Cloud Computing: Issues in Data Mobility and Security

Zaigham Mahmood
School of Computing and Mathematics
University of Derby
Derby, UK
e-mail: z.mahmood@derby.ac.uk

Harjinder Singh Lallie
School of Computing and Mathematics
University of Derby
Derby, UK
e-mail: h.s.lallie@derby.ac.uk

Abstract— **Cloud Computing is a generic term for anything that involves delivering hosted services over the internet, based on a pay-as-you-go approach. Cloud Computing offers numerous benefits and, therefore, many large enterprises have embraced the cloud technologies and infrastructures. Vendors are also developing tools and applications to fulfill the demand and tap into the growing market. Like any new technology and paradigm, there are also numerous issues and concerns, including security and availability of data. This paper explores some of the security issues surrounding data location, mobility and availability as well as issues relating to the security of data at rest. The aim is to provide some useful background information for enterprises preparing to take advantage of Cloud Computing paradigm.**

Keywords- cloud computing; enterprise computing; cloud security; data privacy; data security

I. INTRODUCTION

Cloud Computing is a generic term for anything that involves delivering hosted services and computing resources over the Internet. It is 'a style of computing where massively scalable IT-enabled capabilities are delivered *as services* to external customers using Internet technologies' [1]. According to NIST (National Institute of Standards and Technology, US), it provides 'a convenient, on-demand network access to a shared pool of computing resources' [2, 3]. Here, resources refer to computing applications, software services, platforms and computing infrastructures. Forrester [4] suggests that Cloud Computing refers to 'a pool of abstracted, highly scalable and managed infrastructure capable of hosting end-customer applications and billed by consumption'. It is the latest paradigm in distributed computing that promises to revolutionize IT and business by making computing available as a *utility* over the World Wide Web.

General public have been using Cloud Computing in the form of Internet services like *Hotmail* (since about 1996), *YouTube* (since about 2005), *Facebook* (since about 2006) and *Gmail* (since about 2007). Hotmail is probably the first cloud computing application that allowed the general public to keep their data in the form of text and files at the vendor's servers. Since then, many other services have emerged that allow users to store information (such as text files, photographs, video clips and music) and perform processing without paying any upfront fees. Some well known examples include: *Twitter*, *Myspace*, *Wikipedia* and *Google docs*. These are typically consumer oriented services, different from enterprise-oriented tasks, but the underlying principles are the same i.e. to provide the storage space and processing capability. In the commercial arena, *Amazon.com* was one of the first vendors to provide storage space, computing resources and business functionality following the cloud computing model. *Salesforce.com*, founded in 1999, pioneered the concept of delivering enterprise applications as services to enterprises. In 2002, *Amazon Web Services* provided a suite of cloud-based services and, later in 2006, it launched *Elastic Compute Cloud* (EC2) that allowed companies and individuals to rent computers on which to run their own enterprise applications. The number of cloud services providers and the applications, platforms and infrastructures are increasing at such a rate that, in 2009, Gartner listed Cloud Computing as number 1 in its top 10 strategic technology areas for 2010 [10, 11]. The reasons why more and more companies are planning to take advantage of IT Cloud Computing solutions include:

- reduced costs associated with delivering IT services
- reduced management responsibilities
- increased business efficiency and agility
- easy access to software and hardware resources available elsewhere
- no long term contracts with vendors.

A report on Cloud Computing published in Jan 2010 [12] suggests that: 1) Enterprises are now moving beyond experimentation; 2) they are beginning to develop and deploy management software to deal with scaled Cloud environments; and 3) they are beginning to develop enterprise-level policies and standards for dealing with *Public* and *Hybrid* Clouds.

In the rest of this paper, we first outline the benefits that Cloud Computing offers and briefly discuss the deployment approaches. Then, in Sections III and IV, we discuss, in some detail, the inherent issues with respect to mobility and security of data held on Clouds. The last section presents a brief conclusion.

Copyright (c) IARIA, 2011. ISBN: 978-1-61208-124-3

II. CLOUD COMPUTING

A. The Promise

Large vendors like IBM, Dell, Oracle and Sun have started to take strong positions with respect to Cloud Computing provision [5]. The essential features of this latest paradigm include [2, 6]:

- On-demand services: to enable users to avail of computing capabilities as and when required
- Resource pooling: to allow dynamically assigned computing resources to serve multiple consumers
- Rapid elasticity: to allow services, resources and infrastructures to be automatically provisioned
- Measured provision: to provide a metering capability to determine the usage for pricing purposes
- Effective management: to provide and facilitate easy monitoring, controlling and reporting.

Cloud Computing is an attractive paradigm that can be massively scalable. It provides benefits of efficiency, flexibility and high utilization that, in turn, can result in reduced capital investment costs and lower operational expenditure. The Cloud offerings from service providers and vendors are continuing to mature and increase in number. With this, the cost savings are becoming particularly attractive. There is no doubt that Cloud Computing is making *supercomputing* available to the masses.

B. Deployment Approaches

Cloud Computing can be classified and deployed in a number of ways e.g. as *public*, *private* or *hybrid* clouds.

Public Clouds are networks where services are provided by third parties and hosted and managed by the service providers. The Cloud providers take on the responsibilities of installation, management, provisioning and maintenance. Consumers are charged only for the resources they use following a pay-as-you-go model.

Private Clouds are proprietary networks normally residing within the enterprises for the exclusive use of the organization or for a known group of consumers. In case of Private Clouds, the enterprise is in charge of maintaining the Cloud and also responsible for security and regulatory compliance issues. The issues of data security are, therefore, somewhat reduced.

Hybrid Clouds are a combination of Private and Public Clouds. In this case, the management responsibilities are often split between the enterprise and the Public Cloud providers, which can often become an issue of concern. For mission critical processes, this type of Cloud infrastructure is much more effective because of enhanced control and management by the enterprise itself.

The Cloud model consists of, typically, three components which refer to three types of services: *Software Services*, *Platform Services* and *Infrastructure Services*. These services may be defined as follows:

- Software as a Service (SaaS): referring to prebuilt and vertically integrated applications available for purchase or use by customers as *services*. Here, customers are looking to 'hire' easy-to-consume functionality.
- Platform as a Service (PaaS): referring to application development toolkits and deployment tools (e.g. application servers, portal servers and middleware) which clients make use of to build and deploy their own applications. Here, customers are looking to buy time and cost savings in deploying applications.
- Infrastructure as a service (IaaS): referring to hardware (e.g. servers, storage space, network devices, etc) to enable Cloud Platforms and Applications to operate. Here, customers are looking to hire *computing*. Since, the infrastructure is offered on pay-for-what-you-use basis, it is sometimes referred to as *utility computing*.

C. Inherent Issues

Notwithstanding the benefits that Cloud Computing offers, there are numerous issues and challenges for organizations embracing this new paradigm. Zhen [8] lists a number of major challenges with respect to the following: 1) governance, management and updating of data; 2) management of software services; 3) monitoring of products and processes; 4) reliability and availability of systems and infrastructure and 5) security of information and data. The Expert Group Report [9] mentions a number of issues including: 1) concerns over security with respect to valuable knowledge, information and data placed on an external service; 2) concerns over availability and business continuity; and 3) concerns over data transmission across anticipated broadband speeds. Other shortcomings, as mentioned by various researchers, include: 1) no native security attributes; 2) inadequate or no security provisioning by providers; 3) lack of understanding of cloud legal issues; and 4) the failure to recognize potential liability from either legal issues or a lack of security. Issues with respect to "control" are also real concerns. A closer examination reveals that the major concerns may be broadly classified as those relating to the following:

- Security, including reliability and availability
- Governance and Management.

In this paper, we discuss issues with respect to data mobility, security and availability. Other issues are discussed in a companion papers which is under preparation.

III. DATA MOBILITY AND SECURITY

Cloud Computing provides services with respect to enterprise applications (software components and systems), computing platforms (development tools) and infrastructures (hardware including servers). Befits are huge but the inherent issues are also many. Some of the major issues refer to the security of data. In this respect, there are

many dimensions including: data security, data privacy, data protection, data availability and data transmission. Forrester [22] combines these into three groups: 1) Security and Privacy; 2) Compliance; and 3) Legal and Contractual. Some of these are now discussed in some detail.

A. Data Mobility

Cloud Computing offers a high degree of data mobility in the sense that data stored on the Cloud may reside on a location geographically a long way away from the organization that owns the data. In a majority of cases, the owners and the users know where the data resides; however, this may not be true in all cases. Unless there is a contractual agreement that data should stay in a particular location or reside on a given known server, the Cloud providers may decide to keep it moving from one location to another. There are several reasons for this, including: 1) reducing the cost of storing data; 2) efficiency of retrieval of data; 3) easy availability of data; 4) efficient linking of different data resident on different locations; and 5) resource optimization. Security risks and issues are already big concerns. When data mobility is at a high level then the risks and issues increase many folds especially when data is transferred to another country with different regulations. High levels of data mobility also have negative implications for data security and data protection as well as data availability.

Many factors influence the choice of location for data centres including the cost as the cost of running a centre is high on the list of priorities [19]. The attraction of co-location and distribution of data is particularly justifiable due to the bandwidth efficiencies that this could provide. For instance Amazon's *CloudFront* data centres are located in the following cities: Ashburn Virginia, Dallas/Fort Worth, Los Angeles, Miami, Newark New Jersey, Palo Alto, California, Seattle, St. Louis, Amsterdam, Dublin, Frankfurt, London, Hong Kong and Tokyo. This does not necessarily mean that data stored on the Amazon Cloud may be split across any number of these data centres.

Another influencing factor is the cost relating to high capacity internet access.

B. Data Availability

Data availability is a major legitimate reason for the data to be stored in multiple locations on the Cloud. This answers a core business requirement: that of an uninterruptible service and seamless provision. Furthermore, data availability is such a crucial issue that it is common for Cloud providers to credit customer accounts if the system downtime duration drops below that specified in the SLA (service level agreement). The related issue is that, often, such measures are not specified in the SLAs.

The issue of data availability is exemplified by the outages suffered by Google's *Gmail* service in February 2009 which resulted in embarrassing headlines for the company [17]. In the subsequent service agreement for its *Premier Apps* range of products which also covers *Gmail*, Subsequently, Google promised that customer data availability will be at least 99.9% of the time in any calendar month [18].

C. Cost Relating to Data Mobility

Another reason for data mobility is to reduce the cost of running data centres (by reducing the electricity bills, for example). In Public Clouds, data is often routed to other locations at certain times of the day or year, or when there is a huge climatic temperature fluctuation [20]. The main factor in such considerations is the cost of provision. Qureshi [21] has conducted research into the dynamic routing of data based on the cost of electricity in various regions. This research shows that it is possible to reduce electricity costs by up to 40%. However, as electricity costs rise, Cloud service providers may look for more effective ways of reducing their overheads – at the same time, hopefully, ensuring that there is no compromise on performance and service availability. In this respect, Qureshi's method [21] of dynamically routing data would become an attractive solution. From the point of view of data security and data availability, this would exacerbate the security issues, which are already a major concern when data is being moved between locations. As mentioned before, data mobility or dynamic data routing is also considered as a result of resource optimization. This, in turn, also helps to reduce costs.

Enterprises consuming Cloud services may not be aware of this, however, as they become more knowledgeable, they may decide to request appropriate data relocation and, thus, negotiate lower contract prices for such data services.

D. Data Location Assurance

Data mobility and location concerns, including those relating to security, have been partially addressed by Cloud providers, and two of the largest vendors in the field have started offering solutions to customers. Amazon's *AWS* (Amazon Web Services) provides an option within its *S3* (Simple Storage Service) package to allow customers to specify the regions for the storage and location of their data. It also provides assurance that data will not leave the customer selected regions [23]. Although, the available locations are currently restricted to just three regions: US (Standard), EU (Ireland) and US-West (Northern California), the company has plans to expand into the Asia-Pacific region in 2010. Amazon is marketing this as a way of improving performance and providing a better customer-centric service.

In 2009, Microsoft announced that its Windows *Azure* system would provide its users with the option to specify where in the world they would wish their data to be stored. As well as performance gains, Microsoft also stated legal and regulatory reasons for this facility. This is an attractive facility as different countries have different laws with respect to data privacy and confidentiality and some clients may wish to exploit such differences to their advantage, although

Copyright (c) IARIA, 2011. ISBN: 978-1-61208-124-3

there may be legal implications. However, as with *AWS*, Microsoft has a very restricted choice of geographical locations, currently only two: both within the US. Microsoft has plans to expand this and is especially interested in sites outside the US.

E. Cross Border Data Transition

Cross border data transition can lead to potential legal risks due to different locations having varying policies, regulations and legislation. This means that data protected by legislation in one country may not have the same, or even similar, protection in another country [24]. In an example of this, that appears in Jaeger [19], it is noted that the European Union and United States of America have differing definitions of privacy as a result of disparate privacy policies. Their Data Protection Laws are based on the assumption that the location and responsibility of data is known and understood. Cloud Computing however challenges this presumption.

Presently, a vast majority of data centres are located in the United States [25]. A consequence of this is that data protection and privacy concerns are influenced by the USA Patriot Act 2001, the Foreign Intelligence Surveillance Act (FISA amendments act of 2008), the Electronic Communications Privacy Act (1986), the Privacy Act (1974) and the Homeland Security Act (2002). Under the ruling of these acts, the FBI and similar agencies have the regulatory power to demand access to any data stored on any computer within the USA, even if it is stored on behalf of another jurisdiction [25].

F. Organisations' Response

One of the key recommendations made by Gartner [26] suggests that assurance should be given guaranteeing that customer data will be stored and processed within a certain jurisdiction and that the local laws within that jurisdiction would apply. However, this may conflict with the concept of data privacy particularly in countries such as the US. This means that consumer data stored within the US may be highly vulnerable to disclosure [24] which may pose a potential business and/or economic risk.

World governments and organizations are developing strategies to counter these concerns e.g. the Canadian Government does not allow the use of hosting services that are based in the US [25]. Similarly, SWIFT, an international banking firm, are locating data centres in Switzerland where the data protection regulations are based on EU laws providing specific conditions for the transference of personal data to third parties or abroad [27].

The UK Data Protection Act 1998 requires that personal information is handled in a manner that ensures that key principles and legal obligations are properly adhered to. These principles also include the restriction on transferring data to countries outside of the European Economic Area (the EEA) unless there is a clear and adequate data protection mechanism [28]. The act also makes it an obligation for companies to clearly state where customers' data is being held. However, this may be difficult when providers themselves are unaware of the exact locations.

IV. SECURING DATA AT REST

A valid question with respect to security of data on the Cloud is: how to ensure security of *data at rest*. The obvious answer suggests that data should be encrypted. Unfortunately, this is not as simple as it appears. If the data is being stored by an *IaaS* service (such as Amazon's *Simple Storage Service*, also known as S3), that is not associated with a specific application, then encryption is appropriate and indeed possible and a valid solution. However, data on the Cloud being processed by *SaaS* or *PaaS* applications (such as *Salesforce.com* or *Google Apps*) is generally not considered suitable for encryption. This is because encryption prevents indexing or searching of data, which has implications on availability and access of such data.

Finding an appropriate mechanism that is secure, and allows both addition and multiplication, has proved elusive and many respected cryptologists have suggested that it may not even be possible. However, a number of techniques and schemes have subsequently been put forward favoring the full homomorphic encryption [29]. A fully homomorphic cryptosystem is one where the performance of a mathematical operation on ciphertext is found to have a regular effect on the corresponding plaintext.

One such scheme, developed by Gentry [30], allows data to be processed without being decrypted. This means that a Cloud service provider can perform computations on client's data without exposing the original data. Gentry's proposal has caused great interest amongst cryptographers who have been trying to develop a practical manifestation of the concept of privacy homomorphism for over thirty years [16].

Other research efforts are focusing on methods to limit the amount of data that needs to be decrypted for processing in the Cloud. An example is predicate encryption, a type of asymmetric encryption where different individuals or groups can selectively decrypt some of the encrypted data instead of decrypting all of it [22]. Methodologies are being developed.

V. CONCLUSION

Cloud Computing is 'essentially on-demand access to a shared pool of computing resources'. It helps consumers to reduce costs, reduce management responsibilities and increase business agility. For this reason, it is becoming a popular paradigm and increasingly more companies are shifting toward IT Cloud Computing solutions. Advantages are many but there are also challenges and inherent issues. Generally, these relate to data governance, service management, process monitoring, infrastructure reliability, information security, data integrity and business continuity. This paper focuses on the mobility and availability of data held on the Cloud and discusses the security issues of such data.

Copyright (c) IARIA, 2011. ISBN: 978-1-61208-124-3

In spite of the limitation and issues as discussed in the previous sections, Cloud Computing is becoming an attractive paradigm for large enterprises. In 2008, Forrester [5] predicted that 'cloud computing initiatives could affect the enterprises within 2 to 3 years as it has the potential to significantly change IT'. In 2009, Gartner listed Cloud Computing as number 1 in its top 10 strategic technology areas for 2010 [10, 11]. In another report, Gartner suggested that 'by 2012, 80% of Fortune companies will pay for some cloud computing service and 30% of them will pay for cloud computing infrastructure' [4]. Enterprises are excited about the opportunities that Cloud Computing presents and, as the evidence suggests [4, 5, 10-12], Enterprise Cloud Computing is firmly poised to be the *next big thing* for businesses, large and small.

REFERENCES

[1] David W Cearley, Cloud Computing: Key Initiative Overview, Gartner Report, 2010

[2] Peter Mell and Tim Grance, The NIST Definition of Cloud Computing, version 15, National Institute of Standards and Technology (NIST), Information Technology Laboratory, www.csrc.nist.gov, 7 Oct 2009

[3] Dustin Amrhein and Scott Quint, Cloud Computing for the Enterprise: Part 1: Capturing the Cloud, DeveloperWorks, IBM, 8 Apr 2009,

[4] John Rhoton, Cloud Computing Explained: Implementation Handbook for Enterprises, Recursive Press, 3 May 2010

[5] John M Willis, Cloud Computing and the Enterprise, IT Management and Cloud, [Online] Available at: www.johnmwillis.com/ibm/cloud-computing-and-the-enterprise/, 13 Feb 2008

[6] Caroline Kvitka, Clouds Bring Agility to the Enterprise, [Online] Available at: http://www.oracle.com/technology/oramag/oracle/10-mar/o20interview.html

[7] Michael Sheehan, Cloud Computing Expo: Introducing the Cloud Pyramid, Clod Computing Journal, Aug 2008

[8] Jian Zhen, Five Key Challenges of Enterprise Cloud Computing, Cloud computing journal, 16 Nov 2008

[9] Lutz Schubert, The Future of Cloud Computing, Expert Group Report, [Online] Available at: http://cordis.europa.eu/fp7/ict/ssai/docs/executivesummary-forweb_en.pdf

[10] Dustin Amrhein & Scott Quint, Cloud Computing for the Enterprise: Part 1: Capturing the cloud, Understanding cloud computing and related technologies, DeveloperWorks, IBM, [Online] Available at: www.ibm.com/developerworks/websphere/techjournal/0904_amrhein/ 0904_amrhein.html

[11] Stephen Shankland, Brace yourself for Cloud Computing, CNET News, Oct 2009 http://news.cnet.com/8301-30685_3-10378782-264.html

[12] Ravi Mhatre, Top 5 trends for enterprise cloud computing in 2010, Lightspeed Venuter Partners, Jan 2010

[13] Sharon Sasson, Seven Best Practices for Cloud Computing, Enterprise Systems, August 2008, [Online] Available at: http://esj.com/articles/2009/08/18/cloud-best-practices.aspx

[14] David Linthicum, Cloud Computing? Thank SOA, [Online] Available at: http://www.thecloudtutorial.com/cloud-computing-soa.html

[15] David Linthicum, Cloud Computing and SOA Convergence in Your Enterprise: A Step-by-Step Guide, Addison Wesley, 2009

[16] Katz J., Sahai, A., & Waters, B. (2008) Predicate Encryption Supporting Disjunctions, Polynomial Equations, and Inner Products, Proceedings of the theory and applications of cryptographic techniques 27th annual international conference on Advances in cryptology [Online] Available at http://eprint.iacr.org/2007/404.pdf. (Accessed 6th December 2009)

[17] BBC, Google users hit by mail blackout, BBC News, 24 February 2009. [Online]. Available at http://news.bbc.co.uk/1/hi/technology/7907583.stm (Accessed: November 2009).

[18] Google. Google Apps Service Level Agreement, 2009. [Online]. Available at: http://www.google.com/apps/intl/en/terms/sla.html (Accessed: November 2009).

[19] Jaeger, P. T., Grimes, J. M., Lin, J. & Simmons, S, 'Cloud Computing and Information Policy: Computing in a Policy Cloud?' Journal of Information Technology & Politics, 5(3), 2008

[20] Knight, W, Energy-Aware Internet Routing, 2009. [Online]. Available at: www.technologyreview.com/business/23248/page2/ (Accessed: November 2009)

[21] Qureshi, A, Plugging Into Energy, 7th ACM Workshop on Hot Topics in Networks (HotNets). Calgary, Canada, October 2008

[22] Wang C, Cloud Security Front and Centre, Forrester Report, Nov 2009

[23] Amazon Web Services, Amazon Simple Storage Service FAQs, 2009. [Online]. Available at: http://aws.amazon.com/s3/faqs/#Where_is_my_data_stored (Accessed: 9 December 2009)

[24] European Network and Information Security Agency, (2009) Cloud Computing, Benefits Risks and Recommendations for Information Security, [Online] Available at: http://enisa.europa.eu/

[25] Thompson, B, Storm warning for cloud computing, BBC News, 17 May 2008 [Online]. Available at: http://news.bbc.co.uk/1/hi/7421099.stm (Accessed: November 2009)

[26] Gartner, Assessing the Security Risks of Cloud Computing, 2008,[Online] Available at: http://www.gartner.com/DisplayDocument?id=685308 Last accessed: 7th December 2009

[27] Economist, Computers without borders, 2008, [Online] Available at: Economist (23 October), at http://www.economist.com/ , Last accessed: 6th December 2009

[28] ICO, Review of EU Data Protection Directive: Summary, 2009, [Online]. Available at: http://www.ico.gov.uk/upload/documents/library/data_protection/detailed_specialist_guides/review_of_eu_dp_directive_summary.pdf Date of access: (28 November 2009)

[29] Benaloh J., Verifiable Secret-Ballot Elections. PhD thesis, Yale University, 1987.

Copyright (c) IARIA, 2011. ISBN: 978-1-61208-124-3

[30] Gentry, C, Fully homomorphic encryption using ideal lattices, Annual ACM symposium on theory of computing, Proceedings of the 41st annual ACM symposium on theory of computing, Bethseda, MD, USA, 2009, Session: crypto, pp 169-178, ACM, New York, NY, USA.

[31] Rivest, R., Adleman, L., & Dertouzos, M., On data banks and privacy homomorphisms. In Foundations of Secure Computation, pp. 169–180, 1978.

Copyright (c) IARIA, 2011. ISBN: 978-1-61208-124-3