Cloud Computing: The Limits of Public Clouds for Business Applications

Paul Hofmann • *SAP Labs*
Dan Woods • *CITO Research*

Descriptions of cloud computing often emphasize the silver lining more than the chances of getting wet. Virtualization offers many benefits, but the cloud — especially the public cloud — is unsuited to many business applications and is likely to remain so for many years due to fundamental limitations in architecture and design.

No doubt, cloud computing is a breakthrough technology that will continue to unleash new innovations and bring new efficiencies and advantages to business. It removes infrastructure and capital expense as a barrier to entry and allows startups to scale up cheaply and rapidly. On the other hand, enterprises face limitations in using the cloud for high-performance and mission-critical applications such as ERP. Unfortunately, the cloud's limits are often obscured by all the hype. It's time to stop looking at the cloud as a panacea. This article seeks to clear up some misperceptions and help people make better choices.

The Sunny Side of the Cloud

Certainly, cloud computing offers many attractive benefits to enterprises. The cloud model moves IT infrastructure from an upfront capital expense to an operational one. Companies can use the cloud for large batch-oriented tasks — those involving large spikes in requirements for processing power — that otherwise would be out of reach or require huge investment. Many enterprises provision computing resources for peak loads, which often exceed average use by a factor of 2 to 10. Consequently, server utilization in datacenters is often as low as 5 to 20 percent. One key benefit of cloud computing is that it spares companies from having to pay for these underutilized resources. Cloud computing shifts the IT burden and associated risks to the vendor, who can spread variations over many customers. Organizations can use the cloud to rapidly scale up or down; they can also buy or release IT resources as needed on a pay-as-you-go model. As one group of researchers from the University of California, Berkeley noted, "This elasticity of resources, without paying a premium for large scale, is unprecedented in the history of IT" (www.eecs.berkeley.edu/Pubs/TechRpts/2009/EECS-2009-28.pdf).

The cloud can be a revolutionary technology, especially for small startups, but its benefits wane for larger enterprises with more complex IT needs.

Plug and Play?

Cloud proponents often compare utility computing to electrical utilities. One of the most prominent voices behind this argument is Nicholas Carr, author of *The Big Switch: Rewiring the World, from Edison to Google* (Norton, 2008). Carr hails utility computing as a historic shift similar to the advent of electrical utilities. A century ago, factories provided their own power, but with the emergence of large utilities, electricity became a cheap commodity, enabling businesses to simply plug into the grid. Carr argues that a similar phenomenon is occurring with cloud computing. Private computer systems are being supplanted by services provided via the Internet. "It may take decades for companies to abandon their proprietary supply operations and all the investment they represent," writes Carr. "But in the end the savings offered by utilities become too compelling to resist, even for the largest enterprises. The grid wins."

This utility analogy has taken hold in the public imagination. Although useful, this analogy isn't entirely accurate because it blinds us to the cloud's limitations for enterprises. The reality is that cloud computing simply can't achieve

 1089-7801/10/$26.00 © 2010 IEEE

the same plug-and-play simplicity as electricity.

The Trade-Offs of the Cloud

Enterprises can expect to face many trade-offs when they move IT into the cloud.

Security

Security is one of the biggest challenges to the cloud model, and it's often an emotional one as well. Again, the utility analogy isn't very illuminating here because most companies spend little time worrying about whether their electrical wires are being compromised. In contrast, a violation of data security is a paramount concern to an organization.

Behind the firewall, enterprises have control of their data. In the cloud, they must trust the provider. Many organizations are loath to entrust their sensitive data and their reputation to the public cloud.

For some companies, especially smaller organizations with limited resources, data may be safer with a cloud provider than on premises. But for organizations whose existence depends upon safeguarding customer data, trade secrets, classified information, or proprietary information, public cloud providers don't offer sufficient protection. Most providers find it hard, if not impossible, to meet standards for auditablity and comply with legislation such Sarbanes-Oxley and the Health and Human Services Health Insurance Portability and Accountability Act (HIPAA).

Portability, Lock-In, and Interoperability

As cloud offerings proliferate, there will be ongoing challenges with interoperability, portability, and migration. To be sure, interoperability also is an issue for on-premise applications, but this challenge is magnified in the cloud. In an on-premise model, enterprises can switch hardware or software at any time; in the cloud, they're locked in to a provider and no longer control their own IT.

Cloud providers speak different languages. All the major providers offer unique, and often proprietary, data storage (for example, Google's BigTable, Amazon's Dynamo, and Facebook's Cassandra). Scalable data storage isn't yet a commodity and is unlikely to be so for a long time due to the rapid pace of IT innovation. Scalable relational database management systems (RDBMSs) remain an unsolved scientific problem, posing huge challenges for interoperability. Transporting data between different cloud providers is a nightmare exacerbated by the network's limitations.

Again, the electrical utility analogy isn't illuminating here. Electricity is an interchangeable commodity, meaning the customer can plug into any electrical grid and won't care whether the power comes from a hydroelectric plant, coal plant, or wind farm. Electrons are fungible, but bits of information are not. As long as cloud providers speak different languages, data storage will not be a truly interchangeable commodity.

Cloud Computing Remains Largely Proprietary

Although the world of software has emphasized interoperability in recent years, the clouds haven't yet seen a strong impetus toward standardization. Customers can't easily exchange data from one cloud to another. Where should they store these data, and in what form? Take the example of contact data in Salesforce CRM and Google's Gmail and Calendar services. Salesforce doesn't offer an interface to Gmail or Google Calendar, so companies have to upload their contact data from the Salesforce RDBMS, transport it to Google's AppEngine, and convert it into another format for Google contacts. Cloud users can face severe constraints in moving their data from one cloud provider to another and find themselves locked in.

Absence of Service-Level Agreements

Another problem is the lack of well-defined service-level agreements (SLAs) by cloud providers. What's the guaranteed uptime? What are the repercussions if the provider fails to meet these standards? What happens to customer data if the company moves to a different provider?

Cloud providers offer precious few protections to enterprises that trust all their IT to the cloud. In the article, "Why Cloud Computing Will Never Be Free," Dave Durkee points out that "pricing pressure results in a commoditization of cloud services that deemphasizes enterprise requirements such as guaranteed levels of perfor-

Shape of the Cloud

At its core, cloud computing means providing computing services via the Internet. The "cloud" idea is tightly connected with the "as a service" idea. The public cloud, for example, represents a set of standard resources of varying types that can be combined to build applications. Public clouds offer virtual machines to provide computing power, file systems, data storage systems, network devices, and other elements. They are often referred to as *infrastructure as a service*.

Various forms of public cloud providers and software as a service companies also offer a development *platform as a service* (for example, Google and Salesforce.com). In general, the public cloud has significant limitations when used as an infrastructure as a service to construct business applications. These limitations are challenging enough that the migration to the cloud will primarily consist of a private cloud infrastructure that bears little resemblance to the public cloud.

mance, uptime, and vendor responsiveness" (bit.ly/d1sI84). Furthermore, "in the cloud market space, meaningful SLAs are few and far between, and even when a vendor does have one, most of the time it is toothless. For example, a well-known cloud provider guarantees an availability level of 99.999% uptime, or five minutes a year, with a 10% discount on their charges for any month in which it is not achieved. However, since their infrastructure is not designed to reach five-nines of uptime, they are effectively offering a 10% discount on their services in exchange for the benefit of claiming that level of reliability. If a customer really needs five-nines of uptime, a 10% discount is not going to even come close to the cost of lost revenue, breach of end-user service levels, or loss of market share due to credibility issues."

The lack of enterprise-grade SLAs in the cloud is amplified when customers rely on multiple cloud providers that offer different levels of guarantees. What service does a user receive when cloud provider X offers SLA A and provider Y offers SLA B? To date, there is no scientific solution to the problem of federated SLAs.

Performance Instability

The cloud is often touted as a solution for organizations with large variations in computing demands. Less well known is the performance variability in the clouds themselves.

Researchers in Australia conducted stress tests to demonstrate that Amazon, Google, and Microsoft suffered from variations in performance and availability due to loads. Specifically, the researchers measured how the cloud providers scaled up and responded to the sudden demand of 2,000 concurrent users. In some cases, response times at different points of the day varied by a factor of 20 (www.itnews.com.au/News/153451,stress-tests-rain-on-amazons-cloud.aspx).

Latency and Network Limits

At the risk of sounding ironic, another limitation to using the cloud is the speed of light. As long as we rely on fiber-optic cables, we're limited by network speed (unfortunately, the speed of light isn't amenable to the kind of speed improvements associated with Moore's law).

As applications make ever-more intense use of large volumes of data, data transfer poses an increasing bottleneck. For example, University of California, Berkeley, computer scientists calculated the costs of shipping 10 Tbytes of data from the Bay Area to Amazon in Seattle. Given the average bandwidth, sending this data would take 45 days and cost US$1,000 in network transfer fees. In contrast, shipping 10 1-Tbyte disks overnight would cost only $400. This model of "Netflix for cloud computing" offers a way to avoid some of the latency problems and data transfer costs.

In fact, many companies that must transfer massive amounts of data — say, a pharmaceutical company submitting data to the US Food and Drug Administration to win approval for a new drug — still find it more efficient to ship their computers instead of extracting data and sending it via Internet. This suggests that the cloud won't be a good option for companies that require instant processing of large amounts of data that must be sent over the network. Similarly, the cloud might not be a good option for companies that use data generated by two different cloud applications (one financial and the other supply chain), or data from sensors in a manufacturing plant that must be processed by a business application in the cloud. The cloud isn't suited to stock trading, for example, because it requires speed and split-second precision. Consequently, financial service firms often locate their datacenters as close as possible to stock exchanges.

No Scalable Storage

Cloud computing isn't simply a matter of adding an infinite number of servers. Some problems and processes can't be solved simply by adding more nodes — they require different architectures of processing, memory, and storage.

Most business applications today rely on consistent transactions supported by RDBMSs, which unfortunately do not scale. The cloud lacks scalable storage with an API as rich as SQL, which considers queries as a logical unit. There's no industrial-grade solution for applications that rely on consistent transactions to write on two different nodes at the same time (the famous two-phase commit problem), thus it's difficult for high-volume, mission-critical transactional systems to run in the cloud. Scalable storage with a SQL-like API remains an unsolved research problem (although there are promising attempts under way; www.eecs.berkeley.edu/Pubs/TechRpts/2010/EECS-2010-8.pdf).

Because we have no general solution for scalable data storage and retrieval in the cloud, each platform has its own solution. Amazon Dynamo, Facebook Cassandra, and Google Big Table each rely on key value store, which is scalable but doesn't allow storage of complex table structures like relational databases do. Consequently, these solutions lack the power required for many business applications. Let's say you're a vendor doing inventory management on Amazon. If you have 100 pieces in your inventory but remove half of them, your inventory won't reflect this change for a couple of hours. Needless to say, this sort of key value store database is impractical for enterprise-level applications.

Does the Cloud Stifle Innovation?

Perhaps the cloud's biggest limitation is that it might impair innova-

tion. Implemented properly, ERP represents a significant source of competitive advantage, but if ERP becomes a commodity — the cloud model's central premise — it limits a company's ability to innovate.

IT represents a source of competitive advantage for many organizations. In a 2008 *Harvard Business Review* article (www.scribd.com/doc/13415798/Investing-in-IT-That-Makes-a-Competitive-Difference), Andrew McAfee and Erik Brynjolfsson found that competition within the US economy had accelerated to unprecedented levels in the wake of the mainstream adoption of the Internet and commercial enterprise software. The main catalyst was the massive increase in IT power. As the authors write, "a company's unique business processes can now be propagated with much higher fidelity across the organization by embedding it in enterprise information technology. As a result, an innovator with a better way of doing things can scale up with unprecedented speed to dominate an industry."

The average company's IT investment grew from $3,500 per worker in 1994 to about $8,000 in 2005. During this period, annual productivity growth in US companies roughly doubled. This period of intensive IT investment ushered in an era of greater turbulence, wider gaps between leaders and laggards, and winner-take-all concentration. The key driver of this trend wasn't simply the new array of IT products — rather, IT enabled improvements in operating models and propagated them quickly and widely. This put a premium on deploying powerful technology platforms like ERP, using them to innovate better business processes, and replicating these best practices throughout the enterprise.

But how much can a company innovate when it uses a plain vanilla IT? Real IT innovation comes from tailoring ERP systems to the unique needs of every company. Despite all the hype about enabling innovation, the cloud actually impairs the ability of large enterprises to gain competitive advantage because it's optimized for the cloud provider, not the customer. It's designed for ease of maintenance, scalability, and lowest common denominator functionality. It limits the ability of customers to tailor their software and wring real competitive advantage from their IT systems.

Consider Apple. Its shift from a perpetual license model to the iTunes store's pay-per-use option allowed it to quadruple revenues in four years. The Apple model depends on tight integration between Apple's ERP system and the billing engine, which handles 10 million sales per day. It would be difficult, if not impossible, to set up such a tight integration between the cloud's ERP and Apple's highly proprietary billing software.

General-purpose technologies deliver their full benefit because they spur additional innovations. Electricity gave rise to electric lighting, motors, and machinery. Similarly, IT gave birth to transaction processing, ERP, online commerce, and business model innovations. The cloud limits opportunities for complementarities and co-invention.

Ultimately, the cloud is neither good nor bad: it's just a new paradigm with its own advantages and disadvantages. Over time, some of these concerns will be solved or the risks will be reduced to acceptable levels. For now, these concerns have kept cloud adoption at a modest pace. According to IDC, less than 10 percent of worldwide IT spending will be for cloud computing by 2013 (www.slideshare.net/JorFigOr/cloud-computing-2010-an-idc-update).

The cloud can act as a foundry that allows small startups to overcome IT barriers and bring new offerings to the market, but it's unlikely to serve the needs of larger enterprises. For most organizations, the question of whether to move into the cloud will be a matter of weighing the pros and cons. There's a "sweet spot" for cloud business applications where the trade-off is optimum — and at this point, that spot is around HR and CRM. Not surprisingly, this is where we also see the most common SaaS applications. CRM and HR applications are relatively simple without too much functionality or customization. Thus, we're likely to see more domain-specific business applications like Salesforce for CRM or Workday for HR.

Much of enterprise IT will move toward virtualization, but not necessarily the public cloud. Some companies might virtualize their IT by moving to private clouds, which provide benefits like economies of scale without the drawbacks of a public cloud. For example, large companies such as BP, Intel, or IBM have virtualized their own resources and reaped the advantages of volume, statistical multiplexing, and utilization. In particular, IBM has saved $1.5 billion by consolidating its datacenters from 115 down to 5.

For large companies, the private cloud represents an option to have your cake and eat it too. When compared to the standard components of the public cloud, the custom-made private cloud stands out as a radically different construct. Unfortunately, many people continue to loosely throw around the term "cloud" without realizing that it may refer to very different models and without realizing its limitations.

Dan Woods is CTO and editor of CITO Research, a New York-based technology research and publishing firm. Contact him at dwoods@citoresearch.com.

Paul Hofmann is vice president at SAP Labs in Palo Alto. Contact him at paul.hofmann@sap.com.