UNIT-I

Cloud Computing in a Nutshell

Cloud Computing

Cloud computing is a general term for anything that involves delivering hosted services over the Internet.

"Cloud computing is the use of computing resources (hardware and software) that are delivered as a service over a network (typically the Internet). The name comes from the common use of a cloud-shaped symbol as an abstraction for the complex infrastructure it contains in system diagrams. Cloud computing entrusts remote services with a user's data, software and computation."

Cloud Benefits

- Pay as you go
- Focus on business rather than IT
- Elasticity - Scale up and down based on business need

Cloud Models

- Deployment Models : Public Cloud, Private Cloud, Hybrid Cloud, Community Cloud
- Service Models : SaaS, PaaS, IaaS

Essential Characteristics

- On Demand Self-Service: Allows for provisioning of computing resources automatically as needed.
- Broad Network Access: Access to cloud resources is over the network using standard mechanisms provided through multi-channels.
- Resource Pooling: The vendors' resources are capable of being pooled to serve multiple clients using a multi-tenant model, with different physical and virtual resources in a dynamic way.
- Example of resources include; computation capabilities, storage and memory.
- Rapid Elasticity: Allows for rapid capability provisioning, for quick scaling out and scaling in of capabilities. The capability available for provisioning to the client seems to be unlimited and that it can be purchased as demanded.

- Measured Service: Allows monitoring, control and reporting of usage. It also allows for transparent between the provider and the client.

Basic Roots of Cloud Computing

We can find the basic roots of cloud computing by monitoring the different technologies specially in hardware (virtulization, multicore chips), Internet technologies (Web services, service oriented architecture, web 2.0), distributed computing (cluster, grids) and system management (automatic computing, data center automation) as shown in above figure 1.1, as we discuss one by one technology.

SOA, Web Services, Web 2.0, and Mashups

Web services (WS) open standards have appreciably throw in to advance domain of software assimilation. WS standards have been created on top of existing ubiquitous technologies such as HTTP, XML, thus providing a common mechanism for delivering services, making them ideal for implementing a service-oriented architecture i.e. SOA. In SOA software resources are packaged as "services" which is well defined self contained modules that provide by business functionality [43]. The concept of gluing services initially focused on the enterprise Web but gained space in the consumer realm as well especially with the advent of Web 2.0, in the consumer Web information and services may be programmatically aggregated acting as a building blocks of complex compositions called service mashups. Google make their service APIs publically accessible using standard protocols such as SOAP and REST [44].

Grid & Utility Computing

Grid computing facilitates aggregation of distributed resources and transparently accesses them. Most production grids TeraGrid, EGEE share more computable and storage resources distributed across different administrative domains with their main intension is speeding up a broad range of scientific application such as climate modeling, drug design and protein analysis. Globus Toolkit is a middleware that implements several standard Grid services and over the years has abet the deployment of several service oriented. In utility computing environments users assign a "utility" value to their jobs, where utility is a fixed or time-varying value that captures various QoS constraints (deadline, satisfaction).The service providers attempt their own utility , where said utility may directly correlate with their profit.

Hardware Virtualization

The cloud computing service providers basically provides large scale data center composed of millions of computers. Such data centers are built to serve many consumers and host many isolated applications. For this purpose

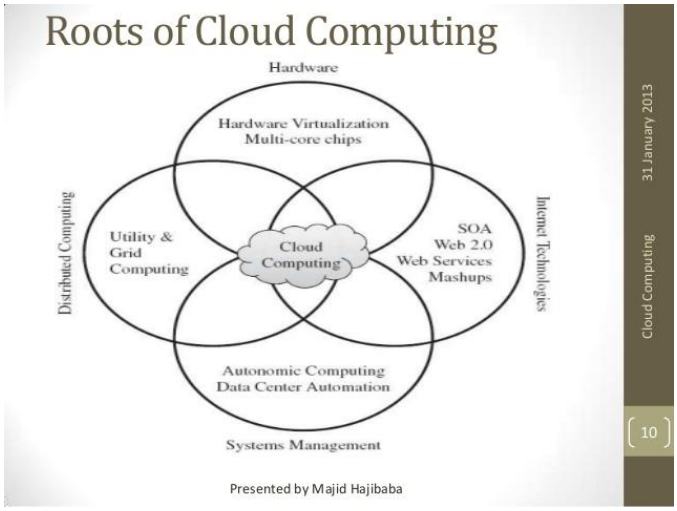
hardware virtualization can be considered. The idea of hardware virtualization comes from distributed operating system, Hardware virtualization allows running multiple operating system call Virtual machines and software stacks on a single physical platform, Virtual machine monitor call hypervisor is hypervisor between guest os and host os as shown in below figure 1.2

Hardware Virtualization

Virtual machine 1, Virtual machine 2 acts as a guest operating system which is run on single physical machine at that time hardware is virtualized such as processor, I/O devices and memory , hypervisor Workload isolation is achieved through API by the service providers, workload migration, VM resume, VM migration, VM pause, VM clone these capabilities also applied using API at that time VM state observation is more important, A number of VMM platform available to handle VMs and Physical machines. The most notables ones is VMWare, Xen and KVM, VirtualBox. In our work we use VirtualBox to create cloud environments and handling all VMs using API of VirtualBox. VirtualBox is a powerful x86 and AMD64/Intel64 virtualization product for enterprise as well as home use. Not only is VirtualBox an extremely feature rich, high performance product for enterprise customers, it is also the only professional solution that is freely available as Open Source Software under the terms of the GNU General Public License (GPL) version 2.

Automatic Computing

Autonomic or self-managing system rely on monitoring probes and sensors, on adaptation engine for computing optimization based on monitoring data, and on effectors to carry out changes on the system. IBM Automatic Computing Initiative has contributing to define four characteristics of automatic systems such as : (i) self-configuration, (ii) self-optimization (iii)self-healing (iv) self-protection, IBM also introduced automatic control loops of automatic managers called MAPE-K i.e. Monitor Analyze Plan Execute-Knowledge [45], [46].



Layers and Types of Clouds

Cloud computing

Cloud computing is defined as delivery of hosted services over the Internet in an on-demand model. A cloud provider offers resources (hardware, software or development stacks) as services over the internet that can be consumed on a pay-as-you-use basis.

Layers Or Services Of Cloud

- Software as a service (SaaS)
- Platform as a Service (PaaS)
- Infrastructure as a Service (IaaS)

Application Cloud (a.k.a Software Cloud) - Software as a Service (SaaS)

"Software as a Service is the hosted delivery of Software that consumers can access over the internet." This definition is rather simplistic. By this definition web based email services such as Gmail, Hotmail etc are SaaS applications. Two features of a SaaS application are scalability and configurability. SaaS applications should be able to quickly scale with demand. In mature SaaS applications, the customer should be able to customize their instance of the software using meta-data. Depending on who you talk to multi-tenancy is yet another feature of SaaS applications. Multiple customers share a single instance of the software and the provider can optimize the resources to suit individual demands.

Google Apps - a SaaS version of desktop applications where the software is hosted on google's servers. Salesforce - SaaS CRM products.

Platform Cloud – Platform as a Service (PaaS)

Platform as a Service offers a complete platform and the tools to develop and deploy applications on the platform. Typically the PaaS vendors also provide tools and other services that enable rapid application development. Applications developed on the platform are tied to the platform. Each platform has its own development model (can be more aptly called quirks) and developers need to be aware of these.

App Engine is a PaaS offering from Google. Applications are developed and deployed on Google's infrastructure. Google App Engine uses big table to store data. This is an example of platform specific development, a change that traditional developers will have to face when they move to App Engine. It does not support relational databases.

Force.com is a PaaS offering from Salesforce. Developers can create applications that run on the force.com platform. It includes database, workflow and UI tools. Applications can be built using their proprietary Apex programming language or using Visual Force (a tag based markup language).
Microsoft Azure is Microsoft's platform as a service offering. It includes an OS with .NET runtime, a database (SQL Azure) and other services that make it easier to couple on-premise and off-premise applications.

Infrastructure Cloud – Infrastructure as a service (IaaS)

Infrastructure as a Service is the hosted delivery of infrastructure services such as servers, networks and other hardware to consumers. IaaS provides consumers access to on-demand, scalable storage and compute power.
Amazon EC2 is an example of IaaS which provides scalable compute capacity in the cloud. Basically, consumers get a machine that they customize and use. This is an example of hardware as a service.

A different way to think about the different layers of cloud computing is to think in terms of control that the customer has has:

- With SaaS you use the application with little customization.
- With PaaS you create applications that run on the platform. You are limited by the platform as to what applications you can create.
- With IaaS you build applications without any tie in to the platform.

Types Of Cloud

Cloud computing is usually described in one of two ways. Either based on the cloud location, or on the service that the cloud is offering.

Based on a cloud location, we can classify cloud as:

- public,
- private,
- hybrid
- community cloud

Based on a service that the cloud is offering, we are speaking of either:

- IaaS (Infrastructure-as-a-Service)
- PaaS (Platform-as-a-Service)
- SaaS (Software-as-a-Service)
- or, Storage, Database, Information, Process, Application, Integration, Security, Management, Testing-as-a-service

When we talk about **public cloud**, we mean that the whole computing infrastructure is located on the premises of a cloud computing company that offers the cloud service. The location remains, thus, separate from the customer and he has no physical control over the infrastructure.
As public clouds use shared resources, they do excel mostly in performance, but are also most vulnerable to various attacks.

Private cloud means using a cloud infrastructure (network) solely by one customer/organization. It is not shared with others, yet it is remotely located. If the cloud is externally hosted. The companies have an option of choosing an on-premise private cloud as well, which is more expensive, but they do have a physical control over the infrastructure.
The security and control level is highest while using a private network. Yet, the cost reduction can be minimal, if the company needs to invest in an on-premise cloud infrastructure.

Hybrid cloud, of course, means, using both private and public clouds, depending on their purpose.
For example, public cloud can be used to interact with customers, while keeping their data secured through a private cloud.

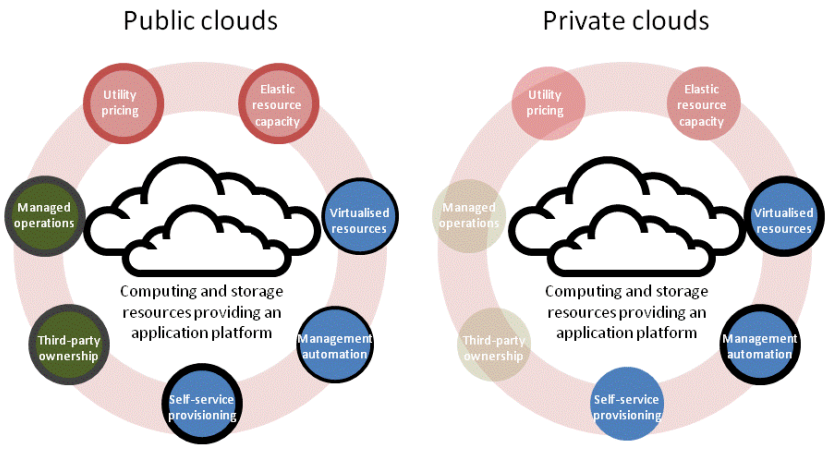


Image 1 – Private vs Public Cloud (Image Source: TalkCloudComputing)

Community cloud implies an infrastructure that is shared between organizations, usually with the shared data and data management concerns. For example, a community cloud can belong to a government of a single country. Community clouds can be located both on and off the premises.

Cloud Service

The most popular services of the cloud are **infrastructure**,**platform**, **software**, or **storage**.

The most common cloud service is that one offering data storage disks and virtual servers, i.e. infrastructure. Examples of Infrastructure-as-a-Service (IaaS) companies are Amazon, Rackspace, Flexiscale.

If the cloud offers a development platform, and this includes operating system, programming language execution environment, database, and web server, the model is known as Platform-as-a-Service (PaaS), examples of which are Google App Engine, Microsoft Azure, Salesforce. Operating system can be frequently upgraded and developed with PaaS, services can be obtained from diverse sources, and programming can be worked in teams (geographically distributed).

Software-as-a-Service (SaaS), finally, means that users can access various software applications on a pay-per-use basis. As opposed to buying licensed programs, often very expensive. Examples of such services include widely used GMail, or Google Docs.

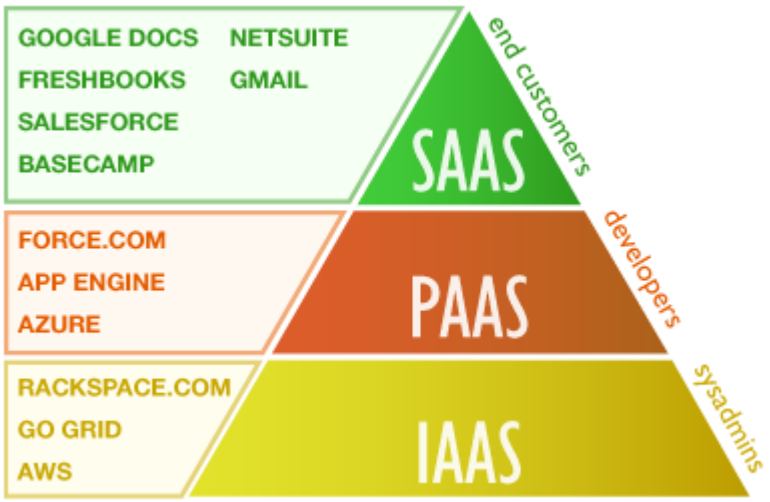


Image 2 – Cloud Services

Public Clouds

A public cloud is basically the internet. Service providers use the internet to make resources, such as applications (also known as Software-as-a-service) and storage, available to the general public, or on a 'public cloud. Examples of public clouds include Amazon Elastic Compute Cloud (EC2), IBM's Blue Cloud, Sun Cloud, Google AppEngine and Windows Azure Services Platform.

For users, these types of clouds will provide the best economies of scale, are inexpensive to set-up because hardware, application and bandwidth costs are covered by the provider. It's a pay-per-usage model and the only costs incurred are based on the capacity that is used.

There are some limitations, however; the public cloud may not be the right fit for every organization. The model can limit configuration, security, and SLA specificity, making it less-than-ideal for services using sensitive data that is subject to compliancy regulations.

Private Clouds

Private clouds are data center architectures owned by a single company that provides flexibility, scalability, provisioning, automation and monitoring. The goal of a private cloud is not sell "as-a-service" offerings to external customers but instead to gain the benefits of cloud architecture without giving up the control of maintaining your own data center.

Private clouds can be expensive with typically modest economies of scale. This is usually not an option for the average Small-to-Medium sized business and is most typically put to use by large enterprises. Private clouds are driven by concerns around security and compliance, and keeping assets within the firewall.

Hybrid Clouds

By using a Hybrid approach, companies can maintain control of an internally managed private cloud while relying on the public cloud as needed. For instance during peak periods individual applications, or portions of applications can be migrated to the Public Cloud. This will also be beneficial during predictable outages: hurricane warnings, scheduled maintenance windows, rolling brown/blackouts.

The ability to maintain an off-premise disaster recovery site for most organizations is impossible due to cost. While there are lower cost solutions and alternatives the lower down the spectrum an organization gets, the capability to recover data quickly reduces. Cloud based Disaster Recovery (DR)/Business Continuity (BC) services allow organizations to contract failover out to a Managed Services Provider that maintains multi-tenant infrastructure for DR/BC, and specializes in getting business back online quickly.

Desired Features of a Cloud

Key Features Of Cloud Computing

The most talked-about term currently in the IT industry is cloud computing. Everyone is thinking about cloud computing from different perspectives. Some emphasize the cost benefits associated with it, while others are still

cautious about security and privacy. It has become extremely important to understand the key defining features of cloud computing.

1. Resource Pooling and Elasticity

In cloud computing, resources are pooled to serve a large number of customers. Cloud computing uses multi-tenancy where different resources are dynamically allocated and de-allocated according to demand. From the user's end, it is not possible to know where the resource actually resides.

The resource allocation should be elastic, in the sense that it should change appropriately and quickly with the demand. If on a particular day the demand increases several times, then the system should be elastic enough to meet that additional need, and should return to the normal level when the demand decreases.

2. Self-Service and On-demand Services

Cloud computing is based on self-service and on-demand service models. It should allow the user to interact with the cloud to perform tasks like building, deploying, managing, and scheduling. The user should be able to access computing capabilities as and when they are needed and without any interaction from the cloud-service provider. This would help users to be in control, bringing agility in their work, and to make better decisions on the current and future needs.

3. Pricing or Per-Usage Metering and Billing

Cloud computing does not have any upfront cost. It is completely based on usage. The user is billed based on the amount of resources they use. This helps the user to track their usage and ultimately help to reduce cost. Cloud computing must provide means to capture, monitor, and control usage information for accurate billing. The information gathered should be transparent and readily available to the customer. This is necessary to make the customer realize the cost benefits that cloud computing brings.

4. Quality of Service or Customization

Cloud computing must assure the best service level for users. Services outlined in the service-level agreements must include guarantees on round-

the-clock availability, adequate resources, performance, and bandwidth. Any compromise on these guarantees could prove fatal for customers.

The decision to switch to cloud computing should not be based on the hype in the industry. A good understanding of the technology enables the user to make smarter decisions. Knowing all the features will empower the business users to understand and negotiate with the service providers in a proactive manner.

,

Challenges and Risks:

Top Five Challenges of Cloud Computing

Cloud computing is steadily gaining acceptance within businesses. It is predicted that by 2018, 59% of the cloud workloads will be generated from Software as a Service (Saas)

Cloud Computing has already started to revolutionise the way we store and access data. We currently see smartphone applications use cloud computing technology to allow users to store and access data they previously couldn't on a smart device.

Although cloud computing is on its way to becoming a huge success and whilst it is clear there is a lot of business value, there are reservations amongst some CIOs about using some cloud technologies. Let's explore some of the challenges and concerns.

1. Security & Privacy

Security is a great concern for CIOs when moving their data to the cloud. Although security in the cloud is generally reliable and proficient, CIOs need to know that the cloud provider they chose to work with has a fully secure cloud environment.

CIOs are becoming more reluctant to hand over important data to a third party provider. With the growth in data breaches and the potential financial penalties and loss of reputation for companies who fall victim, moving your private data to an external provider is more daunting than ever.

A well-established cloud computing vendor will ensure they have the latest sophisticated security systems in place to defend against threats. We have put together a list of questions below that you should ask a cloud provider.

- Where does your data reside?

- Is the data encrypted?
- How do you move data from the cloud?
- What are your security governance policies and procedures?

The cloud provider should be able to answer all the above questions in detail, so that you know exactly where your data is stored and how they will protect your data against internal and external threats. Moreover, how can you retrieve that data if it becomes necessary.

2. Service Quality

Service quality is often one of the most significant factors that businesses cite as a reason for not moving their business applications to the cloud. Often businesses feel as though the SLAs provided by the cloud providers today are not adequate to assure the requirements for running a production application on the cloud, especially those related to availability, performance and scalability.

According to a recent survey 43% of IT decision makers are planning to invest more into cloud computing this year. CIOs need assurance that the company's data will be secure and available, and the service reliable at all times. Ensuring maximum upkeep of the service is paramount for the profitability and sustainability of the business. We've put together a list of ten questions you should ask a provider before signing the contract;

1. What are your minimum service levels?
2. What remedies are in place when a failure occurs?
3. What disaster recovery and business continuity procedures are in place?
4. How portable is my data?
5. What is your change management process?
6. What are your infrastructure and security standards?
7. How quickly do you identify and solve problems?
8. What is your escalation process?
9. What is your exit strategy?
10. What is your termination process?

Without proficient service quality and comprehensive answers to the above questions, businesses will be reluctant to host their critical infrastructure within the cloud.

3. Downtime & Accessibility

Service quality doesn't have to be compromised when your data is in the cloud. Accessing your data when you need it is a basic requirement from many organisations. The challenge with the cloud is that the data is accessed via an internet connection rather than a local connection. So when the network or

internet connection is down, it also means that cloud services are also down; thus data cannot be accessed.

Performance of the cloud infrastructure can be affected by the load, environment and number of users. Ensuring that your cloud infrastructure is resilient to outages is vital. Whilst it is almost impossible to mitigate all server outages, a reputable provider will have robust resilience measures in place to protect your data.

4. Access to data

Cloud-based servers do not always have the most effective or appropriate customer service support systems. CIOs often express their concerns around data ownership and losing control of their data when moving to the cloud, but this shouldn't be an issue. Selecting where and how your data is stored is an important element within the decision making process. Integration is a problem for many organisations. Ensuring that all of the applications are able to seamlessly integrate with one another is also a common challenge.

Important questions you should ask a cloud provider;

- How much control do I have over my data and server?
- How much time does it take to back up my data to the cloud?
- How long does it take to back up my data?
- Where does my data reside?
- How does the service provider secure my data?
- What are their audit procedures?
- What happens in the event of data corruption?
- How can I extract my data if I need to move elsewhere?

5. Transition to the cloud

Many cloud adoption challenges are unknown due to the fact that cloud technology is still in its relative infancy. CIOs are challenged deciding on the best way to transition to the cloud and finding a cloud solution that meets the aims of the businesses, whilst improving efficiencies.

Although transitioning to the cloud is a complex and involved process, there isn't one route to success. CIOs must ensure that the proposed solution compliments their business model. There are various ways businesses can transition to the cloud. Whether it's via private, public or hybrid technologies, identifying the right service model for your business is a vital step.

Questions that a cloud provider should ask you;

- What are your demand patterns?
- When do you get the biggest influx of data?

- How much do you expect your data to grow?
- Do you need control over the region (geographic) where your data resides?
- What is your SLA expectation?

Migrating data poses a number of risks for organisations if not handled correctly. Developing a data migration strategy that integrates seamlessly with the current IT infrastructure is key to overall success.

CIOs are challenged with finding the right service model for their business. Finding a provider that will allow you to create a customised computing environment is vital.

The first step in transitioning to the cloud is being able to identify the challenges and working with your chosen cloud provider to navigate around these barriers in order to facilitate a successful cloud environment for the business. Whether Public, Private or Hybrid, making sure you ask the right questions and understand the risks for your business is imperative.