SATHYABAMA
INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)
Accredited "A" Grade by NAAC | 12B Status by UGC | Approved by AICTE
www.sathyabama.ac.in

SCHOOL OF COMPUTING

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

SIT1304 – CLOUD COMPUTING

UNIT – I – Cloud Computing – SIT1304

UNIT 1

INTRODUCTION TO CLOUD COMPUTING

History of Cloud computing - Cloud Computing Architectural Framework - Types of Clouds - pros and cons of cloud computing - difference between web 2.0 and cloud - key challenges in cloud computing - Major Cloud players - Cloud Deployment Models - Virtualization in Cloud Computing - types of virtualization - Parallelization in Cloud Computing - cloud resource management - dynamic resource allocation - Optimal allocation of cloud models.

1.0 Introduction

Cloud computing is the on-demand availability of computer system resources, especially data storage (cloud storage) and computing power, without direct active management by the user. The term is generally used to describe data centers available to many users over the Internet. Large clouds, predominant today, often have functions distributed over multiple locations from central servers. If the connection to the user is relatively close, it may be designated an edge server.

1.0.1 Difference between Cloud Computing and Web 2.0

Cloud Computing	**Web 2.0**
It is more specific and definite	Programming and business models
It is a way of searching through data.	It is sharing entire pieces of data between different websites.
Cloud computing is about computers.	Web 2.0 is about people.
The internet as a computing platform	Attempt to explore and explain the business rules of that platform
Googleapps are considered in Cloud computing.	A web-based application is considered in Web 2.0.
It is a business model for hosting these services.	It is a technology which allows web pages to act as more responsive applications

1.1. Key challenges in cloud computing

- Bandwidth cost
- Cloud computing saves the hardware acquisition costs but their expenditure on bandwidth rises considerably.
- Sufficient bandwidth is required to deliver intensive and complex data over the network.

- Continuous monitoring and supervision
- It is important to monitor the cloud service continuously as well as to supervise its

performance, business dependency and robustness.

- Security concerns.

To prevent cloud infrastructure damages, some of the measures include tracking unusual behaviour across servers, buying security hardware and using security applications.

1.2 Data access and integration

In cloud where the data stored, how to access it, who is the owner and how to control it. Companies are often concerned about data ownership and loss of data control while moving to cloud.

The integration of existing applications in the cloud for smooth running is another challenge.

1.2.1 Proper usability

Enterprises need to have a good and clear view of how to use the technology to add value to their unique businesses

1.2.2 Migration issues

Migrating data from system to the cloud can pose major risks, if it not handled properly.

It need to develop migration strategy that integrates well with the current IT infrastructure.

1.2.3 Cost assessment

Scalable and on-demand nature of cloud services makes the assessment of cost difficult.

Heavy use of a service for a few days may consume the budget of several months.

Current Cloud Computing Challenges

1.3 Requirement of Cloud

In January 2016, RightScale conducted its fifth annual State of the Cloud Survey on the latest cloud computing trends. They questioned 1,060 technical professionals across a broad cross-section of organizations about their adoption of cloud infrastructure.

1.3.1 Lack of resources/expertise

The organizations are increasingly placing more workloads in the cloud while cloud technologies continue to rapidly advance.

Due to these factors organizations are having a hard time keeping up with the tools. Also, the need for expertise continues to grow.

These challenges can be minimized through additional training of IT and development staff.

1.3.2. Security issues

The start of cloud computing technology: you are unable to see the exact location where your data is stored or being processed.

1.3.3. Compliance

The organization needs to be able to comply with regulations and standards, no matter where your data is stored.

Speaking of storage, also ensure the provider has strict data recovery policies in place.

1.3.4. Cost management and containment

The most part cloud computing can save businesses money.

To avoid purchase of new hardware

Use pay-as-you go models

1.3.5. Governance / Control

Proper IT governance should ensure that IT assets are implemented and used according to agreed upon policies and procedures.

In today's cloud-based world, IT does not always have full control over the provisioning, de-provisioning and operations of infrastructure.

This has increased the difficulty for IT to provide the governance, compliance and risk management required.

1.3.6. Performance

When a business moves to the cloud it becomes dependent on the service providers.

When your provider is down, you are also down.

Make sure your SaaS provider has real time monitoring policies in place to help mitigate these issues.

So, a robust cloud adoption strategy in place when they started to move to the cloud.

The Top 5 Cloud-Computing Vendors:

#1 Microsoft,

#2 Amazon,

#3 IBM,

#4Salesforce,

#5 SAP

#1 Microsoft remains an absolute lock at the top due to four factors: its deep involvement at all three layers of the cloud (IaaS, PaaS and SaaS); its unmatched commitment to developing and helping customers deploy AI, ML and Blockchain in innovative production environments; its market-leading cloud revenue, which I estimate at about $16.7 billion for the trailing 12 months (not to be confused with the forward-projected $20.4 billion annualized run rate the company released on Oct. 26); and the extraordinary vision and leadership of CEO Satya Nadella.

#2 Amazon might not have the end-to-end software chops of the others in the Top 5 but it was and continues to be the poster-child for the cloud-computing movement: the first-moving paradigm-buster and category creator. I believe Amazon will make some big moves to bolster its position in software, and no matter how you slice it, the $16 billion in trailing-12-month cloud revenue from AWS is awfully impressive.

#3 IBM has leapfrogged both Salesforce.com (formerly tied with Amazon for #2 and now in the #4 spot) and SAP (formerly #4) on the strength of its un-trendy but highly successful emphasis on transforming its vast array of software expertise and technology from the on-premises world to the cloud. In so doing, IBM has quietly created a $15.8-billion cloud business (again on trailing-12-month basis) that includes revenue of $7 billion from helping big global corporations convert legacy systems to cloud or cloud-enabled environments. And like #1 Microsoft, IBM plays in all three layers of the cloud—IaaS, PaaS and SaaS—which is hugely important for the elite cloud vendors because it allows them to give customers more choices, more seamless integration, better cybersecurity, and more reasons for third-party developers to rally to the IBM Cloud. Plus, its relentless pairing of "cloud and cognitive" is an excellent approach toward weaving AI and ML deeply into customer-facing solutions.

#4 Salesforce.com falls a couple of spots from its long-time tie with Amazon at #2 but—and this will be the case as long as founder Marc Benioff is CEO—remains a powerful source of digital innovation and disruptive strategy. However, to remain in the rarified air near the top of the Cloud Wars Top 10, Benioff and Salesforce must find a way to extend their market impact beyond their enormously successful SaaS business and become more of a high-impact player in the platform or PaaS space. At this stage, it's simply not possible for Salesforce to become a player in IaaS, so Benioff needs to crank up the genius machine and hammer his way into top contention as a platform powerhouse.

#5 SAP has what all of the other cloud vendors would kill for: unmatched incumbency within all of the world's leading corporations as the supplier of mission-critical business applications that run those companies. It's also fashioned, under CEO Bill McDermott, powerful new partnerships with Amazon and Google to complement its long-standing relationships with IBM and Microsoft, all of which give customers a heightened sense of confidence that SAP will be willing and able to play nice in heterogeneous environments. Plus, SAP's HANA technology is now in full deployment across thousands of businesses, and as it takes root and SAP continues to rationalize its massive product portfolio around HANA in the cloud, SAP has a very bright future ahead of it in the cloud.

So the Cloud Wars are clearly intensifying—particularly among the five most-

powerful players at the top—as business customers are fully embracing the cloud as the best and most- capable approach toward digital transformation.And in such an environment, the long-term winners in the Cloud Wars will be those tech companies that choose to view the world by and create their strategies around what business customers want and need, rather than by the distorted and distorting tech-centric perspectives of the Silicon Valley bubble.

1.4 Cloud Deployment Models

The four types of Cloud Deployment Models identified by NIST.

- Private cloud
- Community cloud
- Public cloud
- Hybrid cloud

1.4.1 Private Cloud

Cloud environments are defined based on hardware location and owner. Private clouds are accessible only to a respective customer residing either on-site or be outsourced by a third party.

The cloud infrastructure is operated solely for an organization.

Contrary to popular belief, private cloud may exist off premises and can be managed by a third party.

Thus, two private cloud scenarios exist, as follows:

- On-site Private Cloud
- Applies to private clouds implemented at a customer’s premises.
- Outsourced Private Cloud
- Applies to private clouds where the server side is outsourced to a hosting company.
- Examples of Private Cloud:
- Eucalyptus
- Ubuntu Enterprise Cloud - UEC (powered by Eucalyptus)
- Amazon VPC (Virtual Private Cloud)
- VMware Cloud Infrastructure Suite
- Microsoft ECI data center.

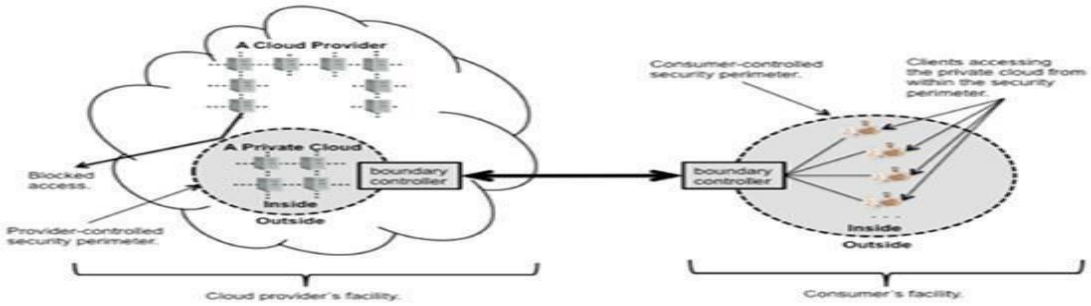


Figure 1.1 Schematic Sketch of Private Cloud

1.4.2 Community Cloud

The cloud infrastructure is shared by several organizations and supports a specific community that has shared concerns (e.g., mission, security requirements, policy, and compliance considerations). Government departments, universities, central banks etc. often find this type of cloud useful.

Community cloud also has two possible scenarios:

On-site Community Cloud Scenario

Applies to community clouds implemented on the premises of the customers composing a community cloud

Outsourced Community Cloud

- Applies to community clouds where the server side is outsourced to a hosting company.
- Examples of Community Cloud:
- Google Apps for Government
- Microsoft Government Community Cloud

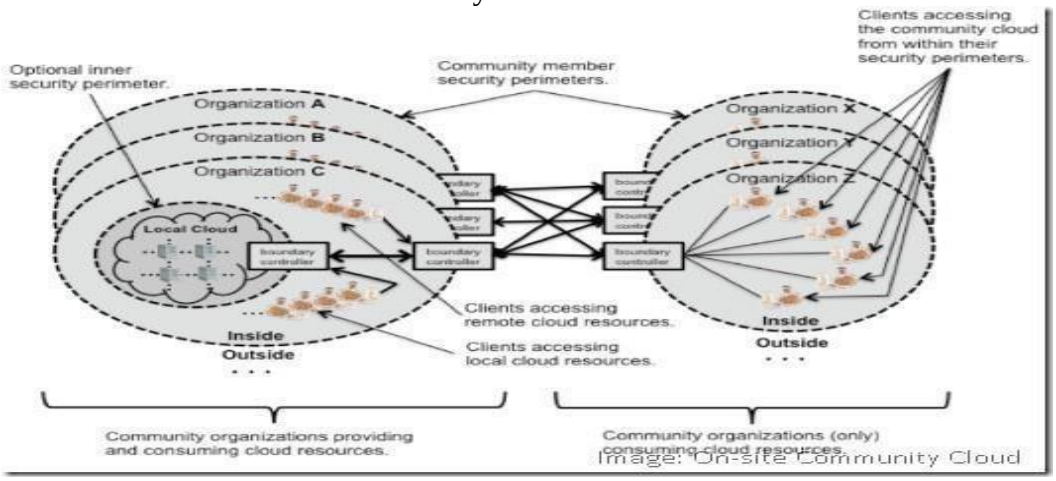


Image: On-site Community Cloud

Figure 1.2. Schematic Sketch of Public Cloud

1.4.3 Public Cloud

The most ubiquitous, and almost a synonym for, cloud computing. The cloud infrastructure is made available to the general public or a large industry group and is owned by an organization selling cloud services.

Examples of Public Cloud:

- Google App Engine
- Microsoft Windows Azure
- IBM Smart Cloud
- Amazon EC2

1.4.4 Hybrid Cloud

The cloud infrastructure is a composition of two or more clouds (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).

Examples of Hybrid Cloud:

- Windows Azure (capable of Hybrid Cloud)
- VMware vCloud (Hybrid Cloud Services)

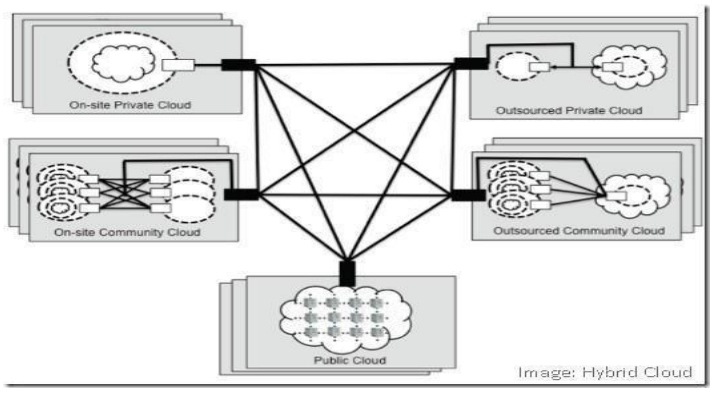


Figure 1.3. Schematic Sketch of Hybrid Cloud

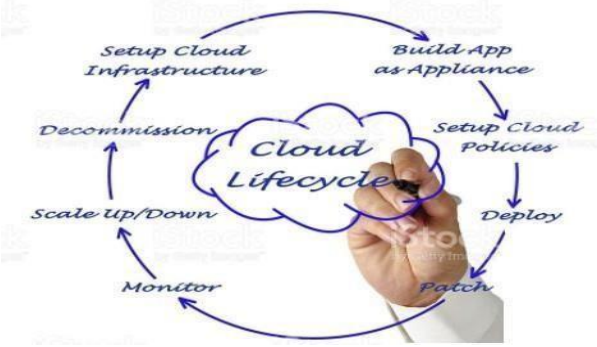


Figure 1.4. Cloud Computing Life Cycle

1.5 Cloud Components

It has three components

- Client computers
- Distributed Servers
- Datacenters

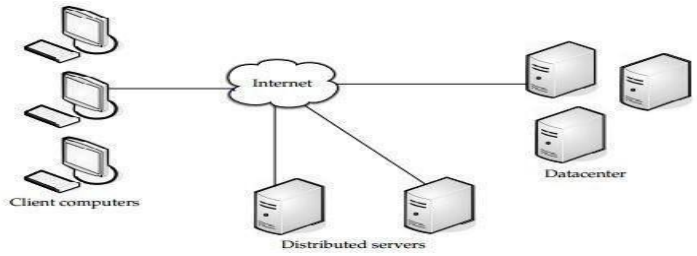


Figure 1.4. Schematic Sketch of Interconnection of Cloud Components

1. Clients

Clients are the device that the end user interacts with cloud.

Three types of clients: 1.) Mobile

2. Thick

3. Thin (Most Popular)

Datacenter

It is collection of servers where application is placed and is accessed via internet.

Distributed servers

Often servers are in geographically different places, but server acts as if they are working next to each other

Who Benefits from Cloud Computing

- Collaborators
- Road Warriors
- Cost-Conscious Users
- Cost-Conscious IT Departments
- Users with Increasing Needs

Who Should not be using Cloud Computing

- The internet impaired
- Offline workers

- The security conscious
- Anyone married to existing applications

Dark Clouds: Barriers to use web-based applications

- Technical issues
- Business model issues
- Internet issues
- Security issues
- Compatibility issues
- Social issues

Cloud Computing Applications

- Clients would be able to access their applications and data from anywhere at any time.
- It could bring hardware costs down.
- The right software available in place to achieve goals.
- Servers and digital storage devices take up space.
- Corporations might save money on IT support.
- The client takes advantage of the entire network's processing power.

Virtualization in cloud computing

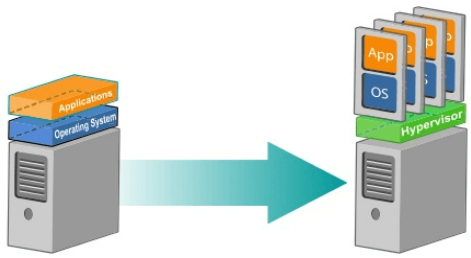


Figure 1.5. Concept of Virtualization

- Virtualization is Nothing but Creating Virtual Version of Something, in Computer Terms it can be OS, Storage Device, Network Resources.
- Virtualization is the ability to run multiple operating systems on a single physical system and share the underlying hardware resources.
- It is the process by which one computer hosts the appearance of many computers.
- Virtualization is used to improve IT throughput and costs by using physical resources as a pool from which virtual resources can be allocated.
- With VMware virtualization solutions you can reduce IT costs while increasing the

efficiency, utilization and flexibility of their existing computer hardware.

- You don't need to own the hardware

Resources are rented as needed from a cloud

You get billed only for what you used

Virtualization Architecture

A Virtual machine (VM) is an isolated runtime environment (guest OS and applications)

Multiple virtual systems (VMs) can run on a single physical system

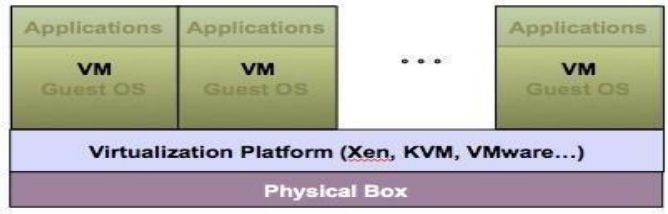


Figure 1.6. Concept of Virtualization

Before Virtualization

- Single OS image per machine
- Software and hardware tightly coupled
- Running multiple applications on same machine often creates conflict
- Inflexible and costly infrastructure

After virtualization

- Hardware-independence of operating system and applications
- Virtual machines can be provisioned to any system
- Can manage OS and application as a single unit by encapsulating them into virtual Machines
- Benefits of Virtualization
- Sharing of resources helps cost reduction.

Isolation: Virtual machines are isolated from each other as if they are physically separated.

Encapsulation: Virtual machines encapsulate a complete computing environment.

Hardware Independence: Virtual machines run independently of underlying hardware.

Portability: Virtual machines can be migrated between different hosts.

What makes virtualization possible?

There is a software that makes virtualization possible. This software is known as a Hypervisor, also known as a virtualization manager.

It sits between the hardware and the operating system and assigns the amount of access that the applications and operating systems have with the processor and other hardware resources.

1.6 Hypervisor

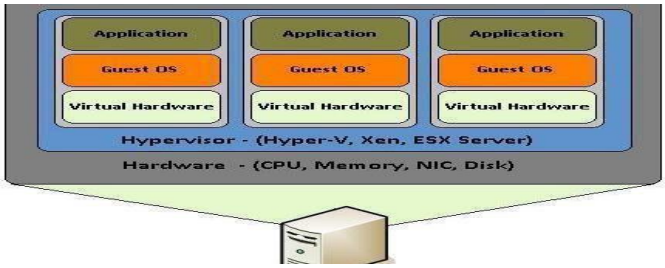


Figure 1.7. Schematic Sketch of Hypervisor

The term hypervisor was first coined in 1956 by IBM

Hypervisor acts as a link between the hardware and the virtual environment and distributes the hardware resources such as CPU usage, memory allotment between the different virtual environments.

A hypervisor or virtual machine monitor (VMM) is computer software, firmware or hardware that creates and runs virtual machines. A computer on which a hypervisor runs one or more virtual machines is called a host machine, and each virtual machine is called a guest machine. A hypervisor is a hardware virtualization technique that allows multiple guest operating systems (OS) to run on a single host system at the same time. The guest OS shares the hardware of the host computer, such that each OS appears to have its own processor, memory and other hardware resources.A hypervisor is also known as a virtual machine manager (VMM).

1.6.1 Types of virtualization

The 7 Types of Virtualization

- Hardware Virtualization.
- Software Virtualization.
- Network Virtualization.
- Storage Virtualization.
- Memory Virtualization.
- Data Virtualization.
- Desktop Virtualization.

1.6.1.1 Hardware Virtualization

Hardware or platform virtualization means creation of virtual machine that act like real computer.

Ex. Computer running Microsoft Windows 7 may host the virtual machine look like a Ubundu

Hardware virtualization also knows as hardware-assisted virtualization or server virtualization.

The basic idea of the technology is to combine many small physical servers into one large physical server, so that the processor can be used more effectively and efficiently.

Each small server can host a virtual machine, but the entire cluster of servers is treated as a single device by any process requesting the hardware.

The hardware resource allotment is done by the hypervisor.

The advantages are increased processing power as a result of maximized hardware utilization and application uptime.

Hardware virtualization is further subdivided into the following types

1.6.1.2 Full Virtualization – Guest software does not require any modifications since the underlying hardware is fully simulated

Para Virtualization – The hardware is not simulated and the guest software run their own isolated domains.

Partial Virtualization – The virtual machine simulates the hardware and becomes independent of it. The guest operating system may require modifications.

1.6.1.3 Software Virtualization

- The ability to computer to run and create one or more virtual environments.
- It is used to enable a computer system in order to allow a guest OS to run.
- Ex. Linux to run as a guest that is natively running a Microsoft Windows OS

Subtypes:

Operating System Virtualization – Hosting multiple OS on the native OS

Application Virtualization – Hosting individual applications in a virtual environment separate from the native OS

Service Virtualization – Hosting specific processes and services related to a particular application

1.6.2 Network Virtualization

- It refers to the management and monitoring of a computer network as a single managerial entity from a single software-based administrator's console.
- Multiple sub-networks can be created on the same physical network, which may or may not is authorized to communicate with each other.
- It allows network optimization of data transfer rates, scalability, reliability, flexibility, and security
- Subtypes:
- Internal network: Enables a single system to function like a network.

1.6.3 Storage Virtualization

- Multiple physical storage devices are grouped together, which look like a single storage device.
- Ex. Partitioning your hard drive into multiple partitions
- Advantages
- Improved storage management in a heterogeneous IT environment
- Easy updates, better availability
- Reduced downtime
- Better storage utilization
- Automated management
- Two types
- Block- Multiple storage devices are consolidated into one
- File- Storage system grants access to files that are stored over multiple hosts

1.6.4 Memory Virtualization

- The way to decouple memory from the server to provide a shared, distributed or networked function.
- It enhances performance by providing greater memory capacity without any addition to the main memory.
- Implementations
- Application-level integration – Applications access the memory pool directly
- Operating System Level Integration – Access to the memory pool is provided through an operating system.

1.6.5 Data Virtualization

Without any technical details, you can easily manipulate data and know how it is formatted or where it is physically located.

It decreases the data errors and workload

The data is presented as an abstract layer completely independent of data structure and database systems

1.6.6 Desktop Virtualization

The user's desktop is stored on a remote server, allowing the user to access his/her desktop from any device or location.

It provides the work convenience and security

It provides a lot of flexibility for employees to work from home or on the go

Since the data transfer takes place over secure protocols, any risk of data theft is minimized

Which Technology to use?

Virtualization is possible through a wide range of Technologies which are available to use and are also Open Source.

They are,

- XEN
- KVM
- OpenVZ

1.7 Parallelization in cloud computing

- Parallel computing is a type of computing architecture in which several processors execute or process an application or computation simultaneously.
- Parallel computing helps in performing large computations by dividing the workload between more than one processor, all of which work through the computation at the same time.

- Most supercomputers implemented parallel computing principles to operate.
- Parallel computing is also known as parallel processing.
- Parallel processing is generally implemented in operational environments/scenarios that require massive computation or processing power.
- The primary objective of parallel computing is to increase the available computation power for faster application processing.
- Typically, parallel computing infrastructure is housed within a single facility where many processors are installed in a server rack or separate servers are connected together.
- The application server sends a processing request that is distributed in small components, which are concurrently executed on each processor/server.
- Parallel computation can be classified as bit-level, instructional level, data and task parallelism.

Cloud Resource Management

- Critical function of any man-made system.
- It affects the three basic criteria for the evaluation of a system:
- Functionality.
- Performance.
- Cost.
- Scheduling in a computing system deciding how to allocate resources of a system, such as CPU cycles, memory, secondary storage space, I/O and network bandwidth, between users and tasks.
- Policies and mechanisms for resource allocation.
- Policy: principles guiding decisions.
- Mechanisms: the means to implement policies
- Cloud resources
- Requires complex policies and decisions for multi-objective optimization.
- It is challenging - the complexity of the system makes it impossible to have accurate global state information.
- Affected by unpredictable interactions with the environment, e.g.,system failures, attacks.
- Cloud service providers are faced with large fluctuating loads which challenge the claim of cloud elasticity

- The strategies for resource management for IaaS, PaaS, and SaaS are different.

Cloud resource management (CRM) policies

- Admission control: prevent the system from accepting workload in violation of high-level system policies.
- Capacity allocation: allocate resources for individual activations of a service.
- Load balancing: distribute the workload evenly among the servers.
- Energy optimization: minimization of energy consumption
- Quality of service (QoS) guarantees: ability to satisfy timing or other conditions specified by a Service Level Agreement

1.7 Dynamic resource allocation

- Cloud Computing environment can supply of computing resources on the basis of demand and when needed
- Managing the customer demand creates the challenges of on-demand resource allocation.
- Effective and dynamic utilization of the resources in cloud can help to balance the load and avoid situations like slow run of systems.
- Cloud computing allows business outcomes to scale up and down their resources based on needs.
- Virtual Machines are allocated to the user based on their job in order to reduce the number of physical servers in the cloud environment
- If the VM is available then job is allowed to run on the VM.
- If the VM is not available then the algorithm finds a low priority job taking into account the job's lease type.
- The low priority job is paused its execution by pre-empting its resource.
- The high priority job is allowed to run on the resources pre-empted from the low priority.
- When any other job running on VMs are completed, the job which was paused early can be resumed if the lease type of the job is suspendable.
- If not, the suspended job has to wait for the completion of high priority job running in its resources, so that it can be resumed.
- There are three types
- Cancellable: These requests can be scheduled at any time after their arrival time
- Suspendable: Suspendable leases are flexible in start time and can be scheduled at any time after their ready time

- Non-Preemptable: The leases associated with such requests cannot be pre-empted at all.

1.8 Optimal allocation of cloud models

The optimal allocation of computing resources is a core part for implementing cloud computing.

High heterogeneity, high dynamism, and virtualization make the optimal allocation problem more complex than the traditional scheduling problems in grid system or cloud computing system.

SATHYABAMA
INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)
Accredited "A" Grade by NAAC | 12B Status by UGC | Approved by AICTE
www.sathyabama.ac.in

SCHOOL OF COMPUTING

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

UNIT – II – Cloud computing – SIT1304

UNIT 2

CLOUD SERVICE MODELS

Software as a Service (SaaS) - Infrastructure as a Service (IaaS)- Platform as a Service (PaaS)- Service Oriented Architecture (SoA) - Elastic Computing - On Demand Computing.

2.1 Types of Services provided by Cloud

- Software as a Service (SaaS)
- Infrastructure as a Service (IaaS)
- Platform as a Service (PaaS)

Service Oriented Architecture

- Elastic Computing
- On Demand Computing

2.2 Cloud Services

Software as a Service

Introduction

SaaS (software-as-a-service). WAN-enabled application services (e.g., Google Apps, Salesforce.com, WebEx).This is a public cloud service model where the application is 100% managed by the cloud provider. SaaS removes the need for organizations to install and run applications on their own computers or in their own data centers.This eliminates the expense of hardware acquisition, provisioning and maintenance, as well as software licensing, installation and support.Software-as-a-Service (SaaS) has evolved from limited on-line software delivery in 1990s to a fully matured “direct-sourcing” business model for enterprise applications.

SaaS is one of the fastest growing concepts: more than 10 million companies will be using SaaS in the next 5 - 10 years; more than 50% of all Fortune 500 companies are already using SaaS.According to influential IT institutes, SaaS is the leading business model of choice for 2008/2009Virtually all big software/service vendors (IBM, Microsoft, Oracle, Cisco) are investing heavily in SaaSWith the continuously increasing bandwidth and reliability of the internet, using web services over the (public) internet has become a viable option.Microsoft Office 365 is available with the Azure cloud platform.The architecture uses an application instance instead of server instances.There is no actual migration of company servers to the cloud.

The SaaS model provides single-tenant and multiple tenant services.The single-tenant dedicates the application instance to the assigned tenant.The multiple tenant application is

shared by multiple tenants.The company can manage the security and storage with the single-

tenant model.The SaaS application is well suited to internet connectivity.The employees along with their partners and customers can access the application with a variety of network access devices.The SaaS billing model is based on either per usage or monthly subscription.The security compliance requirements for some applications prevented deployment to the SaaS cloud.

Some SaaS providers offer Virtual Private SaaS (VPS) services for additional application security.It is a hybrid deployment model that allows peering with an enterprise or VPC database server.The peering is for storage purposes only and used for security compliance. Salesforce.com is a leading SaaS provider with a CRM application to customers.

2.2.1Benefits of SaaS

- Flexible payments
- Scalable usage
- Automatic updates
- Accessibility and persistence
- On demand computing
- Opportunities of SaaS

Software provided as a service by a software vendor to multiple customers with the following main characteristics:

- Standardization of software
- Service including maintenance, support and upgrades
- Web based – usage over the (public) internet
- SaaS offers potential for lowering the Total Cost of Ownership
- Lower operational costs
- No large scale, costly, high risk implementations of applications
- Need few operational resources for application management
- No platform and hardware (maintenance) costs for application servers
- Reduced operational complexity: software delivered as a transparent service through the web
- Minimized software development costs – No lengthy software development and testing cycles
- Lower costs for software use
- No software license and annual maintenance fees
- No expensive software upgrades
- Lower application consultancy and support costs

- SaaS allows corporations to focus on core business activities and responsibilities
- Transparent overview and usage of electronic data and information
- Automation of iterative, manual tasks
- Faster Time to Market – easy to scale software
- More flexibility in changing and modifying application services for business needs – Full-scale integration of business processes
- Control over IT
- Minimized IT Service Management efforts mainly focused on availability –
- Well-defined SLAs between the corporation and the IT vendor
- More predictable cash flow – easier licensing based on access/usage of software
- Increased productivity and improved user satisfaction
- Automatic software upgrades with minimal outage

2.2.2 Limitations

- Businesses must rely on outside vendors to provide the software, keep that software up and running, track and report accurate billing and facilitate a secure environment for the business' data.

2.3Platform as a Service

The Platform as a Service (PaaS) is a way to rent hardware, operating systems, storage and network capacity over the Internet.

PaaS services are,

- Data services
- Application runtime
- Messaging & queueing
- Application management.
- The PaaS is a computing platform that abstracts the infrastructure, OS, and middle-ware to drive developer productivity.
- The PaaS is foundational elements to develop new applications
- E.g., Google Application Engine, Microsoft Azure, Coghead.
- Microsoft Azure
- Pay per role instance
- Add and remove instances based on demand
- Elastic computing!
- Load balancing is part of the Azure fabric and automatically allocated.
- The PaaS is the delivery of a computing platform and solution stack as a service

- The Solution stack is integrated set of software that provides everything a developer needs to build an application for both software development and runtime.

2.3.1 PaaS offers the following

Facilities for application design

- Application development
- Application testing, deployment
- Application services are,
- Operating system
- Server-side scripting environment
- Database management system
- Server Software
- Support
- Storage
- Network access
- Tools for design and development
- Hosting

All these services may be provisioned as an integrated solution over the web

2.3.2 Properties and characteristics of PaaS

- Scalability
- Availability
- Manageability
- Performance
- Accessibility

2.3.3 PaaS Features

- It delivers the computing platform as service
- The capacities to abstract and control all the underlying resources
- It helps to providers any smallest unit of resources
- To provide a reliable environment for running applications and services
- Act as a bridge between consumer and hardware
- Do not need to care about how to build, configure, manage and maintain the backend environment
- It provides a development and testing platform for running developed applications
- Reduce the responsibility of managing the development and runtime environment

2.3.4 Advantages of PaaS

It helps to provide deployment of application without the cost and complexity of buying and managing the hardware and software

It provides all the required to support the complete life cycle of building and delivering web applications and services entirely available from the internet

2.3.5Disadvantages of PaaS

- Less flexible than IaaS
- Dependency on provider
- Adoption of software / system architecture required
- Evolving from different standard.
- Evolving “upwards” from IaaS
- Amazon (Mail, Notification, Events, Databases, Workflow, etc.)
- Evolving “downwards” from SaaS
- Force.com – a place to host additional per-tenant logic.
- Google App Engine
- Evolving “sideways” from middleware platforms
- WSO2, Tibco, vmWare, Oracle, IBM
- Generic PaaS Model
- Infrastructure as a Service
- This service offers the computing architecture and infrastructure i.e. all types of computing resources

All resources are offered in a virtual environment, so that multiple users can access it.

The resources are including,

Data storage

- Virtualization
- Servers
- Networking
- The vendors are responsible for managing all the computing resources which they provided.
- It allows existing applications to be run on a supplier’s hardware.
- User Task in IaaS Cloud
- Multiple user can access Virtual instances

The user responsible for handling other resources such as,

2.3.6 Applications

- Data

- Runtime
- Middleware

Example IaaS service providers

- AWS EC2 / S3 / RD
- GoGrid
- RackSpace

2.3.7 Pros

- The cloud provides the infrastructure with enhanced scalability i.e. dynamic workloads are supported
- It is flexible
- Cons
- Security issues
- Network and service delay

2.3.8 Comparison of cloud services

- Blue indicates the levels owned and operated by the organization / Customer
- White levels are run and operated by the service provider / Operator
- Cloud Computing Services Pros
- Lower computer costs
- Improved performance:
- Reduced software costs
- Instant software updates
- Improved document format compatibility
- Unlimited storage capacity
- Increased data reliability
- Universal document access
- Latest version availability
- Easier group collaboration
- Device independence

2.3.9Cons

- Requires a constant Internet connection
- Does not work well with low-speed connections
- Features might be limited
- Can be slow
- Stored data can be lost

- Stored data might not be secure
- Service Oriented Architecture

2.4Service

A service is a program you interact with via message exchanges A system is a set of deployed services cooperating in a given task.

2.5Architecture

It serves as the blueprint for the system

Team structure

- Documentation organization
- Work breakdown structure
- Scheduling, planning, budgeting
- Unit testing, integration
- Architecture establishes the communication and coordination mechanisms among components

2.5.1 Software Architecture

- It is collection of the fundamental decisions about a software product/solution designed to meet the project's quality attributes (i.e. requirements).
- The architecture includes the main components, their main attributes, and their collaboration (i.e. interactions and behavior) to meet the quality attributes.
- Architecture can and usually should be expressed in several levels of abstraction (depending on the project's size).
- Architecture is communicated from multiple viewpoints

2.5.2 Service Oriented Architecture (SOA)

- It is a design pattern or software architecture which provides application functionality as a service to other applications.
- The basic principles of service-oriented architecture are independent of vendors, products and technologies.
- The services are provided to the other components through a communication protocol over a network.
- Every service has its own business logic
- Consumer interface layer – this layer is used by the customer
- Business process layer – it provides the business process flow
- Service layer – this layer comprises of all the services in the enterprises
- Component layer – this layer has the actual service to be provided

- Operational system layer – this layer contains the data model

SOA – Architecture in details

2.5.3 Principles of SOA

- Service loose coupling – service does not have high dependency
- implementation from outside world
- Service reusability – services can be used again and again instead of rewriting them
- Service statelessness – they usually do not maintain the state to reduce the resource consumption
- Service discoverability – services are registered in registry, so that the client can discover them in the service registry.
- Applications
- Manufacturing – E.g. Inventory management
- Insurance – Take up the insurance of the employees in companies

Companies using SOA

- Banking Sector
- ICICI Bank
- HDFC Bank
- UTI Bank etc..
- Manufacturing Sector
- Apollo Tyres
- Maruthi
- Hyundai

2.5.4 Advantages

- Interoperability
- Programs to run different vendors / locations
- To interact with different networks
- Different operating systems
- Solution: XML
- Scalability
- To extend the processing power of the servers
- Reusability
- If any new systems are introduced, no need to create a new service for every time.
- Parallel application development
- Modular approach

- Easy maintenance
- Greater Reliability
- Improved Software Quality
- Platform Independence
- Increased Productivity

2.5.5 Disadvantages

- Stand alone, non-distributed applications
- Homogenous application environments
- GUI based applications
- Short lived applications
- Real time applications
- One-way asynchronous communication applications

ELASTIC COMPUTING:

Elastic computing is the ability to quickly expand or decrease computer processing, memory and storage resources to meet changing demands without worrying about capacity planning and engineering for peak usage. Typically controlled by system monitoring tools, elastic computing matches the amount of resources allocated to the amount of resources actually needed without disrupting operations. With cloud elasticity, a company avoids paying for unused capacity or idle resources and does not have to worry about investing in the purchase or maintenance of additional resources and equipment.

While security and limited control are concerns to take into account when considering elastic cloud computing, it has many benefits. Elastic computing is more efficient than your typical IT infrastructure, is typically automated so it does not have to rely on human administrators around the clock and offers continuous availability of services by avoiding unnecessary slowdowns or service interruptions.

What is Elastic Cloud Computing and how it Benefits Business

Nowadays cloud computing are well known phenomenon for everyone. Most of the small and large business have switched their data to cloud storage. Moreover, organization also prefer to have elastic computing.

But before proceeding to know more about elastic computing let's have a quick outline of cloud computing. Cloud Computing or Cloud is defined as using various services such as software development platforms, servers, storage, over the Internet.

So, **what does Elastic Cloud Computing mean?**

Elastic computing is nothing but a concept in cloud computing in which computing resources can be scaled up and down easily by the cloud service provider. Cloud service provider gives you provision to flexible computing power when and wherever required. The elasticity of these resources depends upon the following factors such as processing power, storage, bandwidth, etc.

Types of Elastic Cloud Computing

Rather than various types, elastic computing have only one type i.e. **Elasticity, or fully-automated scalability** which removes manual labor for increasing or decreasing resources as everything is controlled by triggers by the system monitoring tools.

Is there any **difference between Scalability and Elasticity?**

Then answer is yes. **Scalability** refers to the ability of system to accommodate larger loads just by adding resources either making hardware stronger (scale up) or adding additional nodes (scale out).

Elasticity refers the ability to fit the resources needed to cope with loads, so that when load increase you scale up by adding more resources and when demand diminishes you shrink back and remove unneeded resources. Elasticity is mostly important in Cloud environment where you pay-per-used resources only.

Benefits/Pros of Elastic Cloud Computing

Elastic Cloud Computing has numerous advantages. Some of them are as follow:-

1. **Cost Efficiency:** - Cloud is available at much cheaper rates than traditional approaches and can significantly lower the overall IT expenses. By using cloud solution companies can save licensing fees as well as eliminate overhead charges such as the cost of data storage, software updates, management etc.

2. **Convenience and continuous availability:** - Cloud makes easier access of shared documents and files with view and modify choice. Public clouds also offer services that are available wherever the end user might be located. Moreover it guaranteed continuous availability of resources and In case of system failure; alternative instances are automatically spawned on other machines.

3. **Backup and Recovery**: - The process of backing up and recovering data is easy as information is residing on cloud simplified and not on a physical device. The various cloud providers offer reliable and flexible backup/recovery solutions.

4. **Cloud is environmentally friendly**:-The cloud is more efficient than the typical IT infrastructure and it takes fewer resources to compute, thus saving energy.

5. **Scalability and Performance**: - Scalability is a built-in feature for cloud deployments. Cloud instances are deployed automatically only when needed and as a result enhance performance with excellent speed of computations.

6. **Increased Storage Capacity**:-The cloud can accommodate and store much more data compared to a personal computer and in a way offers almost unlimited storage capacity.

Disadvantages/Cons of Elastic Cloud Computing:-

1. **Security and Privacy in the Cloud**: - Security is the biggest concern in cloud computing. Companies essentially hide their private data and information over cloud as remote based cloud infrastructure is used, it is then up to the cloud service provider to manage, protect and retain data confidential.
2. **Limited Control**: - Since the applications and services are running remotely companies, users and third party virtual environments have limited control over the function and execution of the hardware and software.
3. **Dependency and vendor lock-in**: - One of the major drawbacks of cloud computing is the implicit dependency on the provider. It is also called "vendor lock-in". As it becomes difficult to migrate vast data from old provider to new. So, it is advisable to select vendor very carefully.
4. **Increased Vulnerability**: - Cloud based solutions are exposed on the public internet therefore are more vulnerable target for malicious users and hackers. As we know nothing is completely secure over Internet even the biggest organizations also suffer from serious attacks and security breaches.

Regardless the disadvantages elastic cloud computing even remains stronger and has great potential for the future. Elastic computing offering better, more fine-tuned and easy to use services and solutions. We can only hope that the advantages will grow more and the disadvantages will be diminished as cloud is the future.

On-Demand Computing:

On-demand computing is a business computing model in which computing resources are made available to the user on an "as needed" basis. Rather than all at once, on-demand computing allows cloud hosting companies to provide their clients with access to computing resources as they become necessary.

On-demand computing is a delivery model in which computing resources are made available to the user as needed. The resources may be maintained within the user's enterprise, or made available by a cloud service provider. When the services are provided by a third-party, the term cloud computing is often used as a synonym for on-demand computing.

The on-demand model was developed to overcome the common challenge to an enterprise of being able to meet fluctuating demands efficiently. Because an enterprise's demand on computing resources can vary drastically from one time to another, maintaining sufficient resources to meet peak requirements can be costly. Conversely, if an enterprise tried to cut costs by only maintaining minimal computing resources, it is likely there will not be sufficient resources to meet peak requirements.

The on-demand model provides an enterprise with the ability to scale computing resources up or down with the click of a button, an API call or a business rule. The model is characterized by three attributes: scalability, pay-per-use and self-service. Whether the resource is an application program that helps team members collaborate or additional storage for archiving images, the computing resources are elastic, metered and easy to obtain.

Many on-demand computing services in the cloud are so user-friendly that non-technical end users can easily acquire computing resources without any help from the organization's information technology (IT) department. This has advantages because it can improve business agility, but it also has disadvantages because shadow IT can pose security risks. For this reason, many IT departments carry out periodic cloud audits to identify greynet on-demand applications and other rogue IT.

Advantages of On-Demand Computing:

The on-demand computing model was developed to overcome the common challenge that enterprises encountered of not being able to meet unpredictable, fluctuating computing demands in an efficient manner.

Businesses today need to be agile and need the ability to scale resources easily and quickly based on rapidly changing market needs.

Because an enterprise’s demand for computing resources can vary dramatically from one period of time to another, maintaining sufficient resources to meet peak requirements can be costly.

However, with on-demand computing, companies can cut costs by maintaining minimal computing resources until they run into the need to increase them, meanwhile only paying for what they use.

Industry experts predict on-demand computing to soon be the most widely used computing model for enterprises.

In fact, IBM’s vice-president of technology and strategy stated, “The technology is at a point where we can start to move into an era of on-demand computing. I give it between two and four years to reach a level of maturity.”

SATHYABAMA
INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)
Accredited "A" Grade by NAAC | 12B Status by UGC | Approved by AICTE
www.sathyabama.ac.in

SCHOOL OF COMPUTING

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

UNIT – III – Cloud Computing– SIT1304

UNIT III

Deployment of application on cloud , Hypervisor, Case Studies, Xen , VMWare, Eucalyptus, Amazon EC2, KVM,Virtual Box, Hyper-V

Deployment of Applications on cloud:

Step 1: Prepare to Deploy

Before you deploy your app to Cloud Foundry, make sure that:

- Your app is *cloud-ready*. Cloud Foundry behaviors related to file storage, HTTP sessions, and port usage may require modifications to your app.
- All required app resources are uploaded. For example, you may need to include a database driver.
- Extraneous files and artifacts are excluded from upload. You should explicitly exclude extraneous files that reside within your app directory structure, particularly if your app is large.
- An instance of every service that your app needs has been created.
- Your Cloud Foundry instance supports the type of app you are going to deploy, or you have the URL of an externally available buildpack that can stage the app.

Step 2: Know Your Credentials and Target

Before you can push your app to Cloud Foundry you need to know:

- The API endpoint for your Cloud Foundry instance. Your username and password for your Cloud Foundry instance.
- The organization and space where you want to deploy your app. A Cloud Foundry workspace is organized into organizations, and within them, spaces. As a Cloud Foundry user, you have access to one or more organizations and spaces.

Step 3: (Optional) Configure Domains

Cloud Foundry directs requests to an app using a route, which is a URL made up of a host and a domain.

- The name of an app is the default host for that app, unless you specify the host name with the `-n` flag.
- Every app is deployed to an app space that belongs to a domain. Every Cloud Foundry instance has a default domain defined. You can specify a non-default, or custom, domain when deploying, provided that the domain is registered and is mapped to the organization which contains the target app space.

Step 4: Determine Deployment Options

Before you deploy, you need to decide on the following:

- **Name**: You can use any series of alpha-numeric characters as the name of your app.
- **Instances**: Generally speaking, the more instances you run, the less downtime your app will experience. If your app is still in development, running a single instance can simplify troubleshooting. For any production app, we recommend a minimum of two instances.
- **Memory Limit**: The maximum amount of memory that each instance of your app can consume. If an instance exceeds this limit, Cloud Foundry restarts the instance.
- **Start Command**: This is the command that Cloud Foundry uses to start each instance of your app. This start command varies by app framework.

- **Subdomain (host) and Domain**: The route, which is the combination of subdomain and domain, must be globally unique. This is true whether you specify a portion of the route or allow Cloud Foundry to use defaults.
- **Services**: Apps can bind to services such as databases, messaging, and key-value stores. Apps are deployed into app spaces. An app can only bind to a service that has an existing instance in the target app space.

Define Deployment Options

You can define deployment options on the command line, in a manifest file, or both together.
See Deploying with App Manifests to learn how app settings change from push to push, and how command-line options, manifests, and commands like `cf scale` interact.

When you deploy an app while it is running, Cloud Foundry stops all instances of that app and then deploys. Users who try to run the app get a "404 not found" message while `cf push` runs. Stopping all instances is necessary to prevent two versions of your code from running at the same time. A worst-case example would be deploying an update that involved a database schema migration, because instances running the old code would not work and users could lose data.

Cloud Foundry uploads all app files except version control files and folders with names such as `.svn`, `.git`, and `_darcs`. To exclude other files from upload, specify them in a `.cfignore` file in the directory where you run the push command. For more information, see the Ignore Unnecessary Files When Pushing section of the Considerations for Designing and Running an App in the Cloud topic.

For more information about the manifest file, see the Deploying with App Manifests topic.

```
# Set the default LANG for your apps
export LANG=en_US.UTF-8
```

Step 5: Push the App

Run the following command to deploy an app without a manifest:

```
cf push APP-NAME
```

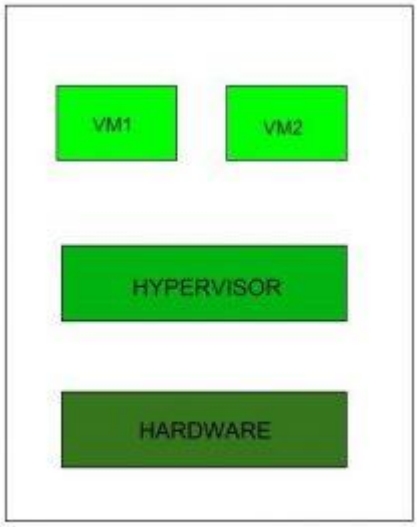
Hypervisor

Hypervisor is a form of virtualization software used in Cloud hosting to divide and allocate the resources on various pieces of hardware.The program which provide partitioning, isolation or abstraction is called virtualization hypervisor. Hypervisor is a hardware virtualization technique that allows multiple guest operating systems (OS) to run on a single host system at the same time. A hypervisor is sometimes also called a virtual machine manager(VMM).

Types of Hypervisor –

TYPE-1 Hypervisor:
Hypervisor runs directly on underlying host system.It is also known as "Native Hypervisor" or "Bare metal hypervisor".It dose not require any base server operating system.It has direct access to hardware

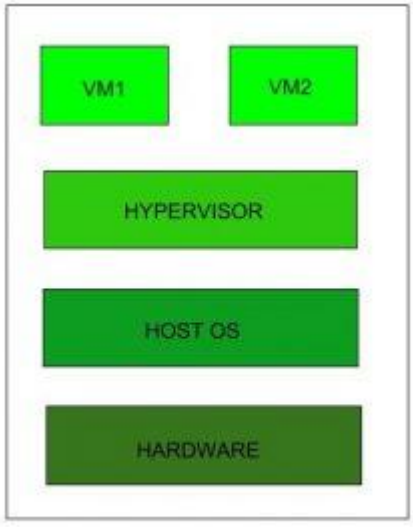
resources.Examples of Type 1 hypervisors include VMware ESXi, Citrix XenServer and Microsoft Hyper-V hypervisor.



TYPE1 HYPERVISOR

TYPE-2 Hypervisor:

A Host operating system runs on undrlying host system.It is also known as 'Hosted Hypervisor".Basically a software installed on an operating system.Hypervisor asks operating system to make hardware calls.Example of Type 2 hypervisor include VMware Player or Parallels Desktop. Hosted hypervisors are often found on endpoints like PCs.



TYPE 2 HYPERVISOR

Choosing the right hypervisor

Type 1 hypervisors offer much better performance than Type 2 ones because there's no middle layer, making them the logical choice for mission-critical applications and workloads. But that's not to say that hosted hypervisors don't have their place – they're much simpler to set up, so they're a good bet if, say, you need to deploy a test environment quickly.One of the best ways to determine which hypervisor meets your needs is to compare their performance metrics. These include CPU overhead, amount of maximum host and guest memory, and support for virtual processors.The following factors should be examined before choosing a suitable hypervisor:

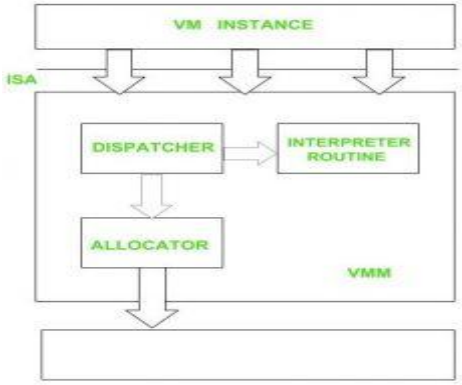
1. Understand your needs: The company and its applications are the reason for the data center (and your job). Besides your company's needs, you (and your co-workers in IT) also have your own needs.Needs for a virtualization hypervisor are:

a. Flexibility
b. Scalability
c. Usability
d. Availability
e. Reliability

f. Efficiency
g. Reliable support

2. The cost of a hypervisor: For many buyers, the toughest part of choosing a hypervisor is striking the right balance between cost and functionality. While a number of entry-level solutions are free, or practically free, the prices at the opposite end of the market can be staggering. Licensing frameworks also vary, so it's important to be aware of exactly what you're getting for your money.
3. Virtual machine performance: Virtual systems should meet or exceed the performance of their physical counterparts, at least in relation to the applications within each server. Everything beyond meeting this benchmark is profit.
4. Ecosystem: It's tempting to overlook the role of a hypervisor's ecosystem – that is, the availability of documentation, support, training, third-party developers and consultancies, and so on – in determining whether or not a solution is cost-effective in the long term.
5. Test for yourself: You can gain basic experience from your existing desktop or laptop. You can run both VMware vSphere and Microsoft Hyper-V in either VMware Workstation or VMware Fusion to create a nice virtual learning and testing environment.

HYPERVISOR REFERENCE MODEL



There are 3 main modues coordinate in order to emiulate the undrelying hardware:

1. Dispatcher
2. Allocator
3. Interpreter

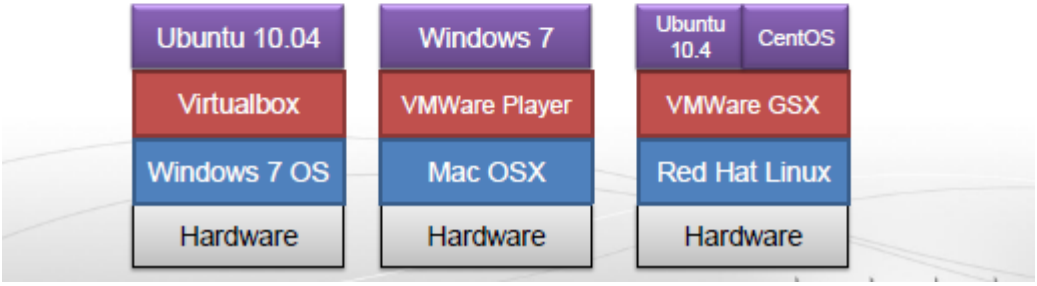
DISPATCHER:
The dispatcher behaves like the entry point of the monitor and reroutes the instructions of the virtual machine instance to one of the other two modules.
ALLOCATOR:
The allocator is responsible for deciding the system resources to be provided to the virtual machine instance.It means whenever virtual machine tries to execute an instruction that results in changing the machine resources associated with the virtual machine, the allocator is invoked by the dispatcher.
INTERPRETER:
The interpreter module consists of interpreter routines.These are executed, whenever virtual machine executes a priviliged instruction.

Virtualization CaseStudies:

- Virtualization environments can be classified as:
  - *Hosted/Dual-Mode Virtual Machine Environments*
    - VirtualBox
    - VMWare GSX/Player/Workstation
  - *System Virtual Machine Environments*
    - Xen
    - VMWare ESX, ESXi

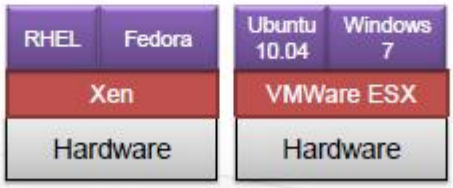
Hosted Virtual Machines

- Hosted Virtual Machines
  - The virtualizing software is installed on top of a traditional operating system.
- Dual-Mode Virtual Machines
  - Portion of the virtualization software runs in the privilege level as the Host operating system
- Examples:
  - VMWare GSX, Player, Workstation



System Virtual Machines

- In a system virtual machine, the virtualizing software (hypervisor) is installed in the place of a traditional operating system.
  - Also known as a "bare-metal" hypervisor
- Guest OSes are installed on top of the hypervisor
  - Eg: Xen, VMWare ESX



Case Study: Xen

Xen

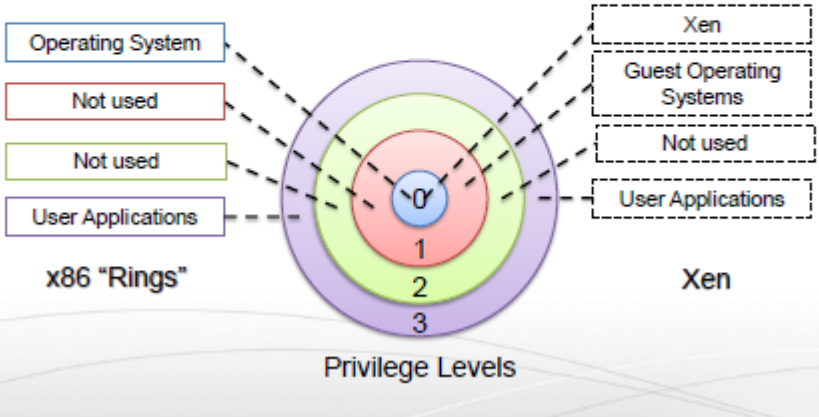
- Xen is a virtual machine monitor that provides services to allow multiple computer operating systems to execute on the same hardware simultaneously.
- Xen was a research project at the University of Cambridge Computer Laboratory in association with Microsoft and Intel research in Cambridge, UK.

Xen approach to Virtualization

- Classically, Xen uses a *paravirtualization* architecture (PVM)
  - Guest OS'es need to be modified in order to work with Xen
  - Guest OS is aware of virtualization – allows better performance and simplifies the hypervisor design.
- As of version 3.0, Xen also supports Hardware-Assisted Virtualization (HVM)
  - Virtualization support was added to the x86 ISA in new processors (such at Intel VT-x or AMD-V)
  - This enables full virtualization without the need for modifying the Guest OS.
- Design Goal: Keep the hypervisor layer as small and as simple as possible.
  - Management tools, device drivers etc. run in a privileged VM
  - This enhances security, resource isolation

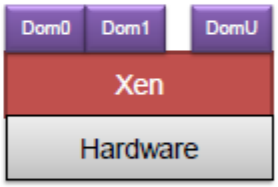
Levels of Protection in x86 and Xen

- Intel's x86 architecture provides levels of privilege for code executing on the processor



Basics of Xen

- Xen hypervisor runs on hardware
  - "Classic" system VM
  - The hypervisor is a thin layer with minimal functionality
- Guest Operating Systems run "on top" of Xen
  - Each virtual machine is known as a **domain**
- Domain 0 is a privileged Virtual Machine
  - Typically some version of NetBSD / Solaris / Linux
  - Privileged access to resources
  - Often hosts device drivers
  - Contains tools to manage other domains on the physical machine

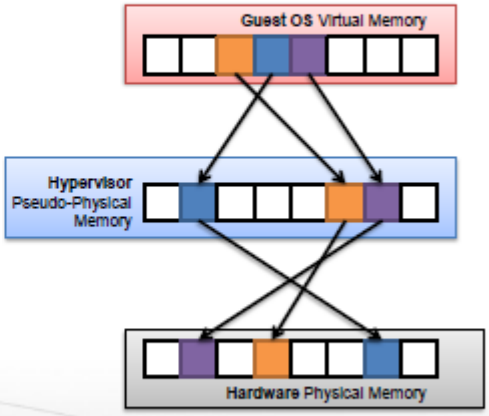


CPU Virtualization in Xen

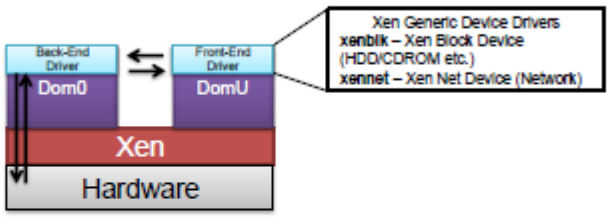
- x86 instruction set was hard to virtualize
  - Certain instructions can be executed only in privileged mode
  - PVM Approach: modify the operating system to replace privileged instructions
- OS code containing privileged instructions are replaced by *hypercalls*
  - The hypercall (similar to a system call) is used to interact with hardware resources
  - The hypervisor will handle a hypercall and will manage the resources between all of the virtual machines
- The modification required to the OS is not very large
  - The OS may contain code that is specific to an particular architecture – which might be privileged instructions. These need to be modified in the PVM approach.
  - Less than 5% of the code to be modified (in case of linux)

Memory Management in Xen

- The Xen hypervisor controls memory management of the system
  - Responsible for physical memory allocation to guest operating systems
- Guest OS has view of "*Real*" (*pseudo-physical*) memory pages
  - Xen hypervisor maps pseudo-physical pages to physical memory on the hardware
- Interaction with Memory depends on OS type
  - For PVM (paravirtualized) – guest OS uses hypercalls to interact with memory
  - For HVM (hardware-assisted) – Xen uses shadow page tables that are accessed through hardware instructions (in Intel VT-x or AMD-V supported CPUs)



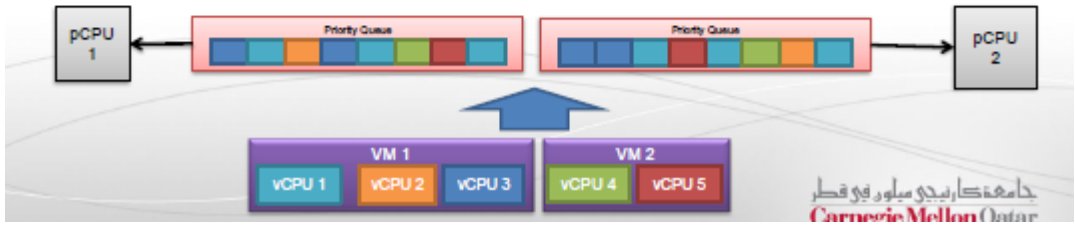
Device Management in Xen



- In a paravirtualized environment, modifying the Guest OS means that existing drivers will not work
  - It is tedious to re-write all device drivers for all the modified guest operating systems
- Instead, Xen uses a *split-driver* model
  - Front-end drivers in DomU, back-end drivers in Dom0
  - The front-end drivers are generic, one type of driver required for each class of device
  - Back-end drivers are device specific and can reuse existing drivers written for Linux

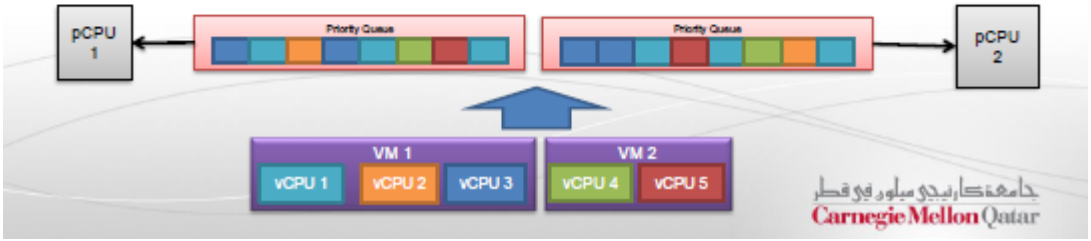
Multiprocessor Virtualization in Xen

- Xen schedules virtual CPUs (vCPUs) against a pool of physical CPUs (pCPUs)
  - Different schedulers available for allocation and scheduling
  - Default scheduler is the Credit Scheduler.
- *Credit Scheduler*
  - Proportional fair-share algorithm
  - Users can assign a weight and a cap for each VM, this influences the scheduling decision and affects the CPU scheduling priority of VMs
  - Assigns credits for all VMs running on the host and debits credits from VMs periodically for each running vCPU.



Credit Scheduler (contd.)

- vCPU priority is calculated based on the credits remaining for each vCPU and is refreshed periodically.
- Positive credit VMs are given status of *OVER* and negative credit vCPUs are given status of *UNDER*.
- Scheduling priority determined in the following order:
  - UNDER VM in the run queue of the local CPU
  - UNDER VM in a run queue of another CPU
  - OVER VM in the run queue of the local CPU
  - OVER VM in a run queue of another CPU



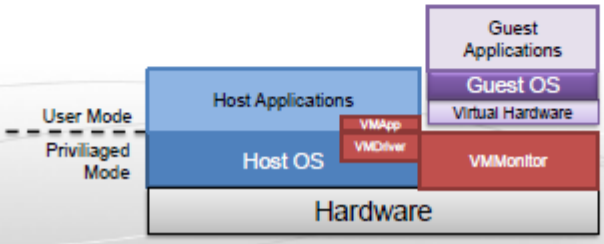
Case Study: VMWare

VMWare Products

- VMWare has two types of Products:
- Hosted VMMs
  - VMWare Workstation, VMWare GSX, VMWare Player, VMWare Fusion etc.
- Bare-Metal Operating Systems
  - VMWare ESX. ESXi

VMWare Dual-Mode Hosted Architecture

- Original VMWare virtual platform approach
  - Host Operating System is loaded first
  - VMWare is installed after the Host OS is installed.
- The virtualization platform contains 3 components:
  - *VMMonitor* (System Level VMM)
    - Runs in privileged mode and has access to hardware resources
  - *VMApp* (User Level VMM)
    - Runs in user mode
    - Makes I/O system calls and has access to the Host OS device drivers
  - *VMDriver* (Pseudo-Driver)
    - Co-ordinates the communication between VMMonitor and VMApp



VMware Bare Metal Architecture

- VMware's ESX/ESXi are system virtual machines
  - Also known as "bare-metal" hypervisors.
- The hypervisor consists of the Virtual Machine Monitor (VMM), a Hardware interface layer and VMKernel.
- A service console allows administrators to manage the hypervisor, virtual machines and hardware on the system.

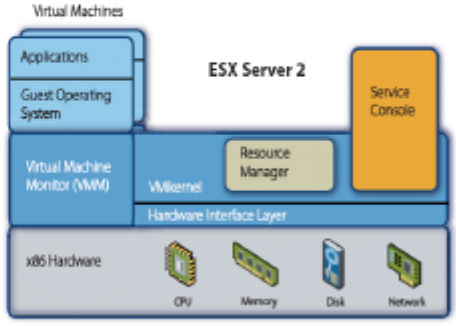
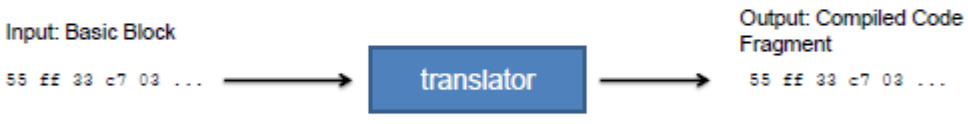


Figure 1: ESX Server architecture

Source: http://www.vmware.com/pdf/esx2_performance_implications.pdf

Processor Virtualization in VMware

- VMWare uses dynamic binary translation
  - Allows code to run on the hardware at near-native speeds.
  - Except for privileged instructions such as traps, interrupts etc.



- Privileged instructions are replaced with a controlled emulation sequence handled by the Hypervisor.
- Further performance improvements
  - Translation of a block of instructions (basic block)
  - Code caching

Multi-Processor Virtualization in VMWare

- VMware features multiple mechanisms to enable efficient scheduling of vCPUs on pCPUs
  - Enables efficient multiplexing of all the vCPUs on a given system, especially if |vCPUs| > |pCPUs|

- *Proportional-share based algorithm*
  - Each VM is allocated ***shares***, by default assignment is 1000 shares per vCPU.
  - Each VM also has an associated ***reservation*** and ***limit*** which are 0 and unlimited respectively by default.
  - The scheduler calculates if each VM is fully utilizing its resources. VMs which do not fully utilize their allocated resources will be accorded higher priority.
  - This way, the scheduler is also designed for fairness as individual VMs cannot constantly consume the Host CPU, their priority will be dynamically adjusted as they receive more CPU time.

Multi-Processor Virtualization in VMWare (contd.)

- *Co-scheduling*
  - Multi-threaded applications usually require all the vCPUs to be active simultaneously in order to make progress during execution
  - Co-scheduling attempts to schedule all the vCPUs of a VM together, as to maximize the performance of multi-threaded applications within the VM

- *CPU Topology Aware Load-Balancing*
  - Balances the vCPU load among multiple processor systems
  - When a vCPU is migrated to another processor, it does no longer has its working set in cache.
  - The vCPU can continue executing, but will have to fetch its working set from memory and *warm-up* the on-chip cache.
  - Frequent migration of the vCPUs across processors can be detrimental to performance.
  - VMware will migrate vCPUs only if they have not consumed too many CPU resources and warmed up the cache.

Memory Virtualization in VMWare

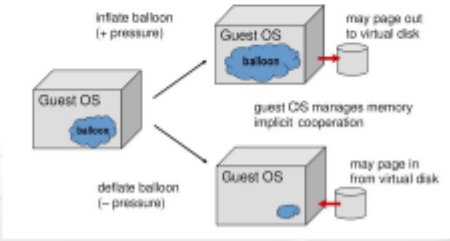
- VMWare virtualizes memory with the following goals:
  - Accurate control over memory resources
  - Efficient over-commitment of memory
  - Exploitation of memory sharing opportunities

- Guest OSes tend to utilize all the memory that is available to them.
  - It's difficult to over-commit memory resources efficiently if the Guest OS does not release memory resources from time to time.

- There are specific mechanisms in VMWare memory management to reclaim memory from guests OSes
  - *Memory Ballooning*
  - *Transparent swapping* of memory pages by the hypervisor
  - *Content-based page sharing* to allow VMs to share identical pages of memory.

Memory Ballooning Technique

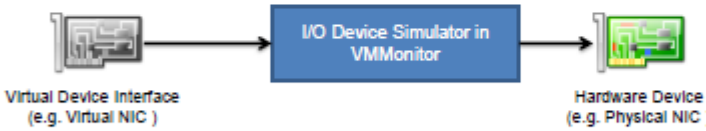
- Mechanism to reclaim memory from Guest operating systems
  - A Guest OS *memory balloon driver* is installed on each Guest OS
  - When the hypervisor needs to claim more memory from a guest OS, it signals the driver to *inflate*, it then allocates more memory, forcing the guest OS to send memory pages to disk
  - When the hypervisor decides to give back memory to the Guest, it signals the driver to *deflate* which then de-allocates memory, allowing the Guest to retrieve memory pages from disk



Source: Virtualization 101: http://labs.vmware.com/download/73/

I/O Virtualization in VMWare

- I/O devices are handled through multiple mechanisms in VMWare
- Devices can be supported via the Split Driver Model
  - VMWare uses a *Virtual Device Interface (VDI)*
  - The hypervisor can either directly access the Hardware through privileged mode

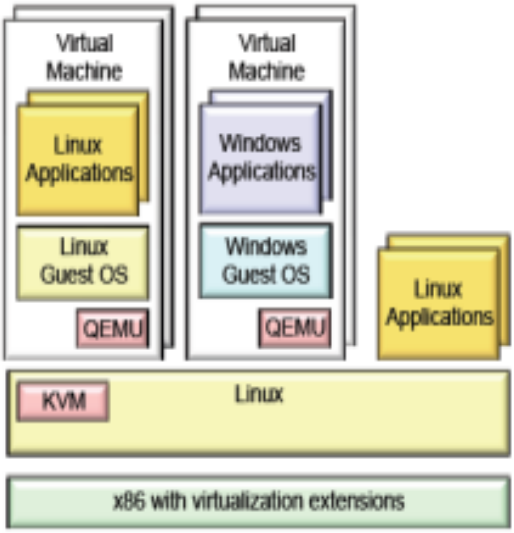


  - In addition, the hypervisor can redirect the I/O request through the native OS drivers, this allows the hypervisor to support more devices (but with reduced performance)

Comparison of Xen and VMWare

	Xen	VMware
Architecture	Bare-metal	Hosted (GSX/Workstation/Player) Bare-metal (ESX/ESXi)
Guest OS Modifications	Required if Host CPU does not support virtualization extensions.	Does not require any Guest OS modification.
CPU Virtualization Technique	Hypercalls – all privileged operations are rewritten in PVM	Binary Translation
CPU Scheduling	Credit Scheduler	Proportional-share
Memory Virtualization	Hypervisor managed page translation	Hypervisor managed page translation
I/O Virtualization	Split-Driver Approach	Split Driver Approach with Direct-Mapped in later versions.

KVM:



KVM

- **Kernel module that turns Linux into a virtual Machine Monitor**
- **Merged into the Linux kernel**

QEMU
- Emulator used for I/O device virtualization

x86 virtualizations extensions
- Intel VT-x
- AMD (AMD-V)

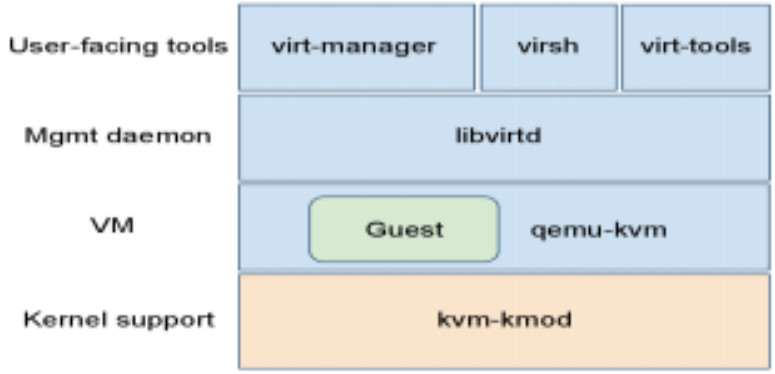
Over the past years x86 virtualization has become widespread through server consolidation and recently it is playing a role at the heart of cloud computing. KVM provides a virtualization solution with world-class performance together with the benefits of an open source platform. This post explains the key components of KVM and how they work together.

Hardware virtualization from Linux kernel

KVM is closely associated with Linux because it uses the Linux kernel as a bare metal hypervisor. A host running KVM is actually running a Linux kernel and the KVM kernel module, which was merged into Linux 2.6.20 and has since been maintained as part of the kernel. This approach takes advantage of the insight that modern hypervisors must deal with a wide range of complex hardware and resource management challenges that have already been solved in operating system kernels. Linux is a modular kernel and is therefore an ideal environment for building a hypervisor.

Full Linux hardware support for network cards, storage, and servers

Since KVM uses the Linux kernel, KVM works with network cards, storage adapters, and other hardware supported by Linux. This gives KVM excellent host hardware support that does not lag behind bare metal operating systems.



Hardware virtualization extensions provide secure and efficient way to run VM code on physical CPU

At the heart of KVM is a Linux kernel module which safely executes guest code directly on the host CPU. This is made efficient by hardware virtualization extensions, introduced in the mid-2000s by both AMD and Intel and available in almost all modern x86 processors. Virtualization extensions added a new mode of execution that allows unmodified guests to run without giving them full access to memory and other resources.

Device emulation in user space

While guest code executes directly on the host CPU in a safe manner, most I/O accesses are trapped instead of sending them directly to host devices. The guest sees an emulated chipset and PCI bus on which both emulated and pass-through adapters can be added. KVM features paravirtualized networking, storage, and memory ballooning drivers that improve efficiency of I/O and allow adjusting the amount of RAM available to a guest at run-time.

Runs with SELinux isolation

Device emulation is performed by the qemu-kvm user space process on the host. This allows the kernel module to stay lean and focus on the most performance-critical aspects while userspace device emulation emulates hardware devices in an isolated process outside of the host kernel. The sVirt feature locks down the qemu-kvm process with SELinux Mandatory Access Control so it can only access files and resources it needs and nothing more.

Secure remote management API

Management tools need to monitor and access guests that might be running on remote hosts or locally. This is done through a set of APIs and utilities that enable applications to manipulate guests and automate management tasks. Libvirt provide the language bindings and command-line utilities for developing applications and scripting common operations.

Each host runs the libvirt daemon, which provides secure remote management APIs but it can also be configured to serve locally only and not be visible over the network. The libvirt daemon maintains guest configurations across reboot and is the central point for setting up networking and storage pools.

Systems management can be added and uses libvirt API

Most administration is done with tools that use the libvirt API, especially the virsh command-line tool which presents guest and host management operations. The graphical virt-manager tool can easily manage local or remote guests. Third-party management tooling such as cloud stacks can be used for higher-level datacenter or cloud management and they typically integrate with libvirt.

11.VMWare

VMware Infrastructure is the industry's first full infrastructure virtualization suite that allows enterprises and small businesses alike to transform, manage and optimize their IT systems infrastructure through virtualization. VMware Infrastructure delivers comprehensive virtualization, management, resource optimization, application availability and operational automation capabilities in an integrated offering.

VMware Infrastructure includes the following components as :

• VMware ESX Server – A production-proven virtualization layer run on physical servers that abstract processor, memory, storage and networking resources to be provisioned to multiple virtual machines

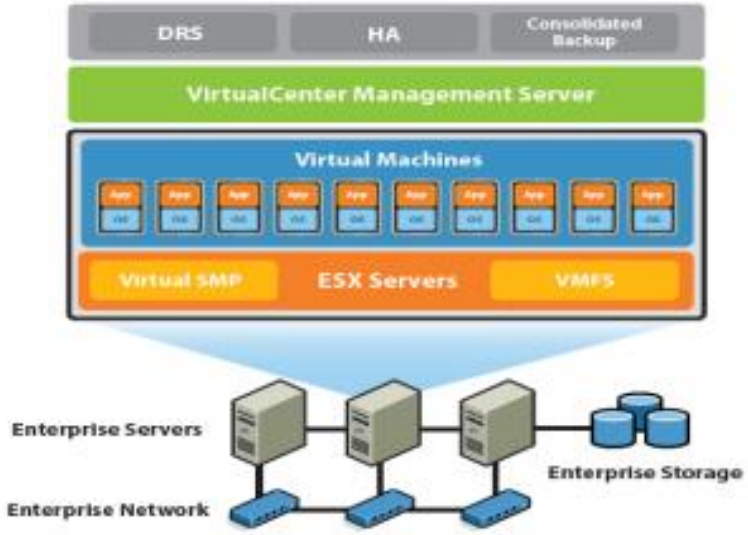
• VMware Virtual Machine File System (VMFS) – A high-performance cluster file system for virtual machines

• VMware Virtual Symmetric Multi-Processing (SMP) – Enables a single virtual machine to use multiple physical processors simultaneously

• VirtualCenter Management Server – The central point for configuring, provisioning and managing virtualized IT infrastructure

• Virtual Infrastructure Client (VI Client) – An interface that allows administrators and users to connect remotely to the VirtualCenter Management Server or individual ESX Server installations from any Windows PC

Figure 1-1: VMware Infrastructure



- Virtual Infrastructure Web Access – A Web interface for virtual machine management and remote consoles access

• VMware VMotion™ – Enables the live migration of running virtual machines from one physical server to another with zero downtime, continuous service availability and complete transaction integrity

• VMware High Availability (HA) – Provides easy-to-use, costeffective high availability for applications running in virtual machines. In the event of server failure, affected virtual machines are automatically restarted on other production servers that have spare capacity

• VMware Distributed Resource Scheduler (DRS) – Intelligently allocates and balances computing capacity dynamically across collections of hardware resources for virtual machines

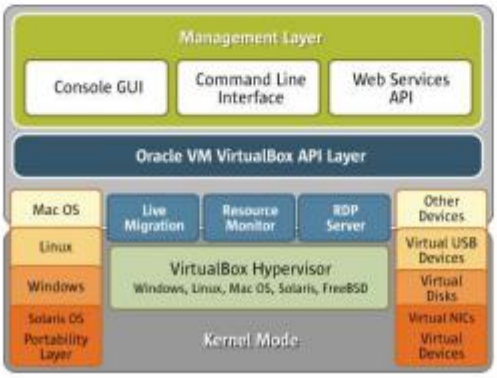
• VMware Consolidated Backup – Provides an easy to use, centralized facility for agent-free backup of virtual machines. It simplifies backup administration and reduces the load on ESX Server installations

• VMware Infrastructure SDK – Provides a standard interface for VMware and third-party solutions to access VMware Infrastructure

12. VIRTUAL BOX

Architecture

VirtualBox uses a layered architecture consisting of a set of kernel modules for running virtual machines, an API for managing the guests, and a set of user programs and services. At the core is the hypervisor, implemented as a ring 0 (privileged) kernel service. The kernel service consists of a device driver named vboxsrv, which is responsible for tasks such as allocating physical memory for the guest virtual machine, and several loadable hypervisor modules for things like saving and restoring the guest process context when a host interrupt occurs, turning control over to the guest OS to begin execution, and deciding when VT-x or AMD-V events need to be handled.



The hypervisor does not get involved with the details of the guest operating system scheduling. Instead, those tasks are handled completely by the guest during its execution. The entire guest is run as a single process on the host system and will run only when scheduled by the host. If they are present, an administrator can use host resource controls such as scheduling classes and CPU caps or reservations to give very predictable execution of the guest machine.

Additional device drivers will be present to allow the guest machine access to other host resources such as disks, network controllers, and audio and USB devices. In reality, the hypervisor actually does little work. Rather, most of the interesting work in running the guest machine is done in the guest process. Thus the host's resource controls and scheduling methods can be used to control the guest machine behavior.

In addition to the kernel modules, several processes on the host are used to support running guests. All of these processes are started automatically when needed.

- VBoxSVC is the VirtualBox service process. It keeps track of all virtual machines that are running on the host. It is started automatically when the first guest boots.
- vboxzoneacess is a daemon unique to Solaris that allows the VirtualBox device to be accessed from an Oracle Solaris Container.
- VBoxXPCOMIPCD is the XPCOM process used on non-Windows hosts for interprocess communication between guests and the management applications. On Windows hosts, the native COM services are used.
- VirtualBox is the process that actually runs the guest virtual machine when started. One of these processes exists for every guest that is running on the host. If host resource limits are desired for the guest, this process enforces those controls.

Interacting with Oracle VM VirtualBox

There are two primary methods for a user to interact with VirtualBox: a simple graphical user interface (GUI) and a very complete and detailed command-line interface (CLI). The GUI allows the user to create and manage guest virtual machines as well as set most of the common configuration options. When a guest machine is started from this user interface, a graphical console window opens on the host that allows the user to interact with the guest as if it were running on real hardware. To start the graphical interface, type the command VirtualBox at any shell prompt. On Oracle Solaris, this command is found in /usr/bin and is available to all users.

13.HYPER-V ARCHITECTURE

Hyper-V is a hypervisor-based virtualization platform and an enabling technology for one of Windows Server 2008 R2's marquee features, Live Migration. With Hyper-V version 1.0, Windows Server 2008 was capable of Quick Migration, which could move VMs between physical hosts with only a few seconds of down-time. With Live Migration, moves between physical targets happen in millisecond, which means migration operations become invisible to connected users. To find out new features and improvements in Windows Server 2008 R2 Hyper-V.

The hypervisor is the processor-specific virtualization platform that can host multiple virtual machines (VMs) that are isolated from each other but share the underlying hardware resources by virtualizing the processors, memory, and I/O devices.

Guest operating systems running in a Hyper-V virtual machine provide performance approaching the performance of an operating system running on physical hardware *if* the necessary virtual server client (VSC) drivers and services are installed on the guest operating system. Hyper-V virtual server client (VSC) code, also known as Hyper-V enlightened I/O, enables direct access to the Hyper-V "Virtual Machine Bus" and is available with the installation of Hyper-V integration services. Both Windows Server 2008 R2 and Windows 7 support Hyper-V enlightened I/O with Hyper-V integration services. Hyper-V Integration services that provide VSC drivers are also available for other client operating systems.

Hyper-V supports isolation in terms of a partition. A partition is a logical unit of isolation, supported by the hypervisor, in which operating systems execute. The Microsoft hypervisor must have at least one parent, or root, partition, running Windows Server 2008 R2. The virtualization stack runs in the parent partition and has direct access to the hardware devices. The root partition then creates the child partitions which host the guest operating systems. A root partition creates child partitions using the hypercall application programming interface (API).

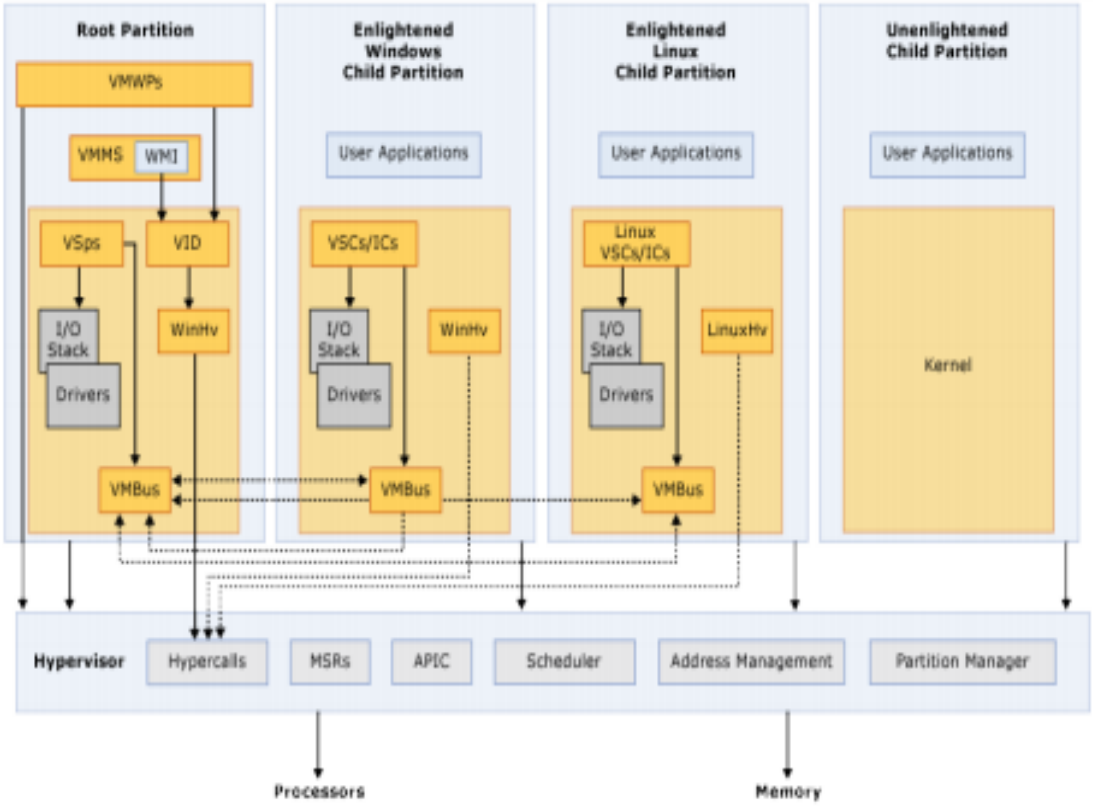
Partitions do not have access to the physical processor, nor do they handle the processor interrupts. Instead, they have a virtual view of the processor and run in a virtual memory address region that is private to each guest partition. The hypervisor handles the interrupts to the processor, and redirects them to the respective partition. Hyper-V can also hardware accelerate the address translation between various guest virtual address spaces by using an Input Output Memory Management Unit (IOMMU) which operates independent of the memory management hardware used by the CPU. An IOMMU is used to remap physical memory addresses to the addresses that are used by the child partitions.

Child partitions also do not have direct access to other hardware resources and are presented a virtual view of the resources, as virtual devices (VDevs). Requests to the virtual devices are redirected either via the VMBus or the hypervisor to the devices in the parent partition, which handles the requests. The VMBus is a logical inter-partition communication channel. The parent partition hosts Virtualization Service Providers (VSPs) which communicate over the VMBus to handle device access requests from child partitions. Child partitions host Virtualization Service Consumers (VSCs) which redirect device requests to VSPs in the parent partition via the VMBus. This entire process is transparent to the guest operating system.

Virtual Devices can also take advantage of a Windows Server Virtualization feature, named Enlightened I/O, for storage, networking, graphics, and input subsystems. Enlightened I/O is a specialized virtualization-aware implementation of high level communication protocols (such as SCSI) that utilize the VMBus directly, bypassing any device emulation layer. This makes the communication more efficient but requires an enlightened guest that is hypervisor and VMBus aware. Hyper-V enlightened I/O and a hypervisor aware kernel is provided via installation of Hyper-V integration services. Integration components, which include virtual server client (VSC) drivers, are also available for other client operating systems. Hyper-V requires a processor that includes hardware

assisted virtualization, such as is provided with Intel VT or AMD Virtualization (AMD-V) technology.

The following diagram provides a high-level overview of the architecture of a Hyper-V environment running on Windows Server 2008.



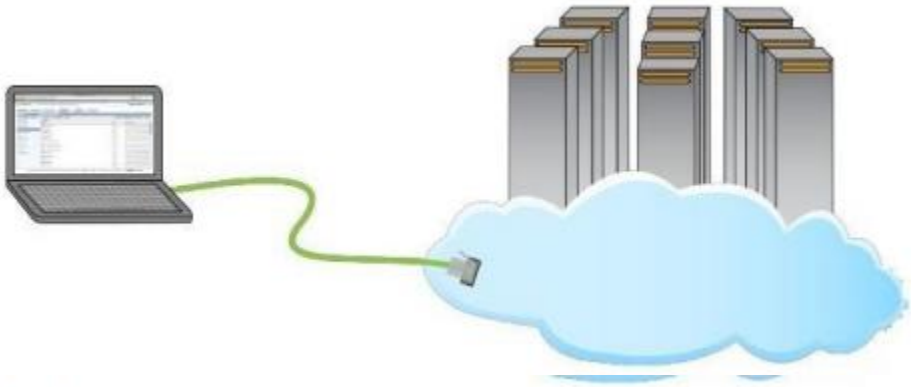
Overview of Hyper-V architecture

- APIC – Advanced Programmable Interrupt Controller. A device which allows priority levels to be assigned to its interrupt outputs.
- Child Partition – Partition that hosts a guest operating system. All access to physical memory and devices by a child partition is provided via the Virtual Machine Bus (VMBus) or the hypervisor.
- Hypercall – Interface for communication with the hypervisor. The hypercall interface accommodates access to the optimizations provided by the hypervisor.
- Hypervisor – A layer of software that sits between the hardware and one or more operating systems. Its primary job is to provide isolated execution environments called partitions. The hypervisor controls and arbitrates access to the underlying hardware.
- IC – Integration component. Component that allows child partitions to communication with other partitions and the hypervisor.
- I/O stack – Input/output stack.
- MSR – Memory Service Routine.
- Root Partition – Manages machine-level functions such as device drivers, power management, and device hot addition/removal. The root (or parent) partition is the only partition that has direct access to physical memory and devices.
- VID – Virtualization Infrastructure Driver. Provides partition management services, virtual processor management services, and memory management services for partitions.

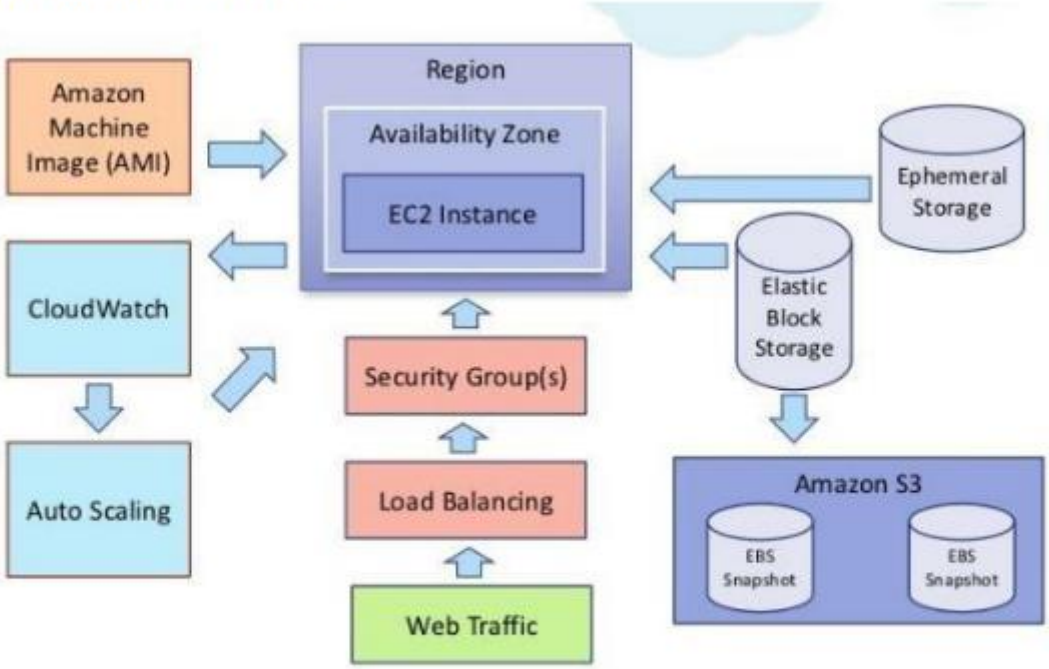
Amazon EC2

a. Introduction

- Amazon Elastic Compute Cloud (Amazon EC2) is a web service that provides secure, resizable compute capacity in the cloud.
- It is designed to make web-scale cloud computing easier for developers.
- EC2 allow users to use virtual machines of different configurations as per their requirement.
- Amazon EC2's simple web service interface allows you to obtain and configure capacity with minimal friction.
- Amazon Web Services is a cloud computing platform that provides flexible, scalable, and cost-effective technology infrastructure for businesses of all sizes around the world.
- Diagram



b. Amazon EC2 Architecture



- Load Balancing
  1. It means to hardware or software load over web servers, that improver's the efficiency of the server as well as the application.
- Elastic Load Balancing

1. It can dynamically grow and shrink the load-balancing capacity to adjust to traffic demands and also support sticky sessions to address more advanced routing needs.

- Auto Scaling
  1. It is particularly effective for those applications that fluctuateon hourly, daily, or weekly usage.
  2. Auto Scaling is enabled by Amazon CloudWatch and is available at no extra cost.
  3. AWS CloudWatch can be used to measure CPU utilization, network traffic, etc.
- Web traffic
  1. Web traffic is the amount of data sent and received by visitors to a website.
- AMI (Amazon Machine Image)
  1. An Amazon Machine Image (AMI) is a special type of virtual appliance that is used to create a virtual machine within the Amazon Elastic Compute Cloud ("EC2").
  2. It serves as the basic unit of deployment for services delivered using EC2.

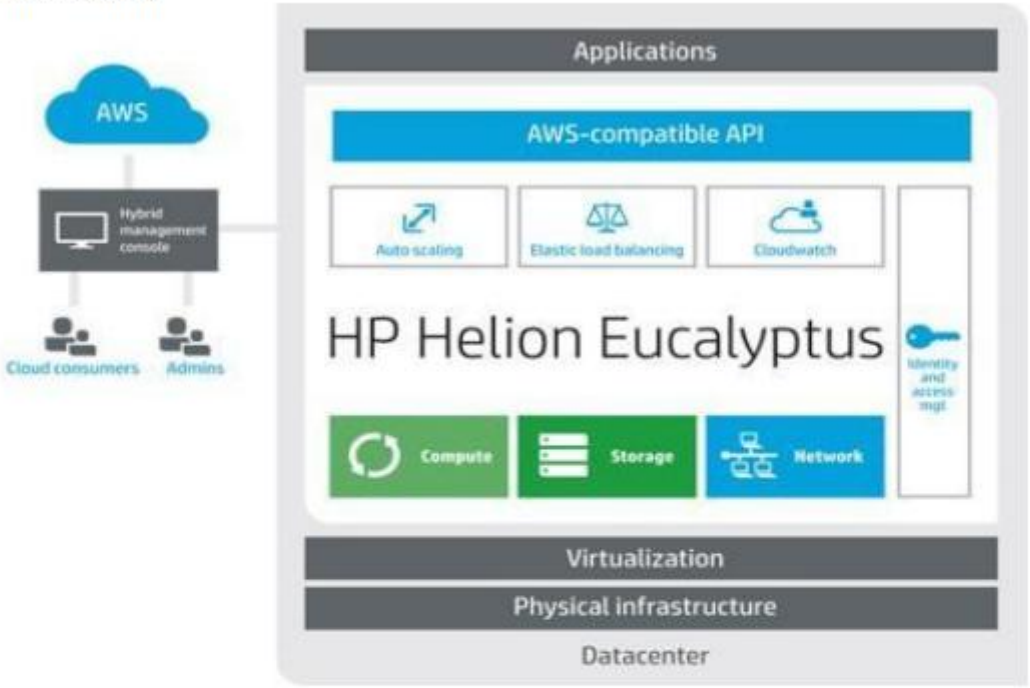
Eucalyptus

a. Introduction

- Eucalyptus is a paid and open-source computer software for building Amazon Web Services (AWS)-compatible private and hybrid cloud computing environments, originally developed by the company Eucalyptus Systems.
- Eucalyptus is an acronym for Elastic Utility Computing Architecture for Linking Your Programs To Useful Systems.

- Eucalyptus enables pooling compute, storage, and network resources that can be dynamically scaled up or down as application workloads change.
- Marten Mickos was the CEO of Eucalyptus.
- In September 2014, Eucalyptus was acquired by Hewlett-Packardand then maintained by DXC Technology.
- On-premise IaaS software
  1. Runs on top of Linux on your own hardware
  2. Numerous control components work together with nodes (virtual hosts) to create a cloud
  3. Commodity hardware focus; abstracts away physical infrastructure into a service-orientated platform
- Open Source
  1. Freely available
  2. downloads.eucalyptus.com
- Subscription-based business model
  1. Support tiers with global coverage
  2. Subscription-only add-ons
- First software release in 2008
- Incorporated in 2009

b. Architecture

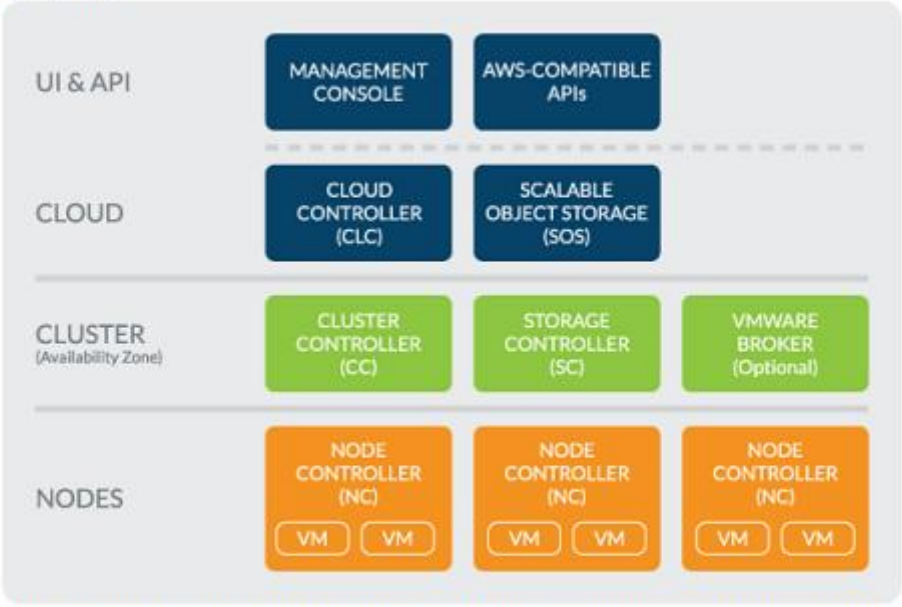


- Images – An image is a fixed collection of software modules, system software, application software, and configuration information.
- When bundled and uploaded to the Eucalyptus cloud, this becomes a Eucalyptus machine image (EMI).
- Instances – When an image is put to use, it is called an instance. The configuration is executed at runtime, and the Cloud Controller decides where the image will run, and storage and networking is attached to meet resource needs

- IP addressing – Eucalyptus instances can have public and private IP addresses. An IP address is assigned to an instance when the instance is created from an image.
- Security – TCP/IP security groups share a common set of firewall rules. This is a mechanism to firewall off an instance using IP address and port block/allow functionality.
- Networking – There are three networking modes. In Managed Mode Eucalyptus manages a local network of instances, including security groups and IP addresses.
- Access Control – A user of Eucalyptus is assigned an identity, and identities can be grouped together for access control.
- Autoscaling – Allows application developers to scale Eucalyptus cloud resources up or down in order to maintain performance and meet SLAs. With auto-scaling, developers can add instances and virtual machines as traffic demands increase.
- Elastic Load Balancing – A service that distributes incoming application traffic and service calls across multiple Eucalyptus workload instances, providing greater application fault tolerance.
- CloudWatch – A monitoring tool similar to Amazon CloudWatch that monitors resources and applications on Eucalyptus clouds. It enables cloud users to view, collect and analyze metrics of their could resources.
- CloudWatch Alarm - currently helps cloud users to take decisions on the resources (e.g instances, EBS volumes, Auto Scaling instances, ELBs) automatically based on the rules defined by the users based on the metrics. Eucalyptus CloudWatch alarm currently works with Auto Scaling policies.

The Components

- Diagram



- The Cloud Controller (CLC) is a Java program that offers EC2-compatible interfaces, as well as a web interface to the outside world.

- Only one CLC can exist per cloud and it handles authentication, accounting, reporting, and quota management.
- The Cluster Controller (CC) is written in C and acts as the front end for a cluster within a Eucalyptus cloud and communicates with the Storage Controller and Node Controller. It manages instance (i.e., virtual machines) execution and Service Level Agreements (SLAs) per cluster.
- The Storage Controller (SC) is written in Java and is the Eucalyptus equivalent to AWS EBS. It communicates with the Cluster Controller and Node Controller and manages Eucalyptus block volumes and snapshots to the instances within its specific cluster.
- The VMware Broker is an optional component that provides an AWS-compatible interface for VMware environments and physically runs on the Cluster Controller.
- The Node Controller (NC) is written in C and hosts the virtual machine instances and manages the virtual network endpoints.

SATHYABAMA
INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)
Accredited "A" Grade by NAAC | 12B Status by UGC | Approved by AICTE
www.sathyabama.ac.in

SCHOOL OF COMPUTING

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

UNIT – IV – Cloud computing – SIT1304

UNIT 4

CLOUD COMPUTING FOR EVERYONE

Cloud data center- Energy efficiency in data center- Mobile Cloud Computing Service Model-Collaboration with services and applications-CRM Management-Project Management-Email-Online Database-Calendar Schedule-Word Processing-Presentation-Spreadsheet-Databases-Network-Social Network and Groupware

CLOUD DATA CENTERS:

A data center (or datacenter) is a facility composed of networked computers and storage that businesses or other organizations use to organize, process, store and disseminate large amounts of data. A business typically relies heavily upon the applications, services and data contained within a data center, making it a focal point and critical asset for everyday operations.

Data centers are not a single thing, but rather, a conglomeration of elements. At a minimum, data centers serve as the principal repositories for all manner of IT equipment, including servers, storage subsystems, networking switches, routers and firewalls, as well as the cabling and physical racks used to organize and interconnect the IT equipment. A data center must also contain an adequate infrastructure, such as power distribution and supplemental power subsystems, including electrical switching; uninterruptable power supplies; backup generators and so on; ventilation and data center cooling systems, such as computer room air conditioners; and adequate provisioning for network carrier (telco) connectivity. All of this demands a physical facility with physical security and sufficient physical space to house the entire collection of infrastructure and equipment.

Data center consolidation and colocation

There is no requirement for a single data center, and modern businesses may use two or more data center installations across multiple locations for greater resilience and better application performance, which lowers latency by locating workloads closer to users.

Conversely, a business with multiple data centers may opt to consolidate data centers, reducing the number of locations in order to minimize the costs of IT operations. Consolidation typically occurs during mergers and acquisitions when the majority business doesn't need the data centers owned by the subordinate business.

Alternatively, data center operators can pay a fee to rent server space and other hardware in a colocation facility. Colocation is an appealing option for organizations that want to avoid the

large capital expenditures associated with building and maintaining their own data centers. Today, colocation providers are expanding their offerings to include managed services, such as interconnectivity, allowing customers to connect to the public cloud.

Data center tiers

Data centers are not defined by their physical size or style. Small businesses may operate successfully with several servers and storage arrays networked within a convenient closet or small room, while major computing organizations, such as Facebook, Amazon or Google, may fill an enormous warehouse space with data center equipment and infrastructure. In other cases, data centers can be assembled in mobile installations, such as shipping containers, also known as data centers in a box, which can be moved and deployed as required.

However, data centers can be defined by various levels of reliability or resilience, sometimes referred to as data center tiers. In 2005, the American National Standards Institute (ANSI) and the Telecommunications Industry Association (TIA) published standard ANSI/TIA-942, "Telecommunications Infrastructure Standard for Data Centers," which defined four tiers of data center design and implementation guidelines. Each subsequent tier is intended to provide more resilience, security and reliability than the previous tier. For example, a tier 1 data center is little more than a server room, while a tier 4 data center offers redundant subsystems and high security.

Data center architecture and design

Although almost any suitable space could conceivably serve as a "data center," the deliberate design and implementation of a data center requires careful consideration. Beyond the basic issues of cost and taxes, sites are selected based on a multitude of criteria, such as geographic location, seismic and meteorological stability, access to roads and airports, availability of energy and telecommunications and even the prevailing political environment.

Once a site is secured, the data center architecture can be designed with attention to the mechanical and electrical infrastructure, as well as the composition and layout of the IT equipment. All of these issues are guided by the availability and efficiency goals of the desired data center tier.

Energy consumption and efficiency

Data center designs also recognize the importance of energy efficiency. A simple data center may need only a few kilowatts of energy, but an enterprise-scale data center installation can demand tens of megawatts or more. Today, the green data center, which is designed for minimum environmental impact through the use of low-emission building materials, catalytic converters and alternative energy technologies, is growing in popularity.

Organizations often measure data center energy efficiency through a metric called power usage effectiveness (PUE), which represents the ratio of total power entering the data center divided by the power used by IT equipment. However, the subsequent rise of virtualizationhas allowed for much more productive use of IT equipment, resulting in much higher efficiency, lower energy use and energy cost mitigation. Metrics such as PUE are no longer central to energy efficiency goals, but organizations may still gauge PUE and employ comprehensive power and cooling analyses to better understand and manage energy efficiency.

Data center security and safety

Data center designs must also implement sound safety and security practices. For example, safety is often reflected in the layout of doorways and access corridors, which must accommodate the movement of large, unwieldy IT equipment, as well as permit employees to access and repair the infrastructure. Fire suppression is another key safety area, and the extensive use of sensitive, high-energy electrical and electronic equipment precludes common sprinklers. Instead, data centers often use environmentally friendly chemical fire suppression systems, which effectively starve a fire of oxygen while mitigating collateral damage to the equipment. Since the data center is also a core business asset, comprehensive security measures, like badge access and video surveillance, help to detect and prevent malfeasance by employees, contractors and intruders.

Data center infrastructure management and monitoring

Modern data centers make extensive use of monitoring and management software. Software such as data center infrastructure management tools allow remote IT administrators to oversee the facility and equipment, measure performance, detect failures and implement a wide array of corrective actions, without ever physically entering the data center room.

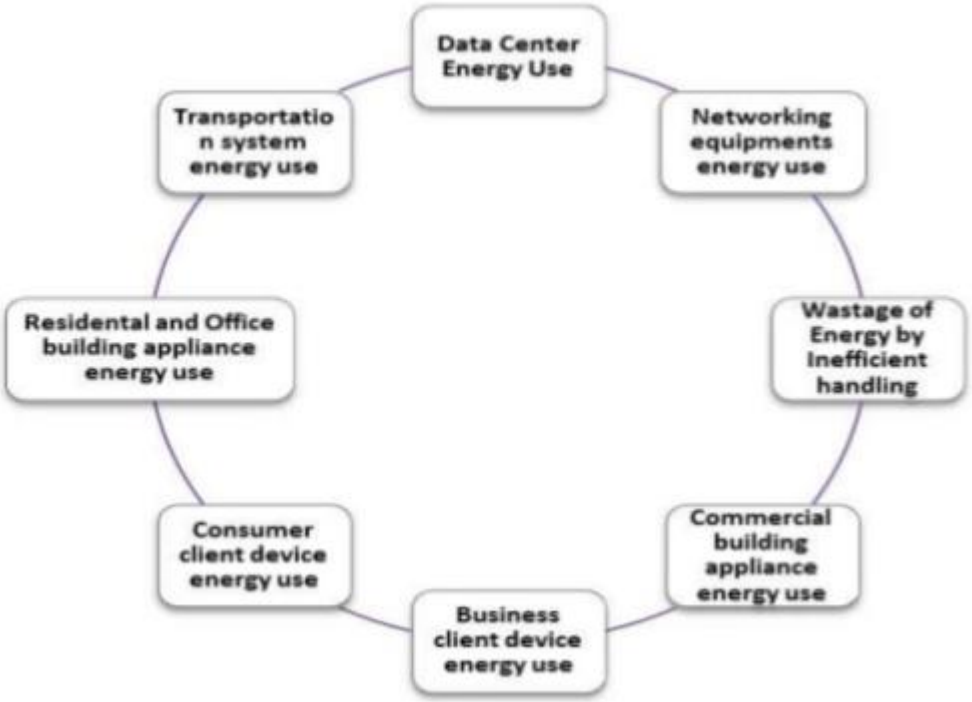
The growth of virtualization has added another important dimension to data center infrastructure management. Virtualization now supports the abstraction of servers, networks and storage, allowing every computing resource to be organized into pools without regard to their physical location. Administrators can then provision workloads, storage instances and even network configuration from those common resource pools. When administrators no longer need those resources, they can return them to the pool for reuse. All of these actions can be implemented through software, giving traction to the term software-defined data center.

Data center vs. cloud

Data centers are increasingly implementing private cloud software, which builds on virtualization to add a level of automation, user self-service and billing/chargeback to data center administration. The goal is to allow individual users to provision workloads and other computing resources on-demand, without IT administrative intervention.

It is also increasingly possible for data centers to interface with public cloud providers. Platforms such as Microsoft Azure emphasize the hybrid use of local data centers with Azure or other public cloud resources. The result is not an elimination of data centers, but rather, the creation of a dynamic environment that allows organizations to run workloads locally or in the cloud or to move those instances to or from the cloud as desired.

Efficiency Model



6

Data Center Energy Use

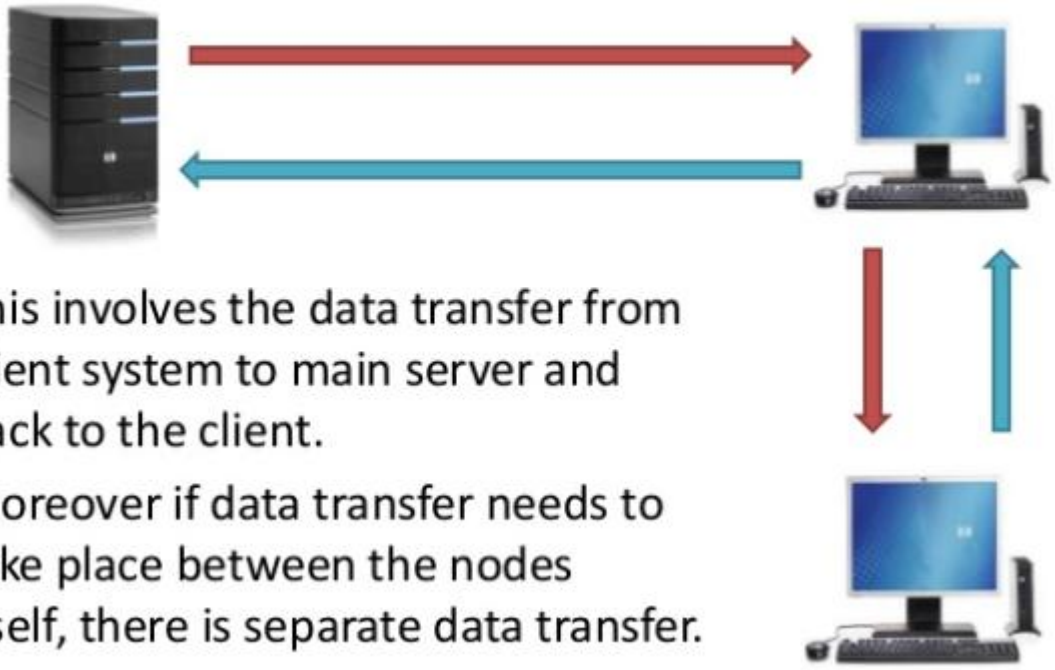
When companies use their dedicated data centers they need to have a lot of resources which primarily involves

- Air Conditioning unit.
- Dedicated Hardware.
- Resource Person
- Power Backup
- Security

Switching to the cloud would mean…

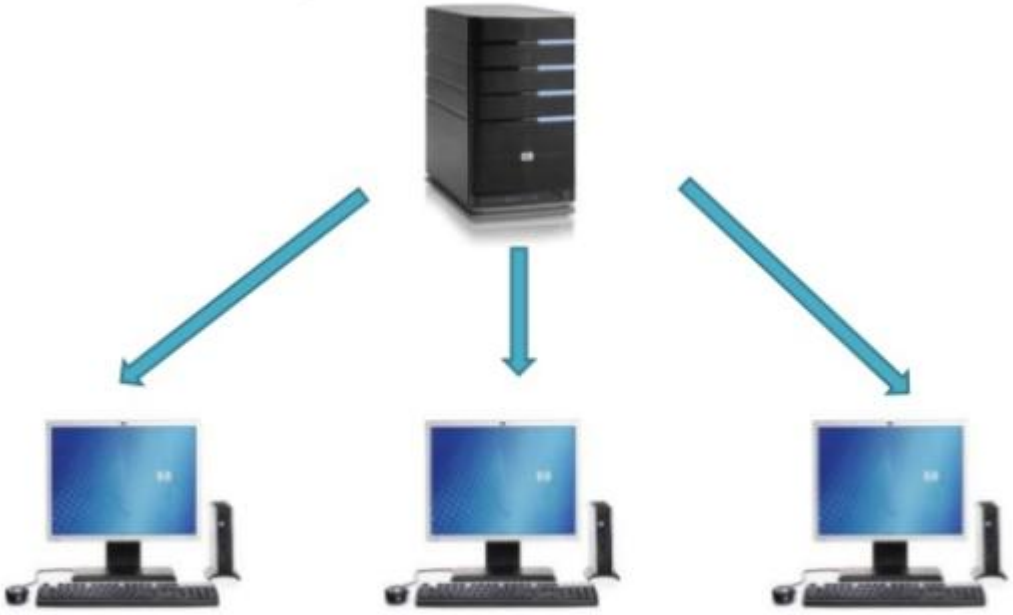
When the cloud is implemented there is only one server at one data center which is maintained by the datacenter and this server is used by many corporations.

Transportation System Energy Use



- This involves the data transfer from client system to main server and back to the client.
- Moreover if data transfer needs to take place between the nodes itself, there is separate data transfer.

Switching to the cloud would mean…



There is direct transfer from client to server thus saving energy that was wasted in transport.

10

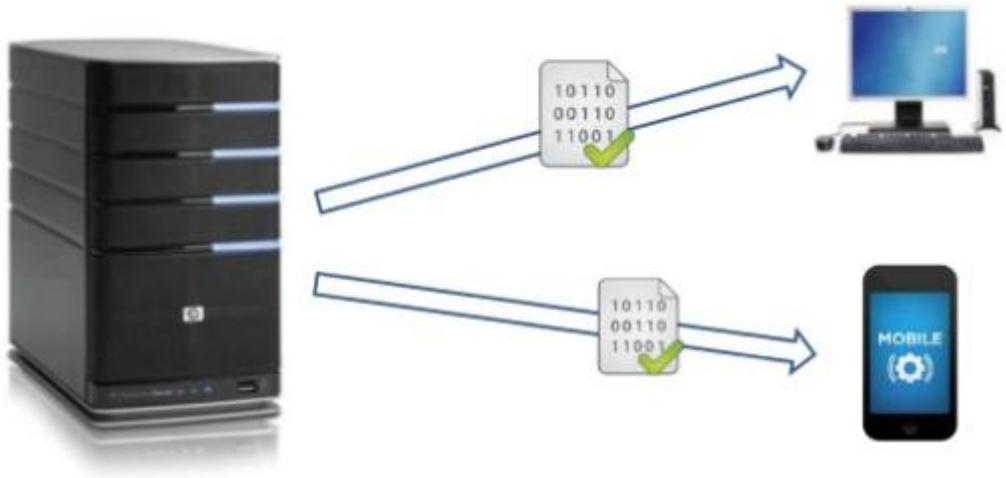
Residential and Office Building appliance energy use

- The individual electrical appliances further used leads to wastage.

No use of any extra appliances on the client side.

Business and Consumer Client device Energy Use

- This is related to the fact that if data transfer needs to be done from client to any other device the entire data is transferred.



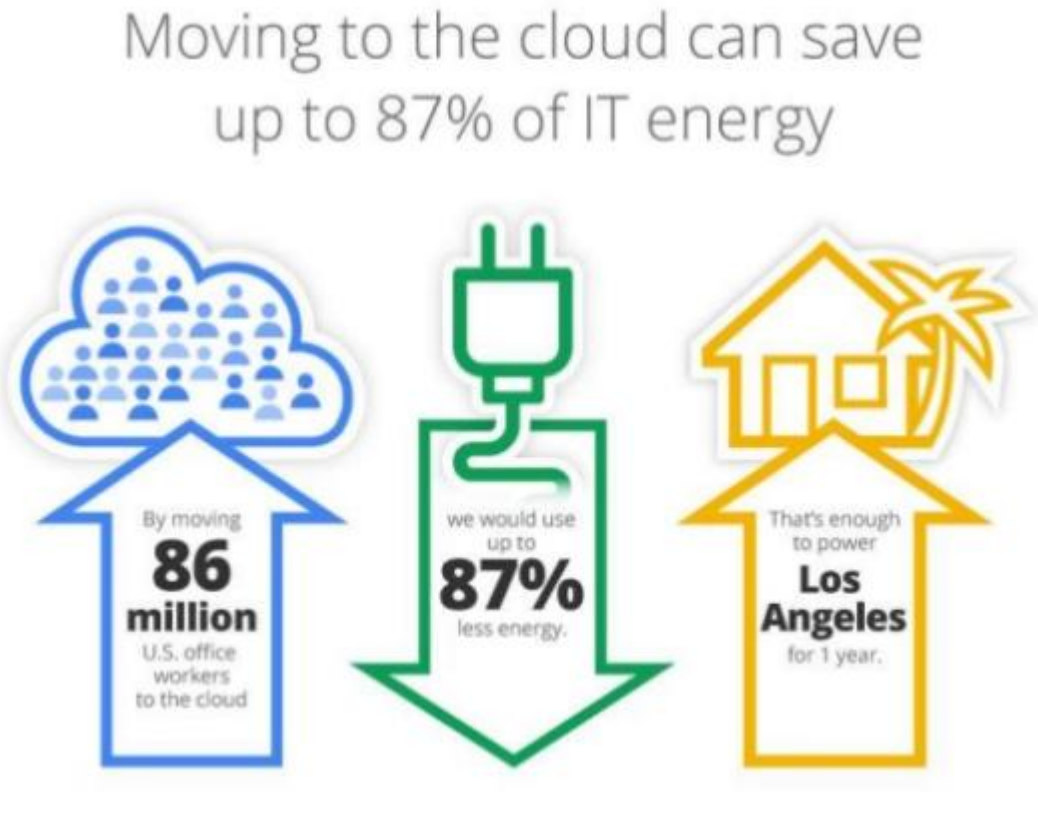
12

Whereas in cloud the data can be easily accessed thorough any device.

13

Network Equipment Energy Use

- There is a lot of networking components involved when there is a network of computers in a building.
- Instead if we go for internet connection and cloud architecture the energy can be saved.



16

- Power consumption is decreased in cloud computing.
- Private and public cloud storage services are more energy efficient than storage on local hard disk drives when files are only occasionally accessed.
- There is an overall decrease in costs involved in computing equipment for an organization.

Mobile Cloud application:

Mobile Cloud Computing, or **MCC**, merges the fast-growing **Cloud Computing Applications**market with the ubiquitous smartphone. One of the most ground-breaking blends of modern-day technologies, MCC has proved itself to be highly beneficial to all the mobile users and cloud-based service-providers as well.

In this technique, user-friendly mobile applications are developed, which are powered by and hosted using the cloud computing technology. The '*mobile cloud*' approach enables the apps

developers to build applications designed especially for mobile-users, which can be used without being bound to the operating system of the device or its capacity to store data. Here, the tasks of data-processing and data storage are performed outside the mobile devices.

The ability of MCC to allow the device to run cloud-based web-applications unlike other native apps differentiates it from the concept of 'Mobile Computing'. Here the users can remotely access the store applications and their associated data anytime on the Internet by subscribing to the cloud services. Although most devices already run a mix of web-based and native apps, the trend these days seems to be shifting more toward the services and convenience that are offered by a mobile cloud.

Researchers are putting in serious efforts in forming a strong and symbiotic platform, coined the *'Third Platform,'* that would bring together the mobile and the cloud. Experts predict this platform to revolutionize further the uprising of MCC which has enabled its users a better means to access and store their data along with latest data synchronization techniques, improved reliability and better performance. All these beneficial aspects have inspired a lot of people to consider MCC for their smart-phones.

Mobile Cloud Computing confirms the impact of certain trends and factors. Here are the factors that have had an astounding impact as far as MCC is concerned.

- **Enhanced broadband coverage:** Better connectivity is being rendered to our mobile devices via 4G, WiFi, femto-cells, fixed wireless etc.
- **Abundant Storage**: Cloud-based mobile apps have proved themselves to be more capable than any smart-phone, especially in terms of the storage space that is offered. Cloud apps' server-based computing infrastructure that is accessible through mobile interface of an app, is quite a contrast to the limited data-storage space and processing power in a mobile device.
- **Budding Technologies**: Advanced technologies like HTML5, CSS3, Hyper-Visor virtual machines for smart-phones, cloudlets and Web 4.0 etc are contributing a lot toward MCC's rising popularity.
- **Latest Trends**: Smart-phones have enabled us with 24/7 access to business applications and other collaborative services has upped the scope to increase the productivity from anywhere, at any given time.

Cloud Computing & The Future Scope of Android

Today, the popularity of this Linux-based operating system is quite apparent after looking at the massive chunk of smart-phone users relying on Android. It has a large community of developers on its platform that develop applications to increase the devices' functionality for their users. Introduction of cloud-computing on this platform has taken the user experience of Android applications to another level altogether. In fact, both, the Android app-developers and smart-phone users are benefiting from the power of cloud computing.

Various layers of Android programming model have smoothly accommodated the scope of creating secure applications that are specially developed for the cloud environment. Also, its open-source policy allows the complex cloud-computing applications to be run by the users anywhere.

For Android app-developers, it's quite different for them to develop applications in the traditional environment and in the cloud computing environment. In the traditional environment, the need to maintain complete infrastructure at the back-end shifts the focus on maintaining the environment instead of making innovative applications. Whereas, in case of apps for the cloud environment, it's the cloud-service providers who manage the infrastructure, software stack and hardware maintenance. This allows developers to write **mobile cloud applications** that profit from cloud computing and can deliver cost-benefits and other such advantages to the users.

Most of us just consider games and other daily-life simplifying apps as the only inspiration for the developers to create Android applications, but a quick reality-check on the app-market reveals that enterprise apps are catching and reaching a market share that attracts significant interest. In fact, research analysts have found mobile-centric applications and interfaces to be among the top 10 technological trends in 2018 and 2019.

Here are two well-known **examples of cloud-based Android applications**:

- **Dropbox:** Operated by Dropbox Inc., this application is a file-hosting service which offers cloud storage. It lets the users access their files in the 'Dropbox' from their Android devices, which can be synced to other computers or mobile devices.
- **Amazon Cloud Player:** One of the most popular applications on Android platform, Amazon Cloud Player is used to store and play MP3 files. Here the 'Cloud Drive' acts as a hard drive set

in the cloud. Users can play their MP3 files via the web or they can conveniently stream them on their Android devices using Amazon 'Cloud' MP3 application.

Android -Mobile Cloud Computing-Robotics – A surreal combination

Organizations and companies have changed their approach towards designing and conceptualizing new products after including cloud-computing in their calculations. Users and developers' newly-earned ability to access the immensely flexible and cost-effective power of cloud computing has helped develop services that must have seemed simply infeasible just a few years back. A perfect example of this is *Voice Search* by Google for mobile devices. 'Voice Search' has enabled users to convey a voice query and have it transcribed accurately on their devices in real time. The credit for this goes to Google's ability to use the vast amount of search data to refine and define such voice queries with cloud infrastructure. Ever since its introduction, smart voice search services have. Today, almost 25% of queries on Android devices are using it.

Robotics and cloud computing can be a great combination which shall add more capabilities and may also help in saving the battery life of the device. And by adding mobile connectivity to this gives robotics new capabilities while using lesser battery power and memory.

Reasons why cloud computing is the future of mobile devices

The interface of MCC has undeniably enabled us to accommodate videos, music files, digital images and more, right into out petite smart-phones. Here are a few reasons that explain why MCC is considered to be the future for mobile devices:

1. **Extended Battery Life:** As the major role of processing is handled by the cloud, mobile devices' battery usage is reduced automatically.
2. **Abundant Storage Space:** Enormous storage capacity that a mobile user can access happens to be the most highlighted USP of the cloud service. Mobile users shall no longer need to worry about their devices' limited storage capacity and spend money on memory cards.
3. **Improved data-synching techniques:** Cloud storage enables the user to store and manage their data by speedy data synchronization between the device and any other desktop or device chosen by

the user. This instantly benefits the users by eliminating their problems of storing all their data files and maintaining a back-up.

4. **Enhanced processing facilities:** The processor of any mobile device determines its speed and performance. However, in the case of mobile cloud computing, most of the processing is performed at the cloud level. This takes the load off the device and thereby enhances its overall performance.
5. **Superior user-experience:** In case of MCC it is always the user who benefits the most by using this platform. The wide range of benefits offered by this platform makes for an optimum productivity and an enhanced user experience.

Scope to embrace new technologies: MCC can easily adjust to the ever-evolving nature of technologies. It is capable enough to perform efficiently with all the upgrades in cloud computing methods and changes in the smart-phones' designs and features.

CRM MANAGEMENT:

Customer relationship management (CRM) describes all aspects of sales, marketing and service-related interactions that a company has with its customers or potential customers. Both business-to-consumer (B2C) and business-to-business (B2B) companies often use CRM systems to track and manage communications through the Web, email telephone, mobile apps, chat, social media and marketing materials.

CRM Information, Tracking and Analytics

Information tracked in a CRM system might include contacts, sales leads, clients, demographic or firmographic data, sales history, technical support and service requests, and more. CRM systems can also automate many marketing, sales and support processes, helping companies provide a consistent experience to customers and prospects, while also lowering their costs.

Some CRM solutions also offer advanced analytics that offer suggested next steps for staff when dealing with a particular customer or contact. Business leaders can also use these analytics to measure the effectiveness of their current marketing, sales and support efforts and to optimize their various business processes.

The Customer Relationship Management Strategy
Customer relationship management is a business strategy that enables companies to improve in the following areas:

Understanding existing customers' needs
Obtaining a 360-degree view of customers and prospects
Retaining customers through better customer experience and loyalty programs
Attracting new customers
Winning new clients and contracts
Increasing profitably
Decreasing customer management costs

Today's CRM Solutions

Many of today's most popular CRM solutions are delivered as cloud-based solutions. Because they have Web-based interfaces, these tools allow sales teams to access customer and lead information from any device in any location at any time of day. These software as a service (SaaS) solutions tend to be more user-friendly than older CRM applications, and some include artificial intelligence or machine learning features that can help organizations make better business decisions and provide enhanced support and service to their customers.

The data captured by CRM solutions helps companies target the right prospects with the right products, offer better customer service, cross-sell and up-sell more effectively, close deals, retain current customers and better understand exactly who their customers are.

The Business Benefits of CRM Systems:

The biggest benefit most businesses realize when moving to a CRM system comes directly from having all their business data stored and accessed from a single location. Before CRM systems became commonplace in the 1990s and 2000s, customer data was spread out over office productivity suite documents, email systems, mobile phone data and even paper note cards and Rolodex entries.

Storing all the data from all departments (e.g., sales, marketing, customer service and HR) in a central location gives management and employees immediate access to the most recent data when they need it. Departments can collaborate with ease, and CRM systems help organization to develop efficient automated processes to improve business processes.

Other benefits include a 360-degree view of all customer information, knowledge of what customers and the general market want, and integration with your existing applications to consolidate all business information.

PROJECT MANAGEMENT:

The truly portable office is not only possible, but is quickly becoming the norm for many businesses. Handheld devices were a big step toward the reality of working anywhere, and now cloud computing has removed the final barrier. With project management software hosted in the cloud, you can have everything you need at your fingertips, anywhere you happen to be.

Of course, every technology has its advantages and drawbacks. In this article, we'll take a look at how companies can move their project management into the cloud, who can benefit, and what pitfalls still remain for this lofty software solution. (For some background reading, check out Project Management 101.)

Project Management Software, Meet the Cloud

Project management software coordinates and automates many of the more tedious functions of managing projects. Most of these software programs include graphs and Gantt chart generation, task assignment, time sheets and milestones. More advanced versions can also regulate resources, monitor budgets, track expenses and calculate costs.

This method of working involves using programs and storing data through an internet connection. Cloud software is hosted by the provider on remote servers, rather than being physically installed on a company's machines. Microsoft OneDrive, Dropbox and your Amazon Kindle digital library are examples of cloud storage. (To learn more about the overall benefits of cloud computing, read

A Beginner's Guide to the Cloud: What It Means for Small Business.)

The Benefits of Cloud Software

Cloud software comes with plenty of advantages. One of the primary benefits for project management is super-easy sharing. Because the software is hosted in the cloud, an entire business team can access the most recent tasks, schedules and progress updates anytime, which makes cloud software ideal for travel and real-time collaboration.

Speaking of access, another great advantage to cloud software is portability. An internet connection is all it takes to access the program anytime, anywhere, from any device, including a desktop computer, laptop, smartphone or tablet.

There's also cost and time to consider. Traditional installed PM software can cost thousands, and may require hardware upgrades. Installation and upgrades are not only expensive, they also slow deployment. Cloud software, on the other hand, doesn't require installation, just a login and password to access the latest version of the software. The pricing structure is also much easier on a company's budget, as most cloud project management software is sold as subscription-based software as a service (SaaS). This means low monthly payments instead of hundreds or thousands upfront. In addition, most project management software doesn't require a long-term commitment.

In a nutshell, using the cloud deals with the following challenges, which all kinds of companies face, whether they work in architecture, construction or telecommunications:

Distribution delays, time-zone problems

With cloud service, everyone can gain access to company data from a laptop, PC, smartphone or tablet from anywhere there's a reliable internet connection.

Project collaboration

Forget cumbersome emails, employees can work on a project over the cloud and submit changes that will be available to the rest of the group in minutes. This allows disparate employees to work in near real time, just as if they were in the same room.

Backup

By having documents in the cloud, companies are protected against hardware and software failure.

Unlimited storage space

The cloud never runs out of space and is accessible from almost anywhere. This allows companies to archive files, allowing team members to continue to access them in the future, even remotely.

What's the Risk?

There's always a catch, right? Despite its strengths, cloud project management software does have a few disadvantages. Chief among them is security, which is an inherent risk for any online transaction. (Read more about security risks in The Dark Side of the Cloud.)

Ever since cloud software began to grow in popularity, companies have cited security as a primary concern. After all, using the cloud means storing all of a company's data – which may include trade secrets, sensitive customer data, and company information – on someone else's servers. Those servers could be vulnerable to hackers, viruses, or even natural disasters or physical theft.

Fortunately, cloud software vendors are aware of these risks, and most use the best available security to protect their servers and their customers' data. After all, without satisfied customers, cloud providers wouldn't be in business. In addition, cloud security has improved over time, and is likely to continue to do so as cloud providers come up with new solutions to protect their clients' data. Even so, companies should find out what security procedures and protocols are in place

before subscribing to any cloud software service.

Downtime is another potential problem. If the cloud provider runs into technical problems, its clients will not be able to access their data. According to the International Working Group on Cloud Computing Resiliency, the uptime of cloud providers ranged between 99.6 percent and 99.9 percent, for an average of 7.5 hours of unavailable time per year. This sounds pretty good, but IWGCR states that it's a long way from the 99.99 percent reliability that's required of a mission-critical system. Many large cloud software vendors – and even some smaller ones – have uptime guarantees, so companies should look for one that carries the least risk. And of course, the most crucial data should also be backed up in-house.

Where Cloud PM Software Works

While large corporations tend to stick to the more traditional installed PM software, cloud versions are being embraced by companies that lack million-dollar IT budgets. The low initial costs, minimal or non-existent IT infrastructure investment, pay-as-you-go rates and anywhere access of cloud software is ideal for:

EMAIL: Email

Email

A robust and easy-to-use webmail for **teams of all sizes**.

A focus on **productivity** allows you to set up your inbox and filing system to best support how you work, along with **intuitive functionality** like drag and drop and auto saving.

You can access your emails anywhere, anytime, and on any device with **IMAP sync** across all major mobile and desktop clients.

Calendar

A calendar that gives you complete flexibility to **create, manage and share** individual, group and resource schedules.

Configure events how you want them with options for **privacy settings, reminders, recurrences, and attachments**.

Always know where you need to be by syncing your calendar across major mobile devices and desktop clients. Easily share and receive events with **ICS files and iCal support**.

Contact

A contact system allows you to **create and access multiple address books** for yourself, your colleagues, and your organization overall.

You can configure what information is collected for a contact to best fit your needs, as well as setting up advanced features such as **auto contact creation** when receiving emails from new contacts.

Also, **easily import** contacts or whole address books, including from your organization's directory.

FileShare

A completely secure filesharing system to support everything from casual sharing to providing the **confidentiality and traceability** for teams and organizations that demand it.

Take advantage of sharing features such as password protection, temporary accounts, expiring links, and maintain a **detailed log of who has downloaded and modified your files**.

Additional security features include **digital signatures, file encryption and built-in antivirus** support to protect yourself and your organization.

WORD PROCESSING:

Just about everyone who uses a computer uses a word processing program. You use your word processor—most likely some version of Microsoft Word—to write memos, letters, thank you notes, fax coversheets, reports, newsletters, you name it. The word processor is an essential part of our computing lives.

The solution, believe it or not, lies in the clouds—in the form of a web-based word processor. That's right, there are a number of web-based replacements for Microsoft's venerable Word program. All of these programs let you write your letters and memos and reports from any computer, no installed

software necessary, as long as that computer has a connection to the Internet. And every document you create is housed on the web, so you don't have to worry about taking your work with you. It's cloud computing at its most useful, and it's here today

1. How Web-Based Word Processing Works

Microsoft Word is a software program that is installed on your computer's hard disk. Web-based word processors, in contrast, are hosted in the cloud, not on your hard drive—as are the documents you create with these applications.

a) Benefits of Web-Based Word Processors

- the most obvious benefit of using a cloud service is that your documents can be accessed wherever you are, from any PC.

- being web based, you can easily share your documents with others.
- being web based is that you can't lose your work

- After you've named the document you're working on, the webbased word processor saves your file on its cloud of servers. From that point on, every change you make to the document gets saved to the cloud servers automatically. Nothing gets lost if you close your web browser, navigate to another website, or even turn off your computer. Everything you do is saved

- on the web.
- web-based applications are free.
- Being free makes it easy to take for a test drive, and even easier to add to your bag of applications.

2. Exploring Web-Based Word Processors

There are a half-dozen or so really good web-based word processing applications, led by the ever-popular Google Docs.

Google Docs

- Google Docs (docs.google.com) is the most
- popular web-based word processor available today.

- Docs is actually a suite of applications that also includes Google Spreadsheets and Google Presentations; the Docs part of the Docs suite is the actual word processing application.

- Like all things Google, the Google Docs interface is clean and, most important, it works well without imposing a steep learning curve.

- Basic formatting is easy enough to do, storage space for your documents is generous, and sharing/collaboration version control is a snap to do.

  - oThis is the home page for all the Docs applications (word processing, spreadsheets, andpresentations); all your previously created documents are listed on this page.

PRESENTATIONS:

Preparing Presentations Online

Working with an online presentation application is no different from working with any other web-based application. Users from multiple locations can access the presentation directly from any Internet-connected computer, making it easy to assemble a presentation via group collaboration.

BrinkPad (www.brinkpad.com) is a Java applet that works inside any web browser. It lets you create, save, and publish your presentations and slide shows on the web.

Empressr

Empressr (www.empressr.com) offers more functionality than BrinkPad and similar applications, via an interface that should be somewhat familiar to PowerPoint users. You can insert text, shapes, tables, or charts onto any slide. You can even create custom slide backgrounds.

Google Presentations

Users can create new presentations and open existing ones from the main Google Docs page (docs.google.com). Open a presentation by clicking its title or icon. Create a new presentation by selecting New, then Presentation. Your presentation now opens in a new window on your desktop.

Preezo

Preezo even looks a lot like Google Presentations. You get the obligatory slide sorter in the leftmost pane, the current slide in the main window, and all available editing and formatting options in a toolbar and series of pull-down menus. Slide transition effects include wipes, fades, splits, and pushes.

Presentation Engine

Presentation Engine (www.presentationengine.com) is an advanced presentation program with an eye toward snazzy graphics and transition effects. It offers a level of graphics sophistication not found in competing online applications.

PreZentit

PreZentit (www.prezentit.com) is a slick-looking application that offers features not found with competing programs.

SlideRocket (www.sliderocket.com) is one of the newest web-based presentation applications.

ThinkFree Show

ThinkFree Show (www.thinkfree.com) is the presentation component of ThinkFree's suite of office applications.

Thumbstacks

Thumbstacks (www.thumbstacks.com) is a bare-bones online presentation program.

Zoho Show

Like Google Presentations, Zoho Show lets you create good-looking text-based slides, but that's about all.

SPREADSHEETS:

Several web-based spreadsheet applications are worthy competitors to Microsoft Excel. Chief among these is Google Spreadsheets, which we'll discuss first, but there are many other apps that also warrant your attention. If you're at all interested in moving your number crunching and financial analysis into the cloud, these web based applications are worth checking out.

Google Spreadsheets

Google Spreadsheets was Google's first application in the cloud office suite first known as Google Docs & Spreadsheets and now just known as Google Docs. (It's also the only app in the suite that Google developed in-house.) As befits its longevity, Google Spreadsheets is Google's most sophisticated web-based application. You access your existing and create new spreadsheets from the main Google Docs page (docs.google.com). To create a new spreadsheet, click the New button and select Spreadsheet; the new spreadsheet opens in a new window.

EditGrid

EditGrid interface is a near-replica of pre-2007 Excel, down to the tabbed sheets, pull-down menus, and toolbars. You even get 50+ keyboard shortcuts, identical to those in Excel.

eXpresso

The spreadsheet itself isn't much to write home about; there are no functions, charts, or advanced formatting options. The collaboration features, however, include notes, email communication, online chat, and sophisticated sharing capabilities. This application is perhaps

best used to first import existing Excel spreadsheets and then share them using eXpresso's collaboration tools.

Num Sum

the program includes no built-in functions and only three rudimentary chart types (line, bar, area), which makes it less than ideal for advanced spreadsheet users.

PeepelWebSheet

WebSheet offers a nice selection of functions and formatting options, it doesn't have a chart feature, so you can't display your data visually.

Sheetster

Sheetster (www.sheetster.com) is a web-based spreadsheet written in JavaScript.

ThinkFreeCalc

ThinkFreeCalc (www.thinkfree.com), like its ThinkFree Write sibling, is a Javabased online application.

Zoho Sheet

Zoho Sheet (sheet.zoho.com) is Zoho's web-based spreadsheet application. Like all Zoho apps, this one is full featured with great sharing and collaboration features.

Preparing Presentations Online

Working with an online presentation application is no different from working with any other web-based application. Users from multiple locations can access the presentation directly from any Internet-connected computer, making it easy to assemble a presentation via group collaboration.

COLLABORATING ON DATABASES

- It used to be that the big three office applications were word processing, spreadsheets, anddatabases.

- It's possible that databases have fallen from the top three replaced either by email or presentation applications, depending on how you look at things), but large businesses especially still have plenty of need for database management applications.

- Today, thanks to cloud computing technology, the underlying data of a database can be stored in the cloud—on collections of web servers—instead of housed in a single physical location.

- This enables users both inside and outside the company to access the same data, day or night, which increases the usefulness of that data.

- It's a way to make data universal.

1. Understanding Database Management

Before we start looking at web-based database management applications, it helps to know a little about how databases themselves work. Although it's convenient to think of a database as simply a collection of data, there's more to it than just that.

a) How Databases Work

- A database does many of the same things that a spreadsheet does, but in a different and often more efficient manner. In fact, many small businesses use spreadsheets for database-like functions.

- Think of it this way. If a spreadsheet is a giant list, a database is a giant filing cabinet. Each —filing cabinetl is actually a separate database file, and contains individual index cards (called *records*) filled with specific information (arranged in *fields*).

- You can use a database application to create and store anything that includes a large amount of data. For example, you can create a database that contains all your favorite recipes or the contents of your CD or video collection.

- For businesses, databases tend to house large amounts of granular data—information about customers, employees, and sales. A database management program not only stores this data

but also automates data entry, retrieval, and analysis. Many businesses build custom applications around their databases, so that the database itself becomes somewhat transparent. Users see

- only the front end that pulls information from the database.

b) How Online Databases Work

- A local database is one in which all the data is stored on an individual computer.

- A networked database is one in which the data is stored on a computer or server connected to a network, and accessible by all computers connected to that network. Finally, an online or web-based database stores data on a cloud of servers somewhere on the Internet, which is accessible by any authorized user with an Internet connection.

- The primary advantage of a web-based database is that data can easily be shared with a large number of other users, no matter where they may be located. When your employee database is in the cloud, for example, the human resources department in your Alaska branch can access employee information as easily as can the HR staff in Chicago—as can HR managers traveling across the country to various college job fairs.

- And, because the data itself is stored in the cloud, when someone at one location updates a record, everyone accessing the database sees the new data. Synchronization is not an issue.

- With these advantages in mind, most online databases are oriented toward quick information sharing among members of workgroups who've assembled to attack a project for a month or two. When accessing data in this manner, ease of use is paramount, which most of these cloud applications address with simple and intuitive interfaces.

SOCIAL NETWORKS

There are three main categories: web email services, instant messaging services, and web conferencing tools. Groups located anywhere in the world can use these tools to communicate with other group members—and further their collaboration on group projects.

2. Evaluating Web Mail Services -Traditional email is anything but cloud based. The type of email program you probably have installed on your PC uses a protocol called the Post Office Protocol (POP). POP email requires the use of a dedicated email client program, such as microsoft Outlook or Outlook Express, and—at the ISP level—email servers to send and receive messages. The problem with traditional POP email is that you're tied to the client program installed on your PC. The messages you receive are stored on that PC, and you usually can't access them when you're traveling or away from that PC. There are none of the —anytime, anywherel advantages you're used to with cloud-based services. Fortunately, there is a better way to manage your email—in the form of webbased email services, also known as web mail or HTTP email. Unlike traditional POP email, web mail can be accessed from any PC using any web browser, and all your messages are stored on the web, not locally. It's just like a cloud service; no special software required. This lets you retrieve and manage your email when you're out of the office or on the road. Not only is web mail more versatile than traditional POP email, it's also easier to set up. All you need to know is your user ID and password, and then you access a page that lets you view the contents of your inbox, read and reply to messages, create new messages, and (in many cases) store messages in folders. You can even, on some services, use your web mail account to access your ISP's POP email. The three largest web mail services today are hosted by Google, Microsoft, and Yahoo!

Gmail

Google's web mail service is called Gmail (mail.google.com), and at first blush it looks a lot like the other services we discuss in this chapter. Gmail is free, it lets you send and receive email from any web browser, and the interface even looks similar to its competitors. But Gmail offers a few unique features that set it apart from the web-based email crowd. First, Gmail doesn't use folders. That's right, with Gmail you can't organize your mail into folders, as you can with the other services. Instead, Gmail pushes the search paradigm as the way to find the messages you want—not a surprise, given Google's search-centric business model.

Gmail does, however, let you —tag‖ each message with one or more labels. This has the effect of creating virtual folders, as you can search and sort your messages by any of their labels. In addition, Gmail groups together related email messages in what Google calls conversations. A conversation might be an initial message and all the replies (and replies to replies) to that message; a conversation might also be all the daily emails from a single source that have a common subject, such as messages from subscribed-to mailing lists. Like most of the other services we discuss here, Gmail is a free service; all you have to do is sign up for an account. Of course, if you already have an account for any other Google service, that account can serve as your Gmail account. When you sign up for your Gmail account, you get assigned your email address (in the form of name@gmail.com) and you get access to the Gmail inbox page. As of June 2008, Gmail offered 6GB of storage for users.

any web browser. Yahoo! also offers a paid service called Yahoo! Mail Plus that lets you send larger messages and offers offline access to your messages via POP email clients. Whether you use the free or the paid version, Yahoo! Mail gives you unlimited storage—which means you can effectively use Yahoo! Mail as an online backup or file-storage system. All you have to do is email yourself those files you want to store, and then place those messages (with attachments) in your designated storage folder.

Windows Live Hotmail

Hotmail was one of the first web-based email services, and it's still one of the largest. But it's not called —Hotmail‖ anymore; Microsoft has moved it into its Windows Live suite of online services and now calls it Windows Live Hotmail. Like most web mail services, Hotmail (we're going to call it by its old, shorter name) can be accessed from any web browser on any PC anywhere in the world, for free. Microsoft gives you 5GB of storage, not quite as much as you get with Gmail (6GB) or Yahoo! Mail (unlimited).

Apple MobileMe Mail

As part of its MobileMe suite of applications, Apple offers MobileMe Mail (www.me.com). What makes MobileMe Mail unique is that it's not limited to just computer users; you can also send and receive emails from your Apple iPhone or iPod touch, via Wi-Fi Internet or cellular network. OtherWeb Mail Services Gmail, Yahoo! Mail, and Windows Live Hotmail are the three largest web mail services (and MobileMe Mail promises to be a competitor), but there are literally hundreds more. Besides these big providers, there are dozens of independent web mail services, plus a plethora of topic-specific websites that offer (among other content and services) their own branded HTTP email. In addition, just about every cloud service provider, such as Zoho, offers web mail as part of its suite; web mail is also part of most web-based desktops.

So if you're looking for a web mail service and don't want to go with one of the big three, here's a short list of some of the other major providers to check out: _ AOL Mail (mail.aol.com) _ BigString (www.bigstring.com) _ Excite Mail (mail.excite.com) _ FlashMail (www.flashmail.com)

GMX Mail (www.gmx.com) _ Inbox.com (www.inbox.com) _ Lycos Mail (mail.lycos.com) _ Mail.com (www.mail.com) _ Zoho Mail (zoho.mail.com)

Evaluating Instant Messaging Services

Email is just one way to communicate online. For many users, instant messaging is a better way to talk; it's more immediate, because you can send text messages in real time to your friends and coworkers. No more waiting for people to respond to your emails—when both parties are online at the same time, it's just like having aoneon-one conversation! Technology-wise, email works a little differently from most Internet applications— and quite different from the cloud services we've been discussing throughout this book. Email (both web based and POP), Usenet, and the World Wide Web operate via a traditional client/server model, with most of the heavy lifting done via a network of dedicated servers. For example, your POP email is stored on and managed by an email server, while all the pages on the web are hosted on millions of individual web servers. Instant messaging, however, doesn't use servers at all. When you send an instant message to another user, that message goes directly to that user's PC; it's not filtered by or stored on any servers. The technical name for this type of connection is peer-to-peer (P2P), because the two computers involved are peers to each other. All instant messaging needs to work is a piece of client software (one for each computer

involved, of course) and the IP addresses of each computer. The messages go directly from one IP address to another, with no servers in the middle to slow things down. (Naturally, the data must still make its way through numerous routers to get to the other PC, but that's part and parcel of any Internet-based application.)

There are several big players in the instant messaging market today, including America Online (with both AOL Instant Messenger and ICQ), Google (Google Talk), Microsoft (Windows Live Messenger), and Yahoo! (Yahoo! Messenger). Unfortunately, most of these products don't work well (or at all) with each other. If you're using Yahoo! Messenger, for example, you can't communicate with someone running AOL Instant Messenger. That means you'll want to use the IM program that all your friends and coworkers are using—so find that out before you download any software. AOL Instant Messenger The most-used instant messaging program is AOL Instant Messenger (www.aim.com), also known as AIM. AOL claims more than 60 million users, which makes it the number-two IM service today, second only to Yahoo! Messenger. For whatever reason,

SATHYABAMA
INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)
Accredited "A" Grade by NAAC | 12B Status by UGC | Approved by AICTE
www.sathyabama.ac.in

SCHOOL OF COMPUTING

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

UNIT – V – Cloud computing – SIT1304

UNIT 5

CLOUD SECURITY

Cloud security-Security threats and solutions in cloud-Auditing protocol-Dynamic Auditing-Storage Security-Privacy preserving-Full homo morphic encryption-Big data security-Cloud availability-DoS attack-Full tolerance management in cloud computing-cloud computing in India

Cloud security is the protection of data, applications, and infrastructures involved in cloud computing. Many aspects of security for cloud environments (whether it's a public, private, or hybridcloud) are the same as for any on-premise IT architecture.

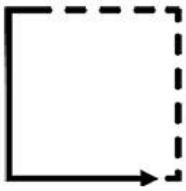
High-level security concerns—like unauthorized data exposure and leaks, weak access controls, susceptibility to attacks, and availability disruptions—affect traditional IT and cloud systems alike. Like any computing environment, cloud security involves maintaining adequate preventative protections so you:

- Know that the data and systems are safe.
- Can see the current state of security.
- Know immediately if anything unusual happens.
- Can trace and respond to unexpected events.

Why cloud security is different

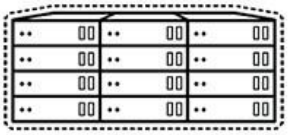
While many people understand the benefits of cloud computing, they're equally deterred by the security threats. We get it. It's hard to wrap your head around something that exists somewhere between amorphous resources sent through the internet and a physical server. It's a dynamic environment where things are always changing—like security threats. The thing is that, for the most part, cloud security is *IT security*. And once you understand the specific differences, the word "cloud" doesn't feel as insecure.

Dissolving perimeters



Security has a lot to do with access. Traditional environments usually control access using a perimeter security model. Cloud environments are highly connected, making it easier for traffic to bypass traditional perimeter defenses. Insecure application programming interfaces (APIs), weak identity and credentials management, account hijacks, and malicious insiders may pose threats to the system and data. Preventing unauthorized access in the cloud requires shifting to a data-centric approach. Encrypt the data. Strengthen the authorization process. Require strong passwords and 2 factor authentication. Build security into every level.

Everything is now in software



"Cloud" refers to the hosted resources delivered to a user via software. Cloud computing infrastructures—along with all the data being processed—are dynamic, scalable, and portable. Cloud security controls need to respond to environmental variables and accompany workloads and data while at rest and in transit, either as inherent parts of the workloads (e.g. encryption) or dynamically through a cloud management system and APIs. This helps to protect cloud environments from system corruption and data loss.

Sophisticated threat landscape

Sophisticated threats are anything that negatively impacts modern computing which—of course—includes the cloud. Increasingly sophisticated malware and other attacks like Advanced Persistent Threats (APTs) are designed to evade network defenses by targeting vulnerabilities in the computing stack. Data breaches can result in unauthorized information disclosure and data tampering. There's no clear solution to these threats, except that it's your responsibility to stay on top of the cloud security practices that are evolving to keep up with emerging threats.

SECURITY THREATS AND SOLUTIONS:

In today's technology-based world, chances are you've heard of the "cloud" or "cloud storage." In fact, nearly 90-percent of all U.S. broadband users acknowledge that cloud computing exists, yet just 29-percent knowingly utilize the technology. However, with cloud technology becoming more applicable and relevant with each passing day, these statistics are expected to soar over the coming years and more people will turn to the cloud for storing data, files, media and more.

Though the prevalence of the cloud has added ease, accessibility, and reliability to computing, it's brought with it a slew of new threats and risks. As you begin to rely more and more on cloud computing, be aware of these 5 security threats and how you can fix them.

Absence of Adequate Encryption

The risk:

Without proper encryption, cloud computing is often subjected to monitoring and eavesdropping by external sources like common Man-in-the-Middle (MitM) attacks. By impersonating critical aspects of the cloud computing process and bypassing authentication methods, hackers are able to access and hijack critical, "secure" data.

The solution:

You can prevent this risk by encrypting communications and data with proper SSL/TLS code. By implementing cryptographic protocols and endpoint authentication, and by using a reliable proxy server, you can shield your cloud data from random attacks.

Feeble Security Management

The risk:

Cloud systems with weak security management will obviously be more subjected to external threats and risks. Poor security includes things such as failure to engage strong authentication measures, identity management and authorization procedures and more.

The solution:

In order to reliably manage the security of your cloud computing system, you have to synchronize data and use proper identity management services so that all systems in the cloud are able to work together. You should also establish an in-house backup system to store critical data that might be inappropriate for the cloud alone. Lastly, it's critical for you to have real-time, end-to-end visibility so that you'll be able to see each level of the network and quickly recognize and resolve security weaknesses.

Poor Data Redundancy

The risk:

Without adequate data redundancy, your entire cloud – and therefore your entire business – will be at risk. You'll be even more vulnerable to external threats if you fail to have additional copies of your data distributed across data centers across the network.

The solution:

In order to implement redundancy, you must use various data centers from both a single cloud provider, multiple public cloud providers, and a hybrid cloud. Distributing data this way will ensure that there is minimal damage in the event of an outage or breach in one data center.

Data Leak

The risk:

Data can be easily compromised, changed, destroyed, or maliciously accessed without the proper security protocols in place. Cloud security service providers that operate in worker-driver BYOD environments are at even more of a risk of data leakages because their systems can be easily breached.

The solution:

When working with a cloud provider, be sure you communicate openly and ask important questions regarding data leak coverage, data loss prevention and more. Never assume this type of security is applied and always ask for agreements in writing. Implemented security measures should always be able to stand up to malfunctions, breaches, errors and more.

Audit Protocols

Audit protocols assist the regulated community in developing programs at individual facilities to

evaluate their compliance with environmental requirements under federal law. The protocols are

intended solely as guidance in this effort. **The regulated community's legal obligations are determined by the terms of applicable environmental facility-specific permits, as well as underlying statutes and applicable federal, state and local law.**
Environmental audit reports are useful to a variety of businesses and industries, local, state and federal government facilities, as well as financial lenders and insurance companies that need to assess environmental performance. The audit protocols are designed for use by persons with various backgrounds, including scientists, engineers, lawyers and business owners or operators. The following protocols provide detailed regulatory checklists and are provided in an easy to understand question format for evaluating compliance:

Comprehensive Environmental Response, Compensation and Liability (CERCLA)

- Protocol for Conducting Environmental Compliance Audits and the Comprehensive Environmental Response, Compensation and Liability Act (12/1/98)

Clean Water Act (CWA)

- Protocol for Conducting Environmental Compliance Audits under the Stormwater Program (1/15/05) Guidance including detailed regulatory checklists to to assess environmental performance in the stormwater program.

- Protocol for Conducting Environmental Compliance Audits for Municipal Facilities under US EPA's Wastewater Regulations (12/1/00).

Emergency Planning and Community Right-to-Know Act (EPCRA)

- Protocol for Conducting Environmental Compliance Audits under the Emergency Planning and Community Right-to-Know Act and CERCLA Section 103 (3/1/01)

Federal Insecticide, Fungicide and Rodenticide Act (FIFRA)

- Protocol for Conducting Environmental Compliance Audits under the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA) (9/1/00)

Resource Conservation and Recovery Act (RCRA)

- Protocol for Conducting Environmental Compliance, Audits of Facilities Regulated under Subtitle D of RCRA (3/1/00)

- Protocol for Conducting Environmental Compliance, Audits of Storage Tanks under the Resource Conservation and Recovery Act (3/1/00)

- Protocol for Conducting Environmental Compliance, Audits of Treatment, Storage and Disposal Facilities under the Resource Conservation and Recovery Act (12/1/98)

Safe Drinking Water Act (SWDA)

- Protocols for Conducting Environmental Compliance, Audits of Public Water Systems under the Safe Drinking Water Act (3/1/00)

Toxic Substances Control Act (TSCA)

- Protocol for Conducting Environmental Compliance Audits of Facilities with PCBs, Asbestos, and Lead-based Paint Regulated under TSCA (3/1/00)

Federal Facilities

- Environmental Audit Program Design Guidelines for Federal Agencies (EPA 300-B-96-011) (4/30/97)

In addition, there is a "how to" manual on designing and implementing environmental compliance auditing programs for federal agencies and facilities.

Contact Us to ask a question, provide feedback, or report a problem.

DYNAMIC AUDITING:

In today's complex business world of technology, disruption and globalisation, stakeholders, regulators and the public are scrutinising the stability and accountability of organisations with intensity.

While financial assurance will always be vital, this challenging new environment means there is room for Audit to play a bigger role. With its access to every corner of a company, and its deep knowledge of a company's processes, finances and strategy, Audit is in a unique position to help organisations see the potential for improvements – which in turn can bring confidence to markets and stakeholders.

Drawing on the deep expertise of our people, along with advanced data and analytics technology, our Dynamic Audit offering can be enhanced with tailored procedures into risk, culture and processes, as well as shining a light on where a company's data and analytics could be harnessed.

As each organisation is unique, each KPMG's Dynamic Audit is tailored to what will deliver the best audit quality and independent insights. In addition to providing a high quality audit opinion, our audit approach can assist with minimising risk, finding opportunities, making better decisions, and helping the organisation stay ahead of the curve in a challenging environment.

STORAGE SECURITY:

Storage security is the collective processes, tools and technologies that ensure that only authorized and legitimate users store, access and use storage resources. It enables better security of any storage resource through the implementation of required technologies and policies on storage access and consumption and the denial of access to all unidentified and potentially malicious users.

Storage security is a broad term that encompasses the implementation and management of security across all layers of a storage environment. This includes storage hardware, software, networks and/or the physical security of storage resources. Typically, storage security primarily deals with implementation at the software or logical layer. This is achieved through several

techniques such as encrypting/encoding data at rest and in motion, firewalling storage servers and implementing enterprise-wide identity and access management (IAM). Besides individuals, storage security also encompasses the management and protection of storage resources from unverified applications and services.

Full Homomorphic encryption

Homomorphic encryption is a form of encryption that allows computation on ciphertexts, generating an encrypted result which, when decrypted, matches the result of the operations as if they had been performed on the plaintext.

Homomorphic encryption can be used for privacy-preserving
outsourced storage and computation. This allows data to be encrypted and out-sourced to commercial cloud environments for processing, all while encrypted. In highly regulated industries, such as health care, homomorphic encryption can be used to enable new services by removing privacy barriers inhibiting data sharing. For example, predictive analytics in health care can be hard to apply due to medical data privacy concerns, but if the predictive analytics service provider can operate on encrypted data instead, these privacy concerns are diminished.

Homomorphic encryption is a form of encryption with an additional evaluation capability for computing over encrypted data without access to the secret key. The result of such a computation remains encrypted. Homomorphic encryption can be viewed as an extension of either symmetric- key or public-key cryptography. *Homomorphic* refers to homomorphismin algebra: the encryption and decryption functions can be thought as homomorphisms between plaintext and ciphertext spaces.

Homomorphic encryption includes multiple types of encryption schemes that can perform different classes of computations over encrypted data.[1] Some common types of homomorphic encryption are partially homomorphic, somewhat homomorphic, leveled fully homomorphic, and fully homomorphic encryption. The computations are represented as either Boolean or arithmetic circuits. *Partially homomorphic encryption* encompasses schemes that support the evaluation of circuits consisting of only one type of gate, e.g., addition or multiplication. *Somewhat homomorphic encryption* schemes can evaluate two types of gates, but only for a subset of circuits. *Leveled fully homomorphic encryption* supports the evaluation of arbitrary circuits of bounded (pre-determined) depth. *Fully homomorphic encryption* (FHE) allows the evaluation of arbitrary circuits of unbounded depth, and is the strongest notion of homomorphic encryption.
For the majority of homomorphic encryption schemes, the multiplicative depth of circuits is the main practical limitation in performing computations over encrypted data.

Homomorphic encryption schemes are inherently malleable. In terms of malleability, homomorphic encryption schemes have weaker security properties than non-homomorphic schemes.

Big data security:

Big Data security is the processing of guarding data and analytics processes, both in the cloud and on-premise, from any number of factors that could compromise their confidentiality.

Big Data Security Overview

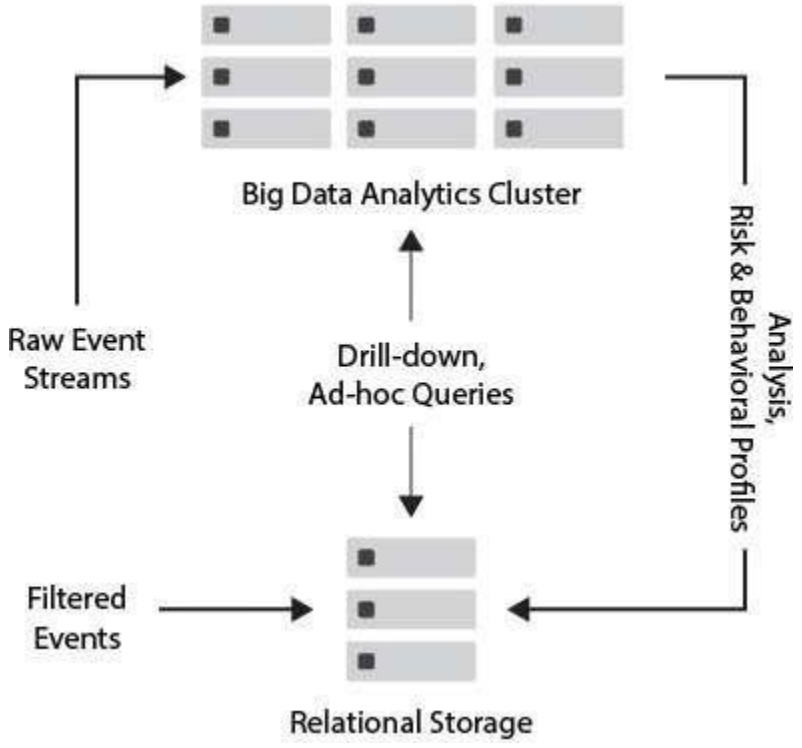
Big data security's mission is clear enough: keep out on unauthorized users and intrusions with firewalls, strong user authentication, end-user training, and intrusion protection systems (IPS) and intrusion detection systems (IDS). In case someone does gain access, encrypt your data in- transit and at-rest.

This sounds like any network security strategy. However, big data environments add another level of security because security tools must operate during three data stages that are not all present in the network. These are 1) data ingress (what's coming in), 2) stored data (what's stored), and 3) data output (what's going out to applications and reports).

Stage 1: Data Sources. Big data sources come from a variety of sources and data types. User-generated data alone can include CRM or ERM data, transactional and database data, and vast amounts of unstructured data such as email messages or social media posts. In addition to this, you have the whole world of machine generated data including logs and sensors. You need to secure this data in-transit from sources to the platform.

Stage 2: Stored Data. Protecting stored data takes mature security toolsets including encryption at rest, strong user authentication, and intrusion protection and planning. You will also need to run your security toolsets across a distributed cluster platform with many servers and nodes. In addition, your security tools must protect log files and analytics tools as they operate inside the platform.

Stage 3: Output Data. The entire reason for the complexity and expense of the big data platform is being able to run meaningful analytics across massive data volumes and different types of data. These analytics output results to applications, reports, and dashboards. This extremely valuable intelligence makes for a rich target for intrusion, and it is critical to encrypt output as well as ingress. Also, secure compliance at this stage: make certain that results going out to end-users do not contain regulated data.



One of challenges of Big Data security is that data is routed through a circuitous path, and in theory could be vulnerable at more than one point.

High-availability is, ultimately, the holy grail of the cloud. It embodies the idea of anywhere and anytime access to services, tools and data and is the enabler of visions of a future with companies with no physical offices or of global companies with completely integrated and unified IT systems. Availability is also related to reliability: a service that is on 24x7 but goes constantly offline is useless. For a service to have true high-availability, it needs not only to be always-on, but also to have several "nines" (99.999[...]) of reliability.

It has long been the case that to build systems with this kind of reliability and availability means large costs for companies. For something like this, it's not enough to simply have a failover cluster of servers in a data center: you must also have multiple redundant energy sources for the data center and even to have replication between multiple geographical locations in case of disasters. With the exception of very large, multinational companies, almost no-one could afford such a setup.

With the advent of infrastructure-as-a-service and platform-as-a-service providers, however, the costs of building such a service have decreased dramatically. It is now possible for most cloud-based service providers, especially for software-based services, to offer very aggressive service level agreements. Before getting there, however, it is necessary to understand what it means.

Understanding high availability

If you want a high-availability service, as a buyer or a seller, the first step is to understand what exactly it means. Let's take a 99.99% SLA, for instance. In practice, this means that in any given month (assuming a 30-day month), the service can only be offline for about 4 minutes and a few seconds, or only about 50 minutes per year. If we look at most cloud service providers today, how many actually deliver on this promise?

There are a few questions to be considered here. As a buyer of the service, do I really need this service level? Am I willing to pay the extra cost that will be associated with this? Are the guarantees being offered enough to cover my expenses in case of failure? The last one is perhaps the most important and most difficult one, because several factors have to be taken into consideration. If your company is going to rely on cloud services to function, what happens when they go offline?

Earlier this year, several high-profile tech companies had troubles when Amazon's EC2 service suffered an outage. The outage, according to Amazon itself, lasted for almost 11 hours. This is much larger than what would be acceptable according to their 99.95% SLA, and they offered a 10-day credit for all affected customers, but does this credit cover the true cost of the outage? As more and more services and applications go to the cloud, this question becomes increasingly important.

Standing on the shoulders of others

As cloud services and applications become more complex and more reliant on the underlying cloud platform, it becomes harder and harder to quickly identify and solve problems. Troubles can arise not only from an individual service, but from the interaction of multiple components and automated systems over distributed networks and data centers, resulting in issues that take a long time to be resolved. Regardless of the quality of a service provider, of their underlying hardware or platform, the chance of failure increases with complexity.

Understanding this complexity and the reliance on the platform is fundamental when defining the required availability level for a service. If you are building a service that needs 99.99% availability, you cannot simply rely on Amazon's EC2, since they only offer 99.95%. It would be necessary to have a different host for that service, or even multiple hosts, to be able to achieve that.

The same thing goes to buyers: if your vendor offers 99.99% availability look at how well they have maintained this service level in the past, and look at the availability of the underlying platform. If you know that it will not be enough, make it clear. I've had clients who demanded that I have replicated servers on geographically distinct locations in case of natural disasters. It might sound crazy, but it may make sense, and if it does, I must be ready to do this.

This is where the existence of multiple providers on the cloud comes in handy. Services can be hosted and replicated to multiple providers, on multiple locations, to greatly reduce the chances of failure. Even downtime related to maintenance can be reduced by spreading a service over

multiple providers: the chance of planned maintenance windows overlapping between providers is very small.

Putting it all together

As we've seen, there are several important factors that need to be considered when discussing availability on the cloud. Several large cloud providers, such as Rackspace and Amazon, offer very aggressive service levels - 100% uptime for Rackspace, 99.95% for Amazon - that are almost impossible to maintain, so we must ask ourselves what happens when the services fail? Are the credits we are entitled to in case of failure enough to cover the true costs of said failure? Is my service provider really capable of delivering on the promised availability level? What is his track record?

As cloud services mature, these questions become as important as price or other factors in choosing the right service provider. Answering them and understanding what is behind them becomes crucial so that everyone can trust and use cloud services without restrictions.

DOS ATTACKS:

A denial-of-service (DoS) is any type of attack where the attackers (hackers) attempt to prevent legitimate users from accessing the service. In a DoS attack, the attacker usually sends excessive messages asking the network or server to authenticate requests that have invalid return addresses. The network or server will not be able to find the return address of the attacker when sending the authentication approval, causing the server to wait before closing the connection. When the server closes the connection, the attacker sends more authentication messages with invalid return addresses. Hence, the process of authentication and server wait will begin again, keeping the network or server busy.

A DoS attack can be done in a several ways. The basic types of DoS attack include:

1. Flooding the network to prevent legitimate network traffic
2. Disrupting the connections between two machines, thus preventing access to a service
3. Preventing a particular individual from accessing a service.
4. Disrupting a service to a specific system or individual
5. Disrupting the state of information, such resetting of TCP sessions

Another variant of the DoS is the smurf attack. This involves emails with automatic responses. If someone emails hundreds of email messages with a fake return email address to hundreds of people in an organization with an autoresponder on in their email, the initial sent messages can become thousands sent to the fake email address. If that fake email address actually belongs to someone, this can overwhelm that person's account.

DoS attacks can cause the following problems:

1. Ineffective services
2. Inaccessible services
3. Interruption of network traffic
4. Connection interference

FAULY TOLERENCE IN CLOUD COMPUTING:

Fault tolerance in cloud computing is largely the same (conceptually) as in private or hosted environments. Meaning that it simply means the ability of your infrastructure to continue providing service to underlying applications even after the failure of one or more component pieces in any layer.

Explicating Fault Tolerance in Cloud Computing

Fault tolerance in cloud computing is about designing a blueprint for continuing the ongoing work whenever a few parts are down or unavailable. This helps the enterprises to evaluate their infrastructure needs and requirements, and provide services when the associated devices are unavailable due to some cause. It doesn't mean that the alternate arrangement can provide 100% of the full service, but this concept keeps the system in running mode at a useable, and most importantly, at a reasonable level. This is important if the enterprises are to keep growing in a continuous mode and increase their productivity levels.

Main Concepts behind Fault Tolerance in Cloud Computing System

- **Replication:** The fault tolerant system works on the concept of running several other replicates for each and every service. Thus, if one part of the system goes wrong, it has other instances that can be placed instead of it to keep it running. Take for example, a database cluster that has 3 servers with the same information on each of them. All the actions like data insertion, updates, and deletion get written on each of them. The servers, which are redundant, would be in the inactive mode unless and until any fault tolerance system doesn't demand the availability of them.
- **Redundancy:** When any system part fails or moves towards a downstate, then it is important to have backup type systems. For example, a website program that has MS SQL as its database may fail in between due to some hardware fault. Then a new database has to be availed in the redundancy concept when the original is in offline mode. The server operates with the emergency database which comprises of several redundant services within.

Techniques for Fault Tolerance in Cloud Computing

- All the services have to be given priority when designing a fault tolerance system. The database has to be given special preference because it powers several other units.
- After deciding the priorities, the enterprise has to work on the mock test. Take for example, the enterprise has a forum website that enables users to log in and posts comments. When the authentication services fail due to some problem, the users will not be able to log in. Then, the forum becomes a read-only one and does not serve the

purpose. But with the fault tolerant systems, remediation will be ensured and the user can search for information with minimal impact.

Major Attributes of Fault Tolerance in Cloud Computing

- **None Point Failure:** The concepts of redundancy and replication defines that fault tolerance can be had but with some minor impacts. If there isn't even a single point failure then the system is not a fault tolerant one.
- **Accept the Fault Isolation Concept:** The fault occurrence has to be handled separately from other systems. This helps the enterprise to isolate it from the existing system failure.

Existence of Fault Tolerance in Cloud Computing

- **System Failure:** This may be either software or hardware issue. The software failure results in system crash or hanging situation that may be due to stack overflow or other reasons. Any improper maintenance of the physical hardware machines will result in hardware system failure.
- **Security Breach Occurrences:** There are several reasons why fault tolerance occurs due to security failures. The hacking of the server negatively impacts the server and results in data breach. Other reasons for the necessity of fault tolerance in the form of security breaches include ransomware, phishing, virus attack etc.

CLOUD COMPUTING IN INDIA:

- Digital transformation has been recognised as being vital to the growth of our nation. This transformation has enjoyed the unanimous approval and contribution from all stakeholders including enterprises, MSMEs, government bodies, and citizens.
- But this level of adoption in a country with a population of over a billion people would need a robust technology base that is capable of collecting and distributing vital data seamlessly.
- Digital India envisions creating high-speed digital highways, that will impact commerce and create a digital footprint for every individual. Technologies based on mobility, analytics, Internet of things and most importantly, cloud technologiesare the building blocks for the digital India mission.
- There is a growing need to manage huge volumes of data, and **making them readily available to the public through digital cloud services.**
- While Data centers have become crucial to this transformation, IT leaders increasingly recognise that current data centers have reached their limits for supporting how state and local governments need to work and provide services.
- Lightening the load to adapt to increasing demand is the principle many government IT managers have in mind as they look to make data centers more efficient, flexible and

capable of delivering new services. Government IT departments are also prioritising investments in data center consolidation and new technologies to enable higher IT service levels.

- **There are three trends that are impacting government IT today:**

- Virtualization and cloud
- Infrastructure as a service (IaaS)
- Flexible infrastructure for application development and delivery

Evolving The Data Center To Hyper-Convergence For Virtualisation And Cloud

The traditional image of a data center is a cavernous room filled with rows of equipment racks and whirring, blinking boxes. This reality is fast disappearing as advances in virtualisation technology pack more capabilities into smaller devices.

Government data centers are evolving to take advantage of virtualisation, especially for servers and storage. The goal is to capture the associated benefits of higher data center efficiency and optimization, as well as reduced capital and operational expenses.

The data center model is also evolving to support private cloud and IT as a service to accomplish IT projects faster and more effectively.

Virtualisation enables a hyper-converged infrastructure that integrates servers and storage in a single appliance. The systems leverage industry-standard hardware and software-defined storage, enabling easy scalability and management.

A well-designed hyper-convergence infrastructure in the data center offers several additional advantages for IT operations and service delivery like:

- Cost reductions for infrastructure, software licenses, cabling and other elements, with predictable budgeting for data center growth
- Easier, on-demand and linear scalability of compute and storage resources, which reduces the need to overprovision resources in anticipation of potential performance demands
- Flexibility to support new IT offerings, such as analytics, that help government employees improve services to constituents
- Simpler management with fewer server and storage silos

Delivering Private Cloud And Infrastructure As A Service

One key to agility — in both government and IT — is **having the right resources ready to go at a moment's notice** but to use them only when they are truly needed. That agility is behind the idea of IaaS on hyper-converged infrastructure: delivering computing, storage and network resources on demand to application developers and users.

This environment operates like a private cloud, where the IT infrastructure can serve more applications and users without the need to add more staff. By creating a private internal cloud, IT

managers also can reduce concerns that come with using untrusted or shared cloud services, including security, compliance and audit trails.

IT can automate many operational tasks around provisioning and orchestration, which makes it easier to activate or repurpose servers as needed. Additionally, automated configuration and management of IT resources means IT staff can focus on strategic, high-value activities.

Any Application At Any Scale

From a smart phone app used by one employee to a complex information system used by hundreds, the ability to deliver any application at any scale is essential for government IT. This scalability requires an infrastructure that can quickly deliver the right resources for computing requirements, storage capacity and application performance.

Yet different applications commonly used for government functions require different types of resources and performance levels. For example, a GIS (Geographic Information Systems) application needs more storage capacity than compute capability, while transaction-oriented applications are often compute-intensive and don't require as much data storage.

An agile government is one that has the information, applications and computing capabilities that keep pace with the fast-paced changes in citizen and employee expectations for services.

By considering the data center trends discussed, IT can make the infrastructure simpler while also delivering services that make government better.

Cloud has demonstrated the capability to digitize the governance system while proving to be cost-effective. The global world is eager to see India embrace the Cloud Computing frontier led by technology capability to improve people's lives.