
Cloud Computing Architecture

Chapter 4 - Cloud Computing Architecture

After completing this unit you should be able to understand

- *Cloud reference model*
- *Architecture of following service models*
 - *IaaS*
 - *PaaS*
 - *SaaS*
- *Types of Clouds and their structure*
- *Economics of the cloud*
- *open challenges of cloud computing*

Introduction



- The term *Cloud computing* is a wide umbrella encompassing many different things; lately it has become a buzzword easily misused to revamp existing technologies and ideas for the public.
- Utility-oriented Data Centers are the first outcome of Cloud computing and they serve as the infrastructure through which the services are implemented and delivered.
- Any Cloud service, whether it is virtual hardware, development platform, or application software, relies on a distributed infrastructure owned by the provider or rented from a third party.
- As it can be noticed from the previous definition, the characterization of a Cloud is quite general: it can be implemented by using a datacenter, a collection of clusters, or a heterogeneous distributed system composed by desktop PC, workstations, and servers.
 - *Cloud computing is a utility oriented and Internet centric way of delivering IT services on demand. These services cover the entire computing stack: from the hardware infrastructure packaged as a set of virtual machines to software services such as development platforms and distributed applications.*

Cloud Reference Model

- Cloud computing supports any IT service that can be consumed as a utility and it is delivered through the network, most likely the Internet. Such characterization includes quite different aspects: infrastructure, development platforms, application and services.

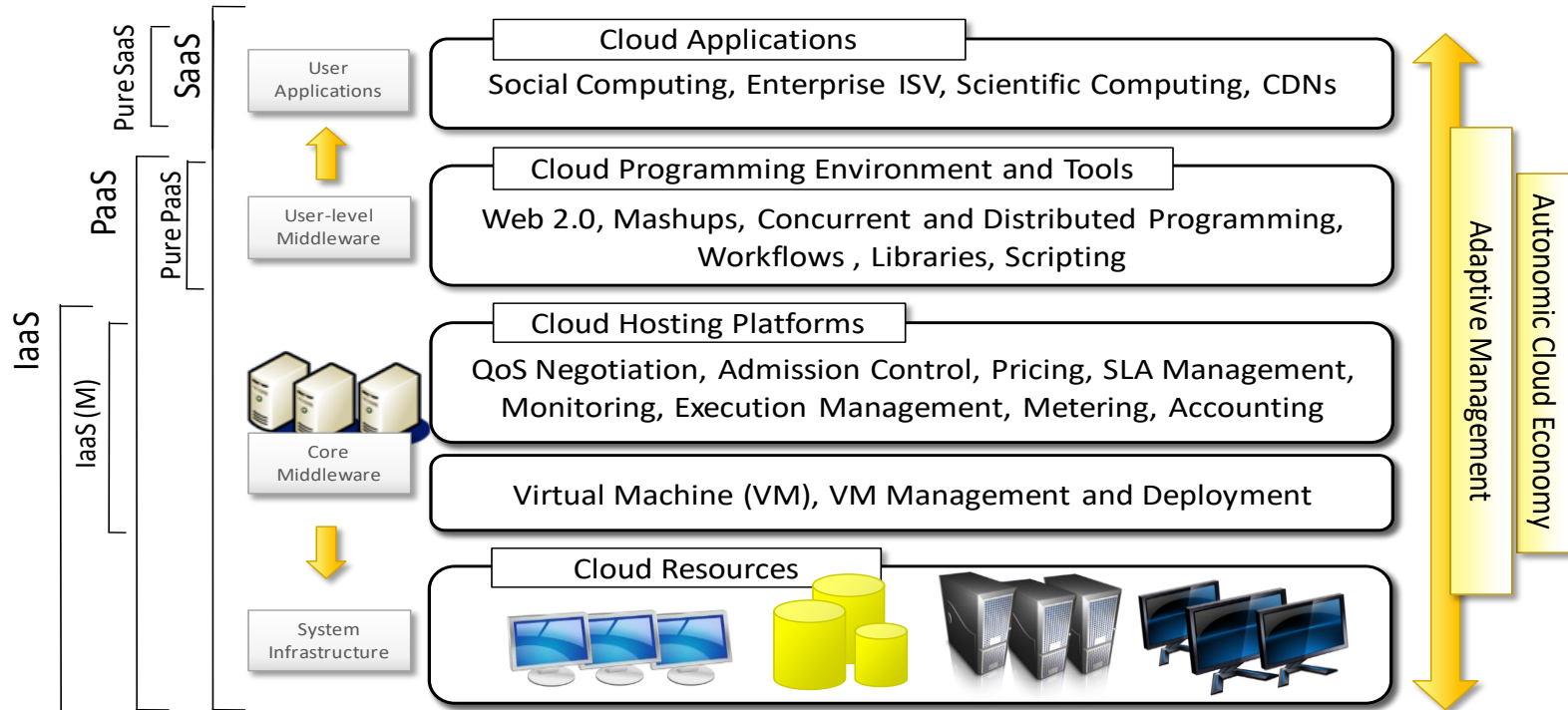


Fig-4.1- Cloud Computing Architecture

Architecture

- It is possible to organize all the concrete realizations of Cloud computing into a layered view covering the entire stack (see Figure in last slide): from the hardware appliances to software systems.
- Cloud resources are harnessed to offer “computing/horse power” required for providing services.
- Often, this layer is implemented by using a datacenter in which hundreds and thousands of nodes are stacked together. Cloud infrastructure can be in heterogeneous nature as variety of resources such as clusters and even networked PCs can be used to build it.
- Moreover, database systems and other storage services can also be part of the infrastructure.
- The physical infrastructure is managed by the core middleware whose objectives are to provide an appropriate runtime environment for applications and to utilize resources at best.
- At the bottom of the stack, virtualization technologies are used to guarantee runtime environment customization, application isolation, sandboxing, and quality of service.
- Hardware virtualization is the most commonly used at this level. Hypervisors manage the pool of resources and expose the distributed infrastructure as a collection of virtual machines.
- By using virtual machine technology it is possible to finely partition the hardware resources such as CPU, memory, and also virtualize specific devices, thus meeting the requirement of users and applications.

Architecture

- The combination of Cloud hosting platforms and resources is generally classified as a *Infrastructure-as-a-Service (IaaS)* solution.
- Infrastructure-as-a-Service solutions are suitable for designing the system infrastructure but provide limited services to build applications.
- Such service is provided by Cloud programming environments and tools, which form a new layer for offering to users a development platform for applications.
- The range of tools include web based interfaces, command line tools, and frameworks for concurrent and distributed programming.
- In this scenario, users develop their applications specifically for the Cloud by using the API exposed at the user-level middleware. For this reason, this approach is also known as *Platform-as-a-Service (PaaS)*, because the service offered to the user is a development platform rather than an infrastructure.
- PaaS solutions generally include the infrastructure as well, that is bundled as part of the service provided to users. In case of *Pure PaaS* only the user level middleware is offered and it has to be complemented with a virtual or physical infrastructure.
- The top layer of the reference model depicted in Figure in previous slide contains services delivered at application level. These are mostly referred as *Software-as-a-Service (SaaS)*.

Architecture

- In most of the cases these are web based applications that rely on the Cloud to provide service to end users.
- The horse power of the Cloud provided by IaaS and PaaS solutions allows independent software vendors to deliver their application services over the Internet.
- Other applications belonging to this layer are those strongly leveraging the Internet for their core functionalities that rely on the Cloud to sustain a larger number of users; this is the case of gaming portals and in general social networking web sites.
- As a vision, any service offered in the Cloud computing style, should be able to adaptively change and expose an autonomic behavior; in particular for its availability and performance.
- The reference model described in Figure in last slide also introduces the concept of *everything as a Service (XaaS)*. This is one of the most important elements of Cloud computing: Cloud services from different providers can be composed together in order to provide a completely integrated solution covering all the computing stack of a system.
- Infrastructure-as-a-Service providers can offer the bare metal in terms of virtual machines where Platform-as-a-Service solutions are deployed.
- Table 4.1 in next slide summarizes the characteristics of the three major categories used to classify Cloud computing solutions.

Cloud Computing Services Classification



Table-4.1: Cloud Computing Services Classification

Category	Characteristics	Product Type	Vendors & Products
<i>SaaS</i>	Customers are provided with applications that are accessible anytime and from anywhere.	Web applications and services (Web 2.0)	SalesForce.com (CRM); Clarizen.com (Project Management); Google Apps...
<i>PaaS</i>	Customers are provided with a platform for developing applications hosted in the Cloud.	Programming APIs and frameworks; Deployment Systems.	Google AppEngine; Microsoft Azure; Manjrasoft Aneka; Data Synapse...
<i>IaaS/HaaS</i>	Customers are provided with virtualized hardware and storage on top of which they can build their infrastructure.	Virtual machines management infrastructure; Storage management; Network management.	Amazon EC2 and S3; GoGrid; Nirvanix...

Infrastructure / Hardware as a Service

- Figure 4.2 provides an overall view of the components forming an Infrastructure-as-a-Service solution.
- It is possible to distinguish three principal layers: the *physical infrastructure*, the *software management infrastructure*, and the *user interface*.
- At the top layer the user interface provides access to the services exposed by the software management infrastructure. Such interface is generally based on Web 2.0 technologies: web services, RESTful APIs, and mash-ups.
- The core features of an Infrastructure-as-a-Service solution are implemented in the infrastructure management software layer.
- In particular, the management of the virtual machines is the most important function performed by this layer. A central role is played by the scheduler, which is in-charge of allocating the execution of virtual machine instances.

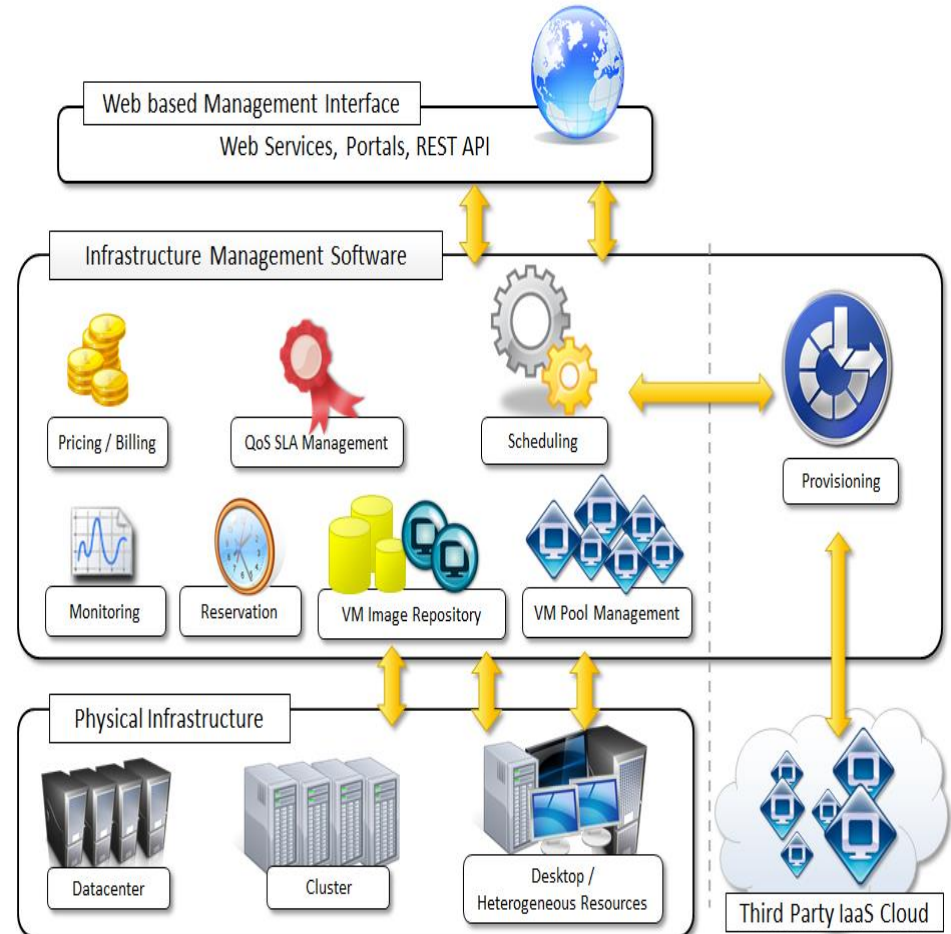


Fig-4.2-iaas

Infrastructure / Hardware as a Service

- A central role is played by the scheduler, which is in-charge of allocating the execution of virtual machine instances. The scheduler interacts with the other components performing different tasks:
 - The *pricing / billing* component takes care of the cost of executing each virtual machine instance and maintains data that will be used to charge the user.
 - The *monitoring* component tracks the execution of each virtual machine instance and maintains data required for reporting and analyzing the performance of the system.
 - The *reservation* component stores the information of all the virtual machine instances that have been executed or that will be executed in the future.
 - The *VM repository* component provides a catalog of virtual machine images that users can use to create virtual instances.
 - A *VM pool manager* component is responsible of keeping track of all the live instances.
- The bottom layer is constituted by the physical infrastructure on top of which the management layer operates.
- At the bottom of the scale it is also possible to consider a heterogeneous environment where different types of resources can be aggregated: PCs, workstations, and clusters.
- In case of complete IaaS solutions all the three levels are offered as service. This is generally the case of Public Clouds vendors, such as Amazon, GoGrid, Joyent, Rightscale,

Platform as a Service

- Platform-as-a-Service (PaaS) solutions provide a development and deployment platform for running applications in the Cloud.
- They constitute the middleware on top of which applications are built. A general overview of the features characterizing the PaaS approach is given in Figure 4.3.
- Application management is the core functionality of the middleware. PaaS implementations provide applications with a runtime environment and do not expose any service for managing the underlying infrastructure.
- They automate the process of deploying applications to the infrastructure, configuring applications components, provisioning and configuring supporting technologies such as load balancers and databases, and managing system change based on policies set by the user.
- From a user point of view, the core middleware exposes interfaces that allow programming and deploying applications on the Cloud.

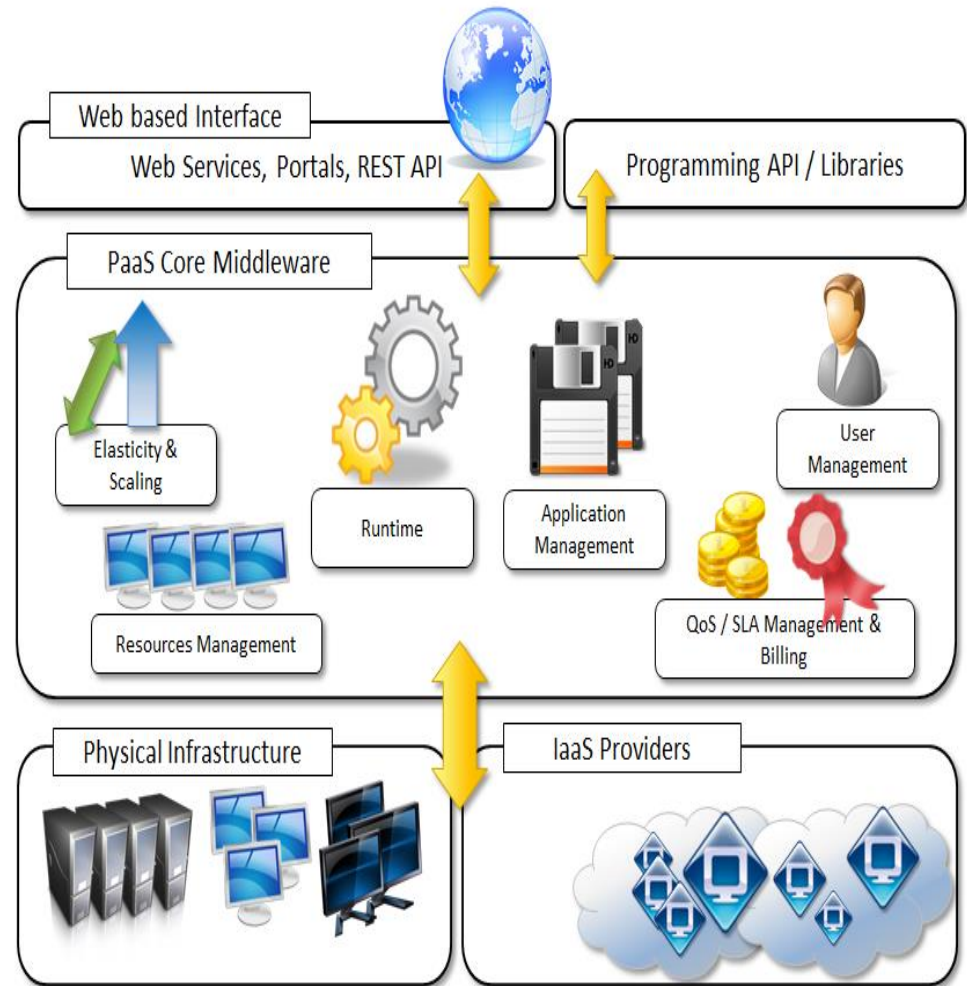


Fig-4.3-paas

Table: Platform as a Service Offering Classification

Table-4.2: Platform as a Service Offering Classification

☐

Essential characteristics of Platform-as-a-Service solution



- *Runtime framework.* It represents the “software stack” of the PaaS model and the most intuitive aspect that comes to the mind of people when referring to Platform-as-a-Service solutions. The runtime framework executes end-user code according to the policies set by the user and the provider.
- *Abstraction.* PaaS solutions are distinguished by the higher level of abstraction that they provide. PaaS the focus is on the applications the Cloud must support.
- *Automation.* PaaS environment automate the process of deploying applications to the infrastructure, scaling them by provisioning additional resources when needed.
- *Cloud services.* PaaS offerings provide developers and architects with services and APIs helping them to simplify the creation and delivery of elastic and highly available Cloud applications.
- The Platform-as-a-Service approach, when bundled with a underlying IaaS solutions, helps even small startup companies to quickly offer to customers integrated solutions on a hosted platform, at a very minimal cost. These opportunities make the PaaS offering a viable option targeting different market segments.

Software as a Service

- Software-as-a-Service (SaaS) is a software delivery model providing access to applications through the Internet as a web-based service.
- The SaaS model is appealing for applications serving a wide range of users and that can be adapted to specific needs with little further customization.
- This requirement characterizes Software-as-a-Service as a “one-to-many” software delivery model where an application is shared across multiple users.
- *“In the software as a service model, the application, or service, is deployed from a centralized data center across a network – Internet, Intranet, LAN, or VPN – providing access and use on a recurring fee basis. Users “rent”, “subscribe to”, “are assigned”, or “are granted access to” the applications from a central provider. Business models vary according to the level to which the software is streamlined, to lower price and increase efficiency, or value-added through customization to further improve digitized business processes.”*
- The analysis carried out by SIIA was mainly oriented to cover Application Service Providers (ASPs) and all their variations, which capture the concept of software applications consumed as a service in a broader sense. ASPs already had some of the core characteristics of SaaS:
 - The product sold to customer is *application access*.
 - The application is centrally managed.
 - The service delivered is *one-to-many*.
 - The service delivered is an integrated solution *delivered on the contract*, which means provided as promised.

Types of Clouds

- Clouds constitute the primary outcome of Cloud computing. They are a type of parallel and distributed system harnessing physical and virtual computers presented as a unified computing resource.
- Clouds build the infrastructure on top of which services are implemented and delivered to customers. Such an infrastructure can be of different types and provides useful information about the nature and the services offered by the Cloud.
- *Public Clouds*: the Cloud is open to the wide public;
- *Private Clouds*: the Cloud is implemented within the private premises of an institution and generally made accessible to the members of the institution or a subset of them;
- *Hybrid or Heterogeneous Clouds*: the Cloud is a combination of the two previous solution and most likely identifies a Private Cloud that has been augmented with resources or services hosted in a Public Cloud;
- *Community Clouds*: the Cloud is characterized by a multi-administrative domain, involving different deployment models (public, private, and hybrid), and it is specifically designed to address the needs of a specific industry.
- Almost all the implementations of Clouds can be classified in this categorization.

Public Clouds

- Public Clouds constitute the first expression of Cloud computing. They are a realization of the canonical view of Cloud computing where the services offered are made available to anyone, from anywhere, and at any time through the Internet.
- From a structural point of view they are a distributed system, most likely constituted by one or more datacenters connected together, on top of which the specific services offered by the Cloud are implemented.
- Any customer can easily sign-in with the Cloud provider, enter his/her credential and billing details, and use the services offered.
- A fundamental characteristic of Public Clouds is multi-tenancy. A Public Cloud is meant to serve a multitude of users and not a single customer. Any customer requires its virtual computing environment that is separated, and most likely isolated, from the other users. This is a fundamental requirement to provide an effective monitoring of user activities, guarantee the desired performance and the other Quality of Service attributes negotiated with users.
- A Public Cloud can offer any kind of service: infrastructure, platform, or applications. For example, Amazon EC2 is a Public Cloud providing infrastructure as a service, Google AppEngine is a Public Cloud providing an application development platform as a service, and Salesforce.com is a Public Cloud providing software as a service.

Private Clouds

- Private Clouds are virtual distributed systems that rely on a private infrastructure and provide internal users with dynamic provisioning of computing resources.
- Differently from Public Clouds, instead of a pay-as-you-go model, there could be other schemes in place, which take into account the usage of the Cloud and proportionally bill the different departments or sections of the enterprise.
- Private Clouds have the advantage of keeping in house the core business operations by relying on the existing IT infrastructure and reducing the burden of maintaining it once the Cloud has been set up.
- From an architectural point of view Private Clouds can be implemented on more heterogeneous hardware
- The physical layer is complemented with infrastructure management software (i.e. IaaS (M)), or a PaaS solution, according to the service delivered to the users of the Cloud.

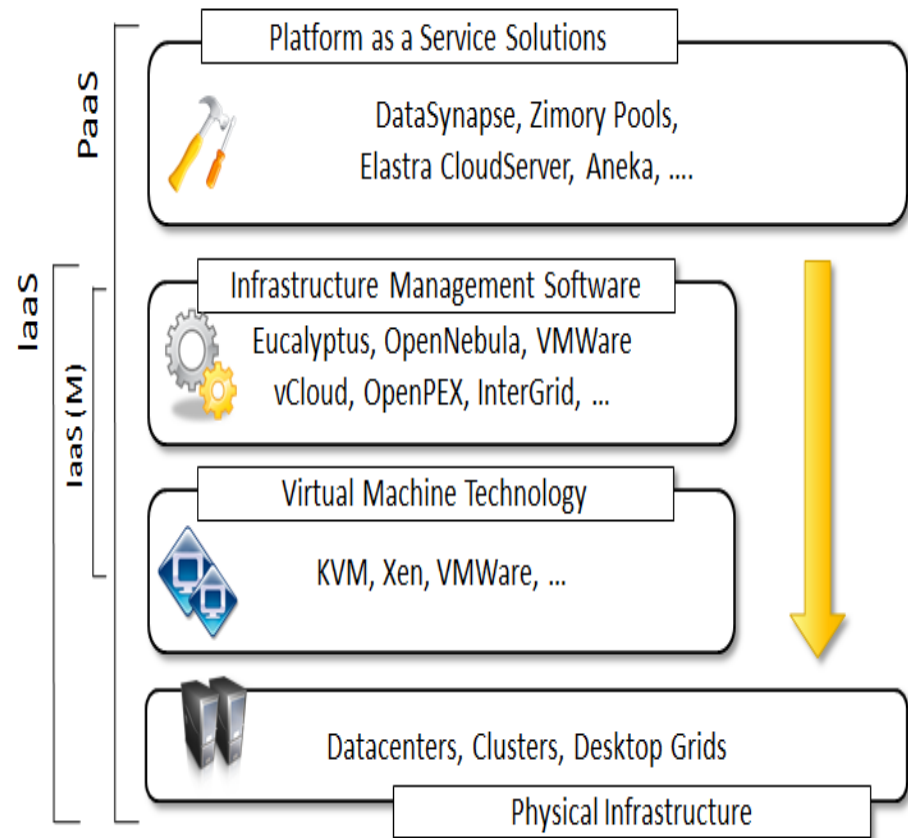


Fig-4.4-Private Clouds Hardware and Software Stack

Hybrid Clouds

- hybrid solution is an interesting opportunity for taking advantage of both of the two worlds. This led to the development and the diffusion of Hybrid Clouds.
- Hybrid Clouds allow exploiting existing IT infrastructures, maintaining sensitive information within the premises, and naturally growing and shrinking by provisioning external resources and releasing them when needed.
- it is a heterogeneous distributed system resulting from a Private Cloud that integrates additional services or resources from one or more Public Clouds. For this reason they are also called Heterogeneous Clouds.
- As depicted in the diagram, dynamic provisioning is a fundamental component in this scenario: Hybrid Clouds address scalability issues by leveraging external resources for exceeding capacity demand. These resources or services are temporarily leased for the time required and then released. This practice is also known as *Cloud-bursting*

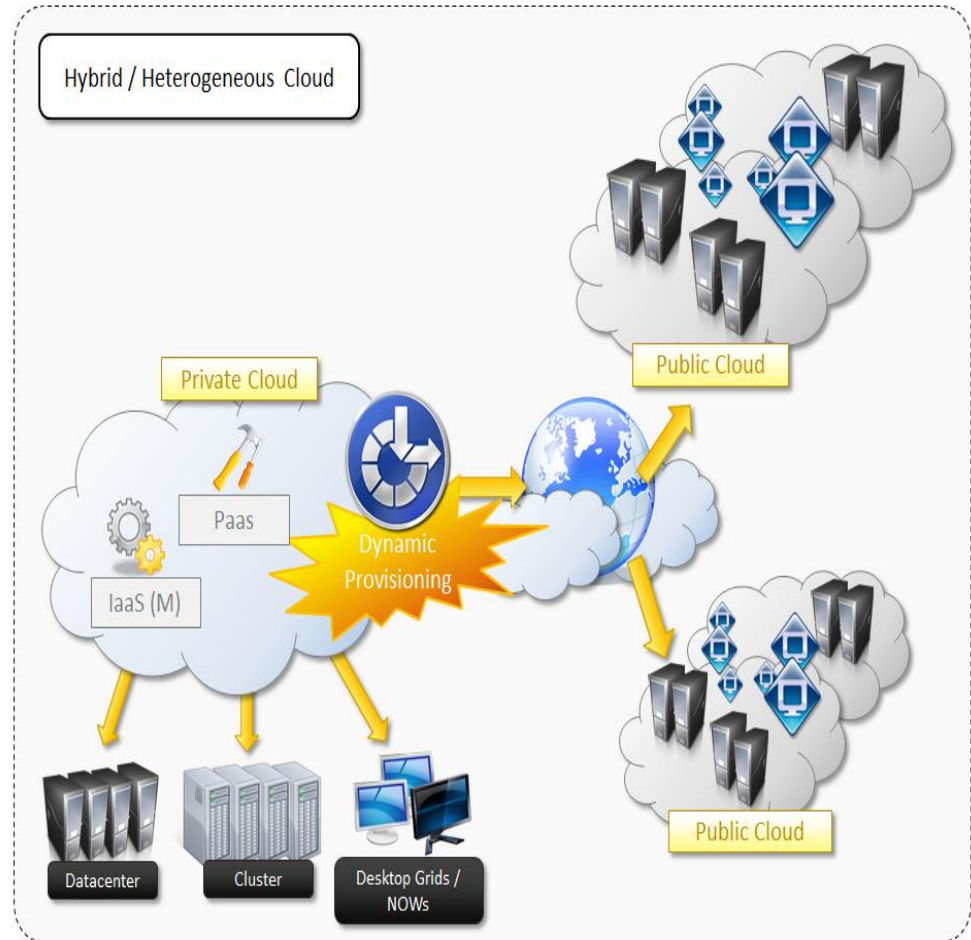


Fig-4.5- Hybrid / Heterogeneous Cloud Overview

Community Clouds

- Community Clouds are distributed system constituted by integrating the services of different Clouds to address the specific needs of an industry, a community, or a business sector.
- The NIST characterizes Community Clouds as follows:
 - *“The infrastructure is shared by several organizations and supports a specific community that has shared concerns (e.g. mission, security requirements, policy, and compliance considerations. It may be managed by the organizations or a third party and may exist on premise or off premise.”*
- From an architectural point of view, a Community Cloud is most likely implemented over multiple administrative domains. This means that different organizations such as government bodies, private enterprises, research organization, and even public virtual infrastructure providers, contribute with their resources to build the Cloud infrastructure.

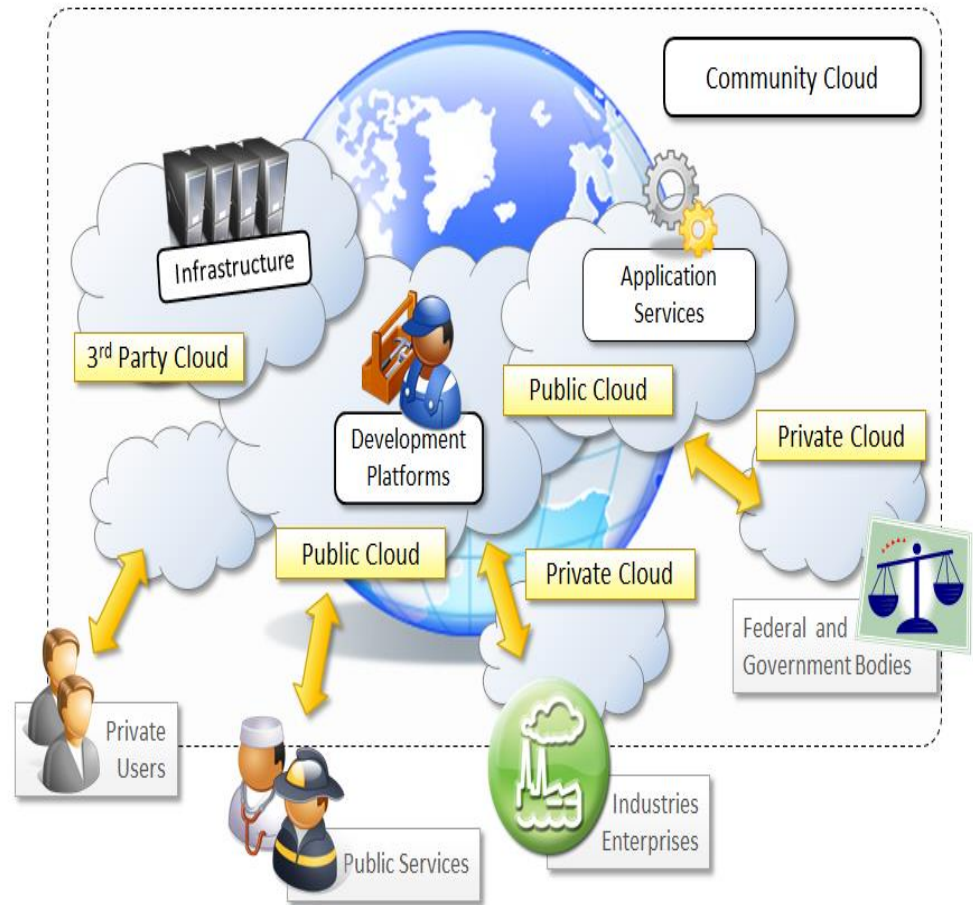


Fig-4.6- Community Cloud

Economics of the Cloud

- The main drivers of Cloud computing are: economy of scale and simplicity of software delivery and its operation. In fact, the biggest benefit of this phenomenon is financial: the *pay-as-you-go* model offered by Cloud providers. In particular, Cloud computing allows:
 - reducing the capital costs associated to the IT infrastructure;
 - eliminating the depreciation or lifetime costs associated with IT capital assets
 - replacing software licensing with subscriptions
 - cutting down the maintenance and administrative costs of IT resources
- A *capital cost* is the cost occurred in purchasing an asset that is useful in the production of goods or the rendering of services.
- The amount of cost savings that Cloud computing can introduce within an enterprise is related to the specific scenario in which Cloud services are used and how they contribute to generate a profit for the enterprise.
- In the case of a small startup starting its business it is possible to completely leverage the Cloud for many aspects such as:
 - IT infrastructure;
 - Software development;
 - CRM and ERP;

Open Challenges

- Being in its infancy, Cloud computing still presents a lot of challenges for the industry and the academia.
- the most important ones are:
- **Cloud Definition**
 - One of the most comprehensive formalization is noted in the National Institute of Standards and Technologies (NIST) working definition of Cloud computing [43]. It **characterizes** Cloud computing as: on-demand self-service, broad network access, resource pooling, rapid elasticity, and measured service; **classifies** services as: SaaS, PaaS, and IaaS; and **categorizes** deployment models as: public, private, community, and hybrid Clouds. The view is inline with our discussion and it shared by many IT practitioners and academics.
 - These characterizations and taxonomies reflect what is meant by Cloud computing at present time, but being in its infancy the phenomenon is constantly evolving and the same will happen to the attempts to capture the real nature of Cloud computing.
- **Cloud Interoperability and Standards**
- **Scalability and Fault Tolerance**
- **Security, Trust, and Privacy**
- **Organizational Aspects**

Review questions

- What does the acronym *XaaS* stand for?
- What are the fundamental components introduced in the Cloud Reference Model?
- What does Infrastructure-as-a-Service refer to?
- Which are the basic components of an IaaS based solution for Cloud computing?
- Provide some examples of IaaS implementations.
- What are the main characteristics of a Platform-as-a-Service solution?
- Describe the different categories of options available in PaaS market.
- What does the acronym SaaS mean? How does it relate to Cloud computing?
- Give the name of some popular Software-as-a-Service solutions?
- Classify the different types of Clouds.
- Give an example of Public Cloud.
- Which is the most common scenario for a Private Cloud?
- What kind of needs is addressed by Heterogeneous Clouds?
- Describe the fundamental features of the economic and business model behind Cloud computing.
- How does Cloud computing help to reduce the time to market for applications and to cut down capital expenses?
- List some of the challenges in Cloud computing.

Conclusions

In this chapter the following things are discussed

- ***Cloud reference model***
- ***Architecture of following service models***
 - ***laas***
 - ***Paas***
 - ***Saas***
- ***Types of Clouds and their structure***
- ***Economics of the cloud***
- ***open challenges of cloud computing***

References

- Rajkumar Buyya, Christian Vecchiola, and Thamarai Selvi, **Mastering Cloud Computing**, McGraw Hill, ISBN-13: 978-1-25-902995-0, New Delhi, India, 2013.
 - Chapter 4- Cloud Computing Architecture
 - Section 4.1 to 4.5
 - This chapter slides text is compiled by:
 - Dr. Sounak Paul, BIT Mesra, Deoghar, India