

A New Approach of CLOUD: Computing Infrastructure on Demand

Kamal Srivastava*
Atul Kumar**

Abstract

Purpose: The paper presents a latest vision of cloud computing and identifies various commercially available cloud services promising to deliver the infrastructure as a service (IaaS).

Design/Methodology/Approach: Cloud computing provides the architectural detail and different types of clouds such as we studied different cloud based architectures like Blue Cloud built on IBM's massive scale computing initiatives, Google Cloud which claims business can get started using Google Apps online pretty much instantly, Salesforce.com consists of development as service, a set of development tools and APIs that enables enterprise developers to easily harness the promise of the cloud computing.

Findings: It was found that Cloud computing is changing the way we provision hardware and software for on-demand capacity fulfillment and changing the way we develop web applications and make business decisions.

Keywords: Cloud Computing, Amazon Elastic Compute Cloud, Google App Engine, Microsoft Azure, Salesforce.com.

Paper Type: Survey

Introduction

The term “cloud”, as used in this white paper, appears to have its origins in network diagrams that represented the internet, or various parts of it, as schematic clouds. “Cloud computing” was coined for what happens when applications and services are moved into the internet “cloud.” Cloud computing is not something that suddenly appeared overnight; in some form it may trace back to a time when computer systems remotely time-shared computing resources and applications. More currently though, cloud computing refers to the many different types of services and applications being delivered in the internet cloud, and the fact that, in many cases, the devices used to access these services and applications do not require any special applications. 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. A cloud computing platform dynamically provisions, configures, reconfigures, and deprovisions servers as needed. Cloud applications are those that are extended to be accessible through the Internet. The datacenter hardware and software is what we will call a Cloud. Cloud computing is changing the way we provide hardware and

* Department of Computer Science, Shri Ramswaroop Memorial College of Engg. & Mgmt. Lucknow, U.P., India. email: srivastava@gmail.com

** Department of Computer Science, Shri Ramswaroop Memorial College of Engg. & Mgmt. Lucknow, U.P., India. email: atulverma16@gmail.com

software for on-demand capacity fulfillment and changing the way we develop web applications and make business decisions. Cloud computing is a computing paradigm in which tasks are assigned to a combination of connections, software and services accessed over a network. This network of servers and connections is collectively known as “the cloud.” Computing at the scale of the cloud allows users to access supercomputer-level power. Users can access resources as they need them.

Understanding Cloud Computing

Cloud computing describes how computer programs are hosted and operated over the Internet. The key feature of cloud computing is that both the software and the information held in it live on centrally located servers rather than on an end-user's computer. How does cloud computing work? The concept is fairly simple. First, consider the traditional means of running application, and application appears to run on a dumb terminal or, these days your PC; practically, this is only front-end of the application. Your computer is connected to a server that actually runs the program or application and returns output to personal computer. The server constitutes the backend and it can be located in the same building as you are or not. With cloud computing application program runs somewhere within the cloud; ideally the user concern only with applications that are available and need not to be aware of the underlying technology or the physical location of the Application's computer. User desktop is connected via internet to a server farm, a collection of remote servers that runs many, many applications at once. Which server or servers an application runs on is determined by the application program already running on the machines; there is an attempt to balance the load so that all of the programs run optimally. There are number of companies that offer cloud computing services like Amazon offers something called Amazon Elastic Compute Cloud (EC2), Google with its own cloud computing offering, Google App Engine and Microsoft offers Microsoft Azure. When a Cloud is made available in a pay-as-you-go manner to the public, we call it a Public Cloud; the service being sold is Utility Computing. We use the term Private Cloud to refer to internal datacenters of a business or other organization that are not made available to the public. Thus, Cloud Computing is the sum of SaaS (Software as a Service) and Utility Computing, but does not normally include Private Clouds. From a hardware point of view, three aspects are new in Cloud Computing (**Vogels, 2008**):

1. The illusion of infinite computing resources available on demand, thereby eliminating the need for Cloud Computing users to plan far ahead for provisioning.

2. 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.
3. The ability to pay for use of computing resources on a short-term basis as needed (e.g., processors by the hour and storage by the day) and release them as needed, thereby rewarding conservation by letting machines and storage go when they are no longer useful. As a successful example, **(Armbrust, et al, 2009)** Elastic Compute Cloud (EC2) from Amazon Web Services (A WS) sells 1.0-GHz x86 ISA “slices” for 10 cents per hour, and a new “slice”, or instance, can be added in 2 to 5 minutes. Amazon's Scalable Storage Service (S3) charges USD 0.12 to USD 0.15 per gigabyte month, with additional bandwidth charges of USD 0.10 to USD 0.15 per gigabyte to move data in to and out of A WS over the Internet.

Commercially Available Cloud Services

- 1) Google: The core of Google's business is all in Cloud Computing. Services delivered over network connections include search, e-mail, online mapping, office productivity (including documents, spreadsheets, presentations, and databases), collaboration, social networking and voice, video, data services. Users can subscribe to these services for free or pay for increased levels of service and support.
- 2) Amazon: As the world's largest online retailer, the core of Amazon's business is ecommerce. While ecommerce itself can be considered Cloud Computing, Amazon has also been providing capabilities which give IT department's direct access to Amazon compute power. Key examples include S3 (Simple Storage Services) and EC2. Any internet user can access storage in S3 and access stored objects from anywhere on the Internet. EC2 is the Elastic Compute Cloud, a virtual computing infrastructure able to run diverse applications ranging from web hosts to simulations or anywhere in between. This is all available for a very low cost per user
- 3) Microsoft: Traditionally Microsoft's core business has been in device operating systems and device office automation software. Since the early days of the Internet Microsoft has also provided web hosting, online e-mail and many other cloud services. Microsoft now also provides office automation capabilities via a cloud (“Office Live”) in an approach referred to as “Software Plus Services” vice “Software as a Service” to allow synchronous/asynchronous integration of online Cloud documents with their traditional offline desktop-resident versions.
- 4) Salesforce.com: The core mission of Salesforce.com has been in

delivery of capabilities centered on customer relationship management. However, in pursuit of this core Salesforce.com has established themselves as thought leaders in the area of Software as a Service and is delivering an extensive suite of capabilities via the Internet. A key capability provided is the site Force.com, which enables external developers to create add-on applications that integrate into the main Salesforce.com application and are hosted on the infrastructure Salesforce.com.

- 5) VMware: Provides several technologies of critical importance to enabling cloud computing, and has also started offering its own cloud computing on demand capability called vCloud. This type of capability allows enterprises to leverage virtualized clouds inside their own IT infrastructure or hosted with external service providers.

New Application Opportunities

- 1) **Mobile interactive applications:** Tim O'Reilly believes that "the future belongs to services that respond in real time to information provided either by their users or by nonhuman sensors" (Li., et al, 2009). Such services will be attracted to the cloud not only because they must be highly available, but also because these services generally rely on large data sets that are most conveniently hosted in large datacenters. While not all mobile devices enjoy connectivity to the cloud 100% of the time, the challenge of disconnected operation has been addressed successfully in specific application domains, so we do not see this as a significant obstacle to the appeal of mobile applications.
- 2) **Parallel batch processing:** Cloud computing presents a unique opportunity for batch-processing and analytics jobs that analyze terabytes of data and can take hours to finish. If there is enough data parallelism in the application, users can take advantage of the cloud's new "cost associativity": using hundreds of computers for a short time costs the same as using a few computers for a long time. For example, Programming abstractions such as Google's MapReduce (Dean & Ghemawat, 2004) and its open-source counterpart Hadoop (Bialecki, Cafarella, Cutting & O'Malley, 2005) allow programmers to express such tasks while hiding the operational complexity of choreographing parallel execution across hundreds of Cloud Computing servers.
- 3) **Analytics:** A special case of compute-intensive batch processing is business analytics. While the large database industry was originally dominated by transaction processing, that demand is leveling off. A growing share of computing resources is now spent on understanding customers, supply chains, buying habits, ranking, and

so on. Hence, while online transaction volumes will continue to grow slowly, decision support is growing rapidly, shifting the resource balance in database processing from transactions to business analytics.

- 4) Extension of compute-intensive desktop applications:** The latest versions of the mathematics software package Matlab and Mathematica are capable of using Cloud Computing to perform expensive evaluations. Other desktop applications might similarly benefit from seamless extension into the cloud. Again, a reasonable test is comparing the cost of computing in the Cloud plus the cost of moving data in and out of the Cloud to the time savings from using the Cloud.

Cloud Architectures and Infrastructure

Cloud computing architecture comprised of two components (hardware and application). These two components have to work together seamlessly or else cloud computing will not be possible. Cloud computing requires an intricate interaction with the hardware which is very essential to ensure uptime of the application. If application fails, the hardware will not be able to push the data and implement certain processes. On the other side, hardware failure will mean stoppage of operation. Applications built on Cloud Architectures are such that the underlying computing infrastructure is used only when it is needed (for example to process a user request), draw the necessary resources on demand (like compute servers or storage), perform a specific job, then relinquish the unneeded resources and often dispose themselves after the job is done. While in operation the application scales up or down elastically based on resource needs. Applications built on Cloud Architectures run in-the cloud where the physical location of the infrastructure is determined by the provider. They take advantage of simple APIs of Internet-accessible services that scale on demand, that are industrial-strength, where the complex reliability and scalability logic of the underlying services remains implemented and hidden inside-the-cloud. The usage of resources in Cloud Architectures is as needed, sometimes ephemeral or seasonal, thereby providing the highest utilization and optimum bang for the buck. Instead of building your applications on fixed and rigid infrastructures, Cloud Architectures provide a new way to build applications on on-demand infrastructures. Cloud Architectures address key difficulties surrounding large-scale data processing. In traditional data processing it is difficult to get as many machines as an application needs. Second, it is difficult to get the machines when one needs them. Third, it is difficult to distribute and coordinate a large-scale job on different machines, run processes on them, and provision another machine to recover if one

machine fails. Fourth, it is difficult to auto scale up and down based on dynamic workloads. Fifth, it is difficult to get rid of all those machines when the job is done. Cloud Architectures solve such difficulties.

A. The on-demand, self-service, pay-by-use model

The on-demand, self-service, pay-by-use nature of cloud computing is also an extension of established trends. From an enterprise perspective, the on-demand nature of cloud computing helps to support the performance and capacity aspects of service level objectives (**Sun Microsystems, 2009**). The self-service nature of cloud computing allows organizations to create elastic environments that expand and contract based on the workload and target performance parameters and the pay by-use nature of cloud computing may take the form of equipment leases that guarantee a minimum level of service from a cloud provider. Virtualization is a key feature of this model. IT organizations have understood for years that virtualization allows them to quickly and easily create copies of existing environments -sometimes involving multiple virtual machines - to support test, development, and staging activities. The cost of these environments is minimal because they can coexist on the same servers as production environments because they use few resources. Likewise, new applications can be developed and deployed in new virtual machines on existing servers, opened up for use on the Internet, and scaled if the application is successful in the marketplace (**Sun Microsystems, 2009**). The ability to use and pay for only the resources used shifts the risk of how much infrastructure to purchase from the organization developing the application to the cloud provider.

B. Cloud computing infrastructure models

There are many considerations for cloud computing architects to make when moving from a standard enterprise application deployment model to one based on cloud computing. There are public and private clouds that offer complementary benefits, there are three basic service models to consider, and there is the value of open APIs versus proprietary ones. IT organizations can choose to deploy applications on public, private, or hybrid clouds, each of which has its trade-offs. The terms public, private, and hybrid do not dictate location. While public clouds are typically “out there” on the Internet and private clouds are typically located.

- 1) Public clouds are run by third parties, and applications from different customers are likely to be mixed together on the cloud's servers, storage systems, and networks. Public clouds are most often hosted away from customer premises, and they provide a way to reduce customer risk and cost by providing a flexible, even temporary extension to enterprise infrastructure.

- 2) Private clouds are built for the exclusive use of one client, providing the utmost control over data, security, and quality of service. The company owns the infrastructure and has control over how applications are deployed on it. Private clouds may be deployed in an enterprise datacenter, and they also may be deployed at a collocation facility.
- 3) Hybrid clouds combine both public and private cloud models. They can help to provide on-demand, externally provisioned scale. The ability to augment a private cloud with the resources of a public cloud can be used to maintain service levels in the face of rapid workload fluctuations. This is most often seen with the use of storage clouds to support Web 2.0 applications. A hybrid cloud also can be used to handle planned workload spikes. Sometimes called “surge computing,” a public cloud can be used to perform periodic tasks that can be deployed easily on a public cloud. Hybrid clouds introduce the complexity of determining how to distribute applications across both a public and private cloud.

C. Architectural layers of cloud computing

Cloud computing can describe services being provided at any of the traditional layers from hardware to applications. In practice, cloud service providers tend to offer services that can be grouped into three categories: software as a service, platform as a service, and infrastructure as a service.

- 1) Software as a service (SaaS) features a complete application offered as a service on demand. A single instance of the software runs on the cloud and services multiple end users or client organizations. The most widely known example of SaaS is salesforce.com, though many other examples have come to market, including the Google Apps offering of basic business services including email and word processing (**Sun Microsystems, 2009**).
- 2) Platform as a service (PaaS) encapsulates a layer of software and provides it as a service that can be used to build higher-level services. There are at least two perspectives on PaaS depending on the perspective of the producer or consumer of the services: Someone producing middleware, application software, and even a development environment that is then provided to a customer as a service. Someone using PaaS would see an encapsulated service that is presented to them through an API. The customer interacts with the platform through the API, and the platform does what is necessary to manage and scale it to provide a given level of service (**Sun Microsystems, 2009**).
- 3) Infrastructure as a service (IaaS) delivers basic storage and compute

capabilities as standardized services over the network. Servers, storage systems, switches, routers, and other systems are pooled and made available to handle workloads that range from application components to high-performance computing applications.

Conclusion and Future Work

In this paper we are presenting that Cloud computing promises significant benefits, but today there are security, privacy, and other barriers that prevent widespread enterprise adoption of an external cloud. In addition, the cost benefits for large enterprises have not yet been clearly demonstrated. The usage of resources in Cloud Architectures is as needed, sometimes ephemeral or seasonal, thereby providing the highest utilization and optimum bang for the buck. In Cloud, for the broader vision of Cloud Interoperability to work, ranging from VM mobility to storage federation to multicast and media streaming interoperability to identity and presence and everything in between, analogous core network extensions (or replacement) technologies need to be invented. Finally, we need improvements in bandwidth and costs for both datacenter switches and WAN routers. While we are optimistic about the future of Cloud Computing but cloud platforms aren't yet at the center of most people's attention. The attractions of cloud-based computing, including scalability and lower costs, are very real. If you work in application development, whether for a software vendor or an end user, expect the cloud to play an increasing role in your future. The next generation of application platforms is here which is "cloud; A computing infrastructure on demand."

References

- Armbrust, Michael., et al. (2009). *Above the Clouds: A Berkeley View of Cloud Computing*. CA: Electrical Engineering and Computer Sciences, University of California at Berkeley. (Technical Report No. UCB/EECS-2009-2). Retrieved from <http://www.eecs.berkeley.edu/Pubs/TechRpts/2009/EECS-2009-28.pdf>
- Bialecki, A., Cafarella, M., Cutting, D., & O'Malley, O. (2005). *Hadoop: A Framework for Running Applications on Large Clusters Built of Commodity Hardware*. Retrieved from <http://lucene.apache.org/hadoop/>
- Dean, J., & Ghemawat, S. (2004). MapReduce: simplified data processing on large clusters. In *Proceedings of the 6th conference on Symposium on Operating Systems Design & Implementation* (Vol. 6) (p. 10). Berkeley, CA, USA: USENIX Association. Retrieved from

<http://dl.acm.org/citation.cfm?id=1251254.1251264>

Li, Hong., et al. (2009). *Developing an Enterprise Cloud Computing Strategy*. Intel Information Technology. Retrieved from http://www.intel.com/en_US/Assets/PDF/whitepaper/icb_cloud_computing_strategy.pdf

Sun Microsystems. (2009). *Introduction to Cloud Computing architecture White Paper (1st Ed.)*. Retrieved from <http://eresearch.wiki.otago.ac.nz/images/7/75/Cloudcomputing.pdf>

Vogels, W. (2008). A Head in the Clouds – The Power of Infrastructure as a Service Infrastructure as a Service. In First workshop on Cloud Computing and In Applications (CCA'08).