Unit-1

Cloud Introduction

Cloud Computing Fundamentals

Vision

1.1 Cloud Computing Definition

Cloud Definition: *The cloud in cloud computing provides the means through which everything — from computing power to computing infrastructure, applications, business processes to personal collaboration — can be delivered to a user as a service wherever and whenever the user needs*. The cloud itself *is a set of hardware, networks, storage, services, and interfaces that enable the delivery of computing as a service.* Cloud services include the delivery of software, infrastructure, and storage over the Internet (either as separate components or a complete platform) based on user demand.

Cloud Computing Definition

Cloud computing is a model for enabling ubiquitous, 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.* Cloud computing refers *to both the applications delivered as services over the Internet and the hardware and system software in the datacenters that provide those services.*

Why Cloud computing?

The cloud embodies the following basic characteristics:

- Elasticity and the ability to scale up and down
- Self-service provisioning and automatic de-provisioning
- Application programming interfaces (APIs)
- Billing and metering of service usage in a pay-as-you-go model
- Security

Elasticity and scalability

The service provider can‘t anticipate how customers will use the service. One customer might use the service three times a year during peak selling seasons, whereas another might use it as a primary development platform for all of its applications. Therefore, the service needs to be available all the time (7 days a week, 24 hours a day) and it has to be designed to scale upward for high periods of demand and downward for lighter ones. Scalability also means that an application can scale when additional users are added and when the application requirements change. This ability to scale is achieved by providing elasticity.

Self-service provisioning

Customers can easily get cloud services without going through a lengthy process. The customer simply requests an amount of computing, storage, software, process, or other resources from the service provider. While the on-demand provisioning capabilities of cloud services eliminate many time delays, an organization still needs to do its homework. These services aren‘t free; needs and requirements must be determined before capability is automatically provisioned.

Application programming interfaces (APIs)

Cloud services need to have standardized APIs. These interfaces provide the instructions on how two application or data sources can communicate with each other. A standardized interface lets the customer more easily link a cloud service, such as a customer relationship management system with a financial accounts management system, without having to resort to custom programming.

Billing and metering of services

A cloud environment needs a built-in service that bills customers. And, of course, to calculate that bill, usage has to be metered (tracked). Even free cloud services (such as Google‘s Gmail or Zoho‘s Internet-based office applications) are metered. In addition to these characteristics, cloud computing must have two overarching requirements to be effective:

- A comprehensive approach to service management
- A well-defined process for security management

Performance monitoring and measuring

A cloud service provider must include a service management environment. A service management environment is an integrated approach for managing the physical environments and IT systems. This environment must be able to maintain the required service level for that organization. In other words, service management has to monitor and optimize the service or sets of services. Service management has to consider key issues, such as performance of the overall system, including security and performance. For example, an organization using an internal or external email cloud service would require 99.999 percent uptime with maximum security. The organization would expect the cloud provider to prove that it has met its obligations.

Many cloud service providers give customers a dashboard — a visualization of key service metrics — so they can monitor the level of service they‘re getting from their provider. Also, many customers use their own monitoring tools to determine whether their service level requirements are being met.

Security

Many customers must take a leap of faith to trust that the cloud service is safe. Turning over critical data or application infrastructure to a cloud-based service provider requires making sure that the information can‘t be accidentally accessed by another company (or maliciously accessed by a hacker). Many companies have compliance requirements for securing both internal and external information. Without the right level of security, one might not be able to use a provider‘s offerings.

1.2 Types of Cloud

Cloud computing is offered in different forms:

- Public clouds: in a public cloud, a business rents the capability and they pay for what they use on-demand. Example- Amazon, Google and IBM
- Public cloud services are available to clients from a third party service provider via the internet. Public clouds provide an elastic, cost

Application Workloads

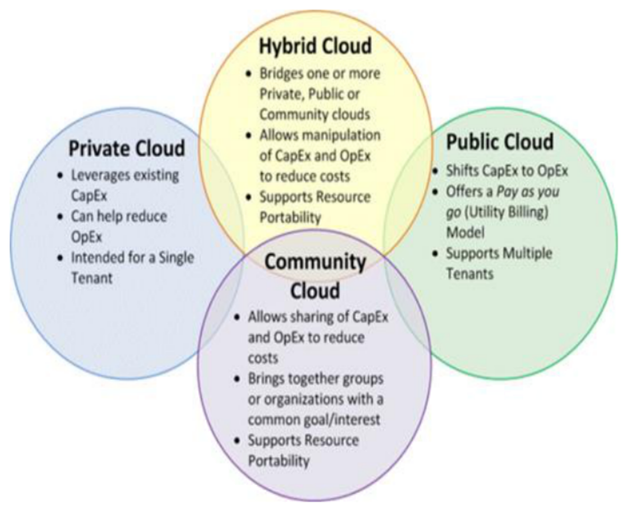
- Public facing Web pages
- Public Wiki‘s, blogs, etc
- Batch processing jobs with limited security requirements
- Data intensive workloads
- Software as a Service applications
- Online storage solutions
- Online backups/restore solutions
- Isolated workloads where latency between application components is not an issue

Drawbacks

- Organizations are hesitant to move to the public cloud primarily due to security and regulatory concerns.
- Others worry about locking data into a single vendor‘s cloud infrastructure.
- While public cloud computing allows the enterprise to avoid many infrastructure expenses, it‘s important to plan for other associated cost areas tied to a deployment, such as vendor management processes, capacity planning, service oriented architecture (SOA),chargeback systems, incident management and service level agreements (SLAs)

- Private clouds: in private clouds, a business essentially turns its IT environment into a cloud and uses it to deliver services to their users. The benefits are

- Elastic on demand capacity
- Self- service provisioning
- Service based access
- Improved security and resiliency



Services

- Virtualization
- Multi-tenancy
- Consistent deployment
- Chargeback and pricing

Drawbacks

- It does not provide the optimum promise of cloud computing such as the lowest price and unlimited elasticity to ramp resources

- Hybrid clouds, which combine both public and private clouds.

Services

- Elasticity
- Pay-as-you-go pricing
- Network isolation & secure connectivity

- Gradually move to public cloud configuration

Drawbacks

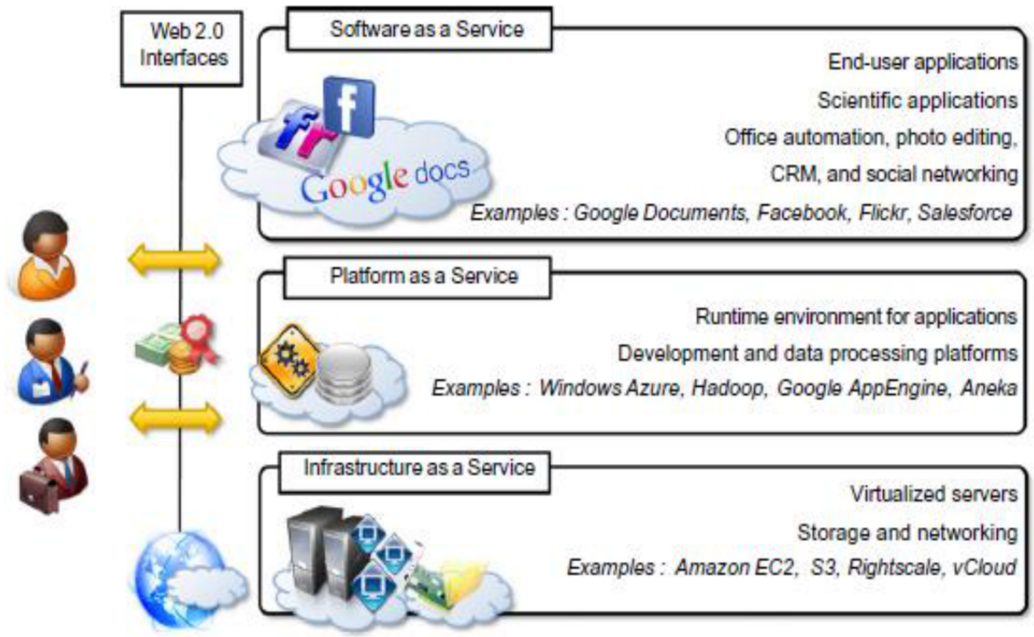
- The complexity of monitoring and managing all portions of the hybrid cloud from a common portal or service desk.
- This requires considerable engineering on the part of the IT organization or the acquisition of third-party vendor services to provide the necessary ―glue‖ to oversee the enterprise environment wealth of information on how to enhance their business value while reducing costs and freeing resources for more strategic initiatives

Other clouds

- Community cloud- it is controlled by a group of organizations that have shared interests such as specific security requirements or a common mission. It 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.
- Shared private cloud- This is shared compute capacity with variable usage based pricing to business units that are based on service offerings, accounts datacenters and it requires an internal profit center to take over or buy infrastructure made available through account consolidations
- Dedicated private cloud- it has IT service catalogue with dynamic provisioning. It depends on standardized SOA assets that can be broadly deployed into new and existing accounts and is a lower cost model
- Dynamic private cloud – it allows client workloads to dynamically migrate to and from the compute cloud as needed. This model can be shared and dedicated. It delivers on the ultimate value of clouds. This is a very low management model with reliable SLAs and scalability.

1.3 Cloud services/ Reference Model

Based upon the services offered, clouds are classified in the following ways: Each layer provides a different service to users.



1. **Infrastructure as a service (IaaS)** - At the base of the stack, Infrastructure-as-a-Service solutions deliver infrastructure on demand in the form of virtual hardware, storage, and networking. It involves offering hardware related services using the principles of cloud computing. These could include some kind of storage services (database or disk storage) or virtual servers. Leading vendors that provide Infrastructure as a service are Amazon EC2, Amazon S3,Rackspace Cloud Servers and Flexiscale.

Service Model	Benefits	Risks	Best Fit
IaaS	✓ On-Demand Infrastructure	✓ Security ✓ Data persistence ✓ Data aggregation ✓ Business Risk of Outages ✓ Service failures can affect multiple tenants and customers	✓ Non-legacy apps ✓ Consolidation Efforts ✓ Hosting for Dev & Test
PaaS	✓ Standardized development environment ✓ Rapid development & testing	✓ Similar risks as above ✓ Vendor Lock-in	✓ New application development ✓ Application development that uses provider building blocks to reduce time-to-market
✓SaaS	✓ Re-usable services ✓ Only requires limited configuration & management	✓ Similar risks as IaaS ✓ Lack of control ✓ Vendor Lock-in	✓ Configuration over customization ✓ Commoditized applications

The IaaS customer rents computing resources instead of buying and installing them in their own data center. The service is typically paid for on a usage basis. The service may include dynamic scaling so that if the customer winds up needing more resources than expected, he can get them immediately (probably up to a given limit). Dynamic scaling as applied to infrastructure means that the infrastructure can be automatically scaled up or down, based on the requirements of the application.

Additionally, the arrangement involves an agreed-upon service level. The service level states what the provider has agreed to deliver in terms of availability and response to demand. It might, for example, specify that the resources will be available 99.999 percent of the time and that more resources will be provided dynamically if greater than 80 percent of any given resource is being used.

Currently, the most high-profile IaaS operation is Amazon‘s Elastic Compute Cloud (Amazon EC2). It provides a Web interface that allows customers to access virtual machines. EC2 offers scalability under the user‘s control with the user paying for resources by the hour. The use of the term *elastic* in the naming of Amazon‘s EC2 is significant. The elasticity refers to the ability that EC2 users have to easily increase or decrease the infrastructure resources assigned to meet their needs. The user needs to initiate a request, so this service provided isn‘t dynamically scalable. Users of EC2 can request the use of any operating system as long as the developer does all the work. Amazon itself supports a more limited number of operating systems (Linux, Solaris, and Windows)

Companies with research-intensive projects are a natural fit for IaaS. Cloud based computing services allow scientific and medical researchers to perform testing and analysis at levels that aren‘t possible without additional access to computing infrastructure. Other organizations with similar needs for additional computing resources may boost their own data centers by renting the computer hardware — appropriate allocations of servers, networking technology, storage, and data center space — as a service. Instead of laying out the capital expenditure for the maximum amount of resources to cover their highest level of demand, they purchase computing power when they need it.

2. **Platform as a Service (PaaS)** involves offering a development platform on the cloud. Platforms provided by different vendors are typically not compatible. Typical players in PaaS are Google Application Engine, Microsofts Azure, Salesforce.com, force.com .

They deliver scalable and elastic runtime environments on demand and host the execution of applications. These services are backed by a core middleware platform that is responsible for creating the abstract environment where applications are deployed and executed.

With Platform as a Service (PaaS), the provider delivers more than infrastructure. It delivers a solution stack — an integrated set of software that provides everything a developer needs to build an application — for both software development and runtime. PaaS can be viewed as an evolution of Web hosting.

In recent years, Webhosting companies have provided fairly complete software stacks for developing Web sites. PaaS takes this idea a step farther by providing lifecycle management — capabilities to manage all software development stages from planning and design, to building and deployment, to testing and maintenance.

The primary benefit of PaaS is having software development and deployment capability based entirely in the cloud — hence, no management or maintenance efforts are required for the infrastructure. Every aspect of software development, from the design stage onward (including source-code management, testing, and deployment) lives in the cloud.

PaaS is inherently multi-tenant and naturally supports the whole set of Web services standards and is usually delivered with dynamic scaling. In reference to Platform as a Service, dynamic scaling means that the software can be automatically scaled up or down. Platform as a Service typically addresses the need to scale as well as the need to separate concerns of access and data security for its customers.

The major drawback of Platform as a Service is that it may lock one in to the use of a particular development environment and stack of software components. Platform as a Service offerings usually have some proprietary elements (perhaps the development tools or even component libraries). Consequently, one may be wedded to the vendor‘s platform and unable to move the applications elsewhere without rewriting them to some degree. If one suddenly becomes dissatisfied with the Platform as a Service provider, one may face very high expenses when one suddenly needs to rewrite the applications to satisfy the requirements of another PaaS vendor.

The fear of vendor lock-in has led to a new variety of Platform as a Service emerging: Open Platform as a Service. This would offer the same approach as Platform as a Service, except that there is no constraint on choice of development software. It avoids the possibility of lock-in.

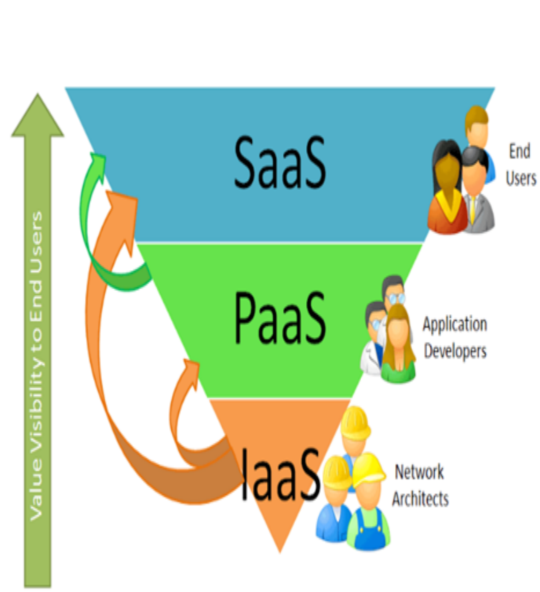
Some examples of Platform as a Service include the Google App Engine, AppJet, Etelos, Qrimp, and Force.com, which is the official development environment for Salesforce.com.

3. **Software as a service (SaaS)** - Software-as-a-Service solutions provide applications and services on demand. Most of the common applications are Office automation, photo editing, CRM, and social networking Examples: Google Documents, Facebook, Flickr, Salesforce Web 2.0. Software as a Service Interfaces are replicated on the provider's infrastructure and made more scalable and accessible through a browser on demand. These applications are shared across multiple users whose interaction is isolated from the other users. The SaaS layer is also the area of social networking Websites, which leverage cloud-based infrastructures to sustain the load generated by their popularity. It includes a complete software offering on the cloud. Users can access a software application hosted by the cloud vendor on pay-per-use basis. This is a well-established sector. The pioneer in this field has been Salesforce.coms offering in the online Customer Relationship Management (CRM) space. Other examples are online email providers like Googles gmail and Microsofts hotmail, Google docs and Microsofts online version of office called BPOS (Business Productivity Online Standard Suite).

One of the first implementations of cloud services was *Software as a Service (SaaS)* — business applications that are hosted by the provider and delivered as a service. SaaS has its roots in an early kind of hosting operation carried out by Application Service Providers (ASPs).

CRM is one of the most common categories of Software as a Service.

The price of the software is on a *per-use* basis and involves no upfront costs from the service provider. Of course, the reality is that the company may have some upfront work to do to get the data loaded into the Software as a Service application database and one may have to deal with ongoing data integration between the internal and cloud data stores. Businesses get the immediate benefit of reducing capital expenditures. In addition, a business gains the flexibility to test new software on a rental basis and then can continue to use and adopt the software, if it proves suitable.



Example: Microsoft Cloud Services

The above classification is well accepted in the industry. David Linthicum describes a more granular classification on the basis of service provided. These are listed below:

1. Storage-as-a-service
2. Database-as-a-service
3. Information-as-a-service
4. Process-as-a-service
5. Application-as-a-service
6. Platform-as-a-service
7. Integration-as-a-service
8. Security-as-a-service
9. Management/Governance-as-a-service
10. Testing-as-a-service
11. Infrastructure-as-a-service

1.4 Benefits and challenges of Cloud Computing

Specifically, cloud computing offers the following key advantages:

- **Optimized server utilization** - as most enterprises typically underutilize their server computing resources, cloud computing will manage the server utilization to the optimum level. Cloud computing can lower IT barriers to innovation, as can be witnessed from the many promising startups, from the ubiquitous online applications such as Facebook and Youtube to the more focused applications like TripIt (for managing one's travel) or Mint (for managing one's personal finances).
- **Cost saving** - IT infrastructure costs are almost always substantial and are treated as a capital expense (CAPEX). However if the IT infrastructure usually becomes an operating expense (OPEX). In some countries, this results in a tax advantage regarding income taxes. Also, cloud computing cost saving can be realized via resource pooling. It dramatically lowers the cost of entry for smaller firms trying to benefit from compute-intensive business analytics that were hither to available only to the largest of corporations. These computational exercises typically involve large amounts of computing power for relatively short amounts of time, and cloud computing makes such dynamic provisioning of resources possible. Cloud computing also represents a huge opportunity to many third-world countries that have been so far left behind in the IT revolution —some cloud computing providers are using the advantages of a cloud platform to enable IT services in countries that would have traditionally lacked the resources for widespread deployment of IT services.
- **Improved Compatibility between Operating Systems** In the cloud, operating systems simply don‘t matter. One can connect the Windows computer to the cloud and share documents with computers running Apple‘s Mac OS, Linux, or UNIX. In the cloud, the data matters, not the operating system.
- **Dynamic scalability** - many enterprises include a reasonably large buffer from their average computing requirement, just to ensure that capacity is in place to satisfy peak demand. Cloud computing provides an extra processing buffer as needed at a low cost and without the capital investment or contingency fees to users.
- Cloud computing makes it easier for enterprises to scale their services – which are increasingly reliant on accurate information – according to client demand. Since the

computing resources are managed through software, they can be deployed very fast as new requirements arise. In fact, the goal of cloud computing is to scale resources up or down dynamically through software APIs depending on client load with minimal service provider interaction.

- **Unlimited Storage Capacity** - Similarly, the cloud offers virtually limitless storage capacity. Consider that when your desktop or laptop PC is running out of storage space. The computer's 200GB hard drive is peanuts compared to the hundreds of petabytes (a million gigabytes) available in the cloud.
- **Universal Access to Documents** - With cloud computing, one doesn't have to take the documents. Instead, they stay in the cloud, where one can access them from anywhere one has a computer and an Internet connection. All the documents are instantly available from wherever one accesses them. There's simply no need to take the documents with us—as long as we have an Internet connection, that is.
- **Latest Version Availability**- And here's another document-related advantage of cloud computing. When one edits a document at home, that edited version is what he'll see when he will access the document at work. The cloud always hosts the latest version of the documents; one is never in danger of having an outdated version on the computer when working on.
- **Shortened development life cycle** - cloud computing adopts the service-orientates architecture (SOA) development approach which has significantly shorter development life cycle that that required by the traditional development approach. Any new business application can be developed online, connecting proven functional application building blocks together.
- **Reduced time for implementation** - cloud computing provides the processing power and data storage as needed at the capacity required. This can be obtained in near-real time instead of weeks or months that occur when a new business initiative is brought online in a traditional way.
- Cloud computing also makes possible new classes of applications and delivers services that were not possible before. Examples include

a) **mobile interactive applications** that are location-, environment- and context-aware and that respond in real time to information provided by human users, nonhuman sensors (e.g. humidity and stress sensors within a shipping container) or even from independent information services (e.g. worldwide weather data);

b) **parallel batch processing**, that allows users to take advantage of huge amounts of processing power to analyze terabytes of data for relatively small periods of time, while programming abstractions like Google's MapReduce or its open source counterpart Hadoop makes the complex process of parallel execution of an application over hundreds of servers transparent to programmers;

c) **Business analytics** that can use the vast amount of computer resources to understand customers, buying habits, supply chains and so on from voluminous amounts of data;

d) **Extensions of compute-intensive** desktop applications that can offload the data crunching to the cloud leaving only the rendering of the processed data at the front-end, with the availability of network bandwidth reducing the latency involved.

- It can provide an almost **immediate access to hardware resources**, with no upfront capital investments for users, leading to a faster time to market in many businesses. Treating IT as an operational expense (in industry-speak, employing an _Op-ex‘ as opposed to a _Cap-ex‘ model) also helps in dramatically reducing the upfront costs in corporate computing. For example, many of the promising new Internet startups like 37 Signals, Jungle Disk, Gigavox, SmugMug and others were realized with investments in information technology that are orders of magnitude lesser than that required just a few years ago. The cloud becomes an adaptive infrastructure that can be shared by different end users, each of whom might use it in very different ways.
- The users are completely separated from each other, and the **flexibility of the infrastructure allows for computing loads to be balanced on the fly** as more users join the system (the process of setting up the infrastructure has become so standardized that adding computing capacity has become almost as simple as adding building blocks to an existing grid). The beauty of the arrangement is that as the number of users goes up, the **demand load** on the system gets more balanced in a **stochastic sense**, even as **its economies of scale expand**.

For all the above benefits of cloud computing, it also incorporates some unique and notable **technical or business risk** as follows:

- Data location - cloud computing technology allows cloud servers to reside anywhere, thus the enterprise may not know the physical location of the server used to store and process their data and applications. Although from the technology point of view, location is least relevant, this has become a critical issue for data governance requirements. It is essential to understand that many Cloud Service Providers (CSPs) can also specifically define where data is to be located.
- Connectivity and Open Access The full potential of cloud computing depends on the availability of high-speed access to all. Such connectivity, rather like electricity availability, globally opens the possibility for industry and a new range of consumer products. Connectivity and open access to computing power and information availability through the cloud promotes another era of industrialization and the need for more sophisticated consumer products.
- Reliability Enterprise applications are now so critical that they must be reliable and available to support 24/7 operations. In the event of failure or outages, contingency plans must take effect smoothly, and for disastrous or catastrophic failure, recovery plans must begin with minimum disruption. Each aspect of reliability should be carefully considered when engaging with a CSP, negotiated as part of the SLA, and tested in failover drills. Additional costs may be associated with the required levels of reliability; however, the business can do only so much to mitigate risks and the cost of a failure. Establishing a track record of reliability will be a prerequisite for widespread adoption.

- Commingled data - application sharing and multi-tenancy of data is one of the characteristics associated with cloud computing. Although many CSPs have multi-tenant applications that are secure, scalable and customizable, security and privacy issues are still often concerns among enterprises. Data encryption is another control that can assist data confidentiality.
- Cloud security policy / procedures transparency - some CSPs may have less transparency than others about their information security policy. The rationalization for such difference is the policies may be proprietary. As a result, it may create conflict with the enterprise‘s information compliance requirement. The enterprise needs to have detailed understanding of

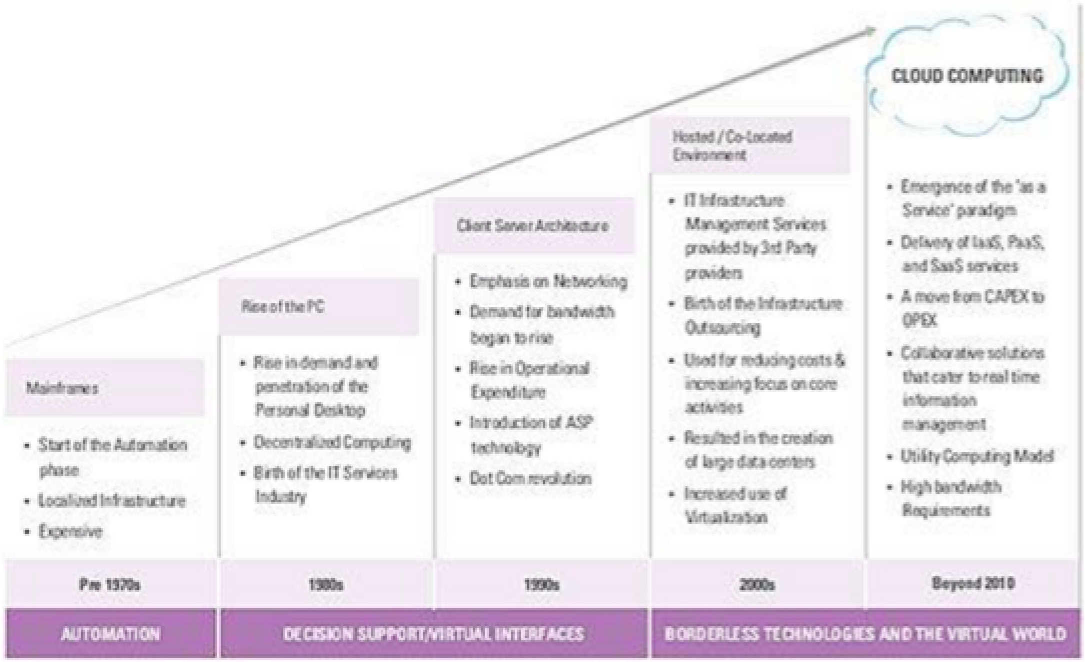
the service level agreements (SLAs) that stipulated the desired level of security provided by the CSPs. Because cloud computing represents a new computing model, there is a great deal of uncertainty about how security at all levels (e.g., network, host, application, and data levels) can be achieved. That uncertainty has consistently led information executives to state that security is their number one concern with cloud computing. The ability of cloud computing to adequately address privacy regulations has been called into question. Organizations today face numerous different requirements attempting to protect the privacy of individuals' information, and it is not clear (i.e., not yet established) whether the cloud computing model provides adequate protection of such information, or whether organizations will be found in violation of regulations because of this new model.

- Cloud date ownership - in the contract agreements it may state that the CP owns the data stored in the cloud computing environment. The CSP may demand for significant service fees for data to be returned to the enterprise when the cloud computing SLAs terminates.
- Lock-in with CSP's proprietary application programming interfaces (APIs) - currently many CSPs implement their application by adopting the proprietary APIs. As a result, cloud services transition from one CSP to another CSP, has become extremely complicated, time-consuming and labour-intensive.
- Compliance requirements - today's cloud computing services, can challenge various compliance audit requirements currently in place. Data location; cloud computing security policy transparency; and IAM, are all challenging issues in compliance auditing efforts. Examples of the compliance requirement including privacy and PII laws; Payment Card Industry (PCI) requirements; and financial reporting laws.
- Disaster recovery - it is a concern of enterprises about the resiliency of cloud computing, since data may be commingled and scattered around multiple servers and geographical areas. It may be possible that the data for a specific point of time cannot be identified. Unlike traditional hosting, the enterprise knows exactly where the location is of their data, to be rapidly retrieved in the event of disaster recovery. In the cloud computing model, the primary CSP may outsource capabilities to third parties, who may also outsource the recovery process. This will become more complex when the primary CSP does not ultimately hold the data.

- Interoperability The interoperability and portability of information between private clouds and public clouds are critical enablers for broad adoption of cloud computing by the enterprise. Many companies have made considerable progress toward standardizing their processes, data, and systems through implementation of ERPs. This process has been enabled by scalable infrastructures to create single instances, or highly integrated connections between instances, to manage the consistency of master and transaction data and produce reliable consolidated information. Even with these improved platforms, the speed at which businesses change may still outpace the ability of IT organizations to respond to these changes. SaaS applications delivered through the cloud provide a low-capital, fast-deployment option. Depending on the application, it is critical to integrate with traditional applications that may be resident in a separate cloud or on traditional technology. The standard for interoperability is either an enabler or a barrier to interoperability, and permits maintenance of the integrity and consistency of a company‘s information and processes.
- Economic Value The growth of cloud computing is predicated on the return on investment that accrues. It seems intuitive that by sharing resources to smooth out peaks, paying only for what is used, and cutting upfront capital investment in deploying IT solutions, the economic value will be there. There will be a need to carefully balance all costs and benefits associated with cloud computing—in both the short and long terms. Hidden costs could include support, disaster recovery, application modification, and data loss insurance. There will be threshold values whereby consolidating investments or combining cloud services makes sense; for example, it might not be efficient or cost effective to utilize multiple autonomous SaaS applications. Each may contract for disaster recovery program services. There is a point where economies of scale mean these functions should be combined in a similar service. Application usage may begin with a low volume of transactions that can be supported with semi-automated master data management. As usage expands and interoperability requirements for the business process become more onerous, a new approach is needed. This evolution may be the most cost-effective approach; however, there is a risk that the business transition costs from one solution to another may change the cost and benefit equation, and hence the solution that should be employed.

- Changes in the IT Organization The IT organization will be affected by cloud computing, as has been the case with other technology shifts. There are two dimensions to shifts in technology. The first is acquiring the new skill sets to deploy the technology in the context of solving a business problem, and the second is how the technology changes the IT role. During the COBOL era, users rarely programmed, the expectations of the user interface varied, and the adaptability of the solution was low. Training was delivered in separate manuals and the user used the computer to solve problems only down predefined paths. With the advent of fourth-generation languages, roles within IT, such as system analyst and programmer, became merged into analyst/programmer, users started to write their own reports, and new applications, including operational data stores, data entry, and query programs, could be rapidly deployed in weeks. IT‘s role will change once again: the speed of change will impact the adoption of cloud technologies and the ability to decompose mature solutions from hype to deliver real value from cloud technology; and the need to maintain the controls to manage IT risk in the business will increase.
- Political Issues Due to Global Boundaries In the cloud computing world, there is variability in terms of where the physical data resides, where processing takes place, and from where the data is accessed. Given this variability, different privacy rules and regulations may apply. Because of these varying rules and regulations, by definition politics becomes an element in the adoption of cloud computing, which is effectively multijurisdictional. For cloud computing to continually evolve into a borderless and global tool, it needs to be separated from politics. **Currently, some major global technological and political powers are making laws that can have a negative impact on the development of the global cloud.** For example, as a result of the USA Patriot Act, Canada has recently asked that its government not use computers in the global network that are operating within U.S. borders, fearing for the confidentiality and privacy of the Canadian data stored on those computers. Providers have been unable to guarantee the location of a company's information on specified set of servers in a specified location. However, cloud computing service providers are rapidly adopting measures to handle this issue. Cloud computing depends largely on global politics to survive. Politics are affecting the scalability of the Internet, the availability of Internet access, the free flow of information, and the cloud-based global economy on a daily basis.

1.5 Evolution of Cloud Computing



- **Mainframes**. These were the first examples of large computational facilities leveraging multiple processing units. Mainframes were powerful, highly reliable computers specialized for large data movement and massive input/output (I/O) operations. They were mostly used by large organizations for bulk data processing tasks such as online transactions, enterprise resource planning, and other operations involving the processing of significant amounts of data. Even though mainframes cannot be considered distributed systems, they offered large computational power by using multiple processors, which were presented as a single entity to users. One of the most attractive features of mainframes was the ability to be highly reliable computers that were —always on‖ and capable of tolerating failures transparently. No system shutdown was required to replace failed components, and the system could work without interruption. Batch processing was the main application of mainframes. Now their popularity and deployments have reduced, but evolved versions of such systems are still in use for transaction processing (such as online banking, airline ticket booking, supermarket and telcos, and government services).
- **Clusters -** Cluster computing started as a low-cost alternative to the use of mainframes and supercomputers. The technology advancement that created faster and more powerful mainframes and supercomputers eventually generated an increased availability of cheap

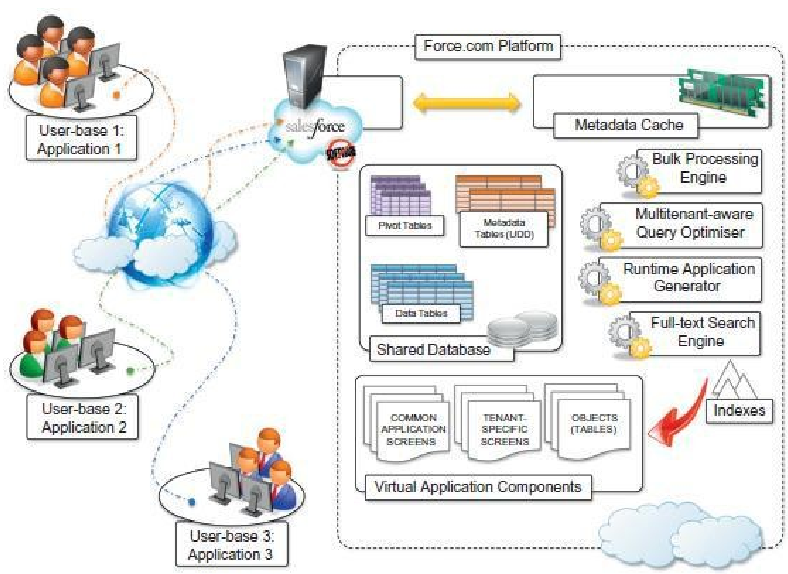
commodity machines as a side effect. These machines could then be connected by a high-bandwidth network and controlled by specific software tools that manage them as a single system. Starting in the 1980s, clusters become the standard technology for parallel and high-performance computing. Built by commodity machines, they were cheaper than mainframes and made high-performance computing available to a large number of groups, including universities and small research labs. Cluster technology contributed considerably to the evolution of tools and frameworks for distributed computing, including Condor, Parallel Virtual Machine (PVM), and Message Passing Interface (MPI). One of the attractive features of clusters was that the computational power of commodity machines could be leveraged to solve problems that were previously manageable only on expensive supercomputers. Moreover, clusters could be easily extended if more computational power was required.

- **Distributed Systems**- A distributed system is a collection of independent computers that appears to its users as a single coherent system. Distributed systems often exhibit other properties such as heterogeneity, openness, scalability, transparency, concurrency, continuous availability, and independent failures. To some extent these also characterize clouds, especially in the context of scalability, concurrency, and continuous availability
- **Grids**. Grid computing appeared in the early 1990s as an evolution of cluster computing. In an analogy to the power grid, grid computing proposed a new approach to access large computational power, huge storage facilities, and a variety of services. Users can ―consume‖ resources in the same way as they use other utilities such as power, gas, and water. Grids initially developed as aggregations of geographically dispersed clusters by means of Internet connections. These clusters belonged to different organizations, and arrangements were made among them to share the computational power. Different from a ―large cluster,‖ a computing grid was a dynamic aggregation of heterogeneous computing nodes, and its scale was nationwide or even worldwide. Several developments made possible the diffusion of computing grids: (a) clusters became quite common resources; (b) they were often underutilized; (c) new problems were requiring computational power that went beyond the capability of single clusters; and (d) the improvements in networking and the diffusion of the Internet made possible long-distance, high-bandwidth connectivity. All these elements led to the development of grids, which now serve a multitude of users across the world.

1.6 Usage scenarios and applications

a) **Customer relationship management (CRM) and enterprise resource planning (ERP)** applications are market segments that are flourishing in the cloud, with CRM applications the more mature of the two. Cloud CRM applications constitute a great opportunity for small enterprises and start-ups to have fully functional CRM software without large up-front costs and by paying subscriptions. Moreover, CRM is not an activity that requires specific needs, and it can be easily moved to the cloud. Such a characteristic, together with the possibility of having access to business and customer data from everywhere and from any device, has fostered the spread of cloud CRM applications. ERP solutions on the cloud are less mature and have to compete with well-established in-house solutions. ERP systems integrate several aspects of an enterprise: finance and accounting, human resources, manufacturing, supply chain management, project management, and CRM. Their goal is to provide a uniform view and access to all operations that need to be performed to sustain a complex organization. Because of the organizations that they target, the transition to cloud-based models is more difficult: the cost advantage over the long term might not be clear, and the switch to the cloud could be difficult if organizations already have large ERP installations. For this reason cloud ERP solutions are less popular than CRM solutions at this time.

Salesforce.com



Salesforce.com is probably the most popular and developed CRM solution available today. As of today more than 100,000 customers have chosen Safesforce.com to implement their CRM solutions. The application provides customizable CRM solutions that can be integrated with additional features developed by third parties. Salesforce.com is based on the Force.com cloud development platform. This represents scalable and high-performance middleware executing all the operations of all Salesforce.com applications. Initially designed to support scalable CRM applications, the platform has evolved to support the entire life cycle of a wider range of cloud applications by implementing a flexible and scalable infrastructure.

At the core of the platform resides its metadata architecture, which provides the system with flexibility and scalability. Rather than being built on top of specific components and tables, application core logic and business rules are saved as metadata into the Force.com store. Both application structure and application data are stored in the store. A runtime engine executes application logic by retrieving its metadata and then performing the operations on the data. Although running in isolated containers, different applications logically share the same database structure, and the runtime engine executes all of them uniformly. A full-text search engine supports the runtime engine. This allows application users to have an effective user experience despite the large amounts of data that need to be crawled.

The search engine maintains its indexing data in a separate store and is constantly updated by background processes triggered by user interaction. Users can customize their application by leveraging the ―native‖ Force.com application framework or by using programmatic APIs in the most popular programming languages.

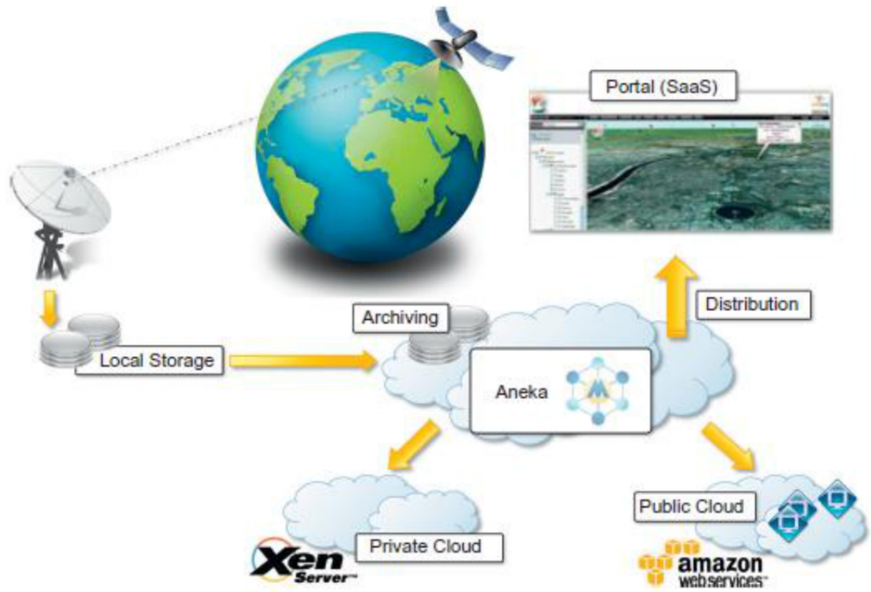
The application framework allows users to visually define either the data or the core structure of a Force.com application, while the programmatic APIs provide them with a more conventional way for developing applications that relies on Web services to interact with the platform. Customization of application processes and logic can also be implemented by developing scripts in APEX. This is a Java-like language that provides object-oriented and procedural capabilities for defining either scripts executed on demand or triggers. APEX also offers the capability of expressing searches and queries to have complete access to the data managed by the Force.com platform.

- Microsoft dynamics CRM- Microsoft Dynamics CRM is the solution implemented by Microsoft for customer relationship management. Dynamics CRM is available either for installation on the enterprise's premises or as an online solution priced as a monthly per-user subscription. The system is completely hosted in Microsoft's datacenters across the world and offers to customers a 99.9% SLA, with bonus credits if the system does not fulfill the agreement. Each CRM instance is deployed on a separate database, and the application provides users with facilities for marketing, sales, and advanced customer relationship management. Dynamics CRM Online features can be accessed either through a Web browser interface or programmatically by means of SOAP and RESTful Web services. This allows Dynamics CRM to be easily integrated with both other Microsoft products and line-of-business applications. Dynamics CRM can be extended by developing plug-ins that allow implementing specific behaviors triggered on the occurrence of given events. Dynamics CRM can also leverage the capability of Windows Azure for the development and integration of new features.

b) Geoscience: satellite image processing

Geoscience applications collect, produce, and analyze massive amounts of geospatial and nonspatial data. As the technology progresses and our planet becomes more instrumented (i.e., through the deployment of sensors and satellites for monitoring), the volume of data that needs to be processed increases significantly. In particular, the geographic information system (GIS) is a major element of geoscience applications. GIS applications capture, store, manipulate, analyze, manage, and present all types of geographically referenced data. This type of information is now becoming increasingly relevant to a wide variety of application domains: from advanced farming to civil security and natural resources management. As a result, a considerable amount of geo-referenced data is ingested into computer systems for further processing and analysis.

Cloud computing is an attractive option for executing these demanding tasks and extracting meaningful information to support decision makers. Satellite remote sensing generates hundreds of gigabytes of raw images that need to be further processed to become the basis of several different GIS products. This process requires both I/O and compute-intensive tasks. Large images need to be moved from a ground station's local storage to compute facilities, where several transformations and corrections are applied. Cloud computing provides the appropriate infrastructure to support such application scenarios.



A cloud-based implementation of such a workflow has been developed by the Department of Space, Government of India. The system integrates several technologies across the entire computing stack. A SaaS application provides a collection of services for such tasks as geocode generation and data visualization. At the PaaS level, Aneka controls the importing of data into the virtualized infrastructure and the execution of image-processing tasks that produce the desired outcome from raw satellite images. The platform leverages a Xen private cloud and the Aneka technology to dynamically provision the required resources (i.e., grow or shrink) on demand. This demonstrates how cloud computing technologies can be effectively employed to offload local computing facilities from excessive workloads and leverage more elastic computing infrastructures.

1.7 Business model around Cloud

The business model canvas has nine basic building blocks and specific relations between those building blocks. The rest of this article describes each of them, and gives a brief example of how they apply to a cloud provider proposition. As the main cloud provider, user will use Amazon Web Services (AWS), in particular EC2 (virtual machines on demand). This is an Infrastructure as a Service offering. The power of the business model canvas approach will become clear if we see how it can distinguish between various cloud service offerings and traditional IT.

Customer segments (CS)

In the Business Model Canvas, ―Customer Segments‖ are the groups of customers that the company ultimately serves, I.e. the ones that consume and pay for the services. In the AWS case, although basically anybody with a credit card can spin up a virtual machine, it looks like Amazon is primarily targeting software developers and (startup) SaaS providers as the main customers. Historically, the Amazon development teams were the first customers. External customers were initially added as an afterthought.

Value Propositions (VP)

The value propositions reflect the customer problems and needs. This is the central element that describes why the customer would ultimately pay for the product or service.The value proposition of cloud computing centers around its five essential characteristics. For example in the AWS EC2 case, the core component of the value proposition is rapid self-service provisioning of virtual machines with pay per use billing. For each individual customer these translate into different business advantages. An example is reduced capital expenditure and reduced risk of over-investing or under-provisioning.

Channels (CH)

Value propositions are delivered to customers through communications, distribution and sales channels. It is often assumed that cloud computing relies solely self-service direct sales, but the reality is much more diverse. SaaS providers in particular are developing extensive partner programs.

AWS primarily employs a self-service direct model, where the delivery is through APIs. AWS also provides a web user interface to those APIs. Interestingly, that interface used to lag in functionality behind the main AWS services, but these days most new features are announced on the API and the Web UI simultaneously. The model is enhanced by premium support.

Customer Relationships (CR)

Customer relations are established and maintained with each specific customer segment. One of the ways that AWS maintains relationships with its customer segments is through conferences.

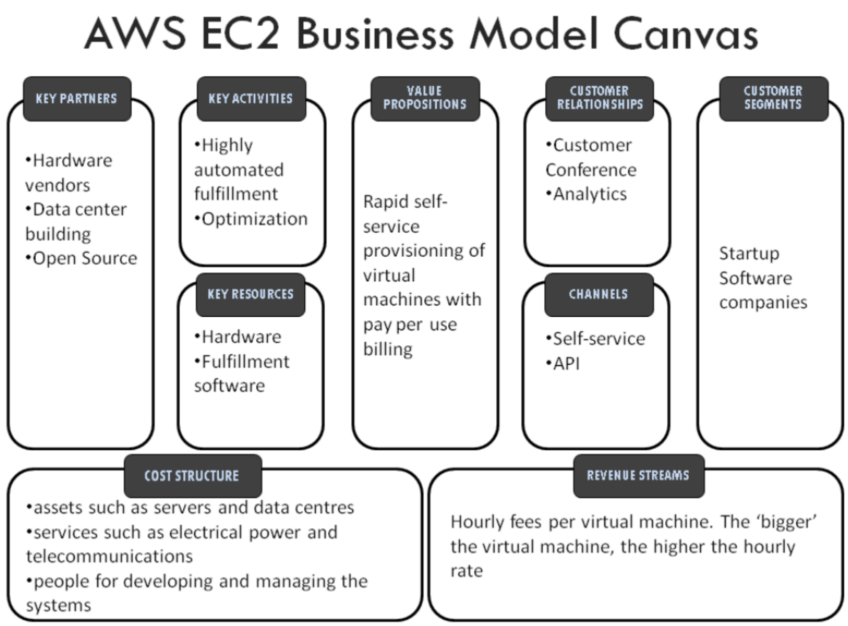
The 2013 re:Invent developer conference attracted 9000 visitors. Additionally, there are vibrant on-line communities. Thus AWS does extensive analytics on the activity that customers have on the platform.

Revenue Streams (RS)

Revenue streams are the result of value propositions that are successfully offered to customers.The structure of revenue streams is where cloud computing differs from earlier IT service models, as they are usage based rather than asset based.AWS basically charges hourly fees per virtual machine. The _bigger‘ the virtual machine, the higher the hourly rate.

Key Resources (KR)

Key resources are the assets required to offer and deliver the previously mentioned elements (e.g. value proposition, customer relationships). AWS owns massive amounts of hardware, estimated at 1 million servers or more. That is housed in dozens of data-centers worldwide. But there is more. The service can only be delivered through advanced and unique fulfillment software and processes. Amazon must have invested very substantially in that.



Key Activities (KA)

The key resources perform key activities. At AWS the key activity, delivery, is highly automated. But at the AWS scale, oversight and resources planning is still a serious effort. Optimizing assets versus utilization is very essential in the IaaS business model. Through economies of scale, AWS is able to spend a lot of effort on these activities.

Key Partnerships (KP)

Some activities are outsourced, and some resources are acquired outside the enterprise. AWS buys immense amounts of hardware, and uses a lot of (open source) software. Building out data centers is also likely to be outsourced.

Cost Structure (CS)

All business model elements result in a cost structure. In more traditional IT service models the revenue streams are tightly coupled to the cost structure. The cloud computing innovation is also about decoupling these. At AWS the main cost elements are in assets such as servers and data centers, in services such as electrical power and telecommunications, and in people for eveloping and managing the systems.

1.8 Major Players of Cloud Computing

Many **players** make up the world of cloud computing:

- The **vendors** providing applications and enabling technology, infrastructure, hardware, and integration
- The **partners** of these vendors that are creating cloud services offerings and providing support services to customers
- The **business leaders** themselves who are either using or evaluating various types of cloud computing offerings

The world of the cloud has lots of **participants:**

- The end user doesn‘t really have to know anything about the underlying technology. In small businesses, for example, the cloud provider becomes the de facto data center. In larger organizations, the IT organization oversees the inner workings of both internal resources and external cloud resources.

- Business management needs to take responsibility for overall governance of data or services living in a cloud. Cloud service providers must provide a predictable and guaranteed service level and security to all their constituents.
- The cloud service provider is responsible for IT assets and maintenance.

Cloud services must enable multi-tenancy - Different companies sharing the same underlying resources. Companies are finding some important new value in cloud services. The cloud can eliminate many of the complex constraints from the traditional computing environment, including space, time, power, and cost.

- Cloud services like social networks (such as Facebook or LinkedIn) and collaboration tools (like video conferencing, document management, and webinars) are changing the way people in businesses access, deliver, and understand information.
- Cloud computing infrastructures make it easier for companies to treat their computing systems as a pool of resources rather than a set of independent environments that each has to be managed.

Vendor/Product	Service Type	Description
Amazon Web Services	IaaS, PaaS, SaaS	Amazon Web Services (AWS) is a collection of Web services that provides developers with compute, storage, and more advanced services. AWS is mostly popular for IaaS services and primarily for its elastic compute service EC2.
Google AppEngine	PaaS	Google AppEngine is a distributed and scalable runtime for developing scalable Web applications based on Java and Python runtime environments. These are enriched with access to services that simplify the development of applications in a scalable manner.
Microsoft Azure	PaaS	Microsoft Azure is a cloud operating system that provides services for developing scalable applications based on the proprietary Hyper-V virtualization technology and the .NET framework.
SalesForce.com and Force.com	SaaS, PaaS	SalesForce.com is a Software-as-a-Service solution that allows prototyping of CRM applications. It leverages the Force.com platform, which is made available for developing new components and capabilities for CRM applications.
Heroku	PaaS	Heroku is a scalable runtime environment for building applications based on Ruby.
RightScale	IaaS	Rightscale is a cloud management platform with a single dashboard to manage public and hybrid clouds.

Cloud Business Leaders

- Microsoft
- Amazon
- Google
- VMware
- IBM
- Gartner

Cloud Business Partners

- Cloud Solutions (Microsoft)
- Cloud Technologies (Microsoft)
- SQLStream (Google)
- Cascadeo (Google)

1.9 Eucalyptus

Eucalyptus is acronym standing for **Elastic Utility Computing Architecture for Linking Your Programs To Useful Systems**. Eucalyptus is a system for implementing on-premise **private and hybrid clouds**, using the hardware and software infrastructure that‘s in place, without modification. In effect, it‘s an add-on capability for data center virtualization to create genuine cloud capability such as self-service provisioning, security, performance management, and end-user customization.

Eucalyptus is **open source**, so the software can be downloaded free and it is also shipped with the Ubuntu 9.04 (and later) distribution of Linux. It is thus becoming the default open-source cloud capability. It is implemented by using commonly available Linux tools and basic Web service technologies. The current interface to Eucalyptus is **compatible with Amazon’s EC2, S3, and Elastic Block Store (EBS)** — a storage area network (SAN) in the cloud — interfaces, so it is possible to create a private cloud by using Eucalyptus with the intention of moving some or all of it onto EC2.

1.10 Nimbus

Nimbus is a toolkit that, once installed on a cluster, provides an infrastructure as a service cloud to its client via WSRF-based or Amazon EC2 WSDL web service APIs. Nimbus is free and open-source software, subject to the requirements of the Apache License, version 2.

Nimbus supports both hypervisors like Xen and KVM and virtual machine schedulers Portable Batch System and Oracle Grid Engine. It allows deployment of self-configured virtual clusters via contextualization. It is configurable with respect to scheduling, networking leases, and usage accounting.

Nimbus overcomes these challenges like users need to find ways to allow

- the on-demand resources to share security and configuration context
- manage the deployment of potentially diverse platform
- Ensure reliability and scalability in the environment, etc.

Its aim is to enable users to move to the cloud quickly and effortlessly, automating and facilitating much of the process.

It also aims to provide a bridge allowing a user to overlay familiar concepts, such as virtual clusters, onto the resources provisioned in the cloud.

- **cloudinit.d** is a tool for launching, controlling, and monitoring cloud applications. If the application is simple or complex, single cloud or multi-cloud, VM based or bare metal, or any combination of the above, cloudinit.d is designed to make the management and coordination of that application easy.
- **Phantom** is a service that provides auto-scaling and high availability for collections of resources deployed over multiple IaaS cloud providers allowing users to develop scalable and reliable applications.
- The **Context Broker** is a service that allows clients to coordinate large virtual cluster launches automatically and repeatable

1.11 Open Nebula

OpenNebula provides the most simple but feature-rich and flexible solution for the comprehensive management of virtualized data centers to enable private, public and hybrid IaaS clouds. OpenNebula interoperability makes cloud an evolution by leveraging existing IT assets, protecting your investments, and avoiding vendor lock-in. Features include

- Openness of the architecture, interfaces, and code
- Flexibility to fit into any datacenter
- Interoperability and portability to prevent vendor lock-in
- Stability for use in production enterprise-class environments
- Scalability for large scale infrastructures
- SysAdmin-centrism with complete control over the cloud
- Simplicity, easy to deploy, operate and use
- Lightness for high efficiency

1.12 CloudSim

It supports for modeling and simulation of large scale Cloud computing data centers & federated clouds. It support for modeling and simulation of virtualized server hosts, with customizable policies for provisioning host resources to virtual machines. It support for modeling and simulation of energy-aware computational resources. It has support for modeling and simulation of data center network topologies and message-passing applications. It has support for dynamic insertion of simulation elements, stop and resume of simulation. It also supports user-defined policies for allocation of hosts to virtual machines and policies for allocation of host resources to virtual machines