Module 1: Cloud Compute Basics

Mohanraj Shanmugam

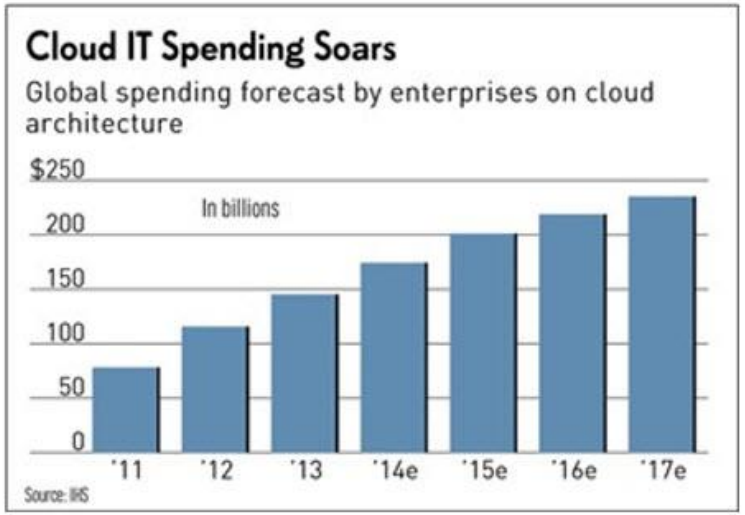
Cloud Computing Basics

Cloud Computing

- Everybody is talking about it , It's a buzzword..!

Why Learn about Cloud Computing?

- Seen an acceleration in adoption of cloud computing and cloud services
- It is mentioned as one of the top 5 technology trends to watch in next 4-5 ye by a report from Gartner
- By 2020 Cloud will be $241 Billion Market
- Most Enterprises are Started or already adopted to Cloud
- Cloud is way to run IT in the Enterprise fo now and future



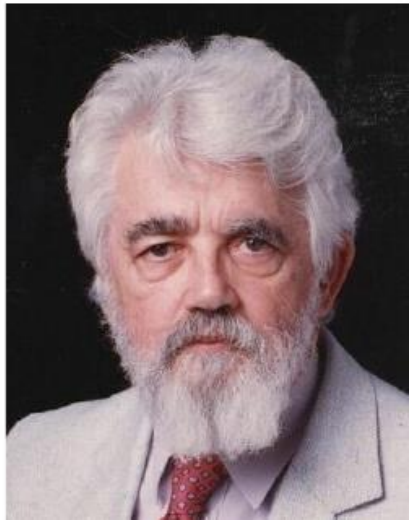
Cloud Computing History

Cloud Computing History

- The idea of computing in a "cloud" traces back to the origins of utility computing, a concept that computer scientist John McCarthy publicly proposed in 1961.

"Computation may someday be organized as a public utility … The computing utility could become the basis for a new and important industry."

John McCarthy (1961)



Cloud Computing History

- In 1969, Leonard Kleinrock, a chief scientist of the Advanced Research Projects Agency Network or ARPANET project that seeded the Internet, stated:

"As of now, computer networks are still in their infancy, but as they grow up and become sophisticated, we will probably see the spread of 'computer utilities' ...".

Cloud Computing History

- Early adapters of Utility Computing or form of cloud computing in 1990's

  - *Yahoo and Google Search Engine*
  - *Email Services(Hotmail, Gmail)*
  - *Open Publishing (Myspace, Facebook, Youtube)*
  - *Social Media like (Twitter , Linkedin)*

Though consumer-centric, these services popularized and validated core concepts that form the basis of modern-day cloud computing.

Cloud Computing History

- In the late 1990s, Salesforce.com pioneered the notion of bringing remotely provisioned services into the enterprise.

- In 2002, Amazon.com launched the Amazon Web Services (AWS) platform, a suite of enterprise-oriented services that provide remotely provisioned storage, computing resources, and business functionality.

Cloud Computing History

- 2006 that the term "cloud computing" emerged in the commercial arena.
- Amazon launched its Elastic Compute Cloud (EC2) services that enabled organizations to "lease" computing capacity and processing power to run their enterprise applications.
- Google Apps also began providing browser-based enterprise applications in the same year, and three years later, the Google App Engine became another historic milestone.

Cloud Computing Definition

Cloud Computing Definition

- *"...a style of computing in which scalable and elastic IT-enabled capabilities are delivered as a service to external customers using Internet technologies.“*

- *"...a standardized IT capability (services, software, or infrastructure) delivered via Internet technologies in a pay-per-use, self-service way.“*

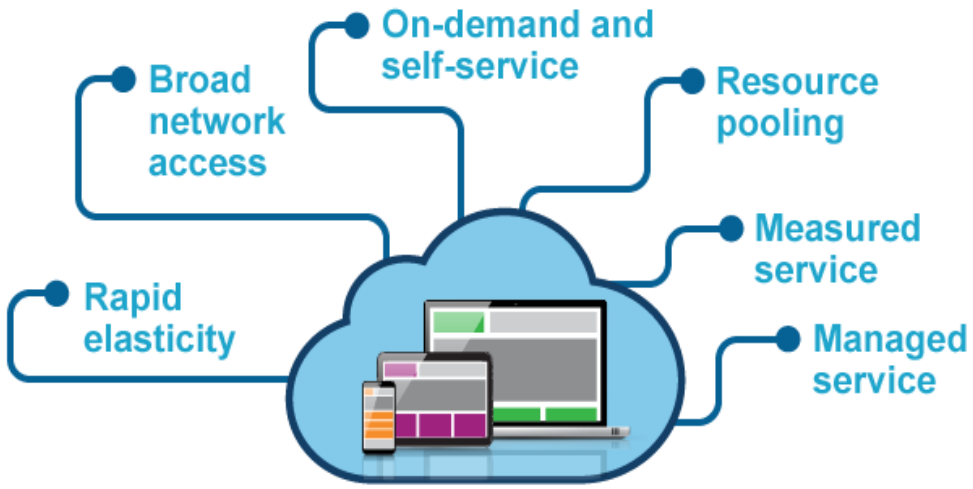
Cloud Computing Definition

- *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.*
- *This cloud model is composed of*
  - *five essential characteristics,*
  - *three service models, and*
  - *four deployment models."*

Cloud Computing Five Essential Characteristics

- On-demand self-service
- Broad network access
- Resource pooling
- Rapid elasticity
- Measured service

Cloud Computing Essential Characteristics



- On-demand self-service
  - A consumer can unilaterally provision computing capabilities, such as server, Storage and network ,
  - As needed automatically without requiring human interaction with each service provider.
- Broad network access
  - Capabilities are available over the network and accessed through standard mechanisms that promote use by heterogeneous thin or thick client platforms
  - (e.g. mobile phones, tablets, laptops, and workstations).
- Resource pooling
  - The provider's computing resources are pooled to serve multiple consumers using a multi-tenant model, with different physical and virtual resources dynamically assigned and reassigned according to consumer demand.
  - There is a sense of location independence in that the customer generally has no control or knowledge over the exact location of the provided resources but may be able to specify location at a higher level of abstraction (e.g., country, state, or datacenter).
  - Examples of resources include storage, processing, memory, and network bandwidth.

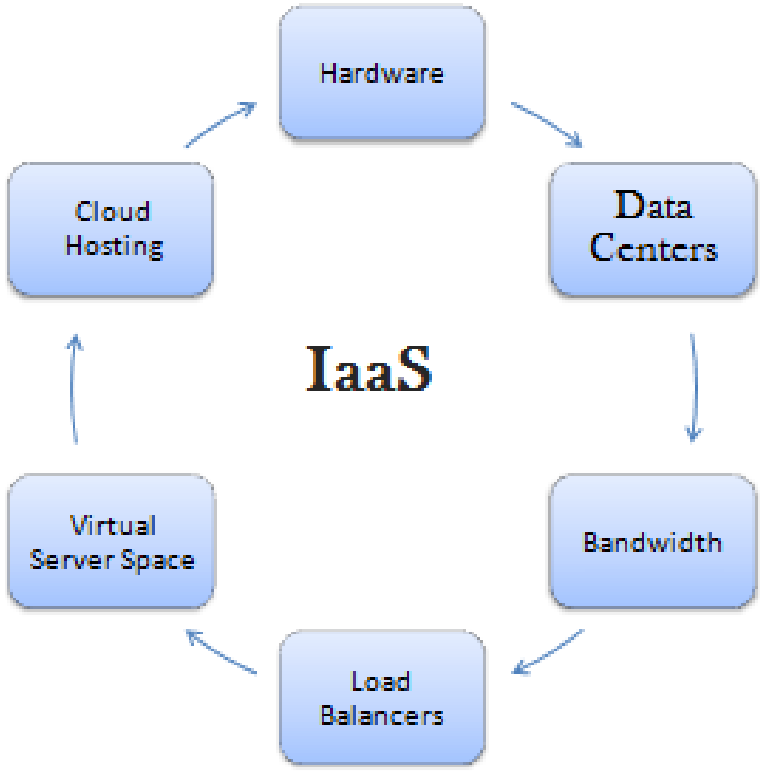
Cloud Computing Essential Characteristics

- Rapid elasticity
  - Capabilities can be elastically provisioned and released, in some cases automatically, to scale rapidly outward and inward commensurate with demand.
  - To the consumer, the capabilities available for provisioning often appear to be unlimited and can be appropriated in any quantity at any time.
- Measured service
  - Cloud systems automatically control and optimize resource use by leveraging a metering capability at some level of abstraction appropriate to the type of service (e.g. storage, processing, bandwidth, and active user accounts).
  - Resource usage can be monitored, controlled, and reported, providing transparency for both the provider and consumer of the utilized service.

Cloud Computing Three Service Models

- Service Model is What are IT services provided by the Cloud
- There are three Service Models:
  - Infrastructure as a Service
  - Platform as a Service
  - Software as a Service

Cloud Service Models



- Infrastructure as a Service (IaaS)
  - The capability provided to the consumer is to provision processing, storage, networks, and other fundamental computing resources
  - where the consumer is able to deploy and run arbitrary software, which can include operating systems and applications.
  - The consumer does not manage or control the underlying cloud infrastructure but has control over operating systems, storage, and deployed applications; and possibly limited control of select networking components (e.g., host firewalls).

Cloud Service Models

- Platform as a Service (PaaS)
  - The capability provided to the consumer is to deploy onto the cloud infrastructure consumer created or acquired applications created using programming languages, libraries, services, and tools supported by the provider.
  - The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, or storage, but has control over the deployed applications and possibly configuration settings for the application-hosting environment.

Cloud Service Models

- Software as a Service (SaaS)
  - The capability provided to the consumer is to use the provider's applications running on a cloud infrastructure.
  - The applications are accessible from various client devices through either a thin client interface, such as a web browser (e.g., web-based email), or a program interface.
  - The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, storage, or even individual application capabilities, with the possible exception of limited user- specific application configuration settings

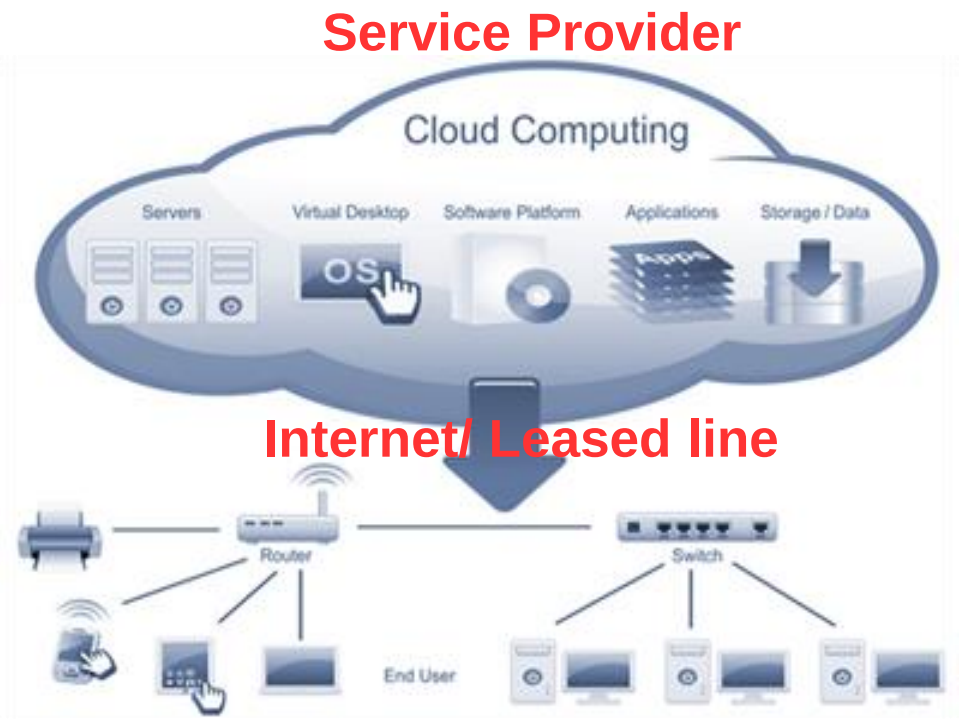
Cloud Computing Four Delivery Model

Cloud Delivery Model is how Cloud services is delivered to the End users

There are Delivery Model Models:

- Public Cloud
- Private Cloud
- Hybrid Cloud
- Community Cloud

Cloud Computing Delivery Models

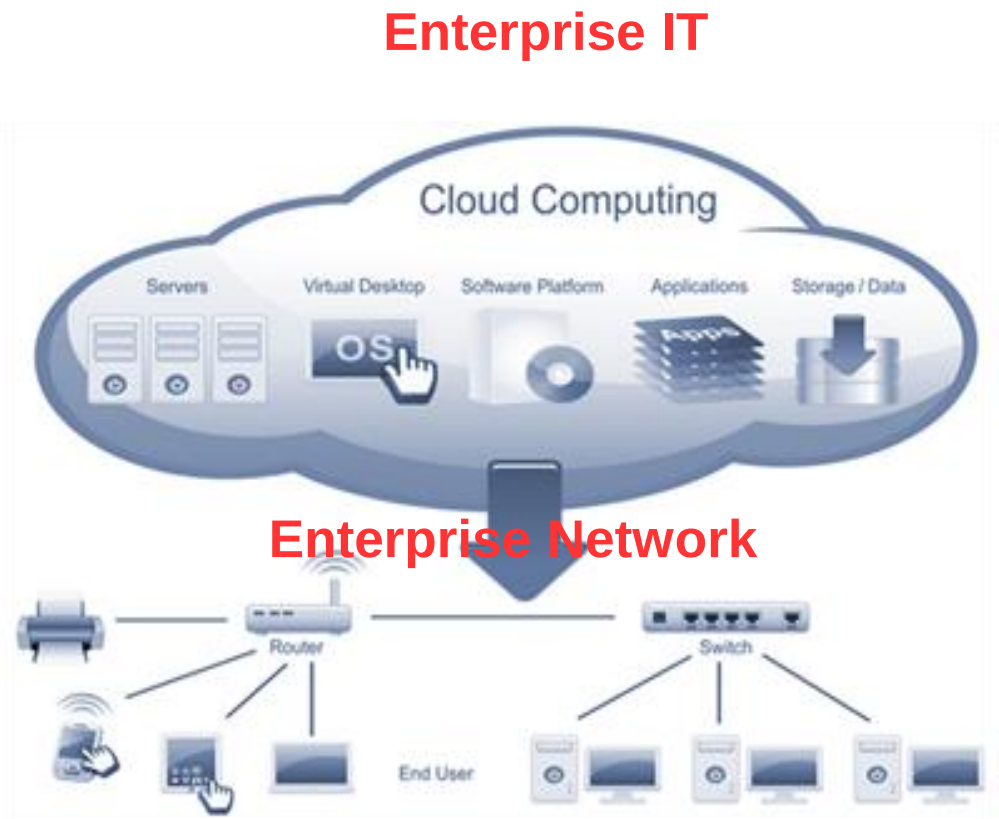


- Public Cloud
  - The cloud infrastructure is provisioned for open use by the general public. It may be owned, managed, and operated by a business, academic, or government organization, or some combination of them. It exists on the premises of the cloud provider.
  - Example:
    - Amazon AWS
    - Microsoft AZURE
    - Google Cloud

Cloud Computing Deployment Models

- Benefits of Public Cloud Computing:
  - Increase speed and agility
    - In a cloud computing environment, new IT resources are only ever a click away, which means you reduce the time it takes to make those resources available to your developers from weeks to just minutes. This results in a dramatic increase in agility for the organization, since the cost and time it takes to experiment and develop is significantly lower.
  - Stop spending money on running and maintaining data centers
    - Focus on projects that differentiate your business, not the infrastructure. Cloud computing lets you focus on your own customers, rather than on the heavy lifting of racking, stacking and powering servers.
  - Go global in minutes
    - Easily deploy your application in multiple regions around the world with just a few clicks. This means you can provide a lower latency and better experience for your customers simply and at minimal cost.

Cloud Computing Delivery Models

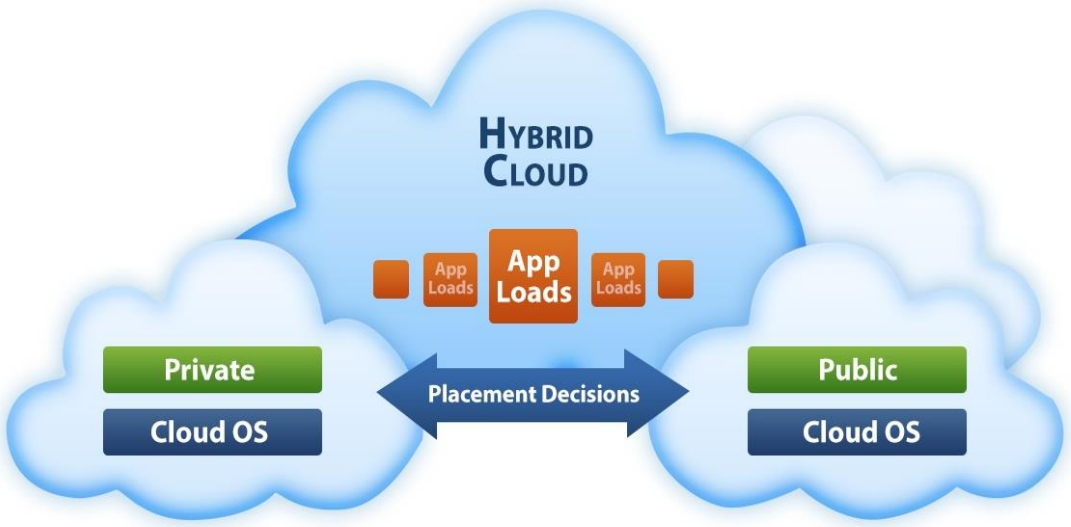


- Private Cloud
  - The cloud infrastructure is provisioned for exclusive use by a single organization comprising multiple consumers (e.g., business units). It may be owned, managed, and operated by the organization, a third party, or some combination of them, and it may exist on or off premises.
- Example
  - Openstack Private Cloud
  - Microsoft Private cloud
  - VMWare Private cloud

Cloud Computing Deployment Models

- Benefits of Private Cloud Computing:
  - Security
    - since the organization's data is tightly secured and controlled on servers that no other company has access too.
    - The servers can either sit within the network boundaries of the organization using them, with those servers managed and maintained by internal IT staff, or they can be located within the data center of a cloud service provider, with an organization accessing the data stored on the servers via private and secure network links.
    - If the servers that form part of private cloud infrastructure are located in company premises though, an organization does not have to worry about the physical security of servers situated in the external data centers of third party cloud providers
  - Flexibility
    - Provide any kind of Platform or software as a service which enterprise uses

Cloud Computing Delivery Models



- Hybrid Cloud
  - The cloud infrastructure is a compos ition of two or more distinct cloud infrastructures (private, community, or public) that remain unique entities, but are bound together by standardized or proprietary technology that enables data and application portability (e.g., cloud bursting for load balancing between clouds).

Example

- Openstack based hybrid cloud with AWS
- Microsoft Hybrid cloud with AZURE and Microsoft Private Cloud
- EMC Hybrid Cloud

Use Cases of Hybrid Cloud

- For example, an enterprise can deploy an on-premises private cloud to host sensitive or critical workloads, but use a third-party public cloud provider, such as Google Compute Engine, to host less-critical resources, such as test and development workloads.
- Hybrid cloud is particularly valuable for dynamic or highly changeable workloads. For example, a transactional order entry system that experiences significant demand spikes around the holiday season is a good hybrid cloud candidate.
- The application could run in private cloud, but use cloud bursting to access additional computing resources from a public cloud when computing demands spike. To connect private and public cloud resources, this model requires a hybrid cloud environment.

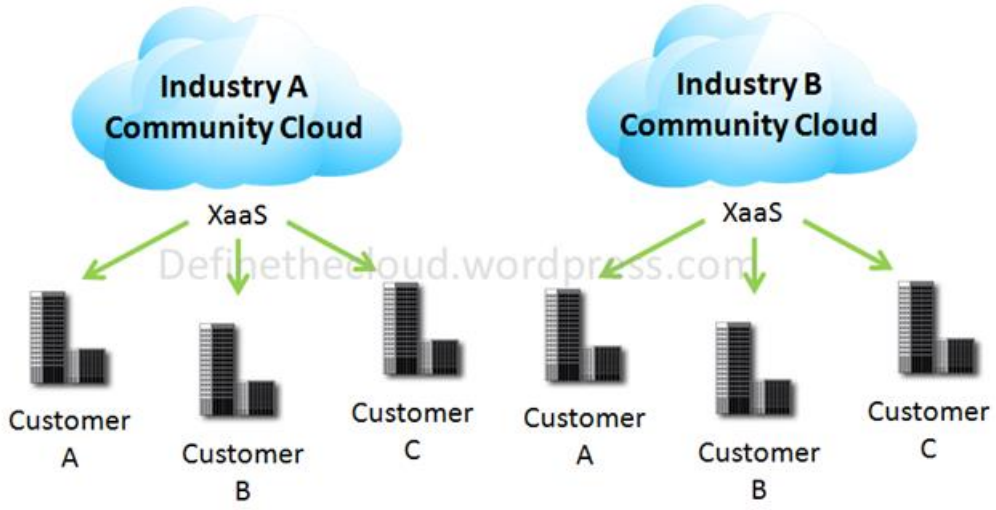
Use Cases of Hybrid Cloud

- Another hybrid cloud use case is big data processing. A company, for example, could use hybrid cloud storage to retain its accumulated business, sales, test and other data, and then run analytical queries in the public cloud, which can scale to support demanding distributed computing tasks

Hybrid Cloud Pros and Cons

- Public cloud's flexibility and scalability eliminates the need for a company to make massive capital expenditures to accommodate short-term spikes in demand. The public cloud provider supplies compute resources, and the company only pays for the resources it consumes.
- Despite its benefits, hybrid cloud can present technical, business and management challenges. Private cloud workloads must access and interact with public cloud providers, so hybrid cloud requires API compatibility and solid network connectivity.

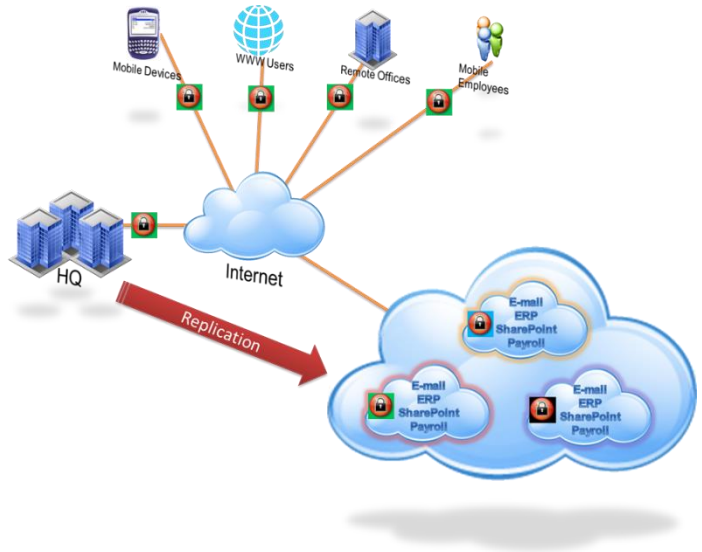
Cloud Computing Delivery Models



- Community Cloud
  - The cloud infrastructure is provisioned for exclu sive use by a specific community of consumers from organizations that have shared concerns (e.g., mission,
    security requirements, policy, and compliance c onsiderations). It may be owned, managed, and operated by one or more of the organizations in the community, a third party, or some combination of them, and it may exist on or off premises.
- Government organizations within the state of California may share computing infrastructure on the cloud to manage data related to citizens residing in California.

Cloud Computing Advantage

Cloud Computing Advantage



- Flexibility
  - Cloud-based services are ideal for businesses with growing or fluctuating bandwidth demands.
  - If your needs increase it's easy to scale up your cloud capacity, drawing on the service's remote servers.
  - Likewise, if you need to scale down again, the flexibility is baked into the service.
  - This level of agility can give businesses using cloud computing a real advantage over competitors

- Disaster recovery
  - Businesses of all sizes should be investing in robust disaster recovery, but for smaller businesses that lack the required cash and expertise, this is often more an ideal than the reality.
  - Cloud is now helping more organizations buck that trend.
  - According to Aberdeen Group, small businesses are twice as likely as larger companies to have implemented cloud-based backup and recovery solutions that save time, avoid large up-front investment and roll up third-party expertise as part of the deal.

Cloud Computing Advantage

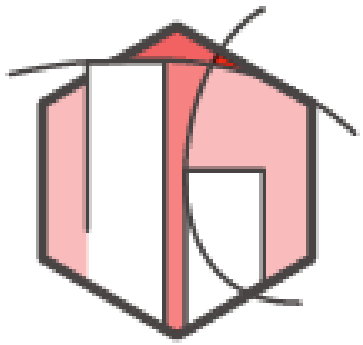
Automatic software updates

- The beauty of cloud computing is that the servers are off-premise, out of sight and out of your hair.
- Suppliers take care of them for you and roll out regular software updates – including security updates – so you don't have to worry about wasting time maintaining the system yourself.
- Leaving you free to focus on the things that matter, like growing your business.

Capital-expenditure Free

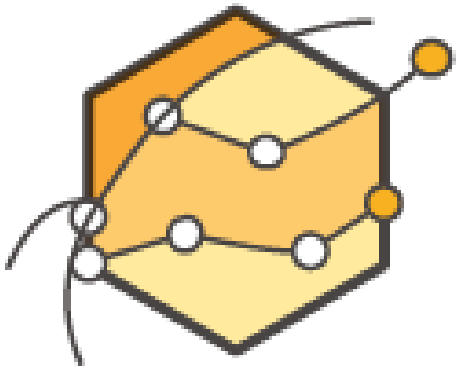
- Cloud computing cuts out the high cost of hardware.
- You simply pay as you go and enjoy a subscription-based model that's kind to your cash flow.
- Instead of having to invest heavily in data centers and servers before you know how you're going to use them, you can only pay when you consume computing resources, and only pay for how much you consume.
- Add to that the ease of setup and management and suddenly your scary, hairy IT project looks at lot friendlier.

Cloud Computing Advantage



Massive economies of scale

- By using cloud computing, you can achieve a lower variable cost than you can get on your own.
- Because usage from hundreds of thousands of customers are aggregated in the cloud, providers such as Amazon Web Services can achieve higher economies of scale which translates into lower pay as you go prices.



capacity

- Eliminate guessing on your infrastructure capacity needs.
- When you make a capacity decision prior to deploying an application, you often either end up sitting on expensive idle resources or dealing with limited capacity.
- With Cloud Computing, these problems go away. You can access as much or as little as you need, and scale up and down as required with only a few minutes notice.

Cloud Computing Advantage

Increase speed and agility

- In a cloud computing environment, new IT resources are only ever a click away, which means you reduce the time it takes to make those resources available to your developers from weeks to just minutes.
- This results in a dramatic increase in agility for the organization, since the cost and time it takes to experiment and develop is significantly lower.

Stop spending money on running and maintaining data centers

- Focus on projects that differentiate your business, not the infrastructure. Cloud computing lets you focus on your own customers, rather than on the heavy lifting of racking, stacking and powering servers.

Go global in minutes

- Easily deploy your application in multiple regions around the world with just a few clicks. This means you can provide a lower latency and better experience for your customers simply and at minimal cost.

Cloud Computing Advantage

Increased collaboration

- When your teams can access, edit and share documents anytime, from anywhere, they're able to do more together, and do it better.
- Cloud-based workflow and file sharing apps help them make updates in real time and gives them full visibility of their collaborations.

Work from anywhere

- With cloud computing, if you've got an internet connection you can be at work.
- And with most serious cloud services offering mobile apps, you're not restricted by which device you've got to hand.
- The result? Businesses can offer more flexible working perks to employees so they can enjoy the work-life balance that suits them – without productivity taking a hit.
- One study reported that 42% of workers would swap a portion of their pay for the ability to telecommute. On average they'd be willing to take a 6% pay cut.

Cloud Computing Advantage

Document control

- The more employees and partners collaborate on documents, the greater the need for watertight document control.
- Before the cloud, workers had to send files back and forth as email attachments to be worked on by one user at a time.
- Sooner or later – usually sooner – you end up with a mess of conflicting file content, formats and titles.
- And as even the smallest companies become more global, the scope for complication rises.
- According to one study, "73% of knowledge workers collaborate with people in different time zones and regions at least monthly".
- When you make the move to cloud computing, all files are stored centrally and everyone sees one version of the truth. Greater visibility means improved collaboration, which ultimately means better work and a healthier bottom line. If you’re still relying on the old way, it could be time to try something a little more streamlined.

Cloud Computing Advantage

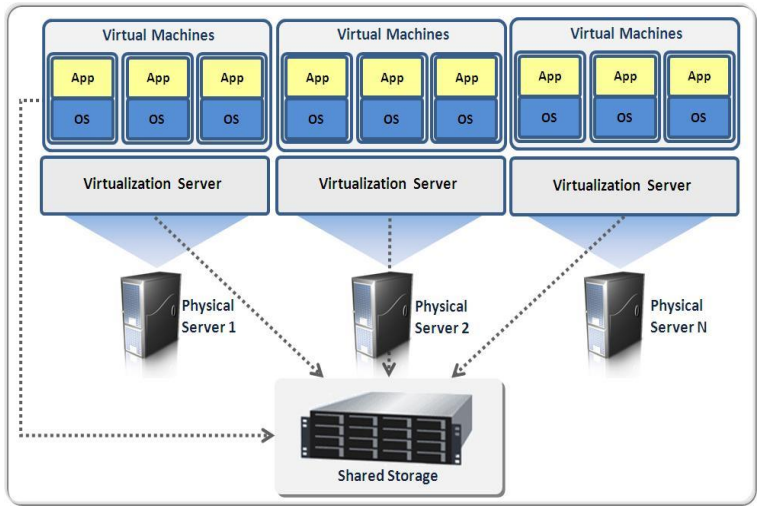
Security

- Lost laptops are a billion dollar business problem.
- And potentially greater than the loss of an expensive piece of kit is the loss of the sensitive data inside it.
- Cloud computing gives you greater security when this happens. Because your data is stored in the cloud, you can access it no matter what happens to your machine.
- And you can even remotely wipe data from lost laptops so it doesn't get into the wrong hands.

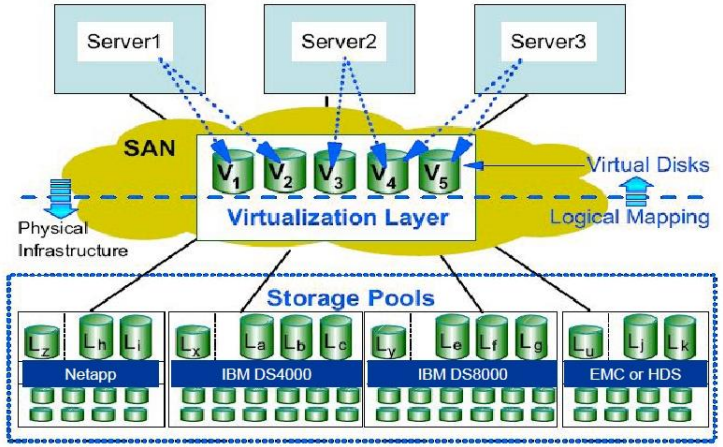
Competitiveness

- Wish there was a simple step you could take to become more competitive? Moving to the cloud gives access to enterprise-class technology, for everyone.
- It also allows smaller businesses to act faster than big, established competitors. Pay-as-you-go service and cloud business applications mean small outfits can run with the big boys, and disrupt the market, while remaining lean and nimble.

Virtualization

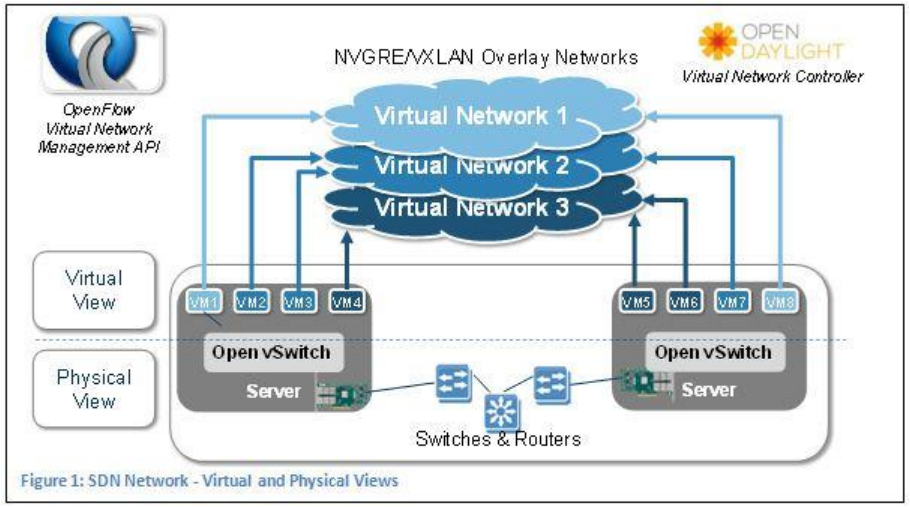


Server Virtualization



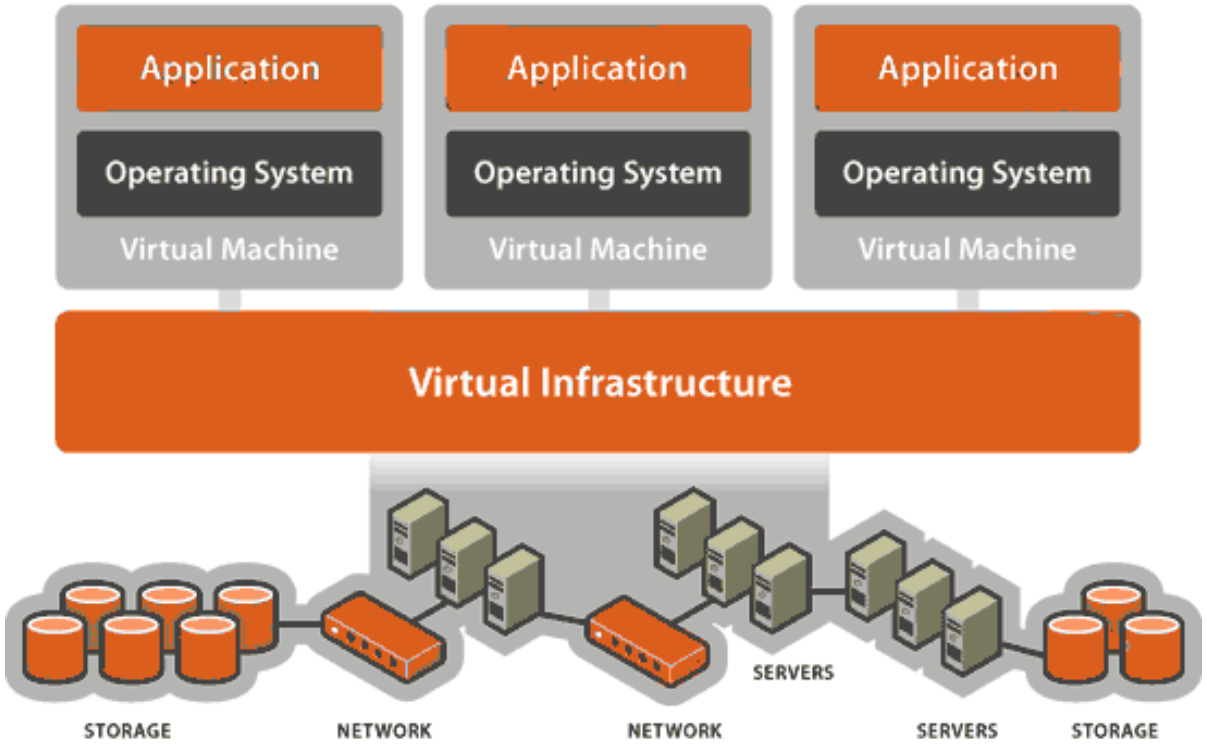
Storage Virtualization

- Virtualization is the creation of a virtual (rather than actual) version of something, such as an operating system, a server, a storage device or network resources.



Network Virtualization

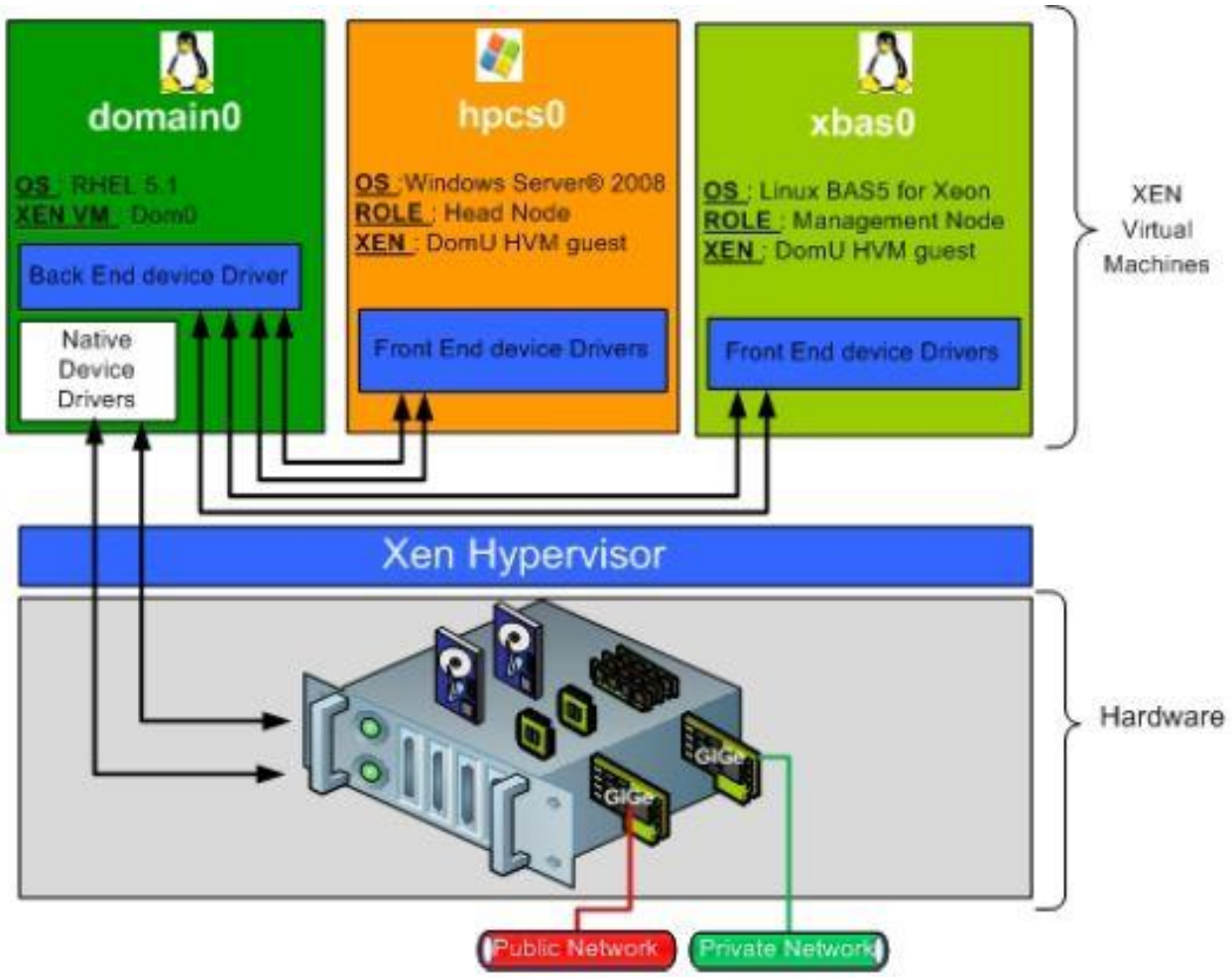
Virtualization



- Virtualization is the creation of a virtual (rather than actual) version of something, such as an operating system, a server, a storage device or network resources.
- Infrastructure Virtualization consist of:
  - Server Virtualization
  - Operating System Virtualization
  - Storage Virtualization
  - Network Virtualization

Server Virtualization

Server Virtualization

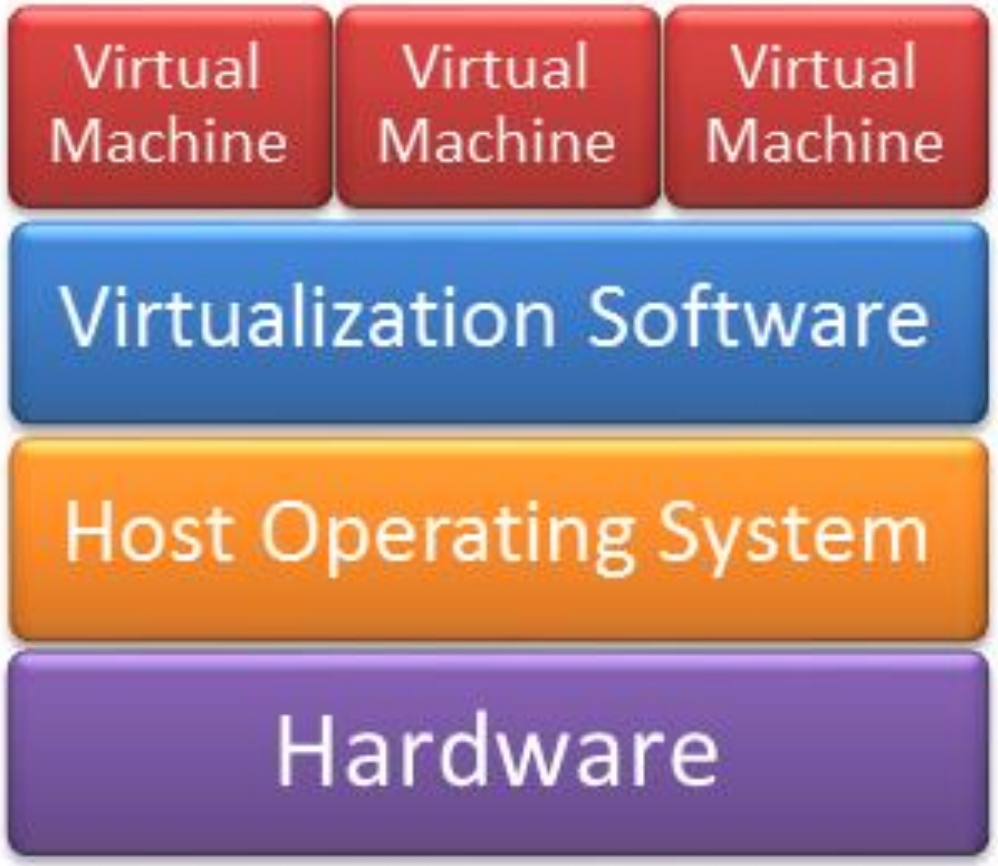


- Server virtualization is creating a Virtual Environment of Server resources like processors, Memory, Motherboard, Disks, and IO devices and operating systems masking from users.
- The virtual environments are called
  - Virtual Machines
  - Guests
  - Instances
  - Containers
  - Emulations
- The OS on which virtual environment running are called
  - Hosts
  - Hypervisor
  - Virtual Machine Monitor
  - Emulators
- Virtualization administrator uses Virtualization Management Console (VMC) to manage virtual machines.

Types of Server Virtualization

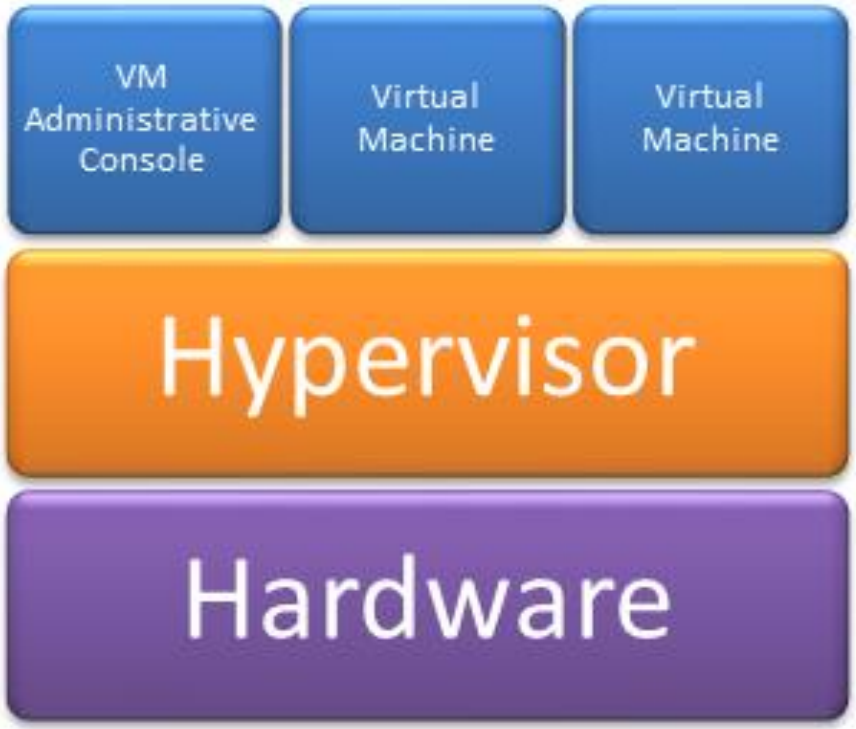
- There are different types of Server Virtualization:
  - *Guest OS/Virtual Machine Model*
  - *Hypervisor Model*
    - *Para Virtualization Model*
    - *Full Virtualization Model*
  - *Hardware Assisted Virtualization Model*
  - *Kernel Base Virtualization Model*

Guest OS/Virtual Machine Model



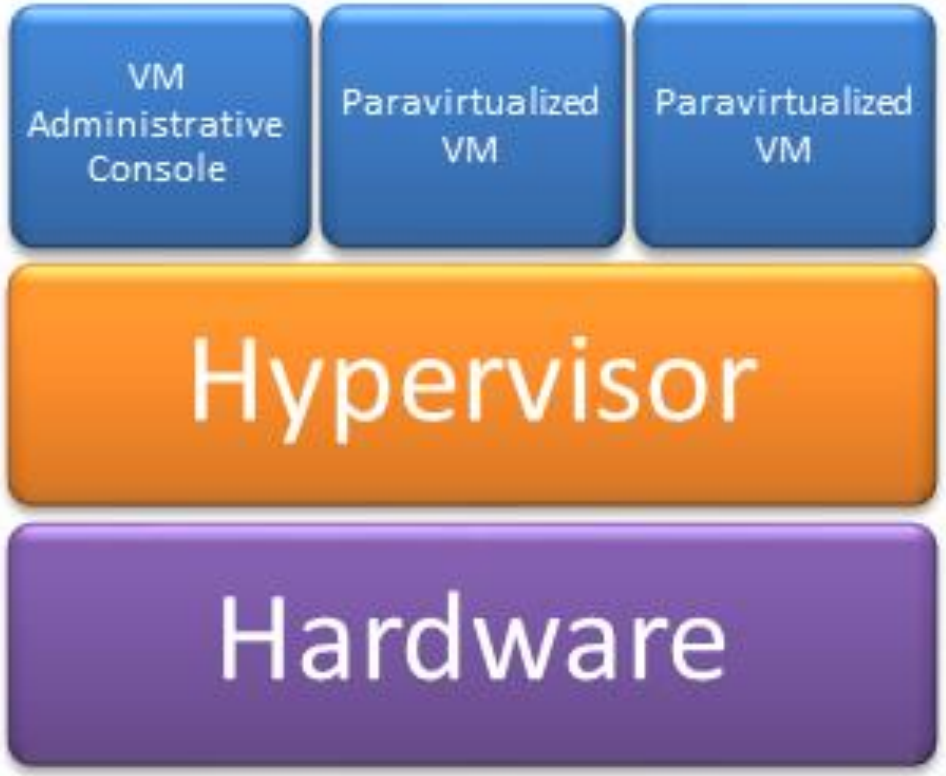
- **Guest OS/Virtual Machine Model**:
  - In this virtualization model each Virtual Machine (VM) runs as a separate instance of operating system within the virtualization software.
  - Virtualization software itself runs in another operating system which is installed on the hardware.
  - The operating system on which the virtualization application runs is often referred to as the “Host Operating System (OS)” because it provides the execution environment for the virtualization application.
  - This virtualization technique does not require any modification in the “guest OS”.
- Examples
  - Parallels Workstation (Mac)
  - VMWare Workstation (VMWare)
  - VMWare GSX Server (VMWare)
  - Oracle Virtual Box (VMWare)
- Use Cases:
  - Labs
  - Proof of Concept
  - Development Environment

Hypervisor based Virtualization



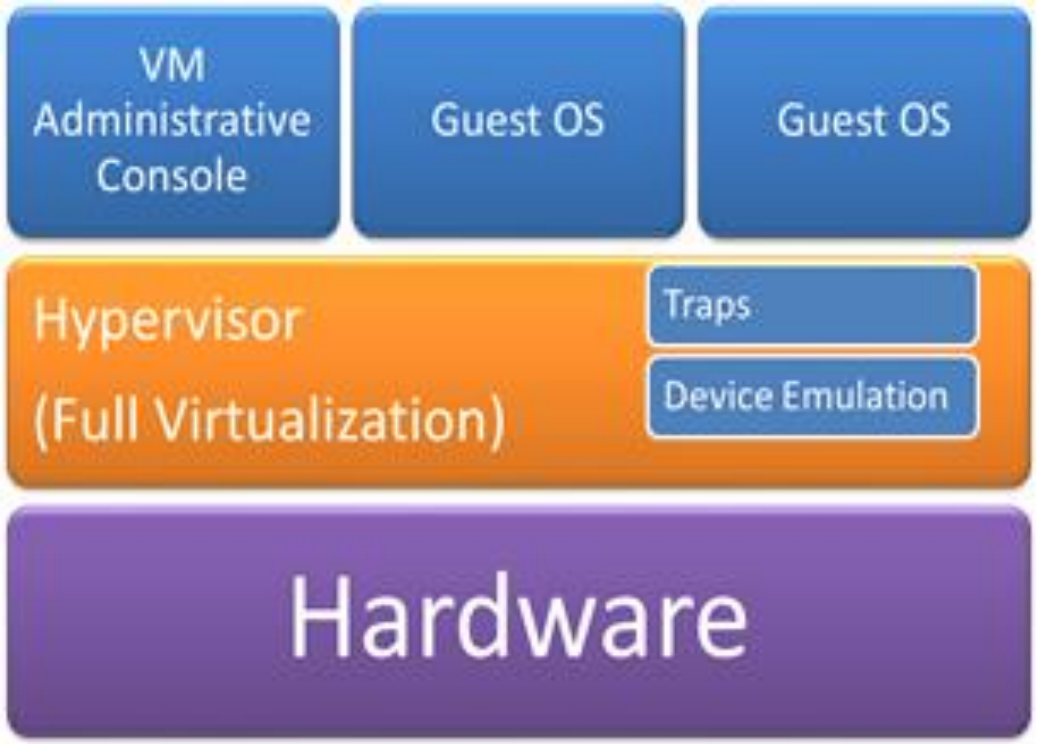
- In Hypervisor based virtualization a thin software layer called virtual machine monitor (VMM) or hypervisor runs on the top of machine's hardware and provides the necessary Virtual drivers and software for virtual devices in the virtual machines.
- It also handles queuing, dispatching and returning the results of hardware requests made by virtual machines.
- Virtual machine console also runs on the top of the hypervisor, whose main purpose is the administration and management of virtual machines.
- Two types of Hypervisor based Virtualization
  - Para Virtualization
  - Full Virtualization

Para Virtualization



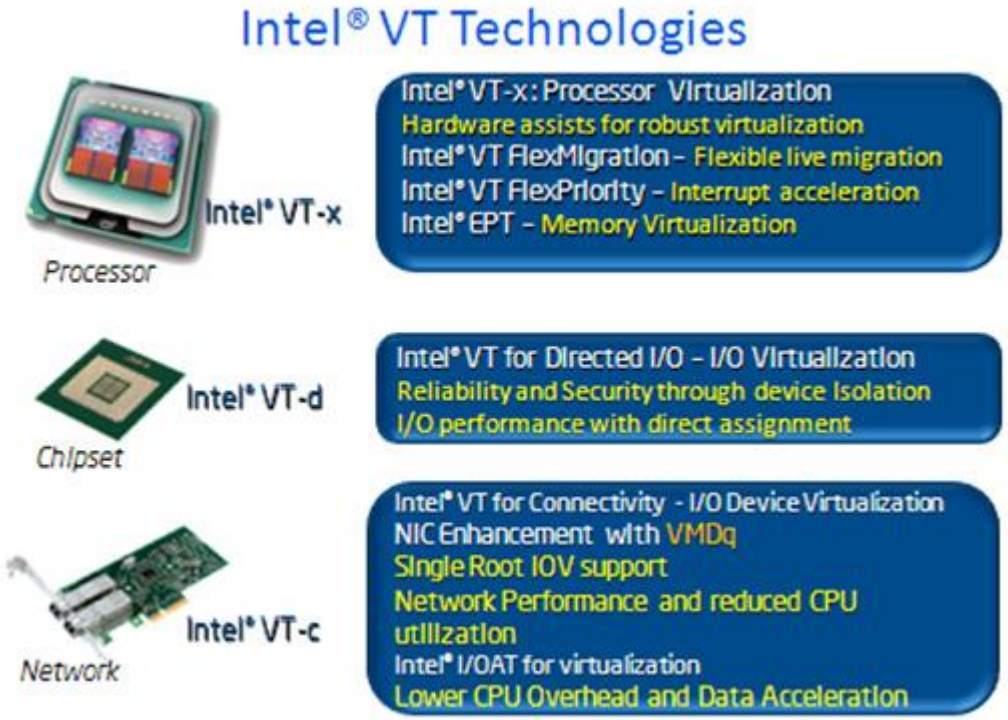
- Para virtualization is also based on hypervisor virtualization model.
- It requires that guest operating system should be recompiled or modified before installation inside the virtual machine to eliminate much of the trapping-and-emulation overhead associated with software implemented virtualization.
- Due to modification in guest operating system it gives performance enhancement over other techniques because the modified guest OS communicates directly with the hypervisor, and eliminates overheads occurred due *emulation process.*
- Examples
  - Xen
  - Hyper-V
- Use Cases:
  - Production Deployments
  - Availability of PV enabled Operation System
  - Direct Access to Hardware is required for Performance
  - Where Availability of Para Virtualization drivers for direct access hardware's

Full Virtualization



- Full virtualization is very similar to paravirtualization model. It contains functionality to emulate the underlying hardware when necessary.
- Full virtualization causes the hypervisor to “trap” the machine operations the OS uses to read or modify the system’s status or perform input/output (I/O) operations.
- After it has trapped them, the hypervisor emulates these operations in software and returns status codes consistent with what the real hardware would deliver.
- That’s why unmodified operating system can run on the top of hypervisor.
- Examples
  - VMWare ESX
- Use Cases:
  - Production Deployments
  - All OS Environment supported by Hypervisor

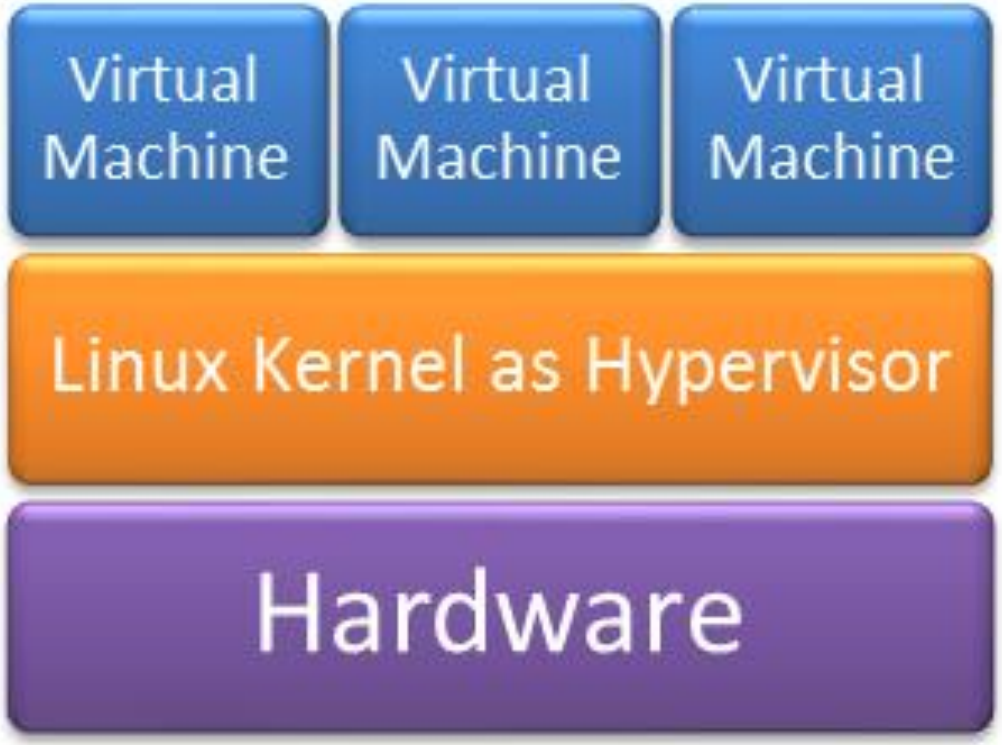
Hardware Assisted Virtualization



- Hardware Assisted Virtualization uses a hypervisor technique but it is only available on systems that support hardware virtualization.
- It relies on hardware extensions to the x86 system architecture to eliminate much of the hypervisor overhead associated with trapping and emulating I/O operations and status instructions executed within a guest OS.
- Hypervisors – based systems such as Xen and VMWare ESX Server, and kernel – level virtualization technologies such as KVM, can take advantage of the hardware support for virtualization that is provided on the latest generation of Intel (Intel VT, aka Vanderpool) and AMD (AMD – V, aka Pacifica) processors.

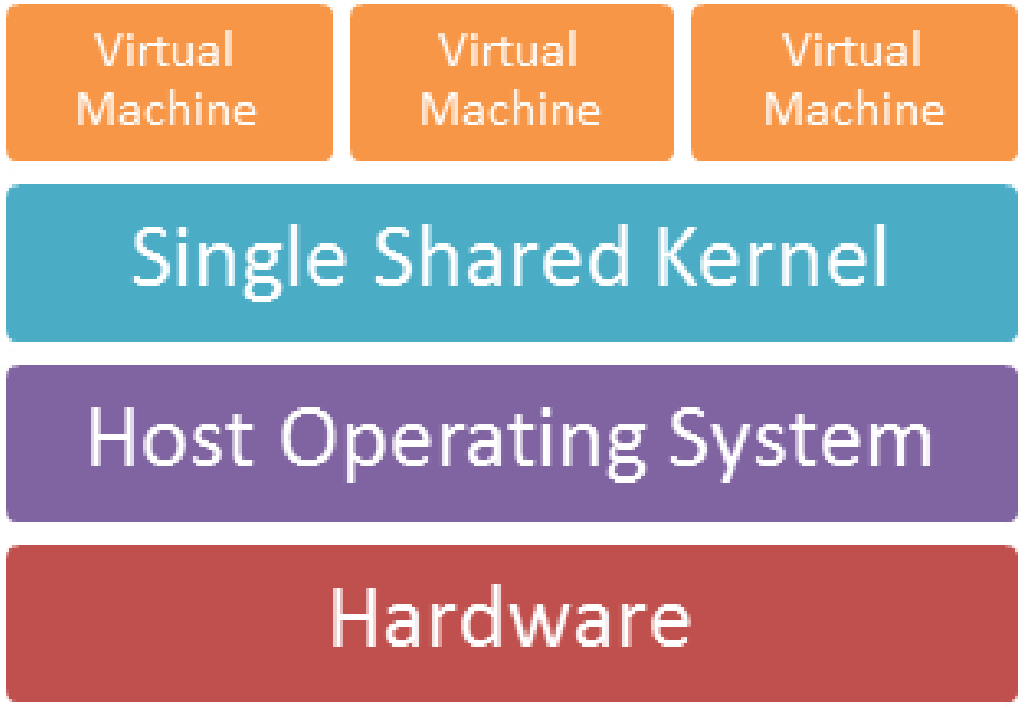
- Use Cases:
  - Production Deployments
  - Availability of Hard Assisted devices
  - All OS Environment supported by Hypervisor

Kernel Level Virtualization



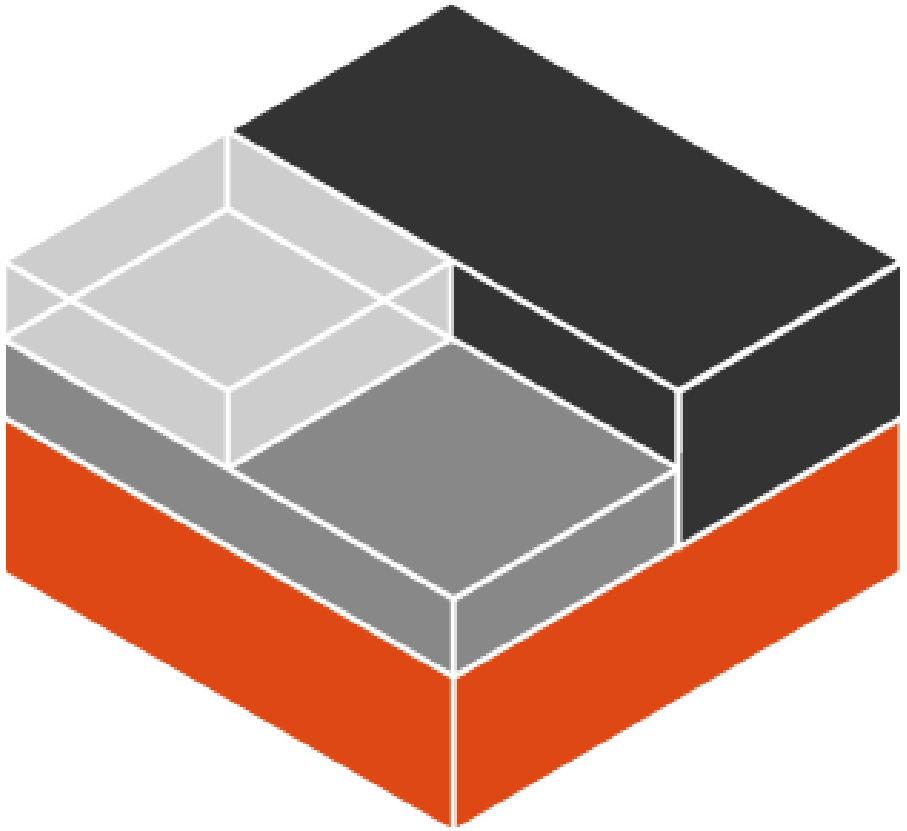
- This type of virtualization does not require a hypervisor, but instead runs a separate version of the Linux kernel and an associated virtual machine as a user – space process on the physical host.
- KVM uses a device driver in the host's kernel for communication between the main Linux kernel and the virtual machines.
- It requires processor support for virtualization (Intel VT or AMD – v) and uses a slightly modified QEMU process as the display and execution container for its virtual machines.
- Examples
  - Kernel Virtual Machine (KVM), which was introduced in the 2.6.20 mainline Linux kernel. It is a kernel based full virtualization
- Use Cases:
  - Production Deployments
  - Availability of Kernel based Virtualized Kernel
  - All OS Environment supported by Hypervisor

Operating System- Level Virtualization



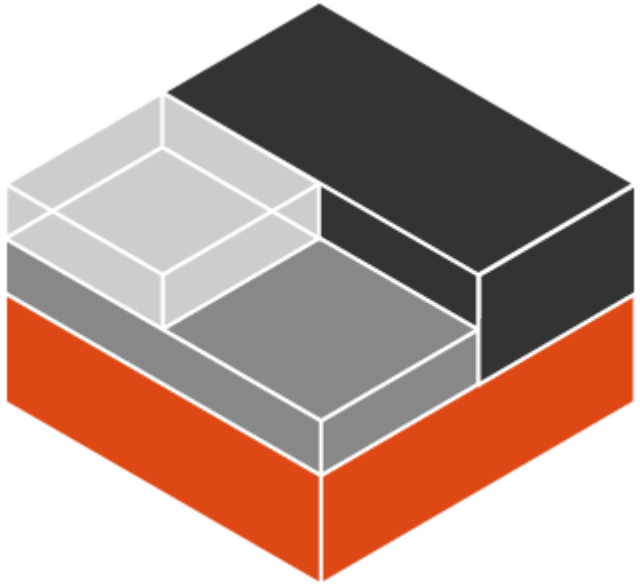
- Operating system – level virtualization describes various implementations of running multiple but logically distinct system environments on a single instance of an operating system kernel.
- . It is also called shared kernel approach because all the virtual instances shared a common kernel of host operating system.
- It is based on the change root **“chroot”** concept that is available on all modern UNIX like systems.
- The chroot mechanism as used by system – level virtualization is an extension of this concept, enabling the system to start virtual servers with their own sets of processes that execute relative to their own filesystem root directories.
- If all of your virtual servers must share a single copy of an operating system kernel, this is system – level virtualization. If different virtual servers can be running different operating systems, including different versions of a single operating system.
- Examples
  - Linux Containers
  - Oracle Solaris Containers and Zones
  - FreeBSD Jailed root.
- Use Cases:
  - DevOps Environment
  - Deployment is based on container vs packages

Linux Container



- **LXC** (Linux Containers) is an operating-system-level virtualization method for running multiple isolated Linux systems (containers) on a control host using a single Linux kernel - Wikipedia
- LXC is a userspace interface for the Linux kernel containment features.
  Through a powerful API and simple tools, it lets Linux users easily create
  and manage system or application containers – Linuxcontaier.org
- It is advanced chroot
- It is advanced BSD Jailed root
- It looks like a VM inside the Container
- But it is a process within the Base OS

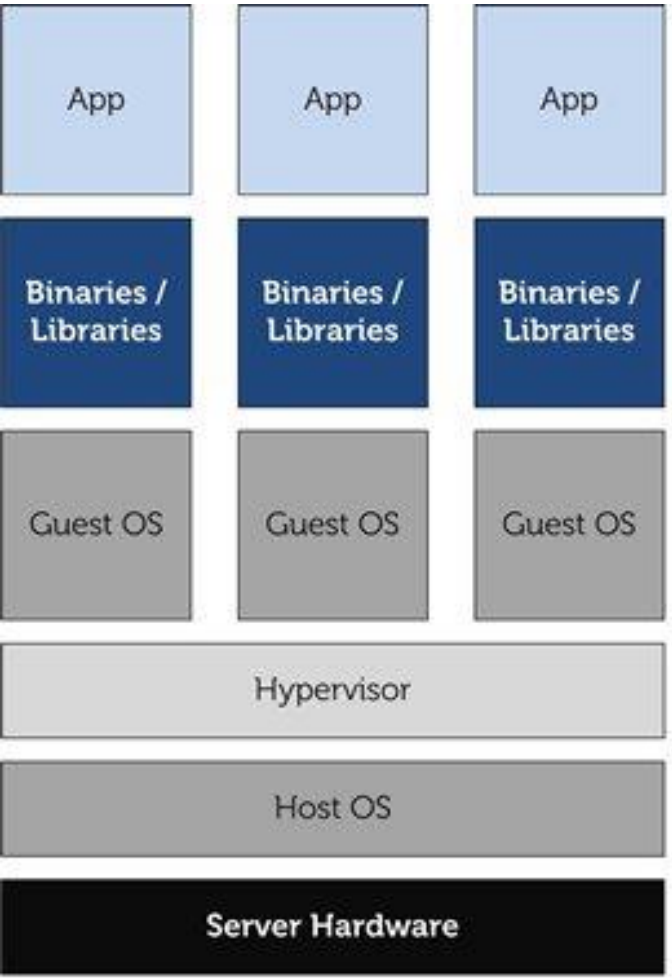
Linux Container



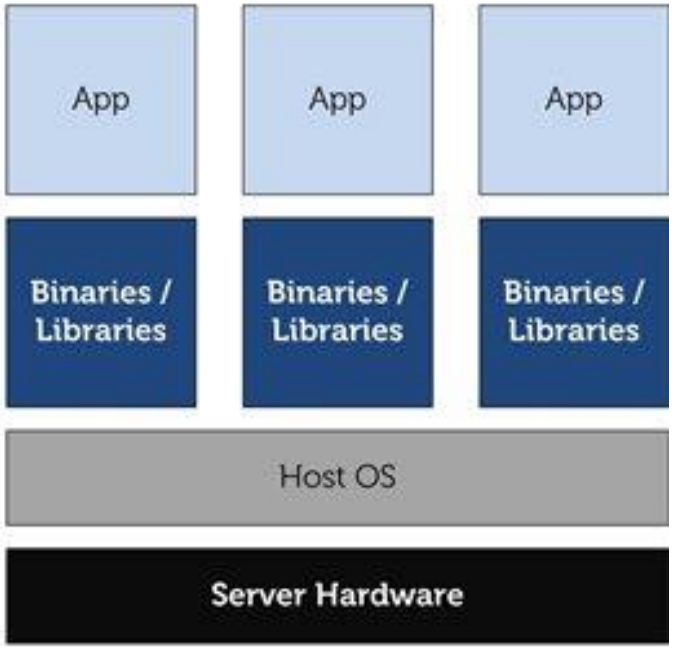
- LXC containers are often considered as something in the middle between a chroot and a full fledged virtual machine.
- The goal of LXC is to create an environment as close as possible to a standard Linux installation but without the need for a separate kernel.
- LXC is free software, most of the code is released under the terms of the GNU LGPLv2.1+ license

Container Vs Virtual Machine

- LXC containers is “lightweight virtualization”.
- Containers are similar to virtual machines (VMs) because they provide isolated computing environments all running independently on the same host.
- However, containers have very low overhead because you are not installing a separate operating system for the container and you do not need a hypervisor running on the host along with its overhead.



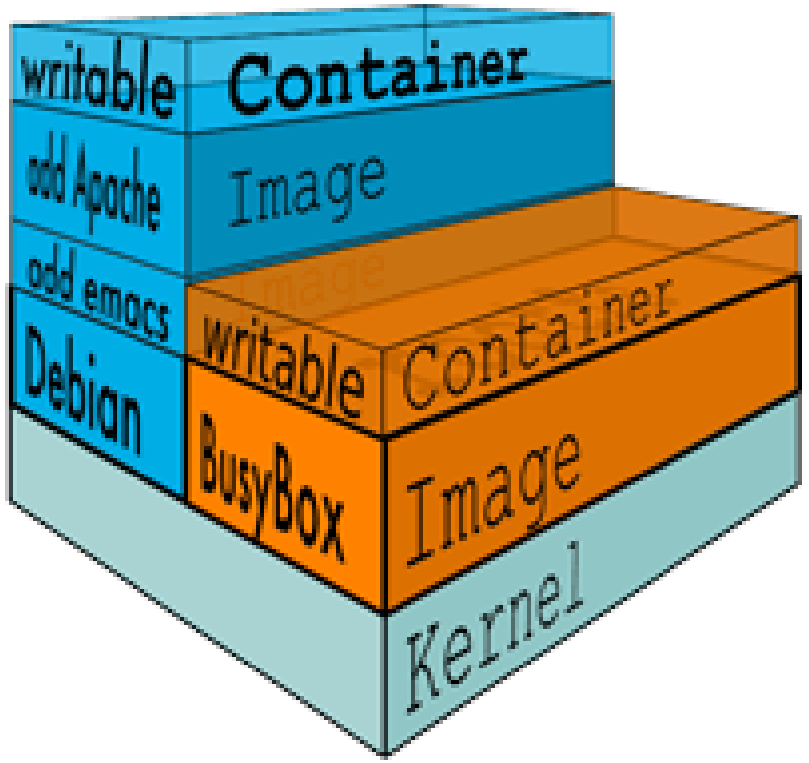
Virtualization



Containers

	Virtualization (i.e. kvm, xen)	LXC Containers
Footprint	Requires a hypervisor and a full operating system image.	Does not require a hypervisor or a separate operating system image.
OS supported	Any OS supported by the hypervisor	Most Linux distros, uses same kernel as host
Typical server deployment	10 – 100 VMs	100 - 1000 containers
Boot time	Less than a minute	Seconds
Physical resources use (i.e. memory, CPU)	Each VM has resource reserved for its own use	Shared by all containers

What is Docker?



- Docker allows you to package an application with all of its dependencies into a standardized unit for software development.
- Docker containers wrap up a piece of software in a complete filesystem that contains everything it needs to run:
  - code,
  - runtime,
  - system tools,
  - system libraries –

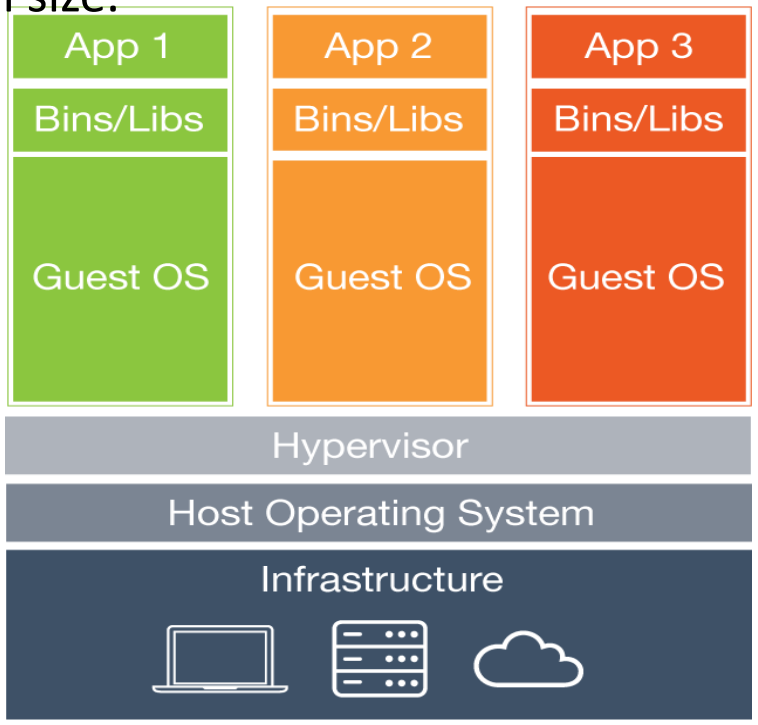
  anything you can install on a server.
- This guarantees that it will always run the same application, regardless of the environment it is running in.

Features of Docker

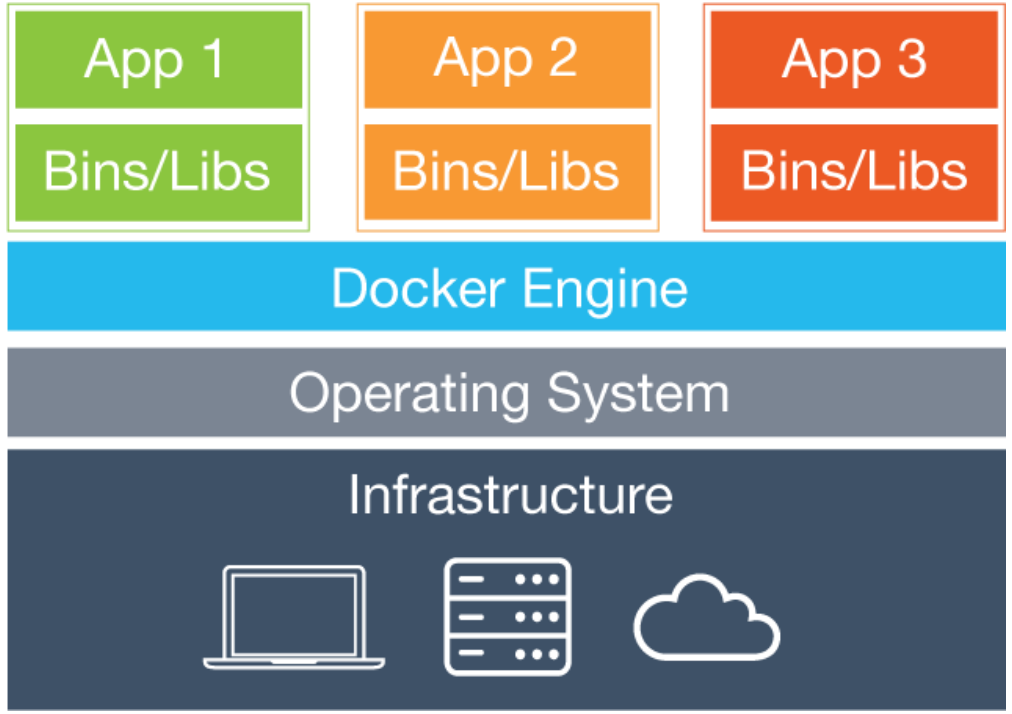
- Lightweight
  - Containers running on a single machine all share the same operating system kernel so they start instantly and make more efficient use of RAM.
  - Images are constructed from layered filesystems so they can share common files, making disk usage and image downloads much more efficient.
- Open
  - Docker containers are based on open standards allowing containers to run on all major Linux distributions and Microsoft operating systems with support for every infrastructure.
- Secure
  - Containers isolate applications from each other and the underlying infrastructure while providing an added layer of protection for the application.

VM Vs Docker

- Virtual Machines
  - Each virtual machine includes the application, the necessary binaries and libraries and an entire guest operating system - all of which may be tens of GBs in size.



- Containers
  - Containers include the application and all of its dependencies, but share the kernel with other containers.
  - They run as an isolated process in userspace on the host operating system.
  - They're also not tied to any specific infrastructure – Docker containers run on any computer, on any infrastructure and in any cloud.



Docker helps to build better Software

- Accelerate Developer Onboarding
  - Stop wasting hours trying to setup developer environments, spin up new instances and make copies of production code to run locally.
  - With Docker, you can easily take copies of your live environment and run on any new endpoint running Docker.
- Empower Developer Creativity
  - The isolation capabilities of Docker containers free developers from the worries of using “approved” language stacks and tooling.
  - Developers can use the best language and tools for their application service without worrying about causing conflict issues.

Docker helps to build better Software

- Eliminate Environment Inconsistencies
  - By packaging up the application with its configs and dependencies together and shipping as a container, the application will always work as designed locally, on another machine, in test or production.
  - No more worries about having to install the same configs into a different environment.

Docker helps to Share Software

- Distribute and share content
  - Store, distribute and manage your Docker images in your Docker Hub with your team.
  - Image updates, changes and history are automatically shared across your organization.
- Simply share your application with others
  - Ship one or many containers to others or downstream service teams without worrying about different environment dependencies creating issues with your application.
  - Other teams can easily link to or test against your app without having to learn or worry about how it works.

Docker helps to Ship Software

- Ship 7X More
  - Docker users on average ship software 7X more after deploying Docker in their environment.
  - More frequent updates provide more value to your customers faster.
- Quickly Scale
  - Docker containers spin up and down in seconds making it easy to scale an application service at any time to satisfy peak customer demand, then just as easily spin down those containers to only use the resources you need when you need it

Docker helps to Ship Software

- Easily Remediate Issues
  - Docker make it easy to identify issues and isolate the problem container, quickly roll back to make the necessary changes then push the updated container into production.
  - The isolation between containers make these changes less disruptive than traditional software models.

Docker is a shipping container system for code

Multiplicity of Stacks

Static website

User DB

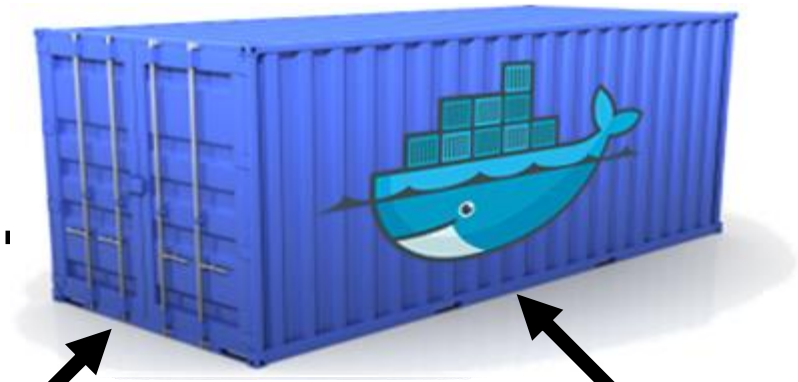
Web frontend

Queue

Analytics DB

Do services and apps interact appropriately?

An engine that enables any payload to be encapsulated as a lightweight, portable, self-sufficient container…



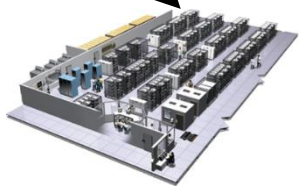
…that can be manipulated using standard operations and run consistently on virtually any hardware platform

Multiplicity of hardware environments

Development VM

QA server

Customer Data Center



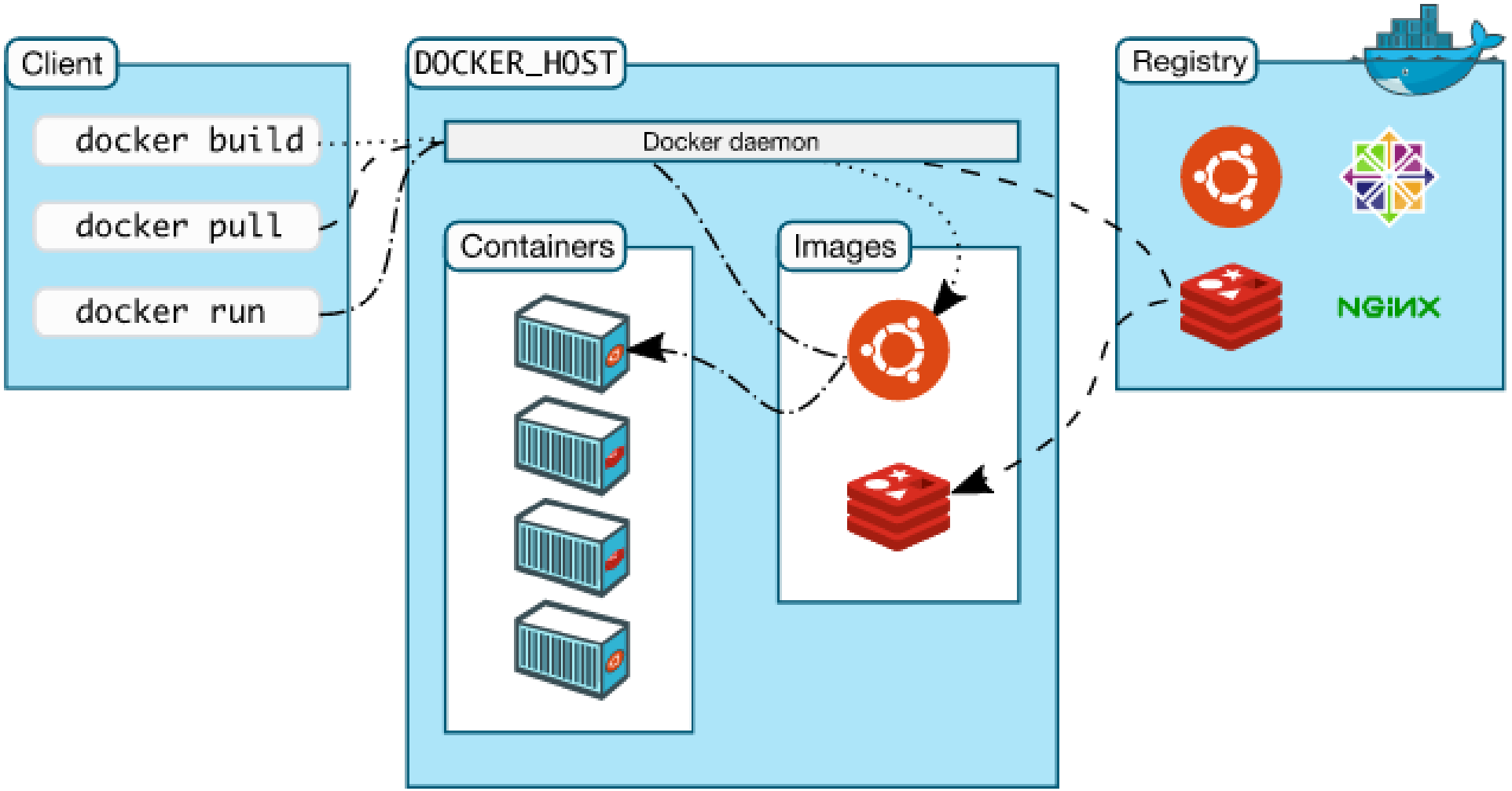
Public Cloud

Production Cluster

Contributor's laptop

Can I migrate smoothly and quickly

Docker Architecture



Cloud Compute Infrastructure

Cloud Compute Infrastructure

- Cloud Compute Infrastructure consist of Below:
  - Compute
  - Storage
  - Network
  - Datacentres and Regions
  - Shared Services
  - Platforms

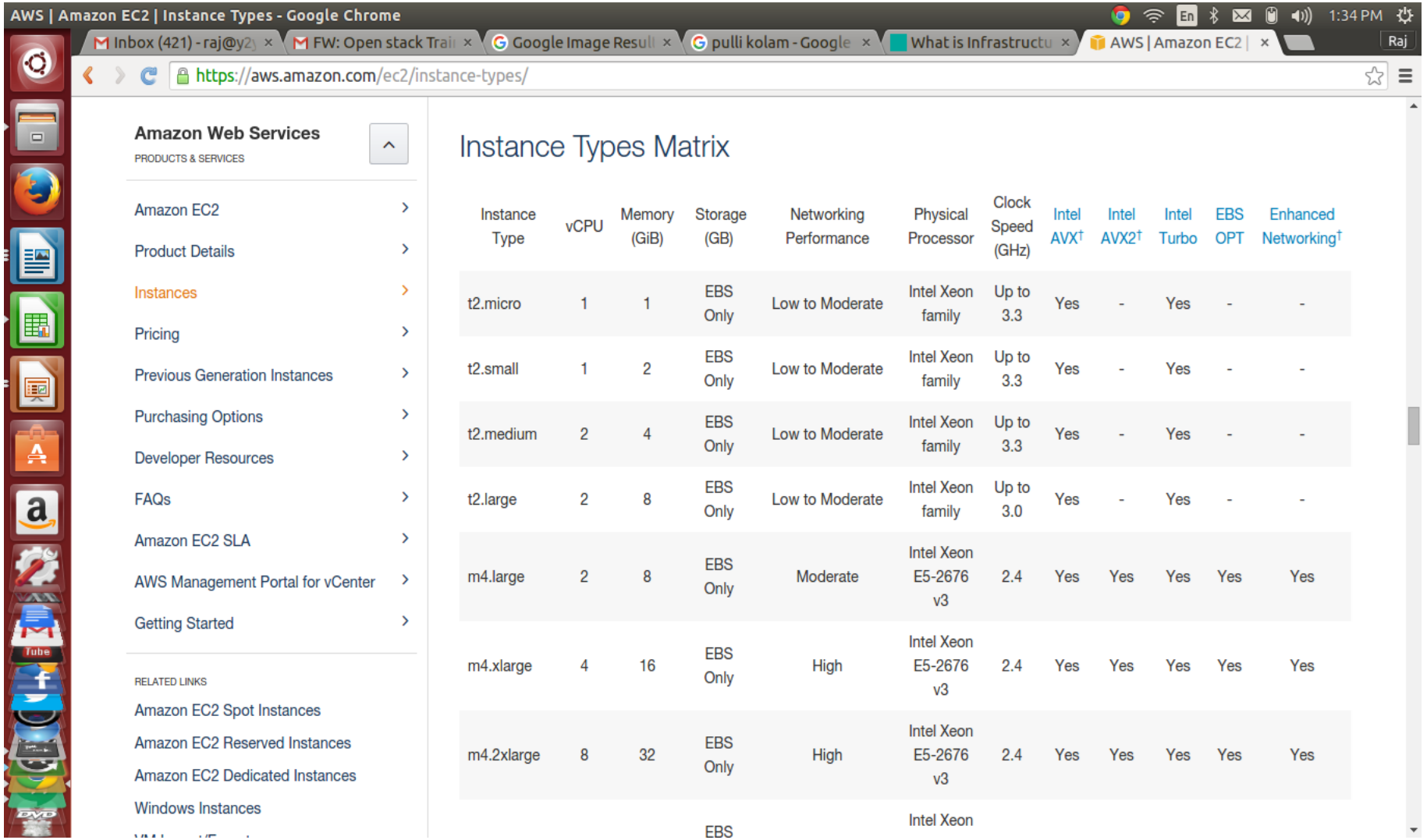
Cloud Compute Infrastructure

- Cloud compute infrastructure is a Virtual Machines on the cloud which provides a
  - Virtual Processor (vCPU)
  - Virtual Memory
  - Virtual disk
  - Virtual Network
  - Virtual Operating System
  - Management dashboard or a API to manage the VM
  - Accessable via Internet or Enterprise Network
  - Will connected to Network to other cloud instance within the network domain
  - On which you can deploy your platform or software for the business application

Cloud Compute Infrastructure

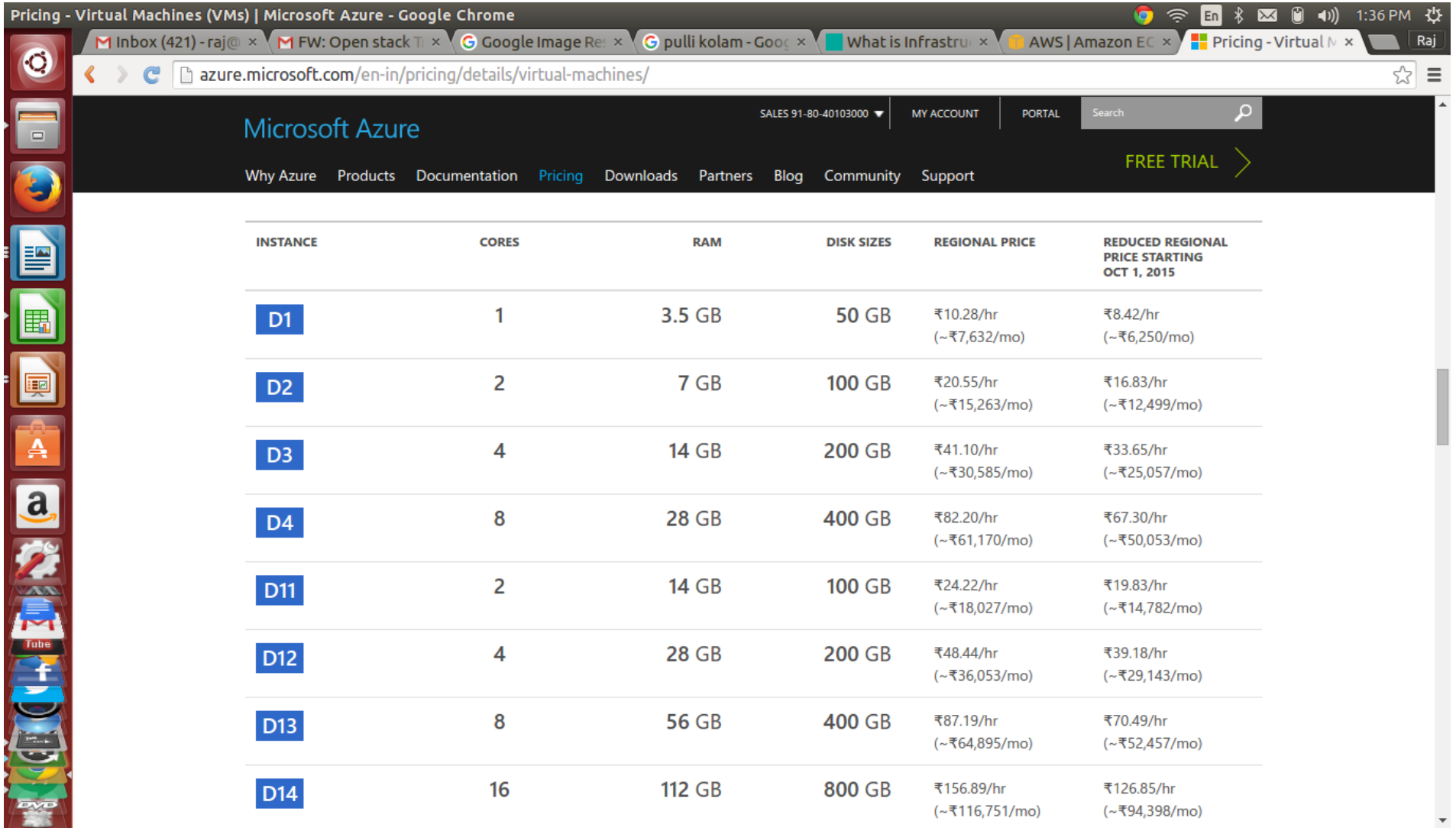
- Characteristics of Cloud Compute Infrastructure
  - Self service
  - Increase or Decrease or stop VM or its resources any time
  - Charged for usage
  - Fully automated provisioning
  - Accessible within the prescribed network boundry

Amazon Compute EC2 Examples



Instance Type	vCPU	Memory (GiB)	Storage (GB)	Networking Performance	Physical Processor	Clock Speed (GHz)	Intel AVX†	Intel AVX2†	Intel Turbo	EBS OPT	Enhanced Networking†
t2.micro	1	1	EBS Only	Low to Moderate	Intel Xeon family	Up to 3.3	Yes	-	Yes	-	-
t2.small	1	2	EBS Only	Low to Moderate	Intel Xeon family	Up to 3.3	Yes	-	Yes	-	-
t2.medium	2	4	EBS Only	Low to Moderate	Intel Xeon family	Up to 3.3	Yes	-	Yes	-	-
t2.large	2	8	EBS Only	Low to Moderate	Intel Xeon family	Up to 3.0	Yes	-	Yes	-	-
m4.large	2	8	EBS Only	Moderate	Intel Xeon E5-2676 v3	2.4	Yes	Yes	Yes	Yes	Yes
m4.xlarge	4	16	EBS Only	High	Intel Xeon E5-2676 v3	2.4	Yes	Yes	Yes	Yes	Yes
m4.2xlarge	8	32	EBS Only	High	Intel Xeon E5-2676 v3	2.4	Yes	Yes	Yes	Yes	Yes
			EBS		Intel Xeon						

Microsoft Azure VM Examples



INSTANCE	CORES	RAM	DISK SIZES	REGIONAL PRICE	REDUCED REGIONAL PRICE STARTING OCT 1, 2015
D1	1	3.5 GB	50 GB	₹10.28/hr (~₹7,632/mo)	₹8.42/hr (~₹6,250/mo)
D2	2	7 GB	100 GB	₹20.55/hr (~₹15,263/mo)	₹16.83/hr (~₹12,499/mo)
D3	4	14 GB	200 GB	₹41.10/hr (~₹30,585/mo)	₹33.65/hr (~₹25,057/mo)
D4	8	28 GB	400 GB	₹82.20/hr (~₹61,170/mo)	₹67.30/hr (~₹50,053/mo)
D11	2	14 GB	100 GB	₹24.22/hr (~₹18,027/mo)	₹19.83/hr (~₹14,782/mo)
D12	4	28 GB	200 GB	₹48.44/hr (~₹36,053/mo)	₹39.18/hr (~₹29,143/mo)
D13	8	56 GB	400 GB	₹87.19/hr (~₹64,895/mo)	₹70.49/hr (~₹52,457/mo)
D14	16	112 GB	800 GB	₹156.89/hr (~₹116,751/mo)	₹126.85/hr (~₹94,398/mo)

Virtual Processor (vCPU)

- A vCPU stands for Virtual Central Processing Unit.
- One or more vCPUs are assigned to every Virtual Machine (VM) within a cloud environment.
- Each vCPU is seen as a single physical CPU core by the VM’s operating system.

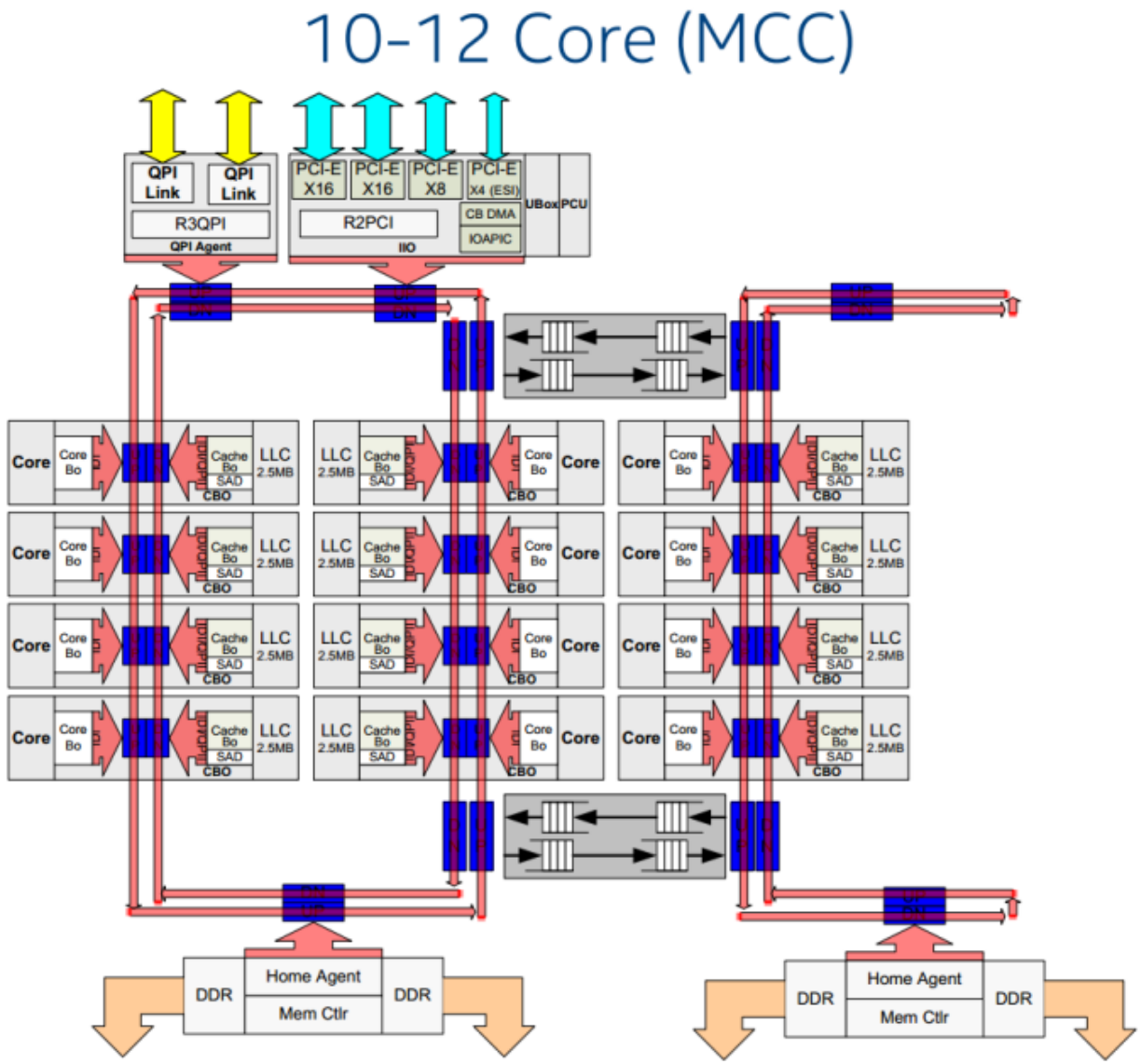
Virtual Processor (vCPU)

- How much processing of 1vCPU gets when i am in the cloud enviroinment...
- It is a complex answer
  - It depends lets understand how the vCPU is transulated in to Physical CPU

Intel xeon 4 core cpu specs

Performance	
# of Cores	4
# of Threads	8
Processor Base Frequency	3.2 GHz
Max Turbo Frequency	3.5 GHz
TDP	140 W

Physical CPU Layout



vCPU Implementation

- Time share allocation
- Single CPU to vCPU Pinning

Time slice allocation

- This is deafault allocation type all the VM will share the time slice of the CPU
- For example if we have 2.4ghz * 10 core we have 24 ghz cpu power
- If we have 10 vms running and each have 2 vcpu allocated then total 20 vcpus allocated
- This 20 vCPU will share the Timeslice of Physical CPU equally
- So each vCPU will 24ghz/20=1.2Ghz when all the VMs running at 100% CPU utilization
- When VM does not use the CPU power, It will be used by the other VMs

vCPU Pinning

- CPU pinning is the ability to run specific VM's virtual CPU (vCPU) on specific physical CPU (pCPU) in a specific host.
- In this pattern one pCPU is pinned to one vCPU in that case the pCPU is dedicated to VM which vCPU is assaigned even it is not utilizing it will not share its timeslice with other VMs

Hyperthreading

- Hyper-threading Technologies used to improve parallelization of computations performed on x86 microprocessors.
- A Physical core is logically divided in to multiple threads and each thread will view as one Physical core to the operating system and it starts scheduling the to the thread
- Most Xeon processor have 2 threads per core that means each physical core will have 2 threads

Hyperthreading

- If One thread is waiting for resources other thread will be processed due to which the effecience of the Per core performence goes up.
- An unoffical thumb of rule we follow to size the thread is One thread can give maximum of 60% of the core performence the means using thread we can get up to 120% of core's performece.

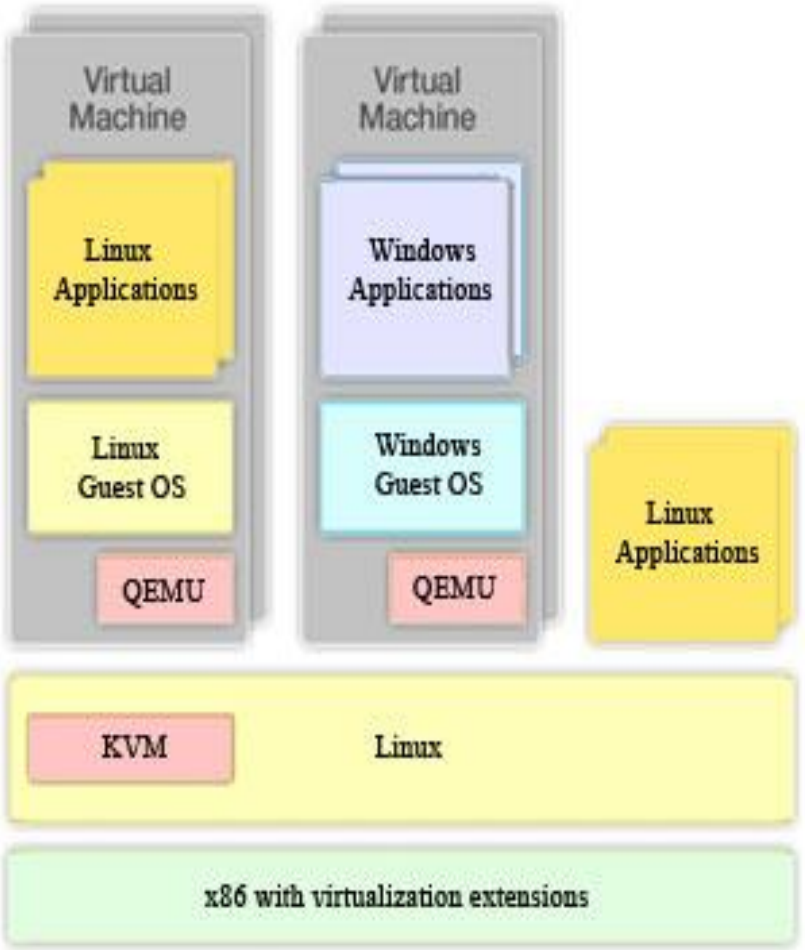
Virtual Threading

- Each threads will show as vCPU if you enable hyper threading.
- Consider you can get 60% of the core performance in thread allocate number for vCPU required

The threads policy will control how the scheduler

- avoid: the scheduler will not place the guest on a host which has hyperthreads.
- separate: if the host has threads, each vCPU will be placed on a different core. ie no two vCPUs will be placed on thread siblings
- isolate: if the host has threads, each vCPU will be placed on a different core and no vCPUs from other guests will be able to be placed on the same core. ie one thread sibling is guaranteed to always be unused,
- prefer: if the host has threads, vCPU will be placed on the same core, so they are thread siblings.

Virtual Memory



- Virtual machine is allocated with Virtual Memory
- Allocate the memory as per the requirement by the VM
- The Memory can be increased and decreased any time based on the requirement as long you have available physical memory

Virtual Memory

- Virtual Memory can be overallocated based on the Hypervisor support
- Overallocation can acheived through
  - Transparent Page Sharing (TPS)
  - Ballooning
  - Swapping
  - Compression

Virtual Memory

- Virtual can support NUMA based on the Processor and Hypervisor support
- Non-uniform memory access (NUMA) is a computer memory design used in multiprocessing, where the memory access time depends on the memory location relative to the processor.

Virtual disk

- Each VM will have a Block Virtual disk allocated to the VM on which Virtual Operating System exists
- The Virtual disk will point to a File in Storage or the storage itself or a partition or a volume of a storage
- Storage can be of three types
  - Local Storage on the computer
  - SAN (Storage Area Network)
  - NAS (Network Area Storage)

Virtual Network

- Virtual Network in a VM connected via Virtual NIC
- Virtual NIC is inturn connected to Physical NIC via Bridging or NATing or a physical device pass through
- Bridging or NATing is created by the Virtual Network with in the cloud
- Virtual Network is created by
  - Virtual Switch
  - Virtual Router
  - Virtual Firewall

Virtual disk file types

- VMWare - VMDK
- Virutalbox – VDI
- HyperV – VHD
- KVM – Raw image(img), qcow2, LVM
- Parallel - HDD

Virtual Operating System

- Virtual Operating system is the OS running on a Virtual Machine manages the Virtual Resources like vCPU, Virtual Memory and Virtual disks.
- The Virtual Operating system you can run depends on the Hypervisor support.
- Virtual Machine runs on a Hypervisor which in turn converts the VM's OS instructions to the physical device instructions
- Popular Hypervisors are:
  - Vsphere – VMWare
  - KVM – Linux
  - Xen – Linux
  - HyperV - Windows

Virtual Machine Management

- Most common virtual machine management
  - Create resources (vCPU, vMEM, vDISK) for VM
  - Modify Resources for the VM
  - Delete resources for a VM

Cloud Management Dashboard

- Cloud Management Dashboard are the portal or a application through which we will manage the Virtual machine on the VM
- Dashboard used by diffrent cloud
  - AWS – AWS Management Console
  - Openstack – Horizon dashboard
  - VMWare – Vcenter portal
  - HyperV - System center portal or Azure portal

Connectivity to Cloud

- Public Cloud
  - User connect to VM on cloud via Internet
- Private cloud
  - User connected to VM on cloud via corporate network
- Hybrid cloud
  - User connect to VM to the Corporate cloud controller which inturn connect to the VM based on the location of the VM either in corporate network or in the internet or via VPC network.
- VMs are connected via ssh or RDP depends on either it is a Linux or Windows

Cloud Storage

Cloud Storage

- Cloud storage is data storage on cloud which can be consumed as when required which is highly available and accessible over network.
- Cloud Stotage would be highly scalable
- It is charged per GB usage
- Diffrent offering based on the performence

Types of Cloud Storage

- Broadly there are three types of Cloud Storage offering based on usage types
  - Block Storage
  - Network Attached Storage (NAS)
  - Object based Stoarge

Cloud Block Storage

- Cloud Block storage is a type of data storage typically provided by a storage-area network (SAN) storage at a backend.
- SAN Storage creates volumes and provides to Cloud Compute to use it for hosting virtual disks
- Required filesystem to be created in operating system to use this storage
- Generally it is replicated across datacentre
- It is higly available across replicated datacentres

Cloud Block Storage usage

- Cloud Block storage is used in following use cases
  - Hosting Virtual operating system
  - Hosting Application
  - Hosting Application data
  - Hosting Database
  - Hosting Clusters

Disks used for Different performance

- SSD – Solid state disks
  - SSDs use NAND-based flash memory, which retains data without power.
- SSHD
  - SSHD combines NAND flash solid-state drive (SSD) with hard disk drive (HDD) technology, with the intent of adding some of the speed of SSDs to the cost-effective storage capacity of traditional HDDs.
- HDD
  - A hard disk drive (HDD) is a data storage device used for storing and retrieving digital information using one or more rigid ("hard") rapidly rotating disks (platters) coated with magnetic material.

Interfaces used for Different performance

- SAS
  - Serial Attached SCSI
    - SAS-1: 3.0 Gbit/s, introduced in 2005
    - SAS-2: 6.0 Gbit/s, available since February 2009
    - SAS-3: 12.0 Gbit/s, available since March 2013
    - SAS-4: 22.5 Gbit/s, under development and expected in 2017
- SATA
  - Serial ATA
    - SATA revision 1.0 (1.5 Gbit/s, 150 MB/s) released on on January 7, 2003.
    - SATA revision 2.0 (3 Gbit/s, 300 MB/s) released on April 2004
    - SATA revision 3.0 (6 Gbit/s, 600 MB/s) release on August 18, 2008
    - SATA revision 3.1(6 Gbit/s, 600 MB/s) mSata and SSD SATA
    - SATA revision 3.2 (16 Gbit/s, 1969 MB/s) SATA Express

Cloud Block Storage Examples

Amazon EBS Pricing

With Amazon EBS, you only pay for what you use. The pricing for Amazon EBS volumes is listed below.

Region: US East (N. Virginia)

Amazon EBS General Purpose (SSD) volumes
- $0.10 per GB-month of provisioned storage

Amazon EBS Provisioned IOPS (SSD) volumes
- $0.125 per GB-month of provisioned storage
- $0.065 per provisioned IOPS-month

Amazon EBS Magnetic volumes
- $0.05 per GB-month of provisioned storage
- $0.05 per 1 million I/O requests

Amazon EBS Snapshots to Amazon S3
- $0.095 per GB-month of data stored

Cloud Block Storage Examples

Amazon EBS Pricing

With Amazon EBS, you only pay for what you use. The pricing for Amazon EBS volumes is listed below.

Region: US East (N. Virginia)

Amazon EBS General Purpose (SSD) volumes

- $0.10 per GB-month of provisioned storage

Amazon EBS Provisioned IOPS (SSD) volumes

- $0.125 per GB-month of provisioned storage
- $0.065 per provisioned IOPS-month

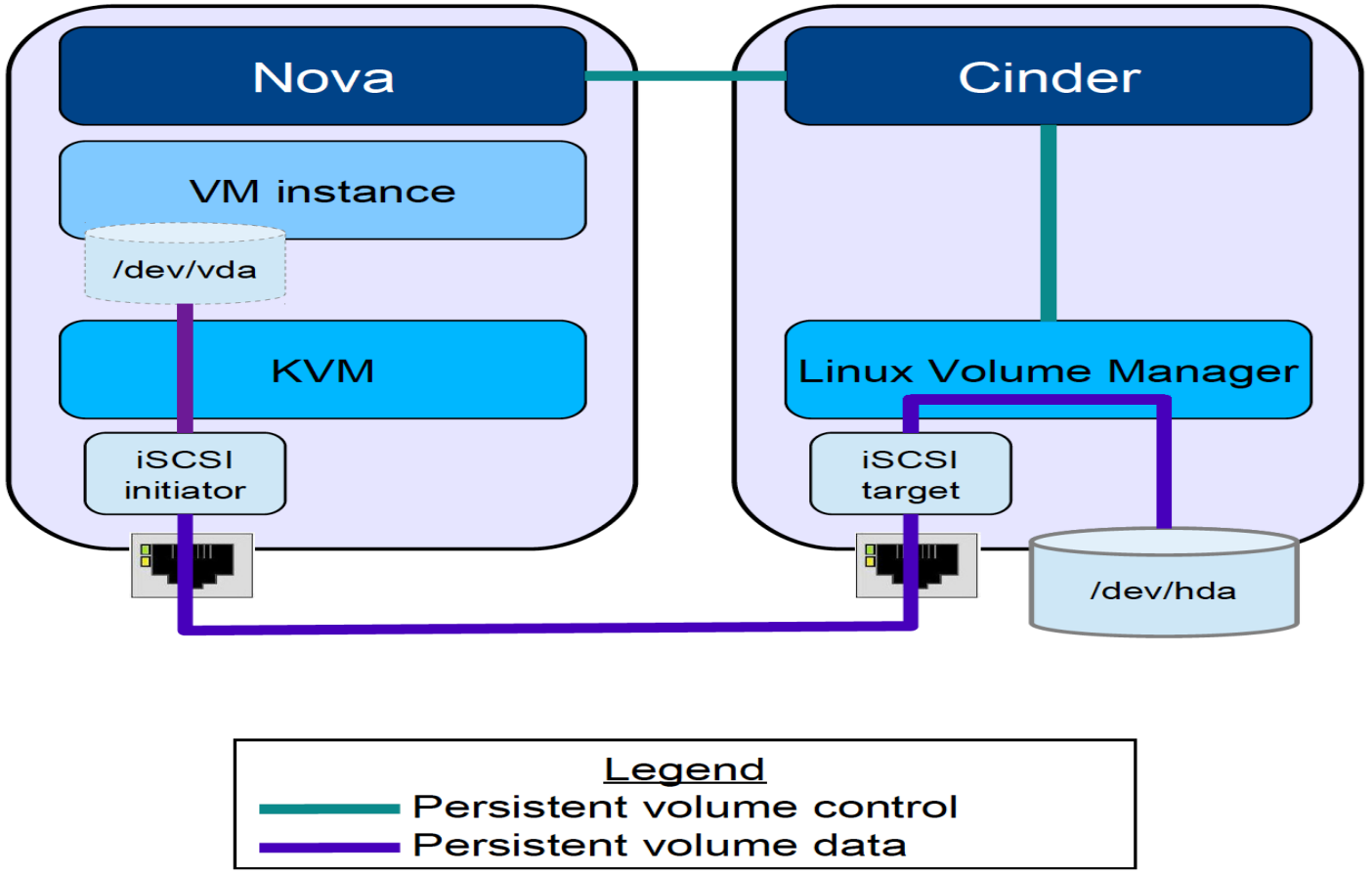
Amazon EBS Magnetic volumes

- $0.05 per GB-month of provisioned storage
- $0.05 per 1 million I/O requests

Amazon EBS Snapshots to Amazon S3

- $0.095 per GB-month of data stored

Openstack Cinder block storage



NAS Storage

- Network-attached storage (NAS) is a file-level computer data storage server connected to a computer network providing data access to a heterogeneous group of clients.
- NAS is specialized for serving files either by its hardware, software, or configuration.
- Generally it is served by NFS for unix and CIFS for windows based system
- For NFS we do not need to create filesystem at host level
- We need NAS clients to access the NAS
- Mode of transfer between the host and storage will be file based

NAS use cases

- File storage
- Used by the operating system using clients
- Can be shared to multiple system
- Home directories
- Application files which needs to be shared across hosts

Cloud NAS Storage Examples

Amazon EFS Pricing

With Amazon EFS, you pay only for the amount of file system storage you use in GB. There is no minimum fee and no set-up costs.

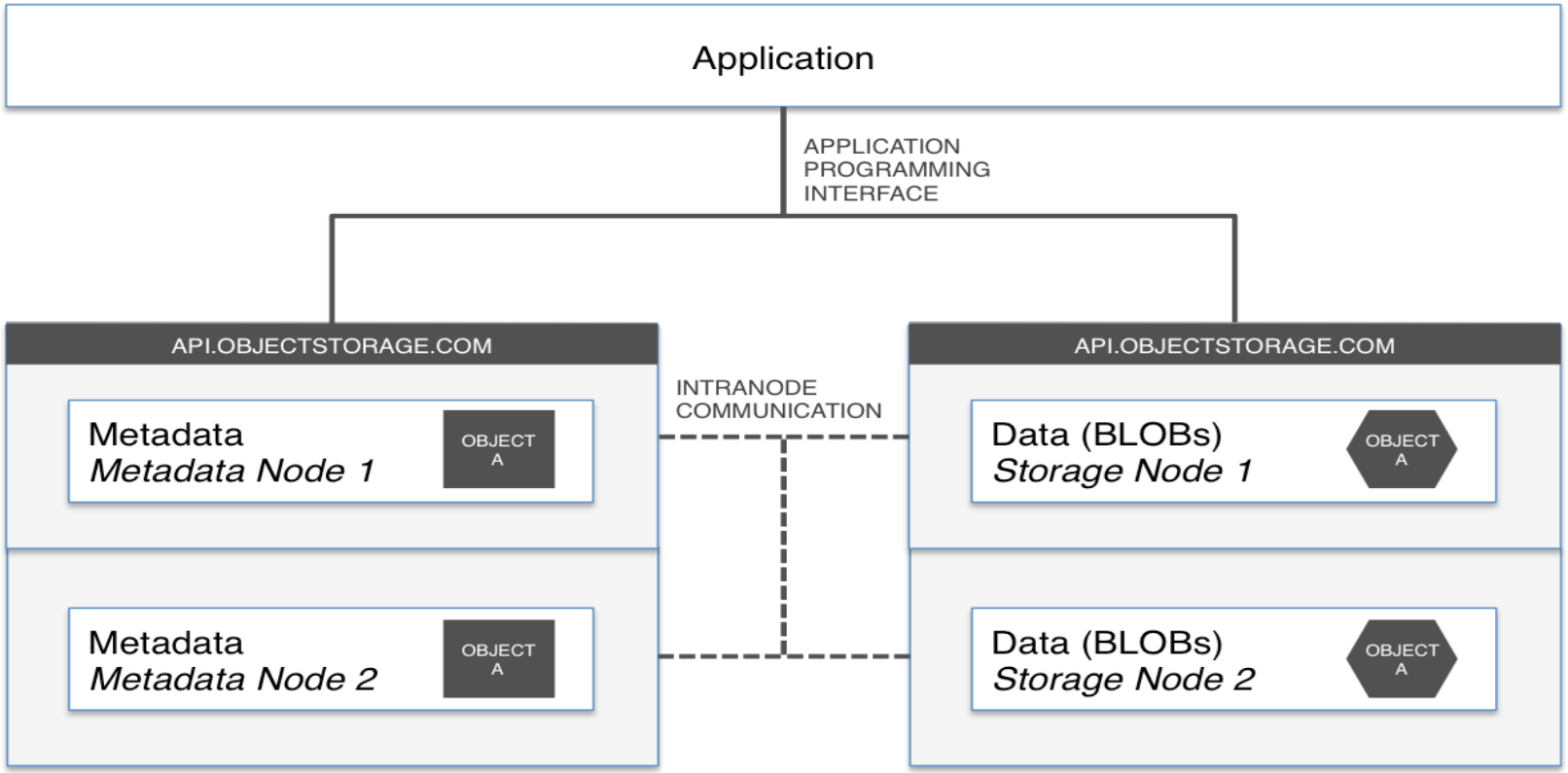
Pricing

- $0.30/GB-month

Object based Storage

- Object Storage is a storage architecture that manages data as objects
- Each object typically includes
  - the data itself,
  - a variable amount of metadata,
  - and a globally unique identifier.
- It uses web API call to access the data in the object storage using its unique identifier
- Usually accessed via client to access all the files for a account
- Usually use a webcall within the application to access the indiviual objects

Object based Storage

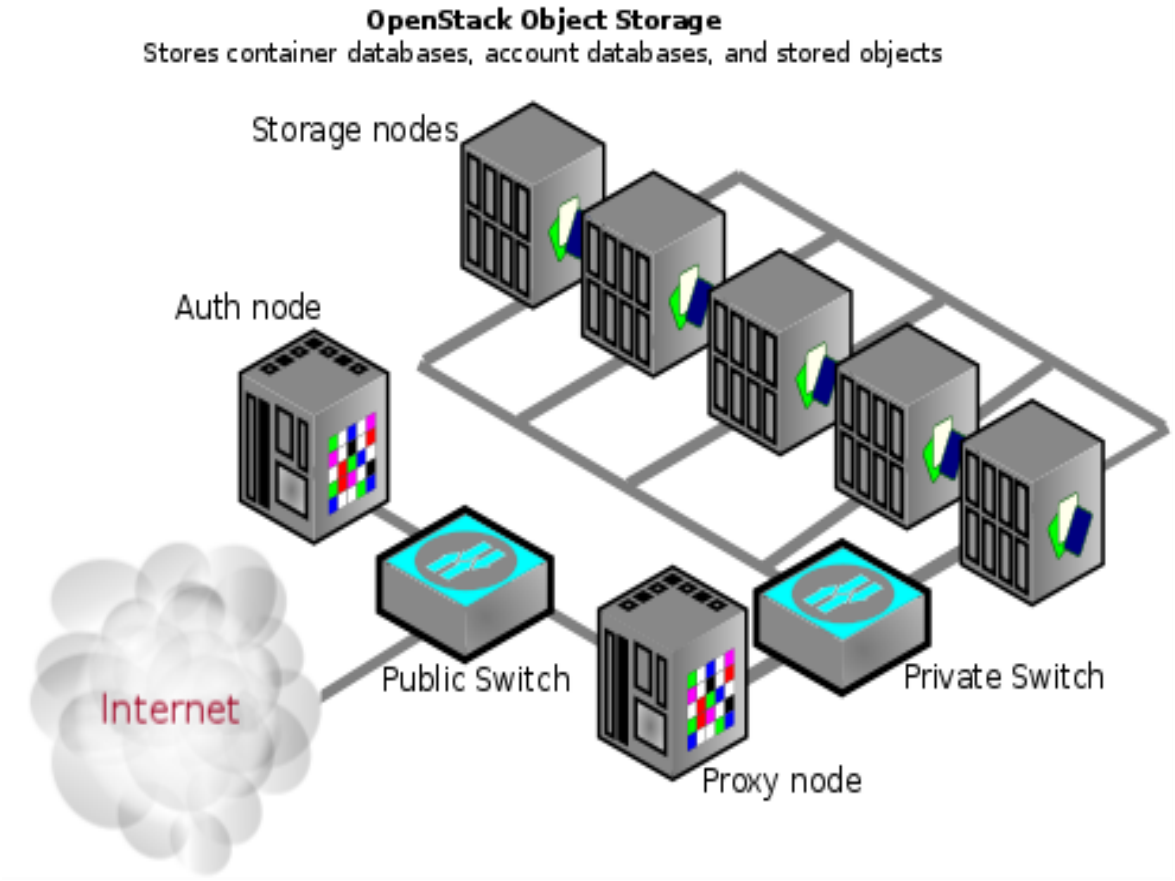


Object Storage Examples

Dropbox

Google Drive

Amazon S3



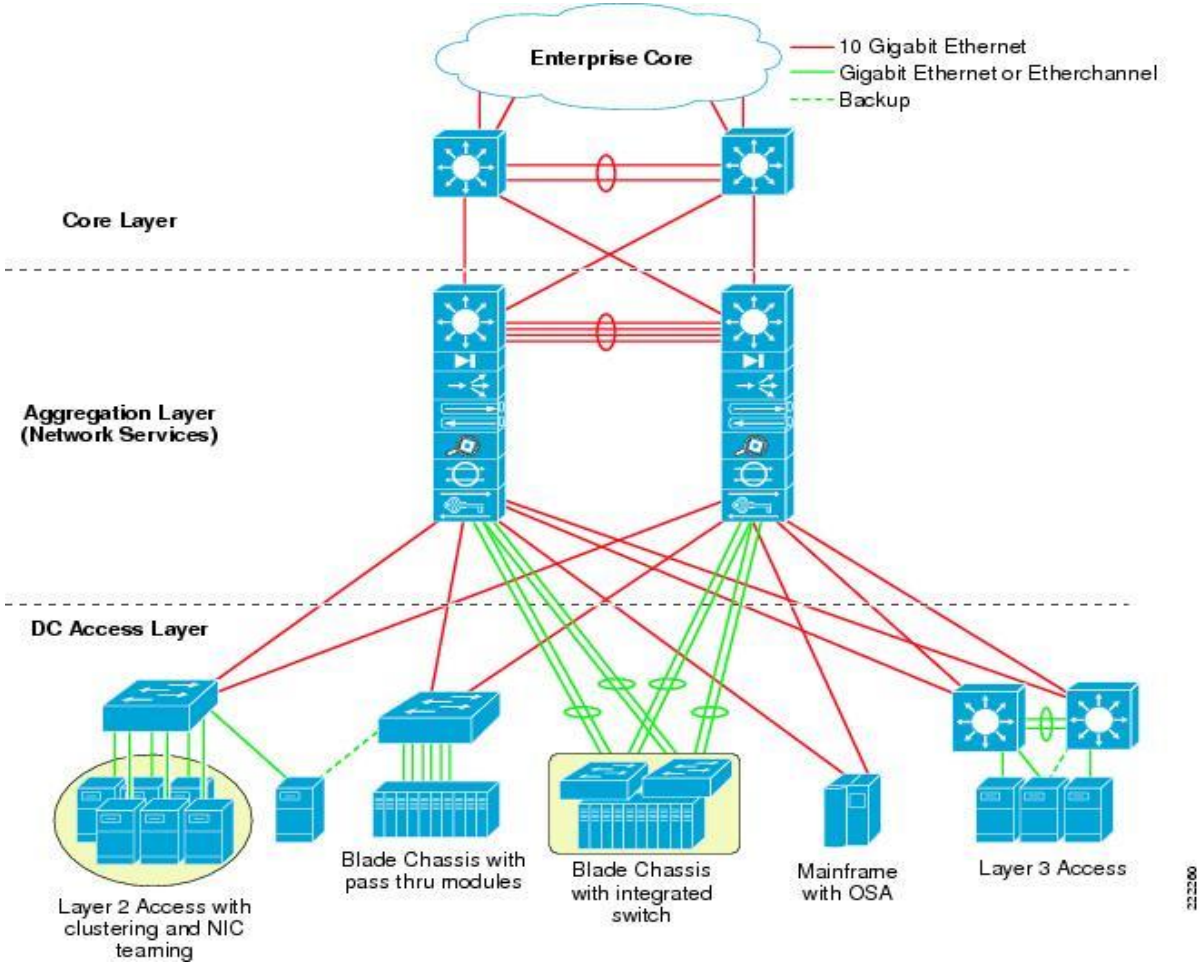
Cloud Networking

- Cloud networking is a new networking paradigm for building and managing secure private networks over the public Internet by utilizing global cloud computing infrastructure.
- In cloud networking, traditional network functions and services including connectivity, security, management and control, are pushed to the cloud and delivered as a service.

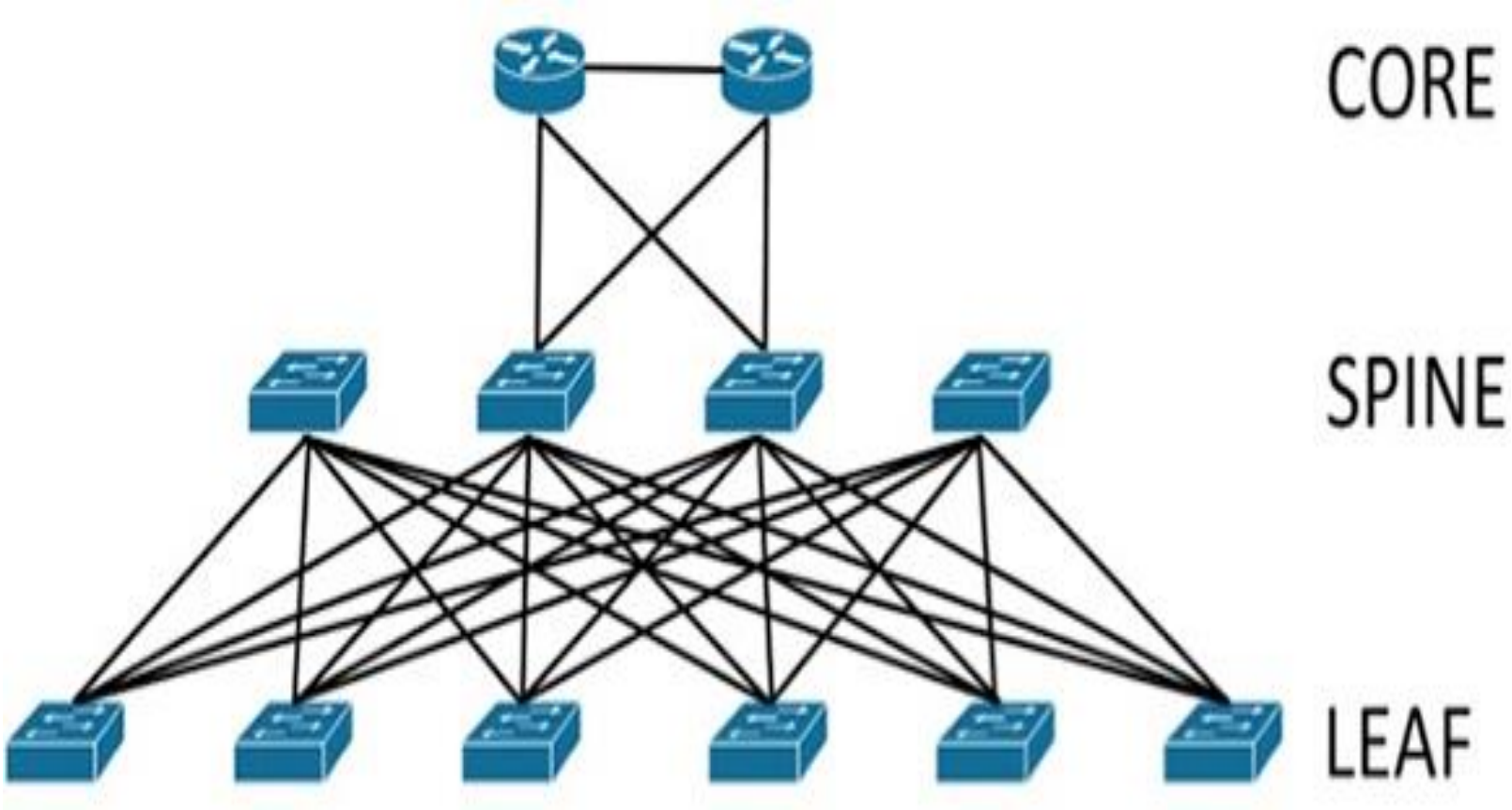
Cloud Networking Types

- There are two categories within cloud networking
  - Cloud-Enabled Networking (CEN)
  - Cloud-Based Networking (CBN)
- CEN moves management and certain aspects of control (such as policy definition) into the cloud, but keeps connectivity and packet-mode functions – such as routing, switching and security services – local and often in hardware.
- CBN moves all core networking functions, including addressing and the actual packet path, into the cloud and eliminates the need for any local hardware other than that which provides an internet connection

Typical Datacentre Networking



Leaf and Spine Model Networking



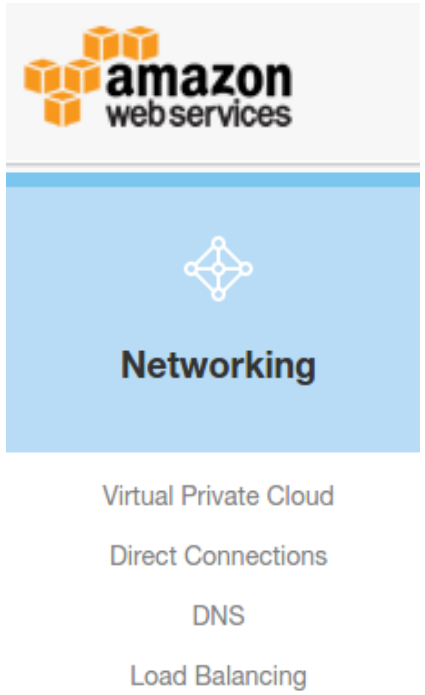
Network Virtualization

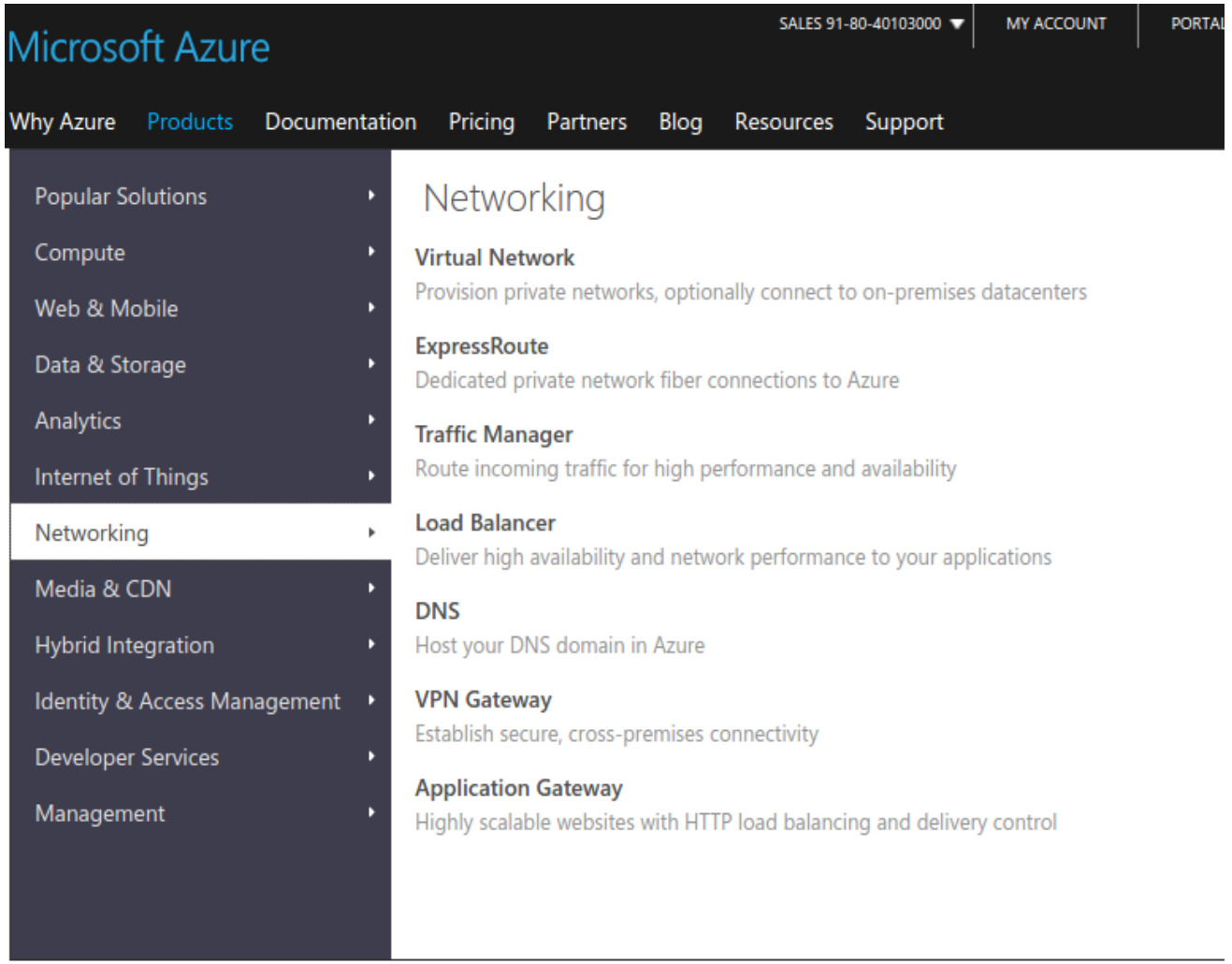
- Network virtualization is the process of combining hardware and software network resources and network functionality into a single, software-based administrative entity, a virtual network.
- There are two types of Network Virtualization
  - External virtualization
    - combining many networks or parts of networks into a virtual unit within a datacentre or spanning across multiple datacentre.
  - Internal virtualization
    - Networking virtualization done with in the server

Cloud Networking

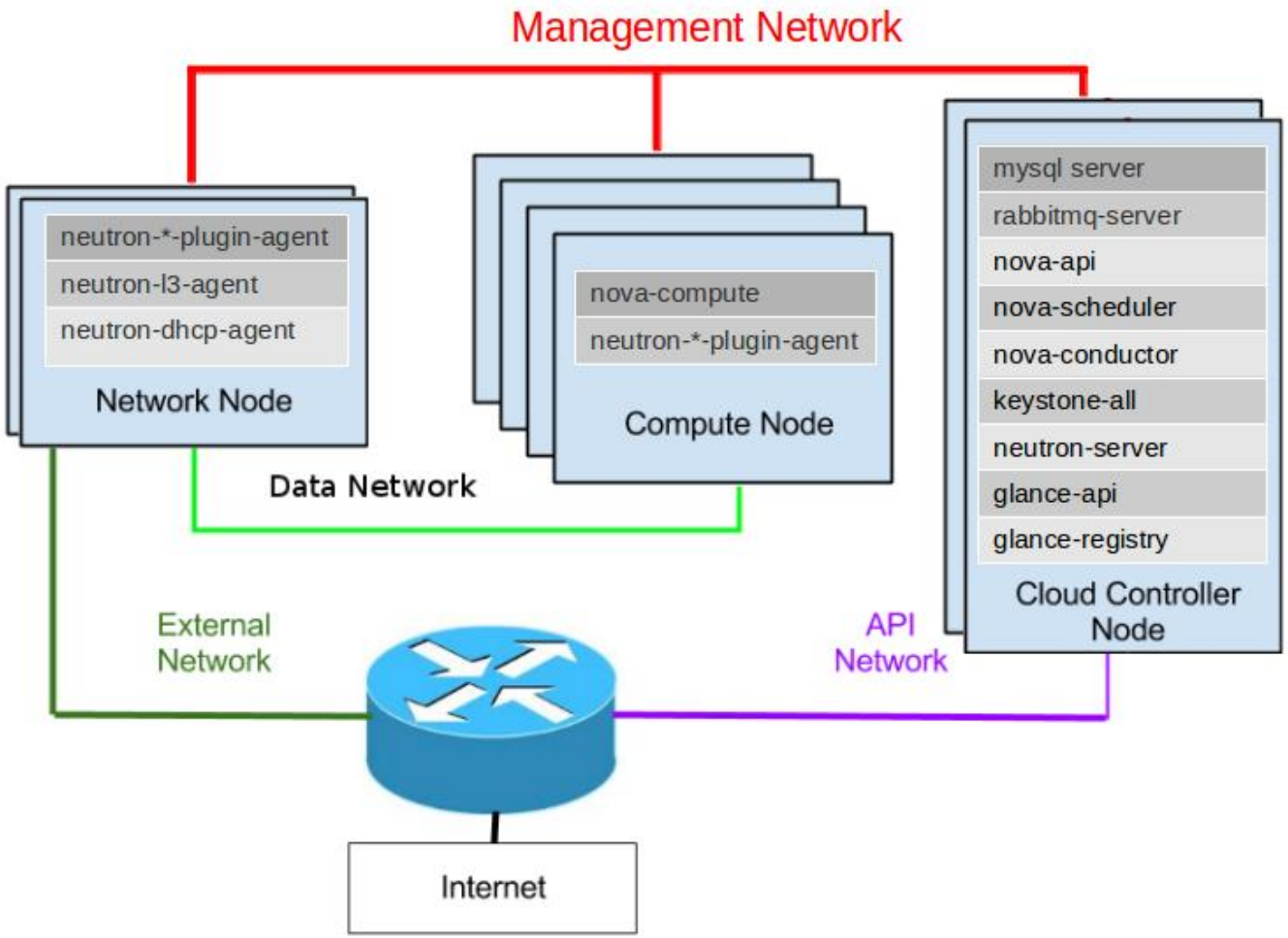
- Most of the Cloud services provide both Internal and External network Virtualization
- Internal network virtualization is used to network between cloud compute
- External network provides segmentation of network, Load balancing, Routing, Security to the network, Connectivity (Internet or Virtual private cloud)

Cloud Networking Examples





Openstack Networking examples



Cloud Management/ Shared Services

Cloud Shared Service

- Cloud Shared service
  - Identity Service
  - Image Service
  - Orchestration service
  - Usage service
  - Database service
  - Other common application services

Identity Service

- Identity service is used to autenticate and authourize the use of a user to the cloud resources
- You create user, groups, roles, permissions in identitiy service to provide right access to the right user for cloud resources
- Identity services are federated across datacentres to manage a single Identity service

Examples of identity service

AWS Identity and Access Management

AWS Identity and Access Management (IAM) enables you to securely control access to AWS services and resources for your users. Using IAM, you can create and manage AWS users and groups, and use permissions to allow and deny their access to AWS resources.

Welcome to Keystone, the OpenStack Identity Service!

Keystone is an OpenStack project that provides Identity, Token, Catalog and Policy services for use specifically by projects in the OpenStack family. It implements OpenStack's Identity API.

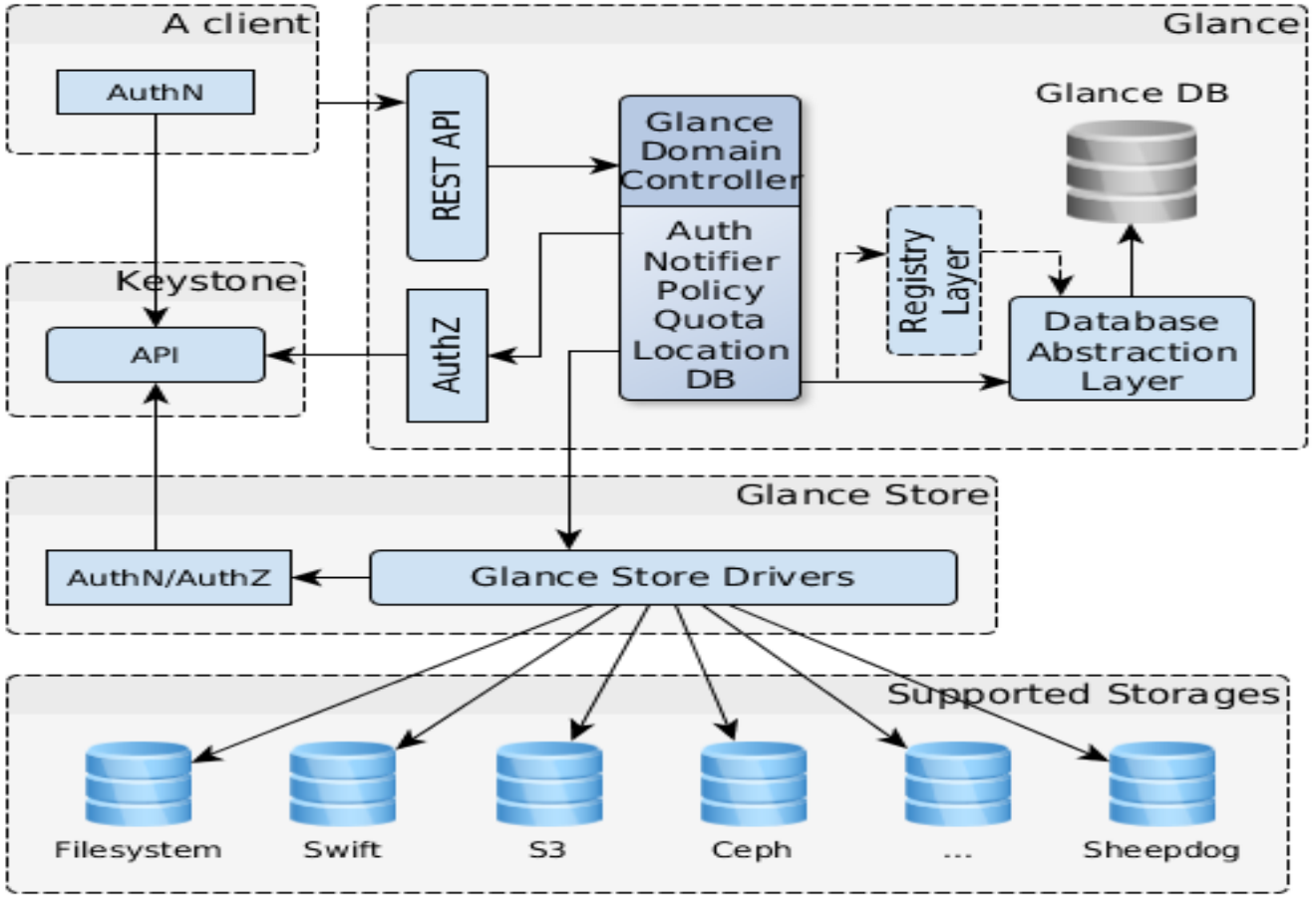
Image Service

- In Cloud Infrastructure the VM's are created from images
- The base images are created by Cloud administrators or providers
- Cloud users can create image out of the Vms which have all the application or databse installed at any time and reuse this image to create VM from them.
- To create any image from a VM we need use a Template process

Image Service

- Image service is to provide image as a service to create cloud compute Vms
- It will provide service to create New image from a VM template or modify the existing image
- Image type depends on the hypervisor type you use in the cloud, Generic image type are:
  - Raw
  - Machine (kernel/ramdisk outside of image, a.k.a. AMI)
  - VHD (Hyper-V)
  - VDI (VirtualBox)
  - qcow2 (Qemu/KVM)
  - VMDK (VMWare)
  - OVF (VMWare, others)

Examples of Image service



Orchestration Service

- Orchestration service helps to describe and automate the deployment of Cloud infrastructure
- It will provide flexible template language that can specify
  - compute,
  - Storage
  - networking configurations
  - detailed post-deployment activity
  - to automate the full provisioning of infrastructure as well as services and applications.

Examples of Orchestration service

AWS CloudFormation

AWS CloudFormation gives developers and systems administrators an easy way to create and manage a collection of related AWS resources, provisioning and updating them in an orderly and predictable fashion.

AWS OpsWorks

AWS OpsWorks is an application management service that makes it easy to deploy and operate applications of all shapes and sizes. You can define the application's architecture and the specification of each component including package installation, software configuration and resources such as storage. Start from templates for common technologies like application servers and databases or build your own to perform any task that can be scripted. AWS OpsWorks includes automation to scale your application based on time or load and dynamic configuration to orchestrate changes as your environment scales.

Examples of Orchestration service

Heat

Heat is the main project in the OpenStack Orchestration program. It implements an orchestration engine to launch multiple composite cloud applications based on templates in the form of text files that can be treated like code. A native Heat template format is evolving, but Heat also endeavours to provide compatibility with the AWS CloudFormation template format, so that many existing CloudFormation templates can be launched on OpenStack. Heat provides both an OpenStack-native ReST API and a CloudFormation-compatible Query API.

Puppet

The Puppet OpenStack Mission: to bring scalable and reliable IT automation to OpenStack cloud deployments.

Puppet OpenStack is open source, openly designed, openly developed by an open community.

Usage Service

- Usage service are used to monitor the individual cloud resources utilization
- Monitoring cloud utilization helps to charge the user for the cloud resources
- It will also collect the performence metrics of the cloud resources
- It will allow cloud operators to view metrics globally or by individual deployed resources

Example of usage service

- Amazon CloudWatch
  - Amazon CloudWatch is a monitoring service for AWS cloud resources and the applications you run on AWS. You can use Amazon CloudWatch to collect and track metrics, collect and monitor log files, and set alarms. Amazon CloudWatch can monitor AWS resources such as Amazon EC2 instances, Amazon DynamoDB tables, and Amazon RDS DB instances, as well as custom metrics generated by your applications and services, and any log files your applications generate.

Example of usage service

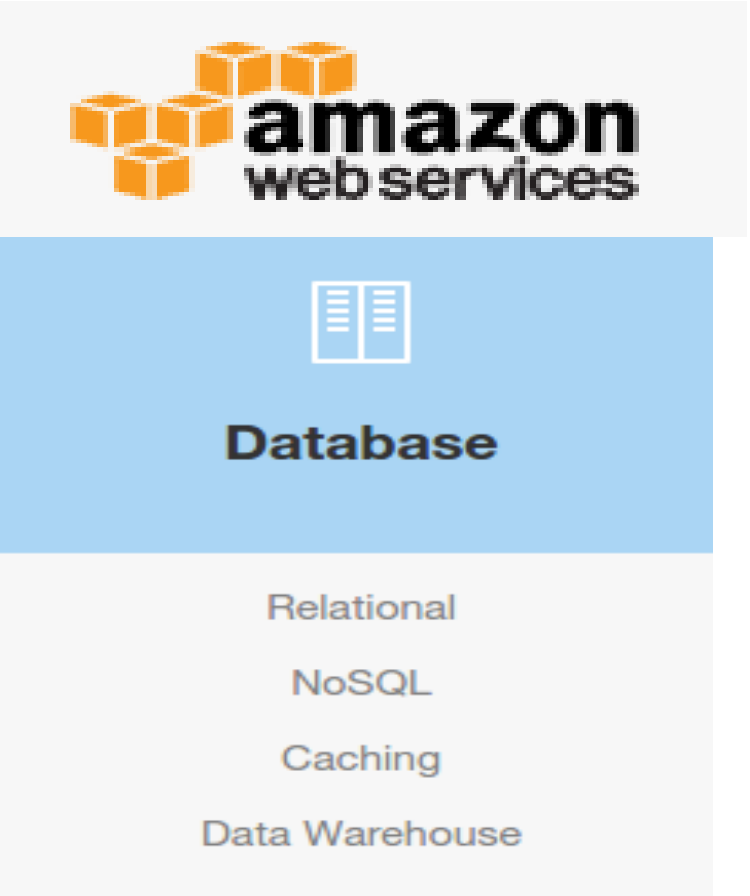
- Openstack Telemetry service

- Efficiently collects the metering data about the CPU and network costs.
- Collects data by monitoring notifications sent from services or by polling the infrastructure.
- Configures the type of collected data to meet various operating requirements. Accessing and inserting the metering data through the REST API.
- Expands the framework to collect custom usage data by additional plug-ins.
- Produces signed metering messages that cannot be repudiated.

Database service

- Database service allowing users to provision database as a service without worrying about provisiong of database application, high availability and security of database.
- Automated provisioning of database as and when requested by user or a DBA
- Most of cloud support relational and Nosql DB

Examples of Database service



- Trove
  - Trove is Database as a Service for OpenStack. It's designed to run entirely on OpenStack, with the goal of allowing users to quickly and easily utilize the features of a relational or non-relational database without the burden of handling complex administrative tasks.

Other PaaS and Saas

- There are numerous Platform as a Service and Software as a Service offered by cloud providers.
- PaaS and SaaS are achieved to fully automated platform and software deployment and scale as and when required.

Cloud Datacentres/Zones

Cloud Datacentres

- Cloud Datacentre are place to host Cloud Datacentres
- The datacentre are designed such that all physical components like power cooling and compute infrasture are highly available
- In case of any of the physical infrastructure goes bad it will seamlessely take over by another device untill it is replaced.

Cloud Availability Zones

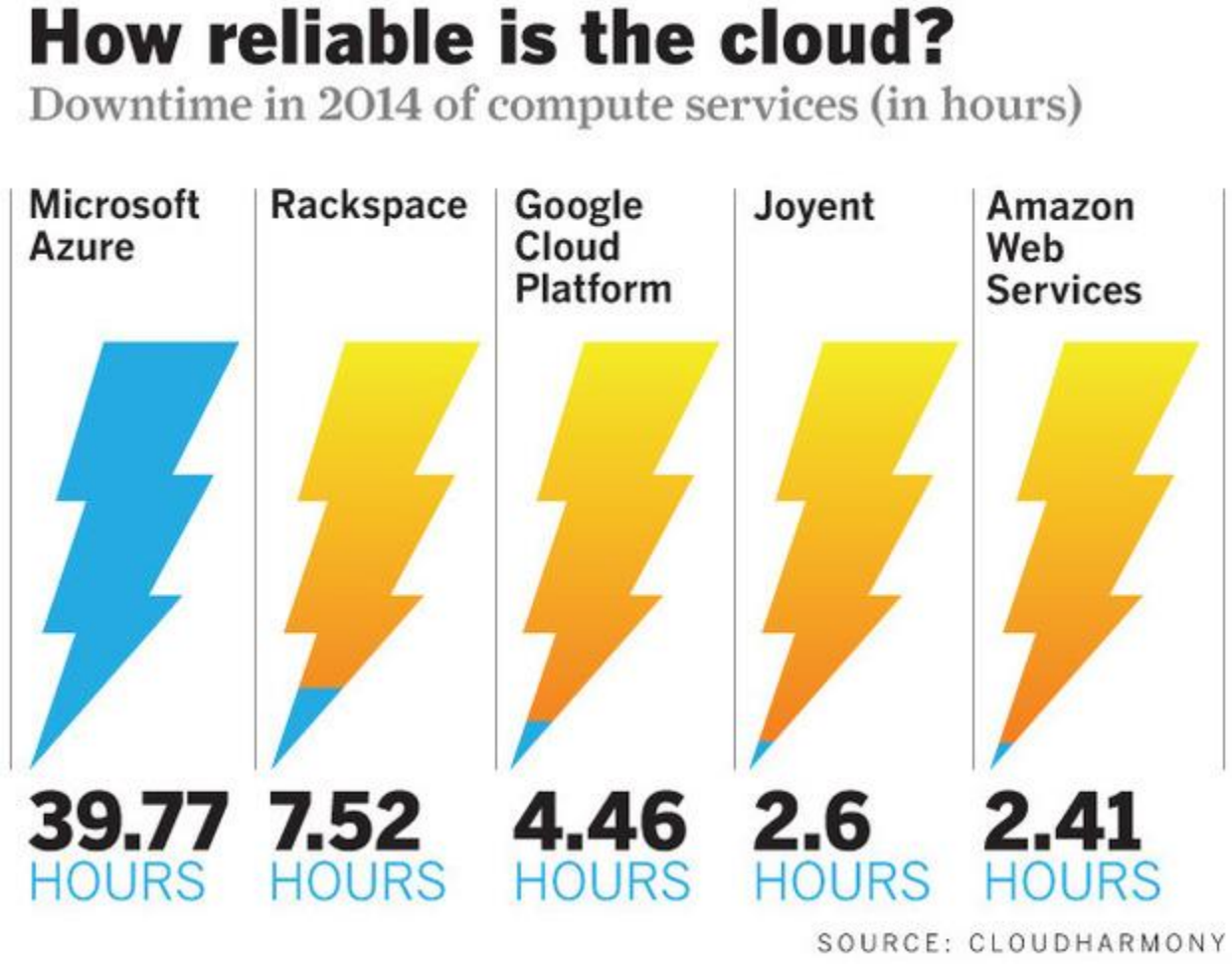
- Cloud Availability zones are created to cater the disaster scnerio
- Generally Availabilty zones are build with in multiple cloud datacentres in case of disaster in one datacenter all the services will be failed over to other datacentre

Cloud Regions

- Create Cloud Regions with multiple availability zones.
- Only few services are replicated across availability zone generally a image service, identity service.
- If you need more available of cloud VM you can deploy Vms in multiple availabilty zone and load abalance across availabilty zones
- No Data is replicated across regions. All your data reside with in the Cloud regions

Cloud Service Attributes

Cloud **Reliability**



- Business-critical services are just that: critical! Businesses cannot afford to have their IT services go down because an outage can impact the business significantly.
- Reliability is about minimizing the probability of a service outage.
- Look for service providers that have a history of minimal cloud outages.

Cloud availability

- Availability of cloud resources deployed in cloud is a key performance indicator for the cloud.
- Availability is related to reliability.
- while reliability is an indication of how often a service goes out (or conversely stays up), availability is about the recovery time when a service actually goes down.
- Ideally, you want a service provider that has the fewest and shortest outages.
- Together, this is generally expressed in a reliability/availability percentage, such as 99.9%. Look for vendors with the highest reliability/availability percentage, keeping in mind that even a 99.9% rating means that the service can go down for as much as 43 minutes every month.
- AWS will use commercially reasonable efforts to make Amazon EC2 and Amazon EBS each available with a Monthly Uptime Percentage (defined below) of at least 99.95%, in each case during any monthly billing cycle (the “Service Commitment”).

Cloud **Scalability**

- A huge benefit of cloud services is their ability to scale to whatever size you need and to be available wherever you need the service.
- Service providers need to be deployed in enough data centers around the world to support the connection of users from anywhere your company does business.
- Look for service providers that can meet the needs of your company's locations.

Cloud **Security**

- One potential disadvantage of moving your IT services to an external cloud is poor security.
- If the cloud service provider does not have the appropriate and necessary security safeguards in place, security can be compromised.
- Look for cloud providers that have multi-factor authentication built into their service as well as full encryption in-transit and at-rest for content.
- Centralized security policies to manage remote/local devices are also important.

Cloud **Quality of Service (QoS)**

- Just delivering a service is not good enough for most businesses, especially when the service has a real-time aspect to it, such as voice/video calls or web meetings.
- The key to QoS is to find service providers that control and manage as much of the solution as possible, including the service technology, bandwidth connectivity, and networking equipment.
- The more a service provider can manage and control, the better the QoS will be.

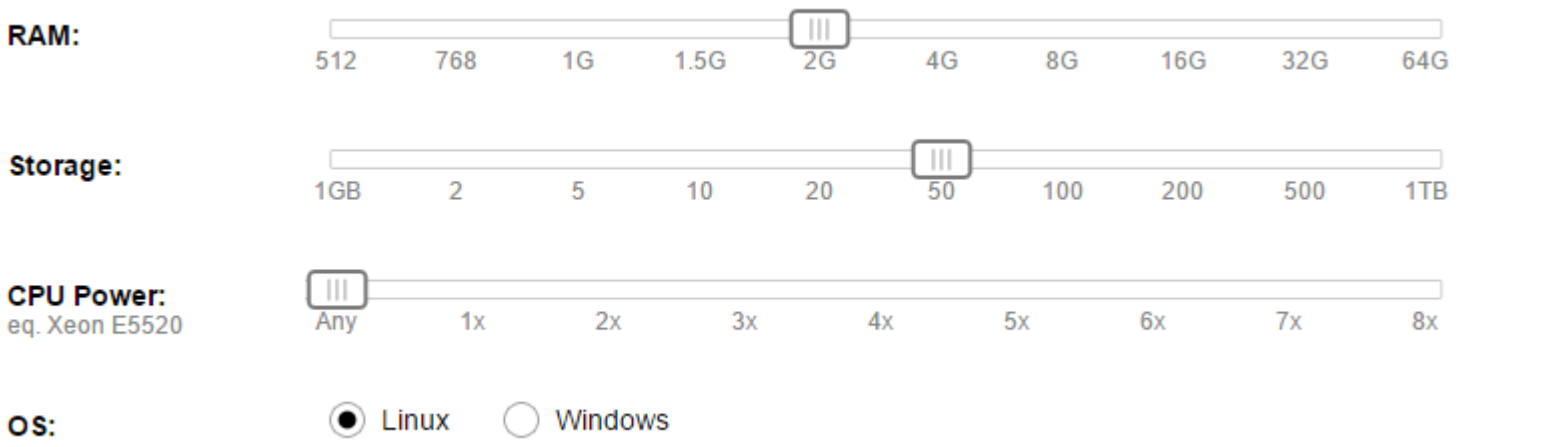
Cloud **Service-Level Agreements (SLA)**

- With most legal agreements, the goal for both parties is to never have to resort to legal enforcement of the agreement.
- However, in many cases, a strong SLA can offset the impact on a business when a service does not perform to the guaranteed level.
- As recently as just this month, Microsoft had an extended outage on its Azure platform that triggered credit compensation in its SLA.
- Compare service provider SLAs to ensure that you get the best protection.

Cloud **Support**

- Support is a key area that is often overlooked when evaluating service providers, but it remains a critical aspect of any service.
- The ability to contact support in real time through a call or IM at any time of the day is essential for companies that conduct business around the world or that have time-critical interactions with customers or prospects.
- In addition, as companies become more mobile and employees more distributed to remote locations, the ability for the service provider to provide real-time 24/7 support becomes even more critical.
- Look for service providers that can provide 24/7 real-time support.

Cloud **Pricing**



Google	Custom Machine 2 GB RAM / 1x VCPU show details	$26
Windows Azure	D1 v2 show details	$53
amazon web services	EC2 m3.medium + 50 GB SSD EBS show details	$54

Cloud **Accessibility**

- Cloud Accessibility is the key to choose the cloud.
- Weather Cloud provider provides
  - Internet Access
  - VPN Access
  - Dedicated Connectivity Access
  - Partner connectivity Access
- Look for cloud providers which gives you the accessibility what your enterprise needs are..

Cloud **Hosting**

- Cloud hosting geography also a key criteria for selection
- Look for vendor which is closer to the user to get a maximum performance
- Look for vendor which can provide the in-country compliances
- AWS has 11 Regions and more regions are in pipeline
- Azure is generally available in 22 regions around the world,
- Google cloud have 4 regions around the world

Cloud **Hosting**

- Cloud hosting geography also a key criteria for selection
- Look for vendor which is closer to the user to get a maximum performance
- Look for vendor which can provide the in-country compliances
- AWS has 11 Regions and more regions are in pipeline
- Azure is generally available in 22 regions around the world,
- Google cloud have 4 regions around the world

Cloud **Provisioning**

- Cloud Provisioning process is the key to select the cloud.
- Does cloud provides the technologies you use for provisioning and automation like Puppet, Chef.
- Does Cloud supports Orchestration will be a key to choose a cloud providers
- Greater the Automation, lesser the provisioning cost.

Cloud **Total Cost of Ownership**

Cost Drivers	Traditional On-Premises Software	Cloud Application
Capital Expenses	• Upfront purchase of software and hardware • May require network infrastructure enhancements, facilities • Need to support third-party monitoring, test tools, security products	• None • Pay-as-you-go subscription pricing • All inclusive: maintenance, support, training, and upgrades all hardware, networking, storage, database, administration
Design and Deployment	• May take months to deploy • Professional services can cost up to 3X the initial software purchase • Difficult for vendor to build best practices • Requires staff or contract labor to research, design, integrate, test, tune, launch, and train	• Deploy in weeks • Lower cost using consistent set of best practices

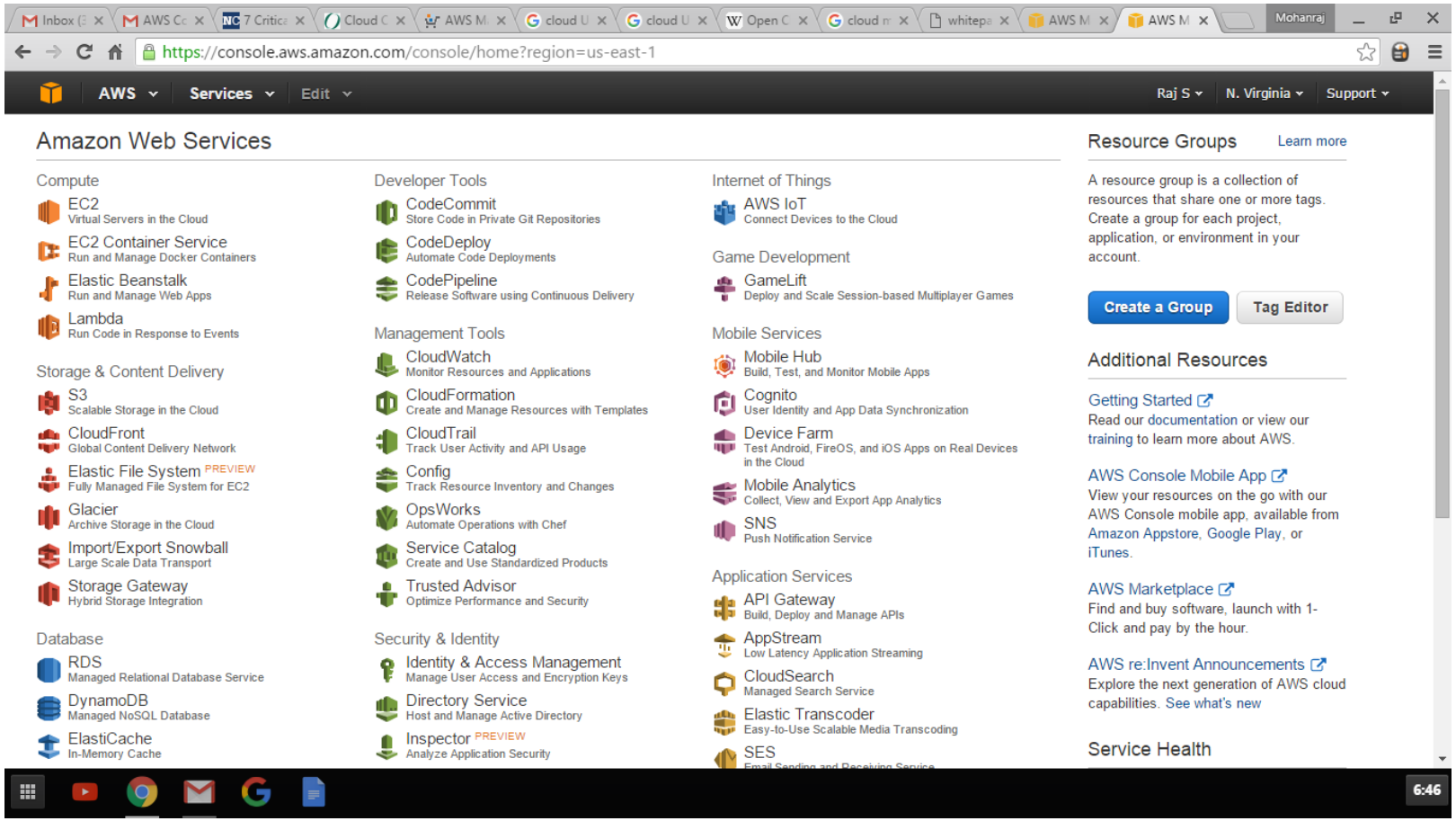
Cloud **Total Cost of Ownership**

Ongoing Infrastructure	• Ongoing software maintenance, upgrades • Ongoing hardware replacement once every three years • Requires network monitoring and management tools • May require additional networking equipment and bandwidth to accommodate incremental traffic	• Vendor provides as part of subscription
Ongoing Ops, Training, Support	• Requires resources to operate, monitor, support, and upgrade the application • Need to hire, train and certify support personnel	• Vendor provides as part of subscription • There may be some training fees • Customer must ensure adequate Internet access and bandwidth

Cloud **User Interface**

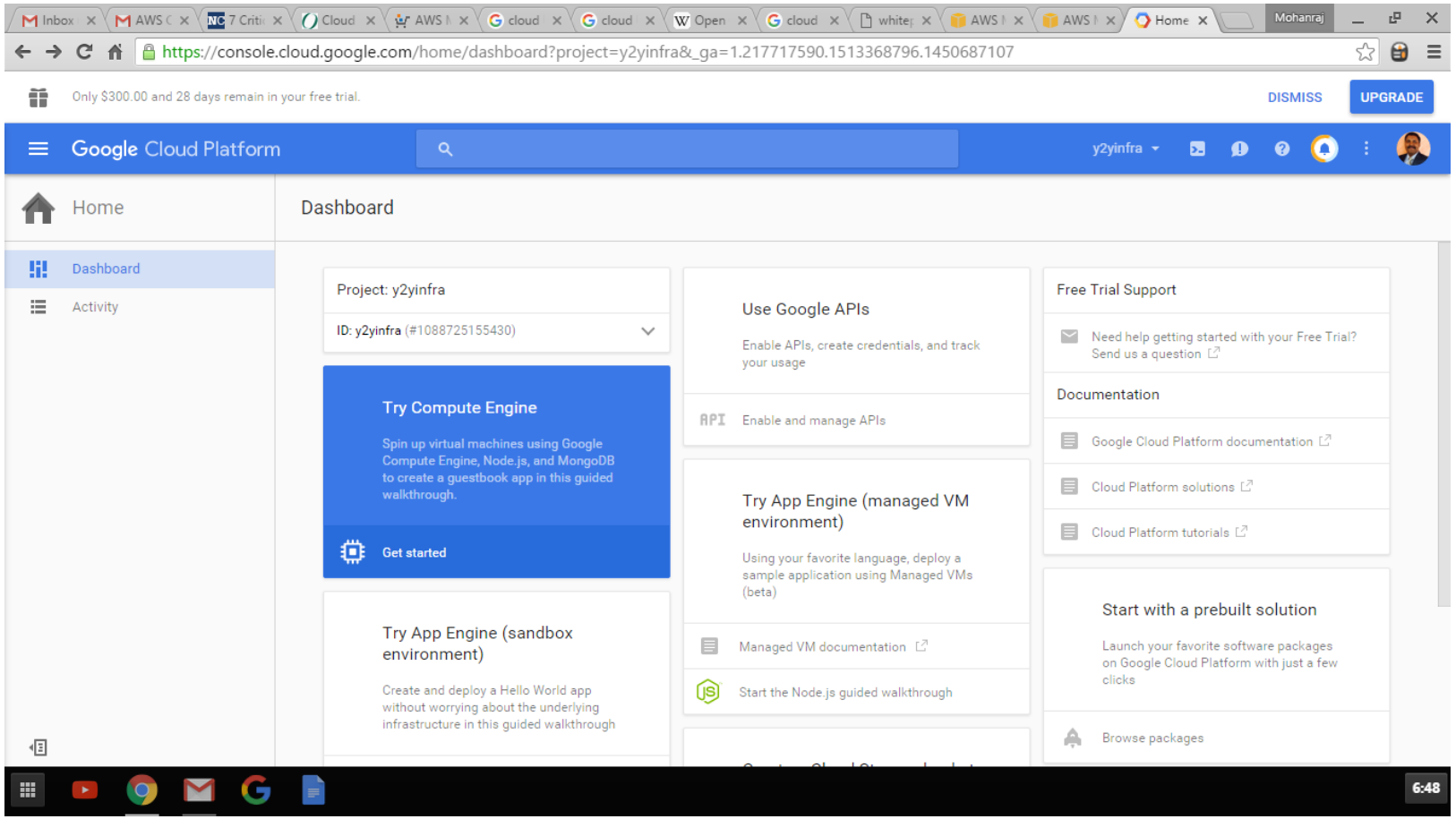
- The Cloud is generally accessed via
  - Dashboard or Management Console
  - Command line Interfaces
  - Using API

Cloud **Management Consoles**



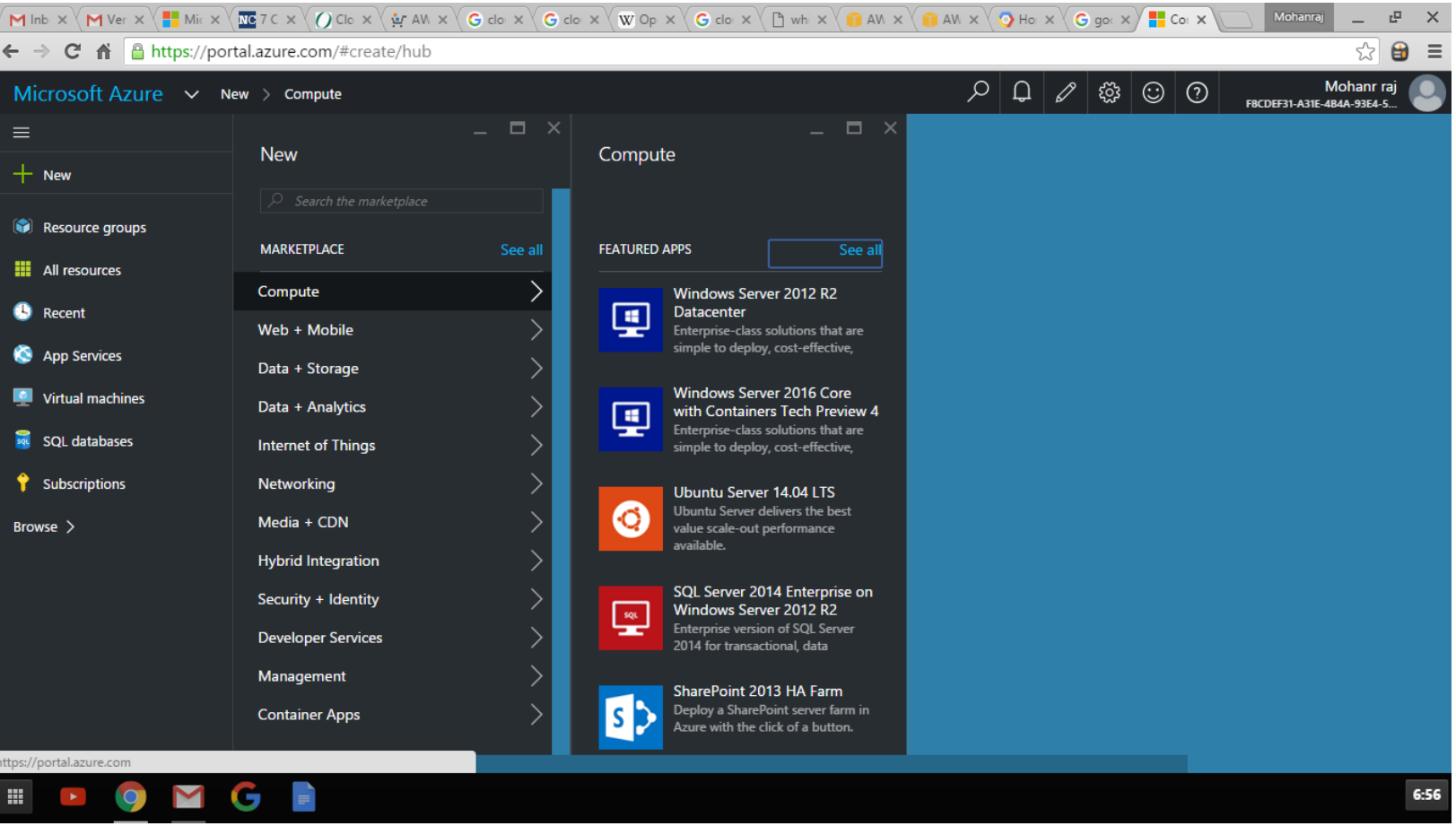
- AWS Console
- Access and manage Amazon Web Services through a simple and intuitive web-based user interface. You can also use the AWS Console mobile app to quickly view resources on the go.

Cloud **Management Consoles**



- Google Cloud Platoform console
- https://console.cloud.google.com

Cloud **Management Consoles**



- Microsoft Azure Portal
- https://portal.azure.com

Cloud **CLI**

- The AWS Command Line Interface (CLI) is a unified tool to manage your AWS services. With just one tool to download and configure, you can control multiple AWS services from the command line and automate them through scripts.
- https://aws.amazon.com/cli/
  - Download Windows, Linux and Mac CLI Tools

Cloud Services **API**

- Mostly all Cloud Software exposes their services as API (Application Programming Interface)
- The Action of create, Delete and Modify of a service can be performed via an API
- A Sample API call will look like this:
  - https://ec2.amazonaws.com/?Action=AttachVolume &VolumeId=vol-1a2b3c4d &InstanceId=i-1a2b3c4d &Device=/dev/sdh &AUTHPARAMS
- Sample Output will be
  - <AttachVolumeResponse xmlns="http://ec2.amazonaws.com/doc/2015-10-01/"> <requestId>59dbff89-35bd-4eac-99ed-be587EXAMPLE</requestId> <volumeId>vol-1a2b3c4d</volumeId> <instanceId>i-1a2b3c4d</instanceId> <device>/dev/sdh</device> <status>attaching</status> <attachTime>YYYY-MM-DDTHH:MM:SS.000Z</attachTime> </AttachVolumeResponse>

Accessing Cloud VMs

- Most Cloud Providers provide two types of Virtual Operating System:
  - Windows
  - Linux
- For Linux VMs will give ssh access to access the VMs
- For Windows VMs will give RDP to access the VMs

Five levels of Redundancy

- Cloud Provides highest level of Redundancy with Five levels of Redundancy,
  - Physical hardware level redundancy
  - Virtual resource redundancy
  - Availability zone redundancy
  - Region redundancy
  - Cloud level redundancy
- It again depends how much level of redundancy you need for an given application
- Each level of redundancy comes with the additional cost.

Physical Hardware level Redundancy

- Here are the techniques used to maintain physical hardware level redundancy
  - All Physical local disk are in RAID to avoid any single disk failure
  - Any Hardware failure the All applications, VMs and Container will be migrated to another hardware using clustering
  - Not storing any data locally, all data are centralized storage and available to group of physical server clustered to make data available instantly on the other server
  - Network Cards are Bonded to have redundancy of the hardware failure
  - HBA connecting storage will have Multipathing to avoid any failures
  - Use of Convergent Hardware like UCS will provide more flexibility at hardware swapping by swapping service profiles

Virtual resource level Redundancy

- Here are the techniques used to maintain virtual resource level redundancy
  - Live migrate the VMs in case of any failure
  - Have virtual Multipathing to access storages
  - Have Multiple controller node for storage and Network to avoid any failure of Network or Storage controllers
  - Have a full baked image to start the VMs in case of VM got corrupted
  - Isolate OS from Data to bring up on application on other physical hardware using the data
  - Create Application with RESTFULL ARCHITECTURE complaint so that it can be load balanced in each layer
  - Load balance the Web centric application
  - VM level clustering for Non web applications

Availability Zone level Redundancy

- Here are the techniques used to maintain Availability Zone level Redundancy
  - Replicate Data across availability zone so that data is available on multiple availability zone to start the application
  - Load balance the REST complaint application across availability zone
  - Have low latency between availability zone for faster replication of Data
  - Have Data replicated across multiple availability Zone
  - Have VM level clustering across availability zones
  - Have Image available at all availability zones across region
  - Create frequent Snapshots to create a Block storage from the recent snapshot possible

Region level Redundancy

- Here are the techniques used to maintain Region level Redundancy
  - Generally Cross region replication is only provided by few Cloud offering like S3 Cross region replication
  - Google cloud offer Cross region load balancing to load balance across regions
  - And use Cloud front or traffic management to route traffic across region in case of any region failure all traffic is routed to available regions
  - Due to high latency between regions do not have clustering or Live data relication across regions.
  - Region level failover have to act a Disaster recovery plan in case of one region fails start the application from other regions using snapshots and cross region s3 bucket

Cloud level Redundancy

- Here are the techniques used to maintain Region level Redundancy
  - Cloud Level redundancy is again a Disaster recovery Methods
  - You have have Private cloud as a your primary cloud for service in case of failure you should DR to Public Cloud or another private cloud
  - this will provide any issues on the cloud level
  - This would require data gets synced between Private cloud and public cloud
  - Have good connectivity to public cloud to access application when there is a failure

Cloud Non Functional Requirement

Cloud availability

- Availability of cloud resources deployed in cloud is a key performance indicator for the cloud.
- Cloud VM Availability is achieved by following types:
  - Load Balancing
  - VM Online Migration
  - Application clustering

Disaster Recovery

- In case of a Datacenter failure the DR policy will kick off in cloud.
- Cloud DR should be automated as much as possible so that DR happens seamlessly in case of disaster
- DR generally done by
  - Migrating Instance from One Datacentre
  - Stroage replication
  - Load balancing across Datacentre
  - Application Geo clustering across datacentre

Capacity

- Cloud Datacentre should be capable of provide capacity to cloud infrastructure as and when required.
- Managing capacity and adding right capcity to the cloud is the key to cloud success
- Capacity addition and deletion needs to seamless and automated

Scalability

- All cloud resources should be scalable globally
- For example
  - VM should be scalable with in datacenter either vertically by increasing the VM size or horizantally by increasing number of Vms and load balanced
  - It should be scalable across datacntres to avoid any disaster
  - It should be scalable across region to be availble near to the user

Performance

- Performence expectation should be clearly laid out for any cloud infrastructure
- There should be multiple performence offering for the user to choose on the cloud
- Perfomence should be monitored to ensure the performence requirement is met if not scale accordigly

Maintainability

- Cloud resouces should be easily Maintable
- Design complexity should be low to have good maintainable cloud envirоinment
- Overengineering in cloud will make Cloud resources less maintanable
- Use common resources to have cloud resources easily maintanable

Monitoring

- Monitoring of individual Cloud Resources and application is key performence indicators
- In case of failure of resources there should be automatic fault tolerent method needs to be implemented

Compliance and Security

- Compliance and Security implementation of data in cloud is key to cloud adoption
- Data needs to rightly encrypted when it is at rest and in transit
- Cloud Infrastructure should comply with most of the compliance requirement like SOX, PCI, HIPPA