The Privacy Implications of Cloud Computing in the Context of Software Reverse Engineering

Article · January 2013

2 authors, including:

Francisca Onaolapo Oladipo
Thomas Adewumi University
93 PUBLICATIONS **461** CITATIONS

SEE PROFILE

The Privacy Implications of Cloud Computing in the Context of Software Reverse Engineering

Onaolapo F. Oladipo and Gloria N. Anigbogu

Abstract— The World-class body, NIST defined cloud computing as as a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources; that can be rapidly provisioned and released with minimal management effort or service provider interaction. On the other hand, the need for eliminating or minimizing the capital cost associated with the deployment of large IT facilities by organizations lead to the business concept of virtual infrastructures provision and management which involves outsourcing the provision of computer resources as a service by a third party, often resulting to a flexible IT usage in a cost efficient and pay-per-use way. While one can acknowledge the original noble concerns of software reverse engineering; that is the understanding of a software system; however with the way the cloud is structured, the data flow has no respect for boundaries and time zones and pinpointing where a specific data resides is difficult, complying with privacy rules in every jurisdiction come into question. Therefore as the technology of cloud computing involves users storing their data with programs hosted on someone else's hardware, it raises a lot of privacy and security concerns in the face of software reverse engineering if the data get into the hands of a reverse engineer with malicious intents. This paper discusses the privacy concerns of outsourcing of IT capabilities in the face of software reverse engineering. The work identifies the security issues associated with reverse engineering in addition to exploring previous research work in cloud computing security issues.

Index Terms— Cloud computing, Data integrity, Privacy, Software Reverse Engineering

1 Introduction

The technology of cloud computing has been around for 'quite a while' and the services are quickly becoming the standard of choice for businesses everywhere. Various authors had come up with several definitions of cloud computing and the 'cloud' while some however believed that it is difficult to provide a precise definition of the terms.

A report by the European Commission group on Information society believed that the concept of cloud computing have proven a major commercial success over recent years and will play a large part in the ICT domain over the next 10 years or more, as future systems will exploit the capabilities of managed services and resource provisioning further. The authors of the report believe that clouds are of particular commercial interest not only with the growing tendency to outsource IT so as to reduce management overhead and to extend existing, limited IT infrastructures, but even more importantly, they reduce the entrance barrier for new service providers to offer their respective capabilities to a wide market with a minimum of entry costs and infrastructure requirements – in fact, the special capabilities of cloud infrastructures allow providers to experiment with novel service types whilst reducing the risk of wasting resources [1].

Cloud computing at its simplest, is a collection of computing software and services available from a decentralized network of services. It is a computing model that is tied to virtualized and scalable resources that are provided as a service over the internet [2]. Cloud computing can basically be explained as doing work on your local computer, but storing the data on a server at another location.

The National Institute of Standard and Technology (NIST) defines cloud computing as a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g Network servers storage, applications and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction[3]; while a similar publication by [4] defined cloud computing as a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources that can be rapidly provisioned and released with minimal management effort or service provider interaction. Reference [5] stated that the concept is entrusting data to information systems that are managed by external parties on remote servers "in the cloud."

Cloud computing can be implemented on private, public, community and hybrid cloud. The Private cloud is an in-house cloud computing option that offers hosted services to a limited number of people from behind an organization's firewalls. Public cloud is any third party services that offers storage and computing power over the internet in a scalable, pay-per-usage fashion. A community cloud consists of a number of organizations that have comparable necessities and also are willing to share infrastructure so as to share benefits of cloud computing; and Hybrid cloud is a cloud model that combines the advantages of public and private cloud computing environment [6].

Cloud computing key characteristics according to [6] is the empowerment of end-users of computing resources

- *F.O. Oladipo is a faculty in the Department of Computer Science, Nnmadi Azikiwe University, Awka Nigeria.*
- *G.N. Anigbogu is with the Department of Computer Science, Nwafor Orizu College of education, Nsugbe, Nigeria.*

own control, as opposed to the control of a centralized IT service. Agility improves with users' ability to reproves technological infrastructure resources. The concept is referred to as cloud computing because the internet is often visualized as a big cloud consisting of a large network of computers connected to each other.

As the technology of cloud computing involves users storing their data with programs hosted on someone else's hardware, it raises a lot of and privacy and security concerns in the face of software reverse engineering. The original concerns of Software Reverse Engineering was with the problem of understanding the architecture of a software application for the purpose of maintenance and re-engineering; it was conceived as a process of examination to unearth the technological principles of a software through the analysis of its structures, functions and operations in order to recreate and not necessarily copying from the original. However, attackers have leveraged on the openness of the concept to explore the vulnerabilities of a software system thereby making the technology an open-ended research area [7].

This paper presents an analysis of the privacy concerns of cloud computing in the context of reverse engineering of application systems and shows that software reverse engineering as an open-ended research area constitutes a threat to the clients' data entrusted to information systems that are managed by external parties on remote servers in the cloud.

2 Cloud Service Models

Cloud computing providers offers their services with dozens of cloud services models, like Desktop, security, data, software, platform infrastructure, IT, testing, hardware, computing, database, storage etc. [2]. As cloud computing is still evolving the providers are free to innovate and offer various services and there are no hard and fast rules governing these services offerings. Among these service models, the NIST put forward the most accepted type of service models. They are: Software as a service (SaaS), Infrastructure as a service (IaaS), and Platform as a service (PaaS). This is illustrated below (Figure 1).

Software as a Service (SaaS) is typically end user applications delivered on demand over a network on a pay per use basis. The software requires no client installation, just a browser and network connectivity. Platform as a Service (PaaS) is used by software development companies to run their software products. Software products need physical servers to run on, with database software, often Web servers too. These are all the platforms that the application runs on. Infrastructure as a Service, IaaS covers a wide range of features, from individual servers, to private networks, disk drives, various long term storage devices as well as email servers, domain name servers as well as messaging systems [9].

Figure 2 illustrates how control and management responsibilities are shared between the provider and the subscriber in the cloud service models. In the SaaS model, the service provider has very high administrative control on the application and is responsible for update, deployment, maintenance and security. The provider exercises final authority over the application; while the end-users have very limited administrative and user level control.

The PaaS model typically includes the development environment programming languages, compilers, testing tools and deployment mechanism. In some cases like Google Apps Engine (GAE), the developers may download development environment and use them locally in the developer's infrastructure, or the developer may access tools in the provider's infrastructure through a browser. Typical subscribers to PaaS independent software vendors, IT service providers or even individual developers who want to develop SaaS. IaaS provides services such as virtual computers, cloud storage, network infrastructure components such as firewalls and configuration services. The system administrators are the subscriber of this service. Usage fees are calculated per CPU hour, data GB stored per hour, network bandwidth consumed, network infrastructure used per hour, value added services used e. g. monitoring, auto-scaling etc. [4].

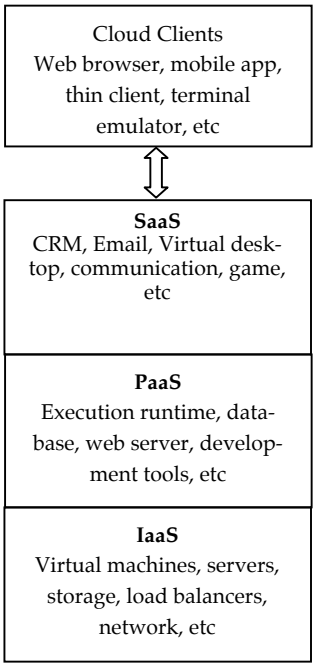

Fig. 1 Cloud Service Models [8]

3 Software Reverse Engineering

Software Reverse Engineering (RE) was defined as

the process of analyzing a subject system to create representations of the system at a higher level of abstraction in order to unravel the complexities of target software or to undermine the assumptions made by the people who created the system and then undermine those assumptions. It may involve going backwards through the development cycle. It is a process of examination to unearth the technological principles of a software system through the analysis of its structures, functions and operations in order to recreate and not necessarily copy from the original. The methods and technologies play an important role in many software engineering tasks, such as program comprehension, system migrations, and software evolution [10].

A report by [11] stated that reverse engineering allows one to learn about a program's structure and its logic thereby leading to some critical insights regarding how a program functions. This kind of insight is extremely useful when the aim is to exploit software. The researchers believed that there are obvious advantages to be had from reverse engineering. For example, one can learn the kind of system functions a target program is using and learn the files the target program accesses. One can also learn the protocols the target software uses and how it communicates with other parts of the target network.

The most powerful advantage to reversing is that it can be enable one to change a program's structure and thus directly affect its logical flow. Technically this activity is called *patching,* because it involves placing new code patches (in a seamless manner) over the original code, much like a patch stitched on a blanket. Patching allows the engineer to add commands or change the way particular function calls work. This enables the addition of secret features, removal or disabling functions, and fixing of security bugs without source code. A common use of patching in the computer underground involves removing copy protection mechanisms. Like any skill, reverse engineering can be used for good and for bad ends [11].

Reverse engineering of software can be accomplished by various methods. The three main groups of software reverse engineering are

1) Analysis through observation of information exchange
2) Disassembly using a disassembler to read and understand the raw machine language of the program in its own terms.

Decompilation using a decompiler, a process that tries, with varying results, to recreate the source code in some high-level language for a program only available in machine code or bytecode [10].

4 Related Research Works

A paper by [12] described the privacy challenges that end user as well as the cloud provider face during the access to the services provided by cloud. The paper discussed the issue of cloud computing and outlined its implications for the privacy of personal information as well as its implications for the confidentiality. A threat model which contained different mitigation techniques to deal with privacy problems in cloud computing was also presented by the authors. Reference [13] presented a study about the risk issues involved in cloud computing. The research highlighted the different types of risks and how their existence can affect the cloud users; discussed the different circumstances in which the risks occur and the measures to be taken to avoid them in addition to laying out measures to be taken while using cloud computing to reduce negative effects on the outcome and maintain data integrity. The authors concluded that with the inherent risks involved in IT infrastructures outsourcing, a proper risk analysis approach will be of great help to both the service providers and the customers as with such an approach, the customers can be guaranteed data security and the service providers can win the trust of their customers.

A model for the integrity checking over the cloud computing with the support of the Third Party Auditor (TPA) using digital signature technique was proposed by [14]. The authors believed that in order to overcome the threat of integrity of the data, the user must be able to use the assistance of a TPA as it has an experience that clouds users does not have, and checks over the integrity that is difficult for the users to check. The proposed model result was shown efficiently with a number of situations that performed by unauthorized attackers. The checking done over two parts the CSP and TPA, without giving any secure data that void the integrity definition and without uploading any secure data to the cloud.

Reference [15] proposed a framework to help safeguard data in the cloud. The paper discussed the security issues of cloud computing and gave some measures to limit the vulnerabilities in the cloud. The frame work offered a secured environment in which the clients need to access the providers' network using secured VPNs. The paper further described the steps taken for user authentication and secure service provision across the cloud deployment platforms of IaaS, PaaS and SaaS.

In their paper [16] believed that deploying cloud computing in an enterprise infrastructure bring significant security concerns; therefore successful implementation of cloud computing in an enterprise requires proper planning and understanding of emerging risks, threats, vulnerabilities, and possible countermeasures. The authors believe that enterprise should analyze the company/organization security risks, threats, and available countermeasures before adopting this technology. In their paper, they discussed security risks and concerns in cloud computing and enlightened steps that an enterprise can take to reduce security risks and protect their re-

sources. They also explained cloud computing strengths/benefits, weaknesses, and applicable areas in information risk management.

An executive summary by the Centre for Technology Innovation opined that there is reason to be optimistic about the gains to be had from a transition of many information services to a cloud architecture as cloud computing makes possible cost savings, scalability, and more efficient use of IT resources, among other things. The authors posited that however, the risks to privacy and security from cloud computing cannot be ignored. They pointed out that not all these risks are new, and that some of them can be mitigated through technology investment and due diligence from the client while others are systematic in nature, and may not be solvable through unilateral innovation. They reported that uncertainty dominates the client's ability to forecast risk and the data subject's expectation of privacy and that transparency would support selection towards a more security-conscious cloud universe, and market competition can enable that shift [17].

An examination of the main security and privacy issues pertinent to cloud computing, as they relate to outsourcing portions of the organizational computing environment was carried out by [18]. The paper pointed out areas of concern with public clouds that require special attention and provides the necessary background to make informed security decisions. In addition to this, the researcher also pointed out that in emphasizing the cost and performance benefits of the cloud, some fundamental security problems have receded into the background and been left unresolved. Several critical pieces of technology, such as a solution for federated trust, are not yet fully realized, impinging on successful deployments. Determining the security of complex computer systems is also a long-standing security problem that overshadows large scale computing in general. Attaining the high assurance qualities in implementations has been an elusive goal of computer security researchers and practitioners, and is also a work in progress for cloud computing.

A report of the World Privacy Forum discussed the issue of cloud computing and outlined its implications for the privacy of personal information as well as its implications for the confidentiality of business and governmental information. The report found that for some information and for some business users, sharing may be illegal, may be limited in some ways, or may affect the status or protections of the information shared. The report discussed how even when no laws or obligations block the ability of a user to disclose information to a cloud provider, disclosure may still not be free of consequences. The report further found that information stored by a business or an individual with a third party may have fewer or weaker privacy or other protections than information in the possession of the creator of the information. The report, in its analysis and discussion of relevant laws, also found that both government agencies and private litigants may be able to obtain information from a third party more easily than from the creator of the information.

A cloud provider's terms of service, privacy policy, and location may significantly affect a user's privacy and confidentiality interests [19]. A technical report by the University of California at Berkeley strived to frame the full space of cloud-computing security issues, attempting to separate justified concerns from possible overreactions. The work examined contemporary and historical perspectives from industry, academia, government, and "black hats" and argued that few cloud computing security issues were fundamentally new or fundamentally intractable. The report also argued that two facets are to some degree new and fundamental to cloud computing: the complexities of multi-party trust considerations, and the ensuing need for mutual auditability [20].

An entity-centric approach for Identity Management (IDM) in the cloud was proposed by [21]. The approach was based on active bundles, —each including a payload of PII, privacy policies and a virtual machine that enforces the policies and uses a set of protection mechanisms to protect themselves; and anonymous identification to mediate interactions between the entity and cloud services using entity's privacy policies. The main characteristics of the approach were: it is independence, ability to give minimum information to the service provider, and its provision of ability to use identity data on untrusted hosts.

5 Cloud Privacy Concerns and the Two Faces of Reverse Engineering

Throughout the years, organizations have experienced and will continue to experience in this cloud computing era numerous system losses which will have a direct impact on their most valuable asset, information [22]. While [23] opined that most of the aims of reverse engineering are closely related but not limited to software maintenance, [11] stated that "Because reverse engineering can be used to reconstruct source code, it walks a fine line in intellectual property law". Reverse engineering for malicious purpose – e.g. theft of intellectual property (such as a competitor's secret formula or process), software tampering, or the discovery and exploitation of vulnerabilities – is facilitated by a number of advanced program analysis tools which also serve the legitimate software development community, e.g. in debugging, software engineering, and understanding malware [24].

It is the position of this paper that software reverse engineering represents a double-ended research area. This agrees with the position of [25] who believed in the open-ended attribute of software reverse engineering. He

described software reverse engineering as the technique of getting the original source code from the binary. He also stated that competitors might use reverse engineering to figure out how certain important features of an application, crackers might use it to see how they can bypass the license policy and game cheats use reverse engineering, as well to cheat. The main threat comes not from actions (or inaction) by the cloud provider (whose commercial reputation relies on keeping customer's data secure); but from other parties (reverse engineers inclusive). The simple act of putting your data in the cloud often makes its confidentiality less reliable than if you had kept it had been kept under the user's own control.

REFERENCES

[1] Electronic Privacy Information Center EPIC- EU Data Protection Directive (2012). Downloaded December 2012 from epic.org/priv acy /intl/eu_data_protection_directiv e.html

[2] Voorsluys, W. Broberg, J. & Buyya, R. (2011). Introduction to Cloud computing. Chapter 1, pages 1–41. Cloud Computing: Principles and Paradigms. John Wiley & Sons, Inc., 2011.

[3] Mell, P. & Grance, T. (2011). NIST Special Publication, National Institute of Standard and Technology Special Publication 800-145

[4] Badger, L., Grance, T., Patt-Corner, R., Voas, J. (2011). Draft Cloud Computing Synopsis and Recommendation. (NIST) National Institute for Standard and Technology. Special Publication 800-146.

[5] Ryan, M. D. (2011). Cloud Computing Privacy Concerns on Our Doorstep. Communications of the ACM. vol. 54 no. 1

[6] Danielson, K. (2008). Distinguishing cloud computing from utility computing. Retrieved January 2012 from http://www.ebizq.net/blogs/saasweek/2008/03/ distinguishing_cloud_computing/

[7] O. F. Oladipo, M. O. Odoh, M. O. Onyesolu, M. O., "Exploring the two faces of software reverse engineering," International Journal of Advanced Research in Computer Science and Software Engineering, Volume 2, Issue 4, April 2012. Pp 366-370

[8] Http://en.wikipedia.org/wiki/cloudcomputing modified on 16 Sept. 2012

[9] C. Czarnecki, "Cloud Service Models: Comparing SaaS PaaS and IaaS", Published November 9, 2011 in Perspectives on Cloud Computing from Learning Tree International. Downloaded December 8 2012 from http://cloud-computing.learningtree.com/2011/11/09/cloud-service-models-comparing-saas-paas-and-iaas/

[10] Osuagwu, O. E., Oladipo, O. F. and Yinka-Banjo, C. (2008). Deploying Reverse Software Engineering as tool for Converting Legacy Applications in critical-sensitive systems for Nigerian Industries. In Proceedings of the 22nd National conference and AGM of the Nigeria Computer Society Conference (ENCTDEV 2008). 24-27 June

[11] G. Hoglund, G. McGraw. (2004). Exploiting Software How to Break Code. Addison Wesley. Chapter three, pp 71-145. ISBN: 0-201-78695-8

[12] P. Metri, G. Sarote, "Privacy Issues and Challenges in Cloud computing", International Journal of Advanced Engineering Sciences and Technologies, Vol No. 5, Issue No. 1, April 2011, Pp. 001 – 006. ISSN: 2230-7818 @ 2011

[13] Chandran S. and Angepat M., "Cloud Computing: Analyzing the risks involved in cloud computing environments," in Proceedings of Natural Sciences and Engineering, Sweden, pp. 2-4, 2010.

[14] D. Attas and O. Batrafi (2011). Efficient integrity checking technique for securing client data in cloud computing. International Journal of Electrical & Computer Sciences IJECS-IJENS Vol: 11 No: 05

[15] A. Mathew (2012). Security And Privacy Issues of Cloud Computing; Solutions And Secure Framework. International Journal of Multidisciplinary Research Vol.2 Issue 4, April 2012, ISSN 2231 5780

[16] A. Bisong and S. M. Rahman (2011). An Overview of The Security Concerns In Enterprise Cloud Computing. International Journal of Network Security & Its Applications (IJNSA), Vol.3, No.1, January 2011. DOI: 10.5121/ijnsa.2011.3103 30

[17] Allan A. Friedman and Darrell M. West (2010). Privacy and Security in Cloud Computing. Issues in Technology Innovations. Executive summary of The Centre for Technology Innivation. No 3, Oct 2010.

[18] W. A. Jansen (2011). Cloud Hooks: Security and Privacy Issues in Cloud Computing. In Proceedings of the 44th Hawaii International Conference on System Sciences.

[19] R. Gellman (2009). Privacy in the Clouds: Risks to Privacy and Confidentiality from Cloud Computing. A report prepared for the World Privacy Forum, February 23, 2009. Downloded December 2012 from www.worldprivacyforum.org/pdf/WPF_Cloud_Privacy_Report.pdf

[20] Y. Chen, V. Paxson, R. H. Katz. (2010). What's New About Cloud Computing Security? Technical Report No. UCB/EECS-2010-5. January 20, 2010. Electrical Engineering and Computer Sciences, University of California at Berkeley. Downloaded December 2012 from http://www.eecs.berkeley.edu/Pubs/TechRpts/2010/EECS-2010-5.html

[21] Angin, P.; Bhargava, B. ; Ranchal, R. ; Singh, N. ; Linderman, M. ; Ben Othmane, L. ; Lilien, L. (2010). An Entity-centric Approach for Privacy and Identity Management in Cloud Computing. in 29th IEEE Symposium on Reliable Distributed Systems, Oct. 31 2010-Nov. 3 2010 Page(s): 177 - 183

[22] Otero, A. R., Otero, C. E., Qureshi, A.(2010). Securing data transfer in the cloud through introducing identification packet and UDT – authentication option field: a characterization. International Journal of Network Security & Its Applications (IJNSA), Vol.2, No.4, October 2010. Retrieved from http://airccse.org/journal/nsa/1010ijnsa01.pdf

[23] Klosch, R.R. (1996). Reverse Engineering: Why and how to reverse engineer software, Proceedings of the California Software Symposium (CSS '96), Los Angeles, California.

[24] 24. Van Oorschot, P. C. (2003). Revisiting Software Protec-

tion. Proceedings of the 6th International Conference on Information Security, ISC 2003, Bristol, UK, pp.1–13, Springer-Verlag LNCS 2851 (2003), Colin Boyd, Wenbo Mao (Eds.).

[25] 25. Chandran, R. (2008). Defend against Reverse Engineering. Palizine Information Security intelligence Magazine, Issue 34, July. Downloaded February 2012 from http://palpapers.plynt.com

Onaolapo F. Oladipo Oladipo obtained a PhD in Computer Science from Nnamdi Azikiwe University, Nigeria where she is currently a faculty member. Her research interests spanned various areas of Computer Science and Applied Computing and she is currently promoting research into deploying contemporary ICT solutions with minimum infrastructures for developing economies. She has published numerous papers detailing her research experiences in both local and international journals and presented research papers in several international conferences. She is also a reviewer for many international journals and conferences. She is a member of several professional and scientific associations both within Nigeria and beyond; they include the European Alliance for Innovation (EAI), British Computer Society (BCS), Nigerian Computer Society (NCS), Computer Professionals (Regulatory Council) of Nigeria, the Global Internet Governance Academic Network (GigaNet), International Association Of Computer Science and Information Technology (IACSIT), the Internet Society (ISOC), Diplo Internet Governance Community, The Institute for Computer Sciences, Social Informatics and Telecommunications Engineering (ICST) and the Africa ICT Network.

Gloria N. Anigbogu is an MSc candidate in the department of Computer Science, Nnamdi Azikiwe University, Nigeria. She is a faculty member at the Nwafor Orizu College of education, Nsugbe, Nigeria and a chattered practitioner of the computing profession. Her research results have been published in both local and international journals.

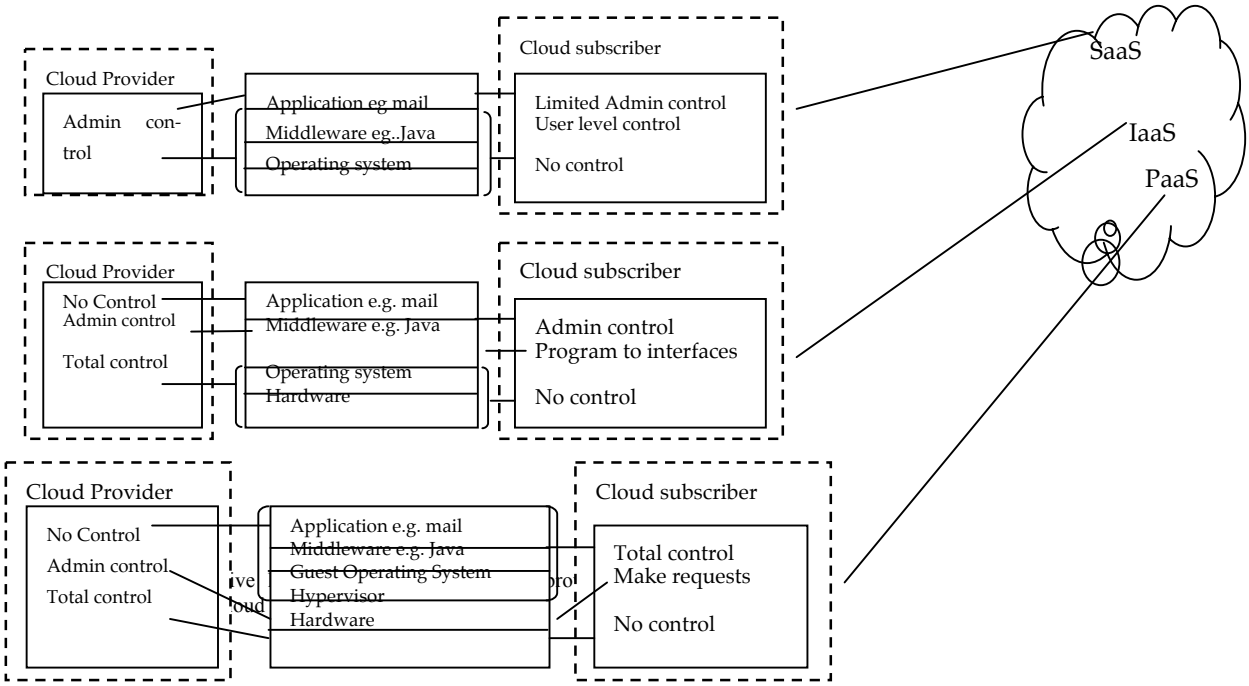

Figure 2 The relative level of control between the provider and the subscriber in Cloud service models [4]