Customer Security Concerns in Cloud Computing

Shirlei A. de Chaves, Carlos B. Westphall, Carla M. Westphall, and Guilherme A. Gerônimo
Network and Management Laboratory
Federal University of Santa Catarina
Florianópolis – SC - Brazil
Emails: {shirlei, westphal, carla, arthur}@lrg.ufsc.br

Abstract—**There is no consensus about what exactly cloud computing is, but some characteristics are clearly repeated. It is a new distributed computing and business paradigm. It provides computing power, software and storage and even a distributed data center infrastructure on demand. In this paper, we investigated what are the main security concerns faced by the customers that are trying to better understand or profit from this new paradigm, especially considering a public cloud and we conclude that data confidentiality, integrity and availability are the biggest ones.**

Keywords - Cloud computing; security; distributed computing.

I. INTRODUCTION

Despite of the fact that industry big players like Google, Amazon, SalesForce, Microsoft and others have products and services under the umbrella of 'cloud computing', 'cloud ready' or other similar denomination, there is no consensus about what exactly cloud computing is. Below we list some definitions made by researchers:

"*Cloud computing is the next natural step in the evolution of on-demand information technology services and products. To a large extent cloud computing will be based on virtualized resources.(…) Cloud computing embraces cyber infrastructure and builds upon decades of research in virtualization, distributed computing, grid computing, utility computing, and more recently networking, web and software services* [1]."

"*A large-scale distributed computing paradigm that is driven by economies of scale, in which a pool of abstracted virtualized, dynamically-scalable, managed computing power, storage, platforms, and services are delivered on demand to external customers over the Internet* [2]."

"(…) *cloud computing is a nascent business and technology concept with different meanings for different people. For application and IT users, it's IT as a service (ITaaS) - that is, delivery of computing, storage, and applications over the Internet from centralized data centers. For Internet application developers, it's an Internet-scale software development platform and runtime environment. For infrastructure providers and administrators, it's the massive, distributed data center infrastructure connected by IP networks* [3]."

""*A Cloud is a type of parallel and distributed system consisting of a collection of inter-connected and virtualized computers that are dynamically provisioned and presented as one or more unified computing resources based on service-level agreements established through negotiation between the service provider and consumers* [4]."

"*a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction* [5]."

From these definitions, it is possible to notice that some characteristics are clearly repeated. It is a new paradigm, not just a distributed computing paradigm, but also a new business paradigm. It is intended to provide computing power, software and storage and even a distributed data center infrastructure on demand. In order to make these characteristics viable, cloud computing makes use of existing technologies, such as virtualization, distributed computing, grid computing, utility computing and Internet. However, even those industry big players have products and services available as also a definition of what are the basic cloud computing underlying technologies, a customer intending to better understand and profit from this new paradigm faces several concerns, especially the ones related to security.

Considering the customer point of view, we have made an extensive research to obtain what are the main security problems pointed in the available literature for cloud computing security, aiming to list and discuss the more recurrent ones. The results and the discussion are presented in Section 3. It is also worth to mention that being the cloud computing security subject under active research, many changes, new events or studies relating to it are coming out in a rapid pace, so this paper does not aim to exhaust it, but to contribute to the discussion.

.

II. CLOUD COMPUTING CATEGORIES

Attempts to cloud computing standardization are being done by some groups, including governments and industry. One effort that can help to avoid

ISBN:978-1-61208-113-7

misunderstandings, by putting everyone to talk the same language, is the definition of cloud computing and its categories. As of this writing, the US National Institute of Standards and Technology (NIST) is one of them, having defined the cloud as composed of four deployment models, three service models and five essential characteristics. The Cloud Security Alliance [6], which formal debut was made at RSA Conference 2009 releasing a white paper entitled "Security Guidance for Critical Areas of Focus in Cloud Computing", has taken these definitions to work through its guidance, explaining that the motivation is "to bring coherence and consensus around a common language so we can focus on use cases rather than semantic nuance [6]".

A. *Deployment models*

The definitions of the deployment models listed next are taken as it is from the NIST definition, although other researches mention this deployment models with similar definitions.

1) *Public Cloud*

The cloud infrastructure is made available to the general public or a large industry group and is owned by an organization selling cloud services.

2) *Private Cloud*

The cloud infrastructure is operated solely for an organization. It may be managed by the organization or a third party.

3) *Hybrid Cloud*

The cloud infrastructure is a composition of two or more clouds (private, community, or public) that remain unique entities but are bound together by standardized or proprietary technology that enables data and application portability (e.g., cloud bursting for load-balancing between clouds).

4) *Community Cloud*

The cloud infrastructure is shared by several organizations and supports a specific community that has shared concerns (e.g., mission, security requirements, policy, and compliance considerations). It may be managed by the organizations or a third party.

B. *Service Models*

The three service models listed below are not exclusively from NIST, being mentioned in several papers, including [2][5].

1) *Infrastructure as a Service (IaaS)*

"The capability provided to the consumer is to provision processing, storage, networks, and other fundamental computing resources where the consumer is able to deploy and run arbitrary software, which can include operating systems and applications [5]."

2) *Platform as a Service (PaaS)*

"The capability provided to the consumer is to deploy onto the cloud infrastructure consumer-created or acquired applications created using programming languages and tools supported by the provider [5]."

3) *Software as a Service (SaaS)*

In this case, is provided "a complete, turnkey application—including complex programs such as those for CRM or enterprise-resource management via the Internet [5]." Or, in the words of NIST, "the capability provided to the consumer is to use the provider's applications running on a cloud infrastructure [5]."

In each of these service models, what can be controlled by the customer varies, but in general, he does not have control over the underlying cloud infrastructure. This is especially true when is the case of a public cloud, the focus of the present paper. In a private cloud, for example, security responsibilities can be taken on by the customer, if he is managing the cloud, but in the case of a public cloud, such responsibilities are more on the cloud provider and the customer can just try to assess if the cloud provider is able to provide security.

The five essential characteristics are related to characteristics already mentioned in the introduction of this paper: on-demand self-service, broad network access, resource pooling, rapid elasticity and measured services.

III. CLOUD COMPUTING SECURITY OVERVIEW

Many cloud computing security problems are still unclear. Being cloud computing such a recent computing paradigm, it is natural that many aspects remain uncovered whereas the paradigm itself is being more developed and understood.

According to [5], there are three main customers' concerns:

Vulnerability to attack: critical business information and IT resources are outside the customers firewall.

Standard security practices: customers want to be confident that such practices are being followed. Most of those practices require disclosure and inspection, which leads to another concern as a customer: will my data be in the same virtual hardware and network resources with other customers, being susceptible to disclosure in someone else's inspection?

Being subject to state or national data-storage laws related to privacy or record keeping: European Union (EU), for example, has privacy regulations that do not permit some personal data to be transmitted outside the EU. In the cloud, data can be stored anywhere in the world; it is important to attend such regulations.

In June 2008, the Gartner Group released a report entitled "Assessing the Security Risks of Cloud Computing" [7]. According to this report, widely commented and cited on the Internet, before jumping into the cloud, the customer should know its unique security risks, considering specially seven security conditions during the process of choosing a cloud provider. These unique security risks are:

Privileged user access: outsourcing means allowing outsourced services to bypass internal controls, including personnel controls. With this in mind, the customer has to obtain as more information as possible about how the possible future provider hires people and what kind of controls their accesses have.

Regulatory Compliance: if the cloud computing provider is not subject of external audits and security certifications, the customer probably should not use its services for non trivial tasks. Customers have to always remember that, unless stated or agreed otherwise, they are responsible for their own data.

Data location: when using the cloud, the customer probably will not know where their data will be stored. Thus, it is recommended checking if the provider will commit to store and process data in specific jurisdictions and if a contractual commitment on behalf of the customer will be made by the provider.

Data segregation: customers should check what is done to separate different customers' provider data, due to the fact that, in a cloud, the environment is shared. Using cryptography, for example, is effective, but do not solve all the problems. It must be checked also if the cryptographic schemes are designed and tested by specialists, because cryptographic accidents are able to make data unusable.

Recovery: the provider capacity of restoring the entire system and how long it would take should be checked by the customer. Any provider that does not replicate its data or infrastructure is prone to total failures.

Investigative support: In order to have confidence that inappropriate or illegal activities will be possible to be investigated, the customer needs a formal commitment from the provider. This commitment should state which kind of investigation will be possible and also gives evidence that similar support was already done by the provider. Otherwise, the customer almost can be sure that such investigations will be impossible.

Long-term viability: if happens that the cloud computing provider be acquired or goes broke, the customer needs to know if the data will still be available and in a format that will allow being imported to a substitute application.

Summarizing the Gartner's report, customers should demand transparency and avoid providers that do not offer clear information about security programs.

In a Tech News from forbes.com, published online in February 02, 2008 [8], by Andy Greenberg is cited that when customers store their data in someone else's software and hardware, "they lose a degree of control over their often-sensitive information". In that article, Greenberg gives the example of an employee of an investment bank that uses Google Spreadsheets to organize a list of bank employees and their social security number. In this example, the responsibility of protecting such information from hackers and internal data breaching is not from the bank, but Google's. Another situation in this same case is that, if government investigators subpoena Google to supply that list, sometimes even with no customer knowledge, Google may attend. Google's privacy politics says that it will share data with the government if it has a "good faith belief" that this is necessary [9]. Greenberg also points that other problems to cloud computing is the cyber crimes. He gives some examples occurred in 2007, like:

- "retailer TJX lost 45 million credit card numbers to hackers";
- "the British government misplaced 25 million taxpayer records";
- "software company Salesforce.com sent a letter to a million subscribers describing how some customers' e-mail addresses and phone numbers had been snagged by cybercriminals - and warning how another wave of phishers were attempting to send malware more broadly to Salesforce.com customers".

Analyzing the articles cited before, it is possible to visualize that the main concerns, if not all, are related to the business level, i.e., customers are worried about how their business process will be affected. This situation is expected because cloud computing can also be seen as a new business model, with many aspects to be fully understood before being adopted with no restrictions or, better, with fewer restrictions. As happen in any new business or technological area, customers and professionals need to be confident on what are they getting into. Foster [2] also considers this business model characteristic in cloud computing, saying that "as for Utility Computing, it is not a new paradigm of computing infrastructure; rather, it is a business model in which computing resources, such as computation and storage, are packaged as metered services similar to a physical public utility, such as electricity and public switched telephone network." Also, it is possible to visualize that major concerns are about data: how and where it will be kept, who will be able to access it and which regulations will have it as subject. Considering that, data confidentiality, integrity and availability will be discussed in more details in the next section.

A. Confidentiality, Integrity and Availability (CIA): the big concern

According to [11],"storing data remotely into the cloud in a flexible on-demand manner brings appealing benefits: relief of the burden for storage management, universal data access with independent geographical locations, and avoidance of capital expenditure on hardware, software, and personnel maintenances, etc." However, users face the situation of losing control of their data. For example, he no longer has physical possession of the outsourced data and may not get to know about data loss and leakage incidents, if the cloud provider for some reason acts unfaithfully and decide not to report the incident [11], just to cite a few. Having that in mind, in the remainder of this section we discuss some of the traditional ways of delivering data confidentiality, integrity and availability.

Cryptography

One could ask if applying some cryptographic and backup schema would not solve at least part of the problem. This is a question certainly being target of studies, especially because as we cited before, cryptographic accidents are able to make data unusable. Trying to contribute to the subject, we bring some questions to this discussion.

- If using cryptography, how the key management is done?
  - One key for each customer?
  - One key to all customers?
  - Multiple keys for the same customer?
- What are the current cryptographic systems more applicable to the cloud computing characteristics, especially data storage?
- Last but not least, in which situations cryptography should be used?

We think that the cloud provider should have a detailed cryptographic plan, explaining what algorithms will be used, how the key management will be done, when encryption will be used and so on. The Cloud Security Alliance Guide [6] provides some guidance in these questions. As stated by them, cloud computing divorces components from location and this creates security issues that result from this lack of any perimeter. Hence there is only one way to secure the computing resources: strong encryption and scalable key management. Also according to [6], cloud customers and providers must encrypt all data in transit, at rest or on backup media, since all communications and all storage may be visible to arbitrary outsiders. Customers and providers want to encrypt their data to ensure integrity and confidentiality as also to avoid having to report incidents to their users (remembering that a provider's customer may have their own customer to report and successively).

According to [10], users are "universally required to accept the underlying premise of trust.", highlighting that although some take trust as synonymous of security, it is not and in security the element of trust is more apparent. Relating to the classic key concepts of information security, the CIA, [10] lists the minimum capabilities that should be offered by the cloud storage provider:

- "a tested encryption schema to ensure that the shared storage environment safeguards all data;
- stringent access controls to prevent unauthorized access to the data; and
- scheduled data backup and safe storage of the backup media [10]."

Wang [11] proposes public auditability for cloud data storage security. Such audit would be done by a third party auditor, called TPA. Knowing that such data in general, due to privacy issues, cannot be subject of disclosure, [11] lists two fundamental requirements for the TPA: 1) efficient cloud data storage auditing without demanding the local copy of data and without additional on-line burden to the cloud user; 2) no new vulnerabilities should be brought to user data privacy by the auditing process. Such requirements are best practices as also are the reasons [11] mention for the TPA being a good choice instead of the own user auditing the correctness of their data: 1) possible large size of stored data; 2) possible user' computer resource constraints; 3) "simply downloading the data for its integrity verification is not a practical solution due to the expensiveness in I/O cost and transmitting the file across the network."

Backup and recovery

Backup is probably the more traditional way of keeping data for recovery purposes. However, being crucial to ensure that a point-in-time data is available to restore business operations and given the special nature of a cloud environment, some questions need to be clearly answered by the provider and understood by the customer:

- Who performs the backup?
- How frequent the backup is performed?
- Who is responsible for storing the backup?
- Which backup format is used? Is it dependent of a specific technology?
- Logical segregation of data is maintained through the backup execution?

Having these questions being done, another important issue is if the provider will be able to meet any specific customer backup requirement. Normally, to have an effective backup and recovery strategy, a careful study of organization's need have to be done. Being the cloud a multi-tenant environment, it is possible that the cloud provider specific backup and recovery plan will not fit completely to the customer's need. Also, as mentioned before, the data should be encrypted on the backup media. According to [6], as a customer and provider of data, it is customer's responsibility to verify that such encryption takes place.

B. *Data format standards*

It seems vital to data availability to have a data format that allows customers to take their data from one provider and leverage it inside another provider's application. This kind of concern, however, is neither new or exclusively of cloud computing, so what was already learnt or developed since the beginning of the Internet and the need of data exchange should be taken into account when addressing this situation. Some standardizations initiatives are in progress, like the Cloud Computing Effort announced on April 27, 2009 by DMTF (Distributed Management Task Force) [12].

We do not know what would be the better in terms of data storage specifically and this is not the focus of this discussion. Maybe dictating a specific format is not a viable idea, at least not in a short time. But, the data interchange should be specified in some standard or well accepted format. The XML (Extensible Markup Language) format was designed to store and transport data. As cited in [13], XML is a technology that started a decade ago and since then great effort has been done by the research and industrial community to support XML and related technologies in RDBMS (Relational Database Management System). Also, being the format subject of standardization and widely adoption, lots of research aiming to secure the format has already been developed and it still is subject of ongoing improvements.

Adopting XML or not, the groups cited here and many others working on cloud computing standardizations should have this in mind: data must be interchangeable.

IV. Final Considerations

Maybe the cloud will evolve and become the largest information system we ever saw, having all sort of data and dealing with all kind of information, all kind of sensitive information. So, much research work is in progress to provide security for cloud computing, especially regarding do data confidentiality, integrity and availability. The general believe, including ours, is that the larger adoption of cloud computing relies on how secure it is and that security should be addressed since the very beginning. Being cloud computing a still evolving paradigm, some new security concerns may appear during the definition process, but the concerns highlighted in the present survey probably will not change. There are, however, a lot of good research and work in progress aiming to mitigate or to solve the security issues and to turn the cloud computing horizon less cloudy. Among these researches are government initiatives, like the cloud security group from US National Institute of Standards and Technology (NIST) and industry initiatives, like the Cloud Computing Security Alliance. Having data confidentiality, integrity and availability a strong legal side, some legal organizations like Strafford Publications are organizing events to discuss the subject, like a Teleconference entitled "Cloud Computing: Managing the Legal Risks" [14], showing that other areas beyond information technology are watching cloud computing growing adoption more closely. Such initiatives bring advantages for the customer that can have more qualified background when analyzing the available cloud computing solutions to migrate his services to a cloud.

References

[1] M. A. Vouk, "Cloud Computing – Issues, research and implementations", In: 30th International Conference on Information Technology Interfaces, pp. 31-40, 2008.

[2] I. Foster, I. Yong Zhao Raicu, and S. Lu, "Cloud Computing and Grid Computing 360-Degree Compared", In: 2008 Grid Computing Environments Workshop, pp. 1-10, 2008.

[3] G. Lin, D. Fu, J. Zhu, and G. Dasmalchi, "Cloud Computing: IT as a Service," IT Professional, vol. 11, no. 2, pp. 10-13, Mar./Apr. 2009.

[4] R. Buyya, Y. Chee Shin, S. Venugopal, "Market-Oriented Cloud Computing: Vision, Hype, and Reality for Delivering IT Services as Computing Utilities", In: 10th IEEE Conference on High Performance Computing and Communications. IEEE Computer Society, 2008, pp. 5-13.

[5] P. Mell, and T. Grance, "Cloud computing definition". NIST, June 2009, http://csrc.nist.gov/groups/SNS/cloud-computing/index.html

[6] Security Guidance for Critical Areas of Focus in Cloud Computing, www.cloudsecurityalliance.org (last access on Dec. 2010).

[7] J. Heiser, and M. Nicolett, Assessing the Security Risks of Cloud Computing. http://www.gartner.com/DisplayDocument?id=685308

[8] http://www.forbes.com/technology/ (last access on Dec. 2010).

[9] http://www.google.com/privacy/privacy-policy.html (last access on Dec. 2010).

[10] L. M. Kaufman, "Data security in the world of cloud computing," IEEE Security & Privacy Magazine, vol. 7, no. 4, pp. 61-64, July 2009. [Online]. Available: http://dx.doi.org/10.1109/MSP.2009.87

[11] C. Wang, Q. Wang, K. Ren, and W. Lou, "Privacy-preserving public auditing for data storage security in cloud computing," in *2010 Proceedings IEEE INFOCOM*. IEEE, March 2010, pp. 1-9.

[12] http://www.dmtf.org/ (last access on Dec. 2010).

[13] R. Zhen Hua Liu Murthy, "A Decade of XML Data Management: An Industrial Experience Report from Oracle", In: IEEE 25th International Conference on Data Engineering, pp. 1351-1362. IEEE Computer Society, 2009.

[14] http://www.straffordpub.com/products/cloud-computing-managing-the-legal-risks-2009-12-09 (last access on Dec. 2010).