cloud computing

Somayyeh Aghababaei
Department of infromation science and technology
University Kebangsaan Malaysia
Bangi,Malaysia
Email: $s_a ghababaei7@yahoo.com$

Abstract

Cloud computing as an utility is a new term of computing which increasingly perceived. This computing utility will provide the basic level of computing service that is considered essential to meet every needs of the general community. To deliver this vision a number of computing paradigms have been proposed ,of which the last one is known as cloud computing. In this paper I want to define cloud computing ,the benefits of cloud computing and the architecture for creating clouds. In addition I discuss about variety of factors might influence companies to become cloud computing providers. I offer 10 obstacles and opportunities for cloud computing. Furthermore, I present new application opportunities .we believe that several important classes of existing applications will become even more compelling with cloud computing and contribute further to its momentum. As the number of cloud resources increase monitoring become a critical requirement so I perform monitoring in this paper. Finally I conclude the significant need of cloud computing which can provide tremendous value to companies of any size.

1. Introduction

Cloud Computing as a utility, has the potential to transform a large part of the IT industry, making software even more attractive as a service and shaping the way IT hardware is designed and Purchased. Cloud computing is a term used to describe both a platform and type of application. A cloud computing platform dynamically provisions, configures, reconfigures, and deprovisions servers as needed. Servers in the cloud can be physical machines or virtual machines. Advanced clouds typically include other computing resources such as storage area networks (SANs), network equipment, firewall and other security devices. Cloud computing also describes applications that are extended to be accessible through the Internet. These cloud applications use large data centers and powerful servers that host Web applications and Web services. Anyone with a suitable Internet connection and a standard browser can access a cloud application. By cloud computing businesses and users able to access applications from anywhere in the world on demand. Thus, the computing world is rapidly transforming towards developing software for millions to consume as a service , rather than to run on their individual computers.

I wish you the best of success.

mds
January 11, 2007

1.1. Cloud computing

For definition is better to say clouds provide on demand resources or services over the Internet, usually at the scale and with the reliability of a data center. Cloud Computing refers to both the applications delivered as services over the Internet and the hardware and systems software in the datacenters that provide those services. The services themselves have long been referred to as Software as a Service (SaaS). The datacenter hardware and software is what we will call a Cloud. When a Cloud is made available in a pay-as-you-go manner to the general public, we call it a Public Cloud; the service being sold is Utility Computing. We use the term Private Cloud to refer to internal datacenters of a business or other organization, not made available to the general public. Thus, Cloud Computing is the sum of SaaS and Utility Computing, but does not include Private Clouds. People can be users or providers of SaaS, or users or providers of Utility Computing. From a hardware point of view, three aspects are new in Cloud Computing. 1. The illusion of infinite computing resources available on demand, thereby eliminating the need for Cloud Computing users to plan far ahead for provisioning. 2. The elimination of an up-front commitment by Cloud users, thereby allowing companies to start small and increase hardware resources only when there is an increase in their needs. 3. The ability to pay for use of computing resources on a short-term basis as needed (e.g., processors by the hour and storage by the day) and release them as needed, thereby rewarding conservation by letting machines and storage go when they are no longer useful..

there are two important differences with the today's interest in cloud computing:.

1. The first difference is the scale. Some companies that rely on cloud computing add capacity by the data center and have infrastructures that scale over several (or more) data centers..

2. The second difference is the simplicity of many of the today's cloud services offerings. Prior to cloud-based computing services, writing code for high performance and distributed computing was relatively complicated and usually required working with grid-based services, developing code that explicitly passed massages between nodes, and employing other specialized methods. Although simplicity is the eye of the beholder, most people feel that the cloud-based storage service APIs and MapReduce style computing APIs are relatively simple compared to previous methods. The impact is the following. By using the Google File System and MapReduce, or the open source Hadoop Distributed File System and the Hadoop implementation of MapReduce, it is rel- atively easy for a project to compute using 10 TB of data over 1000 nodes. Until recently, this would have been out of reach for most projects. .

1.2. Benefits of cloud computing

1. Scalability: IT departments that anticipate an enormous uptick in user load need not scramble to secure additional hardware and software with cloud computing. Instead, an organization can add and subtract capacity as its network load dictates. Better yet, because cloud-computing follows a utility model in which service costs are based on consumption, companies pay for only what they use.

2. Easy Implementation: Without the need to purchase hardware, software licences or implementation services, a company can get its cloud-computing arrangement off the ground in record time and for a fraction of the cost of an on-premise solution.

3. Skilled Practitioners: When a particular technology becomes popular, its not uncommon for a whole slew of vendors to jump on the bandwagon. In the case of cloud computing, however, vendors have typically been reputable enough to offer customers reliable service and large enough to deliver huge datacenters with endless amounts of storage and computing capacity. These vendors include industry stalwarts such as Microsoft, Google, IBM, Yahoo! Inc. and Amazon.com Inc..

4. Frees Up Internal Resources: By placing storage and server needs in the hands of an outsourcer, a company essentially shifts the burden placed on its in-house IT team to a third-party provider. The result: In-house IT departments can focus on business-critical tasks without having to incur additional costs in manpower and training.

5. Quality of Service: Network outages can send an IT department scrambling for answers. But in the case of cloud computing, its up to a companys selected vendor to offer 24/7 customer support and an immediate response to emergency situations. Thats not to suggest that outages dont occur. In February 2008, Amazon.com's S3 cloud-computing service experienced a brief outage that affected a number of companies. Fortunately, service was restored within three hours.

There is plenty of concern surrounding cloud computing and its attendant security risks. What a lot companies fail to understand, however, is that many vendors rely on strict privacy policies, as well on sophisticated security measures, such as proven cryptographic methods to authenticate users. Whats more, companies can choose to encrypt data before even storing it on a third-party providers servers. As a result, many cloud-computing vendors offer greater data security and confidentiality than companies that choose to store their data in house.

1.3. some factors encourage companies to become cloud computing providers

Building, provisioning, and launching such a facility is a hundred-million-dollar undertaking. However, because of the phenomenal growth of Web services through the early 2000s, many large Internet companies, including Amazon, eBay, Google, Microsoft and others, were already doing so. Equally important, these companies also had to develop scalable software infrastructure (such as MapReduce, the Google File System, BigTable, and Dynamo [16, 20, 14, 17]) and the operational expertise to armor their datacenters against potential physical and electronic attacks. Therefore, a necessary but not sufficient condition for a company to become a Cloud Computing provider is that it must have existing investments not only in very large datacenters, but also in large-scale software infrastructure and operational expertise required to run them. Given these conditions, a variety of factors might influence these companies to become Cloud Computing providers: .

]1. Make a lot of money. Although 10 cents per server-hour seems low, for instance the economies of scale in 2006 for medium-sized datacenters(?1000 servers) that very large datacenters (tens of thousands of computer)can purchase hardware ,network bandwidth,, and power for 1=5 to 1=7 the prices offered to a medium-sized (hundreds or thousands of computers) datacenter. Further, the fixed costs of software development and deployment can be amortized over many more machines. Others estimate the price advantage as a factor of 3 to 5 [37, 10]. Thus, a sufficiently large company could leverage these economies of scale to offer a service well below the costs of a medium-sized company and still make a tidy profit. 2. Leverage existing investment. Adding Cloud Computing services on top of existing infrastructure provides a new revenue stream at (ideally) low incremental cost, helping to amortize the large investments of datacenters. Indeed, according to Werner Vogels, Amazons CTO, many Amazon Web Services technologies were initially developed for Amazons internal operations [42]. 3. Defend a franchise. As conventional server and enterprise applications embrace Cloud Computing, vendors with an established franchise in those applications would be motivated to provide a cloud option of their own. For example, Microsoft Azure provides an immediate path for migrating existing customers of Microsoft enterprise applications to a cloud environment. 4. Attack an incumbent. A company with the requisite datacenter and software resources might want to establish a beachhead in this space before a single 800 pound gorilla emerges. Google AppEngine provides an alternative path to cloud deployment whose appeal lies in its automation of many of the scalability and load balancing features that developers might otherwise have to build for themselves. 5. Leverage customer relationships. IT service organizations such as IBM Global Services have extensive customer relationships through their service offerings. Providing a branded Cloud Computing offering gives those customers an anxiety-free migration path that preserves both parties investments in the customer relationship. 6. Become a platform. Facebooks initiative to enable plug-in applications is a great fit for cloud computing, as we will see, and indeed one infrastructure provider for Facebook plug-in applications is Joyent, a cloud provider. Yet Facebooks motivation was to make their social-networking application a new development platform.

Several Cloud Computing (and conventional computing) datacenters are being built in seemingly surprising locations, such as Quincy, Washington (Google, Microsoft, Yahoo!, and others) and San Antonio, Texas (Microsoft, US National Security Agency, others). The motivation behind choosing these locales is that the costs for electricity, cooling, labor, property purchase costs, and taxes are geographically variable, and of these costs, electricity and cooling alone can account for a third of the costs of the datacenter. Table 3 shows the cost of electricity in different locales [10]. Physics tells us its easier to ship photons than electrons; that is, its cheaper to ship data over fiber optic cables than to ship electricity over high-voltage transmission lines. .

1.4. Obstacles and Opportunities for Cloud Computing

I offer a ranked list of obstacles to the growth of Cloud Computing.

The first three are technical obstacles to the adoption of Cloud Computing, the next five are technical obstacles to the growth of Cloud Computing once it has been adopted, and the last two are policy and business obstacles to the adoption of Cloud Computing.

1.4.1. Availability of a Service. Organizations worry about whether Utility Computing services will have adequate availability, and this makes some wary of Cloud Computing. Ironically, existing SaaS products have set a high standard in this regard. Google Search is effectively the dial tone of the Internet: if people went to Google for search and it wasnt available, they would think the Internet was down. Users expect similar availability from new services, which is hard to do. Just as large Internet service providers use multiple network providers so that failure by a single company will not take them off the air, we believe the only plausible solution to very high availability is multiple Cloud Computing providers. The high-availability computing community has long followed the mantra no single source of failure, yet the management of a Cloud Computing service by a single company is in fact a single point of failure. Even if the company has multiple datacenters in different geographic regions using different network providers, it may have common software infrastructure and accounting systems, or the company may even go out of business. Large customers will be reluctant to migrate to Cloud Computing without a business-continuity strategy for such situations. We believe the best chance for independent software stacks is for them to be provided by different companies, as it has been difficult for one company to justify creating and maintain two stacks in the name of software dependability.

1.4.2. Data Lock-In. Software stacks have improved interoperability among platforms, but the APIs for Cloud Computing itself are still essentially proprietary, or at least have not been the subject of active standardization. Thus, customers cannot easily extract their data and programs from one site to run on another. Concern about the difficult of extracting data from the cloud is preventing some organizations from adopting Cloud Computing. Customer lock-in may be attractive to Cloud Computing providers, but Cloud Computing users are vulnerable to price increases (as Stallman warned), to reliability problems, or even to providers going out of business. The obvious solution is to standardize the APIs so that a SaaS developer could deploy services and data across multiple Cloud Computing providers so that the failure of a single company would not take all copies of customer data with it. The obvious fear is that this would lead to a race-to-the-bottom of cloud pricing and flatten the profits of Cloud Computing providers. We offer two arguments to allay this fear. First, the quality of a service matters as well as the price, so customers will not necessarily jump to the lowest cost service. Some Internet Service Providers today cost a factor of ten more than others because they are more dependable and offer extra services to improve usability. Second, in addition to mitigating data lock-in concerns, standardization of APIs enables a new usage model in which the same software infrastructure can be used in a Private Cloud and in a Public Cloud. 9 Such an option could enable Surge Computing, in which the public Cloud is used to capture the extra tasks that cannot be easily run in the datacenter (or private cloud) due to temporarily heavy workloads.

1.4.3. Data Confidentiality and Audit ability. Current cloud offerings are essentially public (rather than private) networks, exposing the system to more attacks. There are also requirements for auditability, in the sense of Sarbanes-Oxley and Health and Human Services Health Insurance Portability and Accountability Act (HIPAA) regulations that must be provided for corporate data to be moved to the cloud. We believe that there are no fundamental obstacles to making a cloud-computing environment as secure as the vast majority of in-house IT environments, and that many of the obstacles can be overcome immediately with wellunderstood technologies such as encrypted storage, Virtual Local Area Networks, and network middleboxes (e.g. firewalls, packet filters). For example, encrypting data before placing it in a Cloud may be even more secure than unencrypted data in a local data center; this approach was successfully used by TC3, a healthcare company with access to sensitive patient records and healthcare claims Similarly, auditability could be added as an additional layer beyond the reach of the virtualized guest OS (or virtualized application environment), providing facilities arguably more secure than those built into the applications themselves and centralizing the software responsibilities related to confidentiality and auditability into a single logical layer. Such a new feature reinforces the Cloud Computing perspective of changing our focus from specific hardware to the virtualized capabilities being provided. A related concern is that many nations have laws requiring SaaS providers to keep customer data and copyrighted material within national boundaries. Similarly, some businesses may not like the ability of a country to get access to their data via the court system; for example, a European customer might be concerned about using SaaS in the United States given the USA PATRIOT Act. Cloud Computing gives SaaS providers and SaaS users greater freedom to place their storage. For example, Amazon provides S3 services located physically in the United States and in Europe, allowing providers to keep data in whichever they choose. With AWS regions, a simple configuration change avoids the need to find and negotiate with a hosting provider overseas.

1.4.4. Data Transfer Bottlenecks. Applications continue to become more data-intensive. If we assume applications may be pulled apart across the boundaries of clouds, this may complicate data placement and transport. At $100 to 150$ per terabyte transferred, these costs can quickly add up, making data transfer costs an important issue. Cloud users and cloud providers have to think about the implications of placement and traffic at every level of the system if they want to minimize costs. This kind of reasoning can be seen in Amazons development of their new Cloudfront service. One opportunity to overcome the high cost of Internet transfers is to ship disks. Jim Gray found that the cheapest way to send a lot of data is to physically send disks or even whole computers via overnight delivery services . Although there are no guarantees from the manufacturers of disks or computers that you can reliably ship data that way, he experienced only one failure in about 400 attempts (and even this could be mitigated by shipping extra disks with redundant data in a RAID-like manner). A second opportunity is to find other reasons to make it attractive to keep data in the cloud, for once data is in the cloud for any reason it may no longer be a bottleneck and may enable new services that could drive the purchase of Cloud Computing cycles. Amazon recently began hosting large public datasets (e.g. US Census data) for free on S3; since there is no charge to transfer data between S3 and EC2, these datasets might attract EC2 cycles. As another example, consider off-site archival and backup services. Since companies like Amazon, Google, and Microsoft likely send much more data than they receive, the cost of ingress bandwidth could be much less. Therefore, for example, if weekly full backups are moved by shipping physical disks and compressed daily incremental backups are sent over the network, Cloud Computing might be able to offer an affordable off-premise backup service. Once archived data is in the cloud, new services become possible that could result in selling more Cloud Computing cycles, such as creating searchable indices of all your archival data or performing image recognition on all your archived photos to group them according to who appears in each photo. A third, more radical opportunity is to try to reduce the cost of WAN bandwidth more quickly. One estimate is that two-thirds of the cost of WAN bandwidth is the cost of the high-end routers, whereas only one-third is the fiber cost [27]. Researchers are exploring simpler routers built from commodity components with centralized control as a low-cost alternative to the high-end distributed routers [33]. If such technology were deployed by WAN providers, we could see WAN costs dropping more quickly than they have historically.

1.4.5. Performance Unpredictability. multiple Virtual Machines can share CPUs and main memory surprisingly well in Cloud Computing, but that I/O sharing is more problematic. Figure 3(a) shows the average memory bandwidth for 75 EC2 instances running the STREAM memory benchmark [32]. The mean bandwidth is 1355 MBytes per second, with a standard deviation of just 52 MBytes/sec, less than 4for 75 EC2 instances each writing 1 GB files to local disk. The mean disk write bandwidth is nearly 55 MBytes per second with a standard deviation of a little over 9 MBytes/sec, more than 16problem of I/O interference between virtual machines. One opportunity is to improve architectures and operating systems to efficiently virtualize interrupts and I/O channels. Technologies such as PCIexpress are difficult to virtualize, but they are critical to the cloud. One reason to be hopeful is that IBM mainframes and operating systems largely overcame these problems in the 1980s, so we have successful examples from which to learn. Another possibility is that flash memory will decrease I/O interference. Flash is semiconductor memory that preserves information when powered off like mechanical hard disks, but since it has no moving parts, it is much faster to access (microseconds vs. milliseconds) and uses less energy. Flash memory can sustain many more I/Os per second per gigabyte of storage than disks, so multiple virtual machines with conflicting random I/O workloads could coexist better on the same physical computer without the interference we see with mechanical disks. The lack of interference that we see with semiconductor main memory in Figure 3(a) might extend to semiconductor storage as well, thereby increasing the number of applications that can run well on VMs and thus share a single computer. This advance could lower costs to Cloud Computing providers, and eventually to Cloud Computing consumers.

1.4.6. Bugs in Large-Scale Distributed Systems. One of the difficult challenges in Cloud Computing is removing errors in these very large scale distributed systems. A common occurrence is that these bugs cannot be reproduced in smaller configurations, so the debugging must occur at scale in the production datacenters. One opportunity may be the reliance on virtual machines in Cloud Computing. Many traditional SaaS providers developed their infrastructure without using VMs, either because they preceded the recent popularity of VMs or because they felt they could not afford the performance hit of VMs. Since VMs are de rigueur in Utility Computing, that level of virtualization may make it possible to capture valuable information in ways that are implausible without VMs.

1.4.7. Scaling Quickly. Pay-as-you-go certainly applies to storage and to network bandwidth, both of which count bytes used. Computation is slightly different, depending on the virtualization level. Google AppEngine automatically scales in response to load increases and decreases, and users are charged by the cycles used. AWS charges by the hour for the number of instances you occupy, even if your machine is idle. The opportunity is then to automatically scale quickly up and down in response to load in order to save money, but without violating service level agreements. Indeed, one RAD Lab focus is the pervasive and aggressive use of statistical machine learning as a diagnostic and predictive tool that would allow dynamic scaling, automatic reaction to performance and correctness problems, and generally automatic management of many aspects of these systems. Another reason for scaling is to conserve resources as well as money. Since an idle computer uses about two-thirds of the power of a busy computer, careful use of resources could reduce the impact of datacenters on the environment, which is currently receiving a great deal of negative attention. Cloud Computing providers already perform careful and low overhead accounting of resource consumption. By imposing per-hour and per-byte costs, utility computing encourages programmers to pay attention to efficiency (i.e., releasing and acquiring resources only when necessary), and allows more direct measurement of operational and development inefficiencies. Being aware of costs is the first step to conservation, but the hassles of configuration make it tempting to leave machines idle overnight so that nothing has to be done to get started when developers return to work the next day. A fast and easy-to-use snapshot/restart tool might further encourage conservation of computing resources.

1.4.8. Reputation Fate Sharing. Reputations do not virtualize well. One customers bad behavior can affect the reputation of the cloud as a whole. For instance, blacklisting of EC2 IP addresses [31] by spam-prevention services may limit which applications can be effectively hosted. An opportunity would be to create reputation-guarding services similar to the trusted email services currently offered (for a fee) to services hosted on smaller ISPs, which experience a microcosm of this problem. Another legal issue is the question of transfer of legal liabilityCloud Computing providers would want legal liability to remain with the customer and not be transferred to them (i.e., the company sending the spam should be held liable, not Amazon).

1.4.9. Software Licensing. Current software licenses commonly restrict the computers on which the software can run. Users pay for the software and then pay an annual maintenance fee. Indeed, SAP announced that it would increase its annual maintenance fee to at least 22computing providers originally relied on open source software in part because the licensing model for commercial software is not a good match to Utility Computing. The primary opportunity is either for open source to remain popular or simply for commercial software companies to change their licensing structure to better fit Cloud Computing. For example, Microsoft and Amazon now offer pay-as-you-go software licensing for Windows Server and Windows SQL Server on EC2. An EC2 instance running Microsoft Windows costs $0.15 per hour instead of the traditional 0.10$ per hour of the open source version.15 A related obstacle is encouraging sales forces of software companies to sell products into Cloud Computing. Payas- you-go seems incompatible with the quarterly sales tracking used to measure effectiveness, which is based on one-time purchases. The opportunity for cloud providers is simply to offer prepaid plans for bulk use that can be sold at discount. For example, Oracle sales people might sell 100,000 instance hours using Oracle that can be used over the next two years at a cost less than is the customer were to purchase 100,000 hours on their own. They could then meet their quarterly quotas and make their commissions from cloud sales as well as from traditional software sales, potentially converting this customer-facing part of a company from naysayers into advocates of cloud computing.

2. Conclusion

The conclusion goes here. this is more of the conclusion

Acknowledgment

The authors would like to thank... more thanks here

References

[1] H. Kopka and P. W. Daly, *A Guide to LaTeX*, 3rd ed. Harlow, England: Addison-Wesley, 1999.