**Introduction to Cloud Computing**

Data Center Approach:

Requirement: Your company wants to host their website.

Solution -

System Administrator’s responsibility.

Arranging all the entire things.

i) Choose the DataCenter / Hosting Provider.

ii) You need to typically send them an inquiry about your requirements.

iii) They will contact you and price negotiations.



When there are any issues, the system administrator has to run.



**Challenges with Data Center Model**

Example 1:-

Due to some big promotion, server capacity needs to be increased from 4GB RAM to 32 GB RAM

Data Center Provider Way:-

Buy a 32 GB RAM stick & install it onto your server

Hosting Provider Way:-

Raise a support ticket and expect a response within 15 minutes to 12 hours for a response.

Get the DC guys to resize your Server.

Cloud Way:-

Stop the Server & change the instance size.

**Introduction to Cloud Computing**

Cloud Computing is a model in which computing resource is available as a service.

3 important characteristic of Cloud Computing :

On-demand & self-serviced [ Any time launch without manual intervention ]

Elasticity. [ Can scale up and down anytime ]

Measured Service [ Pay what you use ]

## **Cloud Computing Models**

There are 3 types of Cloud Computing models

Software as Service [ Google Docs, Office 365 ]

Platform as Service [ Google App Engine ]

Infrastructure as Service [ AWS, Linode, Digital Ocean ]

It is very important to choose the right cloud service provider based on your use-case.

AWS is one of the most comprehensive Cloud providers.

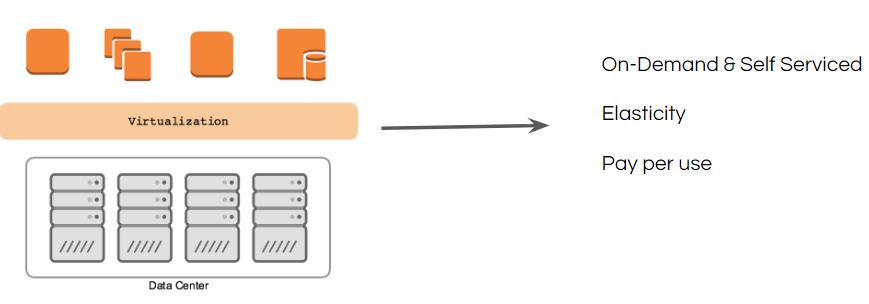
It provides all the type of cloud models

* Software as Service
* Platform as Service
* Infrastructure as Service

But if you just depend on AWS for everything, you will lose a lot of money. Hence many of the organizations opt for Multi-Cloud based approach.

**Architecture of Cloud Environments**

The cloud from behind the scenes is the data center only.



Virtualization Technology plays a very important role in Cloud Computing.

Virtualization allows us to run multiple OS on a single hardware.

There are many virtualization software available like:

* VMware Workstation / vSphere
* KVM
* XEN
* VirtualBox

**On-Demand & Self Service - Characteristics of Cloud**

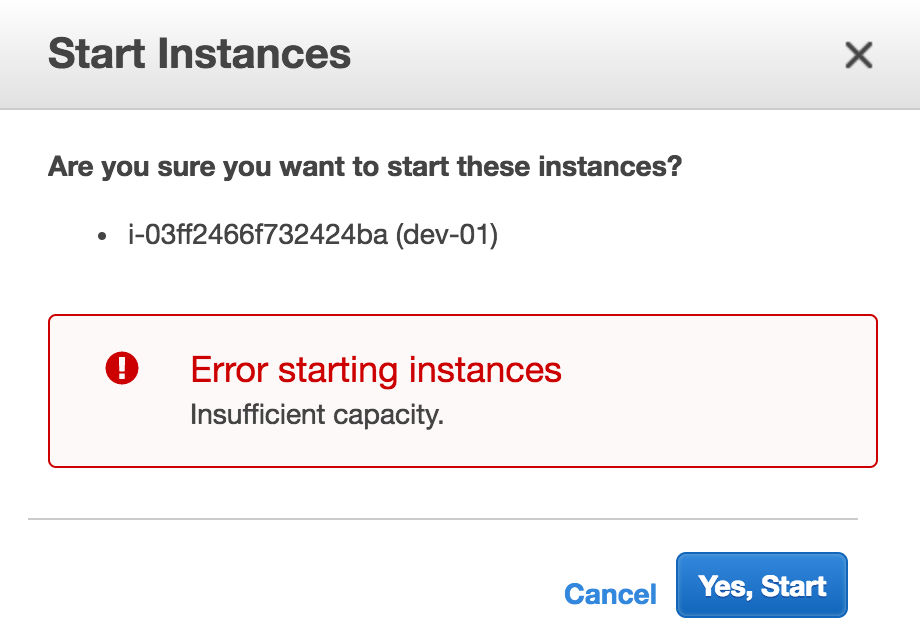
A person can provision resources in the cloud whenever needed, without requiring any human interaction with a service provider.

On-demand makes self-service with automation possible in a seeming less way.

Challenges with On-Demand Model:

On-Demand does not always mean that you will be able to launch instances at any given point of time.

Even a Cloud provider has limits, though it might be high, these limits are definitely reached.



## **Module 4: Elasticity**

Elasticity deals with adding and removing capacity, whenever it is needed in the environment.

Capacity generally refers to mostly processing & memory.

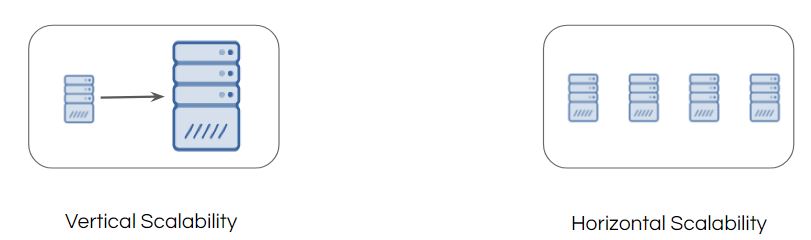
It is like a rubber band.



**4.1 Overview of Scalability**

Horizontal Scalability: Adding or Removing instances from the pool like cluster farm

Vertical Scalability: Adding or Removing resources for existing servers.



**4.2 Overview of Auto Scalability:**

Scaling servers on-demand is the real deal.

It can be achieved through Auto Scaling functionality.

Use Case Scenario:

* Whenever CPU Load > 70%, scale up to two more servers
* Whenever CPU Load < 30%, scale down by two servers.

Here is the sample auto-scaling based configuration:

