Cloud Computing Acceleration Point

Abstract

Cloud computing has likely reached the market adoption acceleration point signaling a period of dramatic growth of this transformational operating model for enterprise information infrastructure. Cloud computing will expand from its base of software services to infrastructure services hosting core business operations and data. The confluence of experience and success with software cloud services, extreme budgetary pressures, dramatic new infrastructure service offerings, and enhanced security and disaster recovery concern is large companies and government agencies to increase the role cloud services play in their business model.

Introduction

Cloud computing has likely reached the market adoption acceleration point signaling a period of dramatic growth of this transformational operating model for enterprise information infrastructure. Cloud computing will expand from its base of software services to infrastructure services hosting core business operations and data. The confluence of experience and success with software cloud services, extreme budgetary pressures, dramatic new infrastructure service offerings, and enhanced security and disaster recovery concern is large companies and government agencies to increase the role cloud services play in their business model.

Market adoption of innovative technologies and processes follows a transition curve\(^1\) illustrated in Exhibit 1. Risk tolerant innovators and early adopters experiment with unproven new technologies and processes based on their assessment of potential returns. They prove the viability of the innovation and help improve its reliability and functionality by identifying weaknesses and defects. With viability and benefit demonstrated, more pragmatic and conservative organizations, representing the bulk of the market, become adopters accelerating installed base expansion. The cloud computing market has reached that acceleration point.

This paper reviews cloud computing concepts, how the cloud computing market was established, cloud infrastructure service offerings that are the basis for accelerated market adoption, and cloud computing security. The discussion addresses reality verses hype, risks, and benefits. This paper is
not intended to survey all cloud offerings, nor all cloud providers, but to put the transition to cloud services in a pragmatic context to aid in decision making. This paper focuses on enterprise infrastructure and not on individual consumer services.

Cloud computing concepts

As with most information technology innovations cloud computing created a great deal of confusion as to its meaning and generated enormous hype about its benefits and risks. This confusion led the National Institute of Standards and Technology (NIST) in September 2011 to issue the following definition of cloud computing:

Cloud computing is a model for enabling ubiquitous, 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.

In conjunction with this definition NIST defined three service models:

- Software as a Service (SaaS). The capability provided to the consumer is to use the provider’s applications running on a cloud infrastructure.

- Platform as a Service (PaaS). The capability provided to the consumer is to deploy onto the cloud infrastructure consumer-created or acquired applications created using programming languages, libraries, services, and tools supported by the provider.

- Infrastructure as a Service (IaaS). The capability provided to the consumer is to provision 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.

NIST completed its definition with four deployment models:

- Private cloud. The cloud infrastructure is provisioned for exclusive use by a single organization comprising multiple consumers (e.g., business units).

- Community cloud. The cloud infrastructure is provisioned for exclusive use by a specific community of consumers from organizations that have shared concerns.

- Public cloud. The cloud infrastructure is provisioned for open use by the general public.

- Hybrid cloud. The cloud infrastructure is a composition of two or more distinct cloud infrastructures.

Although the NIST definitions were intended for Federal Government agencies they are widely used in the private sector. However, the distinction between platform as a service and infrastructure as a service is often overlooked and both are often referred to as infrastructure as a service. The distinction between virtualization (a technology) and cloud (a service) is sufficiently vague that the terms are often used interchangeably.
The U.S. Chief Information Officer’s (CIO) 25 Point Implementation Plan to Reform Federal Information Technology Management includes three points directly addressing cloud computing:

1. **Shift to a “Cloud First” policy:** When evaluating options for new IT deployments, OMB will require that agencies default to cloud-based solutions whenever a secure, reliable, cost-effective cloud option exists.

2. **Stand-up contract vehicles for secure IaaS solutions:** GSA will make a common set of contract vehicles for cloud-based Infrastructure-as-a-Service solutions available government-wide.

3. **Stand-up contract vehicles for commodity services:** GSA will stand up government-wide contract vehicles for cloud-based email solutions [and] other back-end, cloud-based solutions.

The 25 Point plan articulates three broad benefit areas which are further detailed in the U.S. CIO’s Federal Cloud Computing Strategy:

- **Economical:** Cloud computing is a pay-as-you-go approach to IT, in which a low initial investment is required to begin, and additional investment is needed only as system use increases.

- **Flexible:** IT departments that anticipate fluctuations in user demand no longer need to scramble for additional hardware and software. With cloud computing, they can add or subtract capacity quickly and easily.

- **Fast:** Cloud computing eliminates long procurement and certification processes, while providing a near-limitless selection of services.

These cloud computing benefits are only fully realized with public cloud services which are the subject of this paper. Public cloud service providers (CSPs) initially established large information technology infrastructures for their own use providing online services to millions of individual consumers. Then they offered to expand their infrastructure for individual consumers and enterprise cloud computing. The key is that their infrastructure is so large (estimated at over a million servers and an Exabyte of data storage) that, to their customers, the resources appear infinitely flexible and inexhaustible.

Enterprises establishing a private cloud must continue to invest in information technology infrastructure. Pay-as-you-go is achievable at the component organization level through usage-based charge-back. The entire cost of the private cloud infrastructure is still absorbed at the enterprise level, just as in non-cloud environments.

Enterprise private clouds can achieve economies by standardizing hardware platforms and operating systems and by virtualizing the infrastructure to optimize usage. Many organizations built their infrastructure incrementally to serve component organization specific needs resulting in suboptimal usage levels as low as 5-10%. Virtualization, the underlying cloud technology, has reduced infrastructure for some enterprises by factors close to ten.

Private clouds can achieve flexibility and speed allowing component organizations to dynamically expand and contract allocated capacity. However, the sum of concurrent demand from all
component organizations is limited by the enterprise installed infrastructure. The principal difference between private and public clouds is the size of the infrastructure.

**How the cloud computing market was established**

The cloud computing market was established primarily with email. Every enterprise is dependent on email for internal and external communications. Most enterprises manage internal email systems. Individual consumers have turned to email services provided either by their Internet service provider or by free cloud email providers.

Microsoft Hotmail, established in 1996, has been the cloud email leader with 360 million users as of July 2011. Google Gmail, established in 2004, claims to have 425 million users. Yahoo Mail, established in 1997, is reported to have 300 million users. With over a billion individual users of these three cloud services, cloud based email came to be trusted as a reliable, available, capable, and secure software service.

Cloud email providers leveraged their brands and infrastructure to offer cloud based email service to enterprises. At its I/O 2012 conference, Google announced that 5 million businesses, 66 of the top 100 universities, and government agencies in 45 states use Gmail as their enterprise email platform. At the Federal level the National Atmospheric and Oceanic Administration (NOAA) and General Services Administration (GSA) also adopted Gmail.

Two other well known cloud software service providers are ServiceNow and Salesforce. ServiceNow provides cloud based software service for help desk operations. In 2011 ServiceNow reported three year growth rate of 1240%. ServiceNow market growth was aided by significantly lower price structure than its expensive key competitors BMC Remedy and CA Unicenter.

Salesforce provides cloud base customer relationship management software service. It provides all of the tools needed to enable corporate sales staff. Salesforce maintains customer account lists and contacts, marketing leads, marketing materials (e.g., presentations, brochures, data sheets), opportunity pipeline, quotes, sales process automation, and business analytics and forecasting.

These and many other similar functionally focused software service offerings established the cloud computing market. Cloud computing is no longer perceived as technology hype or a futuristic concept. Enterprises are taking a serious look at cloud computing and evaluating the business case and risks of infrastructure cloud services.
Cloud infrastructure service offerings

Infrastructure as a service has been available in different forms since the 1960’s when time-sharing contracts offered access to mainframe computers owned by service providers and operated in the service provider’s data center. Many Federal agencies used time-sharing services. In the 1990s infrastructure as a service was known as outsourcing. An enterprise sold its information technology and even data centers to a service provider who sold the time back to the enterprise on a usage basis.

While many businesses eagerly adopted the outsourcing model to reduce infrastructure costs and focus executive attention on core competencies, few Federal Government agencies followed suit. Federal agencies did not want to depend on a commercial entity for their information technology operations even though they were willing to depend on commercial entities for their telecommunications (telephone and wide area network). They simply did not trust that the outsourcer would routinely provide them with the quality of service necessary for mission success.

Federal agencies also faced the dilemma of how to re-compete outsourced information technology services. All outsourcing contracts have termination (unwind) clauses should the customer decide to switch outsourcers or insource the information technology. In the commercial world, these clauses are rarely exercised. The outsourcer and customer work to resolve any issues and negotiate competitive prices. Contracts are routinely renewed on a sole source basis.

The Federal Acquisition Regulation requires regular re-competition of contracts. Federal agencies potentially face the daunting task of moving their entire information technology infrastructure to a new provider at each contract award. Aside from the complex task of having the new provider purchase the entire infrastructure from the previous provider, risks associated with moving the enterprise infrastructure to the new provider’s data center made the commercial outsourcing model untenable for most Federal agencies.

Twenty-first century infrastructure as a service has significantly reduced the barriers to market adoption. The services offered by Amazon and Google illustrate characteristics of the current cloud infrastructure service marketplace. Amazon, the largest Internet retailer with $48 billion in 2011 sales, is also reported to be the largest cloud service provider with 2011 sales of Amazon Web Services estimated to be $6 billion. Amazon entered the cloud infrastructure services market in 2006.

Google, the provider of a multitude of free Internet software services, announced full service entry into the cloud infrastructure services market on June 28, 2012. Google 2011 revenue was $37.9 billion, primarily from advertising ($36.5 billion). Although six years behind Amazon, the strength of the Google brand, the quality of Google’s infrastructure, and Google’s success dominating most of the markets it enters are harbingers of their role in this burgeoning market.
Amazon and Google built and operate enormous information technology infrastructures to power their operating models. Each company has dedicated data centers located in the United States (Exhibit 2) and around the world. Neither company releases information about the size of their infrastructure. In a 2009 video, Google showcased a new data center housing 45,000 servers. With Google’s 30% annual growth rate and multiple data centers in six domestic locations and five foreign locations, estimates are that Google hosts over one million servers and one Exabyte of data storage. Amazon likely has an infrastructure at similar scale. Enterprise public cloud customers now have access to these massive server and storage farms.

Amazon recognized that it had created an infrastructure of scale an order of magnitude or more larger than most enterprises. Amazon also recognized that its brand as a successful Internet retailer could be leveraged to offer cloud infrastructure services. Google is now counting on its Internet service brand in a similar manner.

Amazon Simple Storage Service (Amazon S3) is storage for the Internet. Amazon Elastic Compute Cloud (Amazon EC2) is a web service that provides resizable compute capacity in the cloud. Amazon Elastic Block Store (EBS) provides block level storage volumes for use with Amazon EC2 instances. Amazon Web Services comprise S3, EBS, EC2, and many additional cloud offerings.

Google offers very similar services including Google Cloud Storage and Google Compute Engine under the brand Google Cloud Platform. At its announcement of Compute Engine, Google demonstrated the power of its infrastructure with a client’s genome mapping calculation. The client’s in house infrastructure with 1000 servers was reported to complete one calculation in ten minutes. Google launched the calculation, allocated 600,000 cores in real time, and was able to complete multiple calculations per second.

Amazon, Google, and many other CSPs’ services are ubiquitous in the sense that they are available wherever one has an Internet connection and provide on-demand access to a shared pool of configurable resources (storage, server, application). They align with the U.S. CIO’s definition of flexibility in that IT departments can add or subtract capacity quickly and easily.
Benefits and risks of public cloud computing

Issuance of cloud computing contracts is likely to take as long as any other Government procurement. However, once GSA issues Government Wide Acquisition Contracts (GWAC) to CSPs, other agencies will be able to obtain these services quickly.

CSP services tend to be pay-as-you-go with low initial investment and additional investment only as needed for agency system usage demand. Whether this turns out to be economical for agencies depends on CSP pricing structure, internal cost structure for comparable services, and how efficiently an agency uses CSP services.

As an example, enterprise scale cloud storage pricing is approximately $1000/TB/year. This is approximately equal to the purchase price of enterprise grade storage. Cloud storage service includes multi-data center replicated storage and includes all networking, data center management, and infrastructure management costs.

Gartner reports the average total cost of ownership for enterprise raw configured storage as $4876 with 50% of enterprises in the range $2200 - $6200/TB/year. For direct comparison of in house useable storage costs with cloud storage costs the Gartner number must be inflated to $13,900/TB/year. For most enterprises, cloud storage offers significant cost benefits. If an enterprise moves their data to a CSP but maintains the empty floor space in their datacenter, they will only realize a fraction of the apparent cost benefit. Similarly, subscribing to more cloud computing capacity than needed results in sub-optimal usage levels and excess cost.

Adequate network bandwidth at the CSP and the customer locations, and response time of CSP storage and processor systems are critical requirements for successful cloud computing operations. Google and Amazon published service level agreements do not include bandwidth or response times. The only service commitment they make is to availability, either 99.9% or 99.95% depending on the service. Given the vast size of their infrastructure and exceptional response time to individual customer requests it is likely that performance will not be an issue for enterprise customers.

The U.S. CIO’s 25 Point Implementation Plan to Reform Federal Information Technology Management describes experiences of two new Internet offerings to support the cloud first strategy: a private sector web-based media production company that grew from 20,000 to 250,000 customers in three days successfully scaling their infrastructure from 50 to 4000 virtual machines; the Federal Government’s Car Allowance and Rebate System (CARS, more commonly known as “Cash-For-Clunkers”) which failed when demand far exceeded projections. Google, as part of its Compute Engine announcement, demonstrated the ability to scale in real time to 600,000 cores for a client’s computational problem.

These examples demonstrate a clear benefit of CSP offerings. However, the benefit applies only to a very small market segment. Relatively few enterprises experience order of magnitude growth spurts in a few days and even fewer have short term computation problems requiring hundreds of thousands of cores for short durations. Most enterprises have relatively steady or slowly growing
compute and storage workloads with daily or weekly cycles. Enterprises launching a new information technology service on the Internet or to a supplier or customer network may well benefit from the flexibility to rapidly scale infrastructure CSPs offer. For most enterprises this is a potential benefit that will rarely if ever be realized.

A significant benefit of cloud computing mentioned briefly in the Federal Cloud Computing Strategy is in continuity of operations (COOP). Cloud storage services include data replication ensuring that the loss of a single data center does not result in loss of persistent data. CSPs with adequate server capacity enable computational workload to shift from a non-operational data center to an operational data center. However, as the Amazon outages discussed later demonstrate, the loss of a CSP data center will impact customers until capacity is restored or transferred. Enterprises leveraging cloud computing still need to address how they will transfer load and all other aspects of COOP planning including staff, facilities, networks, and other mission specific needs.

Moving enterprise application processing to public cloud infrastructure requires software compatibility with the CSP environment. Google Compute Engine\textsuperscript{30} offers Ubuntu and CentOs Linux operating systems. Amazon Elastic Compute Cloud\textsuperscript{31} offers SUSE and Red Had Linux in addition to Microsoft Windows operating systems.

One of the challenges\textsuperscript{32} facing all data center operators is managing energy costs and collateral environmental impact. Many data centers’ power usage effectiveness (PUE = total facility power / information technology equipment power)\textsuperscript{33} exceeds 2.5 implying 2.5 kW of power are consumed by the facility for each 1 kW of power consumed by the information technology equipment.

The Uptime Institute 2012 survey\textsuperscript{35} found an average PUE of 1.8 to 1.89 with 9% of responses from 1,100 data center owners and operators reporting PUE ≥ 2.5. Industry averages are likely higher than 1.89 because of the selective survey population and the 29% of responders in the survey that do not measure PUE.

Amazon has achieved\textsuperscript{36} PUE = 1.45 in its data centers. Google, with somewhat newer data centers, has achieved\textsuperscript{37} PUE = 1.13. Enterprises leveraging cloud services benefit from lower power costs and environmental footprint.

Public cloud computing infrastructure services offer the potential for lower cost, flexibility to meet short or long term workload, support for continuity of operations, high availability, and exceptional power usage effectiveness. The most significant benefit may be relieving senior executives of the burdens associated with managing in house infrastructure enabling them to increase focus on the enterprise core mission. Migrating to a reliable and capable CSP comes with limited risks associated primarily with performance and compatibility.
Security and cloud computing

Security continues to be a concern impeding adoption of public cloud computing because conventions have not yet been developed for the appropriate level and form of security testing of cloud providers’ technology\(^\text{38}\). The Uptime Institute 2012 survey\(^\text{39}\) found security concerns were an impediment for 64% of respondents.

The security certification and accreditation process\(^\text{40}\) necessary to obtain authorization to operate (ATO) has been a long standing barrier to Government adoption of public cloud computing services. The Government is working to dramatically lower this barrier with the Federal Risk and Authorization Management Program (FedRAMP)\(^\text{41}\).

FedRAMP is a government-wide program that provides a standardized approach to security assessment, authorization, and continuous monitoring for cloud products and services. This approach uses a “do once, use many times” framework that will save cost, time, and staff required to conduct redundant agency security assessments.

The purpose of FedRAMP is to:

- Ensure that cloud based services used government-wide have adequate information security;
- Eliminate duplication of effort and reduce risk management costs; and
- Enable rapid and cost-effective procurement of information systems/services for Federal agencies. ...

A CSP follows the process for a provisional authorization under FedRAMP and uses a 3PAO [third party assessor] to assess and review their security control implementations. CSPs then provide documentation of the test results in a completed assessment package to the FedRAMP PMO [Program Management Office]. The security package is then reviewed by the JAB [Joint Authorization Board] and if a CSP system presents an acceptable level of risk, a provisional Authorization [to Operate] is granted. Agencies can then leverage the Provisional ATO and grant their own ATO without conducting duplicative assessments.

In the early days of cloud computing, concerns were raised that the virtualized environment would enable users to penetrate the domains of other users in the cloud. These fears do not appear to have been realized. Google Gmail security breaches have made headlines when accounts of high profile users\(^\text{42}\) were hacked. These breaches all appear related to weak passwords and security questions used by account owners and not to flaws in Gmail security controls\(^\text{43}\).

Amazon suffered two recent major data center outages\(^\text{44}\), one each in 2011 and 2012, that resulted in user service interruptions. More recently, a Yahoo.com security breach\(^\text{45}\) compromised thousands of user account passwords.

These incidents should not deter enterprises from leveraging public cloud computing. They just illustrate that no one has perfect security and enterprises leveraging cloud computing must maintain security vigilance, plan for security breaches, and train all employees on best practices to avoid vulnerabilities. CSP security is likely neither a risk nor a benefit for enterprises that implement best practice security for their in house systems. For organizations concerned about...
OnPoint Consulting, Inc.  Cloud Computing Acceleration Point

vulnerabilities in their security architecture, cloud computing leveraging a proven CSP offers a significant benefit.

Conclusion

Cloud computing software services are a mature market with many proven successful providers. Cloud infrastructure services offer enterprises significant financial and operations benefits. With an expanding base of infrastructure CSPs and growing confidence in the cloud computing paradigm, cloud computing has reached the market adoption acceleration point.

2 Individual consumers refer to people who obtain services for their personal use as contrasted with enterprise consumers such as businesses, organizations, or government agencies.
3 Gartner, Hype Cycle for Emerging Technologies, 2011.
7 1 Exabyte = 10^16 Terabyte (TB) = 10^9 Gigabyte (GB) 10^12 Megabyte (MB) = 10^18 bytes
8 Microsoft, blogs.technet.com, July 5, 2011.
9 Google, googleblog.blogspot.com, June 2012.
10 Venturebeat, Gmail finally blows past Hotmail to become the world’s largest email service, venturebeat.com, June 28, 2012.
12 Enterprise Gmail comprises other collaboration services in addition to email
13 Google, googleenterprise.blogspot.com, June 2011.
14 Rutrell Yasin, “All of GSA’s e-mail now in Google cloud”, Washington Technology, July 26, 2011.
15 ServiceNow, ServiceNow Recognized on 2011 Inc. 500 List With Three-Year Sales Increase of 1,240 Percent, www.servicenow.com, September 1, 2011.
18 TechTarget, Top 10 cloud computing providers of 2012, techtarget.com, April 24, 2012.
23 Amazon, Global Infrastructure, aws.amazon.com.
27 Amazon, Amazon Web Services, aws.amazon.com.
28 The Gartner number is based on raw configured disk space and does not account for RAID (Redundant Arrays of Inexpensive Disks) storage which reduces usable disk space by 20-50% and other factors that reduce the amount of end user data storage. In house storage also must be larger than the amount of data stored to allow for growth. Recognizing that CSPs allocate 2 (or 3) TB of storage for each TB of data, and assuming 70% of raw configured disk space is used, the cost comparison is $1000/TB/year CSP cost to (2/0.7) x $4876 = $13,931/TB/year for enterprise managed storage.
34 It is instructive to note that installing energy-efficient information technology equipment and implementing virtualization reduce data center energy consumption but do not impact PUE.
39 Matt Stansberry, *op. cit.*
44 Amazon, *Summary of the Amazon EC2 and Amazon RDS Service Disruption in the US East Region*, April 29, 2011.