

Trending Toward the Cloud

The First 50 Years of
Cloud Computing



By Robert A. Bye, Executive Vice President, nGenX Corporation



Executive Summary

Anyone who's been involved in the business of computing for any length of time knows that trendy buzzwords almost always accompany new ideas. Even when there is sound business value behind the buzzword, it often triggers a massive hype-cycle in the marketplace as vendors scramble to demonstrate that they are the vanguard of the new idea, causing the term to get so overused and extended that it becomes nearly meaningless.

This accounts for much of the confusion regarding Cloud Computing, which makes it difficult for vendors offering genuinely unique products, services or applications to differentiate themselves from those simply relabeling existing products and services.

Even with this confusion, Cloud Computing is growing at a rapid pace because it is a convenient and environmentally friendly solution businesses of all sizes can use to support their burgeoning IT needs. This market is hitting its stride: according to analyst firm, IDC, spending on cloud services in 2012 will be almost triple its 2007 levels, realizing revenues of \$42 billion industry-wide. IDC further predicts that cloud-related outlays will account for 25% of the growth in IT spending in 2012 and almost a third by 2013.

To really understand the value Cloud Computing can have on a business, it is important to understand how the industry defines the technology. But more important is to understand the practical applications of Cloud Computing and how those applications can lower IT costs while increasing its effectiveness.

Section I Cloud Computing Through the Ages

Cloud Computing is new in some ways; yet, not at all new in others. Although the term itself came about relatively recently as a result of systems engineers depicting networks as clouds when diagramming distributed systems, we've actually been on a journey to the Cloud for more than half a century. Let's take a look at how the ideas behind Cloud Computing have actually been shaping the computer industry for decades.

Cloud Computing was Born as Time-Sharing

Two important benefits of Cloud Computing are the cost efficiency and economy-of-scale that come about as a result of the multi-tenancy, or the shared nature of the Cloud. By increasing the number of users sharing a resource – for example a server – we can reduce both the cost per user for that resource and the likelihood that at any given moment it will be powered up but unused.

One early effort at achieving these benefits was

time-sharing, first initiated by John McCarthy in 1957. By allowing many users to share a single computer, time-sharing dramatically lowered the cost of computing at a time when computers were rare beasts that cost multi-millions of dollars and required care and feeding that was well out of reach for all but the largest organizations.

By providing shared, multi-tenancy computing resources, time-sharing delivered some of the same economy and efficiency benefits that Cloud Computing does today, and inspired the idea of Utility Computing, which for many is synonymous with Cloud Computing.

Everything-as-a-Service Began with CompuServ

Time-sharing remained popular during 60s and 70s, but by 1980, time-sharing began to decline as low cost microprocessors made it possible for much smaller organizations and even individuals to have their own private microcomputers. As computers became more common, CompuServe emerged with a suite of very Cloud-like services for both business

and individual computer users.

Although CompuServe couldn't provide the economy and efficiency of time-sharing to users who were essentially running their own CPUs, they did pioneer the delivery of two other important benefits of today's Cloud: unmatched data storage capacity and the ability to access that data from anywhere. Connecting to CompuServe via a phone line enabled users to edit a file in 16K of memory, upload it to a CompuServe server, and then retrieve it anytime later from wherever they happened to be.

At the same time, CompuServe was also delivering SaaS (Software as a Service) products by making packaged software applications available, and DaaS (Data as a Service) products in the form of stock quotes and other financial data feeds.

From the Cluster to the Grid

In an effort to expand the amount of computing power, sophisticated IT professionals clustered multiple computers to form a single more robust computer. The technique enabled computers to speak with each and act as one. The user neither knew nor cared which CPU ran which program or stored its data. The cluster management software allowed that the currently most accessible computer performed the function.

In the early 1990s, the industry saw this concept grow from a cluster solution to a grid. Like electricity, the analogy suggested, users could access the grid and purchase the needed services on a metered basis. Grid computing expanded the concept of clustering by combining multiple independent clusters not from a single domain to act as one. While not technically Cloud Computing, the concept was a similar fore-runner.

So What Exactly is Cloud Computing Today?

Cloud standards organizations have been working hard to develop a Cloud taxonomy to help answer that question, and though one has yet to be officially agreed upon, the simplest and most common view of Cloud Computing is referred to as the SPI Model, and looks something like this:

- SaaS [Software-as-a-Service] – End-users most often engage the Cloud to gain access to software applications they use in their day-to-day work and personal lives. Salesforce.com is the

most recognized example.

- PaaS [Platform-as-a-Service] – The underlying structure upon which companies create software applications.
- Developers enter the Cloud to facilitate the development of new applications as PaaS eliminates the cost and complexity of buying and managing the underlying hardware and software layers. Amazon E2 or Microsoft Azure are recognized examples.
- IaaS [Infrastructure-as-a-Service] – Data center, managed applications, managed servers. WAN management utilizes the Cloud to relieve its over-burdened computing infrastructure and deploy new IT initiatives without hard-to-secure capital funding.
- The 3 layers of the basic SPI Model are sometimes broken out further, by adding:
- DaaS (Data-as-a-Service) – information services that focus on investments, training and other knowledge, ready to be retrieved. Both consumers and business professionals rely on a host of websites that provide information on a subscription basis. In fact these services are dramatically changing how the world's population gets news, training and data.
- BaaS (Backup as a Service) – off-site data storage/retrieval and disaster recovery. Companies with small or non-existent IT departments turn to data backup services to store their data for later retrieval or as an insurance against man-made or other tragedies.

Cloud Computing is not a single, clearly defined idea. It has many nuances, facets and components businesses can select and combine together, almost cafeteria style, to create a definition that uniquely describes how that company is taking advantage of Cloud Computing's benefits of efficiencies and economies-of-scale. Here is how three major IT analyst firms define the concept:

- IDC: an emerging IT development, deployment and delivery model enabling real-time delivery of products, services and solutions over the internet.
- Gartner, Inc.: A style of computing in which massively scalable IT-related capabilities are

provided “as a service” using internet technologies to multiple external customers.

- Forrester Research, Inc.: A pool of abstracted, highly scalable and managed compute infrastructure, capable of hosting end-customer applications and billed by consumption.

Section II – Life in the Cloud

Users today have become so immersed in their on-line experiences that their desktops, laptops, netbooks, PDAs and mobile phones are simply becoming gateways to their real computer: the internet. It is the internet – the network we now refer to as the Cloud, and arguably the only network that matters – that is the user’s main data and document management vehicle, or computer. For many:

- Connecting with business contacts comes through LinkedIn; connecting with friends, Facebook
- Pictures are stored on Flickr
- Customer relationships are managed through salesforce.com subscriptions
- Data and applications are no longer stored on the hard drive, instead are accessed from an outsourced managed applications service provider via a thin client

These examples make it clear that Cloud Computing is more evolutionary than revolutionary, having evolved from successful prior experiences over several decades, and that adopting Cloud Computing today is neither a radical nor a risky move. Instead the Cloud is becoming ubiquitous in both our personal and business lives.

Viewing the Cloud Through Their Eyes

Different job functions view the Cloud through their own scope, and judge its effectiveness based on that vantage point.

- IT professionals tend to view Cloud Computing either as an escape from having to be tied to on-site, super-cooled, glass walled and very costly data centers; or as a threat to their job security.
- CIOs view it as a solution that allows them to refocus IT resources on more strategic projects;

and as a facilitator to deploying new IT initiatives more quickly and more cost-effectively.

- SMB owners view it as an opportunity to have a sophisticated outsourced IT solution to support their business without the hassle or need to procure capital funding to build, maintain or manage the equipment or staff necessary to support it.
- Non-IT business professionals view it as an improved technology solution that allows them to manage their jobs with the help of multiple devices that all point them to the same applications and data.

Cloud Computing Supports Business in the Shrinking Economy

In today’s tough business and economic climate, if the IT resource can speed up the implementation of vital line-of-business software, improve its management of a growing stockpile of valuable data and/or decrease computer infrastructure costs, it can have a vital impact on the success of the organization.

Ideally, organizations should be able to obtain the necessary IT resources when and as they’re needed, paying only for what’s used. In a healthy economy, this can support growth and speed up new initiatives. In a bad economy, it can mean survival as it can foster efficiency and save costs.

As discussed, Cloud Computing turns up an alphabet soup of terms and concepts describing the many new ideas for bringing efficiency and effectiveness to IT. Taken individually, each of them may provide some specific benefit over the classic or legacy practices that they replace, but it is undeniable that synergies emerges when multiple advances in ways of thinking and working are applied to the current problems and inefficiencies that have traditionally plagued IT departments in every size organization.

Interest in Cloud Computing has been fueled by the need to do more with less, and it has been made a reality in part by recent advances in fiber optics and storage technologies and the availability of low-cost, space saving hardware that enables the construction of mammoth new data centers that are more environmentally friendly and more cost

efficient than ever before.

The Environmental Impact of Data Centers

The main purpose of a data center is to protect and store the data, applications, servers and other equipment collocated there. Keeping the data center cool and protected from the elements requires an enormous amount of energy. If the data center operator isn't mindful of its energy use, the company could spend significantly more on energy for cooling than it costs to initially build the center. Consider these facts:

- Data centers are responsible for nearly 1.6% of all energy use in this country.
- It's not uncommon for the majority of the power used in a data center to be expended on cooling and maintaining the data center, with less than half actually used to power the servers themselves.
- The average in-house data center uses only one third of its computing capacity, yet the lifetime cost to power a single server can reach three to four times the capital expenditure made to purchase it.
- Generating more than 40 Million Metric Tons of CO₂ annually, data centers have the same impact on the environment as cars do.
- The US Environmental Protection Agency (EPA) reported that power consumption by data centers doubled between 2000 and 2006 to 61 Billion Kilowatt Hours (kWh).
- According to the US Department of Energy, the growth in electricity consumed by data centers in the US between 2009 and 2010 is forecast to be the equivalent of ten new power plants. Globally, carbon emissions from data centers are projected to quadruple by 2020.

Many organizations are looking for green IT solutions as a cost savings issue or because they want to promote their business as green, currently a key marketing message. By allowing multiple organizations to share a single data center, Cloud Computing reduces that impact in much the same way as car pooling. Just as several people can share a single vehicle to get to their destination for the same cost as one, the economies-of-scale and fuller utilization achieved in a modern energy ef-

icient data center makes it possible to run multiple Cloud servers with the same energy cost as running a single server in-house.

Section III – Choosing an Infrastructure as a Service Solution from the Cloud

When a business decides that it's time to claim the savings and efficiencies available by taking advantage of the Cloud with an Infrastructure as a Service (IaaS) solution, choosing between all of the available products and services such as collocation, managed servers, managed applications, utility computing and outsourced disaster recovery and data backup can be daunting.

It's important to recognize that every business is unique and selecting the right set of Cloud options requires carefully identifying the specific needs, reviewing existing infrastructure, assessing the financial resources that are available, and then making appropriate choices based on that information.

Consider the following services and how they benefit the typical American business.

Colocation Services

Colocation or data center services provide cabinet or floor space for customer-owned servers and other equipment. The allure of colocation for the customer is the level of security offered by the data center. A simple customer-built and managed data center is typically more cost effective, if security isn't a factor, but you get what you pay for. With the capital markets being as tight as they are in this current economy, many companies are turning to colocation services because either they can't afford the cost or they simply can't procure the capital necessary to build the secure data center needed to keep their company data safe. The current political climate that is focusing on green energy only adds cost to an already expensive proposition. Additionally, many industries have specific regulatory requirements that cover the safety that surrounds stored data, including health care and banking.

Colocation is best suited toward larger companies who employ a fully staffed IT department capable

of managing the server infrastructure stored in the colocation space. Colocation provides complete control and an important level of autonomy that many businesses may want or need.

Managed Server

For the company that needs the superior infrastructure of an outsourced data center, but who doesn't have the expertise or staffing to fully support the remote location, a managed server solution may be a better option. Managed server provides either a dedicated or a virtual, partitioned server accessible remotely. This solution proves to be an ideal middle ground for companies that can support their applications and data, yet either do not want to or cannot support the installation and maintenance of a server infrastructure. Managed servers combine seamlessly to the organization's physically located IT infrastructure. The principal advantage is that, unlike colocation, the customer does not need to purchase, upgrade or maintain hardware or manage the OS. This can make a critical difference for a small to medium-sized business with limited staff and resources, since the time and money saved by not having to buy and manage servers can be made available to support the day-to-day operations of the business.

Managed Applications

Taking a step past managed server is managed applications. Managed application solutions outsource the customer's entire computing infrastructure for the specific application (less computers and peripherals) to the cloud. This solution allows a smaller company with few IT resources access to a fully invested IT team to oversee every aspect of its server and applications infrastructure. These solutions can focus on one specific application or the company's entire software database. Additionally, managed application solutions work for both off-the-shelf and custom applications. The solution is ideal for small companies who are spread over a large geographic area and ones whose security needs rival enterprise-level companies.

Utility Computing

Utility Computing is the packaging of computing resources – such as computation and storage – as a metered service similar to a traditional public utility,

such as electricity, water, natural gas or telephone services. The Utility Computing provider essentially "rents" the computation time, storage, etc. charging the customer for its incremental use, and thereby eliminating wasteful overspending during periods of low demand without having to risk not having enough resources during peak demands. Companies who regularly experience project-driven variations in IT workloads or have unpredictable or frequently changing IT demands based on seasonal or other factors need the ability to scale IT resources both up and down. It is costly and inefficient to maintain a computing infrastructure that supports a company's peak loads when the peaks are hit only periodically.

Outsourced Disaster Recovery, Data Storage and Backup

The flexibility of the Cloud Computing allows for companies to pick and choose the services they draw out of the cloud. Automatic, online data backup services allow businesses to off-load the management of collecting, storing and restoring its data saved specifically for times of unforeseen disasters. These services automatically access the customer's network, drawing down the data for storage in the provider's data center and/or other off-site location. This solution supports smaller companies who are not prepared to manage this process which can be complicated in its details.

Choosing the Right IaaS Provider for Your Business

There are a wide variety of IaaS providers to choose, ranging from very large to very small. Begin by comparing the solutions of several providers, each will have variations in cost, support and the specific offering. Next, review their technical capabilities, including the experience of the staff and the quality of the provider's data center. Consider the level of support they provide and whether their SLA includes 100% uptime? Specifically, ask if the data center and the service are SAS-70, Type II certified. Continue by evaluating the company: how long have they been in business, do they have the financial stability to support company growth and research. Finally, ask to speak with current customers.

Conclusion

Cloud Computing isn't new, in fact computer scientists have been experimenting with its concepts since the 1950s. It is only recently that Cloud Computing has matured, mostly due to the quality and ubiquity of data transmissions and introduction of virtualization and the thin client. The maturity of Cloud Computing happened just in the nick of time as this new economy has companies searching for more effective ways to manage IT, data, software applications and computing infrastructure. Depending on the company's size, geographic scope and reliance on IT, Cloud Computing offers a broad spectrum of solutions to meet IT challenges companies face. While a large company seeking to off-load stress from its data center may look to a colocation service, a very small company may use it to act bigger with enhanced data security and more collaborative and mobile IT solutions.

Whatever the solution, customers of the Cloud can all use it to promote the greening of their company. Data centers are energy hogs, yet larger facilities can easily take advantage of ecological economies-of-scale that simply aren't accessible by smaller facilities or even telco closets. As a new administration and focus in Washington takes hold, environmental-friendly will not only come to have more cache, it may become a necessary standard for doing business with the government.