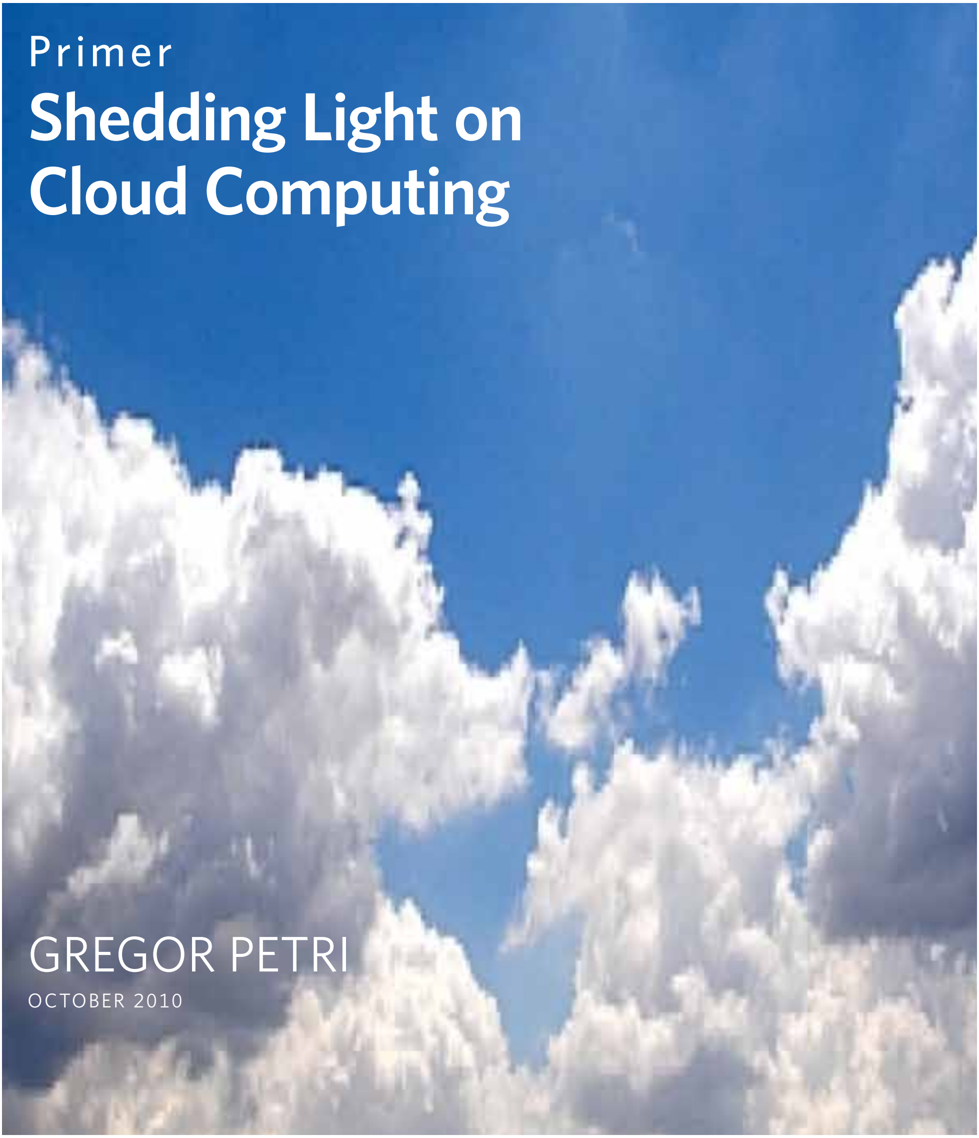
Primer

Shedding Light on Cloud Computing



GREGOR PETRI

OCTOBER 2010

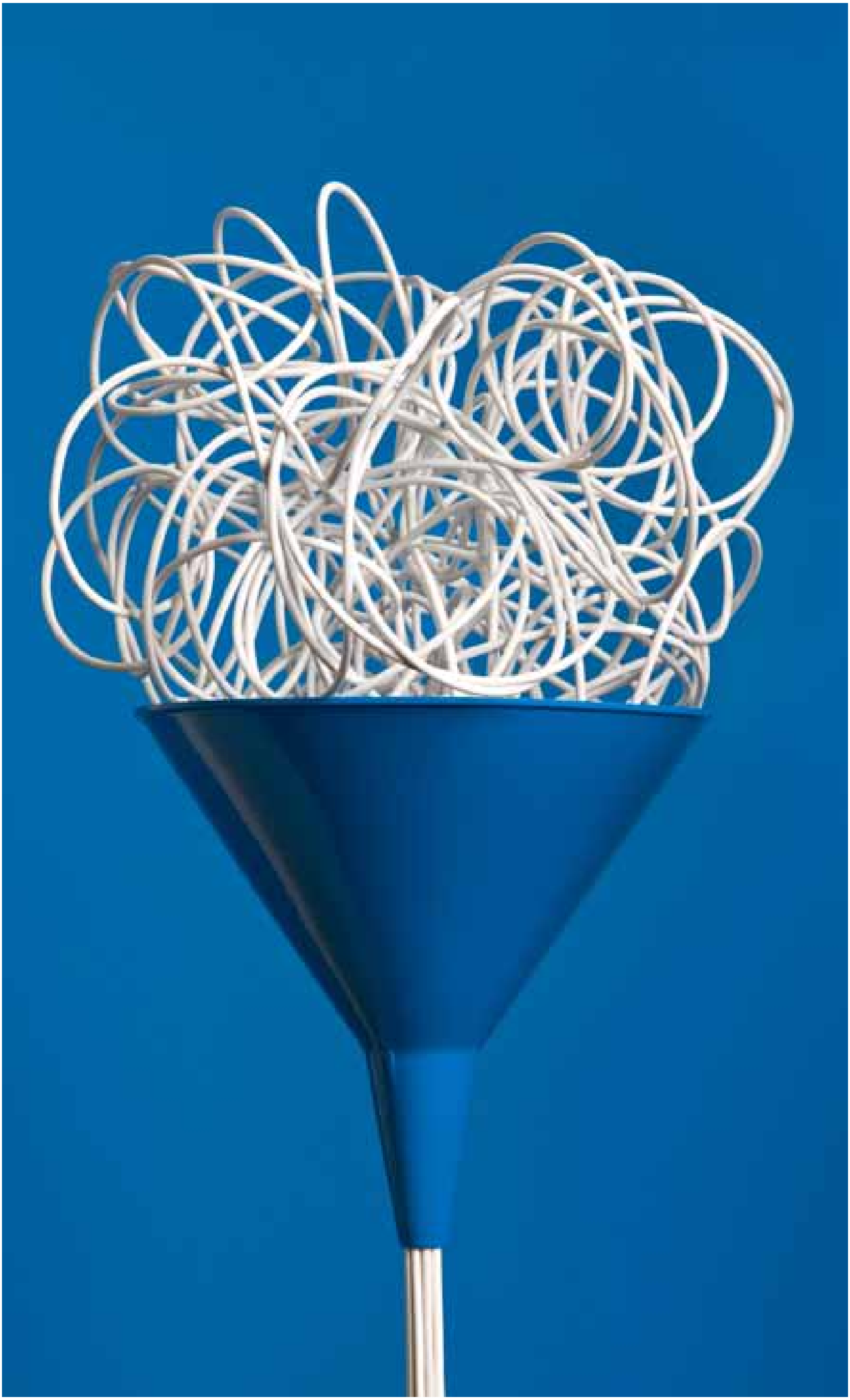
ca technologies

SHEDDING LIGHT ON CLOUD COMPUTING

A Primer
By Gregor Petri

Table of Contents

Note: This primer is also available as free PDF download at www.thecloudacademy.com, as iBook in the Apple iPad bookstore and in ePub format at Smashwords.com for use in a Kindle or other eReader.



THE CLOUD IS THE ANSWER, BUT IT ALSO THE QUESTION

Maybe with the exception of the iPad, Cloud Computing is by far the most discussed IT innovation of recent times. And like the iPad it is enjoying great interest both inside and outside the IT community. Both groups are embarking on a Cloud Computing journey, but from very different starting points. Some see Cloud Computing mainly as a way to make traditional IT more efficient using Infrastructure as a Service and private clouds, while others see it as a way to source solutions "as a service" directly, often bypassing the IT department in the process. Somehow these two groups need to start talking again; otherwise we either get 'strangers passing in the night' or 'a train wreck waiting to happen'.

This primer aims to facilitate this discussion by providing a non-technical, structured overview of what the various types of Cloud Computing are and what the potential benefits and risks are. Risk is consistently cited by CIOs and CFOs as the largest obstacle of cloud computing for their organisations. In this primer we consider the risks in areas such as availability, privacy and regulatory compliance. We also explore why organisations would want to implement cloud computing, to name just a few: cost-savings, increased speed and flexibility and higher performance. Finally we'll look at the impact of Cloud Computing on IT management. Because the cloud, even though it may be the answer, also poses a lot of questions, especially to IT, like how to secure, assure, automate and manage it.

This Primer, which was originally developed for The Cloud Academy, a CA Technologies initiative to further the knowledge around Cloud Computing, can be used by anybody who is affected or will be affected or is simply interested in Cloud Computing.

I hope you enjoy the content. Follow me on twitter @gregorpetri

Welcome to the cloud.

Gregor Petri

Advisor Lean IT & Cloud Computing
CA Technologies

INTRODUCTION

"Cloud" is a collective term for a large number of developments and possibilities. It is not an invention, but more of a "practical innovation", combining several earlier inventions into something new and compelling. Much like the iPod is comprised of several existing concepts and technologies (the Walkman, MP3 compression and a portable hard disk), cloud computing merges several already available technologies: high bandwidth networks, virtualization, Web 2.0 interactivity, time sharing, and browser interfaces.

To understand why there is so much excitement around cloud computing today while the concept is not truly new, it is important to understand that cloud services need very fast and broad networks to work smoothly. Twenty years ago no-one would have conceived of cloud computing because the networks in those days were simply too slow. And even ten years ago no one considered it because every "Enter" or mouse click would have cost $10 in communication costs. But with the global network innovations made over the past few years, these networks are finally ready for primetime, i.e. cloud computing.

In this primer, we will first look at the various definitions of cloud computing and at some of the reasons why organizations would want to implement cloud computing, such as cost-savings, increased speed and flexibility and higher performance.

We will also look at the risks of cloud computing. Risk is consistently cited by CIOs and CFOs as the largest obstacle of cloud computing for their organizations. We will consider the risks in areas such as availability, privacy and regulatory compliance.

Next we delve more deeply into the various types of cloud computing. We will discuss the cloud infrastructure also known as Infrastructure as a Service, and also at the application cloud, also known as Software and Platform as a Service.

When we talk about cloud computing in an infrastructure context, virtualization plays an important role, so we will explore this concept in more depth. In the last part we will explore the impact of cloud computing on organizations, looking forward to what the future holds, along with its impact on IT management.

Cloud Computing Defined

The term cloud computing was first used in an academic context by Prof. Kenneth K Chellapa, who described it in 1997 at the Informs Conference in Dallas as "a computing paradigm where the boundaries of computing will be determined by economic rationale rather than technical limits". This description is clearly broader and less technical than many of the definitions in circulation today. Current definitions, as the one below from Wikipedia all assume some use of the Internet (or at least some Internet technology) to classify something as cloud computing. [1]

Cloud Computing

From Wikipedia, the free encyclopedia
Cloud Computing is a popular phrase that is shorthand for applications that were developed to be rich Internet applications that run on the Internet (or "Cloud"). In the Cloud Computing paradigm, software that is traditionally installed on personal computers is shifted or extended to be accessible via the Internet. These "Cloud applications" or "Cloud apps" utilize massive data centers and powerful servers that host web applications and web services. They can be accessed by anyone with a suitable Internet connection and a standard web browser.

Forrester, Gartner and other analysts all have their own definition, many of which are still evolving. A very pragmatic definition is used by Consulting firm Accenture: "the dynamic provisioning of IT capabilities (hardware, software, or services) from third parties over a network" [10]

The currently most accepted or "official" definition of cloud computing is provided by the North American National Institute for Standard and Technology (NIST), a shortened version is listed below.

The NIST Definition of Cloud Computing (version 15, shortened)
Cloud Computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.

This Cloud model promotes availability and is composed of five essential characteristics (On-demand self-service; Broad network access; Resource pooling; Rapid elasticity; Measured Service), three service models (Cloud Software as a Service (SaaS), Cloud Platform as a Service (PaaS). Cloud Infrastructure as a Service (IaaS) and four deployment models (Private Cloud; Community Cloud; Public Cloud and Hybrid Cloud).

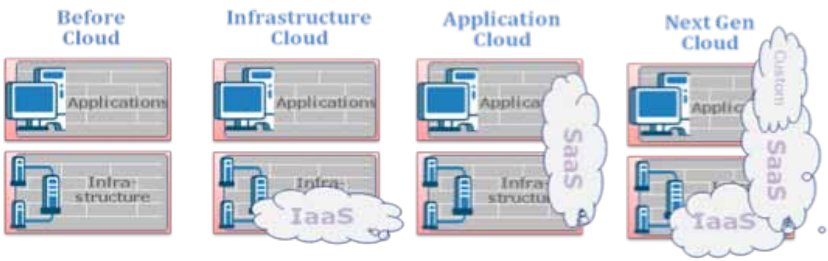
The full NIST definition (version 15) is available at the NIST site and is also included in the appendix of this paper [11]. Even the significantly shortened version is quite a technical mouthful. The objective of this primer is to present an explanation of cloud computing that is a accessible not just to IT personnel, but to the business.

SCENARIOS OF CLOUD COMPUTING

To make the discussion of cloud computing easier we will first look at organizations before cloud computing. Next we will look separately at organizations using infrastructure from the cloud and organizations using applications from the cloud. Separating these makes it easier to understand the various sorts of cloud computing, but it is important to note that organizations will typically have these scenarios in operation at the same time.

In addition many organizations, such as a travel agents, Telco's and banks, will provide some services to their customers over the cloud themselves.

From a management perspective, organizations need to manage all these scenarios in unison. Ideally, they will be able to move IT services transparently across the different models with no hindrance from technical limitations.



Before Cloud

An organization not making use of cloud computing, typically runs its own applications (built or purchased) on a permanently available infrastructure. If an application is running on a specific server today, chances are high that it will be doing the same tomorrow. It can of course be moved, but to do so a 'change request' is needed which has to be approved in advance by a change committee. Applications almost certainly are not assigned dynamically to the server that happens to have the most capacity available. This relatively stable state allows for manual configuration of these servers.

The organization may have outsourced the management of its infrastructure and/or applications, but also at the outsourcer the environment would be relatively stable and dedicated to this customer.

All this stability does not necessarily make this type of environment easy to manage. The inherent complexity of a modern IT environment requires advanced processes, procedures and tools. Often the organization will have turned to best-practice frameworks such as ITIL and COBIT to help govern, manage and secure these large and complex environments.

The Infrastructure Cloud

Often this is called Infrastructure as a Service (IaaS). The original term, first used by Nicholas Carr [7] in 2006 was "Hardware as a Service" (HaaS). With IaaS, organizations source additional infrastructure capacity over the web as a service. Allocating additional storage or processing capacity "over the web" is of course much faster than the supplier bringing and installing new hardware in the basement. The Infrastructure cloud is the domain of IT operations and in most organizations the users are not even aware that their IT department is using cloud services.

Dynamic cloud capacity can cater to unexpectedly large customer demand but also to internal requests such as temporary test servers or an extra SharePoint server for a department's intranet.

The infrastructure allocated to such a request can come from a private cloud (a pool of infrastructure exclusive to your organization) or it can be sourced from an external, public cloud infrastructure provider. Virtualization, in all shapes and sizes, is the most important enabler of this type of cloud computing.

Due to its dynamics, the allocation and de-allocation of capacity is best when fully automated. Often this is done by means of simple scripts, but larger organizations are rapidly turning to more advanced data center automation solutions.

Accurate information about which applications are running on which server and exactly which business processes these applications are supporting, is even more important here than in the 'before cloud' situation-but is also more complex.

IaaS is currently top-of-mind for many IT operation departments. The benefits companies are striving to achieve are increased utilization with reduced capacity requirements, lower cost and lower energy consumption, and also greater scalability and flexibility.

Private versus Public Clouds

With a private cloud we use the same methods and techniques-virtualization and automation-to deploy our applications on dedicated hardware under our direct control.

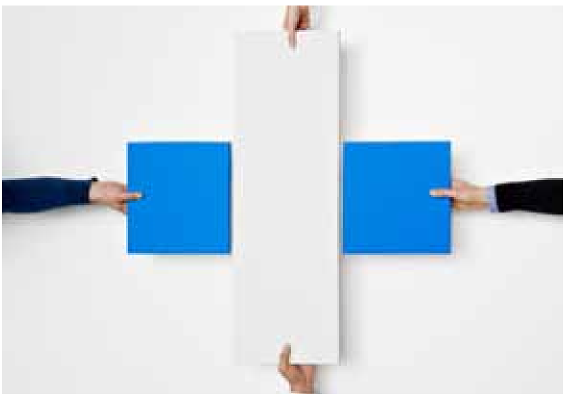
Using a private or dedicated cloud addresses one of the greatest perceived objections against cloud computing: the risk of data theft and service breakdown. It also allows organizations to gain cloud computing experience in a secure and familiar environment. A private cloud typically increases the utilization and lowers the total required capacity. It, however, does not make the cost or the used capacity variable with use as the organization has to make a fixed investment in the used capacity.

The Platform Cloud

Platform as a Service, or PaaS, is the term used to describe software development environments that allow developers to create software applications they can offer as a service to their customers or users. The PaaS platform provider takes care of hosting the created services, for a percentage of the monthly billed revenue, lowering the price of entry to the marketplace for new businesses.

In theory internal IT departments can also use PaaS platforms for building custom applications, but with more and more specialized applications becoming available in the cloud, there is less need to build specific applications for most organizations. Integrating various cloud and non-cloud applications, however, becomes more important. Some newer PaaS platforms are offering these integration capabilities. The term currently used for this is cloud Orchestration. Here the PaaS platform is used to integrate various existing SaaS and non-SaaS applications.

This, in theory, would allow IT departments to compose differentiating and strategic applications or services, by linking together various standard low cost cloud services. Similarly, lean manufacturing enabled Toyota to manufacture thousands of different and exciting car models from standard low cost components (instead of standard black-only Model T-Fords). A PaaS integration platform also prevents lock-in to a particular service provider. The end user or customer only sees the integration or workflow layer, which in principle, makes it possible to transparently replace the underlying cloud services when a business deems it necessary.



The Application Cloud

The Application cloud is also known as Software as a Service, or SaaS. Here we see that organizations do not buy software (on a CD or otherwise) and install it on their own computers. Instead, they simply use their browser to access the software as it is provided over the Internet (software provided as a service).

Today we see organizations deploy SaaS for applications such as Salesforce.com and Google Apps. Other enterprise applications meeting these criteria include Customer Service, HR, Project Management, Web Conferencing, Helpdesks, Wikis, Blogs and other intranet like applications.

With organizations making use of applications from the cloud (or SaaS), it is not uncommon that the IT department is not fully aware of which applications are being used by the organization. Usage of SaaS applications can be hidden from IT. For example, sales might simply charge the use of an application derived from the cloud to their credit card, and it then gets lost along with the myriad of other sales expenses accrued such as client lunches or golf memberships. Also, e-mail and personal productivity cloud services are becoming popular with individual employees because they enable workers to keep in touch easily from their home PC or while on vacation. IT may or may not be aware of the usage of these applications.

But cloud applications are not just being used ‘under the radar’. More and more organizations are consciously opting for selected applications to be operated on Software as a Service basis. Customer Relationship Management (CRM) was one of the first areas to demonstrate that business-critical applications did not necessarily have to operate in-house. What helped the CRM success was that the intended end users (often sales managers) tend not to operate in-house themselves. Also key is that the sales process is less tightly integrated into internal ERP type administrative processes than for example invoicing, purchasing or order in-take.

In the airline industry we already see systems for reserving seats and selling tickets offered “as a service” to multiple airlines, often at a cost of a just few cents per ticket. In fact, very few of today’s low cost carriers run and maintain their own ticketing system because the available SaaS options on the market do it more efficiently and cost effectively.

Until recently a limiting factor was the rudimentary user interface of Web applications compared to graphical PC applications. New browser techniques based on Flash, Ajax, Silverlight and HTML5 are starting to address this challenge.

In general, the term Software as a Service is somewhat misleading, because often the perceived value of the service does not come from the software code, but from the vast amount of content that is delivered as part of the service; content, such as photos of every street (Google Maps), résumés of every potential employee (LinkedIn) or details on all hotels (Expedia). This is content that in-house applications were never able to provide. This content, together with prepackaged implementation guidance and recommended work procedures, further reduces implemention time and increases the total potential value of such SaaS solutions.

This close connection to the business is also a reason why these applications need contextual information, information about the process, the people, and the products - to be able to work effectively. Standardization of technologies, such as networking standards, has proven to be simpler than standardization of company processes and data exchange.

Many believe that the business impact of the application cloud (Software as a Service) will surpass that of the infrastructure cloud by far. Their reasoning is that applications are much closer to users and thus to the business.

Review: Scenarios of Cloud Computing

- **IaaS: Infrastructure as a Service**
  - Domain of IT Operations
  - Allows organizations to source additional capacity over the web as a service Impacts hardware and operating systems
- **PaaS: Platform as a Service**
  - Domain of Application Development
  - Used by aspiring and new cloud vendors
  - Integration capability is biggest promise
- **SaaS: Software as a Service**
  - Domain of users and the business
  - IT department often not aware of usage
  - SaaS also includes content (as a service)
  - Cross application Integration still limited

ADVANTAGES OF CLOUD COMPUTING

Advantages of cloud computing can range from cost reduction to increased flexibility and from decreased complexity to higher added value and thus higher returns on the investment.

Some Simple Cases

Before delving into the underlying sources of value let's list some anonymous but real world examples [2] where cloud computing added significant benefits:

1. Probably the most publicized success story for cloud computing is the New York Times case. The "TimesMachine" project aimed to make their 1922 to 1951 archive available online. But converting the materials into a useable format required a lot more computing capacity than the publisher had imagined and was in a position to procure. As a result, capacity was sourced from Amazon. Through subsequently hosting the document store in the cloud, the NY Times achieved significant additional cost savings versus investing in their own internal capacity expansion.

2. While moving to a new hosting location a national railroad company in Europe consolidated 750 servers into 60 virtualized servers, achieving an unmatched overall virtualization grade of 80%. Measurements during the preparation showed that CPUs were originally running idle on average 96% of the time (in a 24 hour window).

3. A large global provider of IT services built a private cloud environment. This reduced the time needed to make a new server available from forty-six days to thirty-five minutes and they expect to reduce their requirement for capital by 30%.

4. A multinational corporation implemented a CRM as a service solution in a fraction of the time and cost it took to roll out basic ERP financial processes the year before; the number of users involved with the CRM application about twice as high as for the ERP system.

5. A provider of leading edge consumer websites changed its standard deployment method to the cloud. This resulted in an average go live cost of only twenty-five percent of the original cost to their customers.

Financial Advantages

Cloud computing enables the sharing of IT resources, making the total cost of a service variable with use (pay as you go). This enables IT cost to be funded as an expense rather than as an investment.

In general the cost models of cloud computing can be compared to the cost of public transport versus the cost of owning a car. With a public bus or taxi the users only pay a contribution towards the total running costs. If they do not take a trip, there is no cost. When owning a car, one first has to invest, but if one drives more than a certain number of miles (high utilization) the cost typically becomes lower than using taxis, while the level of comfort is still better than typical public transport.

The trick for successful public transport is offering a comparable level of comfort and speed at a significantly lower price. It is the same for cloud computing. If cloud computing felt like using a 1980's green screen and any flexibility (can you stop at the next corner) was ruled out, everyone would just continue to drive their own car. Cloud computing is, however, offering a user experience similar to having your own car, but without the hassle of having to maintain, finance or drive it.

Just like public transport has many hybrids (using private shuttle busses, sharing a car with colleagues on long hauls), cloud computing comes in many forms as well. One example of this is in the form of private and hybrid clouds.

Sharing Cost

IT costs are only partly dependent on usage. An example: regardless of whether we use our PC for three hours or for six hours a day, the total cost is more or less the same. If we, however, share that PC with two people (six hours each) the cost per person suddenly is only half. Such sharing could be inconvenient in the case of the PC, but with a server in the basement it becomes a lot easier.

Servers are already shared by many users, but a typical server only runs one application. Sharing servers between more applications can increase the utilization from a typical fifteen percent to an average of fifty percent or higher. Until now technical limitations, mainly in the form of multiple applications on one server severely impacting each other, prevented IT from sharing servers across multiple applications. And as the cost of hardware was only a small part of the total cost, IT could afford to keep adding vast numbers of individual servers; something now commonly called "server sprawl." One of the main advantages of virtualization is that it isolates applications from impacting each other when running on a shared server. This enables IT to consolidate their machines and increase the sharing of resources and thus their utilization. The impact of this on the total cost of ownership (TCO) is significantly higher than just the reduction in hardware expenses. Namely, the TCO also includes costs of the operating system, database and application licenses and the cost of installing, maintaining, supporting and troubleshooting all these servers. Years ago, sharing capacity with multiple users was widely practiced. The typical shared system was accessed over a slow network without a graphical interface. It was a lot less responsive and a less enjoyable experience than using systems under our own control. Cloud computing, with its fast network, facilities for self service and its rich browser interface removes these technical obstacles to sharing resources and cost (as Kenneth Chelappa intended in his original definition.)

Variable Versus Fixed Cost

Sharing and increasing the utilization of our internal IT Infrastructure reduces the total cost, as we need to buy less total capacity. But it does not yet make the total cost really variable with use. If the total IT department costs sixty million dollars per year to operate, then sixty million will be charged at the end of the year to the user departments- regardless of whether they made much use of the installed systems or not.

This is because IT costs are typically fixed for the year (a fixed number of operators, a fixed number of computers and fixed costs for licenses). Every few years the computers and the licenses will be replaced by new and more efficient ones. But after that replacement the costs are again fixed and not related to daily or weekly usage. Unless of course we only pay for the capacity we actually use.

The first time this variable cost was realized was when Wide Area Networks (WANs) were sourced externally many years ago. Instead of each organization building their own WAN, they procured capacity from a Telco when and where needed. With cloud computing we can now achieve this variable cost for many more IT resources, such as servers, storage and software.

Cloud vendors can of course only make this cost variable if they have other customers who will use and pay for the capacity when we are not using it. Multi-tenancy- multiple customers using the same instance or application- is the key to making this scenario work for the cloud vendor.

OPEX Versus CAPEX

Cloud computing is often promoted using the argument that the cost of cloud computing is an OPEX (operational expenditure) and not a CAPEX (capital expenditure.)

The operational advantage of OPEX is that the decision making process is usually much shorter and less complex than for CAPEX. Risks are lower and easier to anticipate and funding comes from operations and does not have to be financed externally. We should, however, not lose sight of the fact that in most cases someone who buys a house in the longer term is better off than someone who rents a house; apart from avoiding rent increases, a house tends to be yours after 30 years while rent continues for the rest of your life. The question is whether one takes an entrepreneurial approach and invests in IT as a way to reach higher returns, just like one would do in other business resources such as factories or R&D. Or does one source IT just like any other non-primary resource such as catering services or company car schemes. One could say the underlying question, as posed by Nicolas Carr [6] years ago, is: "Do we see IT as a strategic competitive advantage for our industry? Or what part of IT might make a strategic difference? Running the desktops, the network or the servers might not be strategic, but designing friendly applications for end customers might be a strategic differentiator".

But one can also approach this issue of Opex versus Capex more pragmatically. If one could buy and install an application with the certainty that it will be used intensively for the next ten years, then that would be the right capital expenditure to make. But suppose the implementation works out badly and the use is discontinued after two years. In those - not unimaginable - circumstances, sourcing it via the cloud will of course be cheaper and less risky.

Until we get better at predicting the future, it would be useful and pragmatic to be able to start off in the cloud and when necessary or financially more viable bring the projects back in-house.

Supply Side Economics

Cloud computing also offers many cost benefits on the supplier side, which, in a perfect economy, vendors will pass on to their customers. The first benefit to mention is volume discounts. The typical cloud provider will buy infrastructure in very large volume and thu can negotiate much higher discounts than the average end-user organization.

There are also operational and development advantages. The most important operational advantage is that true cloud applications are based on multi-tenancy. All customers make use of the same configuration. Instead of installing and maintaining a server or separate configuration for each customer, all customers use the same one. Patches and bug fixes only need to be applied once, and upgrading to a new release immediately moves all customers to the latest version (eliminating the quite substantial cost of needing to maintain two or three old releases).On the development side the advantage is the number of platforms the vendor has to support. He no longer needs to build and maintain a Windows, Unix and Mainframe version; he offers only a cloud version (using whatever platform is the most efficient for him). Not having to maintain many versions, releases, and platforms makes an enormous cost difference to the supplier.

The flip side of this is a higher need for flexibility. These hundreds of customers will not all want to drive black Model T-Fords, they will want to be able to customize the service to their specific needs. Early cloud application suppliers have invested substantially in building a specific new development platform to do this. Such a platform typically includes a menu system, a user rights administration system, a reporting system and many more items that traditional software vendors had to build before they could start developing their core functionality.

Several of these early cloud providers took their development platform and offered it to other cloud providers as Platform as a Service. New cloud application providers who choose such a Platform as a Service can invest virtually every dollar of their development budget in functionality.

All of the above makes a huge difference compared to traditional software vendors that spend forty percent on sales and marketing, forty percent on maintenance and support and divide the remaining twenty percent across development (including developing their own frameworks and porting their applications to multiple platforms) while still hopefully making some profit.

The cost model for Software as a Service is not comparable to traditional software. With traditional software the customer typically buys the license and pays all the other costs separately. These other costs include hardware, network, storage, operating systems, and installation and support costs. In fact the license cost can be as little as ten percent of the total cost. When providing the application as a service most of these "other costs" are included in the monthly user fee. As a consequence, Software as a Service can be perceived as very expensive, compared to traditional software, because the traditionally hidden or not correctly allocated cost of in-house operators, servers, and storage are not taken into account.

Software as a Service also impacts the traditional business model of software vendors. Once they reach the breakeven point, additional sales lead to high profitability quickly, as traditional software has no cost of manufacturing or cost of raw materials, just the cost of development. This is no longer true for Software as a Service. An additional customer now does require additional hosting (CPU, storage, network capacity etc.) leading to much more traditional (lower) profit expectations.

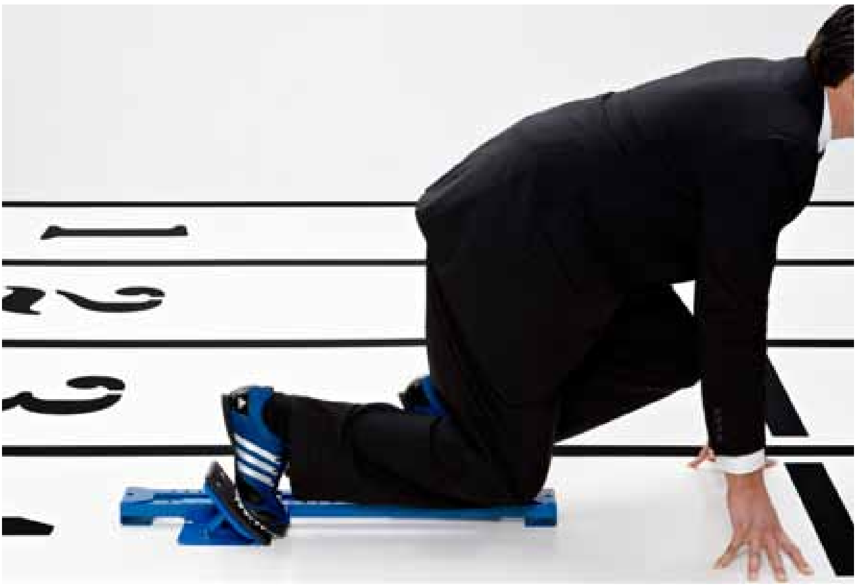
COST:

Time-to-Value

Shorter time to value is maybe the most important reason for organizations to deploy cloud computing, both on the application and the infrastructure side. Many CIOs struggle to explain to their CEO why implementing ERP took years and cost ten times more than the cloud-based CRM application that went live within six months, with a comparable number of users, and a much more significant impact on the business.

Why is Software as a Service often implemented in a fraction of the time required for traditional on-premise applications? One reason may be psychological: “It’s already running in the cloud. So let’s start using it tomorrow or even today.” SaaS is clearly the pragmatic choice when one considers that requesting and provisioning the required in-house resources to start a simple pilot can easily take several months for large organizations.

The same is true for Infrastructure as a Service; getting additional capacity for a large number crunching or data analysis project can take several months. Being able to simply “rent” this capacity for a limited time can be a huge benefit.



Scalability

The more services are offered over the Internet, the more difficult it becomes to predict capacity requirements. Even a company with 100,000 employees can make a reasonable estimate when their users log on to their e-mail system or when they log onto to the BI (Business Intelligence) application to analyze the end-of-month figures.

Once this company offers its applications over the web to a million potential customers to all take part in a promotion or a special offer, estimating the required capacity at any moment in time becomes trickier. The peaks and troughs in required capacity can become a lot more extreme.

Being able to quickly scale and deploy additional servers or storage over the web is an important benefit of the infrastructure cloud. With Software as a Service we also see organizations scale faster from a few users, to thousands of users.

Higher Added Value

Why would cloud applications offer more added value than traditional applications? The main reason is that cloud applications often do not only include software functionality but also information or data- which traditionally the user was expected to create himself. A typical example is LinkedIn, a constantly up-to-date database that offers profiles of virtually every current, former and prospective employee in the world. Many HR departments are now already using systems like LinkedIn, Plaxo and Xing alongside their in-house systems to look up details of their employees because the LinkedIn profiles are often more up to date than the details in their own system. Another example is Expedia, as the average in-house travel department can simply not afford to include every hotel on the globe in its database. The larger scale means cloud providers can invest more in both content and development, while still being able to provide this richer application at equal or even lower cost. The remaining question is whether this richer functionality is important to our use of an application? Do the satellite view and street photos on Google add value or just distract end users?

Review: Advantages of Cloud Computing

- **Financial Advantages**
  - Costs are shared between users/departments
  - Costs become variable
    (as they are shared with other organizations)
  - Cloud costs are an expense not an investment
    (no need to finance / prove ROI upfront)
- **Vendor Advantages**
  - Economies of scale and scope
  - Extreme standardization
  - One size fits all (multi-tenancy)
- **Advantages Beyond cost**
  - Faster time-to-value
  - Scalability
  - Embedded content increases value add

RISKS OF CLOUD COMPUTING

The inherent or perceived risk of cloud computing is likely to be the most restricting factor in its possible success. Risk can occur in areas of availability, privacy, legislation, data theft and user security.

Availability

Concerns about the availability of cloud services are valid: what happens if the cloud supplier closes down tomorrow? On the infrastructure side one could argue that availability should be less of a concern, as we can move workload between different cloud vendors. The standardization that virtualization can bring makes such an approach feasible and practical.

On the application side (Software as a Service) this type of portability, however, is not yet easily available. Organizations must decide per application or area what level of availability is required for them. Just like hospitals typically have an electricity generator in their basement, but local primary schools don't, not all our applications need to be "hot swappable". It is not one size fits all. Various application providers are working on addressing these concerns with high profile efforts such as "SaaS Escrow" where deployable images are kept available at a third party, so they can be easily deployed in case of a calamity.

For smaller organizations the availability of the cloud, i.e. the Internet, may be the greatest concern. A two day Internet breakdown in the north of the Netherlands resulted in a flood of cancellations for a local 'bookkeeping as a service' supplier. Customers collectively ran to the PC store to install a package on their local PC. Larger organizations typically have (or should have) some redundancy built into their Internet access.

Privacy, Legislation

Privacy and legislation are valid concerns that have been far from fully addressed. Some of the aspects that need to be examined in detail for each cloud offering are: the specific terms and conditions offered by the service provider; the flexibility of the arrangements with the service provider; the conditions for exit and termination of the agreement; and the legal and practical implications of moving specific types of data off premise.

Moving data off premise, as many organizations are doing with outsourcing, is one thing. Not having a clear understanding of where the data is (i.e. in which country) is something else. Several cloud providers are starting to offer specific cloud services that address these concerns. For example, Google has committed to providing a cloud environment dedicated for US federal government use. It will cater to specific needs, such as storage inside the US and access restricted to certified Google staff and government employees.

For a readable introduction into the legal implications of cloud computing see the four part article by Tanya Forsheit of the Infolawgroup[7]. The newly formed vendor organization Eurocloud is aiming to streamline some of the required legal and privacy work across European countries and several easy checklists (regarding terms and conditions, licenses, data privacy, ownership of data and logic,etc.) are already becoming available.

Data Theft and Loss

So far most reported 'mega' losses of data were caused by memory sticks and laptops left carelessly on planes, trains and automobiles. In fact many cloud datacenters are physically and procedurally more closed off than many enterprise or government datacenters. But going on national TV with a copy of the SLA (Service Level Agreement) of the cloud provider - who promised this would not happen - is no better remedy than showing the employee handbook, which says one should not use memory sticks offsite.

Customers need to understand the encryption and backup measures the provider is taking. For example many cloud email providers store emails in an encrypted form, so their employees cannot read them. Customers should evaluate these measures on a regular basis and decide whether these precautions are adequate for their needs or not. Forrester recently published an easy one page checklist to do just that[8].

On the other hand, a court order recently caused a public mail provider to close down the mail account of a small entrepreneur, after a bank sent the entrepreneur confidential data by mistake. Here it was not failing technology, but bureaucracy - outside of the service provider's and vendor's control - that caused the problem. Not having a backup – which for email could have been arranged - made this quite awkward for the entrepreneur.

But what if the data is too complex to backup? Suppose we have a project manager who uses a cloud service to plan and monitor the most important project for his company and it is a service he pays for with his credit card. What if somehow the subscription does not get renewed, resulting in the supplier, in line with the published terms and conditions, deleting all the details of this project? Who is in this case liable for any delays the company experiences on this project? How can any such mishaps be anticipated and prevented?

Security First

In addition to concerns around data leaving the company's controlled premises there is also a growing recognition that traditional security management needs some rethinking in a virtualized or cloud environment.

For a while, virtual servers were secured in the same way as physical servers, but this had some practical drawbacks. One drawback was that unlike physical machines, virtual machines typically are not active all the time. Therefore security jobs which used to run in off-peak hours, like a periodic virus or malware scan, now start running as soon as the virtual machine comes online. The same applies to critical updates or patches to the operating system. Having all this security activity start at once can seriously bring down the performance of the overall system.

The fact that virtual systems are only active when needed can also give a false sense of security. The latest compliance scan or report may show all live systems were fully patched and up to date, while all non-active virtual machines were ignored.

With regard to firewalls and port settings cloud introduces some additional significant changes. Traditional physical servers run in or behind a DMZ (Demilitarized Zone) and traditionally security was applied in that context. With virtual servers in a cloud environment the idea is that we can move these virtual servers from inside our own firewall or DMZ to somewhere in the cloud. Security needs to be aware and be able to accommodate this.

A first pragmatic step in this direction was taken by Amazon, who is now offering a VPN option and mandates that all provisioned servers be a logical part of the (Virtual) Private Network of the customer.

Users and Identities

Cloud computing also poses new demands on user management. Just as we allow or deny users access to in-house applications based on their roles and responsibilities, we need this ability in a cloud environment as well.

Software as a Service example: when in-house applications close down a former employee's access to the company intranet and network, access to all internal application is automatically denied. But to be safe, organizations also maintain a central record with roles and responsibilities which notifies all applications that this former employee is no longer allowed any access. How different with the cloud: here the email address is the user ID. Theoretically the former employee could still use his old company's email identity to continue blogging, use social networking sites and maybe even business applications like CRM as a representative of his old company.

Here we need the same approach to user management as we had in-house, preferably offering a single sign-on experience. Single sign-on across cloud services is a good idea anyhow, ensuring greater productivity as users don't have to remember multiple passwords and user IDs.

Administrators and Identities

With regard to Infrastructure as a Service the requirements for Administrator (or Root) security change significantly.

With single physical machines it was a fact of life that the administrator had all the rights over the machine. Many organizations installed some form of access control to prevent the administrator from accidently killing all processes or deleting all users (remove *.* is sometimes just a bit to powerful) or even from viewing all data. In a virtual environment with one "machine" running hundreds or even thousands of virtual servers - often for many different organizations – this becomes mandatory.

In fact many organizations divide between administrators who are allowed to create, move or delete virtual machines and those allowed to access/operate and copy these virtual machines.

Summary

To summarize, risk management is without doubt the area where the cloud draws the most attention. The future is still uncertain when it comes to security in the cloud, and this will need to be addressed by cloud vendors to customer's satisfaction.

When choosing a cloud services provider, one could argue that they should be neither too big, nor too small. If we use, for example, one of the large public e-mail services, all we can do in the event of a breakdown is (along with a few million other users) read the press release notifying us of the mishap, which will be perfectly in line with the published terms and conditions. If, however, we use a niche application from a company with only a few other clients, the only option we have if this supplier gets into difficulty, is to take over the whole set-up including its staff. This may sound farfetched, but has happened several times in the world of traditional software.

The ideal cloud provider should understand and have experience in balancing economies of scale with catering for specific customer demands and warranting continuity. This makes traditional outsourcers and managed service providers more likely candidates as cloud providers for the average enterprise than new mega entrants or niche players.

It is about finding the appropriate levels of security. If a bank has a 24/7 fail-over facility for specific applications, it would be strange if they did not expect the same from their cloud infrastructure or cloud application suppliers for these applications. For other applications the need may be significantly less.

By starting out with a few less business-critical applications, suppliers and consumers can work together towards best practices concerning back-up, recovery, escrow and disaster recovery.

Review: Risks of Cloud Computing

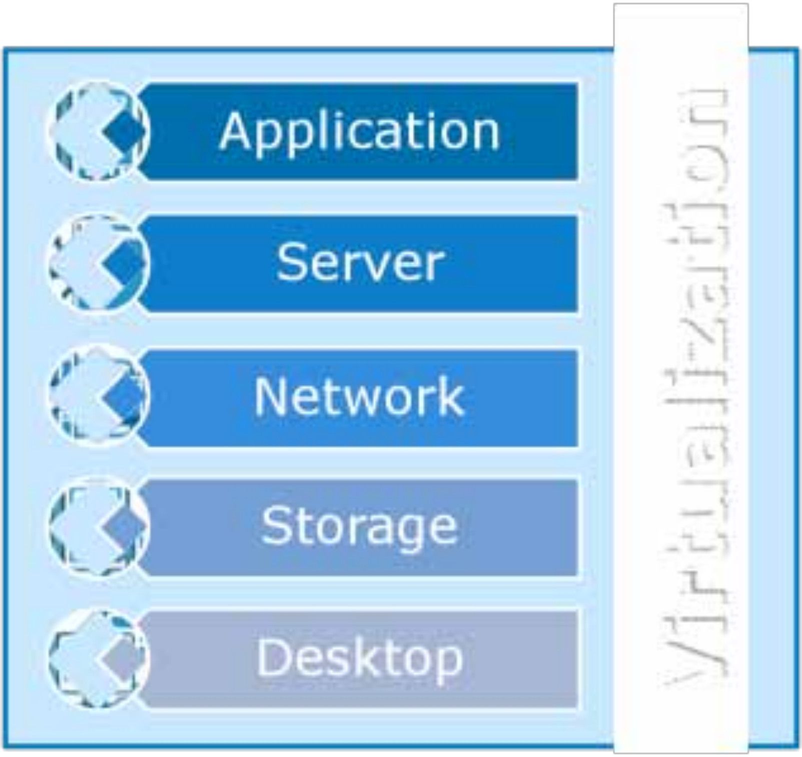
Moving Off Premise Presents Higher Risk

- **Vendors are addressing these by offering:**
  - In continent / in country hosting
  - Private and community clouds
  - Secured IaaS (e.g. using VPN)
  - Cloud specific escrow and failover services
- **Customers are addressing these risks by:**
  - Extending their risk management and including cloud computing in backup, recovery and data protection scenarios and planning.
- **User Management requires rethinking**
  - E.g. employees versus partners and customers
- **Scale of cloud systems**
  - In the cloud any problem is a big problem: one service is used by millions of users, one server runs hundreds of applications.

CLOUD BUILDING BLOCKS:

Virtualization

Virtualization is the key support mechanism to enable Infrastructure as a Service (the Infrastructure cloud.) In the following section, we will briefly discuss the various types of virtualization: network, storage, server, desktop and application virtualization.



Network Virtualization

Today, we see the most widespread use of Infrastructure as a Service in the field of networking, in particular Wide Area Networks (WANs.) Not too long ago most multinational organizations owned and managed their own WAN. This consisted of a vast and expensive network of fixed, leased and dial-up lines that connected the various national and international branches. Sharing the use of an already existing infrastructure with other organizations was found to be much more efficient. Early offerings were based on X25, later on Frame Relay and now increasingly based on standard IP (Internet Protocol.)

The “virtual” part of this movement was that the now rented on-demand network capacity still needed to appear as a separate private network to the customers. To do so it used a “virtualization layer” that behaved as if only the machines in the customer’s offices were connected to the network.

Storage Virtualization

Storage virtualization has been around for years. The simplest example is when someone assumes he has his own drive (D:) but in reality this is just a directory on a larger disk down in the basement. Again, there is a "virtualization layer" that presents part of a larger whole, as a specific dedicated facility to the user.

More advanced use of storage virtualization is what Amazon has been offering since 2002 in the form of the Amazon S3 (Simple Storage Service)[3]. Objects (files, images) can be stored and retrieved using a simple web service interface.

Sites such as Flickr, Slideshare and Twitter now use S3 storage services as well. Some operating systems can even use this S3 service as a backup medium or as a default storage device.

Apple's MobileMe service and Microsoft's equivalent, Skydrive, offer a single disk in the cloud for consumers. With these services, consumers can store their data (emails, pictures, docs, music) in one place with 24/7 access.

Remote virtual storage does imply some changes in how we manage this storage. Traditionally the common way for an application to check whether a file is still available and not corrupted is to open it and read it. With remote storage this would mean transmitting our whole data collection across the network just to be assured it is still there and correct. Several storage vendors are working on a smarter storage API that allows the management application to ask questions like "are you still there and is your checksum still ABC?"

The faster the networks become and the more these storage services meet B2B requirements, this type of storage virtualization is likely to make even greater leaps forward.

Server Virtualization

Server virtualization is currently the most important, and certainly the most discussed type of virtualization. The principle is once again the same: a section of shared physical infrastructure represents itself as a dedicated resource. Today VMware is the best known vendor in this field, but there are several others such as Citrix with Xen in the Unix and Windows environment and also Microsoft is increasingly active in this area.

The Elastic Compute cloud (EC2) [4] service launched by Amazon in 2006 is another service that enables users to rent such virtual servers over the Internet (Infrastructure as a Service.) Quick loading of virtual servers on these “layers” is possible through the use of images, a “kind of backup file” of the whole server. This image file is loaded by the virtualization layer and moments later the virtual server is “available for use.” Physical servers are almost never taken down or used for something else, after being configured. With virtual servers we load and unload these images all the time, based on demand. Thanks to the virtualization layer, we can move such virtual server images easily from a Dell to an HP or IBM server or vice versa. We can even move Linux applications to a mainframe hosting thousands of such Linux images. Originally, virtualization layers added significant performance overhead. But today’s hardware is optimized for running these virtualization layers (also called hypervisors) and the added flexibility far outweighs this now minor overhead.

Application Virtualization

In addition to network, storage and server virtualization there is also application and desktop virtualization. The name application virtualization is somewhat confusing as it refers to something completely different to the cloud applications mentioned earlier.

Application virtualization is all about traditional PC applications working within a so called “virtual box” which allows them to be installed faster. The virtual box including the application is simply loaded as an image. It makes sure that the application does not conflict with other applications, as they are each in their own virtual box. And it does not change the underlying operating systems (no settings in the registry, no DLLs loaded or deleted.)

Desktop Virtualization

Desktop virtualization is especially well known through the efforts of Citrix. In this case the user’s desktop no longer works using the software that is installed on his local desktop machine, but he uses a virtual machine that is installed somewhere on a server (either at the office or in the cloud).

This results in similar advantages in the field of resource sharing. But it also means that we can make use of lighter, less energy-consuming client devices and that our virtual desktop with all its data, applications and settings, can travel with us from one location to another without physically having to take a desktop or laptop.

Sun has been offering such a virtual travelling thin client based desktop for several years. However very few companies made use of them as it required proprietary hardware. Citrix seems more successful, even though the additional license cost is quite high. Users remain resistant to this scenario as they feel their desktop should go home with them in their briefcase.

Automation

Virtualization needs to go hand in hand with automation. A cloud environment implies dynamic scaling capacity based on demand. If we create and configure our virtual machines manually, this process will be too slow. That is where automation comes in.

Automation without virtualization would not work, as the application complexity is prohibitive. The old IT saying is: first organize, then automate. Automating chaos leads to one thing only: automated chaos.

In the case of cloud computing this means that before automating we need to re-structure our applications into a set of independent blocks that can be easily added or removed. Implementing virtualization helps do exactly that.

That is why virtualization and automation have to go hand in hand. Automation releases the benefit of virtualization: dynamic scaling. Implementing virtualization itself reduces the complexity to a level that makes automating feasible.

There are other reasons why automation is good for the business. Apart from cost savings, we want IT to spend more time on the business and less on technology; more time with users and less with plumbing. So for this reason, automation is another crucial cloud benefit.

Review: Cloud Building Blocks

Moving Off Premise Presents Higher Risk

- **Virtualization -Key Enabler Of:**
  - Standardization
  - Scalability
  - Pay per use
- **Automation is required for:**
  - Scale and speed of deployment
  - Dynamics of the environment
  - Cost of deployment
- **Types of Virtualization:**
  - Network: “Buy network capacity on demand”
  - Storage: “Buy storage capacity on demand”
  - Server: “Buy server capacity on demand”
  - Application: “Run Apps were it makes sense”
  - Desktop: “Central desktop travels with user from client to client device”

NEXT GEN CLOUD

With the infrastructure cloud and the application cloud we have discussed the impact of the cloud on a large part of the enterprise computing market. As mentioned earlier, there are two other areas in which IT has invested heavily which will change under the cloud's influence. The first is making desktops, laptops and mobile phones available to employees-something that IT has been doing for years, but which is going to change significantly. The second is the fact that organizations must increasingly offer access to business applications to groups of people other than their own employees.

Over time we expect that enterprises will offer services to their customers and partners directly over the cloud. This is the idea of the extended enterprise; the enterprise is no longer a closed entity in which our employees work exclusively with our own applications. Our employees will use applications provided by supply chain partners while partners and customers may increasingly use applications our organization provides to them. The cloud is the ideal environment to enable this.

At the same time employees no longer use IT technology exclusively for work related purposes. In fact they often use more IT technology and have access to more, newer and faster IT equipment for private and entertainment use, than for work related use.

Over time they will expect and demand that these work seamlessly together, that they can use their fast private laptops and high tech iPhones both privately and at work. The term used for this is Consumerisation.

Consumerisation

Until now the IT department was the provider of equipment such as laptops, mobile phones and sometimes PDAs. The question is to what extent we can expect that, in the future, such personal facilities will be made available by the employer.

The current generation entering the workforce, already has faster phones and cooler laptops than the average IT department is willing or able to supply. A select number of companies have started issuing a personal allowance to their employees to purchase devices of their choice. In that case they are also responsible for their own support. Support from the employer is restricted to some Q&A documents. Technologies such as desktop and application virtualization make this a more feasible approach, as virtualization makes these a lot less dependent on the exact configuration of the underlying hardware.

More and more, we are seeing employees use technical devices such as desktops, applications and mobile devices which they select and own themselves. Not that every product developer comes to the office with his own CAD/CAM system, but many graphic designers already use their own Mac in preference to than anything the IT department may be able to provide him with.

Does this mean less control for the IT department? Certainly, but it also means less work. In this new world order, there will need to be a framework for IT to relinquish control and micro management of employee PCs.

Offering Cloud Services

Cloud services provide clients, partners, suppliers and private individuals with access to applications and services, which were traditionally only available to people employed by an organization. Imagine a sporting goods manufacturer that traditionally ran a planning application to manage its logistical operations. This manufacturer outsourced most of its manufacturing to third parties. It may however still offer this planning application to help the “Extended Enterprise” go to market more efficiently than its competitors.

Or one could think of an electronics manufacturer that differentiates itself with “unmatched customer service.” This manufacturer runs a return-and-repair management system even though 90% of the repairs and returns are handled by third parties.

The traditional idea that internal applications are used primarily by internal staff is no longer sustainable. A number of forward thinking companies therefore no longer provide their business applications exclusively on their own network[9]. They take their applications and offer these on the Internet as a service. They have in effect become cloud service providers, just like a bank offering home banking or a travel agent offering online booking essentially already are providers of services over the cloud.

Review: Next Gen Cloud

- **Extended Enterprise**
  - Partners increasingly expect access to company applications and vice versa
  - Applications will be accessed more from the open Internet than from the company network
- **Customers are addressing these risks by:**
  - Extending their risk management and including cloud computing in backup, recovery and data protection scenarios and planning.
- **Consumerisation**
  - Applications are delivered as a service
  - Users can use standard off-the-shelf devices for accessing these services
  - No specific configuration settings are needed
- **Offering cloud Services**
  - The economy moves from shipping atoms to shipping bits (the information society)
  - These value added information services will be increasingly offered over the cloud using resources (IaaS, PaaS, SaaS) in the cloud

MANAGEMENT ASPECTS OF CLOUD COMPUTING

Manage or Predict?

We know we cannot manage the weather, but merely predict it. Does this also apply to the infrastructure and applications we obtain from the cloud?

Managing a cloud environment is different from a traditional environment. We cannot manage or tune the servers we source from the Amazon Elastic Computing Cloud, neither can we move the CRM application we use as a service to a different server (the service provider takes care of all that). We can move workload between cloud providers based on availability and cost. And we can and should ask the provider for online status information, so our monitoring systems can still alert us about issues before users start calling us.

In most cases the cloud will initially be an additional set of platforms to manage and monitor. For this reason several of the available datacenter automation suites are adding support for Amazon's Elastic cloud and for Acadia VBlocks (very large blocks of virtual servers manufactured by a consortium consisting of Cisco, EMC and VMware) as additional platforms.

A big difference with traditional platforms like Windows, Unix, and Mainframe is that these cloud platforms promise to allow workload to be moved freely around. But was such portability not something that Unix promised us years ago? Back then management suites had to support up to 57 different flavors of UNIX. Let's hope the number of different virtual platforms is more manageable.

ERP for IT?

In a cloud environment, the IT operations manager becomes an IT operations planner. This planning and management process is comparable to distribution or production planning in a large industrial organization. Individual tasks, jobs and workloads are scheduled daily on the basis of a 'master plan'. But the master plan is not static, it changes according to new demand forecasts and because services are added or withdrawn.

Compare this to a company that used to distribute its products with a fleet of owned trucks and directly employed drivers. When this company decides to rent trucks and drivers on a daily or hourly basis or to even use a parcel service to distribute products in certain geographies, they still need to somehow 'plan and manage' their distribution processes. They may no longer own and service trucks, but they do need to know where their goods are at any point in time.

Sensible planning without appropriate information is not possible. Just like a production planner uses the Bill of Material from his ERP system to understand what raw materials or parts are needed, an IT planning system will need a Service Model to determine which applications and resources will be needed. Such a service model would typically be stored in a CMDB/CMS (Configuration Management Database/System). In collaboration with the DTMF (Distributed Management Task Force) a number of suppliers of such IT management systems have initiated a standard (CMDBf) for exchanging information between such systems.

Another interesting parallel is that just like production departments adopted Lean Manufacturing as best practice for efficient manufacturing, more and more IT operations departments are embracing Lean IT as best practice for running their IT operations.

Management Maturity

Recent research by the IT Process Institute ITPI [5] on management of virtualization shows that 72% of the interviewed organizations are engaged in the aggressive implementation of virtualization.

Supporting technologies are important to be able to do so, but it is noteworthy that more than half the companies interviewed at some stage halted their projects to adapt their processes to the more mature standards required by a highly virtual environment. If having good change processes and reliable configuration data in place is important in today's relatively stable datacenters, guess how crucial this will be in a dynamic "provision to order" cloud environment where virtualization and automation enable different virtual resources every day, hour, or even minute.

We all know the stories about IT departments that are afraid of switching off a certain server because they have no idea what it does. Imagine if this was a virtual server that we were paying for by the minute. We better understand exactly which business processes this server is supporting, so we can decide whether it is safe to switch it off or not.

Managing SaaS

We saw earlier that with SaaS, IT may not be aware of which applications the users are deploying as a service. How can IT monitor, manage and secure a portfolio it does not even see? There is no simple answer here.

One thing IT can monitor is the network. Appropriate tooling can - by interpreting the network traffic – report what applications (URLs) are visited and even what the response times are. In the future, using the company network as a source of truth will become less practical as, on the one hand application to application communication will take place directly between virtual machines (bypassing the network), while on the other hand users may no longer be accessing their applications primarily via the corporate network. Increasingly they will use public networks such as their ultrafast fiber connection at home or the free wireless provided at their office campus, local Starbucks, or on commercial airliners.

In a true cloud environment, IT has to get used to the fact that they are less in control. Simply blocking a service will no longer be acceptable. Of course IT can agree to procedures and policies that users are asked to follow. They can also agree with their contracted cloud providers to have real time insight into usage and service levels using standardized reporting API's that feed into the IT management systems. More forensically, IT could "follow the money" by asking the accounting department to report any credit card or other payments to unapproved cloud service providers.

However, a more positive approach, using the carrot instead of the stick, is likely to work better. This would mean a new role for IT, not forbidding or blocking most options, but simply making the use of approved services significantly easier than using non-approved ones. This can be done for example by offering a catalog of prepaid services, all under single sign-on and seamlessly integrated with each other.

A New Role for IT

With SalesForce.com marketing its solutions under the slogan "NO SOFTWARE" and Amazon's Elastic cloud promising: "NO HARDWARE", one may conclude that for the average IT manager this means:"NO JOB!" This is most likely untrue, but the role of IT will change considerably.

Running an efficient factory is important, but the ultimate success of an organization will depend on its product portfolio. The same goes for IT. Running operations efficiently and using virtualization optimally for supplying services, only makes sense if these are the right services.

Choosing the right services may seem trivial, but hundreds of books have already been written on the subject of business and IT alignment. The application cloud with its directly available and prolific functionality, short delivery times and speedy implementations is very attractive for the user organization. But this desire must be weighed against the risks and the integration requirements. If every department runs off and chooses its own set of cloud applications, the end result is far from efficient. How can IT ensure that the end result will be optimal?

The first thing that comes to mind is the management of the services or applications portfolio. IT will need to be involved in selecting which applications and services the organization will use, simply because IT will be held responsible for any associated risks. What if the vendor goes out of business or uses his market power to increase his prices beyond what is reasonable? Vendor lock-in has plagued IT for too long. Let's make sure we prevent cloud lock-in, while we still can. If nothing else, IT will need to provide a disaster recovery and exit strategy for each application. Being responsible for balancing investments and costs against risks and available resources makes Project Portfolio Management ITs' core competency.

Once IT monitors the investment and project portfolio, IT is also best equipped to offer a catalog of available and approved cloud applications (a recent example of this is APPS.GOV, a catalog of pre-approved cloud services for use by any US federal government organization).

And of course IT will be involved in running and implementing these new services. IT's portfolio perspective will help oversee the project and program management of all organizational change efforts. Having a view of both the pipeline of new services and the catalog of available services enables IT to calculate, manage and monitor the integral cost of these services. This is supported by ITIL processes like Service Portfolio and Catalog Management, further defined in the most recent incarnation of ITIL, but basically already foreseen in earlier versions.

Support is another key IT responsibility and it has been a core part of ITIL for years. If the organization uses ten different SaaS applications, does IT want users to go to each individual vendor for support, entering their issues in many different places? Apple reportedly offered 10,000 MacBooks to a corporate client with a recommended support procedure that users visit the Apple Store and line up to talk to the Apple Guru on duty. Would your organization be ready for that? Or do you prefer to keep first line support under one roof, either in-house or by using a Service Desk service?

But IT's responsibilities do not stop here. If the organization uses CRM from one vendor and ERP from another vendor, IT would be expected to connect or integrate these two. In fact "connectability" may need to be the prime criteria for selecting these vendors in the first place. Balancing the need for integration against the risk of vendor lock-in becomes a core capability of IT. Connecting business processes across different (cloud and non-cloud) applications becomes essential and again the cloud, especially in the form of Platforms as a Service, can help.

Review: Management Aspects

- **The Role of IT**
  - Changes significantly
- **Retained Core IT Responsibilities**
  - Application Portfolio Management
  - Risk Assessment and Risk Management
  - Cross Service Integration
  - Cross Service Support and Billing
- **IT Operations**
  - Smart planning and scheduling rather than building a static factory/datacenter
- **IT Service Using Services**
  - IT will procure “services” itself to deliver most of its services

Is Cloud IT a Leaner IT?

Connecting cloud computing and Lean IT may at first glance seem farfetched, but both cloud computing and Lean IT are leveraging the concept that mass produced is almost always cheaper than custom made. The proverbial Ford Model-T was all about using standardization to drive cost down - any color as long as it is black! Cloud computing offers mass produced standardized services to millions of users, and as a result monthly cost per user can be relatively low.

Toyota perfected Lean manufacturing to be able to offer choice at cost comparable to mass production. They did so by using highly standardized components and combining these into unique, desirable automobiles in the last phases of their manufacturing process.

A Lean IT approach can help IT to combine standardized low-cost cloud services into unique, desirable, differentiating and customer relevant services- logical continuation of what IT has been striving for using earlier technologies such as SOA and Object Oriented programming.

It would go too far to explore the relationship between Lean IT and cloud IT here in detail, but one thing IT can and should do today is use the Lean IT mantra of "Maximizing Value" (only do what adds value to the end customer) and "Minimizing Waste" (eliminating steps that do not add value) to guide decisions on cloud computing.

Just take any cloud idea or proposal and evaluate it against these two simple criteria: does the cloud service in question add significant value in the eyes of our end customer? Or, does it minimize waste by eliminating steps that do not add value relevant to our customers?

Conclusion

Reaching conclusions on cloud computing is not easy. To some the whole cloud thing may be a bit overwhelming, providing yet more acronyms and complexity; to others the concept of "the best computer is no computer" sounds very attractive. Although cloud computing combines many existing concepts, it constitutes a fundamental change, comparable to the move from custom made software to standard packages on the application side and as impactful as outsourcing on the operational side.

In general, customers are intrigued by the prospect of using the most economic way to perform computing but not necessarily excited by the idea of moving their data or computing off premise into a cloud environment.

With regard to risk, the most commonly mentioned objection to cloud computing, there are indeed some questions and issues that still need addressing. But the argument that "our systems are so business-critical that we would never risk bringing them under a cloud" is off the mark. If this was the case, companies would never have outsourced or off-shored parts of their operations. It also ignores the enormous investments cloud providers are making to address these still valid concerns.

On the infrastructure side, virtualization and cloud computing add a set of platforms that can significantly increase utilization, scalability and sharing of resources, leading to lower and more variable cost. But in the short term, one should not expect these cloud platforms to replace all current internal platforms, but rather be used in addition.

On the application side, shorter time-to-value and broader functionality at lower initial cost, make Software as a Service very attractive, but the danger of vendor lock-in is also significantly higher here.

The role of IT will certainly change. Technical skills (programming, configuration) will become less important, as maintaining an overview and achieving synergy will become more important. Cloud computing offers too many possibilities and opportunities to ignore, but given the current media hype around cloud computing there is little chance of anyone being able to do so anyhow.

So when and where should your organization start?

Several organizations have started building private clouds. Not as a pilot, but as a first real project. Disentangling existing applications is way too expensive and labor intensive to just have a look; we need to build a business case. Others have started using SaaS applications for some less business critical applications, enjoying the fast time-to-value and generally rich functionality. While for others PaaS may be the fastest way to offer a unique, differentiating portfolio of business services.

Appendix

Cloud Computing

NIST Definition V15[11]

Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. This cloud model promotes availability and is composed of five essential characteristics, three service models, and four deployment models.

On-demand self-service. A consumer can unilaterally provision computing capabilities, such as server time and network storage, as needed automatically without requiring human interaction with each service's provider.

Broad network access. Capabilities are available over the network and accessed through standard mechanisms that promote use by heterogeneous thin or thick client platforms (e.g., mobile phones, laptops, and PDAs).

Resource pooling. The provider's computing resources are pooled to serve multiple consumers using a multi-tenant model, with different physical and virtual resources dynamically assigned and reassigned according to consumer demand. There is a sense of location independence in that the customer generally has no control or knowledge over the exact location of the provided resources but may be able to specify location at a higher level of abstraction (e.g., country, state, or datacenter). Examples of resources include storage, processing, memory, network bandwidth, and virtual machines.

Rapid elasticity. Capabilities can be rapidly and elastically provisioned, in some cases automatically, to quickly scale out and rapidly released to quickly scale in. To the consumer, the capabilities available for provisioning often appear to be unlimited and can be purchased in any quantity at any time.

Measured Service. cloud systems automatically control and optimize resource use by leveraging a metering capability at some level of abstraction appropriate to the type of service (e.g., storage, processing, bandwidth, and active user accounts). Resource usage can be monitored, controlled, and reported providing transparency for both the provider and consumer of the utilized service.

Service Models:

Cloud Software as a Service (SaaS). The capability provided to the consumer is to use the provider's applications running on a cloud infrastructure. The applications are accessible from various client devices through a thin client interface such as a web browser (e.g., web-based email). The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, storage, or even individual application capabilities, with the possible exception of limited user-specific application configuration settings.

Cloud Platform as a Service (PaaS). The capability provided to the consumer is to deploy onto the cloud infrastructure consumer-created or acquired applications created using programming languages and tools supported by the provider. The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, or storage, but has control over the deployed applications and possibly application hosting environment configurations.

Cloud Infrastructure as a Service (IaaS). The capability provided to the consumer is to provision processing, storage, networks, and other fundamental computing resources where the consumer is able to deploy and run arbitrary software, which can include operating systems and applications. The consumer does not manage or control the underlying cloud infrastructure but has control over operating systems, storage, deployed applications, and possibly limited control of select networking components (e.g., host firewalls).

Deployment Models:

Private cloud. The cloud infrastructure is operated solely for an organization. It may be managed by the organization or a third party and may exist on premise or off premise.

Community cloud. The cloud infrastructure is shared by several organizations and supports a specific community that has shared concerns (e.g., mission, security requirements, policy, and compliance considerations). It may be managed by the organizations or a third party and may exist on premise or off premise.

Public cloud. The cloud infrastructure is made available to the general public or a large industry group and is owned by an organization selling cloud services.

Hybrid cloud. The cloud infrastructure is a composition of two or more clouds (private, community, or public) that remain unique entities but are bound together by standardized or proprietary technology that enables data and application portability (e.g., cloud bursting for load-balancing between clouds).

Footnotes:

1. http://en.wikipedia.org/wiki/cloud_computing#cite_note-Informs-6

2. Source: Various magazine articles and case studies.

3. http://en.wikipedia.org/wiki/Amazon_Simple_Storage_Service.

4. http://en.wikipedia.org/wiki/Amazon_Elastic_Compute_Cloud

5. http://www.ca.com/us/press/release.aspx?cid=217789

6. In his earlier article, "IT doesn't matter" Carr compared companies with high and low investment in IT and concluded that there is no correlation between IT investment and the results. His conclusion is that with the introduction of a grid (such as the Internet or now generally called the cloud) the CIO should (have to) leave the building. In the same way that the Chief Electricity Officer left with the introduction of the electric power grid (electricity delivered by the power company instead of from its own power plant). Leading on from the electricity example, Carr gave three guidelines – new rules – in this article for investing in IT: "Spend less", "Follow – Don't lead" and "Focus on vulnerabilities, not opportunities".

7. http://www.infolawgroup.com/2009/11/articles/cloud-computing-1/legal-implications-of-cloud-computing-part-four-ediscovery-and-digital-evidence/

8. Cloud Computing Checklist: How Secure Is Your cloud? October 30, 2009 - http://www.forrester.com/go?docid=55453

9. http://www.opengroup.org/jericho/

10. http://www.accenture.com/Global/Services/Accenture_Technology_Labs/R_and_I/CloudComputing.htm

11. http://csrc.nist.gov/groups/SNS/cloud-computing/

CA, one of the world's largest information technology (IT) management software companies, unifies and simplifies the management of enterprise-wide IT for greater business results. Our vision, tools and expertise help customers manage risk, improve service, manage costs and align their IT investments with their business needs.