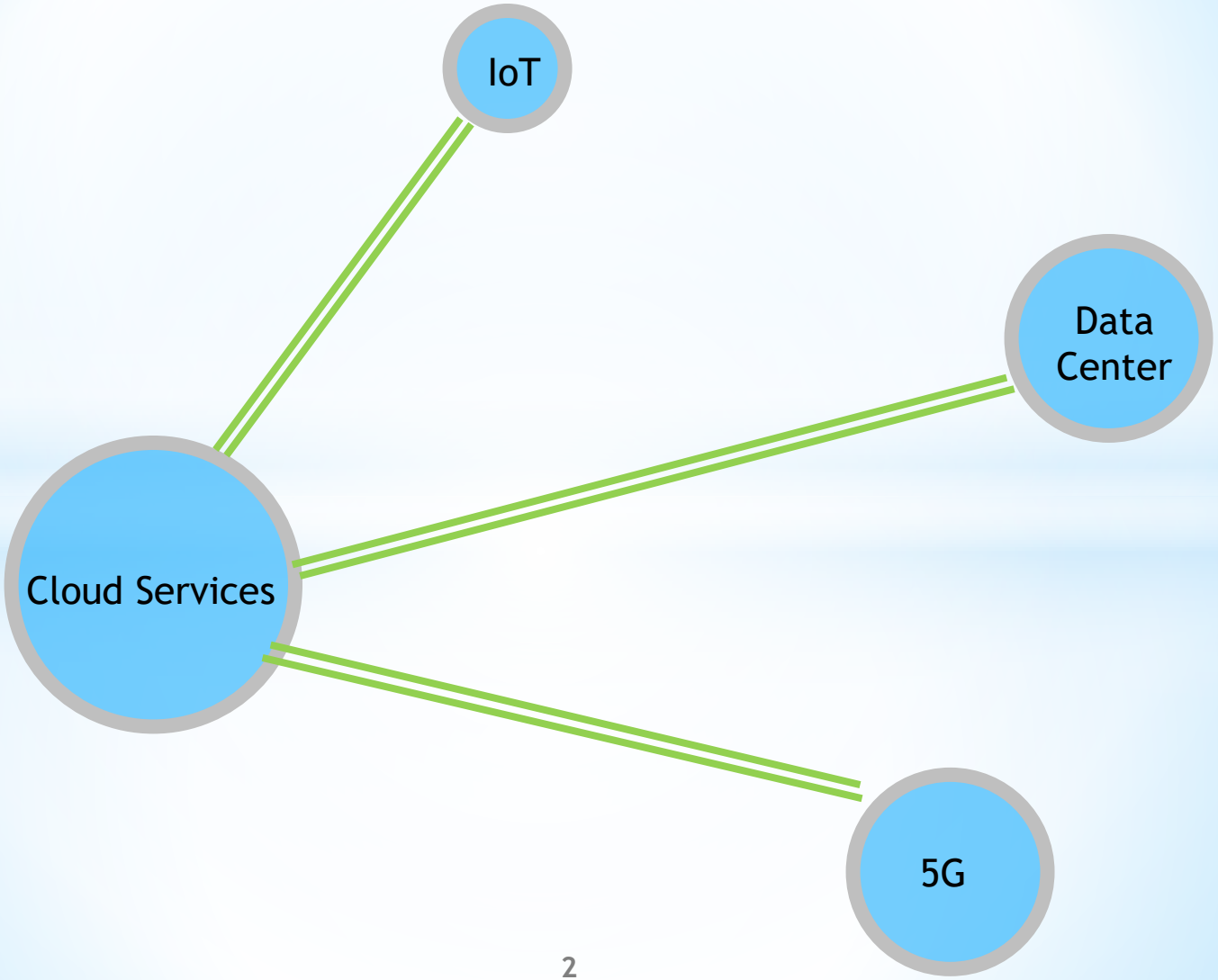
Cloud Computing Essentials

Topic of Current Interest

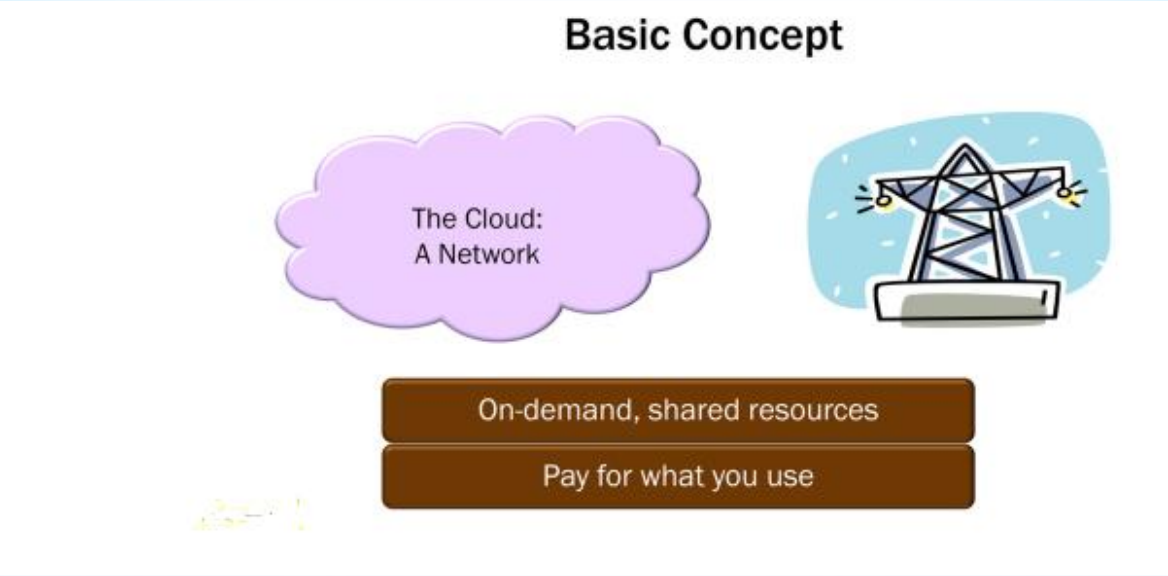
CS Department @ University of Karachi

Cloud Computing ‘Center of Gravity’



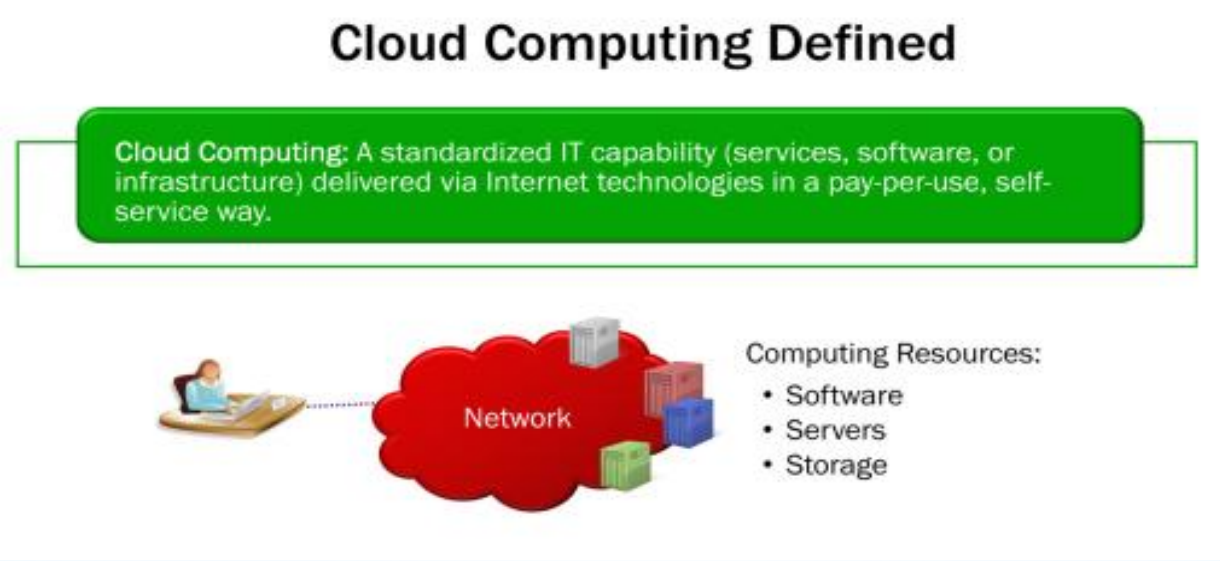
Cloud Computing

Cloud Computing [Accessing shared, metered, self-serve computing resources]



Cloud computing in simplistic terms is Internet-based computing, where the servers and data for a particular service are located on the Internet, rather than on a client device. For example, Google's hosted email service Gmail is considered a cloud application, as are Flickr, salesforce.com, and Microsoft's OneDrive storage solution.

Cloud computing is often compared to obtaining electrical service from a power company. Both feature on-demand, shared resources and users pay for what is used. So instead of buying dedicated servers, users sign a contract to get access to remote processing power and/or remote data storage and then pay for what is actually consumed. This model lets companies with bursty utilization benefit by rapidly or even automatically gaining access to additional processing power when they need it and then releasing it when they don't

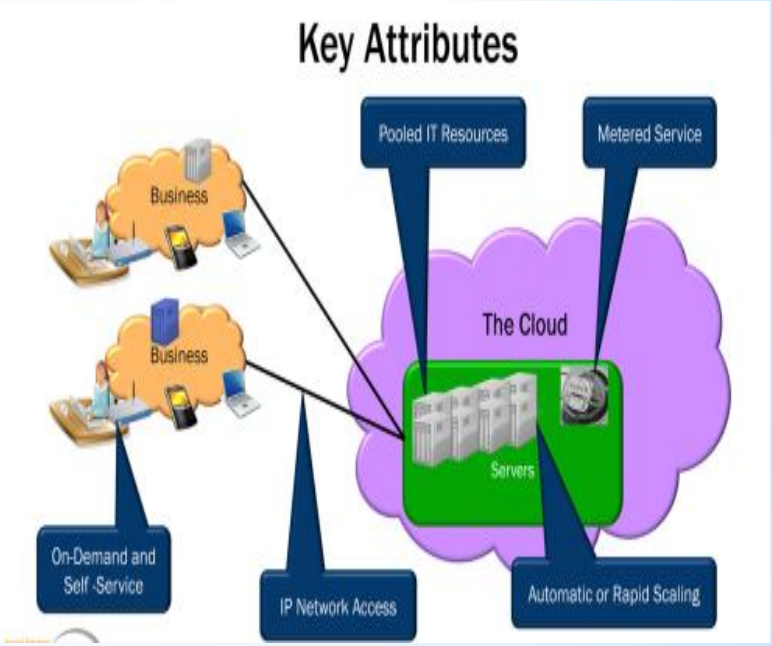


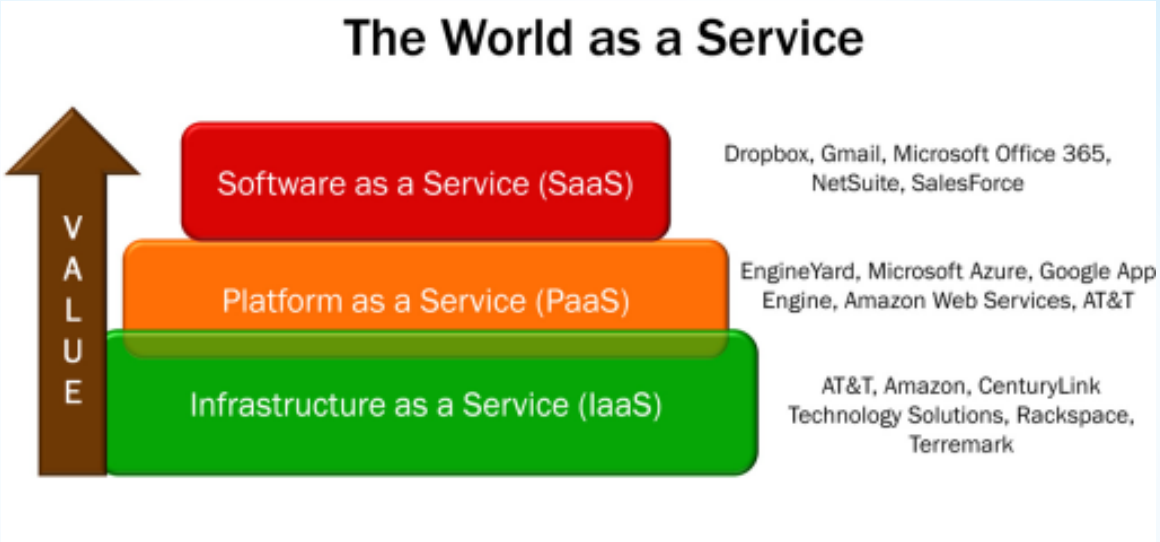
The definition of cloud computing itself can be a bit cloudy. Since cloud computing is a growing trend, many companies are looking for ways to offer their services over the Internet (one form of the cloud), and this diversity of applications can confuse those trying to determine what cloud computing actually means. Forrester provides a helpful working definition:“Cloud computing refers to a standardized IT capability (services, software, or infrastructure) delivered via Internet technologies in a payper-use, self-service way.” It encompasses a wide variety of solutions, from a consumer using a Web-based Internet-security software service, to a business leasing servers hosted by a third party for a new Web service or development project.

Key Attributes

The key attributes of cloud computing solutions can be distilled into five categories.

1. **Network Access:** Resources are accessible over an IP network from thin or thick client devices (e.g., the browser on a laptop or mobile device).
2. **Pooled Resources:** The provider's computing resources (e.g., virtual machines, storage, memory, network bandwidth) are pooled to serve multiple consumers on a multi-tenancy model.
3. **Automatic or Rapid Scaling:** Resources are monitored, automatically provisioned and released as necessary.
4. **Metered Service:** Resource usage (e.g., network bandwidth used, HTTP sessions, data store used) is monitored and charged based on a defined usage fee schedule.
5. **On-Demand and Self-Service:** Cloud services typically provide Web-based interfaces through which customers can request more (or fewer) resources.





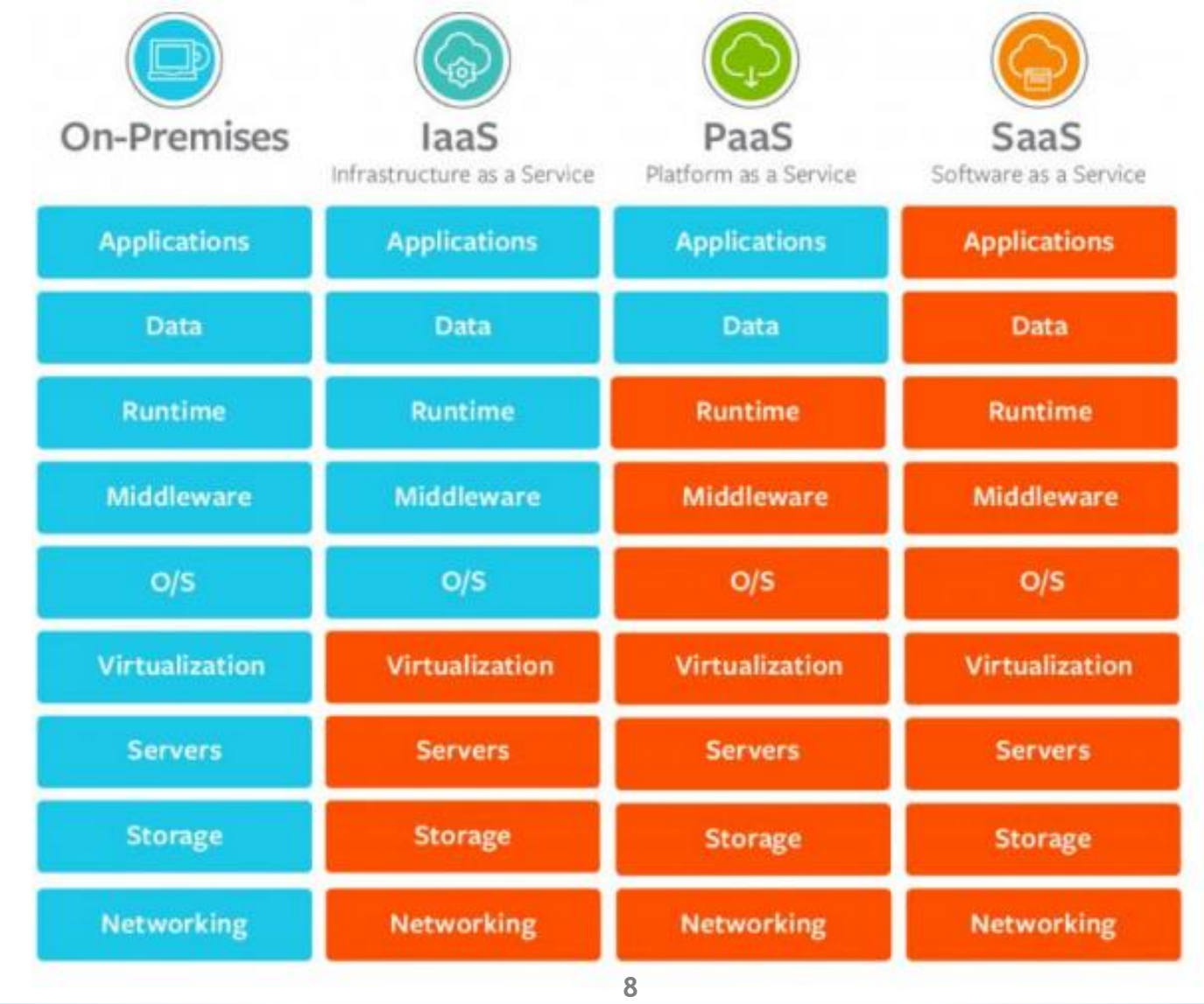
The cloud computing paradigm delivers technology as a service. Three of the core cloud computing offerings are Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS), with the lion's share of cloud offerings falling into IaaS and SaaS.

Infrastructure as a Service refers to offering computing and storage resources from the cloud. Businesses can dynamically request more resources as needed and pay for what they use. IaaS providers generally manage data centers around the world and incorporate availability, redundancy, and security into their solutions.

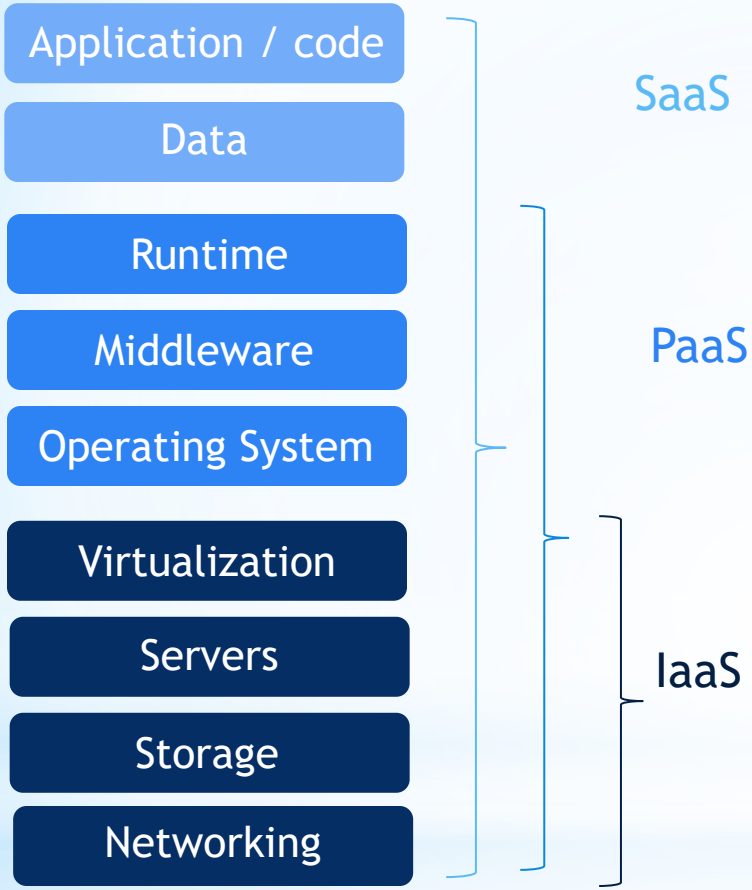
Platform as a Service is a way for cloud providers to make hardware and software resources available to developers for developing applications. By employing a PaaS approach, a development team can be developing within minutes.

Software as a Service means offering software applications that are hosted in the network, rather than installed on individual computing devices. SaaS applications are often accessed through a browser or thin client. The SaaS model offers efficiencies in maintaining and upgrading software as well as keeping application data available in the event of a device crash. A key business consideration related to the SaaS model is whether and how the software can be accessed when a user doesn't have network access.

Coud Computing Model (4 7 9)



Cloud Stack



Minutes: Minutes to have subscription ready before to use business required software

Hours: Minutes to setup and deploy environment focus on your apps and their data. Setup environments and depoy apps very quickly

Days: Minutes to provision VM but time requires to configure software, application, varies, Maintenance/upgrade of OS, Middleware, run, etc.

Cloud Services

Google Apps, Dropbox, Salesforce, Cisco Webex, Gotomeeting

SaaS

CloudFoundary, Azure, Google App Engine, AWS Elastic Beanstalk

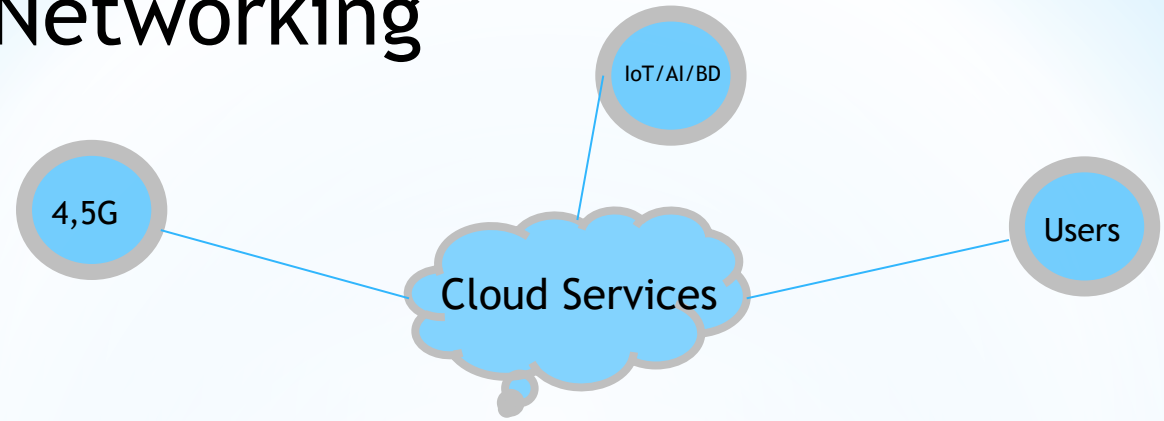
PaaS

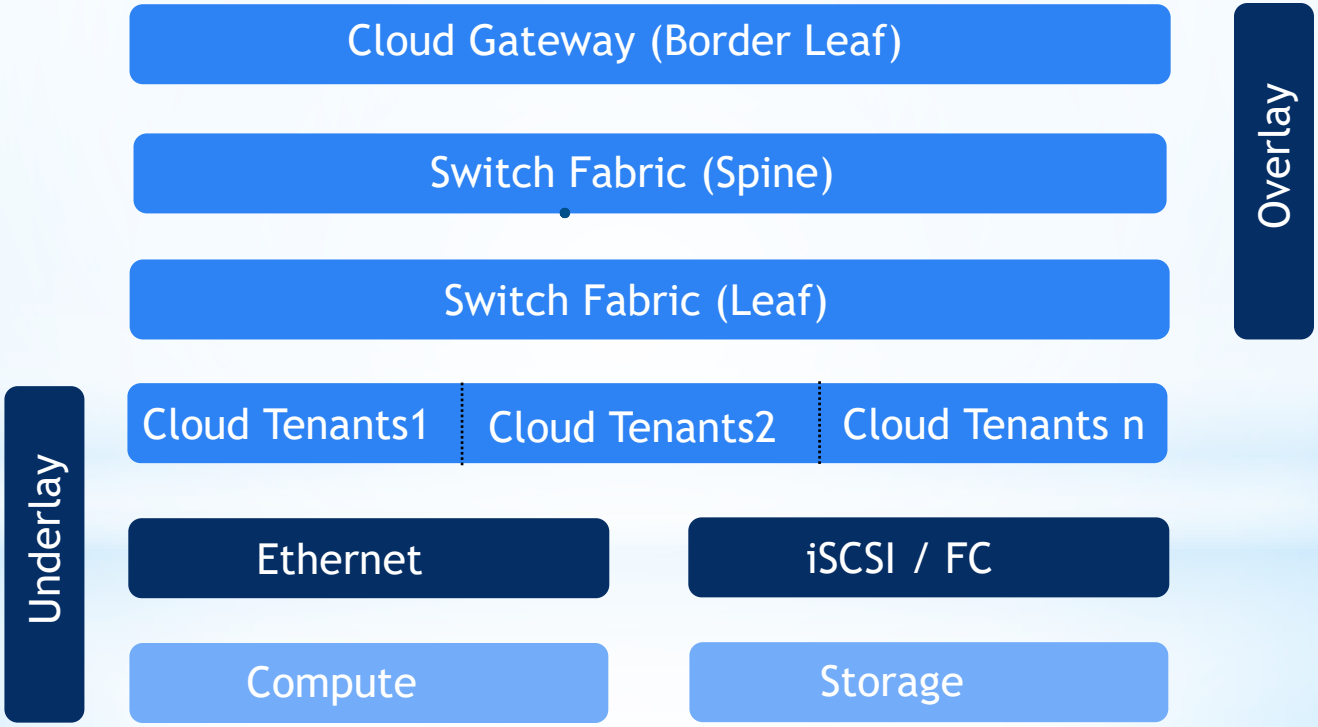
Digital Ocean, Storage, Network, AWS, Rackspace, Google Compute Engine

IaaS

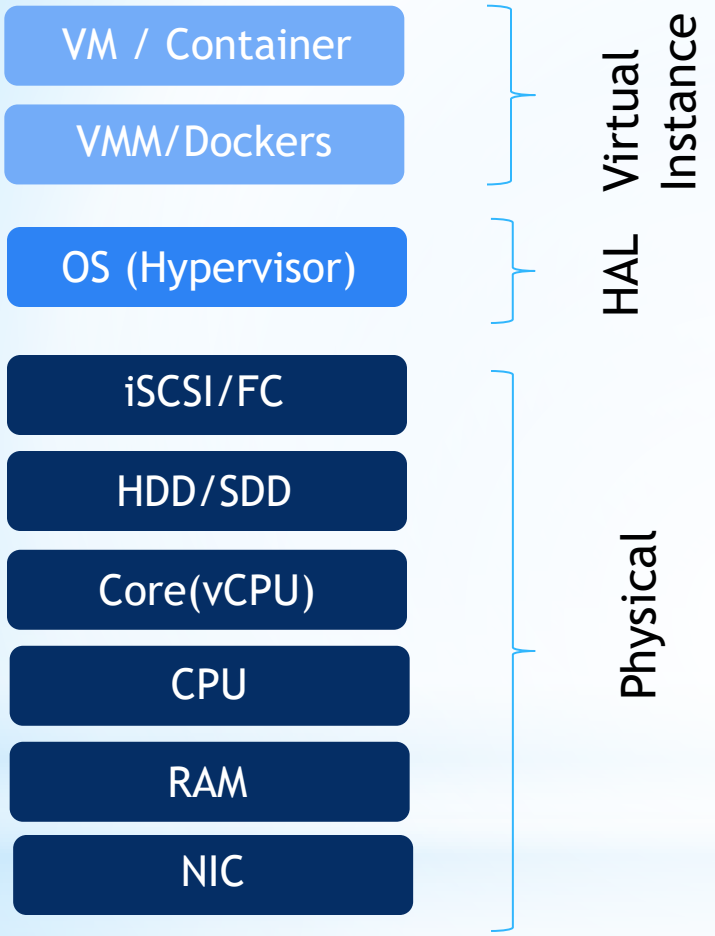
Bare Metal

Networking


IoT/AI/BD
4,5G
Users
Cloud Services


Cloud Gateway (Border Leaf)
Switch Fabric (Spine)
Switch Fabric (Leaf)
Cloud Tenants1
Cloud Tenants2
Cloud Tenants n
Ethernet
iSCSI / FC
Compute
Storage
Overlay
Underlay

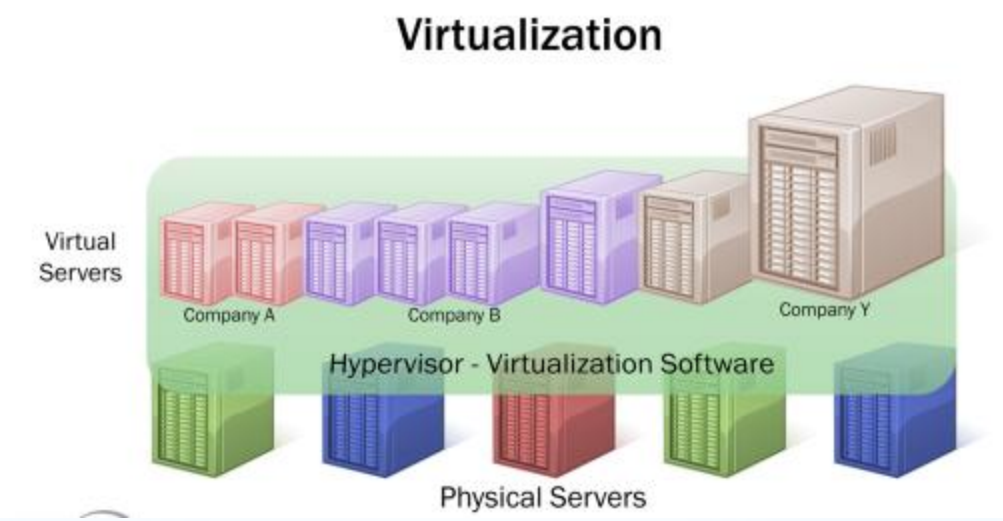
Compute



Minutes: Minutes to have subscription ready before to use business required software

Hours: Minutes to setup and deploy environment focus on your apps and their data. Setup environments and depoy apps very quickly

Days: Minutes to provision VM but time requires to configure software, application, varies, Maintenance/upgrade of OS, Middleware, run, etc.



Virtualization is a key concept in cloud computing that enables resource to truly be pooled. “A virtual computer is a logical representation of a computer in software. By decoupling the physical hardware from the operating system, virtualization provides more operational flexibility and increases the utilization rate of the underlying physical hardware.” (IBM: Virtualization in Education, October 2007.)

Hypervisor type 1

Hypervisor Host Operating System (Hardware Abstraction Layer)

Type 1: Bare Metal, Type 2: Hosted

Hypervisor	Type 1 / 2	Vendor	Description
ESXi	1	VMware	Enterprise grade
XEN / XCP-ng	1	AWS	Used by AWS before Nitro
Nitro	1	AWS	KVM Based
KVM	1	Linux	Opensource
Nutanix-AH	1	Nutanix	Based on XEN Nutanix HCI
Oracle VM Server	1	Oracle	Oracle Infrastructure
HyperV	1 / 2	Microsoft	MS Server, Azure

Hypervisor type 2

Guest Operating System / Type 2 (on top of OS) Hypervisor

- Virtual Machines
- Containers

Hypervisor	Type 2	Vendor	Description
Workstation/Fusion	2	VMware	Workstation for Win/Linux Fusion for MacOS
Oracle Virtual box	2	Oracle	Free, Opensource
QEMU	2	Open Source	Emulator, Often used with KVM
Parallel Desktop	2	Parallel	Opensource
HyperV	2	Microsoft	MS Client

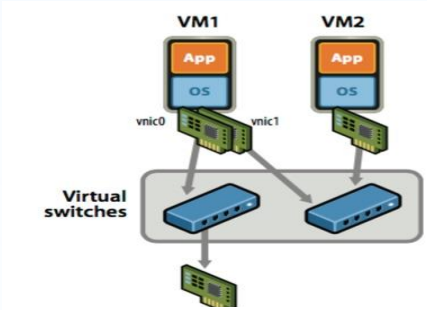
Compute Components

Network Interface Aggregation

- Link Bundling

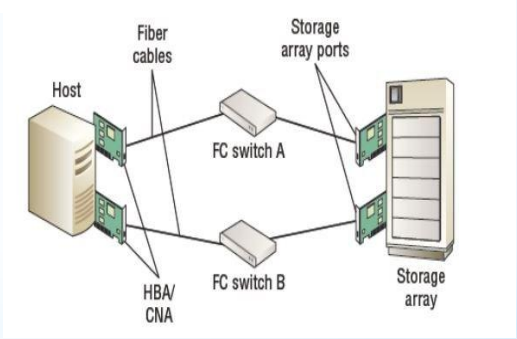
Virtual Network Interface and Switch

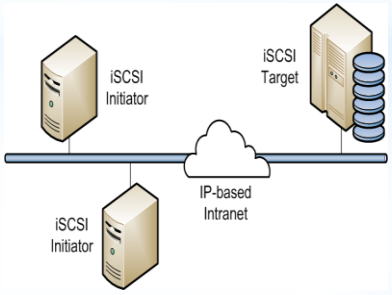
- Link Bundling
- vNIC
- vSwitch
- OpenSwitch

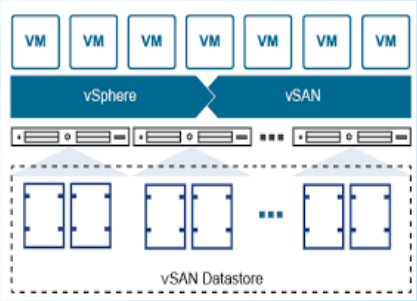


Interface, Devices, & Protocol for Storage

- Storage Area Network (SAN)
- Network Attached Server (NAS)
- virtual SAN
- Fiber Channel
- iSCSI







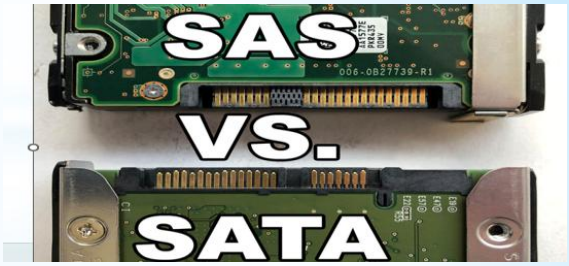
On-Prem Storage

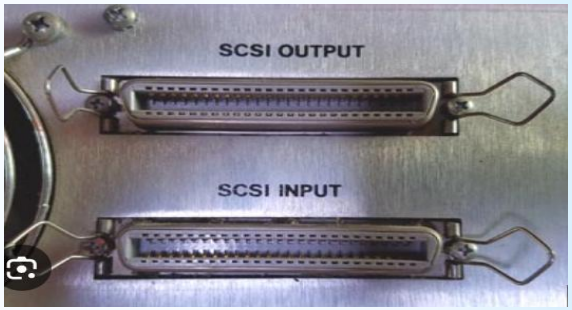
Physical Interface

- **SATA,** Serial ATA, Half Duplex
  from 600MB/s to 6Gbps

- **SAS,** Serial Attached SCSI,
  up to 12Gbps (SAS3) to 22.5Gbps (SAS4)

- **SCSI**, Small Computer System Interface

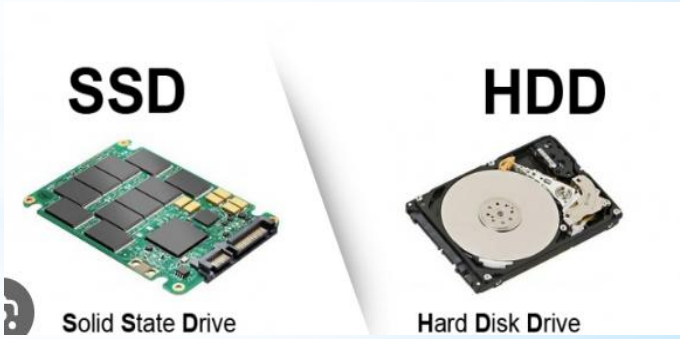




Storage Media

- **SSD**, Solid State Device, low latency,
  Higher IOPS up to 100000+

- **HDD**, Hard Disk Drive, High latency,
  for bulky use, 100-200 IOPS



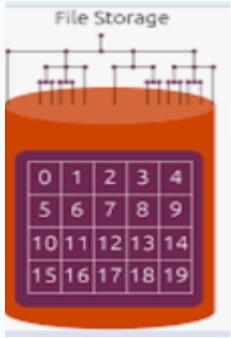
Cloud Storage

Block Storage

- Use for DB VM & have low latency (High IOPS)
- Higher cost per GB
- Data structure is Blocks (no hierarchy)
- Access Raw disk via block interface
- AWS EBS is an example

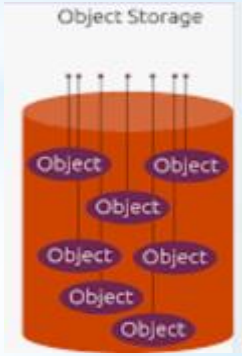
File Storage

- Use for Shared directories
- Moderate cost per GB and moderate latency
- Data structure is File and Folder hierarchy
- Access Shared filesystem vis NFS/SMB
- AWS EFS is an example



Object Storage

- Use for backup, content, unstructured data
- Cost effective, high latency, good throughput
- Data structure Flat name space of objects with IDs
- Access REST API access to objects
- AWS S3 is example



Compute Components

Central Processing Unit (CPU)

- **INTEL**
  - Offers VT-x CPU virtualization, VT-d for I/O virtualization
- **AMD**
  - Offers AMD-V, AMD-vi
- **Processor**
  - Single and Dual Processor socket on mother board
- **Core**
  - Physical processing unit on CPU chip with own L1/L2 cache
  - Multiple cores allow parallel processing
- **vCPU**
  - Software defined unit that a hypervisor assign to VM, single
  - Use single thread of execution with that VM
- **Hyperthreading (Intel) or SMT (AMD)**
  - Allows a single physical core to expose two logical thread to the OS
  - 1 vCPU = 1 Logical thread
  - Example: 4 physical cores with Hyper threading = 8 logical threads = up to 8vCPU assignable

Q & A

Thanks.