Cloud Computer Management
From The Small Business Perspective
Yvette Ghormley, Walden University, USA

ABSTRACT

Cloud computing is a relatively new concept that is attracting much attention from both small businesses and CEOs of large corporations. However, for the novitiate, the decision to implement a cloud computing service and then how to manage it is not straightforward because there is no universal road map. In this review several issues are discussed that should have a bearing on how to bring cloud computing into a business organization and how it should be managed. These issues include, the elasticity performance of cloud computing, the lack of standards in cloud computing, whether cloud computing should be sole sourced, security issues, legal and regulatory compliance of data that may cross national borders, and what the function of the IT department should be in this venture.

Keywords: Cloud Computer Management; Small Business

INTRODUCTION

Cloud computing, the concept of a virtualized technology in which a pool of virtual resources are placed at the disposal of a client, including networks, storage devices, applications, and services, has been a reality for several years. Two of the great advantages of cloud computing are that it offers a pay-as-you-go model in which one pays only for the resources and time used and that those resources can be scaled vertically and horizontally to match demand, which is particularly important for web-based applications in which demand is very unpredictable and can vary over orders of magnitude over time or based on events. However, this is not a “plug and play system,” and the idea of using cloud computing demands considerable up-front acumen to decide how it should be managed in a business operation.

There are three basic service variants to choose from according to NIST (National Institute of Standards and Technology), although lines are blurring and some argue that these definitions are too nebulous to be of much use (Ghormley, 2012). Software as a service (SaaS) aims to provide applications to the user via an appropriate client interface in which the user has no control regarding execution except perhaps user-specific configuration settings. Platform as a service (PaaS) involves the ability to utilize applications developed by the user with tools, programming, and languages supplied by the cloud vendor although the cloud infrastructure and any of its associated operating mechanisms are not under the user’s control. Infrastructure as a service (IaaS), on the other hand, allows the user to install software, operating systems, and applications by harnessing the vendor’s networks, and storage/processing units; Thus in this instance while the consumer can manage the deployed applications, storage mechanisms, and operating systems, the cloud provider still retains control of the basic cloud infrastructure.

Shacklett (2011) suggests three major issues for the small business owner or CEO to initially reflect upon that will help define the extent of cloud computing involvement as well as what services might be useful:

1. Assess the return on investment that would be expected from a cloud service and how it would be managed.
2. Define those areas of the organization in which the company would continue to operate its own systems.
3. Determine those business policies and processes that would be impacted by the introduction of cloud computing services and what levels of training and preparation staff would need for those solutions to work well.
In this review, several issues are discussed that the CEO or business owner should be aware of that can help frame the management of cloud computing.

**ON-DEMAND PERFORMANCE - HOW ELASTIC IS IT?**

Cloud vendors always tout the elasticity of their services as a major reason for adopting cloud computing, a pay-as-you-go mechanism in which surges, whether they are cyclical, time-based, or completely unpredictable can be handled in a transparent fashion at reasonable prices without the need for business to invest in hardware that might well stand idle for long periods of time. But, how “elastic” is the performance of such systems?

The consensus appears that properly executed, elasticity is not a critical issue. For example, looking at Amazon case studies (http://aws.amazon.com/solutions/case-studies/), of which there other over a hundred, there are many stories of successful scale-up to meet peaks of much higher demand. While there continues to be a need for research into improving elasticity in response to demand—e.g., dynamic algorithms to determine how to re-partition, re-replicate, and/or re-distribute data in a cloud, as well as load balancing—other concerns that are connected to elasticity appear to have more immediate importance.

**Does Higher Elasticity in Terms of Higher Performance Mean Higher Prices in the Cloud?**

Although this question is starting to appear in the literature, there are no definitive answers as yet. However, if one looks to the provision of data via wireless technology, legacy agreements in which unlimited data and fixed prices were the norm are rapidly disappearing, in part because 4G networks and even current apps are demanding more and more bandwidth. So, inevitably, pricing agreements are moving to reflect actual data usage. In terms of elasticity, one could envisage a similar tier system based on actual performance in 5-10 years’ time.

**Is the Transfer of Data (Input/Output) the Bottleneck?**

The “fly in the ointment” with tiered pricing based on performance is that in the cloud, performance is somewhat unpredictable because disk IO (input/output) sharing is problematic. Armbrust et al (2010) compared main memory and I/O performance across 75 EC2 instances, and determined that while mean bandwidth for the former was 1,335 Mb/sec, the average disk write bandwidth was 55Mb/sec. This illustrates the problem of I/O interference between virtual machines. Data transfer bottlenecks for some applications can be just as problematic in terms of time and cost. Assuming a bandwidth of 20 Mbits/sec over a WAN (wide area network) link, transfer of 10 TB would take 45 days (Armbrust et al., 2010). Physically shipping disks overnight can improve transfer speeds but not all providers can handle data in this way, and the delay in transferring large datasets to a cloud environment for some applications might be intolerable. Jackson et al. (2010) also conducted extensive experiments in high performance computing (HPC) applications evaluating Amazon’s EC2 against more typical local HPC systems based on a set of predetermined benchmarks. Not surprisingly they demonstrated a strong correlation between the percentage of time an application spent communicating, and its overall performance on EC2. In summary, the more communication was necessary the worse the performance became. In conclusion, a significant impact on performance is possible for HPC cases in which extensive communication must occur between the cloud and the user.

**Vulnerabilities Arising from Inadvertent Access**

A recently announced vulnerability by NIST in which a client could theoretically traverse from one VM (virtual machine) client environment to another when it is being managed by a common hypervisor (Ghormley, 2012) suggests another potential issue of security linked to elasticity (Owens, 2010). While the details of the infrastructure involved (Owens, 2010) are beyond the scope of this review, the attention this disclosure received has caused considerable debate on how to solve the security issue while maintaining elasticity.
LACK OF STANDARDS: HOW DOES IT AFFECT CLOUD COMPUTING?

Most commercial cloud providers use internal data representations, meaning their cloud software infrastructure support as well as their application programming interfaces (APIs) are proprietary (Sehgal et al., 2011). This has come about because there are as yet no cloud computing standards for such processes as or elements as APIs, storage of server images for disaster recovery purposes, and ata import/export (Sultan, 2010). Concerns have arisen over this issue as it forces users to vendor lock-in as one cannot port data and programs from one cloud provider to another. Most importantly, it creates vulnerabilities for the user in terms of possible price increases, reliability problems, and the provider going out of business (Armbrust et al., 2010). These are not fictional musings: on August 8, 2008, the online storage service The Linkup shut down because it had relied on another company, Nirvanix, to store some its data and 20,000 customers lost access to their data—some permanently—for reasons that are not entirely clear (Brodkin, 2008). GNU creator and Free Software founder Richard Stallman is even more vociferous, describing cloud computing as “a trap” in which providers will increase the price of services over time thus forcing customers to pay more because their applications cannot be walked to another provider (Marston, Li, Bandyopadhyay, Zhang, & Ghalsasi, 2011).

Without more history of cloud computing, it is impossible to confirm of refute Stallman’s fears. As a result, several initiatives have started, including the Cloud Computing Interoperability Forum (www.cloudforum.org), whose mission is to promote the creation of a common cloud computing interface, the International Organization for Standardization’s Subcommittee on Distributed Application Platforms and Services (www.iso.org/iso/iso_technical_committee.html?commid=601355), which has a work study group on cloud computing standardization, and the Open Web Foundation (www.openwebfoundation.org), whose members are dedicated to promote the development of open, non-proprietary specifications for web technologies (Marston et al., 2011). Moreover, there is clear recognition of this potential pitfall by the cloud providers themselves. For example, the mission of EuroCloud (www.eurocloud.org), which was formed in 2009, and consists of dozens of cloud computing vendors in Europe, to is to take an active role in the design of cloud industry processes and standards and to build a pan-European contact and knowledge sharing network for companies that have interests in Cloud Computing, either as vendor/provider or facilitator/aggregator for the ecosystem as a whole. Moreover, in the U.S. both Amazon and Microsoft have started to develop APIs based on open source message standards, such as SOAP (Simple Object Access Protocol) and REST (Representational State Transfer) to address interoperability issues, with Amazon making its S3 cloud storage available through SOAP/REST and Microsoft’s Azure cloud supporting those standards and filing for a patent to facilitate cloud hopping (Marston et al., 2011; Sultan, 2010). Finally, Armbrust et al (2010) are optimistic that the industry will see that API standardization in which the same software infrastructure can be used in internal data centers and public clouds will lead to expansion of the cloud computing market, especially surge computing solutions and hybrid clouds.

SHOULD CLOUD COMPUTING BE SOLE SOURCED?

On April 21, 2011 part of Amazon’s EC2 (Elastic Compute Cloud) crashed in the single Availability Zone within the U.S. East Region. The data center affected was located in northern Virginia and is one of five global sites that support EC2. As a result, the online services of several other companies who depended heavily on those computing nodes, including Reddit, HootSuite, Foursquare, and Quora, ceased functioning for several hours. Considerable data for some applications was also irrevocably lost despite Amazon’s back-up systems that apparently were not as robust as they should have been. According to Amazon, the event started when a configuration change was implemented to upgrade the capacity of the primary network, with one of the first steps being the shifting of traffic from of one of the redundant routers in the primary EBS (elastic block store; “volumes”) network (Business Insider, 2011). But this was handled incorrectly, causing a portion of the EBS cluster in this Availability Zone to not have a functioning primary or secondary network. While a quick recovery should have taken place through mirroring, because the issue affected so many volumes simultaneously, the free capacity of the EBS cluster was quickly exhausted, which left many nodes in a “stuck” state, and the cascade propagated. While Amazon has since instructed customers to help themselves in the future by using multiple Availability Zones to create a more fault tolerant service, the incident left many customers with a bitter taste in their mouths, in part because of the lack of communication and the impersonal nature in which it was handled. Other cloud computing providers have also criticized Amazon’s EBS system as being too unreliable. For example, Joyent’s approach distributes data over...
several local nodes to keep it safe and accessible, and stores less frequently used resources more distantly in the cloud (Naone, 2011).

These outage issues, however, are just a new version of an old problem. In the 1980s, when the movement toward sole sourcing in supply chains started in an effort to improve quality, efficiency, and ultimately costs, there were many debates about the wisdom of sole sourcing. Much research on the subject has since taken place. For example, Chiang and Benton (1994) showed that in the case of sole versus dual sourcing except for cases in which ordering cost was high, the lead-time variability or the customer service level low, dual sourcing outperformed sole sourcing under normally distributed demand and shifted-exponential lead times. In their final recommendations, they suggested that material managers should consider splitting purchase orders when two equally qualified suppliers were available. In commenting on W. H. Deming’s Point Four, that sole sourcing is more profitable than competitive sourcing, Richardson and Roumasset (1995) found that in modeling the tradeoff between the costs to set up and coordinate with suppliers and the incentive for performance provided by competition, whether one should sole source depends on parameters such as profit sensitivity to supplier performance. In other words, a prospective cloud computing business owner or CEO should develop a model of their business in which the following might be determined: (a) likelihood of cloud computing outages, and hours of unavailability; (b) likelihood of data loss; (c) economic losses resulting from (a) and (b); and (d) other less tangible costs that might be involved in (a) and (b). These exercises could be stratified with regard to risk. As an example, if one is providing an online service that does not have any major ramifications if it goes down for a few hours, then that application might be sole sourced. However, a healthcare application that involves very confidential information and which is very time sensitive might need at least two cloud computing providers with service-level agreements in place. Armbrust et al. (2010) use the analogy of large Internet service providers using multiple network providers so that failure of a single company does not cripple such services to apply to cloud computing.

SECURITY ISSUES

Results from a recent survey conducted by the Ponemon Institute of 103 cloud service providers in the U.S. and 24 in six European countries ought to give any small business CEO pause before flying into the wild blue yonder of cloud computing (Ponemon Institute, 2011). First, and perhaps foremost, providers do not see any competitive advantage in security, so they do not see security as their responsibility, but rather that of their customers. Second, providers do not believe that their products or services substantially protect and secure the confidential or sensitive nature of customers’ information. This is borne out by the finding that 10 percent or less of operational resources are devoted to security issues and that the majority of cloud providers do not have dedicated security personnel to oversee the security of cloud applications, infrastructure or platforms. The exception might be those providers of private clouds who have a higher commitment to meeting security objectives. In essence, because providers believe that customers primarily migrate to cloud computing systems due to lower costs and faster deployment of applications, they do not see security as playing a role. When data are added from a survey of 642 U.S. and 283 European users conducted a year earlier (Ponemon Institute, 2010), while 69% of cloud providers think that security is the responsibility of cloud users, only 35% of cloud users think they should be saddled with that task. Interestingly, more than twice as many cloud users than cloud providers think that this responsibility ought to be shared (Ponemon Institute, 2011). These disparate beliefs are also apparent in regard to whether certain types of information should be placed in clouds, especially intellectual property, health and financial information, in which cloud users consistently assign a much higher risk than their cloud provider counterparts. Clearly, for any business manager, who provides security must be decided upon before and not after entering the cloud computing arena and a protocol implemented beforehand.

Despite confusion about whose responsibility security should be, Armbrust et al. (2010) state very clearly that the user is responsible for application-level security, while the cloud provider is responsible for physical security, and probably enforcing firewall policies. Responsibility for intermediate layers of the software stack, often termed the middleware (Ghormley, 2012), is a shared function with regard to security, and currently varies according to the provider, with Amazon EC2 users have much more responsibility than AppEngine customers (Armbrust et al., 2010)
According to the Cloud Security Alliance, the top seven threats to cloud computing are: (a) abuse and nefarious use, which includes hackers finding a way to upload malware and using the power of the cloud to spread it; (b) insecure application programming interfaces (APIs), which are those interfaces used by clients to interact with cloud services; (c) malicious insiders, a threat common to all systems that use computers; (d) shared technology vulnerabilities, an example of which for IaaS providers is one client encroaching on another’s “territory”; (e) data loss or leakage, which can occur through accidental deletion of data without a back-up, or loss of encoding keys or unauthorized access; (f) account, service, and traffic hijacking, such as denial-of-service attacks, phishing, or spam campaigns; and (g) an unknown risk profile, which involves such items as code updates, security practices, vulnerability profiles, and intrusion attempts (Cloud Security Alliance, 2010).

Other studies that have analyzed cloud computing security have emphasized the hardware and hypervisors, “Chinese walls” between users, Internet reliability and availability, legal and regulatory issues (see next section), problems with a perimeter security model in which the cloud is outside of the enterprise network, and the need for integration of provider and customer security systems (Chow et al., 2009; Hanna, 2009). Mather, Kumaraswamy, and Latif (2009) highlight infrastructure security, data security and storage, identity and access management, security management (primarily a lack of enterprise-grade access management features), privacy (both user trust and legal/regulatory compliance), audit and compliance (strong internal control monitoring and a robust external audit process). These authors also note that security as a service (SaaS) will likely be a growth industry in cloud computing.

Srinivasamurthy and Liu (2010) review the above quoted sources in detail, as well as several solutions, including the Mirage Management System (Wei, Zhang, Ammons, Bala, & Ning, 2009), the client-based privacy manager (Mowbray & Pearson, 2009), the transparent client protection system (Lombardi & Di Pietro, 2010), as well as methods to ensure secure and efficient access to outsourced data (Wang, Li, Owens, & Bhargava, 2009).

In the final analysis, many analysts think that cloud computing security fears are overblown. The major reason for this point of view is that security is often handled poorly at many companies, especially when it comes to enforcement, deploying security patches, and education of users (Sultan, 2011; and references cited therein). In the case of especially sensitive or high-value data, the use of user-level encryption and auditability added as a layer beyond the reach of the virtualized guest operating system may be necessary however to guarantee further protection (Armbrust et al., 2010).

WHOSE DATA IS IT ANYWAY? LEGAL AND REGULATORY COMPLIANCE

Cloud computing adds a further dimension to the problem of “It’s 11 pm. Do you know where your data are?” Not only is the data likely to be distributed between one or more data centers within a country, but increasingly the data cross national boundaries. Part of the problem is that most cloud provider service level agreements (SLAs) for small and medium enterprises are simple ones in which the user has little or no room for negotiation (Gilbert, 2011). The cloud provider wants full control over where the data are located and the tools to use in processing the data, as well as the ability to change pricing, terminate, or suspend service. Often the SLA will refer to the terms “data controller” and “data processor,” which are referents to the 95/46/EC Directive of the European Commission. By identifying itself as a data processor, the cloud computing service provider essentially shields itself from any liability under the directive while shirting responsibility, compliance, and reporting requirements to the user. However, opinions within the Article 29 Working Party are starting to change, and it is likely that for some forms of cloud computing the provider will be the data controller (Gilbert, 2011).

There is no doubt that some kinds of data may need to reside within site, particularly to meet HIPAA (Health Insurance Portability and Accountability Act), PCI (Payment Card Industry), or Sarbanes-Oxley regulations or for reasons of auditability (Armbrust et al., 2010; McKendrick, 2011). Moreover, when data is stored in another country other than its owner, it is not always clear which country’s privacy laws should be followed by the cloud provider (Marston et al., 2011). Although the development of the U.S.-E.U. Safe Harbor laws has made some inroads to address such problems as these, the issues are far from decided. Nevertheless, the overt histrionics with regard to these concerns that one sees occasionally in the popular press and political arena are really unnecessary. In essence as long as the customer takes the role of being the data controller, and the cloud provider the role of data processor,
provided all relevant questions have been asked and answered satisfactorily in legal and service terms, the question of where data reside should not be a barrier to pursuing cloud applications (Determann, 2011; Gilbert, 2011).

CORE COMPETENCIES: WHAT SHOULD THE IT DEPARTMENT BE DOING?

For a small-to-medium-sized business—defined by the European Commission as less than 250 employees and annual sales of less than 50 million euros (about $65 million at today’s exchange rates) (European Commission, 2005)—that is committed to cloud computing, how should the transformation be brought about and what should the role of their IT departments be?

The first consideration is to recognize that IT personnel may not be thrilled about the idea of adopting cloud computing, recognizing that much of a company’s operations may no longer reside behind the IT’s department’s firewall and that reductions in personnel might be a strong perception (Sultan, 2011). This is essentially a cultural paradigm shift (McKendrick, 2011). Moreover, there may be further resistance within larger companies who have invested much time and money in an enterprise resource planning system despite the recognition that in a world of globalization of both markets and suppliers there is no longer the time or resources to create an internal IT solution for every external business situation or supplier that needs to be brought into the supply chain fold (Shacklett, 2011).

The second consideration is to acknowledge that most IT departments spend 70% to 80% of their budgets just trying to keep existing systems running (Rettig, 2007), which translates to 80% of the staff’s time (Payton, 2010). In a multiyear study of over 400 companies conducted by MIT researchers (Ross, Weill, & Robertson, 2006) it was found that IT departments are not usually viewed as innovative units by business leaders but rather liabilities and areas of heavy cost. Implementation of strategic initiatives can take years and moreover each initiative has to be integrated with existing systems, causing the researchers to comment “Legacy systems cobbled together to respond to each new business initiative create rigidity and excessive costs. Every change becomes a risky, expensive venture” (Ross et al., 2006, p11).

Although outsourcing certain activities within an IT department to cloud computing is potentially attractive in terms of cutting costs and shifting capital expenses to operational expenses—essentially a shift from fixed costs to a subscription-based model (Goodburn & Hill, 2010)—making the transition can be a challenge. It may, for example, involve writing off legacy systems, as well as retraining IT staff to manage cloud-based applications and later the whole organization in how to use them. However, reducing the “maintenance activities” of the IT department by a substantial proportion should allow staff to better focus on how to respond to an organization’s demands in a more agile fashion, which is both a return to core competencies and an improved response in internal customer-supplier relationships.

While core competencies will vary according to the nature of the business the IT department’s role is one of response to business requirements and initiatives—to translate those requirements into practical applications while minimizing cost and maximizing efficiency and at the same time ensuring the systems through which solutions are implemented are compatible with each other. Particular cloud computing services in some instances may be one potential solution but the IT staff should perform due diligence to determine whether it is in fact a rational solution (McKendrick, 2011).

CONCLUSIONS

There is no doubt that implementing cloud computing from a business perspective can be a risky strategy in terms of financial, security, and performance factors. No doubt many of the same concerns faced CEOs during the 1990s when contemplating the acquisition of enterprise software, the implementation of which from the perspective of many employees was not a happy experience. Second-generation enterprise software seems more appealing and is more agile and flexible. Cloud computing promises much along the same lines but is still in its infancy. Nevertheless, the collective experience of many companies who are using it now suggests that it is here to stay and grow despite the challenges of cloud computing management.
AUTHOR INFORMATION

Dr. Yvette Ghormley holds several degrees to include a Ph.D. in Organization and Management with a concentration in E-Commerce, MA in Education and Human Development with a concentration in Education, Technology and Leadership and a MA in Information Technology, Information Security. In addition, she holds graduate certificates in Acquisition and Contract Management and Project Management. Dr. Ghormley has presented at numerous conferences and has published in International journals, texts and handbooks. Dr. Ghormley is well known for using technological initiatives in business and entrepreneurship for non-profit and for-profit enterprises. E-mail: yvette.ghormley@cox.net

REFERENCES


31. Van Gelder, K.
