Emerging Utility and Cloud Computing Models

Ivan I. Ivanov

Empire State College, State University of New York
Long Island Center, Hauppauge, NY 11788, U.S.A.
ivan.ivanov@esc.edu

Abstract. This paper explores emerging opportunities and challenges for end-users and corporations with the shift in the geography of computing services and applications. Recent trends in consolidated Enterprise Information Technology solutions have proven to enhance business efficiency when significant fractions of local computing activities are migrating away from end-user computers and enterprise servers. Data, Information Systems and Enterprise Applications are being removed from desktop PCs and departmental servers and are being integrated and packaged on the Web into "the computing cloud." The shift from locally installed and maintained programs and systems to hosted services in "the compute cloud" affects the entire computational ecosystem, from the CIO to the end-user, software developer, systems analyst, and IT/IS vendor. We may need to take a fresh look at Thomas Watson's pivotal statement from 1943 "I think there's a world market for maybe five computers." There are several distinct innovations towards emerging centralized computational mega-structures such as:

- Grid computing infrastructure, middleware and application services
- Virtualization for more transparent scaling of complex systems
- Utility Computing services and models
- Dynamic Resources and Computing in the Cloud.

Growingly companies seem to achieve business advantage from the recent IT transformations and different forms of Cloud or Utility Computing services. Conversely, many issues remain: how well do the new models protect our investments, business specific challenges, innovations, privacy, and information security?

Keywords. Utility Computing, Cloud Computing, Internal-, External-, Private-, Community-, Public Cloud, Cloud Software as a Service, Cloud Platform as a Service, Cloud Infrastructure as a Service.

1 Introduction

The past decade has witnessed important Information Technology (IT) advances and shifts in geography of computing services, applications and capabilities. The IT transformations and changes stimulate business adaptations and social adjustments to a great extend.

The economic difficulties in 2000s, the cost-effective strength of the Internet, and some new technological advances have made businesses more vigilant and more demanding about the return of their IT investments [6]. The economic and social

Ivanov I.
Emerging Utility and Cloud Computing Models.
DOI: 10.5220/0004463100480056
In *Proceedings of the 3rd International Workshop on Enterprise Systems and Technology (I-WEST 2009)*, pages 48-56
ISBN: 978-989-674-015-3

motivation for efficient consolidated IT solutions such as utility or cloud computing is steadily increasing. Insightful businesses and IT organizations are grasping the ideas, discerning what the emerging technology and models are and how they can use them to create a competitive advantage. Any delays in understanding and adapting novelties might lead to dramatic changes or collapses. One of the many examples from the first decade of the century is with the retail book industry - it was slow in adapting to the Internet and E-commerce opportunities and got “Amazoned.”

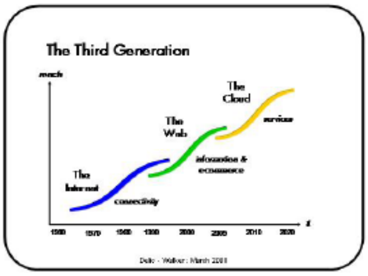

Fig. 1. Internet evolving into Cloud [1].

The magnitude of the Internet revolution broadens with the advancement of the infrastructure, operations and available services though the global network. According to Delic and Walker [1], the evolution of the Internet infrastructure and associated activities (Fig. 1) is heading in the Third Generation – The Cloud Services.

There is a new wave of interest in “Externalization of IT” --anything as a service (including Software as a Service - SaaS, Infrastructure as a Service - IaaS, Platform as a Service - PaaS), On-Demand, Utility and Cloud Computing, Outsourcing-- and many businesses are exploring any alternatives that would allow them to reduce IT operating costs while improving and intensifying information resources. Growing companies appear to achieve business advantage from the recent IT transformations, different forms and delivery models of Utility or Cloud Computing services.

New Utility and Cloud computing services are changing the ways we think of IT, from economics of delivery mode to process and usage models. Many organizations look to the cloud as a potential cost-savings boon by moving internally hosted IT services to external providers. Other IT organizations view the Cloud as a potential disaster recovery plan or as on-demand elastic capacity to boost business continuity and operating service levels. The shift from locally installed and maintained programs and systems to hosted services in “the compute cloud” affects the entire computational ecosystem, from the CIO to the end-user, software developer, systems analyst, and IT/IS vendor. We may need to take a fresh look at Thomas Watson’s pivotal statement from 1943 “I think there’s a world market for maybe five computers.”

Just as it was with that new-fangled "Internet" mania few decades ago, the Cloud and the technologies of Cloud Computing suffer from confusion and hype [2]. Multiple proprietary related definitions, competing vendors and alliances, undefined risks, new business models’ obscure utility and cloud computing services are slowing down adoption. Although there is tremendous interest in the “Externalization of IT” services, most organizations are approaching it cautiously until they have a more complete picture of the risks involved.

Beyond important economic aspects of the emerging utility and cloud computing models, I assume the business sector and society at large will see benefits and opportunities in a global scale of utilizing those services. In this paper, I will describe some of these new business oriented systems and models in more detail.

2 The Evolution of Computing Mega-Structures

As the demand of faster and more powerful computing structures increases, and as the number and variety of applications and services progressively elevates, the need for more capable and dynamic computing environments increases, too. According to the IBM Corporation White Paper "Seeding the Cloud," the evolution toward Cloud Computing started in the late 1980s with the concepts of grid computing [4].

Grid computing specifically refers to leveraging massive number of computers in parallel to solve particular problems, or to run specific applications. The key element of Grid computing is that computers, or nodes, in a grid are able to act independently without centralized control, handling requests as they are made and scheduling others. Grid computing is the underlying technology for utility computing. In a long term, grid computing is heading towards a convergence of utility computing from the pricing and delivery prospective, and Web services-based integration and virtualized technologies to enable multiple, networked computers to be managed as one [13]. Amongst vendors developing and exploiting grid concepts are HP with HP Adaptive Enterprise Initiative, Sun Microsystems Network One, IBM's On-Demand Computing, and Oracle Grid Computing.

In the late 1990s, with virtualization of systems, servers and applications, the model has expanded to higher level of abstraction – a virtual platform, including storage and network resources, and subsequently virtual applications, which have no specific underlying infrastructure. Utility computing has offered clusters as virtual platforms for computing with a metered business model [4]. The utility computing uniquely integrates storage, applications, computing power and network infrastructure as a foundation for business adjustable IT services. It offers companies and private users an access to hosted computing services, scalable and portable business applications through a utility-like, pay-on-demand service over the Internet network. In the ultimate utility computing models, organizations will be able to acquire as much IT services as they need, whenever and wherever they need them [7].

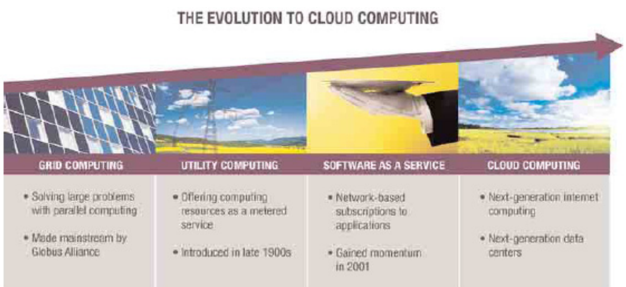

Fig. 2. Evolution toward Cloud Computing [4].

Lately, software as a service (SaaS) has elevated the level of virtualization to applications. The SaaS model has been developed to overcome common enterprise challenges to meet fluctuating demands on IT resources efficiently. Whether referred to a SaaS, utility computing, or hosted services, the idea is basically the same: instead of buying, installing and supporting expensive packaged enterprise applications or systems, users can access and utilize "externalized" applications over the network and pay a fixed subscription fee, or an actual usage fee [14].

The concept of cloud computing has evolved from the concepts of grid, utility and SaaS. In reference to IEEE Computer Society definition cloud computing "is a paradigm in which information is permanently stored in servers on the Internet and cached temporarily on clients that include desktops, entertainment centers, table computers, notebooks, wall computers, handhelds, etc." Cloud computing is an emerging model through which users can gain access to their applications and systems from anywhere, at any time, through their connected devices. These applications and services reside in massively scalable data centers, structured in public or private clouds, where compute resources can be dynamically provisioned and shared to achieve significant economics of scale. The strength of the cloud computing model is its infrastructure management, enabled by the maturity and progress of virtualization technology to manage and better utilize the underlying resources through automatic provisioning, re-imaging, workload balancing, monitoring, change request handling, and dynamic and automated security and resiliency platform [4].

3 Cloud Computing Characteristics and Models

Cloud computing is still an evolving model and definitions, underlying technologies, issues, risks, and benefits have been developed in vigorous debates by the providers, public and private sector users. There are neither standards nor agreed definitions related to the Cloud computing. The most illustrative description of what the cloud is comes from Kevin Marks, of Google "The idea of cloud computing comes from the early days of the Internet, where we drew the network as a cloud. We didn't care where the messages went --they came in one side and out the other-- and we didn't have to worry about the network … the cloud hid it from us."

According to Wyatt Kash, editor in chief for Government Computer News, "A group of leading standards development organizations are working jointly to foster common standards for cloud computing and storage, beginning with the launch a new site called cloud-standards.org." The organizations joining in the collaborative effort include the Cloud Security Alliance, the Distributed Management Task Force, the Open Grid Forum, the Storage Networking Industry Association and the Open Cloud Consortium, with other groups expected to participate. The Clouds Standards Coordination working group plans to focus its efforts on identifying current and emerging practices and products supporting cloud computing. In addition, it expects to help rationalize cloud computing security, deployment and data-exchange formats, taxonomies and reference models, storage standards and service-level agreements [8].

3.1 Cloud Computing Definitions and Characteristics

The National Institute of Standards and Technology, U.S.A. has drafted a definition of Cloud computing to serve as a starting point for government agencies and to maintain the implementers and providers on the same track.

According Peter Mell and Tim Grance from NIST "Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. As the cloud computing industry represents a large ecosystem of many models, vendors, and market niches, this definition attempts to encompass all of the various cloud approaches. This cloud model promotes availability and is composed of five essential *characteristics*, three *delivery models*, and four *deployment models*" [5].

In relation to NIST definition the five key cloud ***Characteristics*** are [9]:

- *On-demand self-service* - a consumer can unilaterally provision computing capabilities, such as server time and network storage, as needed automatically without requiring human interaction with service's provider.
- *Ubiquitous network access* - capabilities are available over the network and accessed through standard mechanisms that promote use by heterogeneous thin or thick client platforms.
- *Location independent resource pooling* - the provider's computing resources are pooled to serve all consumers using a multi-tenant model, with different physical and virtual resources dynamically assigned and reassigned according to consumer demand. The customer generally has no control or knowledge over 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 to quickly scale up and rapidly released to quickly scale down. To the consumer, the capabilities available for provisioning often appear to be infinite and can be purchased 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 utilized service.

3.2 Cloud Computing Delivery and Deployment Models

NIST team has defined three ***Delivery Models*** based on the variety of services requested by the consumer and supported by the provider. Delivery Model Architecture – see Figure 3, must be deployed on top of cloud infrastructure that incorporates the listed above key characteristics [9]:

- *Cloud Software as a Service (SaaS)* - the capability provided to the consumer is to use the provider's applications running on a cloud infrastructure and accessible from various client devices through a thin client interface such as a Web browser. The consumer does not manage or control the underlying cloud infrastructure, network, servers, operating systems, storage, or even individual application capabilities.
- *Cloud Platform as a Service (PaaS)* - the capability provided to the consumer is to deploy onto the cloud infrastructure consumer-created applications using programming languages and tools supported by the provider. The consumer does not manage or control the underlying cloud infrastructure, but the consumer 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 to rent 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 operating systems, storage, deployed applications, and possibly select networking components.

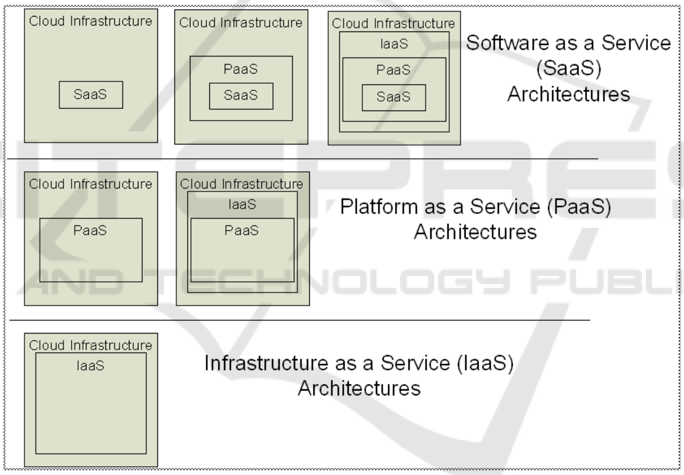

Fig. 3. Delivery Model Architectures [9].

In the same NIST document, Peter Mell and Tim Grance have described four ***Deployment Models*** specified in two types of clouds: *internal* or *external*, depending on where the clouds reside – within or outside of organization's network security parameters:

- *Private cloud* - the cloud infrastructure is owned or leased by a single organization and is operated solely for that organization.
- *Community cloud* - the cloud infrastructure is shared by several organizations and supports a specific community that has shared concerns.
- *Public cloud* - the cloud infrastructure is owned by an organization selling cloud services to the general public or to a large industry group.

- *Hybrid cloud* - the cloud infrastructure is a composition of two or more clouds (internal, community, or public) that remain unique entities but are bound together by standardized or proprietary technology that enables data and application portability.

The NIST drafted definition, characteristics and models are in compliance with currently well positioned platforms such as VMware's vSphere and the open source Eucalyptus software. Similarity and matching descriptions related to types of clouds and delivery models appear with VMware Cloud Business Types, see Figure 4, and NIST defined above models [9].

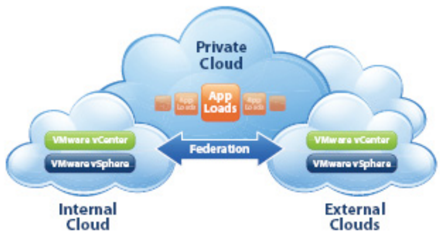

Fig. 4. VMware vSphere and vCenter ***on*** and ***off*** premise clouds [15].

4 Utilizing the Cloud to Tackle Limitations

A growing number of investors and consumers are exploring cloud options and opportunities while Amazon, Google, Oracle, SUN Microsystems, VMware already have appealing offerings and leading companies in the sector such as Microsoft, HP and IBM are investing billions in this initiative. According to leading vendors' and implementers' the new mega-trend - Cloud computing, has become a bona fide enterprise solution when users want to avoid dealing with infrastructure and configuration management. It works well as a consumer service such as SaaS, Web and DB hosting, email, and offsite PCs backup. On the developer side, it lends itself to Web applications design, content delivery, and building online systems and applications where scale, performance, and selection of infrastructure and software environment aren't critical. The provider or the service partners' ecosystem are in complete control of the infrastructure, the platform and the services.

Grid and Utility computing users, conversely, are more inclined to install, arrange, scale and manage online services using the provider's resources being in most cases in full control of the infrastructure and applications. There is immense variety in possible and actual configurations of technologies and infrastructure to support utility and grid computing models. According to Alfredo Mendoza [10], well established and proven technologies like virtualization, advanced application accounting, and dynamic partitioning, that have long existed in mainframes and now are available on newer server architectures in combination with grid computing, web services and hyper-threading technologies are contributing to create grid infrastructures and promote the utility model. The two models fit well where the implementers need to control geographic distribution, scalable compute resources, bandwidth, and performance in order to manage the user experience as the applications are in their own data center. Analyzing the three concepts – grid, utility and cloud computing, their evolution and differences listed above, the final argument actually is: outsourcing IT operations to providers or purchasing your own IT assets.

All three models have proved, in many cases, better and more efficient ways of delivering enterprise IT services, but the final decision may well come down to how confident the user feels about the utility or cloud computing provider. When planning the shift to the clouds there are different options depending on what you want to put there and for how long. Although vendor offers similar services, while moving your computing applications away from your own premises, they might apply quite different pricing models and SLAs. Users have to take into account the current needs, potential growth and ability to change the SLA with forms and cloud services that suit the company best. In her review of recent cloud services, Jarina D'Auria from CIO magazine describes Amazon's Elastic Compute Cloud (EC2) as an environment to run computing resources while keeping the control over the data in the user's hands and emphasizing pay-per use model. The new Google App Engine allows you to build your own virtual application to run Web applications on Google's servers and features include dynamic Web serving, automatic scaling and load balancing, storage sorting, APIs for authenticating users and different pricing scales. The VMware vSphere 4 and vCloud initiative are providing the capability to move physical infrastructures and broader application capability on and off premise via hundreds of cloud services providers. By moving all physical data centers, companies not only save computing and operation costs but also have one silo for storage and resource management , based on vMotion management tools, VMsafe security APIs and data recovery systems [11].

Applying any of the available cloud services, the users will have, in most cases, unlimited computing resources, elastic infrastructure and more opportunities for new applications at a price they are able to pay and with no upfront costs. Strategizing the company final solution it might be most appropriate to build a specific hybrid model incorporating internal private cloud for the most sensitive data and critical applications and to outsource into public or community clouds some less decisive systems and routine applications. Evolving the internal IT infrastructure toward a cloud-like model for supporting the most sensitive and critical business operation at lower risk and higher security will enhance and facilitate the integration with outside applications running into external clouds [12]. In many industries such approach would decrease the risk of total outsourcing and the company will better control its information assets. Amongst higher educational institutions most of them keep and control in their own data centers students' records and financial systems while widely utilized educational services from external providers such as on-line course delivering environment (ANGEL, Blackboard), student consulting and supporting services as Smarthinking, email and e-collaborative tools from Google (Google Apps), or back up and archive services from Amazon (S3) or other vendors.

5 Conclusions

The shift to cloud computing will shake the $3.4 trillion global tech industry, and as the marker researchers Gartner expects the 3.8% shrinking this year will not affect portables, wireless networks and cloud computing over the next few years. Gartner predicts the market for cloud computing services will triple from $46.6 billion last year to 150.1 billion in 2013 [3].

Ultimately, cloud computing is hovering to make a real impact on businesses of all sizes and society at large. It gives to smaller businesses a chance to utilize resources and services previously only IT leading companies could afford. According Vivek Kundra, Federal CIO "The cloud will do for government what the Internet did in the '90s … I believe it's the future … It's a fundamental change to the way our government operates by moving to the cloud. Rather than owning the infrastructure, we can save millions … It's moving technology leaders away from just owning assets, deploying assets and maintaining assets to fundamentally changing the way services are delivered."

As any change this new mega-trend in IT industry and services may take a while for building users' confidence, but like any major concept, it always starts off slowly, but once businesses and society realize the advantages of cloud computing, it will pick up pace rapidly.

References

1. Delic, K., Walker, M.A.: *Emergence of The Academic Computing Cloud*. ACM Ubiquity, Volume 9, Issue 31 (2008)
2. Fingar, P.: *Dot Cloud: The 21st Century Business Platform Built on Cloud Computing*, Meghan-Kiffer Press, FL, U.S.A. (2009)
3. Hamm, S., *Cloud Computing Big Bang For Business*. BusinessWeek, (June 15, 2009
4. IBM Corporation: *Seeding the Cloud: Key Infrastructure Elements for Cloud Computing.* IBM Systems and Technology Group, Route 100, Somers, NY 10589, U.S.A. (2009)
5. InformationWeek: *Government IT On The Leading Edge.* Analytics Alerts, InformationWeekanalytics.com (2009)
6. Ivanov, I.: *Utility Computing Paradigm and SOA Philosophy.* I-WEST'2008, INSTICC PRESS, Portugal (2008)
7. Ivanov, I.: *Utility Computing: Reality and Beyond*. E-Business and Telecommunications, Springer (2008)
8. Kash, W: *Standards groups form alliance to set cloud-computing standards*, http://gcn.com/articles/2009/07/17/standards-groups-set-cloud-computing-standards.aspx (July 19, 2009)
9. Mell, Peter, Grance, Tim: *Draft NIST Working Definition of Cloud Computing*. http://csrc.nist.gov/groups/SNS/cloud-computing/index.html (July 20, 2009)
10. Mendoza, A.: *Utility Computing: Technologies, standards, and strategie*s. Artech House, Norwood, MA 02062, U.S.A. (2007)
11. Nash, K. and D'Auria, J.: *Cloud Control.* CIO magazine (June 1, 2009)
12. Robbins, D.: *Enter the Cloud.* NetworkWorld, U.S.A. (May 18, 2009)
13. The 451 Group: *Grid Technology User Case Study: JP Morgan Chase*. The 451 Group Report, NY, U.S.A. (2003)
14. Turban, E., Volonino, L.: *Information Technology for Management: Improving Performance in the Digital Economy*, John Wiley & Sons, NJ, U.S.A. (2009)
15. VMware: *Eight Key Ingredients for Building an Internal Cloud.* White Paper from VMware, Palo Alto, CA 94304, U.S.A. (2009)