya

Instituto Politécnico Nacional,
Centro de Investigación en Computación,
Mexico

{dinora, eaguirre}@cic.ipn.mx

Abstract. Cloud computing is a type of Internet-based computing that provides different services, is a way to increase capacity or add capabilities dynamically without investing in new infrastructure, training new staff or licensing new software. Cloud computing has gained popularity by providing easy and cost-effective access to IT (Information Technology). However, although it expands existing IT capabilities and although it has gone from being a promising business concept, to one of the fast growing segments; as more and more information about individuals and companies is placed in the cloud, concerns begin to grow about how secure the environment is. Security is one of the main problems that reduces the growth of cloud computing and the complications with privacy and data protection continue to affect the market. Users of cloud services should be vigilant to understand the risks of data breaches in this new environment. This article presents a survey of the different security risks that represent a threat to the cloud due to the nature of the service delivery models of a cloud computing system. On the other hand, works carried out with respect to security are also explored, such as Vigenère, Playfair, Hill and Vernan encryption, proxy re-encryption and Railfence cipher. With this work we hope to provide useful information about cloud computing security.

Keywords. Cloud computing, security, data integrity, data privacy, data protection, virtualization, elgamal encryption.

1 Introduction

Cloud computing is the result of decades of research in virtualization, distributed computing, Grid computing, utility computing and also involves work in networks, web and software services. It implies a service-oriented architecture with a lower IT overhead for the end user. Today, Small and Medium Enterprises (SMEs) in Mexico and other parts of the world, are increasingly realizing that simply by accessing the cloud can get quick access to the best commercial applications or increase dramatically its infrastructure resources, all at an insignificant cost.

Gartner [4] defines cloud computing as “a massively scalable computing style and where IT-enabled capabilities are delivered 'as a service' to external customers who use Internet technologies.” On the other hand, providers of the cloud currently enjoy a deep opportunity in the market and should ensure that they have the right security aspects, to take responsibility for managing the information of companies and not suffer, if things go wrong. Benefits, such as rapid implementation, payment for use, lower costs, scalability, rapid provisioning, rapid elasticity, ubiquitous access to the network, greater resistance, hypervisor protection against network attacks, low cost disaster recovery and storage solutions data, on-demand security controls, real-time detection of system manipulation and rapid reconstitution of services. Although the cloud offers these advantages, until the risks are better understood, many of the main actors will be tempted to restrain themselves from using it [14].

Cloud computing moves the software of the application and the bases of data to large data centers, where for the moment, the management of data and services are not considered 100% reliable. However, this raises many new security challenges [11]. These challenges include, among others, accessibility vulnerabilities, virtualization vulnerabilities, web application vulnerabilities such as SQL (Structured Query Language) administration and cross-site scripting, physical access problems, privacy and control derived from

third parties that have physical control of data, issues related to identity and credential management, issues related to data verification, manipulation, integrity, confidentiality, loss and theft of data, points related to device authentication and property protection intellectual.

This paper aims to describe various security problems of cloud computing due to its models of service provision, as well as to explore work carried out with respect to security; where, the underlying technology of the cloud by itself provides a significant security risk. This document is organized as follows: Section 2 presents the similarity between Grid computing and the cloud and then, in section 3, the common security problems posed by cloud service delivery models are described, summarizing the threats to security presented by the 'Software as Service' (SaaS) model, 'Platform as a Service' (PaaS) and the 'Infrastructure as a service' model (IaaS). Section 4 explores some of the current solutions that address in part the security challenges posed by the cloud. Finally, conclusions derived from this study are provided.

2 Grid Computing vs. the Cloud

To make cloud computing work, three things are necessary: thin clients, Grid computing and computer services. Grid computing links disparate computers to create a single infrastructure, taking advantage of unused resources. Computer services are what you pay for, that is, shared servers, as well as paying for a public service (such as electricity, gas, etc.).

With Grid, computing it is possible to provide computing resources, such as services that can be turned on and off. Cloud computing goes a step further by providing resources on demand. This eliminates the need for over-acquisition of equipment to meet the demand of millions of users.

Cloud computing is an evolution of Grid computing and provides provisioning of resources on demand (see Figure 1). Grid computing may or may not be in the cloud, depending on the type of users that use it. If users are system administrators and integrators, they care how the cloud is maintained, updated, installed and virtualize servers and applications. If users are consumers, they are not interested in how the system works.

Grid computing requires the use of software that can divide and group pieces of a program as if it were a large system through thousands of computers. A problem with the grid is that if one piece of software in one node fails, others may fail as well. This can be mitigated if that component has some other backup in another node, but problems can still arise if the components need other pieces of software to accomplish one or more tasks in the grid. The images of large systems and the associated hardware to operate and maintain them can contribute to large acquisition costs and operating expenses [13].

2.1 Similarities and Differences

Cloud computing and Grid computing (see Figure 1) are scalable. Scalability is achieved through the load balancing of application instances, which are executed separately in different operating systems connected through Web Services. The CPU and the network width are reserved and released on demand. The storage capacity of the system increases and decreases depending on the number of users, instances and the amount of data that is transferred at any given time Both types of computing support multitasking, that is, many clients can perform different tasks, accessing one or multiple instances of applications.

By sharing resources for a large group of users, infrastructure costs and peak load capacity are reduced. Cloud and Grid computing have Service Level Agreements (SLA) to guarantee availability (for example, 99%). If the service is below the agreed level, the consumer will receive service credit for late receipt of data. For example, Amazon S3 has a web services interface for storing and retrieving data in a cloud. You can set a maximum limit of objects that are stored in S3. You can store objects as small as 1 byte and as large as 5GB or even several terabytes. S3 uses the concept of "buckets" as containers for each storage location of the objects. The data is stored securely using the same storage infrastructure that Amazon uses for its web e-commerce.

While grid storage works well for storing many data, it is not economically convenient to store objects as small as 1 byte.

Fig. 1. Cloud Computing vs. Grid

In data grids, the amount of data distributed must be large to maximize the benefit.

3 Common Security Issues Raised by Cloud Service Delivery Models

Cloud services can be identified according to software (SaaS), platforms (PaaS) or infrastructure as a service (IaaS). Next, the security considerations of infrastructures and cloud services are considered. [15].

3.1 Abuse and Misuse of Cloud Computing

This threat mainly affects the IaaS and PaaS service models and is related to a log of access to these infrastructures / platforms that is not very restrictive. For example, anyone with a valid credit card can access the service, with the consequent proliferation of spammers, creators of malicious code and other criminals who use the cloud as a center of operations; the IaaS have hosted Botnets that have hosted their control centers in cloud infrastructures or complete ranges of addresses of cloud infrastructures blocked by sending junk mail.

3.2 Interfaces and Unsafe API

In general, service providers (IaaS, SaaS, PaaS) in the cloud offer a series of interfaces and API (Application Programming Interface) to control and interact with resources. In this way, the entire organization, control, provision and monitoring of cloud services is done through these APIs or interfaces.

Given that everything (authentication, access, data encryption, etc.) is carried out through these tools, it is necessary that the interfaces are designed in a secure manner, thus avoiding security problems, both those that are intended and those that are they happen accidentally.

For example: allow anonymous access, reuse of tokens, authentication without encryption, limitations in managing records (activity logs) and monitoring.

3.3 Internal Threats

As in all information systems, the threat posed by the users themselves is one of the most important, given that they naturally have access to the company's data and applications. In a cloud environment this is not at all different since security incidents caused by disgruntled employees and accidents due to error or ignorance can also be triggered.

In addition, in many cases, it is the service provider itself (SaaS) that manages the registration and cancellation of users, resulting in security breaches when the consumer of the service does not inform the provider of the personnel losses in the company. As is logical, these incidents have an important impact on the image of the company and on the assets that are managed.

Service providers (SaaS) must provide consumers with the service of means and methods to control internal threats.

3.4 Problems Derived from Shared Technologies

This threat affects the IaaS models, since in an Infrastructure as a Service model the physical components (CPU, GPU, etc.) were not designed specifically for architecture of shared applications. There have been cases in which virtualization hypervisors could access the physical resources of the host, thus causing security incidents.

To avoid this type of incidents, it is recommended to implement an in-depth defense with special attention to computing, storage and network resources. In addition, a good security strategy must be generated that correctly manages the resources so that the activities of one user cannot interfere with those of the rest. For example, exploits or malware those get access to the resources of the virtualization host team.

3.5 Loss or Leakage of Information

There are many ways in which data can be compromised. For example, the deletion or modification of data without having a backup of the originals, supposes a loss of information, misuse of encryption and software keys or weak authentication, authorization and auditing.

In the cloud, the risk of data being compromised increases as the number of interactions between them is multiplied due to the architecture itself. This results in loss of company image, economic damages and if it is about leaks, legal problems, breaches of rules, etc.

3.6 Session or Service Hijacking

In a cloud environment (any of the SaaS, PaaS, IaaS models), if an attacker obtains credentials from a user's environment, they can access activities and transactions, manipulate data, return falsified information, or redirect clients to malicious sites.

3.7 Risks due to Lack of Knowledge

One of the pillars of cloud infrastructures is to reduce the amount of software and hardware (IaaS, SaaS, PaaS) that companies must acquire and maintain, in order to focus on the business. This, although it has an impact on both economic and operational cost savings, cannot be a reason for the deterioration of security due to lack of knowledge of this infrastructure (IaaS).

To assist in making decisions about the security measures that have to be implemented in a cloud environment, it is convenient to know, at least in part, the technical information of the platform (PaaS). Data such as who shares the infrastructure (IaaS) or unauthorized access attempts (SaaS) can be very important when deciding the security strategy. The lack of information of this type can lead to security breaches unknown by the affected.

3.8 Risks Detected by Gartner

Gartner, Inc. is a US company that conducts research and analysis of Information Technology (IT), based in Stamford, Connecticut, United States. It is known as Gartner Group until 2001. Its products and services include a complete collection of research and advice for technology users and vendors, contract review, detailed analysis of IT performance and peer networking opportunities through a forum Online, executive coaching and events. Gartner has an extensive database of market information and performs benchmarking analysis on IT, finance, sales, marketing and operations.

Since 1993, the company has made more than 32 acquisitions and investments. In 2017, the company acquired L2, a company that compares the digital performance of brands and CEB, a talent management and best practices company. The company serves 300,000 professionals in all business functions in more than 11,000 organizations and more than 110 countries. Gartner's Latin American operations are headquartered in Brazil. In other parts of the region it has offices in Mexico, Uruguay and Chile - covering Peru - and Venezuela, which covers activities in Colombia, Panama, the Dominican Republic and Puerto Rico. Based in Stamford, Gartner was founded in 1979.

From his position as IT analyst, he has also recently produced the report "Top Security and Risk Management Trends" [2] on the main risks in cloud computing. The following is an extract of the

key recommendations and trends of the aforementioned report.

3.8.1 Recommendations

Security and risk management leaders seeking to capitalize on these trends should:

– Build a security program that can link security strategy with business initiatives.
– Ensure that digital business plans include both data protection costs and data liability considerations.
– Exploit the emerging cloud security services providers to improve security and reduce administration overhead.
– Seek solutions that leverage machine learning but look for proof of value over other approaches.
– Ensure that digital business projects take into account the emerging geopolitical cyber landscape and concentration of digital resources.

3.8.2 Key Trends

– Senior business executives are finally aware that cybersecurity has a significant impact on the ability to achieve business goals and protect corporate reputation.
– Legal and regulatory mandates on data protection practices are impacting digital business plans and demanding increased emphasis on data liabilities.
– Security products are rapidly exploiting cloud delivery to provide more agile solutions.
– Machine learning is providing value in simple security tasks and elevating suspicious events for human analysis.
– Security-buying decisions are increasingly based on geopolitical factors, along with traditional buying considerations.
– Dangerous concentrations of digital power are driving decentralization efforts at several levels in the ecosystem.

3.9 Key Aspects of Security in the Cloud According to NIST

The National Institute of Standards and Technology (NIST) is an agency of the Technology Administration of the United States Department of Commerce. The mission of this institute is to promote innovation and industrial competition in the United States through advances in metrology, standards and technology in ways that improve economic stability and quality of life.

In this sense, it created on December 9, 2011 and updated on November 10, 2018 one of its guides "Guidelines on Security and Privacy in Public Cloud Computing" [5] in which it proposes security reinforcements focusing on a classification particular. It is summarized below.

3.9.1 Governance

Governance involves the control and supervision of policies, procedures and standards for the development of applications, as well as the design, implementation, testing and monitoring of distributed services. The cloud, due to its diversity of services and its wide availability, amplifies the need for good governance.

Ensuring that systems are secure and that risks are managed is a challenge in any environment in the cloud. It is a security requirement to properly install audit mechanisms and tools to determine how data is stored, how it is protected and how it is used both to validate services and to verify compliance with policies.

On the other hand, special attention must be paid to the roles and responsibilities involved in risk management. It is highly recommended to start a risk management program that is sufficiently flexible to deal with an environment of variable risks and in continuous evolution.

3.9.2 Fulfillment

Compliance requires compliance with standard specifications, standards or established laws. Legislation and regulations regarding privacy and security vary greatly depending on the countries with differences, sometimes at the national, regional or local level, making complicity in the cloud very complicated.

Location of the Data

One of the main problems of services in cloud computing is the lack of information about how the infrastructure has been implemented, so the subscriber has virtually no information on how and

where the data is stored or how it is protected. The possession of security certifications or the performance of external audits by the supplier mitigates, in part, the problem, although it is not a solution either.

When information moves through different countries, its legal and regulatory frameworks change and this obviously affects the way data are treated. For example, data protection laws impose additional obligations to the data handling and processing procedures that are transferred to the US.

The main concern of compliance lies in knowing the limits in which it stops applying the legislation of the country that collects the data and begins to apply the legislation of the destination country of the same as if the legislation in the destination involves some risk or additional benefit. In general, technical, physical and administrative safeguards, such as access controls, apply.

Electronic Research

Electronic research deals with the identification, collection, processing, analysis and production of documents in the discovery phase of a judicial proceeding. Organizations also have obligations for the preservation and generation of documents, such as complying with audits and requests for information. These documents not only include emails, email attachments and other data stored in the systems, but also metadata such as creation and modification dates.

The capabilities of a cloud provider and the available research tools can make it difficult to fulfill the obligations of the organization. For example, if the storage elements of a provider do not store the original metadata and intentional damage occurs (deletion of data, loss, material alteration or blocking of evidence that is basic to the investigation) the lack of this metadata has a negative impact on the investigation.

Confidence

In cloud computing, the organization gives direct control over many aspects of security, conferring an unprecedented level of trust on the cloud provider. Trust like the following:

– Access from within. Data stored outside the boundaries of an organization is protected by firewalls and other security controls that carry in themselves an inherent level of risk. Internal threats are a problem known to most organizations and although their name does not reflect it, it also applies to cloud services. These threats can be caused by the old ones and by the current employees as well as by the associated companies, the technical assistance and other actors that receive access to the corporate networks and data to facilitate the operations. The incidents can be both intentional and unintentional and of very different types including fraud, sabotage of information resources, theft of confidential information. When the data is transferred to a cloud environment operated by a provider, internal threats extend not only to the provider's staff but also to service consumers. An example of this was demonstrated by a denial of service performed by an internal attacker. The attack was made by a subscriber who created 20 accounts and launched a virtual machine instance for each of them, in turn each of these accounts continued to create 20 accounts each, causing exponential growth to take resources to the limits of failure.
– Data ownership. When a contract with a supplier is established, the rights to the data must be clearly defined and thus create a first trust framework. There is an important controversy surrounding the ambiguous terms that social networks use in their privacy and data ownership policies. The contract must state clearly that the organization retains ownership of all its data, but must also ensure that the provider does not acquire rights or licenses through the agreements to use the data to its own benefit.
– Complex services. The services of the cloud in themselves are usually formed by the collaboration and union of other services. The level of availability of a cloud service depends on the availability of the services that comprise it. Those services that depend on third parties for their operation must consider the establishment of a control framework with said third parties to define the responsibilities and obligations, as well as the remedies for possible failures. Responsibility and performance guarantees can become a

serious problem in complex services. An example of this is a social network that subscribed cloud storage services and closed by losing access to a large amount of data from 20,000 subscribers. The problem was that old data, new applications and databases were in different providers of cloud services.

– Visibility. Migration to public cloud services transfers control of security systems to the providers that operate the organization's data. The administration, procedures and controls used in the cloud must keep a certain analogy with those implemented in the internal organization itself to avoid potential security holes. Cloud providers are usually quite jealous to give details of their security and privacy policies, since such information could be used to carry out an attack. In general, the details of the network and the monitoring levels of the systems are not part of the service agreements. The transparency of the way in which providers operate is a vital issue for the supervision of an organization's security and privacy systems. To ensure that policies are met during the life cycle of the systems, service agreements must include some clauses to obtain visibility of the security controls and the processes used by the providers. Ideally, the organization should have control over certain aspects such as the definition of the limits that trigger alerts, the detail levels of the reports, etc. so that they adapt to the needs of the company.
– Risk management. With cloud-based services, many components of information systems are beyond the direct control of the subscribing organization. Many people feel better with a risk always.

3.9.3 Architecture

The architecture of a cloud infrastructure includes both hardware and software. Virtual machines are used as software distribution units associated with storage devices. The applications are created through the programming interfaces. They usually encompass multiple components of the infrastructure that communicate with each other through these interfaces. This global communication of the infrastructure can lead to security failures:

– Surface of the attack. The hypervisor of the virtual machines is superimposed as an extra layer of software between the operating system and the hardware resources and is used to run the multi-user virtual machines. The inclusion of these hypervisors means adding an extra point of attack with respect to traditional architectures. Examples of possible incidents are the disclosure of sensitive data when migrating virtual machines or executing arbitrary code on the host computer when exploiting vulnerabilities in virtualization products.
– Protection of the virtual network. Most virtualization products support the creation of virtual network switches and network configurations as part of the environment. Likewise, they support the creation of private subnets for communication between virtual machines hosted on the same server. This traffic cannot be monitored by the typical physical elements of the network (firewalls, intrusion prevention systems, etc.). Therefore, the security precautions of the network in these internal connections must be taken to avoid attacks from within these virtual networks.
– Auxiliary data. The most normal thing is that the security in these environments is focused on the data managed by the application, but there are other data not considered so critical whose alteration, theft or disclosure can produce serious security incidents (for example: customer databases). , payment files, user information, etc.). Another problem may occur when not protecting access to the repositories of the templates of virtual machines that contain the default settings of the same. Sharing this type of data is a fairly common practice in cloud environments. Sharing this type of data can offer the attacker platforms to prepare his attack by checking the vulnerabilities that may be inherent to them, although there may also be the opposite process: attackers who try to replace an image with another with malicious content installed.
– Client protection. Browsers, as a fundamental part of cloud environments (since accesses

are typically made via the web), can carry extensions or plugins with important security breaches. Maintaining the logical and physical security of the client's part can be complicated especially for mobile environments since, due to their size and portability; they can be lost or stolen. Likewise, desktop systems are not updated systematically, causing vulnerabilities to produce important security breaches. Another aspect to consider is the possible presence of Trojans, back doors or viruses that can obtain confidential information or monitor the victim. The solution could start by reinforcing the security check of the clients. For example, banks begin to develop measures for their clients' browsers to secure data by encrypting communications and applying mechanisms to avoid the interception of keystrokes.

– Server protection. The servers in cloud environments must be protected both physically and logically, so that they are segmented and separated so that access cannot occur from unauthorized areas. In the same way, in the part of the client it is necessary to assure the patching of the servants so that they do not have vulnerabilities that compromise the security of the surroundings. Another important aspect in the protection of the servant part is the availability, reason why it will be advisable to have systems that enjoy this, mainly in those servers that lodge critical parts of the infrastructure.

3.9.4 Identity and Access Control

Sensitive data and privacy have become the main concern regarding the protection of organizations and unauthorized access to information resources. The alternative of using two authentication methods, one for internal organization and another for clients, can be very complicated. The federated identity, which has become popular with the growth of service-oriented architectures, can be one of the solutions, being able to be implemented in several ways following the SAML (Security Assertion Markup Language) standard or the OpenID standard.

Authentication

A growing number of cloud service providers support the SAML standard, using it to manage users and authenticate them before granting access to applications and data. This standard provides an environment for the exchange of information for authentication between cooperating domains. The requests and responses in this standard are mapped using the SOAP protocol (Simple Object Access Protocol). SOAP messages are digitally signed. In this way, when a user has a public key certificate for a cloud environment, their private key can be used to sign the SOAP requests.

Validation of an authentication via SOAP messages is complicated and must be treated with care to avoid possible attacks. For example, the success of XML Wrapping attacks against cloud infrastructures has already been demonstrated. This type of attacks manipulate the SOAP requests by introducing wrappers in the SOAP security header, moving the body to said envelope and replacing the body with one executing an action defined by the attacker. As the original message is not deleted, the signature is verified but the execution of the malicious action is performed.

Access control

The SAML standard alone is not enough to provide both authentication and access controls in cloud services. It is necessary to implement also a control of access to resources. For this, the XACML (eXtensible Access Control Markup Language) standard can be used to control access to resources. This standard complements SAML for cooperative domain authentication and authorization exchange environments. However, messages between XACML entities may be susceptible to attacks by third parties, so they must be protected against attacks of information exposure, repetition, deletion and modification.

Software isolation

To achieve high efficiency rates, providers must ensure both a dynamic provision of service and the isolation of service subscribers. The concurrence of users is performed in cloud environments by multiplexing the execution of virtual machines for

different users on the same physical server. Even so, the applications that run in these environments allow being pockets of attack. The most outstanding are:

– Hypervisor complexity. The security of a computer system depends on the quality of the software running in its core, which controls the location and execution of processes. A virtual machine monitor is designed to run multiple machines concurrently in the same physical host providing isolation between the different virtual machines. Normally these hypervisors are less complex than an operating system, so the analysis and improvement of security is, in theory, simpler. The reality is that the evolution of these hypervisors has turned them into much more complex and extensive elements, similar to operating systems. A clear example of this is Xen which contains a modified Linux kernel to isolate the input / output operations and includes KVM (Kernel-based Virtual Machine), so that it transforms the Linux kernel into a hypervisor. It is important for the provider to understand the use of virtualization in order to understand the associated risks that may occur.
– Attack vectors. The concurrence of multiple users sharing physical resources through different virtual machines can produce new sources of threats. One of the main threats is the possibility that malicious codes can leave the virtual machines and interfere with the hypervisor or with other virtual machines. The possibility of hot migration of virtual machines from one physical server to another, and other features provided by hypervisors to facilitate management, increases the size and complexity of the software and therefore adds new areas to carry out attacks. The programming interface of cloud services is usually a common goal to discover vulnerabilities that can be exploited. These types of vulnerabilities are usually buffer overflows, which allow the attacker to execute arbitrary code or failures that allow denial of service that may affect the virtual machine or the host server itself. Indirect attacks, such as those demonstrated by some developers, may also be caused by a weakness in the process of migration of virtual machines that allowed an attacker to obtain administrative control of the machine through a man-in-the-middle attack to modify the code of authentication. On the other hand, memory modifications during migration that allows the installation of specific rootkits for virtual machines can suppose another vector of attack. Another option can be the monitoring of the use of resources in a shared server to collect information and in that way, to obtain information on when is the best time to carry out an attack.

3.9.5 Data Protection

The data that is stored in cloud environments usually reside in equipment shared by multiple clients. Therefore, organizations that manage sensitive data in the cloud must be concerned about the way, in which these data are accessed and ensure that they are stored securely.

Data isolation

Data in cloud environments can take many forms depending on the activity to which they are dedicated. If, for example, the activity is the development of applications, the data will be found in the form of programs, scripts and configuration data. On the other hand, if an application that has already been developed resides in the platform, the data will be of the registers type, contents created and used by the application or user information, etc.

One of the main problems of the cloud environments is the authentication of the identity of the users. Access controls are usually based on identity verification.

A particular case is the typical database environments of the cloud environments that are composed of a single database manager system with one instance per virtual machine. The security and configuration of these instances falls on the subscribing part of the service. The segregation of the data is usually done through labels for the data, which provides a false appearance of exclusive use of the instances.

On the other hand, the data must be protected when they are at rest, in transit. Likewise, access to the data must be controlled. Communications standards and public-key certificates allow data

transfers to be protected using cryptography. Rest security is not so simple because the procedures are not clear because most systems are proprietary, which, in turn, hinders interoperability between different providers with different systems.

The management of cryptographic keys rests mainly with the subscriber of the service. The generation of these keys is usually done using hardware security modules or HSM (Hardware Security Module). Data protection in use is an emerging area of cryptography with little practical material to offer, letting the trust mechanisms be the greatest safeguard.

Data sanitation

Sanitation is the elimination of sensitive data from a storage medium when it is no longer used in the environment or is to be reused in another environment or situation. Sanitation also applies to backup data and residual data that remain when the service ends.

In cloud environments, this task can be very complicated since the data of several clients share storage, so sanitation must be carried out with great caution. In addition, through specific techniques and equipment, data can be recovered from storage media previously erased, which makes this sanitation a critical task.

3.9.6 Availability

Availability can be interrupted temporarily or permanently. Denial of service attacks, equipment failures and natural disasters are all threats to availability.

3.9.7 Response to Incidents

The work of the provider is basic in the response activities in the event of a security incident. This includes verification, analysis of the attack, containment, collection of evidence, application of remedies and restoration of the service.

The collaboration between providers and subscribers for the detection and recognition of incidents is essential for security and privacy in the cloud computing, since the complexity of the services can make the task of detection more difficult. It is necessary to understand and negotiate the incident response procedures before signing a service contract. The location of the data is another aspect that can prevent an investigation, so it is another of the points that must be negotiated in contracts.

The solution negotiated must be aimed at mitigating the incident in a time that limits the damage and improves recovery times. The teams for the resolution should be mixed (provider and subscriber) since the solution can involve one of the parties individually or both together and the incident can even affect other subscribers who share the infrastructure.

4 Current Solutions that Address in Part the Security Challenges Proposed by the Cloud

There are several research works in the area of security in the cloud. Several groups and organizations are interested in developing security solutions and standards for the cloud. Cloud Security Alliance (CSA) is bringing together solution providers, non-profit organizations and individuals to discuss current and future best practices for securing information in the cloud [1]. The cloud standards website is collecting and coordinating information on the cloud-related standards that are being developed the most. The Open Web Application Security Project (OWASP) maintains a list of the main vulnerabilities of the cloud-based or SaaS models that is updated as the threat landscape changes [9]. On the other hand, Open Grid Forum (OGF) publishes documents that contain security and infrastructure and information specifications for Grid computing developers and researchers [8].

4.1 Data Classification and Data Encryption Techniques Used in Cloud Computing

On the other hand, there are algorithms that have been used to try to improve security problems in the cloud, among which are those described in sections 4.1.1, 4.1.2, 4.1.3, 4.1.4 and 4.1.5.

4.1.1 Vigenère Encryption

It is a typical example of a polyalphabetic cipher whose invention was wrongly attributed to Blaise de Vigènere and dating from the sixteenth century. The key is constituted by a sequence of symbols of the alphabet K = {k0, k1, ..., kd-1}, of length d

and that it uses, the following congruent linear transformation of cipher: Ek (mi) = mi + k (i mod d) (mod n) being my the i-th symbol of the clear text and n the cardinal (length) of the input alphabet. As a key, you can use any word of a length for example between 6 and 8 characters that does not have repeated letters.

To see this better, suppose that with the Spanish alphabet of 27 symbols, you want to encrypt the plain text "PLAN" and that for encryption, the word "SOL" is used as the key.

The first letter of the message, the P will be encrypted with the first letter of the key, S, which indicates that the monoalphabetic substitution E ("P") = E (16) = (16 + 19) mod 27 = 8 = "I", since if A occupies position 0, S occupies position 19 of the alphabet. The letter L of the message will be encrypted using the letter O of the key and the letter A of the message will be encrypted using the letter L of the key. For the last letter of the message (N), the first letter of the key (S) is used again.
Therefore, the result is:

- Message P L A N
- Key S O L S
- Encryption I Z L F

To facilitate operations with this cryptosystem, the so-called Vigenère table is available (see Figure 2), which is formed by a square matrix of 27x27 in the case of an alphabet of 27 letters such as Spanish. The first row of the matrix is formed by the alphabet starting with the letter A and ending with the letter Z, the second by the alphabet that begins with B and ends in A and thus, until the last row, the 27th, which begins by the letters ZAB ... and ends with the letter Y [7].

Cryptanalysis of the Vigenère encryption

To cryptanalyze this type of encryption, it is necessary to perform independent statistical analyzes by grouping the cryptogram symbols in different groups according to the k_i used to code them; each group will be encoded with the same encryption alphabet. To estimate the length d of the key, (that is, the number of different alphabets used in the encryption) the periodicity of the common patterns that may appear in the encrypted text is sought. Obviously, for cryptanalysis, it is necessary at least d times more amount of

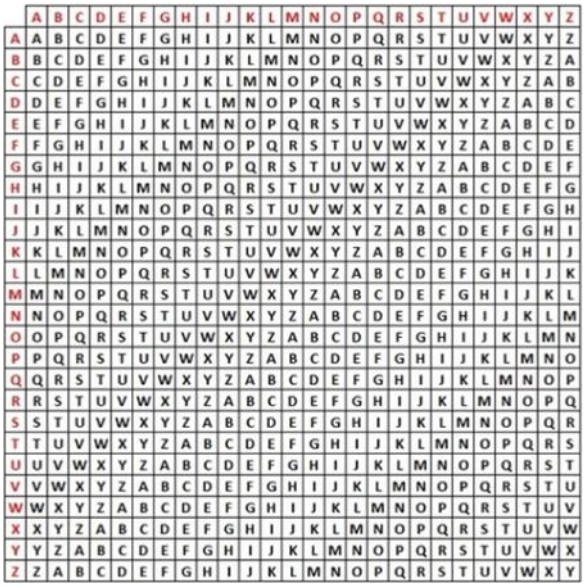

Fig. 2. Vigenère encryption in the case of a 27-letter alphabet [7]

```
N O R I A
B C D E F
G H K L M
P Q S T U
V W X Y Z
```

Fig. 3. Keyword encryption matrix NORIA

encrypted text than with monoalphabetic methods [7].

4.1.2 Playfair encryption

This algorithm arises around 1850 and is developed by Wheatstone who implements it on the disk developed by him and named Wheatstone Disc, however, the procedure used called Playfair in honor of his friend Lord Playfair [12].

It is an algorithm that seeks to increase the security of the encryption by avoiding a frequency analysis, for which it uses polygrams, that is, performs the process of encryption by blocks of characters, in this case diagrams, using it a matrix of 5 x 5 which contains the 26 letters of the English alphabet and starting the matrix with the sequence corresponding to the keyword [12], so that if the keyword is NORIA, the resulting cipher matrix is the shown in Figure 3.

A	B	C	D	E	F	G	H	I	J	K	L	M	N
0	1	2	3	4	5	6	7	8	9	10	11	12	13
Ñ	O	P	Q	R	S	T	U	V	W	X	Y	Z	
14	15	16	17	18	19	20	21	22	23	24	25	26	

Fig. 4. Keyword encryption matrix NORIA

Additionally, the following encryption rules taken from [12] should be considered.

Rules to encrypt 2 characters m1 m2:

– If m1 and m2 of the same row, choose c1 and c2 on your right (circularly and for each letter).
– If m1 and m2 of the same column, get c1 and c2 from below (circularly and for each letter).
– If m1 and m2 of different rows and columns, form the square and obtain c1 and c2 of opposite diagonal, always from right to left.
– If m1 = m2, insert character without meaning between m1 and m2 to avoid its repetition, then apply rules 1-3.
– If the number of letters is odd, add one with no meaning at the end of the text.

Therefore, if we have the clear text for example: AT AQ UE CE RO HO RA SX (the X is put at the end because the text is odd number of letters, you must put a letter without meaning to fill and return to parity). You could also put letters without meaning at the end of each word to avoid confusion or make the resulting text clearer.

Thus, the resulting Cryptogram is: IU OU TF DF IR QC IN XR.

4.1.3 Hill Encryption

Hill's cipher was invented, based on linear algebra, by the American mathematician Lester S. Hill in 1929 and is explained in his article Cryptography in an Algebraic Alphabet, published in The American Mathematical Monthly.

It is a cryptographic system of polyalphabetic substitution, that is, the same sign, in this case the same letter, can be represented in the same message with more than one character. Thus, in the example we are going to analyze below, the letter A of the original message is represented in the encoded message in three different ways, such as C, K and I [6, 12].

Hill's encryption consists, first, of the association of each letter of the alphabet with a number. The simplest way to do it is with the ordered natural association, although other different associations could be made. In the following example only the 27 letters of the alphabet are used (see Figure 4), but other common symbols could also be added, such as the blank "_", the dot "." Or the comma ",", the interrogation " ? ", 10 basic figures, etc. As in the previous correspondence, between letters / signs and numbers, only 27 numbers appear, you must work with the whole numbers "module 27".

That is, the integers are considered 0, 1, 2, ..., 26 and the rest is identified with them in a cyclical way. Thus, 27 is equal to 0, 28 to 1, 29 to 2, etc. and the same with negative numbers, so that - 1 equals 26, - 2 equals 25, etc.

In addition, the arithmetic operations (addition, subtraction, multiplication and division) are reduced to the set of whole numbers modulo 27 naturally, that is, when operating two integers (module 27) the result is also considered module 27. For example, if the multiplication of the numbers 6 and 13, module 27, takes place, the result will give 24 (module 27), since 6 mod 13 = 78 and 78 = 2 mod 27 + 24. Or the inverse of 2, that is, the number a such that 2 mod a is equal to 1 (module 27), is 14, since 2 mod 14 = 28, which is equal to 1, module 27 [6, 12].

In Hill's cipher, a square matrix of numbers A is used as a key, which determines the linear transformation Y = A · X, where Y, X are column vectors and A and X are multiplied with matrix multiplication (see the following image). Let's see an example. Consider the 3 x 3 square matrix (although in general we can consider square matrices of any size) and the corresponding linear transformation Y = A · X [6, 12]:

$$\begin{pmatrix}1 & 2 & 3\\ 0 & 4 & 5\\ 1 & 0 & 6\end{pmatrix} \quad \begin{pmatrix}y_1\\ y_2\\ y_3\end{pmatrix}=\begin{pmatrix}1 & 2 & 3\\ 0 & 4 & 5\\ 1 & 0 & 6\end{pmatrix}\cdot\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}$$

$$\left.\begin{aligned} y_1 &= 1\cdot x_1+2\cdot x_2+3\cdot x_3\\ y_2 &= 0\cdot x_1+4\cdot x_2+5\cdot x_3\\ y_3 &= 1\cdot x_1+0\cdot x_2+6\cdot x_3\end{aligned}\right\}$$

4.1.4 Vernam Encryption

Developed and published in 1920 by G.S Vernam of the laboratories Bell and AT & T and specified in the patent 1310719 of the E.U.A, Vernam

represents the limit case of Vigénere's encryption [12].

Vernam [12] proposes to convert the plain text into a string of bits (which can be represented in ASCII code) so that this gives a greater robustness to the encryption since until now the algorithms studied saw the character as the unit smaller and can be represented in ASCII code) so that this gives greater strength to the encryption, since until now the algorithms studied saw the character as a smaller and invisible unit during the process of encryption-decryption and Vernam comes to show , that each character can be represented and used during the process in smaller fractions, also suggests using as key a binary random or pseudo-random sequence of the same size as the clear message and in a relevant way that the key was of a single use (encrypt to broadcast). Then apply an OR operation Exclusive, bit by bit with these two strings to obtain the cryptogram, so the encrypted text cannot provide any information to the cryptanalyst.

Encryption Process

Take as an example the encryption of the surname of the creator of the algorithm: VERNAM and as the first step, we obtain for each character its numerical equivalent in ASCII and the corresponding coding:

V	E	R	N	A	M
86	69	82	78	65	77
01010110	01000101	01010010	01001110	01000001	01001101

Now consider the random sequence 00110101 00001011 11010101 01111111 11001010 01101001 that could have been obtained from various sources, such as sequence generator and proceeds, to perform the Exclusive OR.

Mcla	01010110	01000101	01010010	01001110	01000001	01001101
K	00110101	00001011	11010101	01111111	11001010	01101001
Crypto	01100011	01001110	10000111	00110001	10001011	00100100

As it is observed, the use of the encoded text to obtain the cryptogram makes it more robust and more difficult to decrypt for the cryptanalyst since even using the same carnation sequence, the encryption will be different if instead of uppercase letters in the message in of course, this would have been written only in lowercase or interspersed in the message in clear, this would have been written only in lowercase or interspersed uppercase and lowercase.

Decryption Process

To perform the decryption process only requires knowing the sequence used as the key and the corresponding cryptogram, the operation that reverts the encryption is the same as that used to encrypt, this is an Exclusive OR and determine which ASCII characters correspond to the sequence.

Crypto	00100111	00001111	11011000	00000111	01011000	01100001
Mcla	01110111	01101010	10101010	01100011	11111010	00001111
K	01010000	01100101	01110010	01100100	10100010	01101110
	80	101	114	100	162	110
	P	e	r	d	ó	n

4.1.5 Proxy re-encryption schemes

During a proxy re-encryption scheme, a proxy server will convert simple text to encrypted text under a PKA public key [10] and then this encrypted text will be encrypted again under another PKB public key using the RKA encryption key -> B and in this way, the plain text is transformed into an encrypted text. In this scheme, messages are encrypted before they are stored on the storage server.

If a user wants to share their messages, they send a re-encryption key to the storage server. The storage server has the encrypted messages and then re-encrypted with the re-encryption key for the demanding user. Therefore, your system has information confidentiality and a secure way of communication. Proxy encryption is based on the concept of a partial trust proxy that uses an encryption key to translate an encrypted text under the public key of the owner of the data into other encrypted text that can be decrypted by another user's private key.

The data is never decrypted before being re-encrypted, so the proxy can never reveal the plain text at any time. However, the problem with this technique is that it does not handle the case in

which a revoked user and the proxy conspire, which can reveal all the private keys of the rest of the user in the group. Also, another important problem with this is that, since it uses public key cryptography of Elgamal, it does not allow the encryption or decryption of very large data [10].

4.1.6 RailFence Cipher

A RailFence cipher is a type of written or encrypted code that allows users to transform text for encoding purposes, using only a pen and paper.

In the RailFence encryption, the letters are not changed, but only changed according to their position in the message.

This type of encryption is often called transposition encryption, because the letters are simply transposed in terms of their location. Transposition ciphers, such as RailFence encryption, are relatively weak coding forms and can be easily broken, especially with current technology. These types of ciphers date back to the American Civil War, where soldiers would use the code to send encrypted messages. In a RailFence cipher, the writer takes a message and writes it in descending lines or "lanes". The encryption of the railway guide is sometimes called zigzag encryption if the writer uses a zigzag or W pattern to represent the text. To encode the text, the user takes the letters of the top line or rail and puts them together. Then the second line and the third line are written. The result is a coded text line. For example, when using the phrase "hello world" and a series of three lanes, the result (for a linear descent) would be HLODEORLWL [3].

5 Conclusions and Future Works

The demand of potential consumers through the Internet presents a challenge for developers and project teams. It is good to have this alternative in mind to face possible problems, which could easily be solved with techniques such as Cloud Computing. The fact that cloud environments proliferate exponentially forces potential users better understanding of these environments and their main problems. The term cloud computing is broad and its definition is not precise.

Therefore, when it comes to choosing services in the cloud, it is necessary to be clear about the type of infrastructure that supports it and the type of service offered. After the analysis carried out in this study, a global view of this problem is obtained and common conclusions are drawn from all points of view.The security and ownership of the data is one of the key aspects. The reports show a great concern for the ownership and treatment of the data given that these infrastructures can manage the data in multiple countries, which can generate conflicts regarding the legal framework in which they are treated.

It is also suggested that these environments, when handling a large amount of data, can be the object of information leaks, whether intentional or fortuitous. Regulatory compliance is also one of the pillars of security in cloud environments. In this case, the problem arises due to the lack of transparency of these infrastructures, so it is highly recommended that the subscriber of the service be clearly informed of how the environment is managed.

Quantity of software from different providers is involved in the creation of a cloud service. That is, they are complex environments, so special attention must be paid to possible vulnerabilities and the implementation of retouching procedures. Identity and access control are another important aspect. In general, most infrastructures are shared by multiple companies or users and the poor definition of access controls can lead to unauthorized access to confidential data. The definition of a good identity policy and access control based on minimum privilege policies is essential in cloud environments.

Finally, there is a common denominator to all these mentioned aspects. It is about service agreement agreements. All the recommendations regarding this matter indicate that they should be reviewed and created specifically, detailing the controls, the regulations, the protection measures, the recovery periods of the service, etc.

Cloud computing is one of the emerging paradigms and the security of data in the cloud is the most important problem that acts as a barrier in the implementation of cloud computing. This survey on security and the possible techniques of classification of existing data used in cloud computing allows establishing that the objective of data classification is to institute the level of security required for the information and protect the data by

providing a level of security enough according to the risk levels. In this document, a survey is made on the existing encryption techniques, which protect the data throughout the life cycle from the beginning to the end in cloud computing.

This brief study can also allow comparisons of existing encryption techniques used in cloud computing, looking for and analyzing their advantages and limitations. With all the work reviews, we can conclude that all the techniques, among which are those shown in this article, available in the current market, are useful for encrypting data in real time. Each technique is unique and useful in its own way, which may be suitable for different applications based on the specific requirements of that application. In addition, each technique has its own advantages and disadvantages.

Perhaps in a future study, we can add other techniques and establish what are the best security techniques for data in the environment in the cloud say that address the maximum security and even, aim to propose a scheme that will contain the security features of these, while overcoming the disadvantages and the open problems in them.

Acknowledgements

We thank the support of Instituto Politecnico Nacional (IPN), SIP-IPN.

References

1. **Cloud Security Alliance (2015).** *Best Practices for Mitigating Risks in Virtualized Environments.* © 2015 Cloud Security Alliance – All Rights Reserved, This paper is based on TR 30: 2012, "Technical Reference for virtualisation security for servers", developed by the Information Technology Standards Committee.
2. **Firstbrook, P., Perkins, E., Wheatman, J., Mahdi, D., Zhang, J., & Olyaei, S (2018).** *Top Security and Risk Management Trends.* Gartner, Research, ID Number: G00337028.
3. **Ritwik, G., Binod, K.M., & Prashant, L. (2017).** *Enhancing the performance of Data Encryption Standard algorithm by using Rail Fencing.* Vol. 2, No. 3.
4. **Heiser, J. (2015).** *Roundup of Cloud Application Security and Governance Research.* Gartner, Research, ID Number: G00281221.
5. **Jansen, W. & Grance T. (2011).** Guidelines on Security and Privacy in Public Cloud Computing. *Special Publication (NIST SP), Created National Institute of Standards and Technology, (NIST).*
6. **Krishna, A.V.N. & Babu, A. (2007).** A modified hill cipher algorithm for encryption of data in data transmission. *Computer Sciences and Telecommunications,* Vol. 3: pp. 78–83.
7. **Al-Amin, M.A. & Abdulrahman, O. (2016).** Vigenere Cipher: Trends, Review and Possible Modifications. *International Journal of Computer Applications,* Vol. 135, No. 11. DOI: 10.5120/ijca 2016908549.
8. **Open Grid Forum (2017).** https://www.ogf.org/ogf/ doku.php/documents/documents.
9. **OWASP (2014).** https://www.owasp.org/ index.php/ Category:OWASP_Cloud_%E2%80%90_10_.
10. **Qin, Z., Xiong, H., Wu, S., & Batamuliza, J. (2016).** A Survey of Proxy Re-Encryption for Secure Data Sharing in Cloud Computing. *IEEE Transactions on Services Computing.* DOI: 10.1109/TSC.2016. 2551238.
11. **Ren, K., Wang, C., & Wang, Q. (2015).** Security Challenges for the Public Cloud. *IEEE Published by the IEEE Computer Society.*
12. **Stallings, W. (2005).** *Cryptography and Network Security: Principles and Practices.* Prentice Hall, pp. 35.
13. **Stanoevska-Slabeva, K. & Wozniak, T. (2010).** *Grid and Cloud Computing-A Business Perspective on Technology and Applications.* Springer-Verlag, Berlin, Heidelberg.
14. **Viega, J. (2009).** Cloud computing and the common man. *IEEE Computer Society,* Vol. 42, No. 8, pp. 106–108. DOI: 10.1109/MC.2009.252.
15. **National Institute of Standards and Technology**. **(2011).** NIST Special Publication 500-291. *NIST Cloud Computing Standards. Natl. Inst. Stand. Technol. Spec. Publ.,* No. 108, pp. 500–291.

Article received on 22/01/2019; accepted 08/08/2019.
Corresponding author is Sandra Dinora Orantes Jiménez.