

Must-Haves For Cloud Computing: 7 Checkpoints For Success



Title: Must-Haves For Cloud Computing:
7 Checkpoints For Success
Author(s): Xtravirt (Paul Davey)
Target Audience: Technology - All
Current Revision: 1.0 (Sept 2009)
First Published: Sept 2009
Product(s): Cloud Computing
UID: XD10128

Content Overview:

- Cloud Computing Overview
- The Cloud Proposition
- The Cloud Jigsaw - 7 key checkpoints
- Case Studies

1.0 Overview

In today's economic climate, CIO's are focussed more than ever on opportunities to improve IT service levels for their business whilst facing restrictions from budget cuts and headcount reductions. However, where there are challenges there are also opportunities. Cloud computing introduces a new methodology which allows businesses to scale out IT services on demand. This provides businesses the ability to meet increasing resource demand, whilst offsetting capital-intensive costs for deploying new services.

There has been a recent explosion of cloud service companies as world markets start to incorporate cloud computing into their near- and long-term strategies, and it can be difficult to sift through vendor offerings to determine which are the required technologies for cloud enablement. Companies such as Univa UD, who have a long history of success in distributed and high performance computing, are a mature example of this paradigm; they are a respected vendor who has shifted successfully into the cloud market with offerings built on proven products that address cloud computing requirements.

This paper discusses the cloud computing proposition and focuses on 7 key checkpoints for any business adopting a cloud strategy. It also demonstrates real world insight of an already successful cloud solution provider, and offers concrete steps to those wishing to move forward with a cloud strategy.

2.0 Cloud Computing Overview

Cloud computing is an IT infrastructure resource and delivery solution where computing resources are provided over the internet. The key elements of any cloud environment are generally agreed to be:

- Boundless support for applications
- The pooling of resources that can then be shared
- Support for a virtualised infrastructure
- A services approach to application delivery
- A metered, pay-per-use charging model

Depending on the underlying technology of the cloud offering, sophisticated features may include the facility for resources to dynamically grow or shrink on demand, or the ability to create Service Level Agreements (SLAs) based on application level requirements.

The inherent level of transparency in a cloud environment means that businesses do not necessarily need to know what constitutes the infrastructure, or to some extent the geographic location of that infrastructure; rather, they are more concerned that appropriate SLAs are in place to provide equivalency of protection as if the customer still operated the systems internally.

Commoditised cloud components include SaaS (Software as a Service), PaaS (Platform as a Service) and IaaS (Infrastructure as a Service).

Online resources for the virtualization industry

We simplify complex concepts and make them valuable, practical assets. Free white papers, articles, how-to-guides, e-books, presentation tools and much more.

- **SaaS:** "Software as a Service" is a model of software deployment whereby a provider licenses an application to customers for use as a service on demand. SaaS software vendors may host the application on their own web servers and deliver it over the Internet or download the application to the consumer device, disabling it after use or after the on-demand contract expires. The on-demand function may be handled internally to share licenses within a firm or by a third-party application service provider (ASP) sharing licenses between firms^[1]. This is typically the most easily understood variant of cloud, since many individuals today utilise SaaS offerings whether they realise it or not. Examples include Salesforce.com, Facebook, Google Analytics, or any webmail application.
 - **PaaS:** "Platform as a service" is the delivery of a computing platform and solution stack as a service. It facilitates deployment of applications without the cost and complexity of buying and managing the underlying hardware and software layers. This provides all of the facilities required to support the complete lifecycle to build and deliver web applications and services entirely available from the Internet^[2], with no software downloads or installation for developers, IT managers or end-users. It's also known as cloudware. PaaS could be the new acronym that defines a web-oriented model where more than just specific vertical services are delivered as SaaS (e.g. CRM, ERP, etc). Examples include Google AppEngine and Salesforce.com's Force.com.
 - **IaaS:** "Infrastructure as a Service" is the delivery of computer infrastructure (typically a platform virtualisation environment) as a service. These 'virtual infrastructure stacks' are an example of the "everything-as-a-service" trend and shares many of the common characteristics. Rather than purchasing servers, software, data centre space or network equipment, clients instead purchase resources as a fully outsourced service. The service is typically billed on a utility computing basis and amount of resources consumed (and therefore the cost) will typically reflect the level of activity. It is an evolution of web hosting and virtual private server offerings^[3]. Examples include Amazon EC2, Amazon S3 and GoGrid.
- These commoditised cloud components are used to form three models of cloud; Private, Hybrid and Public.
- **Private (Internal) Cloud:** describes a private network within a business's data centre which provides cloud-like services, meeting the five cloud criteria (boundless applications, pooled set of resources, virtualised environment, services approach, metering for usage) within a secure internal environment. Private clouds may require a higher initial investment and ongoing costs due to up front purchase of services and systems, and which also require continual maintenance^[4]. Early adopters for private clouds therefore typically include large organisations who already operate substantial data centres, as well as service providers who need to power an external cloud offering by cloud-enabling their internal environment. An example of this could be an internal virtual server infrastructure which operates an internal chargeback model based on usage.
 - **Public Cloud:** A public (or external) cloud is a hosted resource provided by a 3rd party provider (such as Univa or Amazon EC2). Companies utilise services offered by the public cloud provider, and pay for only what they use^[5]. This scenario

is attractive to smaller organisations who may not have the means to support a private cloud internally, or larger organizations looking for a solution to meet occasional peak demand.

- **Hybrid Cloud:** Hybrid describes multi-connected private clouds, or a combination of private and public^[6]. In this scenario, a company enabled their private cloud environment to burst on demand into a public cloud service.

1. http://en.wikipedia.org/wiki/Software_as_a_service
2. http://en.wikipedia.org/wiki/Platform_as_a_service
3. http://en.wikipedia.org/wiki/Infrastructure_as_a_service
4. <http://www.univaud.com/about-cloud/internal-private.php>
5. <http://www.univaud.com/about-cloud/external-public.php>
6. <http://www.univaud.com/about-cloud/mixed-hybrid.php>

3.0 The Cloud Proposition

3.1 Why Go Cloud?

Traditional IT infrastructures are comprised of internally centralised computing resource purchased through capital expenditure (CapEx). Industry data has long shown that IT infrastructure resources (servers, storage, and networking) have been considerably underutilised and over-specified to handle predicted growth. This inefficiency reduces a business's return on investment (ROI).

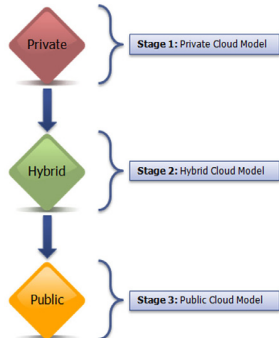


Fig 1-1 Cloud model progression

By utilising cloud based computing, no upfront hardware purchase is required as consumers only pay for resource consumed, or, in a private scenario, they use the resources they already had in place. This allows businesses a straightforward and cost efficient entry point into scalable computing platforms and applications whilst lowering management overhead. A Service Level Agreement (SLA) to ensure a guaranteed level of uptime and accessibility is part of a master agreement between the business and cloud provider (the "provider" can be a 3rd party or an internal IT team).

Where spikes in resource occur, allocated resource can be manually or dynamically expanded, providing a highly efficient cost model. It is likely that many companies will initially implement a private cloud model to ensure the best ROI is realised from existing hardware and software platforms. This model can naturally develop over time from a hybrid model to a public cloud model.

3.2 Financial Speak

Cloud computing introduces the 'pay-as-you-use' cost model, whereby customers pay only for resources they actually use. For some companies this offers significant advantages over traditional in-house purchasing of hardware and management. For example, if a typical server costs £3,500 and utilisation is 20%, then 80% of the server value (nearly £3000) is not realised. Imagine multiplying this by thousands of servers. The cloud model enables businesses to quickly introduce new technology that meets business requirements, whilst negating the requirement for high upfront costs.

Low initial investment for introducing the service together with price points in line with resource requirement allows a business to match their cash flow with the cost of the system. Companies can incorporate a chargeback model for private cloud resources, billing business units for the resource that is being utilised.

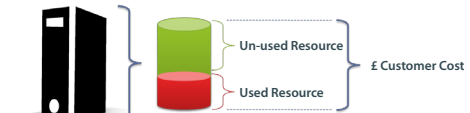


Fig 1-2 CapEx Purchasing

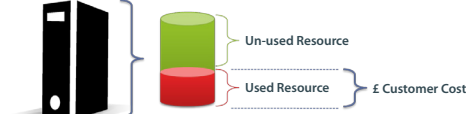


Fig 1-3 Pay-as-you-use

3.3 Instant Satisfaction

An essential part of cloud computing is the ability to dynamically expand computing resource on-demand. Traditionally, customers may have purchased highly specified servers to ensure demand can be met at peak times. More recently, virtualisation has taken a step further by enabling a pool of computing resource for rapid scalability to combat bottlenecks. Where additional resource is needed, it can quickly and automatically be added to a resource pool ensuring that applications and platforms run efficiently regardless of peaks and troughs in demand.

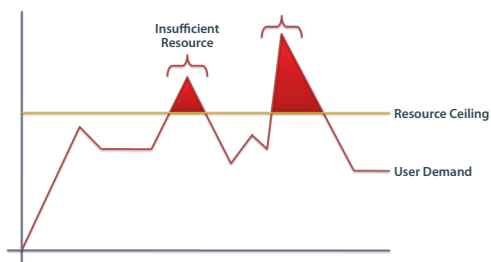


Fig 1-4 Resource Starvation

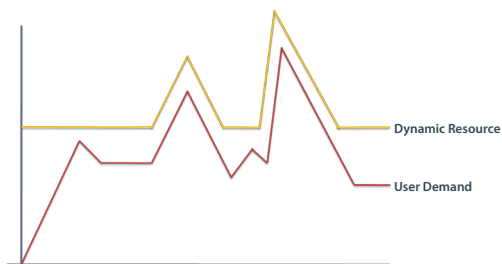


Fig 1-5 Dynamic Cloud Resourcing on-Demand

4.0 The Cloud Jigsaw

Successful implementation of a cloud strategy requires that the customer and cloud provider work in close partnership. There are many considerations which a business needs to address, and these can be summarised into the following 7 key considerations, which can be used as checkpoints for success of any cloud venture:

- Security
- SLA/OLA
- Service Governance
- Performance
- Licensing
- Backup and Disaster Recovery
- Compliance and Data Ownership

4.1 Security

Security is often cited as the greatest perceived barrier to public cloud computing, although public cloud vendors and their technology partners are addressing this successfully with offerings like Cloud VPN.

Whereas enterprises may be familiar with managing their own data security, cloud computing solutions may require joint management, and the following aspects should be considered:

- Ask the cloud provider for a list of people with privileged access to your data
- Query if there have there been any security breaches with the cloud provider in the past, and

if so, what the nature of these were

- Request security auditing be carried out on data access and request a copy of these reports
- Consider the use of data encryption, query if it is available and who would have the ability to decrypt it
- Confirm if there is any data which cannot be maintained by a 3rd party provider for security and compliance reasons even if it is encrypted
- Discuss operating procedures should a security breach occur

4.2 SLA/OLA

The outsourcing of data, platforms or applications, either in part or full, should always contain agreed and documented Service Level Agreements (SLAs) as well as Operational Level Agreements (OLAs) which address:

- Minimum and maximum service levels across all tiers of service provisioning. This should include:
 - Agreed SLAs for planned and unplanned downtime
 - The process in which the cloud provider will notify the company of planned downtime, and mechanisms to accept or defer
 - Contractual penalties for any unplanned downtime suffered outside of the agreed SLA
 - Agreements for events which the cloud provider has no control over. For example, natural disasters at the cloud provider's data centre
- OLAs which record engagement details between the cloud provider support teams including contact details during business and non business hours
- Definition of individual support tiers contained within OLAs, including individual responsibilities for service, process and delivery timeframes

In addition to the above, service providers should be able to offer SLAs based on application and user requirements. These may include:

- Application response time
- Application availability
- Issue resolution

Typically an application service governance component is required to enable this (refer to section 4.3 Service Governance).

4.3 Service Governance

Service Governance comprises the technologies that manage the configuration policies of the cloud, providing automated provisioning actions across physical and virtual resources that ensure configuration policies are adhered to. Sometimes referred to as the “brain of the cloud,” this component provides the infrastructure and application SLA management, an essential piece of a cloud environment.

Key elements required from any service governance tool include:

- Customer inputs, including business priorities, IT service descriptions, service quality, application performance levels, and cost policies
- Assessment of real-time data feeds from a range of 3rd party systems, and in Univa’s case, automated decisions are made within the infrastructure to ensure policy consistency
- Meta-resource management which is enabled by service governance tools and federates local resource provisioning and management tools so they can be grouped into a common, larger shared pool of resources
- Contention management to manage the prioritisation of SLAs and workloads

Key benefits include:

- Service provision optimisation
- Higher levels of application performance
- Reduced costs
- The ability to offer performance guarantees as a service differentiator

4.4 Performance

Performance plays a key role in cloud architecture. Whilst performance of the server instances themselves should be monitored as of right, there are other aspects to consider such as:

- Minimum levels of performance (outlined in the SLA) should be agreed, and guarantees provided via service governance
- The availability of real-time performance metrics of the cloud infrastructure. This is useful for the customer to monitor and react to capacity thresholds
- Geographic data placement in a cloud may be significant to some businesses. For example, a customer with an EMEA client base and their cloud services sited in the United States may be exposed to increased latencies; however these issues can

be normally be alleviated with service based SLAs.

- Connection speeds and latencies to the cloud will affect the perceived performance of the cloud. This will be a function of the connection speed at both the cloud and the business end. Any bottleneck in a VPN connection to/from the cloud for example will reduce the perceived performance

4.5 Licensing

Various software licensing methodologies are available. Options are typically dependent upon the cloud provider, type of cloud and the service required.

Key points to consider are:

- Ultimately, licensing compliance is the responsibility of the customer
- Given the dynamic nature of cloud infrastructures, confirm who is responsible for operating system licenses and their management

As the rate of adoption for cloud services continues to grow, the trend for software vendors will be to update their licensing models to ensure easier compliance for customers and cloud providers. It should be noted that Microsoft has already modified their licensing model to allow customers to take better advantage of Microsoft technologies on cloud platforms.

4.6 Backup and Disaster Recovery

Backup and restore operations are a common challenge for system administrators. Considerations when applying this to cloud based services should include:

- Confirm how the data is to be protected and whether the cloud provider can supply any additional features and functionality not previously available to the customer
- Define schedules, processes and SLAs for data and service restoration
- Define disaster recovery service priorities, staging sequencing and responses to partial recovery scenarios
- Confirm processes and frequency of disaster recovery testing

Note: data archival and lifecycle management are also typically associated components. Confirm with the cloud provider how they can support and integrate these functions and strategies where relevant.

Any business consumer of cloud services should have confidence that their cloud provider can restore their data in the event of a catastrophic disaster at the cloud

providers' site. These requirements should be outlined in the SLA and OLA documents between the provider and customer.

4.7 Compliance and Data Ownership

Cloud computing models can transcend national and geographic borders. Where this occurs, consideration should be given to the following:

- Complexities which can arise for businesses that are bound by regulation to store data in specific geographical locations. Companies can address this issue by ensuring agreements are in place to keep their data inside stated geographical boundaries within the cloud
- Security considerations which may also affect the ability to store sensitive or protected data with a 3rd party. Ensure that steps are taken to have these aspects sanctioned by the business and any relevant regulatory bodies

Whilst the storage of data and access is the undertaking of the cloud provider, the customer remains the data owner, and therefore responsible for ensuring legislation compliance.

5.0 A Real World Example – Univa UD

5.1 Overview

For over a decade, Univa UD, a global cloud solution provider, has delivered application and infrastructure enablement software to increase the efficiency of customers' IT environments while reducing complexity and time to value. A provider of private, hybrid and public cloud software with a global presence, Univa have created 'enablement software' facilitating their customers to transition to a cloud computing model without the negative 'rip out and replace' approach.

Univa outlines some of their key differentiators as follows:

- **Extensive solution offerings.** Univa provides a comprehensive range of solutions including:
 - Supply of both infrastructure management and application service governance products that can be used in all types of cloud. In particular, their Reliance product is the only available product that addresses all 3 key capabilities for service governance: SLA automation and management, contention management, and meta-resource management.

- Univa products support all 3 mainstream cloud scenarios - private, public and hybrid clouds
- **Application and SLA focus.** Univa's Reliance is one of the few available commercial products to enable application-centric service governance, making intelligent decisions about resource provisioning and allocation based on the needs of the application or SLA.
- **Heritage.** Univa UD was one of the first cloud providers, and their experience and reliability has been proven through hundreds of global deployments^[7].

7. <http://www.univaud.com/about/resources/case-studies.php>

6.0 Case Studies

6.1 Case Study 1: A Private Cloud Implementation - Corus

The Customer

Corus is Europe's second largest steel producer with annual revenues of around £12 billion and a crude steel production of over 20 million tonnes, primarily in the UK and the Netherlands. Corus is a subsidiary of Tata Steel, the world's sixth largest steel producer. With a combined presence in nearly 50 countries, the Tata Steel Group including Corus, Tata Steel Thailand and NatSteel Asia, has approximately 80,000 employees across five continents and a crude steel production capacity of over 28 million tonnes.

The Requirement

Corus sought a solution for their Automotive Engineering team to create an internal environment for more easily submitting, executing, tracking, and monitoring HPC workload; thereby providing their users all necessary functionality to run HPC applications.

The Solution

Univa UD was selected to provide the solution, and implemented a private cloud model for Corus utilising the Univa UniCluster and UniPortal products. The UniCluster product, comprised of multiple application tiers into a single stack, is a systems management product that enables the high performance computing environment required by Corus in order to run their high-throughput engineering applications. The UniPortal product provided Corus with a single point of entry through a simplified web dashboard to manage

their cloud infrastructure. The dashboard enabled users to manage workloads as well as automate repetitive daily tasks. The solution also bundled Sun Grid Engine for workload management and Oracle Enterprise Linux.

Solution Support

Univa UD acts as the single support entity for the private cloud. They offer support for all of the systems and tiers of the cloud stack, including the components from Sun and Oracle, simplifying usability and ongoing maintenance for Corus.

Solution Benefits

Through partnership with Univa and the adoption of a cloud model, Corus realised the following benefits:

- Flexibility and usability at enterprise scale
- Elasticity of resources ensuring peak computing demands are met
- Running the UniCluster and UniPortal products on Oracle Enterprise Linux providing a bulletproof HPC cluster solution with a single path for all support issues.

6.2 Case Study 2: A Public Cloud

Implementation - Pathwork Diagnostics

The Customer

Pathwork Diagnostics is a molecular diagnostics company that develops high-value diagnostic tests to assist oncologists in determining treatments for cancer care. Using proprietary technologies, Pathwork utilises the power of genomics information to develop FDA-cleared microarray based tests that provide physicians with clinically actionable information unavailable using alternative methods, so that standard-of-care, cancer-specific treatment can begin.

The Requirement

Pathwork analyses large libraries of gene expression profiles from tumor specimens and selects optimal models for testing. Such analysis requires massive amounts of computational power at certain peak times. Since the requirement for this computational power occurred only at periodically, Pathwork required a solution which avoided further investment in on-site resources.

The Solution

Pathwork turned to public cloud services and choose

Amazon Elastic Compute Cloud (EC2) for pay-per-use computing. They then selected Univa UD's UniCloud product to build high-performance computing (HPC) clusters in the EC2 cloud in order to perform their compute-intensive research.

Solution Benefits

- Saved hundreds of thousands of dollars by avoiding the expenses of purchasing and maintaining HPC hardware in-house
- Satisfied peak demands in computational capacity without purchasing new hardware
- Enabled key research innovations which would otherwise have been unfeasible

7.0 Summary & Recommendations

This paper discussed the cloud value proposition, available cloud models, key considerations and case studies which can be used to form the basis of consideration for any next steps. The information has been written and presented in a format which abstracts industry propaganda to allow businesses to make informed factual decisions.

The 7 key checkpoints for success are tangible assets which can be applied to any business contemplating adoption of cloud based services. These are:

- Security
- SLA/VOLA
- Service Governance
- Performance
- Licensing
- Backup and Disaster Recovery
- Compliance and Data Ownership

Strategists and decision makers will recognise that cloud can offer substantial advantages such as pay-per-use and scale on-demand, but also appreciate that such a platform touches many aspects of IT and business, and therefore robust strategy and planning are at the forefront of any cloud project.

Selecting a cloud partner to work with is a difficult choice, and a clear set of evaluation criteria should be established in order to validate the decisions made. These could include:

- Credentials and experience. A provider who has been a supplier of distributed computing solutions for a number of years should be well

placed to demonstrate history and progression in the rapidly moving cloud industry

- Case studies. Proof of tangible benefits for a range of clients in the providers target market will illustrate credibility
- Leadership and authority. A prospective partner who dedicates investment to research and development of cloud technologies can demonstrate a greater level of integrity and grass roots awareness

Univa and Amazon are two examples of experienced cloud partners mentioned in this paper who tick these boxes, and their credentials substantiate their positions as leaders in their field.

Cloud is no longer an empty buzz word – it is a concrete reality. The significant investment by key participants such as Microsoft, VMware, Cisco, Google and Amazon clearly demonstrates and underpins a trend for future growth and mainstream adoption. This points to a bright future and a silver lining on cloud.

Univa UD

For more information on cloud computing Univa UD can be contacted as follows:

EMEA

Patrick Malaperiman
Email: patrickm@univa.com
Tel +44 1189 242744



UNIVA UD

USA

Chicago
Univa, Inc.
1001 Warrenville Road
Suite 550
Lisle, IL 60532
USA
Tel: +1 630.563.8600
Fax: +1 630.563.8601

Austin

Univa, Inc.
9737 Great Hills Trail
Suite 300
Austin, TX 78759
USA
Tel: +1 800.370.5320
Fax: +1 512.331.6235

Web: <http://www.univaud.com>

About Xtravirt

Xtravirt is a knowledge-based company that delivers its expertise in virtualization online and in person. We have developed a reputation for astute leadership and expertise through our work with an impressive array of organisations. It is this real-world experience that drives our ability to provide independent, current and free advice online.

We work with organisations whose IT staff are frustrated with how hard it is to find detailed information and skills around virtualization. We help our clients deliver the true benefits of virtualization, resulting in cost and time savings.

For more information contact:

Dorset House, Regent Park
297 Kingston Road, Leatherhead
Surrey KT22 7PL
t +44 (0) 1372 824 296
f +44 (0) 1372 824 576
e information@xtravirt.com
w www.xtravirt.com

© Copyright 2009 Xtravirt Ltd. All rights reserved. The information contained herein is subject to change without notice. Xtravirt Ltd shall not be liable for technical or editorial errors or omissions contained herein. Xtravirt and the Xtravirt logo are registered trademarks of Xtravirt Ltd. The names of actual companies and products mentioned herein may be the trademarks or registered trademarks of their respective owners.

This document remains the property of Xtravirt Ltd. Contents may not be copied, reproduced or modified without written consent from Xtravirt Ltd.

References

1. Nil

Tags

cloud, hybrid, private, public, saas, paas, iaas, amazon ec2, univa ud

Useful Links

1. Beginners Guide to Cloud, <http://xtravirt.com/xd10107>
2. Microsoft Cloud Licensing Model, <http://www.microsoft.com/applicationplatform/default.aspx>
3. Cloud for the CFO, http://www.cio.com/article/459819/How_to_Expain_Cloud_Computing_to_Your_CFO
4. Univa UD, <http://www.univaud.com/index.php>