Privacy Issues In Cloud Computing

Leela Krishna Ganapavarapu[1], Sireesha Prathigadapa[2], Kiranmayi Bodapati[3]
[1,2,3]*Malaysia*

Abstract*: Cloud computing is the delivery of computing as a service rather than a product. It provides shared resources, software, and information to computers and other devices over a network. The increasing network bandwidth and reliable yet flexible network connections make it even possible that users can now subscribe high quality services from data and software that reside solely on remote data centres. We can store and retrieve the data as we like using cloud computing. The cloud computing paradigm changes the way in which information is managed, especially where personal data processing is concerned. End-users can access cloud services without the need for any expert knowledge of the underlying technology. This is a key characteristic of cloud computing, which offers the advantage of reducing cost through the sharing of computing and storage resources, combined with an on-demand provisioning mechanism based on a pay-per-use business model. These new features have a direct impact on the IT budget and cost of ownership, but also bring up issues of traditional security, trust and privacy mechanisms.Privacy, in this Article, refers to the right to self-determination, that is, the right of individuals to 'know what is known about them', be aware of stored information about them, control how that information is communicated and prevent its abuse. In other words, it refers to more than just confidentiality of information. Protection of personal information (or data protection) derives from the right to privacy via the associated right to self-determination. Every individual has the right to control his or her own data, whether private, public or professional. Privacy issues are increasingly important in the online world. It is generally accepted that due consideration of privacy issues promotes user confidence and economic development. However, the secure release, management and control of personal information into the cloud represent a huge challenge for all stakeholders, involving pressures both legal and commercial. This study analyses the challenges posed by cloud computing and the standardization work being done by various standards development organizations (SDOs) to mitigate privacy risks in the cloud, including the role of privacy-enhancing technologies (PETs).*

Key Words: *Cloud Storage Services, data integrity, dependable distributed storage, data dynamics, Cloud Computing, Privacy.*

I. Introduction

Few years ago, people used to carry their documents around on disks. Then, more recently, many people switched to memory sticks. Cloud computing refers to the ability to access and manipulate information stored on remote servers, using any Internet-enabled platform, including smart phones. Computing facilities and applications will increasingly be delivered as a service, over the Internet. We are already making use of cloud computing when, for example, we use applications such as Google Mail, Microsoft Office365 1 or Google Docs. In the future, governments, companies and individuals will increasingly turn to the cloud. "Cloud computing is 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."

This definition, the cloud model promotes availability and is composed of five essential characteristics, three delivery models and four deployment models.

The five key characteristics of cloud computing are on-demand self service, ubiquitous network access, location-independent resource pooling, rapid elasticity and measured service, all of which are geared towards seamless and transparent cloud use. Rapid elasticity enables the scaling up (or down) of resources. Measured services are primarily derived from business model properties whereby cloud service providers control and optimize the use of computing resources through automated resource allocation, load balancing and metering tools.

As per the definition provided by the National Institute for Standards and Technology (NIST) (Badger etal., 2011), "*cloud computing* is 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". It represents a paradigm shift in information technology many of us are likely to see in our lifetime. While the customers are excited by the opportunities to reduce the capital costs, and the chance to divest themselves of infrastructure management and focus on core competencies, and above all the agility offered by the on-demand provisioning of computing, there are issues and challenges which need to be addressed before a ubiquitous adoption may happen.

Cloud computing refers to both the applications delivered as services over the Internet and the hardware and systems software in the data centers that provide those services. There are four basic cloud delivery models, as outlined by NIST (Badger et al., 2011), based on who provides the cloud services.

The agencies may employ one model or a combination of different models for efficient and optimized delivery of applications and business services. These four delivery models are: (i) *Private cloud* in which cloud services are provided solely for an organization and are managed by the organization or a third party. These services may exist off-site. (ii) *Public cloud* in which cloud services are available to the public and owned by an organization selling the cloud services, for example, Amazon cloud service. (iii) *Community cloud* in which cloud services are shared by several organizations for supporting a specific community that has shared concerns (e.g., mission, security requirements, policy, and compliance considerations). These services may be managed by the organizations or a third party and may exist offsite. A special case of community cloud is the Government or G-Cloud. This type of cloud computing is provided by one or more agencies (service provider role), for use by all, or most, government agencies (user role). (iv) *Hybrid cloud* which is a composition of different cloud computing infrastructure (public, private or community). An example for hybrid cloud is the data stored in private cloud of a travel agency that is manipulated by a program running in the public cloud.

From the perspective of service delivery, NIST has identified three basic types of cloud service offerings. These models are: (i) *Software as a service* (SaaS) which offers renting application functionality from a service provider rather than buying, installing and running software by the user. (ii) *Platform as a service* (PaaS) which provides a platform in the cloud, upon which applications can be developed and executed. (iii) *Infrastructure as a service* (IaaS) in which the vendors offer computing power and storage space on demand.

From a hardware point of view, three aspects are new in the paradigm of cloud computing (Armbrust etal., 2009). These aspects of cloud computing are: (i) The illusion of infinite computing resources available on demand, thereby eliminating the need for cloud computing users to plan far ahead for provisioning. (ii)The elimination of an up-front commitment by cloud users, thereby allowing companies to start small and increase hardware resources only when there is an increase in their needs. (iii) The ability to pay for use of computing resources on a short-term basis as needed and release them when the resources are not needed, thereby rewarding conservation by *letting machines and storage go when they are no longer useful*. In a nutshell, cloud computing has enabled operations of large-scale data centers which has led to significant decrease in operational costs of those data centers. On the consumer side, there are some obvious benefits provided by cloud computing. A painful reality of running IT services is the fact that in most of the times, peak demand is significantly higher than the average demand. The resultant massive over-provisioning that the companies usually do is extremely capital-intensive and wasteful. Cloud computing has allowed and will allow even more seamless scaling of resources as the demand changes. In spite of the several advantages that cloud computing brings along with it, there are several concern sand issues which need to be solved before ubiquitous adoption of this computing paradigm happens. First in cloud computing, the user may not have the kind of control over his/her data or the performance of his/her applications that he/she may need, or the ability to audit or change the processes and policies under which he/she must work. Different parts of an application might be in different place in the cloud that can have an adverse impact on the performance of the application. Complying with regulations may be difficult especially when talking about cross-border issues – it should also be noted that regulations still need to be developed to take all aspects of cloud computing into account. It is quite natural that monitoring and maintenance is not as simple a task as compared to what it is for PCs sitting in the Intranet.

Second, the cloud customers may risk losing data by having them locked into proprietary formats and may lose control over their data since the tools for monitoring who is using them or who can view them are not always provided to the customers. Data loss is, therefore, a potentially real risk in some specific deployments. Third, it may not be easy to tailor *service-level agreements* (SLAs) to the specific needs of a business. Compensation for downtime may be inadequate and SLAs are unlikely to cover the concomitant damages. It is sensible to balance the cost of guaranteeing internal uptime against the advantages of opting for the cloud. Fourth, leveraging cost advantages may not always be possible always. From the perspective of the organizations, having little or no capital investment may actually have tax disadvantages. Finally, the standards are immature and insufficient for handling the rapidly changing and evolving technologies of cloud computing. Therefore, one cannot just move applications to the cloud and expect them to run efficiently. Finally, there are latency and performance issues since the Internet connections and the network links may add to latency or may put constraint on the available bandwidth.

II. ARCHICTECTURE OF CLOUD COMPUTING

In this section, we present a top-level architecture of cloud computing that depicts various cloud service delivery models. Cloud computing enhances collaboration, agility, scale, availability and provides the potential for cost reduction through optimized and efficient computing. More specifically, cloud describes the use of a collection of distributed services, applications, information and infrastructure comprised of pools of compute, network, information and storage resources (CSA Security Guidance, 2009). These components can be rapidly orchestrated, provisioned, implemented and decommissioned using an on demand utility-like model of allocation and consumption. Cloud services are most often, but not always, utilized in conjunction with an enabled by virtualization technologies to provide dynamic integration, provisioning, orchestration, mobility and scale. While the very definition of cloud suggests the decoupling of resources from the physical affinity to and location of the infrastructure that delivers them, many descriptions of cloud go to one extreme or another by either exaggerating or artificially limiting the many attributes of cloud. This is often purposely done in an attempt to inflate or marginalize its scope. Some examples include the suggestions that for a service to be cloud-based, that the Internet must be used as a transport, a web browser must be used as an access modality or that the resources are always shared in a multi-tenant environment outside of the “perimeter.” What is missing in these definitions is context.

From an architectural perspective, given this abstracted evolution of technology, there is much confusion surrounding how cloud is both similar and different from existing models and how these similarities and differences might impact the organizational, operational and technological approaches to cloud adoption as it relates to traditional network and information security practices. There are those who say cloud is a novel sea-change and technical revolution while other suggests it is a natural evolution and coalescence of technology, economy and culture. The real truth is somewhere in between.

There are many models available today which attempt to address cloud from the perspective of academicians, architects, engineers, developers, managers and even consumers. The architecture that we will focus on this chapter is specifically tailored to the unique perspectives of IT network deployment and service delivery.

Cloud services are based upon five principal characteristics that demonstrate their relation to, and differences from, traditional computing approaches (CSA Security Guidance, 2009). These characteristics are: (i) abstraction of infrastructure, (ii) resource democratization, (iii) service oriented architecture, (iv) elasticity/dynamism, (v) utility model of consumption and allocation.

Abstraction of infrastructure: The computation, network and storage infrastructure resources are abstracted from the application and information resources as a function of service delivery. Where and by what physical resource that data is processed, transmitted and stored on becomes largely opaque from the perspective of an application or services’ ability to deliver it. Infrastructure resources are generally pooled in order to deliver service regardless of the tenancy model employed – shared or dedicated. This abstraction is generally provided by means of high levels of virtualization at the chipset and operating system levels or enabled at the higher levels by heavily customized file systems, operating systems or communication protocols.

Resource democratization: The abstraction of infrastructure yields the notion of resource democratization-whether infrastructure, applications, or information – and provides the capability for pooled resources to be made available and accessible to anyone or anything authorized to utilize them using standardized methods for doing so.

Service-oriented architecture: As the abstraction of infrastructure from application and information yields well-defined and loosely-coupled resource democratization, the notion of utilizing these components in whole or part, alone or with integration, provides a services oriented architecture where resources may be accessed and utilized in a standard way. In this model, the focus is on the delivery of service and not the management of infrastructure.

Elasticity/dynamism: The on-demand model of cloud provisioning coupled with high levels of automation, virtualization, and ubiquitous, reliable and high-speed connectivity provides for the capability to rapidly expand or contract resource allocation to service definition and requirements using a self service model that scales to as-needed capacity. Since resources are pooled, better utilization and service levels can be achieved.

Utility model of consumption and allocation: The abstracted, democratized, service-oriented and elastic nature of cloud combined with tight automation, orchestration, provisioning and self-service then allows for dynamic allocation of resources based on any number of governing input parameters. Given the visibility at an atomic level, the consumption of resources can then be used to provide a metered utility cost and usage model. This facilitates greater cost efficacies and scale as well as manageable and predictive costs.

Cloud Service Delivery Models

Three archetypal models and the derivative combinations thereof generally describe cloud service delivery. The three individual models are often referred to as the “SPI MODEL”, where “SPI” refers to Software, Platform and Infrastructure (as a service) respectively (CSA Security Guidance, 2009).

Software as a Service (SaaS): The capability provided to the consumer is to use the provider’s applications running on a cloud infrastructure and accessible from various client devices through a thin client interface such as web browser. In other words, in this model, a complete application is offered to the customer as a service on demand. A single instance of the service runs on the cloud and multiple end users are services. On the customers’ side, there is no need for upfront investment in servers or software licenses, while for the provider, the costs are lowered, since only a single application needs to be hoste and maintained. In summary, in this model, the customers do not manage or control the underlying cloud infrastructure, network, servers, operating systems, storage, or even individual application capabilities, with the possible exception of limited user-specific application configuration settings. Currently, SaaS is offered by companies such as Google, Sales force, Microsoft, Zoho etc.

Platform as a Service (PaaS): In this model, a layer of software or development environment is encapsulated and offered as a service, upon which other higher levels of service are built. The customer has the freedom to build his own applications, which run on the provider’s infrastructure. Hence, a capability is provided to the customer to deploy onto the cloud infrastructure customer-created applications using programming languages and tools supported by the provider (e.g., Java, Python, .Net etc.). Although the customer does not manage or control the underlying cloud infrastructure, network, servers, operating systems, or storage, but he/she has the control over the deployed applications and possibly over the application hosting environment configurations. To meet manageability and scalability requirements of the applications, PaaS providers offer a predefined combination of operating systems and application servers, such as LAMP (Linux, Apache, MySql and PHP) platform, restricted J2EE, Ruby etc.

Some examples of PaaS are: Google’s App Engine, Force.com, etc.

Infrastructure as a Service (IaaS): This model provides basic storage and computing capabilities as standardized services over the network. Servers, storage systems, networking equipment, data center space etc. are pooled and made available to handle workloads. The capability provided to the customer is to rent processing, storage, networks, and other fundamental computing resources where the customer is able to deploy and run arbitrary software, which can include operating systems and applications. The customer does not manage or control the underlying cloud infrastructure but has the control over operating systems, storage, deployed applications, and possibly select networking components (e.g., firewalls, load balancers etc.). Some examples of IaaS are: Amazon, GoGrid, 3 Tera etc. Understanding the relationship and dependencies between these models is critical. IaaS is the foundation of all cloud services with PaaS building upon IaaS, and SaaS-in turn – building upon PaaS. Architecture of cloud layer model is depicted in Figure 1.

Challenges to privacy in cloud computing:

The promise to deliver IT as a service is addressed a large range of consumers, from small and medium-sized enterprises (SMEs) and public administrations to end-users. According to industry analysts, the ICT sector is poised for strong growth of cloud services. Users are creating an ever-growing quantity of personal data. IDC predicts that the "digital universe" – the amount of information and content created and stored digitally – will grow from 1.8 zeta bytes (ZB) in 2011 to over 7 ZB by 2015.

This expanding quantity of personal data will drive demand for cloud services, particularly if cloud computing delivers on the promises of lower costs for customers and the emergence of new business models for providers.

Among the main privacy challenges for cloud computing is:

a) Complexity of risk assessment in a cloud environment
b) Emergence of new business models and their implications for consumer privacy
c) Achieving regulatory compliance.

Complexity of risk assessment

The complexity of cloud services introduces a number of unknown parameters. Service providers and consumers are cautious, respectively, about offering guarantees for compliance-ready services and adopting the services. With service providers promoting a simple way to flow personal data irrespective of national boundaries, a real challenge arises in terms of checking the data processing life cycle and its compliance with legal frameworks.

In a cloud service, there are many questions needing to be addressed in order to determine the risks to information privacy and security:

- Who are the stakeholders involved in the operation?
- What are their roles and responsibilities?
- Where is the data kept?
- How is the data replicated?
- What are the relevant legal rules for data processing?
- How will the service provider meet the expected level of security and privacy?

To address these issues, the Madrid Resolution states that every responsible person shall have transparent policies with regard to the processing of personal data. Stakeholders need to specify requirements for cloud computing that meet the expected level of security and privacy. In Europe, the European Network and Information Security Agency (ENISA) provides recommendations to facilitate understanding of the shift in the balance of responsibility and accountability for key functions such as governance and control over data and IT operations and compliance with laws and regulations.

Emergence of new business models and implications for consumer privacy

A report by the Federal Trade Commission (FTC) on "*Protecting consumer privacy in an era of rapid change*" analyses the implications for consumer privacy of technological advances in the IT sphere. According to FTC, users are able to collect, store, manipulate and share vast amounts of consumer data for very little cost.

These technological advances have led to an explosion of new business models that depend on capturing consumer data at a specific and individual level and over time, including profiling, online behavioural advertising (OBA), social media services and location-based mobile services.

III. Conclusion

- Cloud computing is still in its infancy. This is an emerging technology which will bring about innovations in terms of business models and applications. The widespread penetration of smart phones will be a major factor in driving the adoption of cloud computing. However, cloud computing faces challenges related to privacy and security.
- The global dimension of cloud computing requires standardized methodologies and technical solutions to enable stakeholders to assess privacy risks and establish adequate protection levels. From a business point of view, privacy should represent an opportunity for cloud providers to promote brand image and differentiate services. However, privacy challenges require the involvement of a wide range of stakeholders to cover multidisciplinary approaches benefiting all areas of society. Robust privacy protection needs interoperable built-in privacy components capable of ensuring compliance with principles such as data minimization in complex architectures. Privacy standards will play an important role in fostering the adoption of cloud services by promoting social responsibility and addressing privacy challenges. The implementation of PETs is seen as a good mechanism by data protection authorities to protect the data subject's rights and meet privacy principle objectives. In cloud services, the implementation of PETs will depend on the availability of standards to assess privacy risks and describe means of ensuring data protection compliance. PETs can ensure that breaches of the data protection rules and violations of individuals’ rights are not only forbidden and subject to sanctions, but are also a technically daunting undertaking. The embedding of privacy by design features when designing technologies is increasingly supported by regulators and is also being included in the reform of the EU Data Protection Directive.
- Cybercriminal activities impacting cloud computing environments −for example, fraud and malicious hacking are threats that can undermine user confidence in the cloud. Cloud computing providers face multiple, and potentially conflicting, laws concerning disclosure of information. Achieving a better understanding of jurisdictional issues is critical and should be tackled through enhanced dialogue.
- ITU could have an enabling role to play in developing technical standards, guidelines and methodologies for implementing privacy by design principles, including assessment of risks to personal information in the cloud. These can be used as best practices by service providers in order to ensure compliance with legal frameworks for personal information protection. ITU could consider organizing an event on this topic to promote the standards work being done in this area. ITU-T SG 17 has taken the initiative, through a number of study Questions, to work on specific topics related to cloud security. However, a good deal of work

remains to be done in the area of cloud privacy. Cloud security is set to form a major part of SG 17's future work, while extensive collaboration with other standardization bodies and industry groups would help to expedite progress and avoid duplication of effort.

References

[1] Grilo A, Jardim-Goncalves R (2011) Challenging electronic procurement inthe AEC sector: A BIM-based integrated perspective. Automation in Construction 20: 107–114

[2] Vorakulpipat C, Rezgui Y, Hopfe CJ (2010) Value creating construction virtual teams: A Case study in the construction sector. Automation in Construction 19: 142–147

[3] International Standard (ISO) 16739:2005 - Industry Foundation Classes

[4] Succar B (2009) Automation in Construction 18: 357–375

[5] Royal Institute of British Architects (RIBA) Plan of Work. Available at: http://www.ribaplanofwork.com/. [Last accessed: Jan 17, 2012]

[6] Serror M (2007) Shared computer-aided structural design model for construction industry (infrastructure). Comput Aided Des 40: 778–788

[7] Singh V (2010) A theoretical framework of a BIM-based multi-disciplinary platform. Automation in Construction 20: 131–144

[8] Rezgui Y, Cooper G, Brandon P (1998) Information management in a collaborative multiactor environment: the COMMIT Approach. J Comput Civil Eng 12: 136–145

[9] Rezgui Y, Hopfe C, Vorakulpipat C (2010) Generations of knowledge management in the architecture, engineering and construction industry: an evolutionary perspective. Adv Eng Inform 24: 219–228

[10] British Standards Institute (BSI): British Standard 1192:2007 – Collaborative production of architectural, engineering and construction information, Code of practice

[11] Cooper G, Cerulli C, Lawson BR, Peng C, Rezgui Y (2010) Tracking decision-making during architectural design. Electron J Inf Technol Construction 10: 125–139

[12] Kim H, Chaudhari S, Parashar M, Martyy C (2009) Online Risk Analytics on the Cloud International Workshop on Cloud Computing In: conjunction with the 9th IEEE International Symposium on Cluster Computing and the Grid, pp pp 484–489

a. CometCloud – available at: http://nsfcac.rutgers.edu/CometCloud/.

[13] IBM WebSphere Cast Iron Cloud Integration – available at: http://www-01.ibm.com/software/integration/cast-iron-cloud-integration/. [Last accessed: Jan 10, 2013]

[14] Autodesk Revit Architecture – available at: http://www.autodesk.com/ products/autodesk-revit-family/overview. [Last accessed: Jan 15, 2012]

[15] 3D Modelling Software from Google (version 8) – available at: http:/www.sketchup.com/. [Last accessed: Jan. 15, 2012]

[16] bimserver.org, http://www.bentley.com/en-US/Products/projectwise+project+team+

a. collaboration/