

On An Ontological View Of Cloud Computing

Harry Katzan, Jr., Savannah State University, USA

ABSTRACT

Cloud computing is an architecture for providing computing service via the Internet. Use of the term “cloud” is a metaphor for the representation of the Internet used in most systems diagrams. In this case, the Internet is the transport mechanism between a client and a server located somewhere in cyberspace, as compared to having computer applications residing on an “on premises” computer. Adoption of cloud computing practically eliminates two ongoing problems in IT service provisioning: the upfront costs of acquiring computational resources and the time delay of building and deploying software applications. This paper gives an ontological view of the subject in order to serve as a point of reference in the discipline and to facilitate ongoing technical development.

Keywords: Software-as-a-service, cloud computing, software architecture, platform-as-a-service, infrastructure-as-a-service, private cloud, public cloud, community cloud, hybrid cloud.

CLOUD COMPUTING CONCEPTS

*C*loud computing is a means of providing computer facilities via the Internet, but that is only half of the picture. The other half is that it is also a means of accessing those same computer facilities via the Internet from different locations. When a large bank, for example, moves to cloud computing for online operations, it necessarily considers both halves of the equation. The adjective “cloud” reflects the diagrammatic use of a cloud as a metaphor for the Internet. In telecommunications, a *cloud* is the unpredictable part of a network through which business and personal information passes from end-to-end and over which we do not have direct knowledge or control.¹ Most of us have been using cloud-computing facilities in one form or another for years through ordinary email and the World Wide Web. Recently, the term has come to reflect the use of software and the running of computer applications via the Internet where the computer infrastructure and software are not “on premises.”

Cloud computing is an evolving concept, even though major financial institutions, consulting organizations, and software vendors have invested heavily in the technology and associated business practices, and the U.S. Government has endorsed the model for federal computer operations. Surprisingly, there are few papers in the academic/professional literature on the subject, leading to divergent views on what the cloud-computing paradigm actually represents. The National Institute of Standards and Technology has stepped into the fray, and this paper essentially covers a *public domain*² definition of the topic. (Mel 2009a, Mel 2009b) In this paper, we will be concerned with concepts and terminology.

CLOUD COMPUTING UTILITY

In principle, all utilities are the same and characterized by four key factors: necessity, reliability, usability, and scalability. *Necessity* refers to the notion of a preponderance of users depending on the utility to satisfy everyday needs. *Reliability* refers to the expectation that the utility will be available when the user requires it.

¹ See *Privacy in the Clouds* by Cavoukin for more information on this subject.

² Public domain is an intellectual property designation for content not owned by anyone and usually paid for with taxpayer money. Proper attribution is expected but not required with public domain content.

Usability refers to the requirement that the utility is easy and convenient to use – regardless of the complexity of the underlying infrastructure. *Scalability* refers to the fact that the utility has sufficient capacity to allow the users to experience the benefits of an expandable utility that provides economy of scale. (Rappa 2004) Certainly, modern Internet facilities for search operations that engage thousands of servers satisfy these characteristics, and one would expect cloud services to do the same. (Katzan 2009) Cloud computing is different from a conventional utility because it deals with information and computers that may contain sensitive information leading to considerations of security and privacy.

CLOUD ARCHITECTURE

A cloud architecture is a collection of three categories of information resources for achieving business agility, availability, collaboration, and elasticity in the deployment and use of cloud services that include software, information, and a cloud infrastructure. The *software category* includes system software, application software, infrastructure software, and accessibility software. The *information category* refers to large collections of data and the requisite database and management facilities needed for efficient and secure storage utilization. The *category of cloud infrastructure* includes compute³ resources, network facilities, and the fabric for scalable consumer operations. We are going to adopt an ontological formulation for the description of a cloud framework that necessarily includes three classes of information: terminology, architectural requirements, and a reference model. The description adopts the National Institute of Standards and Technology (NIST) cloud-computing paradigm. (Mell 2009b, Brunette 2009)

Ontology

Ontology is a specification of “what is.” In philosophy, use of the term reflects the study of being (or existence) and describes and delineates a collection of basic categories, and also defines the entities and classes of elements within a category. In service science, ontology is a specification of a conceptualization used to enable knowledge sharing. Since ontology concerns existence, an ontological definition of a subject – perhaps a service category – reflects a materialization of a concept obtained through a shared reality. We are going to adopt Gruber’s definition of ontology, as “a set of representational primitives with which to model a domain of knowledge or discourse.” (Gruber 2008) Most forms of ontology are expressed in an ontology language and share structural similarities, such as individuals, classes, attributes, relations, function, restrictions, rules, axioms, and events. (Sowa 2000) The components determine whether a specific ontology is a domain ontology or an upper ontology. In a *domain ontology*, a specific type would be relevant to a particular category, such as in a consumer or business cloud category. In an *upper ontology*, a type would be applicable to all ontologies in the universe of discourse. In the cloud ontology, presented in the following sections, we are going to be developing an upper ontology for cloud services.

Cloud Characteristics

In a cloud environment, the essential cloud characteristics fall in the former category. The NIST lists the characteristics as: (Mell 2009a)

On-demand self-service. A consumer can unilaterally provision computing capabilities, such as server time and network storage, automatically without requiring human interaction with a service’s provider.

Broad network access. Capabilities are available over the network that are accessed through standard mechanisms that promote use by heterogeneous thin or thick client platforms (e.g., mobile phones, laptops, and PDAs).

Resource pooling. The provider’s computing resources are pooled to serve multiple consumers using a multi-tenant model, with different physical and virtual resources dynamically assigned and reassigned according to consumer demand. There is a sense of location independence in that the customer generally has no control or knowledge over

³ The use of the word compute is correct in this instance and is refers in cloud computing to a facility for computer processing.

the exact location of the provided resources but may be able to specify location at a higher level of abstraction (e.g., country, state, or datacenter). Examples of resources include storage, processing, memory, network bandwidth, and virtual machines.

Rapid elasticity. Capabilities can be rapidly and elastically provisioned, in some cases automatically, to quickly scale out and to rapidly be released to quickly scale in. To the consumer, the capabilities available for service provisioning appear to be unlimited and can be obtained in any quantity at any time.

Measured Service. Cloud systems automatically control and optimize resource use by leveraging a metering capability at some level of abstraction appropriate to the type of service (e.g., storage, processing, bandwidth, and active user accounts). Resource usage can be monitored, controlled, and reported providing transparency for both the provider and consumer of the utility service.

The essential characteristics are more than a description of an “application layer.” (Youseff 2009) The characteristics give a salient set of reasons why a consumer would want to move to the cloud.

CLOUD SERVICE PROVISIONING

The salient features of the cloud computing business model are summarized as follows:

- The ownership of the software is transferred to the cloud service provider.
- The responsibility for hardware, application software, storage facilities, and professional services resides with the provider.
- Systems software is available from a trusted vendor for supporting cloud services.
- Data centers are available for sustaining the operational structure and for supporting the requisite fabric needed to utilize server farms.

Accordingly, the business model provides the economy of scale needed to target the long tail by providers and reduces up-front and operational costs for the client.

Cloud Application Architecture

A cloud application architecture includes a continuum of service levels, based on the capability of handling multiple clients and software configurations. Two levels are identified here. The number of levels in any specific operational environment is based on the cloud platform and its characteristics.

Level One. At the first level, the server hosts a separate instance of the software for each customer, but the instance is a configurable version of the same code base, reducing maintenance costs and contributing to reduced software costs. This instance is known as *single tenant* software.

Level Two. At the second level, all of the customers within an application client domain address a single instance of an application running on a server. Each client/instance is totally independent of other client/instances running on the same server. Authentication and authorization security policies insure the separation of user data. This instance is known as *multiple tenant* software.

The choice among architectural levels is determined by the provider/client’s business, architectural, and operational requirements.

Cloud Platforms

A *cloud platform* is an operating system that runs in the cloud and supports the software-as-a service concept. A cloud platform resides in a cloud data center and exists as a powerful computing facility, a storage system, an advanced operating system, support software, and the necessary fabric to sustain a server farm and scale up to support millions of Internet clients. A cloud platform is as much about operating in the cloud, as it is about

developing applications for the cloud. A cloud platform provides the facility for an application developer to create applications that run in the cloud; and, in so doing, the application developer uses services that are available from the cloud. Cloud platforms are a lot like enterprise-level platforms, except that they are designed to scale up to Internet-level operations supporting millions of clients.

A cloud platform can be conceptualized as being comprised of three complementary groups of services: foundations, infrastructure services, and application services. The *foundation* refers to the operating system, storage system, file system, and database system. *Infrastructure services* include authorization/authentication/security facilities, integration between infrastructure and application services, and online storage facilities. *Application services* refer to ordinary business services that expose “functional” services as SOA components. Cloud platforms are a lot like enterprise-level platforms, except that they are designed to scale up to support Internet-level operations.

Service Models

The cloud service models give an ontological view of what a cloud service is. A cloud service system is a set of elements that facilitate the development of cloud applications. (Youseff op cit.) Here is a description of the three layers in the NIST service model description: (Mel op cit.)

Cloud Software as a Service (SaaS). The capability provided to the consumer is the use of the provider’s applications running on a cloud infrastructure. The applications are accessible from various client devices through a thin client interface such as a web browser (e.g., web-based email). The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, storage, or even individual application capabilities, with the possible exception of limited user-specific application configuration settings.

Cloud Platform as a Service (PaaS). The capability provided to the consumer is that of deploying onto the cloud infrastructure consumer-created or acquired applications developed through the use of programming languages and tools supported by the provider. The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, or storage, but has control over the deployed applications and possibly application hosting environment configurations.

Cloud Infrastructure as a Service (IaaS). The capability provided to the consumer is the capability of provisioning 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. The consumer does not manage or control the underlying cloud infrastructure but has control over the operating system, storage, and deployed applications, as well as limited control over selected networking components (e.g., host firewalls).

The three service model elements should be deployed in a cloud environment with the essential characteristics in order to achieve a desired cloud state. Figure 1 gives an idea of how the service models are interrelated.

Service Deployment Models

The essential elements of a cloud service system are given above. In order to develop enterprise-wide applications, a domain ontological viewpoint has to be assumed with deployment models from the following list: (Mel op cit.)

Private cloud. The cloud infrastructure is operated solely for an organization. It can be managed by the organization or a third party and may exist on-premises or off-premises.

Community cloud. The cloud infrastructure is shared by several organizations and supports a specific operational community that has shared concerns (e.g., mission, security requirements, policy, and compliance considerations). It may be managed by an organizations or a third party and may exist on-premises or off-premises.

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

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).

Most cloud software service application domains will be synthesized from a combination of the deployment models.

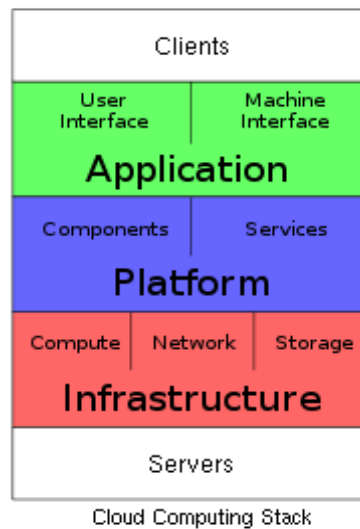


Figure 1. Service Model Architectures. (Wikipedia 2010)

PRIVACY AND SECURITY

Many examples of cloud computing conveniently characterize it as analogous to the electric utility. Historically, the private generators of the twentieth century were replaced by the electricity grids of today. Although similar in some respects, cloud computing is different in one important way. The cloud will typically handle information about which privacy and security are of paramount concern. The key issues therein are identity, accountability, and privacy. The side issues are identity attributes (such as age, gender, and race), accountability (for the security of online activity), and anonymity (in order to protect free speech). The key consideration may turn out to be a matter of control. From an organizational perspective, control over information should remain with the end user. From a personal perspective, the person should have the wherewithal to control his or her identity as well as the release of identity attributes. Who owns the data? Is it the person about whom the data pertains? Is it the organization that prototypically stores its data? Or, is it the cloud provider that physically controls and stores the data somewhere out in cyberspace? As an example, is your financial information (as personal data) your property or is it your bank's business property? They are not rhetorical questions, but issues that must be resolved before moving to a cloud operational environment.

QUICK SUMMARY

1. Cloud computing is a means of accessing computer facilities via the Internet. (The *cloud* is a metaphor for the Internet.)
2. Cloud service facilities are characterized by four key factors: necessity, reliability, usability, and scalability.
3. Software-as-a-service (SaaS) is software deployed as a hosted service and accessed over the Internet.
4. For the cloud client, business service is a balance between control and economy of scale.
5. A cloud platform is based on an operating system that runs in the cloud and provides an infrastructure for software development and deployment.
6. Most cloud software service application domains will be synthesized from a combination of the deployment models.

ACKNOWLEDGMENT

Thanks to Margaret Katzan for reading the manuscript and the reviewers for insightful comments.

REFERENCES AND SELECTED READING

1. Anderson, C. 2006. *The Long Tail*. New York: Hyperion.
2. Brunette, G. and R. Mogull (ed). 2009. *Security Guidance for Critical Areas of Focus in Cloud Computing V2.1*. Cloud Security Alliance, December 2009.
3. Cavoukian, A. 2009. *Privacy in the Clouds*. Toronto: Information and Privacy Commission of Ontario www.ipc.on.ca.
4. Chappell, D. 2008. A Short Introduction to Cloud Platforms. Microsoft Corporation.
5. Ferrario, R. and N. Guardino. 2008. Towards an Ontological Foundation for Services Science. Proceedings of the Future Internet Symposium, Vienna Austria, 28-30 September 2008.
6. Gruber, T. 2008. Ontology. *Encyclopedia of Database Systems*, Liu, L. and M. Ozsu (Eds.), Springer-Verlag,
7. Katzan, H. 2009. Cloud Computing Economics: Democratization and Monetization of Services. *Journal of Business & Economics Research*, 7(6):1-11.
8. Mell, P. and T. Grance. 2009a. The NIST Definition of Cloud Computing. National Institute of Standards and Technology, Information Technology Laboratory, Version 15, 10-7-09. <http://www.csrc.nist.gov/groups/SNS/cloud-computing/index.html>
9. Mell, P., Badger, L., and T. Grance. 2009b. Effectively and Securely Using the Cloud Computing Paradigm. National Institute of Standards and Technology, Information Technology Laboratory, 10-7-09. <http://www.csrc.nist.gov/groups/SNS/cloud-computing/index.html>
10. Rappa, M. 2004. The utility business model and the future of computing services. *IBM Systems Journal*, 43(1):32-41.
11. Sowa, J. 2000. *Knowledge Representation: Logical, Philosophical and Computational Foundations*, Brooks Cole Publishing.
12. Wikipedia. 2010. Cloud Computing. http://en.wikipedia.org/wiki/Cloud_computing. Downloaded 01/29/2010.
13. Youseff, L., Butrico, M., and D. Da Silva. 2009. Toward a Unified Ontology of Cloud Computing. (Available from the following: lyouseff@cs.uscb.edu, butrico@us.ibm.com, and dilmasilva@us.ibm.com).