Cloud computing — The business perspective

Sean Marston a, Zhi Li a, Subhajyoti Bandyopadhyay a,⁎, Juheng Zhang a, Anand Ghalsasi b

a University of Florida, FL, United States
b Persistent Systems, Inc., CA, United States

A B S T R A C T

The evolution of cloud computing over the past few years is potentially one of the major advances in the history of computing. However, if cloud computing is to achieve its potential, there needs to be a clear understanding of the various issues involved, both from the perspectives of the providers and the consumers of the technology. While a lot of research is currently taking place in the technology itself, there is an equally urgent need for understanding the business-related issues surrounding cloud computing. In this article, we identify the strengths, weaknesses, opportunities and threats for the cloud computing industry. We then identify the various issues that will affect the different stakeholders of cloud computing. We also issue a set of recommendations for the practitioners who will provide and manage this technology. For IS researchers, we outline the different areas of research that need attention so that we are in a position to advice the industry in the years to come. Finally, we outline some of the key issues facing governmental agencies who, due to the unique nature of the technology, will have to become intimately involved in the regulation of cloud computing.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

The emergence of the phenomenon commonly known as cloud computing represents a fundamental change in the way information technology (IT) services are invented, developed, deployed, scaled, updated, maintained and paid for. Computing as we know today reflects a paradox — on one hand, computers continue to become exponentially more powerful and the per-unit cost of computing continues to fall rapidly, so much so that computing power per se is nowadays considered to be largely a commodity [24,32]. On the other hand, as computing becomes more pervasive within the organization, the increasing complexity of managing the whole infrastructure of disparate information architectures and distributed data and software has made computing more expensive than ever before to an organization [42]. The promise of cloud computing is to deliver all the functionality of existing information technology services (and in fact enable new functionalities that are hitherto infeasible) even as it dramatically reduces the upfront costs of computing that deter many organizations from deploying many cutting-edge IT services [46]. All such promise has led to lofty expectations — Gartner Research expects cloud computing to be a $150 billion business by 2014, and according to AMI partners, small and medium businesses are expected to spend over $100 billion on cloud computing by 2014.

The impetus for change right now is seen predominantly from a costs perspective (even though, as we discuss later in the document, the promises from a technological functionality perspective are equally attractive), as organizations increasingly discover that their substantial capital investments in information technology are often grossly underutilized. One recent survey of six corporate data centers found that most of the servers were using just 10–30% of their available computing power, while desktop computers have an average capacity utilization of less than 5% [51]. Equally pertinent are the maintenance and service costs that have proved to be a drain on scarce corporate resources. A recently conducted survey by Gartner Research indicated that about two-thirds of the average corporate IT staffing budget goes towards routine support and maintenance activities [21], which does seem anachronistic in an age of globalized
and cutthroat competition — as the CEO of a cloud platform provider commented recently, “If you woke up this morning and read in The Wall Street Journal that, say, Overstock.com has stopped using UPS and FedEx and the U.S. mail, and had bought fleets of trucks and started leasing airport hubs and delivering products themselves, you would say they were out of their minds. Why is that much more insane than a health care company spending $2 billion a year on information technology?” [30].

Cloud computing represents a convergence of two major trends in information technology — (a) IT efficiency, whereby the power of modern computers is utilized more efficiently through highly scalable hardware and software resources and (b) business agility, whereby IT can be used as a competitive tool through rapid deployment, parallel batch processing, use of compute-intensive business analytics and mobile interactive applications that respond in real time to user requirements [29]. The concept of IT efficiency also embraces the ideas encapsulated in green computing, since not only are the computing resources used more efficiently, but further, the computers can be physically located in geographical areas that have access to cheap electricity while their computing power can be accessed long distances away over the Internet. However, as the term business agility implies, cloud computing is not just about cheap computing — it is also about businesses being able to use computational tools that can be deployed and scaled rapidly, even as it reduces the need for huge upfront investments that characterize enterprise IT setups today.

There are perhaps as many definitions as there are commentators on the subject (including one by the National Institute of Standards and Technology (NIST) that is nearly 800 words long [38]), but none of them seem to identify all the key characteristics of cloud computing. In coming up with our definition, we tried to encapsulate the key benefits of cloud computing from a business perspective as well as its unique features from a technological perspective. Our formal definition of cloud computing is as follows: “It is an information technology service model where computing services (both hardware and software) are delivered on-demand to customers over a network in a self-service fashion, independent of device and location. The resources required to provide the requisite quality-of-service levels are shared, dynamically scalable, rapidly provisioned, virtualized and released with minimal service provider interaction. Users pay for the service as an operating expense without incurring any significant initial capital expenditure, with the cloud services employing a metering system that divides the computing resource in appropriate blocks.” Fig. 1 shows a schematic of the cloud computing model. It shows how the computing resources in the cloud can be accessed from a variety of platforms through the Internet.

We note that our definition does not explicitly require that the services be provided by a third-party, but emphasizes more on the aspects of (1) resource utilization, (2) virtualized physical resources, (3) architecture abstraction, (4) dynamic scalability of resources, (5) elastic and automated self-provisioning of resources, (6) ubiquity (i.e. device and location independence) and (7) the operational expense model. Cloud computing can be provisioned using an organization’s own servers, or it can be rented from a cloud provider that takes all the capital risk of owning the infrastructure.

We have several objectives in this article. In the first part (Sections 2 and 3 that follow), we give an overview of the core concepts of cloud computing and its key advantages. The second part of the article starts with Section 4, where we make a case for the need for a roadmap for IS professionals and IS researchers in understanding and evaluating cloud computing. In Section 5, we analyze the strategic imperatives of the cloud computing industry as a whole. The analysis is presented in a SWOT framework, so that we understand both the opportunities and challenges to the fledgling industry. Section 6 introduces the various stakeholders in cloud computing and discusses the relevant issues that they will need to consider. Section 7 discusses the nature as well as the role of the regulatory bodies that oversee cloud computing. Section 8 then assimilates the information from the prior section to come out with two sets of recommendations — one for the practitioners and the other for the researchers in IS in the cloud computing area. The final section concludes with a call to arms to the IS community to start the many relevant discussions on the “business” side of cloud computing.

2. The key advantages of cloud computing

Many of the incipient ideas in cloud computing are not exactly new (in fact, it was as far back as 1965 that Western Union dreamt up the future role of the company as a nationwide “information utility” as part of the company’s strategic plans [26]), which have led several observers such as Oracle’s CEO Larry Ellison to declare the whole concept as a product of hype [7]. While it is true that several of the above ideas were indeed present for a long time, we nonetheless argue that their confluence today in an environment where information can be accessed independent of device and location

---

1 An oft-cited example is that of The New York Times, which used 100 Amazon EC2 instances and a Hadoop application to process 4 TB of raw image TIFF data (stored on Amazon S3 servers) into 11 million finished PDF documents in the space of 24 h at a computational cost of about $240 (excluding the sunk bandwidth costs).

2 Interestingly, Oracle seems to have bought into the hype as it has created its own “Application Grid” system, which allows Oracle databases to be deployed into the Amazon EC2 cloud.
represents a major shift in computing as we know it. Specifically, cloud computing offers the following key advantages:

1. It dramatically lowers the cost of entry for smaller firms trying to benefit from compute-intensive business analytics that were hitherto available only to the largest of corporations. These computational exercises typically involve large amounts of computing power for relatively short amounts of time, and cloud computing makes such dynamic provisioning of resources possible. Cloud computing also represents a huge opportunity to many third-world countries that have been so far left behind in the IT revolution — as we discuss later, some cloud computing providers are using the advantages of a cloud platform to enable IT services in countries that would have traditionally lacked the resources for widespread deployment of IT services.

2. It can provide an almost immediate access to hardware resources, with no upfront capital investments for users, leading to a faster time to market in many businesses. Treating IT as an operational expense (in industry-speak, employing an ‘Op-ex’ as opposed to a ‘Cap-ex’ model) also helps in dramatically reducing the upfront costs in corporate computing. For example, many of the promising new Internet startups like 37 Signals, Jungle Disk, Gigavox, SmugMug and others were realized with investments in information technology that are orders of magnitude lesser than that required just a few years ago. The cloud becomes an adaptive infrastructure that can be shared by different end users, each of whom might use it in very different ways. The users are completely separated from each other, and the flexibility of the infrastructure allows for computing loads to be balanced on the fly as more users join the system (the process of setting up the infrastructure has become so standardized that adding computing capacity has become almost as simple as adding building blocks to an existing grid). The beauty of the arrangement is that as the number of users goes up, the demand load on the system gets more balanced in a stochastic sense, even as its economies of scale expand.

3. Cloud computing can lower IT barriers to innovation, as can be witnessed from the many promising startups, from the ubiquitous online applications such as Facebook and YouTube to the more focused applications like Triplt (for managing one’s travel) or Mint (for managing one’s personal finances).

4. Cloud computing makes it easier for enterprises to scale their services — which are increasingly reliant on accurate information — according to client demand. Since the computing resources are managed through software, they can be deployed very fast as new requirements arise. In fact, the goal of cloud computing is to make resources up or down dynamically through software APIs depending on client load with minimal service provider interaction [19].

5. Cloud computing also makes possible new classes of applications and services that were not possible before. Examples include (a) mobile interactive applications that are location-, environment- and context-aware and that respond in real time to information provided by human users, nonhuman sensors (e.g. humidity and stress sensors within a shipping container) or even from independent information services (e.g. worldwide weather data); (b) parallel batch processing, that allows users to take advantage of huge amounts of processing power to analyze terabytes of data for relatively small periods of time, while programming abstractions like Google’s MapReduce or its open-source counterpart Hadoop makes the complex process of parallel execution of an application over hundreds of servers transparent to programmers; (c) business analytics that can use the vast amount of computer resources to understand customers, buying habits, supply chains and so on from voluminous amounts of data; and (d) extensions of compute-intensive desktop applications that can offload the data crunching to the cloud leaving only the rendering of the processed data at the front-end, with the availability of network bandwidth reducing the latency involved.

3. Core technological concepts and terminology

While the evolution of cloud computing will take several years or even a decade to fully unfold, the three core technologies that will enable it — virtualization, multitenancy and Web services — are rapidly taking shape.

Virtualization is the technology that hides the physical characteristics of a computing platform from the users, instead presenting an abstract, emulated computing platform [52]. This emulated computing platform for all practical purposes behaves like an independent system, but unlike a physical system, can be configured on demand, and maintained and replicated very easily. The computing infrastructure is much better utilized, leading to lower upfront and operational costs (one side benefit of virtualization is the savings in real estate for the data centers). While the concept of virtualization has been prevalent since the 1960s, it is only in the recent past that computing power and networking resources have caught up to deliver the level of seamless performance within an emulated system that users have grown accustomed to on personal computers.

A related concept is that of multitenancy, whereby a single instance of an application software serves multiple clients. This allows better utilization of a system’s resources (in terms of memory and processing overhead), the requirements of which could otherwise be considerable if the software instance had to be duplicated for each individual client.

A Web service is defined by the W3C as “a software system designed to support interoperable machine-to-machine interaction over a network” [53]. The definition encompasses many different systems, but in common usage the term refers to clients and servers that communicate over the HTTP protocol used on the Web. Web services help standardize the interfaces between applications, making it easier for a software client (e.g. a web browser) to access server applications over a network.

From an end-user’s perspective, the cloud computing industry often speaks about different delivery models of cloud computing, all of which refer to the different layers of the cloud computing architecture. The most commonly heard term perhaps is Software as a Service or SaaS, in which the application runs on the cloud, eliminating the need to install and run the application on the client computer. Examples of SaaS include enterprise-level applications such as Salesforce, Netsuite or Google Apps to personal applications such as GMail, TurboTax Online, Facebook, or Twitter. A Platform as a Service, or PaaS, facilitates the development and deployment of applications without the cost and complexity of buying and managing the underlying hardware and software layers. Examples of PaaS include Microsoft’s Azure Services Platform, Salesforce’s Force.com, Google App Engine, Amazon’s Relational Database Services and Rackspace Cloud Sites. The third ‘model’ of cloud computing is Infrastructure as a Service or IaaS, whereby storage and compute capabilities are offered as a service. Amazon’s S3 storage service and EC2 computing platform, Rackspace Cloud Servers, Joyent and Terremark are some prominent examples of IaaS.

Table 1 lists some of the key players currently in the cloud computing arena. It includes established players such as IBM or Microsoft, as well as small but promising upstarts such as the Irish cloud computing provider Vordel. We also summarize some of their main contributions (in terms of products, services or innovations) to
Table 1

<table>
<thead>
<tr>
<th>Name</th>
<th>Contribution</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>The “established” players</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IBM</td>
<td>IBM provides cloud computing services called Blue Cloud, which offers companies access to tools that allow them to manage large scale applications and database via IBM’s Cloud. The company offers consulting services to help companies integrate their infrastructure into the cloud. Recent partnership with Google to work with several universities in order to promote new software development methods which will help students and researchers address the challenges of the cloud applications of the future.</td>
<td>As the leading provider IT consulting and infrastructure, IBM “maters”. It helps that the company has adopted an aggressive cloud computing strategy: it is spending over $400 million on two cloud computing facilities and is also expanding the number of researchers in the area of cloud computing. Thanks to its consulting and services division, IBM can push the adoption of cloud computing, especially in the corporate arena.</td>
</tr>
<tr>
<td>Google</td>
<td>Google’s App Engine offers client organizations access to Google’s cloud-based platform that provide tools to build and host web applications. Its premier SaaS offering is Google Apps, a set of online office productivity tools including e-mail, calendaring, word processing and a simple Web site creation tool. Its recent acquisition of Postini, which offers a set of e-mail and Web security services, makes it a credible player in the area of electronic corporate communications.</td>
<td>A commanding online presence and a successful business model makes Google a company that is automatically paid heed to in the online world. Just by its sheer presence, Google Apps is hastening the industry’s move from packaged software to Web-hosted services, and App Engine provides a credible alternative in the platform-as-a-service market. Google’s financial clout makes it a player that can “dig in” for the long haul.</td>
</tr>
<tr>
<td>Microsoft</td>
<td>The company has slatted Windows Azure, the “cloud operating system” Paas to appear in early 2010. Additionally they are creating the Azure Services Platform to run on the Windows Azure operating systems giving client organizations access to several online Microsoft services like Live, .Net, SQL, SharePoint, and Microsoft’s Dynamic CRM.</td>
<td>The ubiquitous presence of the software behemoth in the current computing market give them the credibility to be the complete infrastructure provider to both the SaaS and IaaS markets.</td>
</tr>
<tr>
<td>AT&amp;T</td>
<td>AT&amp;T provides two cloud services: Synaptic Hosting, through which client companies will be able to store Windows serve, Linux client server applications and web applications on AT&amp;T’s cloud; and Synaptic Storage, enabling clients to store their data on AT&amp;T’s cloud.</td>
<td>AT&amp;T provides one key component of the requisite infrastructure — the network backbone, and has the necessary experience in billing for it (i.e. they have an established revenue model). Adding data services just helps augment its competitive position. A possible competitor is Verizon, which is ramping up its cloud offerings.</td>
</tr>
</tbody>
</table>

**Key technology providers**

<table>
<thead>
<tr>
<th>Name</th>
<th>Contribution</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apache</td>
<td>Apache’s Hadoop is an open-source software framework that has inspired the development of database and programming tools for cloud computing (Apache Hadoop has been discussed further in the main text). Recently started offering specialized storage solutions for cloud applications. Recently introduced their vCloud initiative, which allows client organizations to run their in-house applications on a cloud and be interoperable with other cloud services from other providers that run within the vCloud ecosystem.</td>
<td>Powerful hackers like IBM, Facebook, Yahoo! and others make Hadoop a formidable development ecosystem for cloud computing applications.</td>
</tr>
<tr>
<td>EMC</td>
<td>Provides two key components in cloud computing — storage and virtualization software (thanks to its recent acquisition of VMware). Recently started offering specialized storage solutions for cloud applications. Recently introduced their vCloud initiative, which allows client organizations to run their in-house applications on a cloud and be interoperable with other cloud services from other providers that run within the vCloud ecosystem.</td>
<td>EMC provides two key and complementary products for building “private clouds” for enterprises — storage solutions and virtualization software. It is also directing a lot of resources towards its cloud computing initiatives. The vCloud initiative has many hosting and cloud computing vendors on the common VMWare platform, thereby essentially locking their clients into using their technology.</td>
</tr>
<tr>
<td>Cisco</td>
<td>A relatively late entrant in the cloud computing space, Cisco is actively working on a set of standards that will allow portability across providers. One crucial aspect of that work is ensuring workload portability from one autonomous system to another, which includes the consistent execution of the workload on the new system (i.e. the execution of the complete IT policy associated with that workload). As the networking infrastructure provider behemoth, it makes sense for Cisco to work on cloud computing standards, since they can afford to be completely agnostic of the cloud provider. Their recent forays into the $55 billion server market give them the credibility to be the complete infrastructure provider to both the SaaS and IaaS markets.</td>
<td></td>
</tr>
</tbody>
</table>

**The “innovators”**

<table>
<thead>
<tr>
<th>Name</th>
<th>Contribution</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amazon</td>
<td>Offers its Amazon Web Services, a suite of services which include the Elastic Compute Cloud (EC2), for computing capacity, and the Simple Storage Service (S3), for on-demand storage capacity. In addition to these core offerings, Amazon offers the SimpleDB (a database Web service), the CloudFront (a Web service for content delivery) and the Simple Queue Service (a hosted service for storing messages as they travel between nodes). Like IBM and Google, Amazon is working with universities, by giving access to their large databases and their engineers teaching classes on web-scale development.</td>
<td>One of the pioneers in cloud computing, and one of the first to offer pay-as-you-go access to virtual servers and data storage space. These days it is one of those companies that are most associated with cloud computing. CEO Jeff Bezos has his share of Wall Street critics, but is arguably one of the few visionaries in the online world. Managing one of the largest online operations gives it the ideal platform to test its own technological advancements in the area before selling them to client organizations. An upcoming competitor is Rackspace, which provides a similar suite of IaaS products, and has a long history of offering hosted data center services.</td>
</tr>
<tr>
<td>SalesForce.com</td>
<td>SalesForce.com is the first well-known and successful SaaS application. Riding on its coattails, the company has now introduced Force.com, an integrated set of tools and application services that independent software vendors and corporate IT departments can use to build any business application and run it on the same infrastructure that delivers the Salesforce CRM applications. More than 100,000 business applications already run on the Force.com platform. It includes the company’s Apex programming language.</td>
<td>Providing the requisite proof of concept from its successful CRM application, Force.com is influencing organizations — both vendors and corporate customers — in deciding in favor of putting more resources behind their cloud computing efforts, and providing the necessary tools to implement their ideas. A promising competitor is Netsuite, which offers a more ‘complete’ business software suite including ERP, CRM, accounting and e-commerce tools.</td>
</tr>
<tr>
<td>Enomaly</td>
<td>Enomaly’s Elastic Computing Platform (ECP) integrates enterprise data centers with commercial cloud computing offerings, letting IT professionals manage and govern both internal and external resources from a single console, while making it easy to move virtual machines from one data center to another. Enomaly’s software could prove crucial as enterprises grapple with the problem of managing a wide array of computing resources on and off the cloud. Enomaly has received financial backing from Intel, which gives it a shot at long-term viability.</td>
<td></td>
</tr>
</tbody>
</table>

**The “enablers”**

<table>
<thead>
<tr>
<th>Name</th>
<th>Contribution</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>CapGemini</td>
<td>Capgemini is the first major professional services firm to pursue a partnership on Google Apps Premier Edition (GAPE) for enterprises. It uses Google’s software as a service initiative to target some among large enterprises. Its GAPE service offerings reside within its well-established and mature Desktop Outsourcing Services practice. Capgemini has a long-standing partnership with Microsoft since 1997 to help client organizations implement the latter’s solutions. It would not risk upsetting that relationship if it did not sense among its customer base a genuine interest in Google’s enterprise initiatives. We feel that this also represents the start of a trend whereby SaaS providers begin relationships with existing integrators (or develop new ones) to develop a professional services ecosystem that can appeal to larger companies that usually have more complicated requirements.</td>
<td></td>
</tr>
</tbody>
</table>
this nascent industry and then present our opinion of their roles in shaping the industry in the next few years. We have purposefully kept the list short, but have tried to be eclectic at the same time. Our aim is to highlight some organizations whose products or ideas may shape the cloud computing arena over the next few years, even if some of the companies themselves – as it inevitably happens in new industries – do not survive.

We end this section with a brief overview of the different cloud deployment models within organizations. A public cloud is characterized as being available from a third party service provider via the Internet, and is a cost-effective way to deploy IT solutions, especially for small or medium sized businesses. Google Apps is a prominent example of a public cloud that is used by many organizations of all sizes. A private cloud offers many of the benefits of a public cloud computing environment, such as being elastic and service based, but is managed within an organization. Private clouds provide greater control over the cloud infrastructure, and are often suitable for larger installations. A private cloud can actually be handled by a third-party provider, e.g. the upcoming Government Cloud product from Google that will be certified under the Federal Information Security Management Act (FISMA) to store both applications and data of companies themselves – as it inevitably happens in new industries – do not survive.

We end this section with a brief overview of the different cloud deployment models within organizations. A public cloud is characterized as being available from a third party service provider via the Internet, and is a cost-effective way to deploy IT solutions, especially for small or medium sized businesses. Google Apps is a prominent example of a public cloud that is used by many organizations of all sizes. A private cloud offers many of the benefits of a public cloud computing environment, such as being elastic and service based, but is managed within an organization. Private clouds provide greater control over the cloud infrastructure, and are often suitable for larger installations. A private cloud can actually be handled by a third-party provider, e.g. the upcoming Government Cloud product from Google that will be certified under the Federal Information Security Management Act (FISMA) to store both applications and data of companies themselves – as it inevitably happens in new industries – do not survive.

We end this section with a brief overview of the different cloud deployment models within organizations. A public cloud is characterized as being available from a third party service provider via the Internet, and is a cost-effective way to deploy IT solutions, especially for small or medium sized businesses. Google Apps is a prominent example of a public cloud that is used by many organizations of all sizes. A private cloud offers many of the benefits of a public cloud computing environment, such as being elastic and service based, but is managed within an organization. Private clouds provide greater control over the cloud infrastructure, and are often suitable for larger installations. A private cloud can actually be handled by a third-party provider, e.g. the upcoming Government Cloud product from Google that will be certified under the Federal Information Security Management Act (FISMA) to store both applications and data of companies themselves – as it inevitably happens in new industries – do not survive.

We end this section with a brief overview of the different cloud deployment models within organizations. A public cloud is characterized as being available from a third party service provider via the Internet, and is a cost-effective way to deploy IT solutions, especially for small or medium sized businesses. Google Apps is a prominent example of a public cloud that is used by many organizations of all sizes. A private cloud offers many of the benefits of a public cloud computing environment, such as being elastic and service based, but is managed within an organization. Private clouds provide greater control over the cloud infrastructure, and are often suitable for larger installations. A private cloud can actually be handled by a third-party provider, e.g. the upcoming Government Cloud product from Google that will be certified under the Federal Information Security Management Act (FISMA) to store both applications and data of companies themselves – as it inevitably happens in new industries – do not survive.

We end this section with a brief overview of the different cloud deployment models within organizations. A public cloud is characterized as being available from a third party service provider via the Internet, and is a cost-effective way to deploy IT solutions, especially for small or medium sized businesses. Google Apps is a prominent example of a public cloud that is used by many organizations of all sizes. A private cloud offers many of the benefits of a public cloud computing environment, such as being elastic and service based, but is managed within an organization. Private clouds provide greater control over the cloud infrastructure, and are often suitable for larger installations. A private cloud can actually be handled by a third-party provider, e.g. the upcoming Government Cloud product from Google that will be certified under the Federal Information Security Management Act (FISMA) to store both applications and data of companies themselves – as it inevitably happens in new industries – do not survive.

We end this section with a brief overview of the different cloud deployment models within organizations. A public cloud is characterized as being available from a third party service provider via the Internet, and is a cost-effective way to deploy IT solutions, especially for small or medium sized businesses. Google Apps is a prominent example of a public cloud that is used by many organizations of all sizes. A private cloud offers many of the benefits of a public cloud computing environment, such as being elastic and service based, but is managed within an organization. Private clouds provide greater control over the cloud infrastructure, and are often suitable for larger installations. A private cloud can actually be handled by a third-party provider, e.g. the upcoming Government Cloud product from Google that will be certified under the Federal Information Security Management Act (FISMA) to store both applications and data of companies themselves – as it inevitably happens in new industries – do not survive.
With so many sweeping changes over the horizon, the role of IS researchers in the new environment cannot be stressed enough. If organizations are to reap the full benefits of cloud computing, we passionately believe that it is imperative for IS researchers, as the experts and the thought leaders in the area, to be proactively involved in every discussion surrounding the technology from its very outset. As Agarwal and Lucas [1] note about the IS community, “our strength as a scholarly community derives partly from our study of the first-order, second-order and third-order effects of IT that span multiple functional areas and business processes” (p.390). With our background in the underlying technology and the associated business issues, IS researchers can bring forth a holistic perspective that has often been lacking in many technology discussions. We also note that while there is an impressive amount of literature on cloud computing in computer science, there is still a dearth of literature in the IS area that look at cloud computing. One of the goals of this paper is to start that process by presenting a starting list of the various issues at the intersection of the business and the technology involved in cloud computing.

We pursue several objectives in the remainder of this paper. First, from a practitioner’s perspective, we strategically analyze the cloud computing industry. Second, we identify the various stakeholders — whether they are the providers of cloud computing or the consumers and the regulators who have to deal with the technology. Third, and perhaps most importantly, from an IS researcher’s perspective, our aim is to bring forth the issues that are likely to be important to these stakeholders, and thereby suggest some of the research topics that we should start exploring and be in a position to advise the community in due course. In a bid to ensure that our prescription is not biased from just one perspective, this group of authors consists of members from the academia along with a senior executive from a software company that is currently developing applications for some of the largest players in the cloud computing arena. We also carried out in-depth interviews with various industry executives, in order to get both a cloud computing provider perspective as well as an enterprise user perspective.6

5. Cloud computing — a SWOT analysis

5.1. Strengths

We covered much of the key benefits earlier in Section 2, and we will therefore keep the discussion in this section restricted to ideas that were not explored there. The ability to scale up services at a very short notice obviates the need for underutilized servers in anticipation of peak demand. When an organization has unanticipated usage spikes in computing above its internally installed capacity, it has the ability to request more computing resources on the fly. Cloud computing offers organizations the ability to effectively use time-distributed computing resources. One example is that of an internet photo website Smugmug. The company has relatively stable computing workloads throughout the year; however during the months of December and January the required resources spike to five times the usual workload. Cloud computing allows the company to meet the excess requirements during the two months without incurring the costs of hosting a traditional infrastructure for the rest of the year.

In 2000, over 45% of capital equipment budget was spent on IT, however on average only 6% of the server capacity is utilized. Assuming a 3-year lifespan of a server, the infrastructure and energy costs alone exceed the purchase price of a server. Cloud computing leads to reduced infrastructure costs and energy savings as well reduced upgrades and maintenance costs. Economies of scale for datacenters cost savings can lead to a 5– to 7-time reduction in the total cost of computing [3].

One of the components of maintenance costs is the management of technology, which is potentially made much simpler by using a cloud computing service. Preset configuration of servers and virtual machines can be put in place with appropriate applications, security, and data. (With so much of personal computing moving to virtual servers or to the cloud, it is perhaps no wonder that the market for enterprise servers is expected to double by 2013 [8].) This allows for a more secure environment with the company having better control of the resources on their network. Cloud computing services allow an organization to control when, where, and how employees have access to the organization’s computer systems, all managed over a simple web-based interface (for example, Amazon Web Services (AWS) can be managed easily through the AWS Management Console). Employees like the arrangement too, since they are able to make full use of the company’s computer systems using less powerful devices such as a smartphone or a netbook.

5.2. Weaknesses

There are many issues that need to be resolved before cloud computing can be accepted as a viable choice in business computing. As pointed out in the previous section, organizations will be justifiably wary of the loss of physical control of the data that is put on the cloud. Hitherto, providers have been unable to guarantee the location of a company’s information on specified set of servers in a specified location. However, cloud computing service providers are rapidly adopting measures to handle this issue. For example, Amazon Web Services recently announced the Amazon Virtual Private Cloud that allows a business to connect its existing infrastructure to a set of isolated AWS compute resources via a VPN connection. To satisfy the European Union data regulations, AWS now allows for companies to deploy its SimpleDB structured storage physically within the EU region. The Government Cloud product from Google that we alluded to earlier is also a response to allay concerns from government entities over the location of their data.

Large organizations will also be wary of entrusting mission-critical applications to a cloud computing paradigm where providers cannot commit to the high quality of service and availability guarantees that are demanded in such environments. For example, Amazon Web Services Service Level Agreement (SLA) currently commits to an annual uptime percentage of 99.95% over the trailing 365 days, which might be enough for most small- and medium-sized organizations, but will be deemed insufficient for mission-critical applications for large organizations. Even though many in-house IT services often fail to live up to such uptime standards, such failures are not held up for media scrutiny, unlike the much-publicized failures of prominent cloud computing service providers.

5.3. Opportunities

One of the significant opportunities of cloud computing lies in its potential to help developing countries reap the benefits of information technology without the significant upfront investments that have stymied past efforts. In fact, cloud computing might do to computing in developing countries what mobile phones did to communications — allow the governments and local firms to benefit from the effective use

---

6 On the provider side, we interviewed Ken Comee, the President and CEO of Cast Iron Systems, a leading software provider today for SaaS application integration with existing enterprise solutions; Bob Chung, Director of Microsoft Alliance, Full Armor, a leading software and services company providing enterprise endpoint management on various Windows platforms; and Krishna Kumar and Randy Guthrie, Developer Evangelists at Microsoft Corporation. To get an enterprise-user perspective of cloud computing, we interviewed Mike Manis, Vice President, Global Technology Shared Services at Office Depot, one organization that has been in the vanguard of using virtualization technologies. Our knowledge and understanding also benefited from conversations with several other executives in the cloud computing industry who did not want to be named for the purposes of this article. While we benefited enormously from all these conversations, any opinion expressed in this document is entirely our responsibility.
of information technology. A recent survey by the Forrester Group indicates that SaaS is a priority for 74% of Chinese firms, with 29% planning to pilot SaaS projects in the next 12 months. In contrast, the survey found that a majority of European or American firms is interested but have no plans to pursue SaaS [47]. An impressive example of the power of cloud computing in developing countries comes from Ethiopia, where the government has commissioned the cloud computing provider FullArmor to remotely manage 250,000 laptops with teachers throughout the country. The laptops will contain sensitive teacher and student data, and information like syllabi and class material will be managed centrally. In order to prevent security breaches, if a laptop drifts outside a virtual “fence”, its contents can be remotely ‘wiped’ (i.e. made unusable) through cloud-based interfaces.

Much like developing countries, small businesses represent another huge opportunity for cloud computing. All of a sudden, small businesses can exploit high-end applications like ERP software or business analytics that were hitherto unavailable to them. While it can be argued that some of the more involved features of such applications might not be available on their cloud-based counterparts, such omissions will matter very little for their intended customers [49].

Mashups represent another opportunity in cloud computing. In web development, a mashup is a web page or application that combines data or functionality from two or more external sources to create a new service in originally unintended ways. An example of a mashup is the use of cartographic data to add location information to real estate data, thereby creating a new and distinct Web service that was not originally provided by either source. The new type of mashup that we are beginning to see combines different cloud computing services and integrates them into a single service or application. Amazon’s GrepTheWeb is a good example for cloud computing service compositions within the domain of a single provider.

In an age where businesses are looking to burnish their ‘green’ credentials, cloud computing appeals to large IT infrastructures that want to reduce their carbon footprint. According to a Forrester survey, over 41 percent of people in the IT departments believe energy efficiency and equipment recycling are important factors that need to be considered. In the same survey, 65 percent believed reduction of energy related operating costs as the driving factor for implementing Green IT [18]. Moving to the cloud will allow organizations to not only reduce their IT infrastructure, but, since it is much cheaper to transport computing services than energy, it will also represent a smarter use of energy.

In his much-heralded book, The Innovator’s Dilemma, Clayton Christensen pointed to disruptive technologies as innovations that upset the existing order of things in a particular industry [12]. Such disruptive technologies are usually lower-functionality innovations that appeal to customers who are not served by the current industry, but which quickly leapfrog the market incumbents in terms of functionality, innovation and price to upend the latter. Cloud computing today shows all the characteristics of a disruptive technology. We believe that many of the innovative services that will be developed on the cloud – such as the education applications being developed for Ethiopian schools – will soon make many cloud computing applications functionally richer than their in-house counterparts.

5.4. Threats

One of the biggest threats to cloud computing is the possibility of backlash from entrenched organizations. While we believe that many forward-looking organizations will see cloud computing as an opportunity to migrate to better computing practices that open up exciting opportunities for the in-house IT staff, there will probably be many other IT departments who will view it as a threat to their corporate IT culture (in terms of data security, IT audit policies, etc.) or just in terms of job security. Although small businesses have been quick to adapt and even welcome cloud computing, larger corporate customers have voiced a plethora of concerns about handing over their operations to another company. Another legitimate concern has centered on cloud providers going bankrupt, especially in a shrinking economy. Yet another concern is security — in an ongoing survey conducted by the research firm IDC, almost 75 percent of IT executives and CIOs report that security is their primary concern, followed by performance and reliability [56]. The cloud computing industry continues to make rapid strides in all these areas, but it will still be interesting to see how all these threats play out over the next few years in this nascent industry.

Several concerns have centered on the lack of standards. The cloud has been described as “a trap” by GNU creator and Free Software Foundation founder Richard Stallman — one where companies like Google will force customers into locked, proprietary systems that will gradually cost more and more over time. It is therefore encouraging to note that the International Organization for Standardization’s (ISO) technical committee for information technology has just announced the formation of a new Subcommittee on Distributed Application Platforms and Services (DAPS) that includes working a Study Group for standardization of cloud computing, with the goal of pursuing “active liaison and collaboration with all appropriate bodies...to ensure the development and deployment of interoperable distributed application platform and services standards in relevant areas.” More informally, industry professionals have coalesced to form several bodies like the Open Web Foundation (formed in 2008) that promote the development and protection of open, non-proprietary specifications for web technologies. Anticipating the backlash against proprietary cloud computing platforms, cloud computing providers are also proactively promoting standards. The recent formation of EuroCloud (in 2009), backed by more than 30 leading cloud computing vendors, to promote the development of standards in cloud computing across the EU that coordinate with local issues at the national level of individual countries is a welcome step. Even individual providers have promoted standards — for example, Google has formed the Data Liberation Front, an engineering team within Google whose goal is to make it easier for users to move their data in and out of Google products [55]; and Microsoft has recently filed a patent for a method that promises to streamline the process of moving from one cloud to another, and in many cases completely automate the process [13].

Perhaps the biggest factor that will impede the adoption of the cloud computing paradigm is regulation at the local, national, and international level. Regulation can range from data privacy and data access to audit requirements and data location requirements. When corporate data are moved to the cloud, regulations such as Sarbanes-Oxley and the Health and Human Services Health Insurance Portability and Accountability Act (HIPAA) with their defined requirements for physical data audit will come into play. Such and other requirements at the local, national and international level (e.g. many nations have laws requiring SaaS providers to keep customer data and copyrighted material within national boundaries) might negate many the benefits of cloud computing [14]. Regulation as an issue that is important enough to warrant a separate discussion in itself, and Section 7 is devoted to that discussion.

6. Stakeholders in cloud computing

In a traditional computing setup, the main stakeholders are the providers and consumers: the consumers use, own, maintain, and upgrade the systems while the providers deal with the sale, installation, licensing, consulting and maintenance of the technology involved. Cloud computing changes the roles of the traditional stakeholders and adds new ones. These stakeholders include not only the providers and the recipients of the service, but also, due to
the unique nature of the delivery model for the service, the regulators who need to understand the impact of the location of the infrastructure of the service providers. We discuss these stakeholders briefly below.

6.3. Enablers

Computing service providers own and operate cloud computing systems to deliver service to third parties. The providers will perform the maintenance and the upgrades on the system which consumers were in charge of when they owned the systems. They will also be responsible for maintaining the software used on the cloud, along with the pricing of the cloud services. Most cloud computing provider companies today have been large scale datacenters and software infrastructure. The different providers that we witness today have developed competencies around the different components (software, platform, and infrastructure) that make up the cloud computing service.

6.2. Providers

Cloud computing service providers own and operate cloud computing systems to deliver service to third parties. The providers will perform the maintenance and the upgrades on the system which consumers were in charge of when they owned the systems. They will also be responsible for maintaining the software used on the cloud, along with the pricing of the cloud services. Most cloud computing provider companies today have been large scale datacenters and software infrastructure. The different providers that we witness today have developed competencies around the different components (software, platform, and infrastructure) that make up the cloud computing service.

7. The role of regulation in cloud computing

In 2007, it was independently verified that the Internet service provider Comcast was slowing down network traffic within its servers that originated from the popular peer-to-peer (P2P) networks [35]. After initially denying any such behavior, Comcast then defended its actions by claiming that the traffic from P2P networks were slowing down other network traffic. The United States Federal Communications Commission (FCC) later declared Comcast's actions to be illegal, thus providing further fuel to the ‘net neutrality’ debate that is currently making the rounds in the US Congress and Senate [11]. We introduce this incident as an example to show that as the Internet becomes the backbone for transmission of all types of digital content, the government will increasingly find itself in the role of an arbiter in debates that involve the Internet and its use. The advent of cloud computing represents such a situation, and the success of this computing paradigm will depend to a large extent on how the regulatory bodies – both national and international – design laws to regulate it. While this article is too small to do justice to this aspect of the debate, we briefly touch upon some of the salient issues.

Developing countries can possibly have an advantage here, since in many cases they would not have to deal with an existing computing infrastructure. For example, when FullArmor proposed the cloud-based nationwide school information systems in Ethiopia, it was possible to re-imagine on a clean slate how data could be delivered to schools nationwide by fully exploiting the unique advantages of the cloud. A similar undertaking in a developed country would have to work through a smorgasbord of regulations at the national, state and local level, making such an undertaking a very daunting exercise.

Consumers and businesses today not only own their data, but they also control how that data is physically housed. The distributed nature of cloud computing alters many notions about residency and ownership of data and information. In converting to cloud computing, companies are essentially handing over their data to third-party service providers, who store and process such data in the cloud [27] and whose physical location could be anywhere in the world. This could potentially be a problem. For example, if some private data is stored in a country other than its owner, which country’s privacy laws would be followed by the cloud’s parent organization? Issues like these make it necessary for an active and informed role of national and international regulatory agencies. Some progress has already been made in this area through the development of US–EU Safe Harbor laws, but more needs to be done. As of now, providers like Amazon Web Services make sure that they comply with local laws through local infrastructure, and allowing their customers to select their "availability zones".

Unlike traditional Internet services, standard contract clauses may deserve additional review because of the nature of cloud computing. The parties to a contract should pay particular attention to their rights and obligations related to notifications of breaches in security, data transfers, creation of derivative works, change of control, and access to data by law enforcement entities. Because the cloud can be used to outsource critical internal infrastructure, and the interruption of that infrastructure may have wide ranging effects, the parties should carefully consider whether standard limitations on liability adequately represent allocations of liability, given the parties’ use of the cloud, or responsibilities for infrastructure. Since this is an issue that will crop up in every new contract that is drawn between an organization and a cloud provider, a consistent set of guidelines at a national level will be very helpful (e.g., establishing the obligation of cloud providers to notify customers of data security breaches or formulating the liability exemptions for cloud providers for different categories of security breaches).

Many other issues will come to fore because private information can be stored in a country that is different from that of the owner of that information. For example, can electronic evidence be suppressed
from a court of law in one country simply because that information is stored in another country and the two countries do not have an established “information exchange” treaty in place? What happens if the copyright law in the country where the data is physically stored does not permit access of some types of content, and therefore private consumers and organizations might not be able to access some necessary information (that they have legally procured elsewhere) when inside the country: how will a cloud computing provider ensure availability of information in different countries? Any company transferring its computing activities to the cloud risks running afoul of different countries’ laws governing data protection, most notably in the EU, which arguably has the world’s most stringent data protection laws. In fact, the central tenet of cloud computing – that the data could reside in a place which the organization might be unaware of – conflicts with the EU’s requirements that a company know where the data in its possession is at all times.

Governments therefore have to be proactive in dealing with cloud computing. Not only will they have to clarify their individual position regarding data, residency, privacy and related issues, but we also recommend that an international regulatory body be formed as soon as possible with the twin role of formulating cross-border issues and consulting individual governments in formulating their own cloud-related laws. The pressing need is to mitigate issues of moral hazard — businesses or private consumers should not be able to shop for lenient laws with respect to their data at their convenience. The formation of EuroCloud is thus a very welcome development, since it indicates that cloud computing providers are working proactively with the EU and the individual governments within the EU to remove potential future roadblocks.

The government may audit the communication or data in clouds, just as it does in financial markets, e.g. with the Sarbanes–Oxley regulation in the United States. Cloud computing will need an analogous set of regulation that unambiguously defines the responsibilities of an organization with respect to its cloud data. However, such regulation will also need to be sensitive with regards to privacy laws. Cloud computing raises new privacy issues that require clear standards for custodians of this information who receive government requests for access to that information. Privacy advocates have maintained that the US federal standards for accessing data stored in cloud computing applications (for example, photos, calendars, address books, and other personal or business information) need to be reassessed and modified to reduce their risks of harming privacy [4].

One issue in cloud computing that will have a definite regulatory impact is cloud computing forensics. Currently if a server or desktop is used to perform any illegal action, officials can seize the hardware for examination. However in a cloud computing system all servers and desktop are virtualized, being brought up and shut down as needed. There is currently no regulation in place to determine how to keep track of the use of the cloud system and what is required to be audited and logged (in other words, monitoring and traceability of incidents). Since the computing infrastructure will not be housed within premises, the legal system at both the locations of the user and the cloud computing provider will need to be examined, and if necessary, developed in tandem. International differences in relevant regulations including data protection and privacy need to be highlighted in order to prevent scenarios where an individual or organization cannot take advantage of the differences in regulations in two countries.

The logistical implications of cloud computing – huge data centers that have easy access to the Internet backbone and cheap power, with redundancy built in for both resources – will also necessitate government intervention. With potentially large amounts of sensitive, private and critical information hosted in a few physical locations, governments (as well as inter-governmental agencies for the coordination between governments) need to ensure that such sites are physically secure from terrorist attacks. The US government’s latest initiatives to combat cyber-terrorism [43] could in future also include the mandate to develop industry-wide guidelines for securing cloud data centers.

8. Recommendations

We conclude with a set of recommendations to the two sets of audiences that this article is aimed at: the business professionals who will need to prepare to migrate to cloud computing; and the IS researchers who will be looked upon to provide guidance to businesses on the new terrain. For the practitioners, our suggestions are relatively broad, given the fluid nature of the technology at this point in time. For the IS researchers, our recommendations are more concrete, since the fundamental forces that are influencing the cloud computing phenomenon are relatively clear.

8.1. Recommendations for business professionals

8.1.1. Which applications to move to the cloud

The discussion in the preceding sections makes it clear that there remain a significant number of challenges that need to be addressed before cloud computing becomes robust enough for large enterprises. It is also clear that not all applications are currently ripe for moving to the cloud. General-purpose applications (like office, email, collaboration technologies) are prime candidates, since there are rarely any instances of application requirements in such technologies that are specific only to an organization. Standalone applications like the popular CRM Salesforce.com might be easy to deploy on a cloud, but it is a different proposition altogether to migrate a smorgasbord of internally developed applications, third-party software and legacy applications with all their intricate interdependencies to the cloud. The cloud-based applications will also have to cross the threshold in terms of security and reliability – either actual or perceived – of their traditional counterparts.

8.1.2. Who should move to the cloud — the case for the SMEs

As of now, cloud computing makes eminent sense for SMEs; however, there are significant technical, operational and organizational issues which need to be tackled before clouds are used extensively at the enterprise level [50]. Current cloud computing services are often not cost-effective for larger enterprises, especially those that have achieved best-of-breed efficiencies from their computing operations. McKinsey Consulting found that a ‘typical’ data center of a large organization can operate at significantly lower costs than what would be required to outsource it to a cloud service like Amazon.com’s EC2 (though this price can be significantly lowered through pre-payment schemes, and with Linux systems). Further, McKinsey estimates that though the cloud service would lower labor costs, the extent of decrease (10–15%) is modest. Finally, many organizations set their service level agreement (SLA) uptimes at 99.99% or higher, which cloud providers currently are not prepared to match.

The equations are very different for SMEs that do not have the wherewithal to set up the initial infrastructure that is necessary to realize the cost structures of large data centers. The prices and the SLAs from the leading cloud providers are far better than what most SMEs can realize with their modest investment levels. Even more significantly, cloud computing needs no upfront investment, which will allow cash-strapped SMEs more flexibility with the use of their capital. Having much less of legacy IS infrastructure to contend with, it will also be much easier for SMEs to move to the cloud (and in many cases, the cloud might be the first instance when they try a new functionality, e.g. ERP, because the traditional alternative would have been too expensive in the first place). Finally, smaller organizations will have much less of the ingrained “attitudinal” issues that Carr [10] noted to deal with in moving their IS infrastructure to a cloud.
8.1.3. Large enterprises and the cloud

Large enterprises can still benefit from utilizing some of the core technological components of the cloud, e.g., virtualization. There are significant cost savings that can be realized by virtualizing end user computing, server storage, network operations, etc., and large organizations can learn from best practices to achieve significantly higher server utilization rates and lower total cost of ownership (TCO). In other words, large organizations can essentially implement “private clouds”, which offer many of the advantages of public clouds with one crucial difference: they will still incur a capital expenditure. However, by creating a high level of transparency about the TCO of the IS infrastructure at a business unit level, large organizations can inculcate a sense of the operational expenses on IT at a business unit level, and thereby drive down unnecessary IT expenses.

We believe that CIOs and CTOs should proactively develop an overall “cloud strategy” in order to determine a time-based plan about which of their applications they can move to the cloud, and the timeframes associated with each of them. We also recommend that large organizations should set up a small group (the ‘cloud committee’) that is distinct from the current IT setup and that continually evaluates developments in the cloud computing area (the independence is necessary in order to ensure that current practices do not unduly influence the group’s deliberations). If, as we believe, cloud computing can deliver new value for customers through innovative applications along the value-chain, the importance of cloud computing will be measured not only in terms of cost savings but increasingly in terms of the competitive advantages that it can deliver [6]. Therefore, the group should also be entrusted with the responsibility of managing the change within the IT department as the latter becomes geared more towards delivering innovative technology solutions and less on maintaining the current infrastructure. Instead of reinventing the proverbial wheel in every organization, we recommend that such groups across organizations should collaborate and share best practices. In fact, a cross-organizational body like the one we envisage would be ideally positioned to advise the technology industry in search of new ideas in this area. We believe that such cross-fertilizations can be very useful for vendors and clients alike.

Just as with any new technology, we will recommend a studied and deliberate approach in adopting cloud computing, and this will be especially true for larger enterprises. In this exercise, the cloud computing enablers should help manage the entire process, from helping develop pre-transition plans, to executing the phased rollover while keeping open backup systems in case of emergencies. Cloud computing enablers will also help implement organization-wide IS policies across diverse cloud computing services from multiple vendors. We also believe that given the extent of change and stakes involved, a risk-appropriate transition strategy for the larger enterprises would be to first move towards efficient use of their existing IT resources through virtualization. As a rule of thumb, no critical on-premise applications should be the first set of applications to be moved to the cloud. After the existing hardware infrastructure has been more efficiently utilized through partitioning and virtualization, the internal ‘cloud committee’ should recommend appropriate new projects that do not have any legacy component as possible candidates for cloud computing. As cloud computing technologies and application development tools become more mature, the organization can then explore the migration of their IT infrastructure to a ‘private cloud’ that is under the control of the organization. Depending on their level of comfort with the cloud infrastructure providers, organizations can look at either outright ownership of the ‘private cloud’ or leasing it from a provider of repute and financial stability (e.g., IBM, HP or Amazon). For many very large organizations, the economies of scale might mean that they will not find it economical to migrate their operations to a public cloud for many years to come [23]. Finally, some mission-critical applications will never be candidates for migration, regardless of the economics, just as hospitals sometimes maintain their own power generators for some of their critical services.

8.1.4. The cloud providers’ strategy

From a providers’ perspective, we believe that their interests will be best served by thinking of their end customers and how their needs will be met, rather than developing cloud applications just because they can be. They should also give a lot of thought in thinking about migration strategies for existing applications. Two approaches are possible: the first, and the more difficult strategy would be to develop a comprehensive migration strategy for all existing applications. The second and perhaps a more pragmatic approach – would be to develop a “divide-and-conquer” strategy, whereby potential customers can be enticed to try some of the novel characteristics of the cloud-based application, with the hope that increasing familiarity would lead to a greater degree of acceptance in future. For example, by their very design, cloud-based applications allow information sharing, something that has not been a design consideration for many traditional applications. Thus, while desktop-based office applications have far richer functionality than their cloud-based counterparts, the latter shine when it comes to document-sharing. Several people can work on different parts of the same document at the same time in Google Docs, something that cannot be achieved simply within Microsoft Office. This allows for a much richer environment for collaborative document creation (especially if such an endeavor is carried out in conjunction with a collaboration application like Skype or Google Chat). In such a setup, the initial structure of a document or presentation can be created collaboratively in, say, Google Docs, and then if required (for example, if the document is meant for external consumption), the final version can be finally downloaded into a desktop application for carrying out the decorative flourishes.

Since the industry is in such a fluid state (in terms of the technology, the business models and even the overall industry structure), we believe that the best opportunities for the cloud computing service providers lie in the small and medium segments of the market. Developing economies that do not have a large legacy IT infrastructure would also be prime candidates for moving to the cloud. Not only are such clients more receptive to cloud computing, but servicing these clients would crucially allow the providers to gain experience and help develop credibility in approaching the larger enterprises in future. The short term focus should lie in developing a customer base rather than on profitability. Successful implementations over a period of time can help allay fears about data security and availability [44]. Since security is a priority concern for many cloud customers, many of them will make buying choices on that basis. In a crowded marketplace, this might be a very effective market differentiator, which in turn can be a strong driver for cloud providers to improve their security practices.

One area where cloud computing providers can definitely approach larger enterprises is to offer their “elastic” infrastructure in handling heavy-duty computational work that would otherwise require huge investments in computational infrastructure whose costs would be difficult to recoup (two examples are of Eli Lilly, which is carrying out genomic research within the cloud [57], and NASA’s Jet Propulsion Laboratory (JPL) that uses Amazon EC2’s Cluster Compute instances to process high resolution satellite images of several giga-pixels that provide guidance and situational awareness to its robots).

Regardless of the strategy, it is very important for the cloud providers to manage end user expectations, so that this promising and fledgling technology does not become a victim of its own hype. We believe that some of the responsibility of managing expectations lies with the users too — especially with the afore-mentioned ‘cloud

---

7 Microsoft has claimed that the next version of Microsoft Office, Office 2010, will enable multiple users to work on the same document, but such a functionality will require additional server-based requirements.
committee who will recommend future IT projects for migration to the cloud.

8.1.5. Development of standards

Cloud computing providers should continue spending resources in the development of standards that promote interoperability within diverse cloud computing services. Our interactions with different industry executives lead us to believe that without a framework that allows transparent movement of data between organizations who might be with different providers, the success of cloud computing might be very limited. A related issue is manageability – the large-scale deployment of cloud computing in an organization will require not only the development of standards, but also standardized interfaces and automation tools for managing them in a mix-and-match environment. Finally, we feel that in the current “formational” stages, cloud computing providers should share best-of-breed operational practices openly, since that would help develop better products and promote lower prices, leading to greater acceptability and growth of the market for the industry as a whole.

8.2. A suggested research agenda

Our discussion suggests several potential streams of research for IS researchers. We purposely do not concentrate on the core technological issues in cloud computing, since that is outside the purview of IS research. The broad IS research agenda in cloud computing can be divided into six categories: (1) Cloud computing economics; (2) Cloud computing and IT strategy/policy issues (including security); (4) Technology adoption and implementation issues; (5) Cloud computing and green IT; and (6) Regulatory issues. We should emphasize that these topics are not islands in themselves, and many research issues will often have points of contact with several of them.

In going through the specific topics of research, we quickly realized that there were two fundamental dimensions under which these topics could be classified. Some of the research topics, such as cloud computing pricing strategy can be thought of to be more of a “business” issue, while a topic like security standards is more of a “technology” issue. Fig. 2 summarizes some of the specific research topics we identified and places them in a “business-technology” framework. Fig. 2 also has the advantage of placing the research topics visually with respect to one another in the “business-technology” framework, so that it is easy to see the linkages between the various research topics. For example, cloud computing pricing strategy (topic number 1 in Fig. 2) has several points of contact with optimal risk transfer and SLA contract design (topic number 9 in Fig. 2). We will elaborate on these linkages in greater detail after discussing the different research topics below.

8.2.1. Cloud computing economics

Research topics will include pricing strategies for cloud computing providers, and issues in industrial organization such as revising the boundary of the firm with altered transaction costs. There are many different potential pricing policies available to providers, such as a flat fee, a pay-per-use fee, or a two-tier mix of flat and pay-per-usage fee [2]. These can be analyzed in conjunction with capacity investment decisions and QoS guarantees. Researchers can examine which of these policies are best suited for a company given a particular market structure. Researchers can draw from the wealth of extant literature on pricing web services [48] and pricing of grid computing resources [58]. Research methodologies might include industrial organization and game theory. Recently, the IaaS provider Enomaly opened Spotcloud, a cloud computing clearinghouse and marketplace for cloud service providers to sell their unused cloud capacity to sellers looking for cloud services at the best possible price. Amazon too allows for selling its unused cloud capacity at spot prices. Such models

Fig. 2. The circled numbers in the above business-technology framework refer to the research topics indicated below (the topics are grouped under the six main research categories that we mention in the main text, and the dashed figures indicate the linkages between the various topics):

Research agenda
Cloud computing economics research
1. Cloud service pricing strategy
2. The role of enablers effect on the cloud computing provider economic value and the entire value chain
Strategy research in cloud computing
3. Corporate culture impact
4. Partnership/3rd party relational impact
IS policy research
5. Managing and implementing consistent IS policy across the usage of multiple cloud providers
6. Optimal software management for both a cloud provider and a Cloud user
7. The design of IT auditing policies, forensics, and evidence gathering methods
8. Security standards and issues
9. Optimal risk transfer and SLA contract design
10. International regulation and policy
Technology adoption and implementation research
11. The design of an optimal set of rules to decide the types of applications that should be moved to a cloud
12. The design of an optimal set of rules to implement the adoption of a private or public cloud by an organization
13. Developing a methodology to assess risk from adopting cloud computing
14. Researching best of breed application solutions within an industry
Government policy/regulation research
15. The identification of pertinent issues which need to be addressed that are created by the use of cloud computing for pricing open up new vistas of research in auctioning for resources whose availability is uncertain. Another promising research area is investigating the role of enablers, and the circumstances under which they would end up delivering economic value to a cloud computing provider or to the entire value chain [37,39].

8.2.2. Strategy issues in cloud computing

The introduction of cloud computing will potentially bring about a large change in the corporate IT structure [16], resulting in a host of intra-organizational issues that would need to be addressed. What type of cultural change would be needed, how a corporation will address this change, and how would the corporation achieve employee acceptance of the change are all important topics that
need to be examined. The effect of cloud computing on corporate culture will play an important role in its eventual success or failure. Researchers in this topic will find several points of contact with the IT outsourcing literature.

From a supplier perspective, the fundamental question that needs to be addressed is how a current technology vendor would move its offerings to (and modify them for) the cloud. Specifically, questions that will require research inputs include: How do providers utilize their strengths in deciding on future opportunities and target markets? What sort of strategy should they adopt in training their internal workforce even as they service current customers? What sort of strategy should they adopt for their current retail and other partners? Some of them might have been reliable partners in the current business model, but will be either unwilling or unable to evolve in the new environment. Creating new partnerships will create new opportunities as well as risks [41], and investigating relevant issues would provide fertile areas of research. Finally, technology vendors would require pointers that will help identify their ideal partners to serve as enablers for their technologies.

8.2.3. IS policy issues in cloud computing

IS policy issues will cover a broad swath of topics, from data privacy and data security to data ownership and audit [54], and organizations will look towards IS researchers for a set of principles and guidelines in these sensitive areas. Some specific areas where IS departments will look for guidance include (1) managing and implementing a consistent IS policy across multiple cloud computing providers [34]; (2) managing the move of software subscription, which traditionally is the domain of the IS department, to individual departments that require the cloud application instances; and (3) developing IT audit policies that are consistent with local, regional, national and international regulations.

Security on the cloud will be a major research topic in itself (for a good review on information assurance strategies, see [25]). This is not a novel issue for the IS community — security on grid computing has been studied for some time now (a very good review can be found in [15]). In some ways, cloud computing will offer several security benefits, with standardized interfaces, benefits of scale (i.e., the same amount of investment can provide better protection in terms of filtering, patch management, deployment of standard IS policy, etc.) and rapidity of response to security attacks. However, in some other ways, thanks to its very nature, cloud computing will increase some other security risks to the client organization. This is true especially in the early stages of cloud computing, when providers are yet to prove themselves to be “battle-ready”. In some cases, it might be difficult for the client to effectively check the data handling practices of the cloud provider and thus to be sure that the data is being handled in a lawful way. The standardized interfaces of the cloud providers might not be amenable to very fine levels of security policy administration details for different users and processes, thereby increasing the risks of breach. Further, early-stage SLAs might leave gaps in security defenses resulting in uncovered liabilities for the client. With multiple tenancy and reuse of hardware resources, adequate and timely deletion of data might be impossible to be carried out (or even be verified) when a client desires. This introduces an extra element of risk in terms of security and legal compliance for sensitive data. Finally, though usually less likely, damages caused by malicious insiders will often be far greater in the cloud, since cloud architectures involve certain roles which are extremely high-risk by their nature and scope (e.g., system administrators and managed security service providers). All these issues will necessitate research into topics such as optimal risk transfer and SLA contract design [5,22]; effects of different forms of breach reporting on security; forensics and evidence gathering mechanisms; incident handling (monitoring and traceability); international differences in relevant regulations including data protection and privacy [31,36], and so forth.

8.2.4. Technology adoption and implementation issues

Organizations are also looking for guidance in developing technology roadmaps, in order to decide (a) which applications are best positioned for moving to the cloud (i.e. how are applications to be divided between in-house and on the cloud) and (b) how to implement the changes in the least disruptive manner. The traditional in-house analysis, design and development of information technology will need to incorporate cloud computing as one of the viable choices. Another fruitful area of research will lie in developing methodologies for assessing the risks involved in adopting cloud services and comparing offerings from different providers to select the best provider(s) for a particular project. The security check-list should cover all aspects of security requirements including legal issues, physical security, policy issues and technical issues [20]. Once again, the existing literature on technology adoption [28] and specifically moving to grid computing [33,45] should provide a good starting point. Researchers can take an empirical approach to explore the gradual adoption of cloud computing, and develop econometric models to identify the strength of the relationship between relevant variables of interest. Even though secondary data sources are not yet established for such a nascent industry, our informal discussions with industry executives indicate that in many cases, the industry players would be happy to share their data in order to understand industry trends. Finally, organizations would also seek best practices for integrating the cloud applications with their in-house applications. For those applications that cannot be currently moved to a cloud, organizations will still benefit from empirical research that outlines the best-of-breed solutions [9] that can be set as benchmarks.

8.2.5. Regulatory issues

As we discussed in the previous section, there is an urgent need for the government and international agencies to be proactive in dealing with the unique challenges presented by the cloud computing environment (for a discussion on regulatory issues in grid computing see [40]). IS researchers can help in identifying the pertinent issues (some of which we identified earlier), in order to enable thoughtful legislation on the subject.

While the research topics discussed above are classified under different categories, several of these topics have points of contact between each other. In Fig. 2, these are shown as clusters that enclose the linked topics. The relative proximity of the topics in Fig. 2 also serves a purpose — they indicate the probability that research in one topic have fallout in the other area: for example, research on cloud computing pricing strategy will in all probability have to intimately consider issues in optimal risk transfer and SLA contract design. Similar conclusions can be drawn about the research topics in the other ‘clusters’. The different clusters are summarized below:

1) Cloud service pricing strategy
9) Optimal Risk transfer and SLA contract design
4) Partnership/3rd party relational impact
12) The design of an optimal set of rules to implement the adoption of a private or public cloud by an organization
14) Researching best of breed application solutions within an industry
5) Managing and implementing consistent IS policy across the usage of multiple cloud providers
11) The design of an optimal set of rules to decide the types of applications that should be moved to a cloud
7) The design of IT auditing policies, forensics, and evidence gathering methods
8) Security standards and issues
10) International regulation and policy
13) Developing a methodology to assess risk from adopting cloud computing
12) The design of an optimal set of rules to implement the adoption of a private or public cloud by an organization
15) The identification of pertinent issues which need to be addressed that will affect the use of cloud computing

9. Conclusion

Cloud computing is here to stay. The specific roadmap for the technology might be still unclear, but the fundamental economic and business forces that shape the computing industry point to a logical conclusion: many computing applications are general-purpose in nature, and therefore offer tremendous economies of scale if their supply can be consolidated. While the researchers and practitioners in the computer science community are making rapid strides in realizing this conclusion in technological terms, we believe that an equally important discussion needs to start from a business perspective. In this article we have tried to identify the different issues that will affect the various stakeholders in this radically different computing paradigm. In the process, we have attempted to come up with a set of recommendations for businesses and academics alike, with a passionate call to the latter community to start researching proactively in the issues that will affect the various stakeholders. Unlike some of the previous waves in computing, cloud computing also demands a thoughtful and coordinated response from governmental agencies, and we have listed down some of the key issues in this respect.

The famous physicist Niels Bohr once said, “Prediction is very difficult, especially about the future.” The fluid and uncertain environment that surrounds cloud computing – be it in terms of the technology, its adoption, the industry structure or the regulatory regime that will eventually need to be in place – makes such an exercise much that difficult. We however hope that our attempt to look at the phenomenon through a business perspective reduces some of the uncertainty surrounding the phenomenon, and more importantly jumpstarts the discussions on the various issues that we have identified.

The famous physicist Niels Bohr once said, “Prediction is very difficult, especially about the future.” The fluid and uncertain environment that surrounds cloud computing – be it in terms of the technology, its adoption, the industry structure or the regulatory regime that will eventually need to be in place – makes such an exercise much that difficult. We however hope that our attempt to look at the phenomenon through a business perspective reduces some of the uncertainty surrounding the phenomenon, and more importantly jumpstarts the discussions on the various issues that we have identified.

References


Sean R. Marston is a PhD candidate in the Department of Information Systems and Operations Management in the Warrington College of Business Administration at the University of Florida. He holds a B.E. in Computer Engineering and an MBA from Georgia Institute of Technology. His current research is in the area of digital distribution, e-commerce, and cloud computing.

Zhi Li is currently a PhD student in the department of Information Systems and Operations Management in the University of Florida, Gainesville. His research interests include economics of Information Systems; Artificial Intelligence and Data Mining; Systems Analysis and Design; and Electronic Business.

Subhajyoti Bandyopadhyay is an Associate Professor in the department of Information Systems and Operations Management in the University of Florida, Gainesville. He received his Ph.D. in MIS from Purdue University in 2002. His work has been published in several journals in the areas of Information Systems, Operations Management and Marketing. His current research interests include economics of information systems and information systems policy issues, especially in the area of Net Neutrality, national broadband policy and health informatics.

Juheng Zhang is a Ph.D. student at the Department of Information Systems and Operations Management at University of Florida. She has previously published in Decision Support Systems. Her research interests are data mining, cloud computing, and quantum computing.

Anand Ghalsasi is the Senior Director of Strategic Relationships at Persistent Systems, an award-winning independent software vendor that has developed several cloud-based applications with IBM, Microsoft, Oracle, Google and other global technology firms.