CLOUD COMPUTING

by:
DR. BELA SHRIMALI
COMPUTER ACIENCE AND ENGINEERING DEPARTMENT,
NIRMA UNIVERSITY

Syllabus of UNIT-1

UNIT-1

Introduction to cloud and virtualization:

→Cloud Computing in a Nutshell, Layers and Types of Clouds

→ Desired Formats of Cloud

→Cloud Infrastructure Management, Challenges and Risks.

Cloud computing Principles and Paradigms :
By Rajkumar Buyya, James Broberge, Andrzej Goscinski

Introduction

- The ACM *Computing Curricula 2005* defined "computing" as

"In a general way, we can define computing to mean any goal-oriented activity requiring, benefiting from, or creating computers. Thus, computing includes designing and building hardware and software systems for a wide range of purposes; processing, structuring, and managing various kinds of information; doing scientific studies using computers; making computer systems behave intelligently; creating and using communications and entertainment media; finding and gathering information relevant to any particular purpose, and so on. The list is virtually endless, and the possibilities are vast."

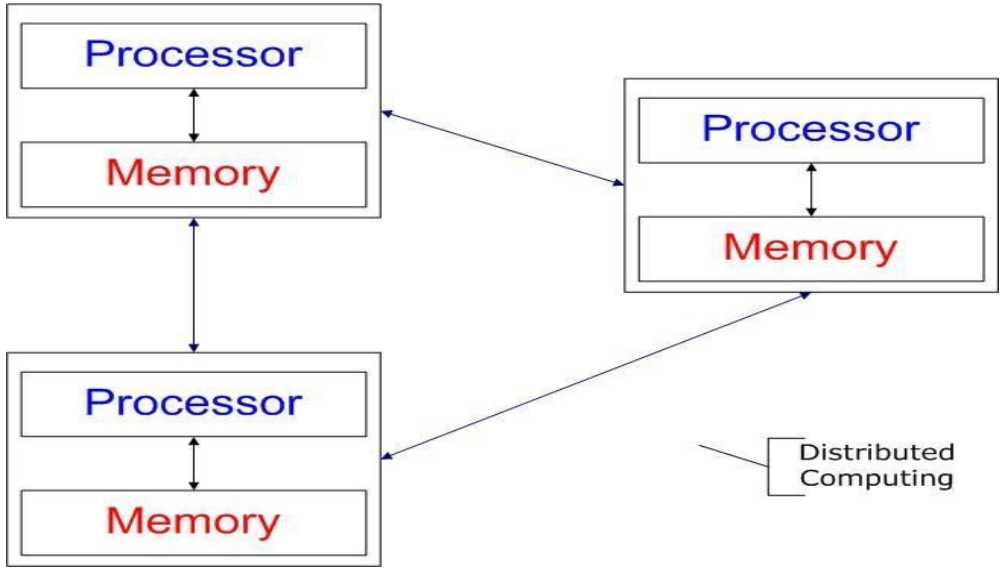
Trends in Computing

- Distributed Computing
- Grid Computing
- Cluster Computing
- Utility Computing
- Cloud Computing

Distributed Computing/System?

- Distributed computing
  - Field of computing science that studies distributed system.
  - Use of distributed systems to solve computational problems.
- Distributed system
  - Wikipedia
    - There are several autonomous computational entities, each of which has its own local memory.
    - The entities communicate with each other by message passing.
  - Operating System Concept
    - The processors communicate with one another through various communication lines, such as high-speed buses or telephone lines.
    - Each processor has its own local memory.

Example?



Grid Computing?

- Pcwebopedia.com
  - A form of networking. unlike conventional networks that focus on communication among devices, grid computing harnesses unused processing cycles of all computers in a network for solving problems too intensive for any stand-alone machine.
- IBM
  - Grid computing enables the virtualization of distributed computing and data resources such as processing, network bandwidth and storage capacity to create a single system image, granting users and applications seamless access to vast IT capabilities. Just as an Internet user views a unified instance of content via the Web, a grid user essentially sees a single, large virtual computer.
- Sun Microsystems
  - Grid Computing is a computing infrastructure that provides dependable, consistent, pervasive(present everywhere) and inexpensive access to computational capabilities

Electrical Power Grid Analogy

Electrical Power Grid

- Users (or electrical appliances) get access to electricity through wall sockets with no care or consideration for where or how the electricity is actually generated.
- **“The power grid”** links together power plants of many different kinds

Grid

- Users (or client applications) gain access to computing resources (processors, storage, data, applications, and so on) as needed with little or no knowledge of where those resources are located or what the underlying technologies, hardware, operating system, and so on are
- **“The Grid"** links together computing resources (PCs, workstations, servers, storage elements) and provides the mechanism needed to access them.

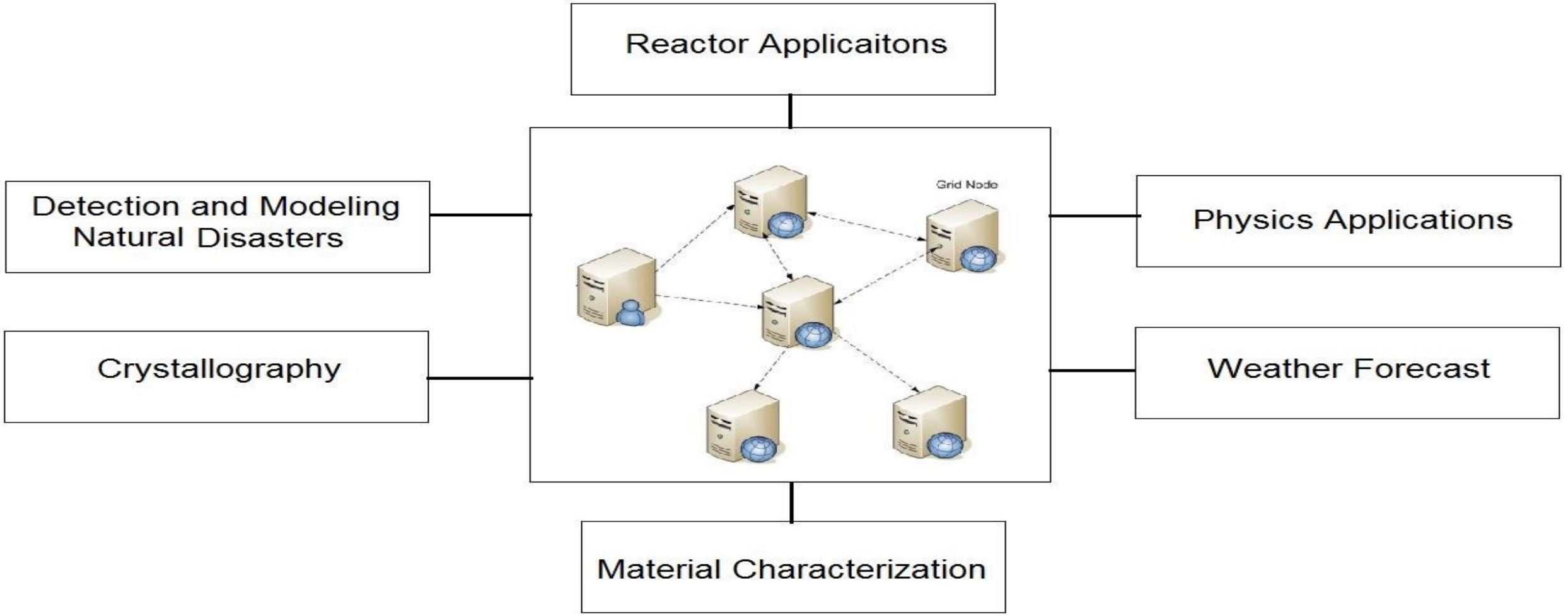
Grid Computing

1. Share more than information: Data, computing power, applications in dynamic environment, multi-institutional, virtual organizations
2. Efficient use of resources at many institutes. People from many institutions working to solve a common problem (virtual organisation).
3. Join local communities.
4. Interactions with the underneath layers must be transparent and seamless to the user.

Need of Grid Computing?

- Today's Science/Research is based on computations, data analysis, data visualization & collaborations
- Computer Simulations & Modelling are more cost effective than experimental methods
- Scientific and Engineering problems are becoming more complex & users need more accurate, precise solutions to their problems in shortest possible time
- Data Visualization is becoming very important
- Exploiting under utilized resources

Who uses Grid Computing ?



What is Cluster Computing?

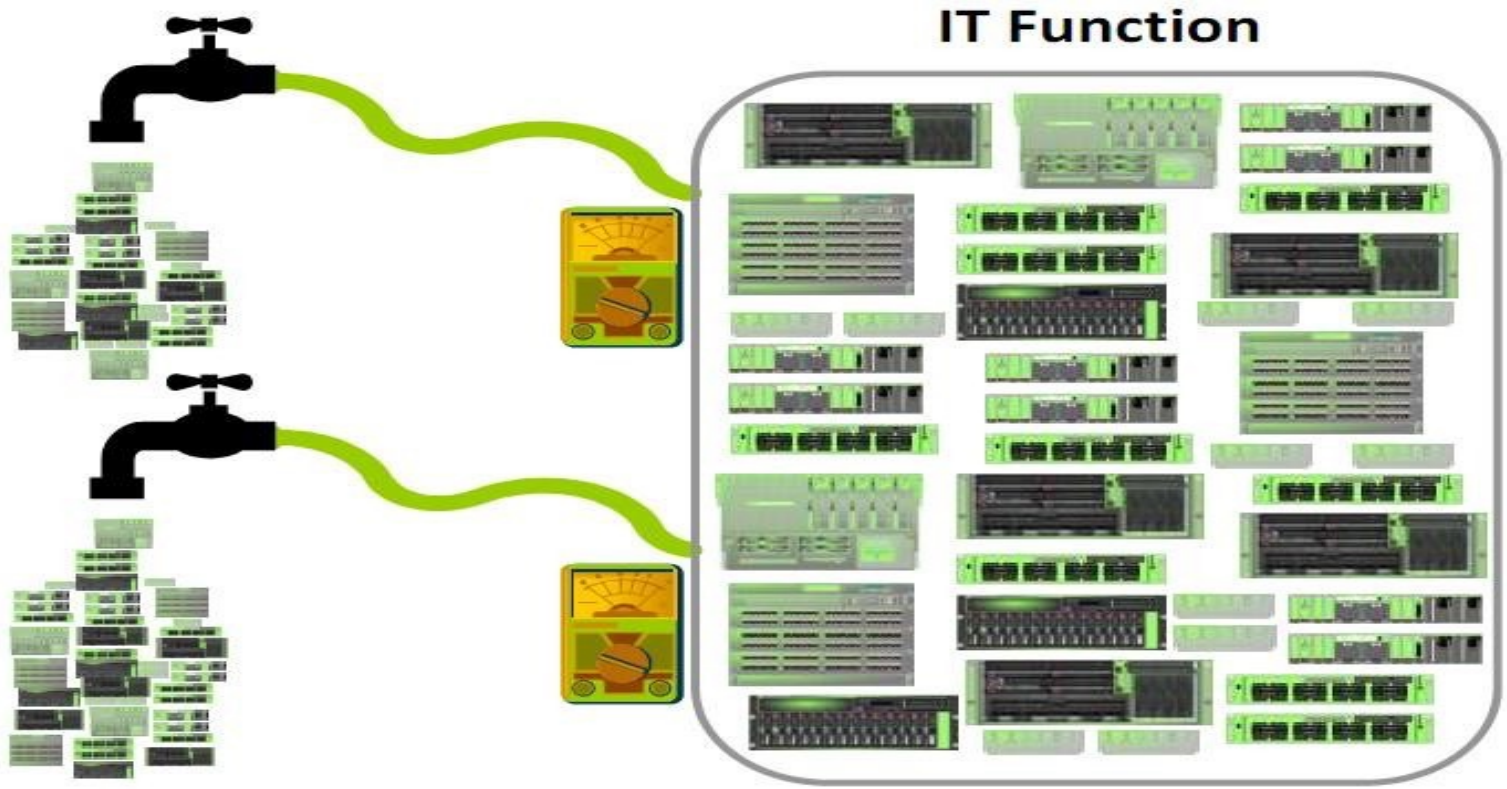
- A cluster is a type of parallel or distributed computer system, which consists of a collection of inter-connected stand-alone computers working together as a single integrated computing resource .
- Key components of a cluster include multiple standalone computers (PCs, Workstations, or SMPs), operating systems, high-performance interconnects, middleware, parallel programming environments, and applications.

“Utility” Computing ?

- Utility Computing is purely a concept which cloud computing practically implements.
- Utility computing is a service provisioning model in which a service provider makes computing resources and infrastructure management available to the customer as needed, and charges them for specific usage rather than a flat rate.
- This model has the advantage of a low or no initial cost to acquire computer resources; instead, computational resources are essentially rented.
- The word *utility* is used to make an analogy to other services, such as electrical power, that seek to meet fluctuating customer needs, and charge for the resources based on usage rather than on a flat-rate basis. This approach, sometimes known as *pay-per-use.*
- *Other example?*

“Utility” Computing ?

- "Utility computing" has usually envisioned some form of virtualization so that the amount of storage or computing power available is considerably larger than that of a single time-sharing computer.



Source: Alan McSweeney, “The Economics of Utility Computing”

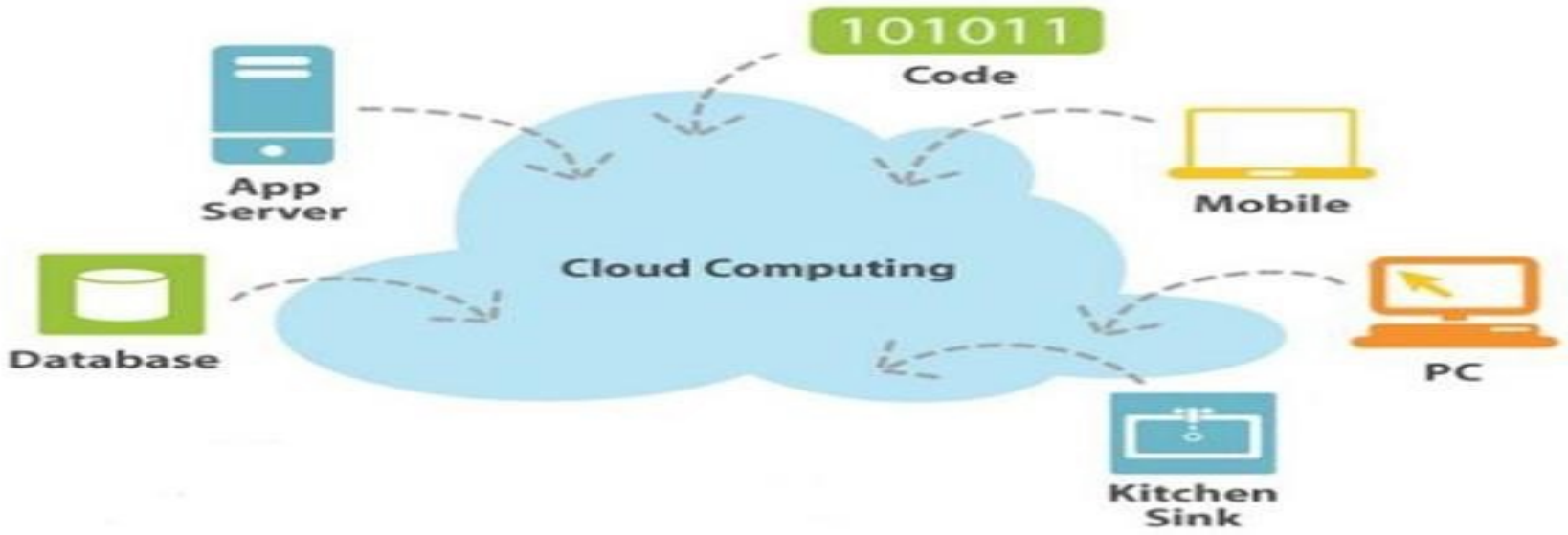
“Utility” Computing ?

a) Pay-for-use Pricing Business Model
b) Data Center Virtualization and Provisioning
c) Solves Resource Utilization Problem
d) Outsourcing
e) Web Services Delivery
f) Automation

Cloud Computing

US National Institute of Standards and Technology defines Computing as

" 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. "



Source: http://www.smallbiztechnology.com/archive/2011/09/wait-what-is-cloud-computing.htm30l/

Cloud Definition

It denotes a model on which a computing infrastructure is **viewed as a "cloud," from which businesses and individuals access applications from anywhere in the world on demand .**

By Rajkumar Buyya:-

"Cloud is a parallel and distributed computing system consisting of a collection of inter-connected and virtualised computers that are dynamically provisioned and presented as one or more unified computing resources based on service-level agreements (SLA) established through negotiation between the service provider and consumers.

Another Definition of Cloud Computing

A report from the University of California Berkeley summarized the key characteristics of cloud computing as:

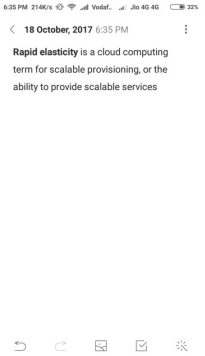
(1) The illusion of infinite computing resources;

(2) The elimination of an up-front commitment by cloud users; and

(3) The ability to pay for use ...as needed...

Essential Characteristics

- **On-demand self-service**
  - A consumer can unilaterally provision computing capabilities, such as server time and network storage, 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.



Cloud Characteristics

- **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.
- **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.

Common Characteristics

- Massive Scale
- Resilient Computing
- Homogeneity
- Geographic Distribution
- Virtualization
- Service Orientation
- Low Cost Software
- Advanced Security

Advantages of Cloud Computing

- **Lower computing costs**:
  - No need of a high-powered and high-priced computer to run cloud computing's web-based applications.
  - Since applications run in the cloud, not on the desktop PC, your desktop PC does not need the processing power or hard disk space demanded by traditional desktop software.
  - When you are using web-based applications, your PC can be less expensive, with a smaller hard disk, less memory, more efficient processor...
  - In fact, your PC in this scenario does not even need a CD or DVD drive, as no software programs have to be loaded and no document files need to be saved.

Advantages of Cloud Computing

- **Improved performance:**
  - With few large programs hogging your computer's memory, you will see better performance from your PC.
  - Computers in a cloud computing system boot and run faster because they have fewer programs and processes loaded into memory.
- **Reduced software costs:**
  - Instead of purchasing expensive software applications, you can get most of what you need for free.
    - most cloud computing applications today, such as the Google Docs suite.
  - better than paying for similar commercial software
    - which alone may be justification for switching to cloud applications.

Advantages of Cloud Computing

- **Instant software updates**
  - Another advantage to cloud computing is that you are no longer faced with choosing between obsolete software and high upgrade costs.
  - When the application is web-based, updates happen automatically available the next time you log into the cloud.
  - When you access a web-based application, you get the latest version without needing to pay for or download an upgrade.

- **Improved document format compatibility.**
  - You do not have to worry about the documents you create on your machine being compatible with other users' applications or OS.
  - There are less format incompatibilities when everyone is sharing documents and applications in the cloud.

Advantages of Cloud Computing

- **Unlimited storage capacity**
  - Cloud computing offers virtually limitless storage.
  - Your computer's current 1 Tera Bytes hard drive is small compared to the hundreds of Peta Bytes available in the cloud.
- **Increased data reliability**
  - Unlike desktop computing, in which if a hard disk crashes and destroy all your valuable data, a computer crashing in the cloud should not affect the storage of your data.
    - if your personal computer crashes, all your data is still out there in the cloud, still accessible
  - In a world where few individual desktop PC users back up their data on a regular basis, cloud computing is a data-safe computing platform. For e.g. Dropbox, Skydrive

Advantages of Cloud Computing

  - That is not a problem with cloud computing, because you do not take your documents with you.
  - Instead, they stay in the cloud, and you can access them whenever you have a computer and an Internet connection
  - Documents are instantly available from wherever you are.
- **Latest version availability**
  - When you edit a document at home, that edited version is what you see when you access the document at work.
  - The cloud always hosts the latest version of your documents as long as you are connected, you are not in danger of having an outdated version.

Advantages of Cloud Computing

- **Easier group collaboration**
  - Sharing documents leads directly to better collaboration.
  - Many users do this as it is an important advantages of cloud computing
    multiple users can collaborate easily on documents and projects
- **Device independence**
  - You are no longer tethered to a single computer or network.
  - Changes to computers, applications and documents follow you through the cloud.
  - Move to a portable device, and your applications and documents are still available.

Disadvantages of Cloud Computing

- **Requires a constant internet connection**
  - Cloud computing is impossible if you cannot connect to the Internet.
  - Since you use the Internet to connect to both your applications and documents, if you do not have an Internet connection you cannot access anything, even your own documents.
  - A dead Internet connection means no work and in areas where Internet connections are few or inherently unreliable, this could be a deal-breaker.
- **Does not work well with low-speed connections**
  - Similarly, a low-speed Internet connection, such as that found with dial-up services, makes cloud computing painful at best and often impossible.
  - Web-based applications require a lot of bandwidth to download, as do large documents.

Disadvantages of Cloud Computing

- **Features might be limited**
  - This situation is bound to change, but today many web-based applications simply are not as full-featured as their desktop-based applications.
    - For example, you can do a lot more with Microsoft PowerPoint than with Google Presentation's web-based offering
- **Can be slow**
  - Even with a fast connection, web-based applications can sometimes be slower than accessing a similar software program on your desktop PC.
  - Everything about the program, from the interface to the current document, has to be sent back and forth from your computer to the computers in the cloud.
  - If the cloud servers happen to be backed up at that moment, or if the Internet is having a slow day, you would not get the instantaneous access you might expect from desktop applications.

Disadvantages of Cloud Computing

- **Stored data might not be secured**
  - With cloud computing, all your data is stored on the cloud.
    - The questions is How secure is the cloud?
  - Can unauthorized users gain access to your confidential data ?

- **Stored data can be lost!**
  - Theoretically, data stored in the cloud is safe, replicated across multiple machines.
  - But on the off chance that your data goes missing, you have no physical or local backup.
    - Put simply, relying on the cloud puts you at risk if the cloud lets you down.

Disadvantages of Cloud Computing

- Not clear that you can run compute-intensive HPC applications that use MPI/OpenMP!
- Scheduling is important with this type of application
  - as you want all the VM to be co-located to minimize communication latency!

- **General Concerns**
  - Each cloud systems uses different protocols and different APIs
    - may not be possible to run applications between cloud based systems
  - Amazon has created its own DB system (not SQL 92), and workflow system (many popular workflow systems out there)
    - so your normal applications will have to be adapted to execute on these platforms.

ROOTS OF CLOUD

COMPUTING[https://www.researchgate.net/publication/282294494_Cloud_Computing_Based_e-Learning_Opportunities_and_Challenges_for_Tertiary_Institutions_in_Nigeria/figures?lo=1&utm_source=google&utm_medium=organic

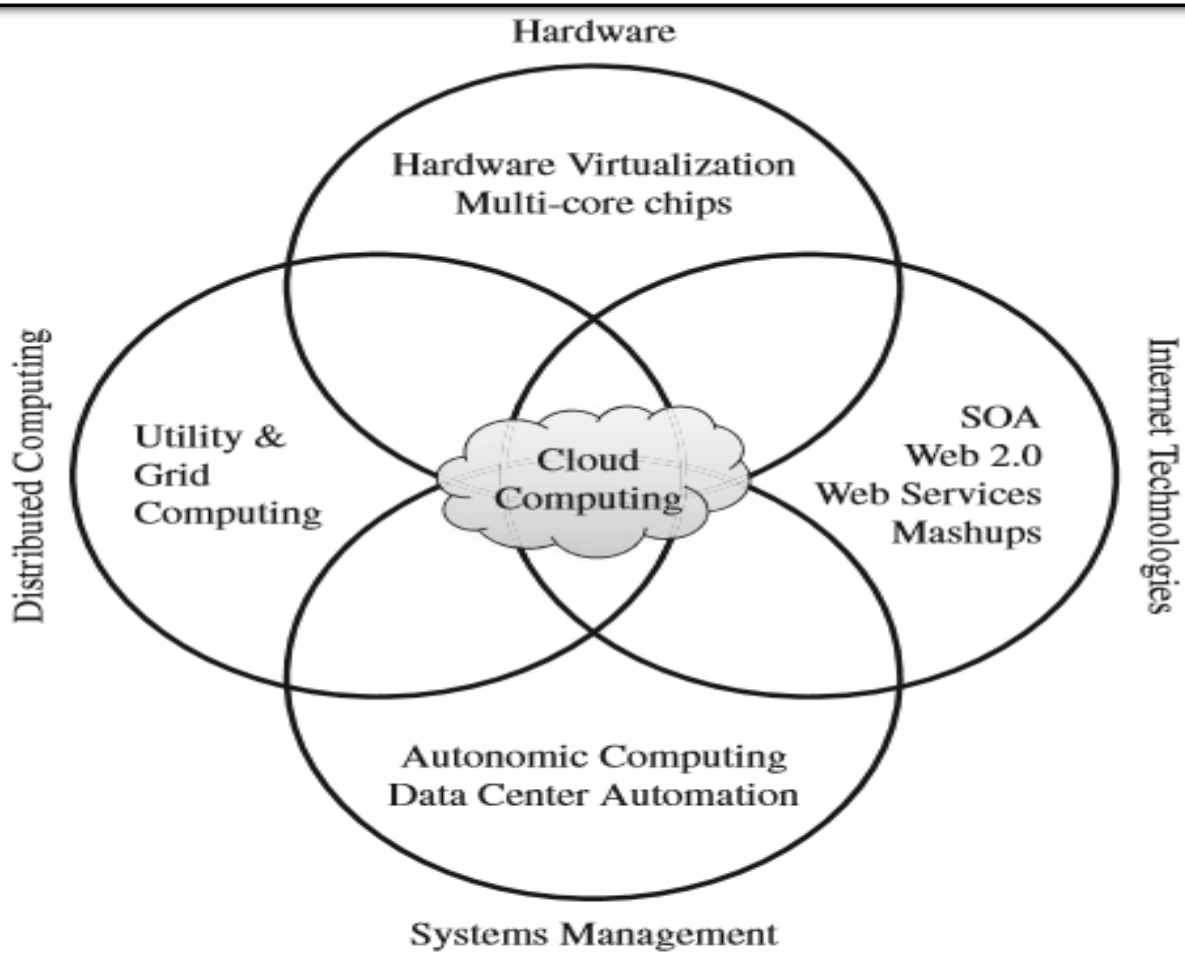


FIGURE 1.1. Convergence of various advances leading to the advent of cloud computing.

SOA, Web Services, Web 2.0, and Mashups

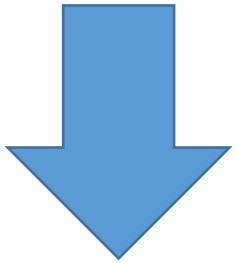
- Part of the Software applications glossary: A **mash-up** (sometimes spelled as one word, **mashup**) is a Web page or application that integrates complementary elements from two or more sources. **Mash-ups** are often **defined** by the type of content that they aggregate.(it's an application!!!!!!!)

==========

- A **service-oriented architecture** (**SOA**) is a style of software design where services are provided to the other components by application components, through a communication protocol over a network.(it's a service!!!!!!!!)

Hardware virtualization

- Cloud computing services are usually backed by large-scale data centers composed of thousands of computers.



Hardware Virtualization can be considered as a perfect fit to overcome most Operational Issues of data centre building and maintenance.

Virtual Machines [https://www.dataveneta.it/en/products/virtual-machines-player]

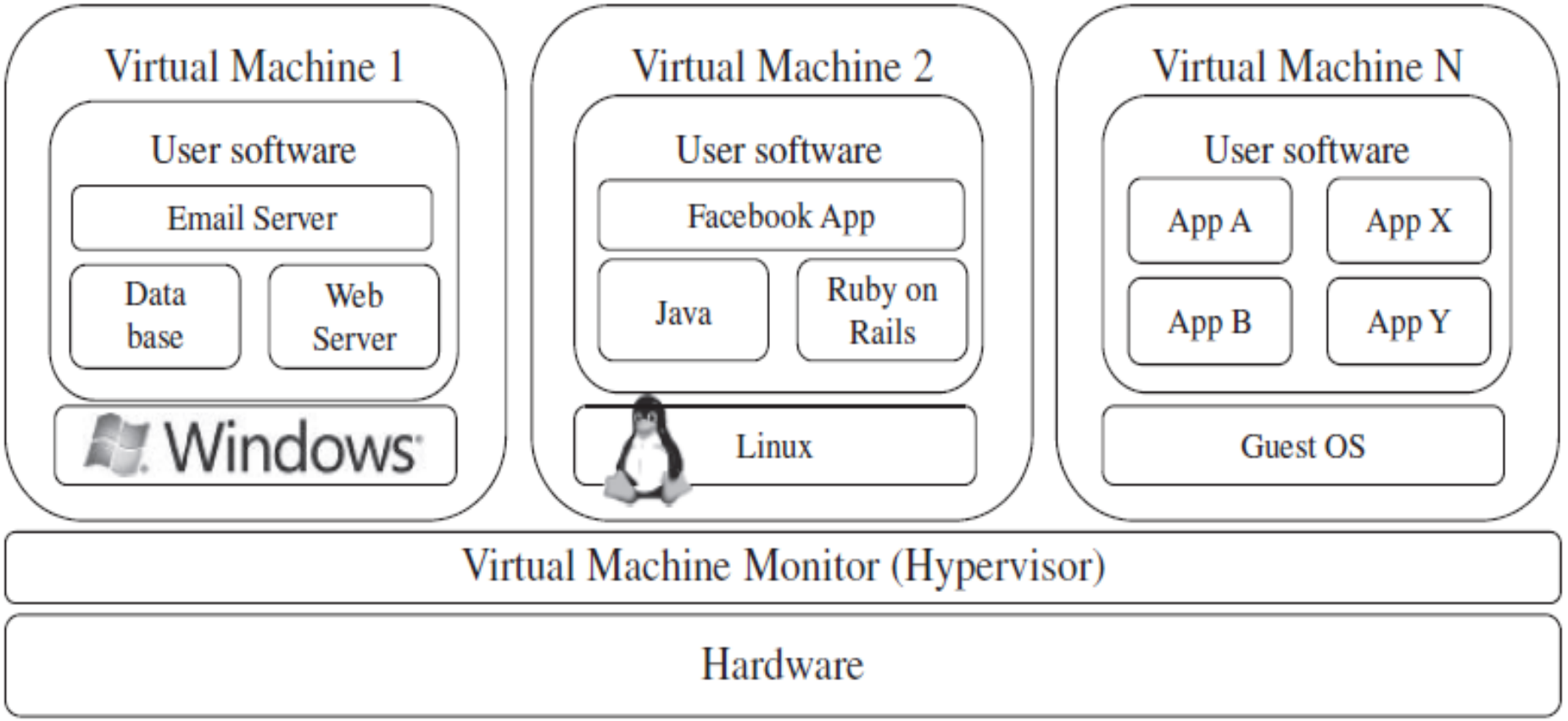


FIGURE 1.2. A hardware virtualized server hosting three virtual machines, each one running distinct operating system and user level software stack.

Hardware Virtualization……continued….

- Workload isolation is achieved since all program instructions are fully confined inside a VM, which leads to improvements in security. Better reliability is also achieved because software failures inside one VM do not affect others.
- Workload migration, also referred to as application mobility , targets at facilitating hardware maintenance, load balancing, and disaster recovery
- A number of **VMM (Virtual Machine Monitor)** platforms exist that are the basis of many utility or cloud computing environments. The most notable ones, **VMWare, Xen, and KVM**, are outlined in the following sections.

- Xen. The Xen hypervisor started as an open-source project and has served as a base to other virtualization products, both commercial and open-source. It has pioneered the para-virtualization concept, on which the guest operating system, by means of a specialized kernel, can interact with the hypervisor, thus significantly improving performance. In addition to an open-source distribution, Xen currently forms the base of commercial hypervisors of a number of vendors, most notably Citrix XenServer and Oracle VM.

- KVM. The Kernel-Based Virtual machine (KVM) is a Linux virtualization subsystem. Is has been part of the mainline Linux kernel since version , thus being natively supported by several distributions. In addition, activities such as *memory management and scheduling are carried out by existing kernel features, thus making KVM simpler and smaller than hypervisors that take control of the entire machine.*

Autonomic Computing

https://www.slideshare.net/sandpoonia/9-the-semantic-grid-and-autonomic-grid

LAYERS AND TYPES OF CLOUDS

https://www.researchgate.net/publication/273897590_The_Challenges_of_Cloud_Computing_Management_Information_System_in_Academic_Work/figures?lo=1

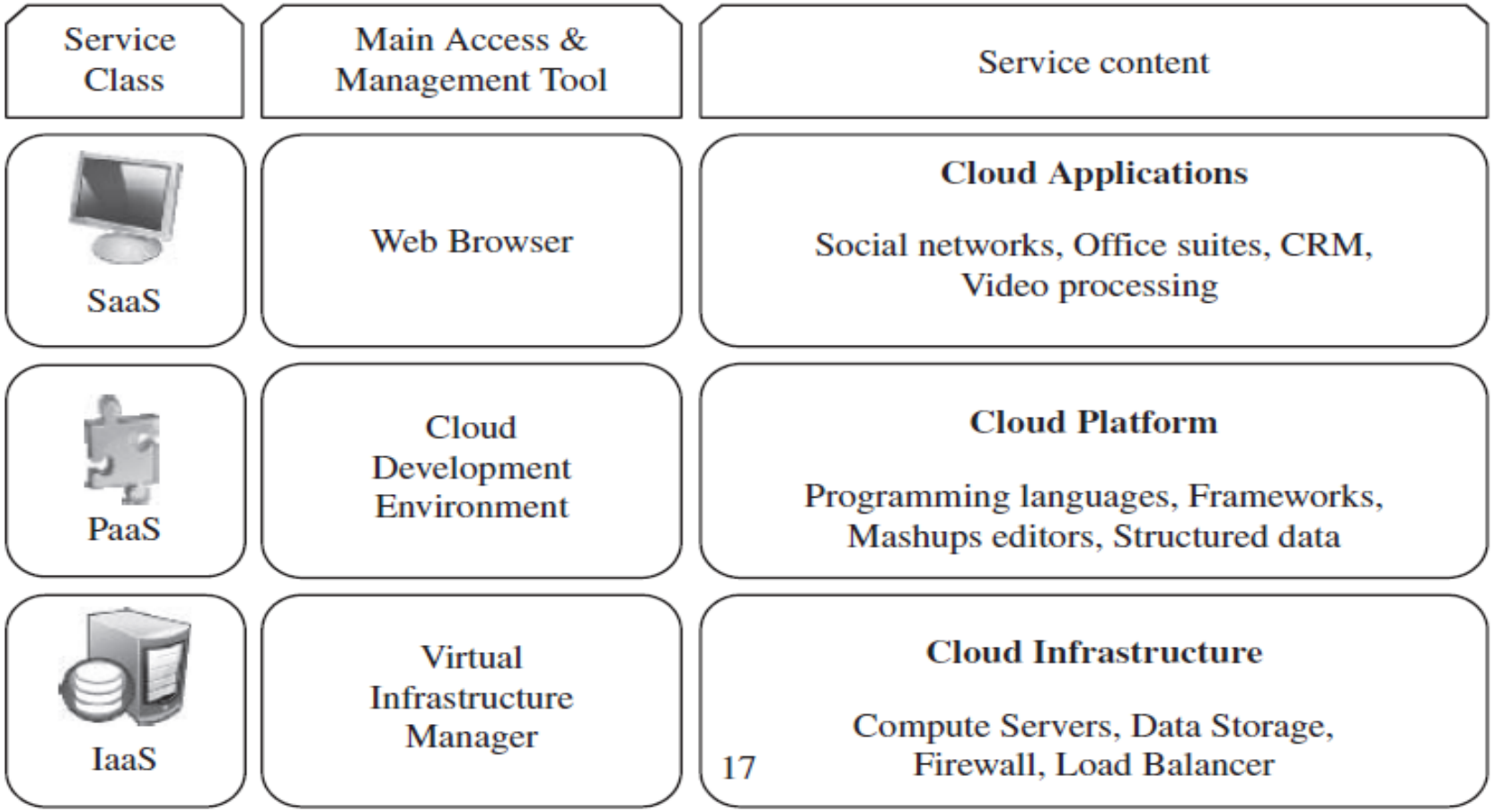


FIGURE 1.3. The cloud computing stack.

Infrastructure as a Service(IaaS)

- Offering virtualized resources (**computation, storage, and communication**) on demand is known as Infrastructure as a Service (IaaS).
- A cloud infrastructure enables **on-demand provisioning of servers running several choices of operating systems and a customized software stack**. Infrastructure services are considered to be the bottom layer of cloud computing systems .
- Amazon Web Services mainly offers IaaS, which in the case of its **EC2 (Elastic Cloud Compute) service means offering VMs with a software stack that can be customized similar to how an ordinary physical server would be customized**.
- Users are given privileges to perform numerous activities to the server, such as: **starting and stopping it**, **customizing it by installing software packages**, **attaching virtual disks to it**, and configuring access permissions and firewalls rules etc.

Platform as a Service (PaaS)

- A cloud platform offers an environment on which *developers create and deploy applications and do not necessarily need to know how many processors or how much memory that applications will be using. In addition, multiple programming models and specialized services (e.g., data access, authentication, and payments) are offered as building blocks to new applications*.
- ***Google AppEngine***, an example of Platform as a Service, offers a scalable environment for developing and hosting Web applications, which should be written in specific programming languages such as Python or Java, and use the services' own proprietary structured object data store.

Software as a Service (SaaS)

- Applications reside on the top of the cloud stack. Services provided by this layer can be accessed by end users through Web portals. Therefore, consumers are increasingly shifting from **locally installed computer programs to on-line software services that offer the same functionally**. Traditional desktop applications such as word processing and spreadsheet can now be accessed as a service in the Web. This model of delivering applications, known as Software as a Service (SaaS), alleviates the burden of software maintenance for customers and simplifies development and testing for providers .
- *Example:*- Salesforce.com, which relies on the SaaS model, offers business productivity applications (CRM) that reside completely on their servers, allowing costomers to customize and access applications on demand.

Deployment Model: Cloud Computing

https://www.javatpoint.com/types-of-cloud

Public + Private + Community + Hybrid

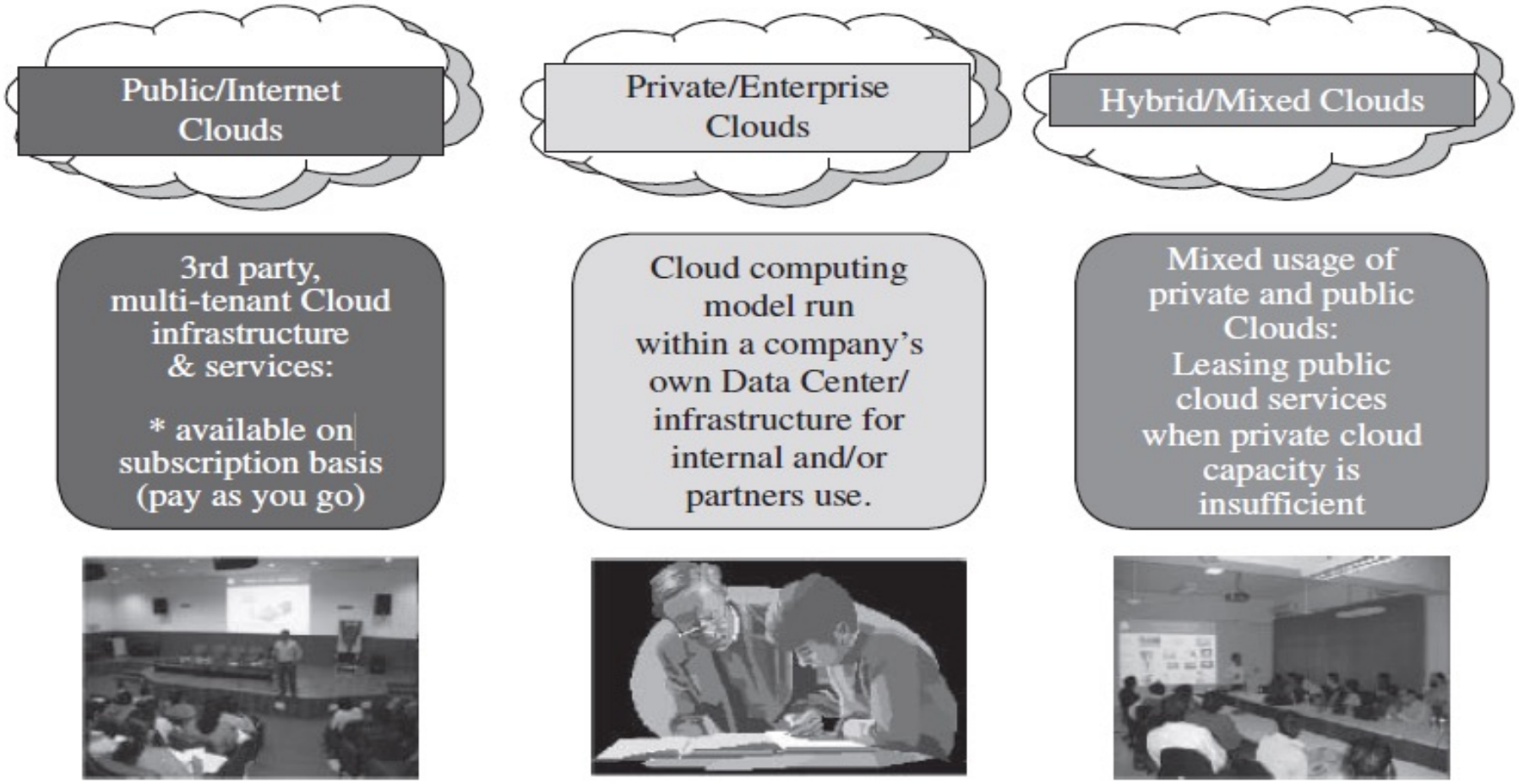


FIGURE 1.4. Types of clouds based on deployment models.

Deployment Model…….continued

- Public cloud as a “cloud made available in a pay-as-you-go manner to the general public” and private cloud as “internal data center of a business or other organization, not made available to the general public.”
- A community cloud is “shared by several organizations and supports a specific community that has shared concerns (e.g., mission, security requirements, policy, and compliance considerations) .”
- A hybrid cloud takes shape when a private cloud is supplemented with computing capacity from public clouds . The approach of temporarily renting capacity to handle spikes in load is known as “cloud-bursting”

Public Cloud

The public cloud is the one in which cloud infrastucture services are provided to general public or large industry group over internet. In this cloud model, the infrastructure is not owned by user but by the organization which provides the cloud services.

The storage backup and retrieval services in this model are provided without any cost or as subscription or based on used basis.

Example of Public cloud:

•Amazon elastic compute cloud (EC2) •IBM SmartCloud Enterprise•Google AppEngine •Windows Azure Services Platform

- Advantages and Disadvantages ??

Private Cloud

https://www.capgemini.com/gb-en/2018/06/how-hybrid-cloud-is-fueling-digital-transformation/

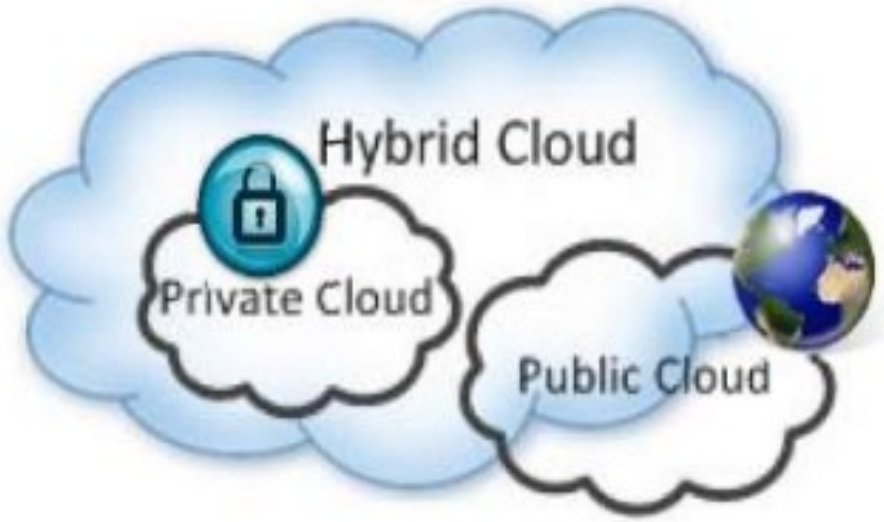
The private cloud is the one in which cloud infrastructure is set aside for exclusive use by single organization. It is owned, managed and operated by organization, third party or combination of both. The cloud infrastructure in this model is provisioned on the premises of organization but hosted in data center owned by third party.

In private cloud, organizations will have advantages over public cloud as it provides greater flexibility of control over cloud resources to them. Moreover private cloud is useful in the storage applications where in security, latency and regulatory issues are of utmost concern.

- Advantages and Disadvantages ??

Hybrid Cloud

https://www.capgemini.com/gb-en/2018/06/how-hybrid-cloud-is-fueling-digital-transformation/



As the name suggests, hybrid cloud is the combination of other cloud models viz. public cloud, private cloud or community cloud. This model takes advantages of all the models which are part of it. Hence it will have scalability, cost effectiveness and data security all in one model.

The disadvantage of this model is difficulty in implementing such a storage solution.

- Advantages and Disadvantages ??

Community Cloud

https://www.capgemini.com/gb-en/2018/06/how-hybrid-cloud-is-fueling-digital-transformation/

The model type community cloud shares the cloud infrastructure across several organizations to support specific community having common concerns. In this model, cloud infrastructure is provided on the premises or at the data center owned by third party. This is managed by participating organizations or third party.

Community cloud takes benefits of both public cloud (e.g. minimal shared infrastructure costs, pay per use basis billing) as well as private cloud (e.g. added privacy level, policy compliance) .

Deployment Model: At a glance

Cloud Storage Type	Host	Owner	Access	Users
Public cloud	service provider	service provider	Internet	public as individuals, organizations
Private cloud	Enterprise (Third Party)	Enterprise	Intranet, VPN	Business organizations
Hybrid cloud	Enterprise (Third Party)	Enterprise	Intranet, VPN	Business organizations
Community cloud	Community (Third party)	Community	Intranet, VPN	Community members

Desired Feature of a Cloud

Certain features of a cloud are essential to enable services that truly represent the cloud computing model and satisfy expectations of consumers, and cloud offerings must be :-

(i) self-service,

(ii) per-usage metered and billed,

(iii) elastic,

and (iv) customizable.

Per usages metering and billing

- Cloud computing eliminates up-front commitment by users, allowing them to request and use **only the necessary amount**. Services must be priced on a short term basis (e.g., by the hour), allowing users to release (and not pay for) resources as soon as they are not needed .
- For these reasons, clouds must implement features to allow efficient trading of service such as **pricing, accounting, and billing** .
- Metering should be done accordingly for different types of service (e.g., **storage, processing, and bandwidth**) and usage promptly reported, thus providing greater transparency.

Elasticity

- Cloud computing gives the illusion of infinite computing resources available on demand . Therefore users expect clouds to rapidly provide resources in any quantity at any time. In particular, it is expected that the additional resources can be
- (a) provisioned, possibly automatically, when an **application load increases** and
- (b) released when **load decreases** (scale up and down)

Customization

- In a multi-tenant cloud a great disparity between user needs is often the case. Thus, resources rented from the cloud must be highly customizable.
- In the case of infrastructure services, **customization means allowing users to deploy specialized virtual appliances and to be given privileged (root) access to the virtual servers.**
- **Other service classes (PaaS and SaaS) offer less flexibility and are not suitable for general-purpose computing, but still are expected to provide a certain level of customization.**

Cloud Infrastructure Management:-

1) Challenge IaaS providers face when building a cloud infrastructure is managing physical and virtual resources, namely servers, storage, and networks, in a holistic fashion.

VIM - **Virtual Infrastructure Management-** rapid and dynamic!!!

-> Infrastructure sharing software.

->Virtual infrastructure engine.

Features of VIMs

1) **Virtualization Support**. The multi-tenancy aspect of clouds requires multiple customers with disparate requirements to be served by a single hardware infrastructure. Virtualized resources (CPUs, memory, etc.) can be **sized and resized** with certain flexibility.
2) **Self-Service, On-Demand Resource Provisioning**.

This feature enables users to **directly obtain services from clouds**, such as spawning the creation of a server and tailoring its software, configurations, and security policies, **without interacting with a human system administrator**. This capability "**eliminates the need for more time-consuming, labor-intensive, humandriven procurement processes familiar to many in IT".**

Features of VMs……..continued

3) Multiple Backend Hypervisors. Different virtualization models and tools offer different benefits, drawbacks, and limitations.

This characteristic is more visible in open-source VM managers, which usually provide pluggable drivers to interact with multiple hypervisors . In this direction, the aim of **libvirt** is to provide a uniform API that VM managers can use to manage domains (a VM or container running an instance of an operating system) in virtualized nodes using standard operations that abstract hypervisor specific calls.

Features of VMs……..continued

4) Storage Virtualization. Virtualizing storage means abstracting logical storage from physical storage. By consolidating all available storage devices in a data center , ***it allows creating virtual disks independent from device and location.***

- Storage devices are commonly organized in a storage area network (SAN) and attached to servers via protocols such as Fibre Channel, iSCSI(Internet Small Computer Systems Interface), and NFS; a storage controller provides the layer of abstraction between virtual and physical storage.

Interface to Public Clouds. Researchers have perceived that extending the capacity of a local in-house computing infrastructure by borrowing resources public clouds is advantageous. In this fashion, institutions can make good use of their available resources and, in case of spikes in demand, extra load can be offloaded to rented resources .

A VI manager can be used in a hybrid cloud setup if it offers a driver to manage the life cycle of virtualized resources obtained from external cloud providers. To the applications, the use of leased resources must ideally be transparent.

Virtual Networking. Virtual networks allow creating an isolated network on top of a physical infrastructure independently from physical topology and locations .

A virtual LAN (VLAN) allows isolating traffic that shares a switched network, allowing VMs to be grouped into the same broadcast domain. Additionally, a VLAN can be configured to block traffic originated from VMs from other networks.

Similarly, the VPN (virtual private network) concept is used to describe a secure and private overlay network on top of a public network (most commonly the public Internet).

Dynamic Resource Allocation.

Increased awareness of energy consumption in data centres has encouraged the practice of dynamic consolidating VMs in a fewer number of servers. In cloud infrastructures, where applications have **variable and dynamic needs**, capacity management and demand prediction are especially complicated.

This fact triggers the need for **dynamic resource allocation aiming at obtaining a timely match of supply and demand** .

Energy consumption reduction and better management of SLAs can be achieved by dynamically remapping VMs to physical machines at regular intervals. **Machines that are not assigned any VM can be turned off or put on a low power state**. In the same fashion, **overheating can be avoided by moving load away from hotspots** .

A number of VI (Virtual Infrastructure) managers include a dynamic resource allocation feature that continuously monitors utilization across resource pools and reallocates available resources among VMs according to application needs.

Virtual Clusters. Several *VI managers* can holistically manage groups of VMs. This feature is useful for provisioning computing virtual clusters on demand, and interconnected VMs for multi-tier Internet applications.

Reservation and Negotiation Mechanism. *When users request computational resources to available at a specific time, requests are termed advance reservations (AR).*

This is especially useful in clouds on which resources **are scarce(limited)**; since not all requests may be satisfied immediately, they can benefit of VM placement strategies that support queues, priorities, and advance reservations .

High Availability and Data Recovery. The high availability (HA) feature of VI managers aims at *minimizing application downtime and preventing business disruption*. A few VI managers accomplish this by providing a failover mechanism, which detects failure of both physical and virtual servers and restarts VMs on healthy physical servers. This style of HA protects from host, but not VM, failures.

- For **mission critical (above will not work)**applications, when a failover solution involving restarting VMs does not suffice, additional levels of fault tolerance that rely on redundancy of VMs are implemented. In this style, redundant and synchronized VMs (running or in standby) are kept in a secondary physical server. The HA solution monitors failures of system components such as servers, VMs, disks, and network and ensures that a duplicate VM serves the application in case of failures .

Data Recovery…Continued…..

Data backup in clouds should take into account the high data volume involved in VM management. *Frequent backup of a large number of VMs, each one with multiple virtual disks attached, should be done with* ***minimal interference*** *in the systems performance*.

In this sense, some VI managers offer data protection mechanisms that perform incremental backups of VM images. The backup workload is often assigned to proxies, thus offloading production server and reducing network overhead.

CHALLENGES AND RISKS

- Issues to be faced include user privacy
- Data security, data lock-in
- Availability of service
- Disaster recovery
- Performance
- Scalability
- Energy-efficiency
- Programmability.

Security, Privacy, and Trust

- Security and privacy affect the entire cloud computing stack, since there is a massive use of **third-party services** and infrastructures that are used to host important data or to perform critical operations. In this scenario, the **trust toward providers** is fundamental to ensure the desired level of privacy for applications hosted in the cloud.
- **Legal and regulatory issues** also need attention. When data are moved into the Cloud, providers may choose to locate them anywhere on the planet. The physical location of data centers determines the set of laws that can be applied to the management of data.
- For example, specific cryptography techniques could not be used because they are not allowed in some countries. Similarly, country laws can impose that sensitive data, such as patient health records, are to be stored within national borders.

Data Lock-In and Standardization

- A major concern of cloud computing users is about having their **data locked-in by a certain provider**. Users may want to move data and applications out from a provider that **does not meet their requirements**. However, in their current form, cloud computing infrastructures and platforms do not employ standard methods of storing user data and applications. Consequently, **they do not interoperate and user data are not portable**. The answer to this concern is standardization. In this direction, there are efforts to create **open standards for cloud computing**.
- **The Cloud Computing Interoperability Forum (CCIF) was formed by organizations such as Intel, Sun, and Cisco in order to "enable a global cloud computing ecosystem whereby organizations are able to seamlessly work together for the purposes for wider industry adoption of cloud computing technology." The development of the Unified Cloud Interface (UCI) by CCIF aims at creating a standard programmatic point of access to an entire cloud infrastructure**.
- In the hardware virtualization sphere, the **Open Virtual Format (OVF**) aims at facilitating packing and distribution of software to be run on VMs so that virtual appliances can be made portable—that is, seamlessly run on hypervisor of different vendors.

Availability, Fault-Tolerance, and Disaster Recovery

- It is expected that users will have certain **expectations about the service level to be provided** once their applications are moved to the cloud. These expectations include **availability of the service**, **its overall performance**, and what measures are to be taken **when something goes wrong in the system or its components**.
- In summary, users seek for a warranty before they can comfortably move their business to the cloud.→ ***Trust Management in the Cloud***
- ***SLAs***, which include QoS requirements, must be ideally set up between customers and cloud computing providers to act as warranty. An SLA specifies details of the service to be provided, including availability and performance guarantees. Additionally, *metrics must be agreed upon by all parties, and penalties for violating the expectations must also be approved*

Resource Management and Energy-Efficiency

→ One important challenge faced by providers of cloud computing services is the **efficient management of virtualized resource pools**.

→ Physical resources such as

→ *CPU cores, disk space, and network bandwidth must be sliced and shared among virtual machines running potentially heterogeneous workloads.*

CHALLENGES AND RISKS……continued…..

- The multi-dimensional nature of virtual machines complicates the activity of finding **a good mapping of VMs onto available physical hosts while maximizing user utility**. Dimensions to be considered include: **number of CPUs**, **amount of memory**, **size of virtual disks**, and **network bandwidth**.
- Dynamic VM mapping policies may leverage the *ability to suspend, migrate, and resume VMs as an easy way of pre-empting low-priority allocations in favor of higher-priority ones*.

CHALLENGES AND RISKS……continued…..

Migration of VMs also brings additional challenges such as **detecting when to initiate a migration, which VM to migrate, and where to migrate**.

In this case, an additional concern is the trade-off between the negative impact of a live migration on the performance and stability of a service and the benefits to be achieved with that migration

CHALLENGES AND RISKS……continued…..

- Another challenge concerns the outstanding amount of data to be managed in various VM management activities. Such data amount is a result of particular abilities of virtual machines, including the ability of traveling through space (i.e., migration) and time (i.e., check pointing and rewinding), operations that may be required in load balancing, backup, and recovery scenarios.
- In addition, dynamic provisioning of new VMs and replicating existing VMs require efficient mechanisms to **make VM block storage devices (e.g., image files) quickly available at selected hosts**.

CHALLENGES AND RISKS……continued…..

- Data centers **consumer large amounts of electricity**. According to a data published by HP, 100 server racks can consume 1.3MW of power and another 1.3 MW are required by the cooling system, thus costing USD 2.6 million per year. Besides the monetary cost, data centers significantly impact the environment in terms of CO2 emissions from the cooling systems .
- In addition to optimize application performance, **dynamic resource management can also improve utilization and consequently minimize energy consumption in data centers**. This can be done by judiciously consolidating workload onto smaller number of servers and **turning off idle resources**

- Chapter-I Part-I ends here.
- Next part is dedicated to the Virtualization.