MODULE 1

Cloud Computing

The term cloud refers to a network or the internet. It is a technology that uses remote servers on the internet to store, manage, and access data online rather than local drives. The data can be anything such as files, images, documents, audio, video, and more

There are the following operations that we can do using cloud computing:

- Developing new applications and services
- Storage, back up, and recovery of data
- Hosting blogs and websites
- Delivery of software on demand
- Analysis of data
- Streaming videos and audios

With increase in computer and Mobile user's, data storage has become a priority in all fields. Large and small scale businesses today thrive on their data & they spent a huge amount of money to maintain this data. It requires a strong IT support and a storage hub. Not all businesses can afford high cost of in-house IT infrastructure and back up support services. For them Cloud Computing is a cheaper solution. Perhaps its efficiency in storing data, computation and less maintenance cost has succeeded to attract even bigger businesses as well.

Cloud computing decreases the hardware and software demand from the user's side. The only thing that user must be able to run is the cloud computing systems interface software, which can be as simple as Web browser, and the Cloud network takes care of the rest. We all have experienced cloud computing at some instant of time, some of the popular cloud services we have used or we are still using are mail services like gmail, hotmail or yahoo etc.

Following are the benefits of cloud computing:

1. Lower IT infrastructure and computer costs for users
2. Improved performance
3. Fewer Maintenance issues
4. Instant software updates
5. Improved compatibility between Operating systems
6. Backup and recovery
7. Performance and Scalability
8. Increased storage capacity
9. Increase data safety

Characteristics of Cloud Computing

1) Agility

The cloud works in a distributed computing environment. It shares resources among users and works very fast.

2) High availability and reliability

The availability of servers is high and more reliable because the chances of infrastructure failure are minimum.

3) High Scalability

Cloud offers "on-demand" provisioning of resources on a large scale, without having engineers for peak loads.

4) Multi-Sharing

With the help of cloud computing, multiple users and applications can work more efficiently with cost reductions by sharing common infrastructure.

5) Device and Location Independence

Cloud computing enables the users to access systems using a web browser regardless of their location or what device they use e.g. PC, mobile phone, etc. As infrastructure is off- site (typically provided by a third-party) and accessed via the Internet, users can connect from anywhere.

6) Maintenance

Maintenance of cloud computing applications is easier, since they do not need to be installed on each user's computer and can be accessed from

different places. So, it reduces the cost also.

7) Low Cost

By using cloud computing, the cost will be reduced because to take the services of cloud computing, IT company need not to set its own infrastructure and pay-as-per usage of resources.

8) Services in the pay-per-use mode

Application Programming Interfaces (APIs) are provided to the users so that they can access services on the cloud by using these APIs and pay the charges as per the usage of services.

Disadvantages of Cloud Computing

1) Internet Connectivity

As you know, in cloud computing, every data (image, audio, video, etc.) is stored on the cloud, and we access these data through the cloud by using the internet connection. If you do not have good internet connectivity, you cannot access these data. However, we have no any other way to access data from the cloud.

2) Vendor lock-in

Vendor lock-in is the biggest disadvantage of cloud computing. Organizations may face problems when transferring their services from one vendor to another. As different vendors provide different platforms, that can cause difficulty moving from one cloud to another.

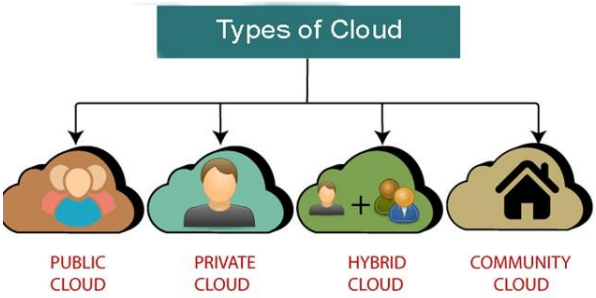
3) Limited Control

As we know, cloud infrastructure is completely owned, managed, and monitored by the service provider, so the cloud users have less control over the function and execution of services within a cloud infrastructure.

4) Security

Although cloud service providers implement the best security standards to store important information. But, before adopting cloud technology, you should be aware that you will be sending all your organization's sensitive information to a third party, i.e., a cloud computing service provider. While sending the data on the cloud, there may be a chance that your organization's information is hacked by Hackers.

Types of Cloud



Public Cloud

- Public Cloud provides a **shared platform** that is accessible to the **general public** through an Internet connection.
- Public cloud operated on the **pay-as-per-use model** and administrated by the **third party**, i.e., Cloud service provider.
- In the Public cloud, the same storage is being used by multiple users at the same time.
- Public cloud is **owned, managed,** and **operated** by businesses, universities, government organizations, or a combination of them.
- Amazon Elastic Compute Cloud (EC2), Microsoft Azure, IBM's Blue Cloud, Sun Cloud, and Google Cloud are examples of the public cloud.

Advantages of Public Cloud

1) Low Cost

Public cloud has a lower cost than private, or hybrid cloud, as it shares the same resources with a large number of consumers.

2) Location Independent

Public cloud is location independent because its services are offered through the internet.

3) Save Time

In Public cloud, the cloud service provider is responsible for the manage and maintain data centers in which data is stored, so the cloud user can save their time to establish connectivity, deploying new products, release product updates, configure, and assemble servers.

4) Quickly and easily set up

Organizations can easily buy public cloud on the internet and deployed and configured it remotely through the cloud service provider within a few hours.

5) Business Agility

Public cloud provides an ability to elastically re-size computer resources based on the organization's requirements.

6) Scalability and reliability

Public cloud offers scalable (easy to add and remove) and reliable (24*7 available) services to the users at an affordable cost.

Disadvantages of Public Cloud

1) Low Security

Public Cloud is less secure because resources are shared publicly.

2) Performance

In the public cloud, performance depends upon the speed of internet connectivity.

3) Less customizable

Public cloud is less customizable than the private cloud.

Private Cloud

- Private cloud is also known as an **internal cloud** or **corporate cloud**.
- Private cloud provides computing services to a **private internal network (within the organization)** and **selected users** instead of the general public.
- Private cloud provides a **high level of security** and **privacy** to data through firewalls and internal hosting. It also ensures that operational and sensitive data are not accessible to third-party providers.
- HP Data Centers, Microsoft, Elastra-private cloud, and Ubuntu are the example of a private cloud.

Advantages of Private cloud

1) More Control

Private clouds have more control over their resources and hardware than public clouds because it is only accessed by selected users.

2) Security & privacy

Security & privacy are one of the big advantages of cloud computing. Private cloud improved the security level as compared to the public cloud.

3) Improved performance

Private cloud offers better performance with improved speed and space capacity. Disadvantages of Private Cloud

1) High cost

The cost is higher than a public cloud because set up and maintain hardware resources are costly.

2) Restricted area of operations

As we know, private cloud is accessible within the organization, so the area of operations is limited.

3) Limited scalability

Private clouds are scaled only within the capacity of internal hosted resources.

4) Skilled people

Skilled people are required to manage and operate cloud services.

Hybrid Cloud

- Hybrid cloud is a combination of **public and private** clouds.
- **Hybrid cloud = public cloud + private cloud**
- The main aim to combine these cloud (Public and Private) is to create a unified, automated, and well-managed computing
- environment.
- In the Hybrid cloud, **non-critical activities** are performed by the **public cloud** and **critical activities** are performed by the **private cloud**.
- Mainly, a hybrid cloud is used in finance, healthcare, and Universities.
- The best hybrid cloud provider companies are **Amazon, Microsoft, Google, Cisco,** and **NetApp**.

Advantages of Hybrid Cloud

1) Flexible and secure

It provides flexible resources because of the public cloud and secure resources because of the private cloud.

2) Cost effective

Hybrid cloud costs less than the private cloud. It helps organizations to save costs for both infrastructure and application support.

3) Cost effective

It offers the features of both the public as well as the private cloud. A hybrid cloud is capable of adapting to the demands that each company needs for space, memory, and system.

4) Security

Hybrid cloud is secure because critical activities are performed by the private cloud.

5) Risk Management

Hybrid cloud provides an excellent way for companies to manage the risk.

Disadvantages of Hybrid Cloud

1) Networking issues

In the Hybrid Cloud, networking becomes complex because of the private and the public cloud.

2) Infrastructure Compatibility

Infrastructure compatibility is the major issue in a hybrid cloud. With dual-levels of infrastructure, a private cloud controls the company, and a public cloud does not, so there is a possibility that they are running in separate stacks.

3) Reliability

The reliability of the services depends on cloud service providers.

Community Cloud

Community cloud is a cloud infrastructure that allows systems and services to be accessible by a group of several organizations to share the information. It is owned, managed, and operated by one or more organizations in the community, a third party, or a combination of them.

Example: Our government organization within India may share computing infrastructure in the cloud to manage data.

Advantages of Community Cloud

1) Cost effective

Community cloud is cost effective because the whole cloud is shared between several organizations or a community.

2) Flexible and Scalable

The community cloud is flexible and scalable because it is compatible with every user. It allows the users to modify the documents as per their needs and requirement.

3) Security

Community cloud is more secure than the public cloud but less secure than the private cloud.

4) Sharing infrastructure

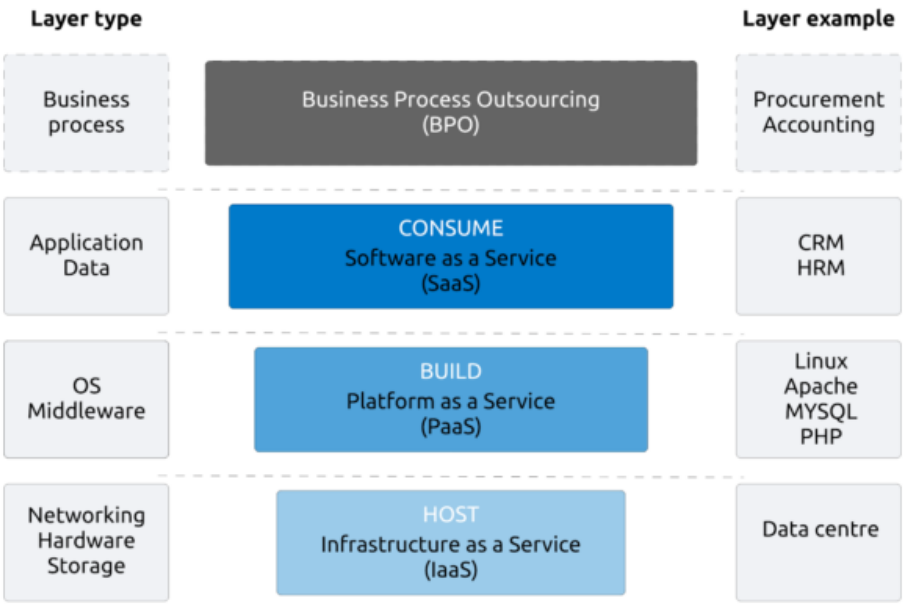
Community cloud allows us to share cloud resources, infrastructure, and other capabilities among various organizations.

Disadvantages of Community Cloud

There are the following disadvantages of Community Cloud -

- Community cloud is not a good choice for every organization.
- Slow adoption to data
- The fixed amount of data storage and bandwidth is shared among all community members.
- Community Cloud is costly than the public cloud.
- Sharing responsibilities among organizations is difficult.

The 4 layers of Cloud



Features of a Cloud Infrastructure Management

Provisioning and configuration: Developers, systems engineers and other IT professionals use these tools to set up and configure the hardware and software resources they need. This would include Spinning up a new server, Installing an operating system or other software, Allocating storage resources and other cloud infrastructure needs.

This also includes features for enabling and managing self-service provisioning, in which end users use a dashboard or other mechanisms for standing up their own resources as needed, based on predetermined rules.

Visibility and monitoring: Cloud infrastructure management tools allow operators to "see" their environments. More importantly, they include or integrate with monitoring tools that: Check system health, Deliver real-time alerts and notifications, Create reporting and analytics.

Resource allocation: Related to cost optimization, resource allocation features enable granular control over how users consume cloud infrastructure, including self-service provisioning. This is similar to budgeting: dividing up shared resources appropriately and in some cases creating criteria for going over budget.

Cost optimization: Managing costs is a critical capability of cloud infrastructure management tools. Without this component, enterprises run an increased risk of "sticker shock" when the cloud bill arrives. Proactively monitoring costs via strategies such as turning off unused or unnecessary resources is key to maximizing the ROI of cloud infrastructure.

Automation: Cloud infrastructure management tools sometimes offer automation capabilities for various operational tasks, such as configuration management, auto-provisioning and auto-scaling.

Security: Cloud infrastructure management tools are another part of a holistic cloud security strategy. They are one mechanism for properly configuring a cloud provider's native security controls based on a particular setup and needs.

Cloud Service Models

There are the following three types of cloud service models -

1. Infrastructure as a Service (IaaS)
2. Platform as a Service (PaaS)
3. Software as a Service (SaaS)

Infrastructure as a Service (IaaS)

IaaS is also known as **Hardware as a Service (HaaS)**. It is a computing infrastructure managed over the internet. The main advantage of using IaaS is that it helps users to avoid the cost and complexity of purchasing and managing the physical servers.

Characteristics of IaaS

There are the following characteristics of IaaS -

- Resources are available as a service.
- Services are highly scalable.
- Dynamic and flexible.
- GUI and API-based access.
- Automated administrative tasks.

Example: DigitalOcean, Linode, Amazon Web Services (AWS), Microsoft Azure, Google Compute Engine (GCE), Rackspace, and Cisco Metacloud.

Platform as a Service (PaaS)

PaaS cloud computing platform is created for the programmer to develop, test, run, and manage the applications.

Characteristics of PaaS

- Accessible to various users via the same development application.
- Integrates with web services and databases.
- Builds on virtualization technology, so resources can easily be scaled up or down as per the organization's need.
- Support multiple languages and frameworks.
- Provides an ability to "**Auto-scale**".

Example: AWS Elastic Beanstalk, Windows Azure, Heroku, Force.com, Google App Engine, Apache Stratos, Magento Commerce Cloud, and OpenShift.

Software as a Service (SaaS)

SaaS is also known as "**on-demand software**". It is a software in which the applications are hosted by a cloud service provider. Users can access these applications with the help of internet connection and web browser.

Characteristics of SaaS

- Managed from a central location.
- Hosted on a remote server.
- Accessible over the internet.
- Users are not responsible for hardware and software updates. Updates are applied automatically.
- The services are purchased on the pay-as-per-use basis.

Example: BigCommerce, Google Apps, Salesforce, Dropbox, ZenDesk, Cisco WebEx, ZenDesk, Slack, and GoToMeeting.

IaaS	Paas	SaaS
It provides a virtual data center to store information and create platforms for app development, testing, and deployment.	It provides virtual platforms and tools to create, test, and deploy apps.	It provides web software and apps to complete business tasks.
It provides access to resources such as virtual machines, virtual storage, etc.	It provides runtime environments and deployment tools for applications.	It provides software as a service to the end-users.
It is used by network architects.	It is used by developers.	It is used by end users.
IaaS provides only Infrastructure.	PaaS provides Infrastructure+Platform.	SaaS provides Infrastructure+Platform +Software.

Roots of cloud computing

Roots of cloud computing mean the existing technologies that are cooperating in cloud computing for delivering services to end-users.

1- System managements
2- Distributed computing
3- Internet technologies
4- Hardware virtualization

1- System Management

Autonomic/Automatic computing

The basic idea behind this is to make our cloud software system so effective that our machine can operate without any human interface and perform any critical function without any outside help.

Vital elements of automatic/autonomic computing systems

If any machine can do the below things by itself, then we called it autonomic computing.

* Known itself what next task to do.
* Configure resources on their own.
* automatically optimize the function parameter which is given to it.
* If any problems come, then it heals by itself.
* Automatically adopt the changing value.

2- Distributed computing

Grid computing

It is a distributed computing, In grid computing groups of the computer's resources, which are located at anywhere interconnected with each other with the help of the Internet for performing some bigger tasks. And the computer resources can be heterogeneous (it means it can be of a different type).

The work of grid computing is that it divides any big computing task into smaller tasks and distribute them.

Essential points in grid computing:

For performing grid computing, there will be a requirement for special grid software.
At least one computer in a group works as a server.

Utility computing

It is a pay per use model (you have to pay according to your usage)
It provides you on-demand facility as metered services (metered means which was charged)

Users use the computing resources present on the cloud server and pay according to these resource usages.

3- Internet Technologies

Web 2.0
In cloud, computing the Internet plays a very centric role. With the help of the web, cloud computing provides its essential services to the users. This web uses various technologies to deliver those services. The advancements in technologies make web so rich platform for developing the application and all this happen because of web 2.0. It provides flexibility to the user and the developer to easily share their responses on the web. Some examples of web 2.0 are google map, all social media websites, and so on.

SOA (service-oriented architecture)
SOA is an architectural approach which provides services with the help of the Internet to the users for developing an application

* service provider and service consumer plays a vital role in SOA

Web services
Web services are an arrangement of the software used for transferring the data between the web application (for example, any social media website) so that users excess it easily on the Internet.

4- Hardware virtualization in cloud computing

This was the another root of cloud computing and before moving to **hardware virtualization**, let us understand virtualization.

Virtualization- creating a version of something which was virtual, not real, and the advantage of this is that multiple users can easily share the single resources.

There are two crucial things in hardware virtualization:

* host machine - on which our virtual machine run

* Guest machine - Virtual machine which we will create considered as a guest machine

And there is a program called a hypervisor, which manages all the resources in the host server without hardware virtualization. We can't even think about cloud computing because any resources which we were using in cloud computing are virtual, not physical.

Challenges and Risks in Cloud Computing

Security issues: The security risks of cloud computing have become a reality for every organization, be it small or large. That’s why it is important to implement a secure BI cloud tool that can leverage proper security measures.

Cost management and containment: It sometimes difficult to define and predict quantities and costs.

Lack of resources/expertise: Organizations are increasingly placing more workloads in the cloud while cloud technologies continue to rapidly advance. Due to these factors, organizations are having a tough time keeping up with the tools. Also, the need for expertise continues to grow.

Governance/Control: Proper IT governance should ensure IT assets are implemented and used according to agreed-upon policies and procedures; ensure that these assets are properly controlled and maintained, and ensure that these assets are supporting your organization's strategy and business goals.

Compliance: Every time a company moves data from the internal storage to a cloud, it is faced with being compliant with industry regulations and laws.

Managing multiple clouds: The state of multi-cloud has grown exponentially in recent years. Companies are shifting or combining public and private clouds and, as mentioned earlier, tech giants like Alibaba and Amazon are leading the way.

Performance: When a business moves to the cloud the performance becomes dependent on the service providers.

Building a private cloud: Although building a private cloud isn't a top priority for many organizations, for those who are likely to implement such a solution, it quickly becomes one of the main challenges facing cloud computing – private solutions should be carefully addressed.

Segmented usage and adoption: Most organizations did not have a robust cloud adoption strategy in place when they started to move to the cloud. Instead, ad-hoc strategies sprouted, fuelled by several components. One of them was the speed of cloud adoption.

Migration: This is a process of moving an application to a cloud. Some challenges faced during storing data in cloud are:

- Extensive troubleshooting
- Security challenges
- Slow data migrations
- Migration agents
- Cutover complexity
- Application downtime

Migrating into Cloud

Migrating an application to the cloud is not an easy task. It is important to strictly adhere to the seven-step model to ensure that the process is robust and error free. The seven quick stages of migration into the cloud are outlined below.

1. ASSESSMENT
Migration starts with an assessment of the issues relating to migration, at the application, code, design, and architecture levels. Moreover, assessments are also required for tools being used, functionality, test cases, and configuration of the application. The proof of concepts for migration and the corresponding pricing details will help to assess these issues properly.

2. ISOLATE
The second step is the isolation of all the environmental and systemic dependencies of the enterprise application within the captive data center. These include library, application, and architectural dependencies. This step results in a better understanding of the complexity of the migration.

3. MAP
A mapping construct is generated to separate the components that should reside in the captive data center from the ones that will go into the cloud.

4. RE-ARCHITECT
It is likely that a substantial part of the application has to be re-architected and implemented in the cloud. This can affect the functionalities of the application and some of these might be lost. It is possible to approximate lost functionality using cloud runtime support API.

5. AUGMENT
The features of cloud computing service are used to augment the application.

6. TEST
Once the augmentation is done, the application needs to be validated and tested. This is to be done using a test suite for the applications on the cloud. New test cases due to augmentation and proof-of-concepts are also tested at this stage.

7. OPTIMISE
The test results from the last step can be mixed and so require iteration and optimization. It may take several optimizing iterations for the migration to be successful. It is best to iterate through this seven-step model as this will ensure the migration to be robust and comprehensive.

MODULE 2

Software as a Service (SaaS)

Software as a service (SaaS) is a software distribution model in which a cloud provider hosts applications and makes them available to end users over the internet. In this model, an independent software vendor (ISV) may contract a third-party cloud provider to host the application. Or, with larger companies, such as Microsoft, the cloud provider might also be the software vendor.

SaaS is one of three main categories of cloud computing, alongside infrastructure as a service (IaaS) and platform as a service (PaaS). A range of IT professionals, business users and personal users use SaaS applications. Products range from personal entertainment, such as Netflix, to advanced IT tools. Unlike IaaS and PaaS, SaaS products are frequently marketed to both B2B and B2C users.

SaaS works through the cloud delivery model. A software provider will either host the application and related data using its own servers, databases, networking and computing resources, or it may be an ISV that contracts a cloud provider to host the application in the provider's data center. The application will be accessible to any device with a network connection. SaaS applications are typically accessed via web browsers.

SaaS architecture

SaaS applications and services typically use a multi-tenant approach, which means a single instance of the SaaS application will be running on the host servers, and that single instance will serve each subscribing customer or cloud tenant. The application will run on a single version and configuration across all customers, or tenants.

The History of SaaS

The Timeline of Computing

1950s: Discrete apps on the first commercially available mainframe systems
1960s: Discrete apps on midrange or minicomputers
1970s: Discrete apps on personal or microcomputers
1980s: Application suites on DOS-based personal computers and computer networks
1989: Invention of the World Wide Web
1990s: Enterprise app suites move to client-server architectures and mainframe computing
1990s: First SaaS application introduced, managed hosting services, ASPs (Single-tenant)
1998: 1st Multi-tenant SaaS app
2000s: SaaS app suites and ecosystems such as Salesforce begin to emerge
2010: Salesforce led the way to prove that SaaS is the future of software delivery
2010-Now: Refined SaaS products that include almost everything that was available as a packaged software

SaaS has taken various forms since the 60s and has not only survived over five decades, but has also transformed into a technology that's still growing at a rapid rate. The growth of SaaS and cloud computing is showing no signs of slowing down anytime soon. Experts believe that SaaS growth has not even reached its peak yet and is still in its middle age. What this essentially means is that the SaaS industry is mature enough, but the best of what it has to offer is yet to come.

SaaS advantages

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.

Other benefits of the SaaS model include:

Flexible payments: Rather than purchasing software to install, or additional hardware to support it, customers subscribe to a SaaS offering. Transitioning costs to a recurring operating expense allows many businesses to exercise better and more predictable budgeting. Users can also terminate SaaS offerings at any time to stop those recurring costs.

Scalable usage: Cloud services like SaaS offer high Vertical scalability, which gives customers the option to access more or fewer services or features on demand.

Automatic updates: Rather than purchasing new software, customers can rely on a SaaS provider to automatically perform updates and patch management. This further reduces the burden on in-house IT staff.

Accessibility and persistence: Since SaaS vendors deliver applications over the internet, users can access them from any internet-enabled device and location.

Customization: SaaS applications are often customizable and can be integrated with other business applications, especially across applications from a common software provider.

SaaS challenges and risks

SaaS also poses some potential risks and challenges, as 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's data.

Issues beyond customer control: Issues can arise when providers experience service disruptions, impose unwanted changes to service offerings or experience a security breach -- all of which can have a profound effect on the customers' ability to use the SaaS offering. To proactively mitigate these issues, customers should understand their SaaS provider's SLA and make sure it is enforced.

Customers lose control over versioning: If the provider adopts a new version of an application, it will roll out to all of its customers, regardless of whether or not the customer wants the newer version. This may require the organization to provide extra time and resources for training.

Difficulty switching vendors: As with using any cloud service provider, switching vendors can be difficult. To switch vendors, customers must migrate very large amounts of data. Furthermore, some vendors use proprietary technologies and data types, which can further complicate customer data transfer between different cloud providers. Vendor lock-in is when a customer cannot easily transition between service providers due to these conditions.

Security: Cloud security is often cited as a significant challenge for SaaS applications.

Integration Challenges: Application Programming Interfaces (API) are Insufficient. Many SaaS providers have responded to the integration challenge by developing APIs.

SaaS security and privacy

The cybersecurity risks associated with software as a service are different from those associated with traditional software. With traditional software, the software vendor is responsible for eliminating code-based vulnerabilities, while the user is responsible for running the software on a secure infrastructure and network. As a result, security is more the responsibility of the independent software vendor and third-party cloud provider.

Despite the rapid adoption of cloud-based models for fully serviced software products, organizations still have certain reservations about SaaS products when it comes to security and privacy.

These concerns include:

- Encryption and key management.
- Identity and access management (IAM).
- Security monitoring.
- Incident response.
- Poor integration into broader, company-specific security environments.
- Fulfillment of data residency requirements.
- Data privacy.
- Cost of investing in third-party tools to offset the SaaS security risk.
- Lack of communication with technical and security experts during the sales process.

SaaS vs IaaS vs PaaS

SaaS is one of the three major cloud service models, along with IaaS and PaaS. All three models involve cloud providers that deliver their own hosted data center resources to customers over the internet. Where the models differ is in the completeness of the product. SaaS products are complete and fully managed applications. IaaS is largely outsourcing data center resources, and PaaS delivers a development platform and other tools hosted by the provider's data center.

SaaS application users do not have to download software, manage any existing IT infrastructures or deal with any aspect of the software management. Vendors handle maintenance, upgrades, support, security and all other aspects of managing the software.

IaaS is used by companies that want to outsource their data center and computer resources to a cloud provider. IaaS providers host infrastructure components such as servers, storage, networking hardware and virtualization resources. Customer organizations using IaaS services must still manage their data use, applications and operating systems (OSes).

PaaS provides a framework of resources for an organization's in-house developers. This hosted platform enables developers to create customized applications. The vendor manages the data center resources that support the tools. Customer organizations using PaaS services do not have to manage their OSes, but must manage applications and data use.

Approaching the SaaS Integration Enigma

- Integration as a Service (IaaS) is all about the migration of the functionality of a typical enterprise application integration into the cloud for providing for smooth data transport between any enterprise and SaaS applications.
- Cloud middleware will be made available as a service.
- Due to varying integration requirements and scenarios, there are a number of middleware technologies and products such as JMS compliant message queues and integration backbones such as EAI, ESB, EDB, CEP, etc.
- For application integration, it is *Enterprise Application Integration (EAI)*
- For service integration, it is *Enterprise Service Bus (ESB)*
- For data integration, it is *Enterprise Data Bus (EDB)*
- *Complex Event Processing (CEP)* connects decoupled systems where Events are coming up fast.
- Cloud infrastructure is not very useful without SaaS applications that run on top of them, and SaaS applications are not very valuable without access to the critical corporate data that is typically locked away in various corporate systems. That brings out the importance of SaaS integration.
- Integration is not easier either to implement as successful untangling from the knotty situation is a big issue.
- The constraining attributes of SaaS applications are:
    - Dynamic nature of the SaaS interfaces that constantly change.
    - Dynamic nature of the metadata native to a SaaS provider such as salesforce.com.
    - Managing assets that exist outside of the firewall.
    - Massive amounts of information that need to move between SaaS.
    - On-premise systems daily and the need to maintain data quality and integrity.

Integration is more complicated. The most probable reasons are:

- New integration scenarios.
- Access to the cloud may be limited. Access to cloud resources (SaaS, PaaS, and the infrastructures) is more limited than local applications. Once applications move to the cloud, custom applications must be designed to support integration because there is no longer that low level of access.
- Enterprises putting their applications in the cloud or those subscribers of cloud-based business services are dependent on the vendor to provide the integration hooks and APIs. But service providers fail to provide support.
- **Dynamic Resources:** Due to the dynamism factor that is sweeping the whole could ecosystem, application versioning and infrastructural changes are liable for dynamic changes. These would clearly impact the integration model.
- **Performance:** Clouds support application scalability and resource elasticity the network distances between elements in the cloud are no longer under our control. The round-trip latency is an issue that slows down the cloud integration.

New Integration Scenarios

1. Within a Public Cloud:

Two different applications are hosted in a cloud. The role of the cloud integration middleware (say cloud-based ESB or internet service bus (ISB)) is to seamlessly enable these applications to talk to each other. The possible sub-scenarios include these applications can be owned by two different companies. They may live in a single physical server but run on different virtual machines.

2. Homogeneous Clouds:

The applications to be integrated are posited in two geographically Separated cloud infrastructures. The integration middleware can be in cloud 1 or 2 or in a separate cloud. There is a need for data and protocol transformation and they get done by the ISB.

3. Heterogeneous Clouds:

One application is in public cloud and the other Application is private cloud. This is the currently dominating scene for cloud integration

The Integration Methodologies

Traditional enterprise integration tools deployed on-premise with special connectors to access cloud-based applications:

- This is a most likely approach for IT organizations that have already standardized on an integration tool for integrating their enterprise applications, and now looking to extend that capability to include cloud applications.
- Many integration vendors are also providing subscription based licensing options to help customers license their technologies for term-based projects.
- Examples include data integration technologies like Informatica's PowerCenterproduct, message based EAI/ESB technologies, and appliance based integration tools.

Traditional enterprise integration tools hosted in the cloud:

- This approach is similar to the first option, except that the equipment for installing the integration software is hosted by a third party provider so that the customer does not have to worry about procuring and managing the hardware, or installing the integration software.
- This is a good fit for IT organizations or systems integrators who have the skills and resources to build integration jobs and manage them on behalf of their end users or customers.
- It is a good fit for cloud-to-cloud integrations, but requires a secure VPN tunnel to access on-premise data.
- An example of a hosted integration technology is Cloud Edition on Amazon EC2.

Integration-as-a-Service or On Demand Integration offerings:

- These are SaaS applications that are designed to deliver integration securely over the internet, and are able to integrate cloud applications to *on-premise systems*, *cloud-to-cloud applications*, or *on-premise to on-premise applications* without requiring the users to setup a VPN connection or provide any special access through their firewall.
- This approach is a good fit for companies who care the most about *ease of use*, *ease of maintenance*, *time to deployment*, and are on a tight budget.
- It is appealing to small and mid-sized companies, as well as large enterprises with a departmental application deployment.
- good example is Informatica's On Demand Integration Services**.**

No single approach is a perfect fit for all scenarios, and the right approach for any company will depend on some of the following factors:

- Who will build and maintain integration jobs – In-house IT resources, the application administrator/business analyst or both.
- The scope and complexity of the project – Is it a pilot project, a small departmental project or a strategic enterprise deployment.
- The size of the company – What is their budget for an integration tool and do they prefer a buy vs. lease approach.

SaaS Integration Products and Platforms

Jitterbit:

- Jitterbit is a fully graphical integration solution that provides users a versatile platform and a suite of productivity tools to reduce the integration efforts sharply.
- Jitterbit Integration Environment: Point-and-click GUI that enables to quickly configure, test, deploy and manage integration projects on the Jitterbit server.
- Jitterbit Integration Server: A powerful and scalable run-time engine that processes all the integration operations, fully configurable and manageable from the Jitterbit application.

Bungee Connect:

- For professional developers, Bungee Connect offers an application development and deployment platform that enables highly interactive applications integrating multiple data sources and facilitating instant deployment.
- Bungee automates the development of rich UI and eases the difficulty of deployment to multiple web browsers.

OpSource Connect:

- OpSource services provide both enterprise-class cloud and managed server hosting that delivers the security, control, high-performance and ease of integration that enterprise customer's demand.

SnapLogic:

- SnapLogic is a capable, clean solution for data integration that can be deployed in enterprise as well as in cloud landscapes.

Bluewolf:

- Bluewolf proactively alerts its customers of any issues with integration and helps to solves them quickly.
- The Bluewolf Integrator integrates with Salesforce, BigMachines, Oracle, SAP, Microsoft SQL server, MySQL, and supports flat files, such as CSV, XHTML and many more.

B2B integration

Business-to-business (B2B) integration is the automation of business processes and communication between two or more organizations. It allows them to work and trade more effectively with their customers, suppliers and business partners by automating key business processes. B2B integration software provides the architecture needed to digitize information and quickly route it through an organization's trading ecosystem.

A B2B integration platform helps companies integrate all their complex B2B and electronic data interchange (EDI) processes across their partner communities in a single gateway. The platform collects data from source applications, translates the data into standardized formats and then sends the documents to the business partner using the appropriate transport protocol. B2B integration software is available for on-premises use or integration services can be accessed through hosted cloud services.

To achieve goals like increasing revenue, speeding time to market and improving efficiencies, organizations need a successful business network — and that requires a modern B2B integration solution. With the right B2B tools, organizations can digitally connect and communicate quickly and reliably. This can reduce the time it takes to get new products and services to market and help companies achieve the nimbleness and agility they need to compete.

Enterprise Cloud Computing

Enterprise Cloud Computing refers to a computing environment residing behind a firewall that delivers software, infrastructure and platform services to an enterprise. Cloud computing also typically delivers Web services, providing access to components that can be easily combined to rapidly create composite web applications to meet the ever-changing needs of a business operation. Web services rely on service-oriented architecture that provides software developers with interfaces that leverage functionality contained within existing web applications, resulting in reduced web application development time frames and lower software development costs for an enterprise.

Benefits

Enterprise Cloud Computing offers many benefits to an organization, including superior speed and performance for IT resources, more efficient utilization of IT resources, lower IT infrastructure costs, lower IT operational costs and increased capacity to handle peaks in demand for IT resources, like web applications and services. Cloud computing within an enterprise also provides a safer computing environment, through the use of virtual servers, that reduces the threat of an onsite intruder attack on the physical storage devices within a data center. Enterprise cloud computing also provides the capacity for flexible data security policies, where security decisions can be made based on a variety of factors including: the user's role within the enterprise, the user's current access location, the type of data or application being accessed by the user and the type of device being used.

Implementation

With Enterprise Cloud Computing, the bottlenecks that typically occur with the configuration, expansion and replacement of traditional on-premise IT systems and components can be eliminated, since the IT infrastructure can be expanded or contracted on demand through virtualization. Cloud computing eliminates the typical challenges presented by localized power grid interruptions, physical data loss due to catastrophic events and malicious on-site attacks to the IT infrastructure within an organization. The cloud computing framework provides an optimal environment for faster, safer and cheaper delivery of IT services within an enterprise.

Issues for Enterprise applications on Cloud

The need for quick adaptation: Today, in the field of enterprise applications development, the need to be flexible and respond instantly to changes is especially acute.

More strict security requirements: Enterprise information system is a complex structure. As the system grows, ensuring information security and protecting business-critical resources is becoming increasingly difficult.

Processing and storage of large amounts of data: The amount of data that companies generate daily is enormous. Even simple storing of all this information, not to mention analysing it, is itself a daunting task. Most of this data is not structured, which makes it difficult to find and analyze. As part of enterprise applications development, companies have two options for storing data: locally on their own or leased servers or in the cloud e.g. in the network. Both options have their advantages and disadvantages.

Integration with other systems: When using several business applications at the enterprise, it is often necessary to ensure their interaction and integration into the corporate information environment. System integration is a great way to optimize performance, as it provides the ability to view and update relevant information in real time.

Need for quality post-release support: If the corporate application does not work even for several minutes, it directly affects the income and productivity of the business. Thus, in case of an application failure, it is crucial that the software company can detect and fix the problem in hours, or even better in minutes.

Lack of skills in the development team: Finding specialists with skills that meet current needs is the main task for companies that need enterprise applications.

Challenges faced when Transitioning to the Cloud

Here is a closer look at some challenges faced during a cloud transition:

Managing Disruption
Disruptive cloud technologies are how companies like Uber and Airbnb have changed the transportation and travel industries. By leveraging new approaches and Big Data, these companies have transformed the industries they have entered. A cloud migration usually brings with it big changes to business processes and collaborations within a company.

Security and Compliance
For better or worse, companies with in-house systems were in control and responsible for security, both physical and digital, of the systems, machines, and data stored there. With the cloud, security is relegated to a third-party provider. Usually, these companies have robust systems to protect digital and physical systems. That's why it's imperative that companies be clear on what those protocols are when selecting a system. Similarly, compliance issues can be a headache in certain sectors such as healthcare, public sector, retail. Ensuring compliance among cloud providers is a new obligation for those companies shifting services.

Interoperability
For any new system, the ability to intersect with other areas of the company is of great importance. In cloud computing, it's no different. As data migrates into and out of the cloud, it's important for companies to understand how cloud systems will operate with on-premises systems that haven't or won't migrate. The same goes for IT staff, who will be interacting with cloud provider staff to ensure smooth operation and coordinate patching, updates, upgrades, and new services coming online.

Bandwidth and Performance
Costs are often a key driver for cloud implementation. With cloud solutions, companies can save considerably on hardware, storage, space, and cooling expenses. There may be efficiencies gained in use or number of staff for IT, data entry, and reporting functions. However, in an increasingly mobile world, the demand for data and applications may result in increased bandwidth costs that may not be fully predictable. Lack of proper bandwidth allocation can result in performance issues, disrupting credibility among employees and customers.

The evolution of enterprise cloud strategy

Adopting cloud computing is no longer a futuristic goal for enterprises but the norm. The Cloud is one of the principal drivers for digital transformation. Many organizations credit the cloud for their success in scaling heights in their business. Leveraging an enterprise cloud environment provides businesses with increased velocity in performance, superior security, and reduced cost.

History of cloud evolution

The early hints of cloud computing can be traced to way back to the 1950s when mainframes were shared by multiple users using time-sharing schedules. The central mainframe was connected to 'dumb terminals' which had no power of its own. Thus, began a primitive cloud computing system. In the 1960s, American scientist C.R. Licklider invented the ARPANET (Advanced Research Projects Agency Network) where resources were shared among a system of interconnected computers that were not in the same location. Among other things, the internet and the world wide web traces its origins back to the ARPANET.

Then came a series of ever-advancing cloud innovations. Virtualization was a new concept introduced by IBM in 1972, through a new operating system approach called the 'virtual machines' (VM). A virtual or an instance of a network or storage could be created thus allowing multiple instances of virtual machines to co-exist within a piece of infrastructure. VMware took the VM and transformed it into a thing of power and extreme utility.

In so many ways, this was the stepping stone for modern-day cloud computing. By the 1990s, business users could share resources in the same infrastructure using virtual private networks. Amazon introduced the Elastic Compute Cloud (EC2) in the 2000s. This allowed companies to hire virtual space to run their applications and programs. In the same decade, Google launched cloud document services which enabled users to edit and transfer documents from anywhere.

Over the last decade, cloud computing has seen explosive growth. This growth has greatly encouraged cloud transition, at the enterprise as well as SMB levels. This also saw the rise of providers who delivered SaaS (Software as a Service) offerings. Not only SaaS but this decade also saw the emergence of PaaS (Platform as a Service), IaaS(Infrastructure as a Service), BaaS(Backup as a Service), DRaaS(Disaster Recovery as a Service), CaaS(Container as a Service), and, now, XaaS(Everything as a Service). Literally, every service is available on the cloud, making it the lifeline of every enterprise.

Key trends shaping enterprise cloud strategy

Digital transformation lays the foundation for many innovations in cloud computing. The cloud provides an edge for companies looking to transform and gain competitive advantage. Cloud strategies are quickly penetrating into data analytics, application architecture blueprints, and managed infrastructures.

Let's check out a few vital trends transforming enterprises and their cloud strategies:

The rise of multi-cloud strategies
Multi-cloud strategies will play a key role in cloud transitions for enterprises. Typically, enterprises will use a private cloud to host vital data and public cloud for less critical workloads. This will offset the risk involved and greatly reduce IT expenditure. 451 Research suggests that nearly 69% of enterprises will have either multi-cloud or hybrid cloud IT environments by the end of this year.

Containerization on the cloud
The concept of containerization has paved the path for major changes in the cloud. Containers on the cloud make application management scalable, agile, lightweight, and serverless. With more and more enterprises adopting Docker containers, Gartner says, "more than 50% of global organizations will be running containerized applications in production by 2020"

Improved SaaS ecosystems
The Software as a Service market has greatly evolved. The growth and improvement have been driven by the need for companies to cut costs and improve efficiency. With better vendor management and decentralized application management, SaaS ecosystems are becoming richer, ever-more open, and uber-functional.

Orchestrated cloud framework
Contemporary multi-cloud environments have complex heterogeneous technology, security systems, and network capabilities. Cloud orchestration manages this complex workload on private and public clouds to deliver a seamless infrastructure. Workloads can be automated based on the user's specific needs, making the infrastructure highly scalable and adaptable. Task-automation and business workflow automation deliver a resilient cloud ecosystem.

Containerized microservices expansion
Enterprises are increasingly adopting containerized microservices for application development and deployment, Docker containers being the most popular. This allows resource-controlled, scalable, agile, and adaptable operating environments. Containerized microservices allows scaling out applications independently with more processing power and network bandwidth, based on-demand, without hampering other applications.
It has been a long journey for cloud computing since the days of bulky time-sharing mainframes to the present-day AWS frameworks and hybrid cloud environments. As businesses grow, so will the challenges. We may see another phase of innovations in cloud computing to address those challenges. But even today, the evolution of the cloud is already redefining the technology ethos. It's time for big changes and the businesses that embrace the Cloud will gain massively.

Cloud Supply Chain

The aim of every business is to get their products or services to target customers. Even if a business produces top notch products or services, it will be redundant if it cannot get them across to the people in need of these products or services. This is where supply chain comes in.

The supply chain refers to the system, collection and connection of resources, both human, material, financial, information, that are used in a product or service to customers from the producer. The supply chain includes activities, organizations, planning, management, etc. It also involves the coordination and communication between every part of the supply channel which includes middle men and intermediaries.

Given the importance of the supply chain to businesses, there have been several attempts to improve the efficiency of supply chains. Several new technologies have been utilized to increase the effectiveness, reach and performance of supply chains. One prominent one, however, is cloud computing. This technology has been on the rise in terms of application areas, and supply chain is one such areas.

Some organizations have been taking advantage of cloud computing in specific supply chain activities. There are certain activities which are most commonly focused on in supply chain. These activities include forecasting and planning, logistics, service and spare parts management, as well as sourcing and procurement. Many of these activities have been greatly improved by supply chain managers with the use of cloud computing. We will now proceed to look in detail at how all of these activities have been affected by cloud computing.

- **Forecasting and Planning**: This is one of the most integral parts of supply chain. Every company wants to be able to forecast demand and supply, so as to plan for production and supply. Cloud computing can be of help here by collecting and unifying information / data from customers, retailers, wholesalers, and even the producers themselves. Some form of analytics can then be performed on the data in order to generate accurate forecasts and enable companies plan well for the future.
- **Logistics:** One of the things that plagues supply chain managers is logistics. From inventory to transport management, poor logistics can be the downfall of even the best supply chain. Cloud computing can help with logistics by providing tracking operations, automatic inventory management and transport route optimization. The sharing and unification of data on the cloud can also help the various entities involved in logistics plan effectively.
- **Service and Spare parts Management:** The management of spare parts and servicing can easily be improved with cloud computing. Spare parts can be easily ordered at a minimum level of availability in order to prevent shortage. Also, the time for servicing can be automatically determined utilizing information on the cloud, leading to a reduction in downtime. RFID can also be used to track inventory location ensuring accountability.
- **Sourcing and Procurement**: Acquisition and procurement of materials in critical to any supply chain. Cloud platforms can act as a master database that will contain all the information of suppliers. These platforms can also automatically order from the best supplier once a set minimum level is reached. The development of contracts automatically is another benefit of using cloud computing in this supply chain activity.

Benefits of Cloud Computing in Supply chain

- **Intelligence and Automation:** The analytics and cognitive intelligence provided by cloud computing can totally transform the supply chain. Cloud platforms will provide insights into opportunities in the supply chain that can be harnessed. Potential risks can also be spotted by predictive analytics. It does not just stop in the suggestions and identification as the automation allows it to execute with operational efficiency. Automatic orders and supplies based on certain triggers are some examples of how cloud computing can automate supply chain.

- **Real time Visibility:** These days, several actions happen at the same time. Often times, the number of occurrences are too many to keep up with, resulting to an overload of information. However, cloud computing can ensure you stay up to date with the happenings in the supply chain. Real time visibility of occurrences and changes will help with the making of quick decisions about the supply chain, giving you advantage and improving operations.
- **Improved Scalability:** Cloud computing helps with the ease of scaling supply chain operations. Whether you have an increased number of staff or customer base, cloud computing will help with the scaling of supply chain operations in a very efficient and affordable manner. In periods of high demand, the supply chain can easily scale to cater for your needs.
- **Speed:** The speed with which companies can now perform supply chain operations and activities because of the efficiency provided by cloud computing is another benefit. The proactive identification and prediction of opportunities and risks will ensure that companies swiftly swing to action. Also, there is an enhanced responsiveness to increase in demand or other things.
- **Cost Efficiency:** The increased efficiency provided by cloud computing means that companies can save costs previously resulting from bottlenecks, logistics issues and other mishaps. The best part is that cloud computing does not require investments on in-house servers or software. Only a subscription is needed to improve the efficiency of your supply chain with cloud computing.

MODULE 3

Virtual Machine

A virtual machine (VM) is a digital version of a physical computer. Virtual machine software can run programs and operating systems, store data, connect to networks, and do other computing functions, and requires maintenance such as updates and system monitoring.

Multiple VMs can be hosted on a single physical machine, often a server, and then managed using virtual machine software. This provides flexibility for compute resources (compute, storage, network) to be distributed among VMs as needed, increasing overall efficiency. This architecture provides the basic building blocks for the advanced virtualized resources we use today, including cloud computing.

Now, with the emergence of virtualization technology and the cloud computing IaaS model, it is just a matter of minutes to achieve the same task. All you need is to provision a virtual server through a self-service interface with small steps to get what you desire with the required specifications.

Provisioning this machine in a public cloud like Amazon Elastic Compute Cloud (EC2), or using a virtualization management software package or a private cloud management solution installed at your data center in order to provision the virtual machine inside the organization and within the private cloud setup.

Previously, whenever there was a need for performing a server upgrade or performing maintenance tasks, you would exert a lot of time and effort, because it is an expensive operation to maintain or upgrade a main server that has lots of applications and users. Now, with the advance of the revolutionized virtualization technology and migration services, these tasks are very easy and need no time to accomplish.

Provisioning a new virtual machine is a matter of minutes, saving lots of time and effort. Migrations of a virtual machine is a matter of milliseconds.

Virtual Machine Life Cycle

- The cycle starts by a request delivered to the IT department, stating the requirement for creating a new server for a particular service.
- This request is being processed by the IT administration to start seeing the servers resource pool, matching these resources with requirements.
- Starting the provision of the needed virtual machine.
- Once it provisioned and started, it is ready to provide the required service according to an SLA.
- Virtual is being released; and free resources.

VM Provisioning Process

The common and normal steps of provisioning a virtual server are as follows:

- Firstly, you need to select a server from a pool of available servers (physical servers with enough capacity) along with the appropriate OS template you need to provision the virtual machine.
- Secondly, you need to load the appropriate software (operating System you selected in the previous step, device drivers, middleware, and the needed applications for the service required).
- Thirdly, you need to customize and configure the machine (e.g., IP address, Gateway) to configure an associated network and storage resources.
- Finally, the virtual server is ready to start with its newly loaded software.

To summarize, server provisioning is defining server's configuration based on the organization requirements, a hardware, and software components. Normally, virtual machines can be provisioned by manually installing an operating system, by using a preconfigured VM template, by cloning an existing VM, or by importing a physical server or a virtual server from another hosting platform. Physical servers can also be virtualized and provisioned using P2V (Physical to Virtual) tools and techniques (e.g., virt- p2v). After creating a virtual machine by virtualizing a physical server, or by building a new virtual server in the virtual environment, a template can be created out of it. Most virtualization management vendors (VMware, XenServer, etc.) provide the data center's administration with the ability to do such tasks in an easy way.

The Anatomy of Cloud Infrastructures

Here we focus on the subject of IaaS clouds and, more specifically, on the efficient management of virtual machines in this type of cloud.

There are many commercial IaaS cloud providers in the market, such as those cited earlier, and all of them share five characteristics:

- They provide on-demand provisioning of computational resources.
- They use virtualization technologies to lease these resources.
- They provide public and simple remote interfaces to manage those resources.

- They use a pay-as-you-go cost model, typically charging by the hour.
- They operate data centers large enough to provide a seemingly unlimited number of resources to their clients.

Distributed Management of Virtual Infrastructures

- Managing VMs in a pool of distributed physical resources is a key concern in IaaS clouds, requiring the use of a virtual infrastructure manager.
- OpenNebula is capable of managing groups of interconnected VMs—with support for the Xen, KVM, and VMWare platforms—within data centers and private clouds that involve a large amount of virtual and physical servers.
- OpenNebula can also be used to build hybrid clouds by interfacing with remote cloud sites.

The primary target of OpenNebula is to manage VMs. Within OpenNebula, a VM is modeled as having the following attributes:

- A capacity in terms of memory and CPU.
- Aset of NICs attached to one or more virtual networks.
- A set of disk images. In general it might be necessary to transfer some of these image files to/from the physical machine the VM will be running in.
- A state file (optional) or recovery file that contains the memory image of a running VM plus some hypervisor-specific information.

The life cycle of a VM within OpenNebula follows several stages:

Resource Selection: Once a VM is requested to OpenNebula, a feasible placement plan for the VM must be made. OpenNebula's default scheduler provides an implementation of a rank scheduling policy, allowing site administrators to configure the scheduler to prioritize the resources that are more suitable for the VM, using information from the VMs and the physical hosts.

Resource Preparation: The disk images of the VM are transferred to the target physical resource. During the boot process, the VM is contextualized, a process where the disk images are specialized to work in a given environment. Different techniques are available to contextualize a worker node, including use of an automatic installation system (for instance, Puppet or Quattor), a context server, or access to a disk image with the context data for the worker node (OVF recommendation).

VM Termination: When the VM is going to shut down, OpenNebula can transfer back its disk images to a known location. This way, changes in the VM can be kept for a future use.

OpenNebula manages a VMs life cycle by orchestrating **three** different management areas:

1. **Virtualization** by interfacing with a physical resource's hypervisor, such as Xen, KVM, or VMWare, to control (e.g., boot, stop, or shutdown) the VM.

2. **Image management** by transferring the VM images from an image repository to the selected resource and by creating on-the-fly temporary images.

3. **Networking** by creating local area networks (LAN) to interconnect the VMs and tracking the MAC addresses leased in each network.

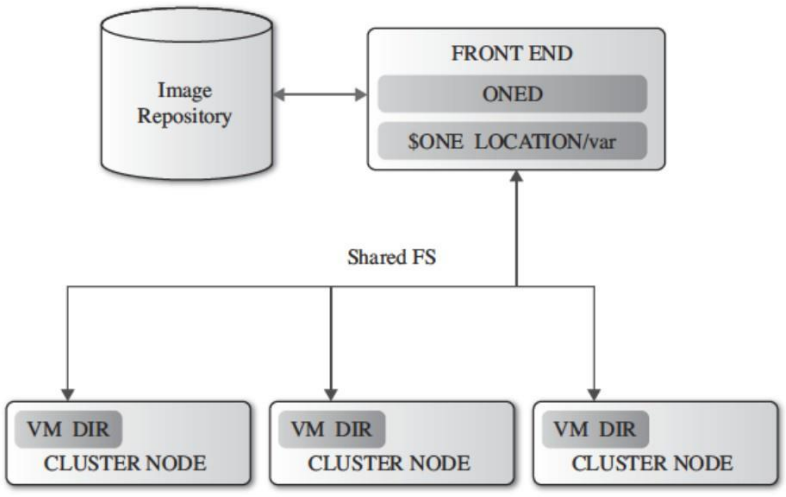


FIGURE 6.1. Image management in OpenNebula.

- **Image Repositories** refer to any storage medium, local or remote, that hold the base images of the VMs.
- **Virtual Machine Directory** is a directory on the cluster node where a VM is running. This directory holds all deployment files for the hypervisor to boot the machine, checkpoints, and images being used or saved.

Any given VM image goes through the following steps along its life cycle:

1. **Preparation** implies all the necessary changes to be made to the machine's image.
2. **Cloning** the image means taking the image from the repository and placing it in the VM's directory in the physical node where it is going to be run before the VM is actually booted.
3. **Save/Remove** the images and all the changes.

Scheduling Techniques for Advanced Reservation of Capacity

While a VI manager like OpenNebula can handle all the minutiae of managing VMs in a pool of physical resources, scheduling these VMs efficiently is a different and complex matter. **Immediate provisioning model** is used by commercial cloud providers, such as Amazon, since their data centers' capacity is assumed to be infinite. **Best effort provisioning** is used where requests have to be queued and prioritized. **Advance provisioning** is used where resources are pre-reserved so they will be guaranteed to be available at a given time period.

Enhancing Cloud Computing Environments Using A Cluster as A Service

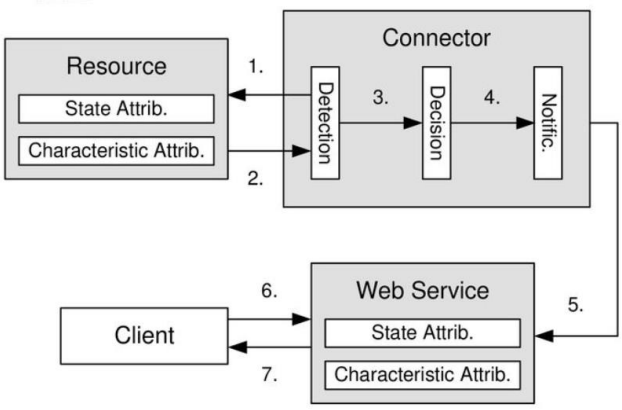
RVWS DESIGN

Dynamic Attribute Exposure

There are two categories of dynamic attributes addressed in the RVWS (Resource Via Web Service) framework:

1. **State attributes** cover the current activity of the service and its resources, thus indicating readiness.
2. **Characteristic attributes** cover the operational features of the service, the resources behind it, the quality of service (QoS), price and provider information.

To keep the stateful Web service current, a Connector is used to detect changes in resources and then inform the Web service. The Connector has three logical modules: Detection, Decision, and Notification. The Detection module routinely queries the resource for attribute information. Any changes in the attributes are passed to the Decision module that decides if the attribute change is large enough to warrant a notification. This prevents excessive communication with the Web service. Updated attributes are passed on to the Notification module, which informs the stateful Web service that updates its internal state. When clients requests the stateful WSDL document, the Web service returns the WSDL document with the values of all attributes at the request time.



Exposing Resource Attributes

Cluster as A Service: The Logical Design

Simplification of the use of clusters could only be achieved through higher layer abstraction that is proposed here to be implemented using the service-based Cluster as a Service (CaaS) Technology. The purpose of the CaaS Technology is to ease the publication, discovery, selection, and use of existing computational clusters.

CaaS Overview

The exposure of a cluster via a Web service is intricate and comprises several services running on top of a physical cluster.

A typical cluster is comprised of three elements: **nodes, data storage, and middleware**.

The middleware virtualizes the cluster into a single system image; thus resources such as the CPU can be used without knowing the organization of the cluster.

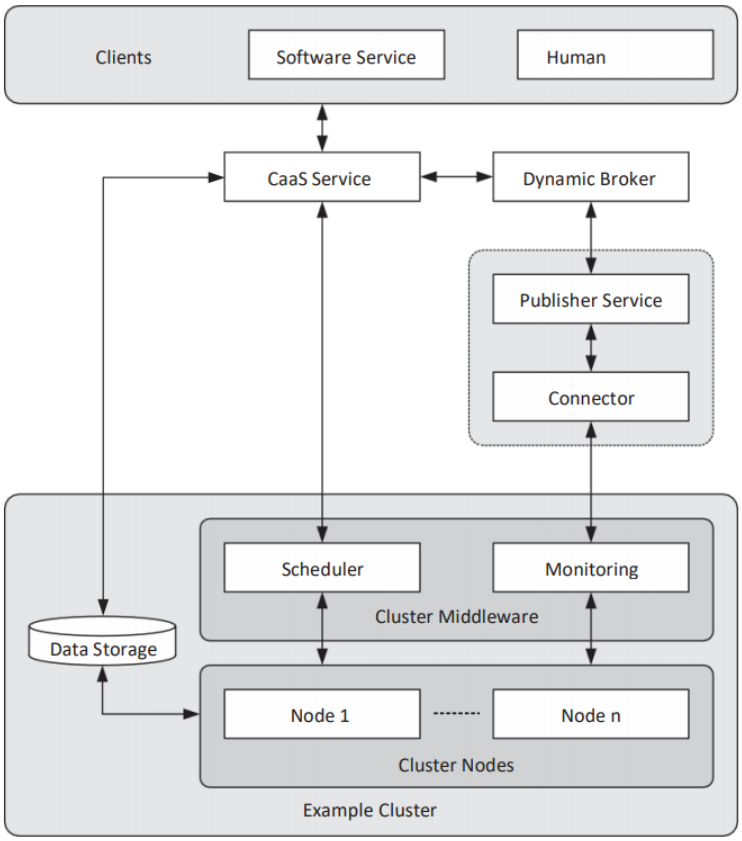
To make information about the cluster publishable, a Publisher Web service and Connector were created using the RVWS framework. The purpose of the publisher Web service was to expose the dynamic attributes of the cluster via the stateful WSDL document. Furthermore, the Publisher service is published to the Dynamic Broker so clients can easily discover the cluster.

To find clusters, the CaaS Service makes use of the Dynamic Broker. While the Broker is detailed in returning dynamic attributes of matching services, the results from the Dynamic Broker are too detailed for the CaaS Service. Thus another role of the CaaS Service is to summarize the result data so that they convey fewer details.

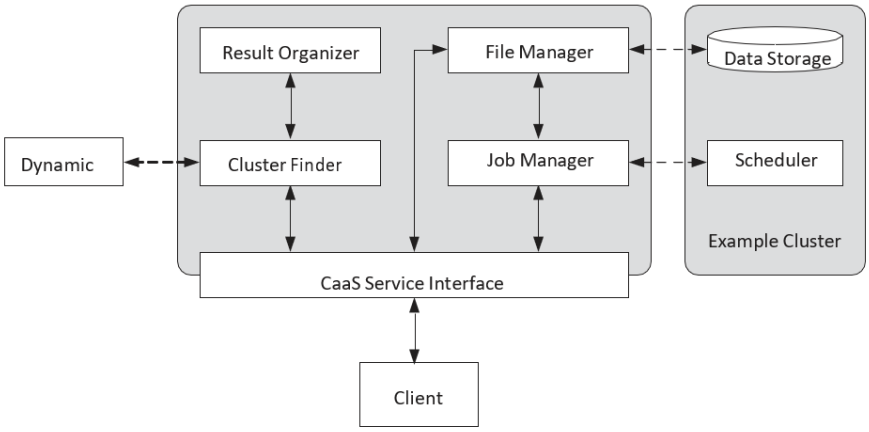
Ordinarily, clients could find required clusters but they still had to manually transfer their files, invoke the scheduler, and get the results back. All three tasks require knowledge of the cluster and are conducted using complex tools.

The role of the CaaS Service is to provide easy and intuitive file transfer tools so clients can upload jobs and download results and also offer an easy to use interface for clients to monitor their jobs. The CaaS Service does this by allowing clients to upload files as they would any Web page while carrying out the required data transfer to the cluster transparently.

Because clients to the cluster cannot know how the data storage is managed, the CaaS Service offers a simple transfer interface to clients while addressing the transfer specifics. Finally, the CaaS Service communicates with the cluster's scheduler, thus freeing the client from needing to know how the scheduler is invoked when submitting and monitoring jobs.



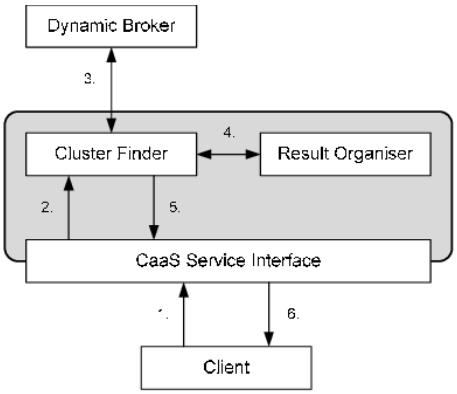
Complete CaaS system.



CaaS Service Design

Cluster Discovery

Before a client uses a cluster, a cluster must be discovered and selected first. To start, clients submit cluster requirements in the form of attribute values to the CaaS Service Interface (1). The requirements range from the number of nodes in the cluster to the installed software (both operating systems and software APIs). The CaaS Service Interface invokes the Cluster Finder module (2) that communicates with the Dynamic Broker (3) and returns service matches (if any).

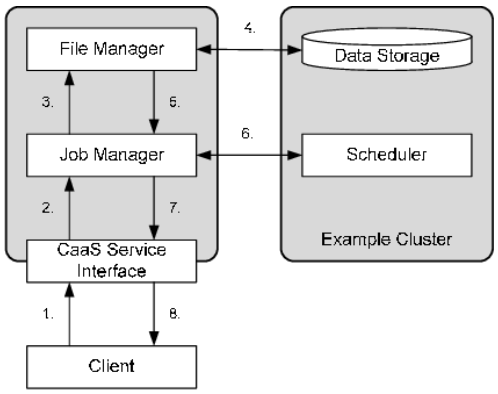


Cluster Discovery

To address the detailed results from the Broker, the Cluster Finder module invokes the Results Organizer module (4) that takes the Broker results and returns an organized version that is returned to the client (5—6). The organized results instruct the client what clusters satisfy the specified requirements. After reviewing the results, the client chooses a cluster.

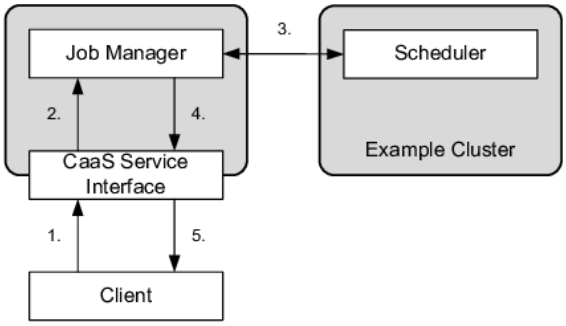
Job Submission.

After selecting a required cluster, all executables and data files have to be transferred to the cluster and the job submitted to the scheduler for execution. As clusters vary significantly in the software middleware used to create them, it can be difficult to place jobs on the cluster. To do so requires knowing how jobs are stored and how they are queued for execution on the cluster.



Job Submission

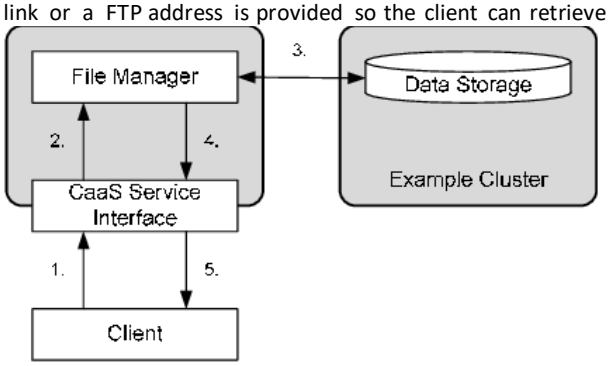
Job Monitoring. During execution, clients should be able to view the execution progress of their jobs. Even though the cluster is not the owned by the client, the job is. Thus, it is the right of the client to see how the job is progressing and (if the client decides) terminate the job and remove it from the cluster.



Job Monitoring

Result Collection. The final role of the CaaS Service is addressing jobs that have terminated or completed their execution successfully. In both cases, error or data files need to be transferred to the client.

Clients start the error or result file transfer by contacting the CaaS Service Interface (1) that then invokes the File Manager (2) to retrieve the files from the cluster's data storage (3). If there is a transfer error, the File Manager attempts to resolve the issue first before informing the client. If the transfer of files (3) is successful, the files are returned to the CaaS Service Interface (4) and then the client (5). When returning the files, URL link or a FTP address is provided so the client can retrieve the files.



Result Collection

Secure Distributed Data Storage in Cloud Computing

CLOUD STORAGE: FROM LANs TO WANs

Cloud computing has been viewed as the future of the IT industry. It will be a revolutionary change in computing services. Users will be allowed to purchase CPU cycles, memory utilities, and information storage services conveniently just like how we pay our monthly water and electricity bills. However, this image will not become realistic until some challenges have been addressed. In this section, we will briefly introduce the major difference brought by distributed data storage in cloud computing environment. Then, vulnerabilities in today's cloud computing platforms are analyzed and illustrated.

Most designs of distributed storage take the form of either storage area networks (SANs) or network-attached storage (NAS) on the LAN level, such as the networks of an enterprise, a campus, or an organization. SANs are constructed on top of block-addressed storage units connected through dedicated high-speed networks. In contrast, NAS is implemented by attaching specialized file servers to a TCP/IP network and providing a file-based interface to client machine . For SANs and NAS, the distributed storage nodes are managed by the same authority. The system administrator has control over each node, and essentially the security level of data is under control. The reliability of such systems is often achieved by redundancy, and the storage security is highly dependent on the security of the system against the attacks and intrusion from outsiders. The confidentiality and integrity of data are mostly achieved using robust cryptographic schemes.

Existing Commercial Cloud Services

In normal network-based applications, user authentication, data confidentiality, and data integrity can be solved through IPSec proxy using encryption and digital signature. The key exchanging issues can be solved by SSL proxy. These methods have been applied to today's cloud computing to secure the data on the cloud and also secure the communication of data to and from the cloud. The service providers claim that their services are secure.

Amazon's Web Service provides Infrastructure as a Service (IaaS) with different terms, such as Elastic Compute Cloud (EC2), SimpleDB, Simple Storage Service (S3), and so on. They are supposed to ensure the confidentiality, integrity, and availability of the customers' applications and data. The downloading process is similar to the uploading process. The user creates a manifest and signature file, e-mails the manifest file, and ships the storage device attached with signature file. When Amazon receives these two files, it will validate the two files, copy the data into the storage device, ship it back, and e-mail to the user with the status including the MD5 checksum of the data. Amazon claims that the maximum security is obtained via SSL endpoints.

Microsoft Windows Azure. The Windows Azure Platform (Azure) is an Internet-scale cloud services platform hosted in Microsoft data centers, which provides an operating system and a set of developer services that can be used individually or together. The platform also provides scalable storage service.

Technologies for Data Security in Cloud Computing

Database Outsourcing and Query Integrity Assurance: Researchers have pointed out that storing data into and fetching data from devices and machines behind a cloud are essentially a novel form of database outsourcing.

Data Integrity in Untrustworthy Storage: One of the main challenges that prevent end users from adopting cloud storage services is the fear of losing data or data corruption. It is critical to relieve the users' fear by providing technologies that enable users to check the integrity of their data.

Web-Application-Based Security: Once the dataset is stored remotely, a Web browser is one of the most convenient approaches that end users can use to access their data on remote services. In the era of cloud computing, Web security plays a more important role than ever.

Multimedia Data Security: With the development of high-speed network technologies and large bandwidth connections, more and more multi- media data are being stored and shared in cyber space. The security requirements for video, audio, pictures, or images are different from other applications.

MODULE 4

Workflow management system

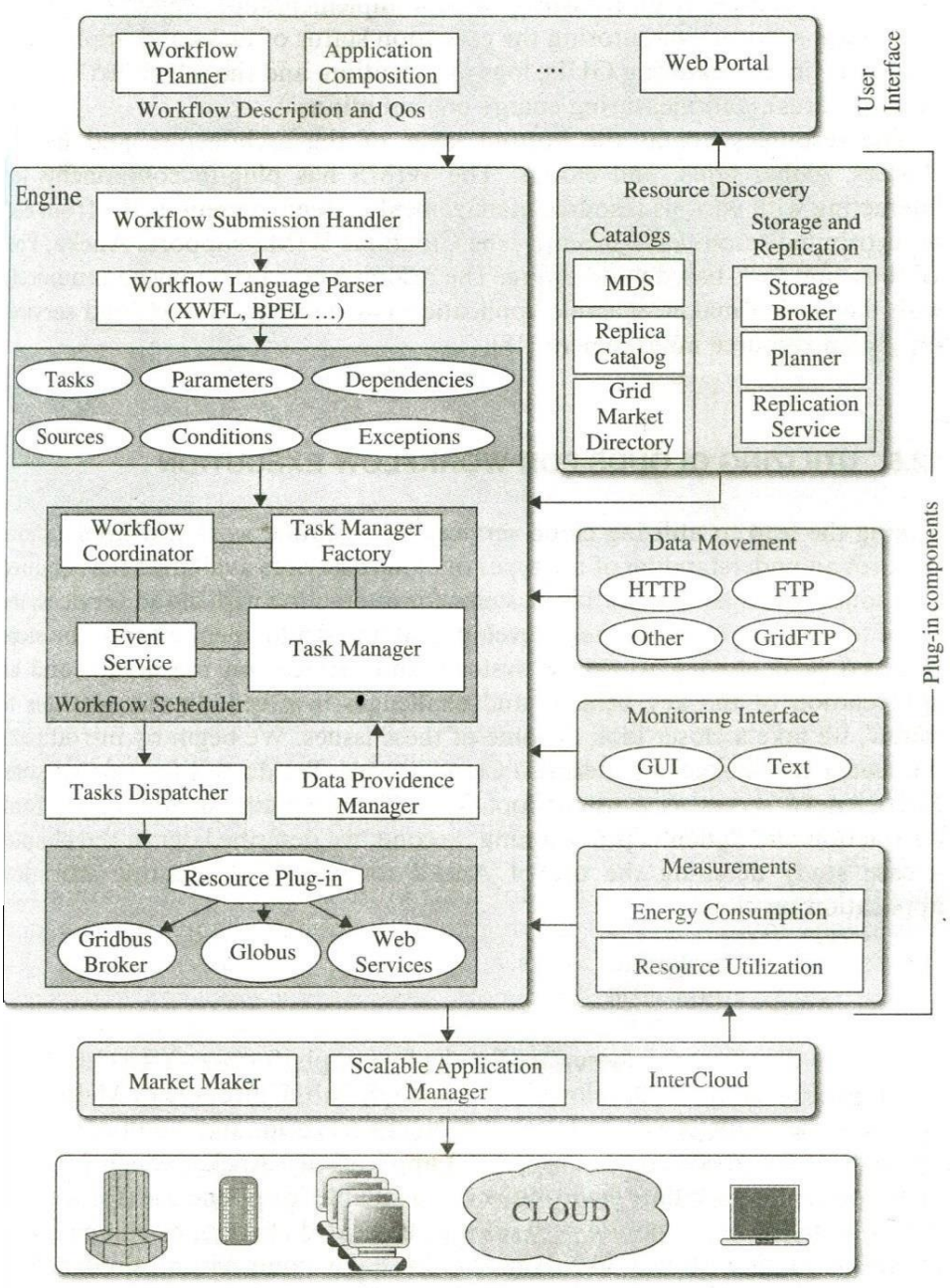
A workflow management system is defined as a process consisting of services of steps that simplifies the execution and management of cloud application.

Architecture of Workflow management system

Scientific applications are modeled as work flows consisting of tasks, data elements, control sequences and data dependencies.
Workflow management systems are responsible for managing and executing these workflows. The Cloud bus workflow management system consists of components that are responsible for
Landlines tasks, data and resources.

The architecture consists of three major parts.

- The user Interface
- The Core System
- The Plug-ins



Architecture of Workflow Management System

User Interface

The user interface allows end users for work with

- Workflow Composition
- Workflow Execution Planning
- Submission
- Monitoring

These features are delivered through a web portal or through stand-alone applications that is installed of the user's end.

The Core System

The core components are responsible for managing the execution of workflows. They facilitate the transaction of high-level workflow description into task and data objects.

These objects are used by the execution sub system. The scheduling component applies user selected scheduling policies and plans to the work flow of various stages on their execution.

Plug-ins

The Plug-ins support workflow execution on different environments and platforms. The plug-ins are used for querying task transferring the execution status of tasks and applications and measuring the energy consumption.

The resources are at the bottom layer of the architecture which includes clusters global grids and clouds. The resources managers may communicate with the:

- Market Maker
- Sealable Application Manager
- Intercloud Services for global resource manager

Aneka

Aneka is a software platform and framework for developing distributed applications on the cloud. Aneka provides developers with a rich set of API's for using the resources by expressing application logic with a variety of programming abstraction. It is a workflow management tool.

Aneka is a distributed middleware for deploying platform as a service. Aneka, which is both a development and runtime environment, is available for public use for a cost. It can be installed on corporate networks or dedicated clusters or it can be hosted on Infrastructure clouds like Amazon Ec2 to support work flow management.

Aneka was developed on Microsoft.Net framework 2.0 and is compatible with other implementation. Aneka can run on popular platforms such as Microsoft windows, Linux and Macosx.

The runtime environment consists of a Aneka containers running on physical or virtualized models. Each of these containers can be configured to play a specific role such as scheduling or execution.

The Aneka service stack provides services for infrastructure management, application execution management, accounting, licensing and security.

Dynamic resource Provisions:

This service enables horizontal scaling depending on the overall load in the cloud. The platform is elastic in nature and can provision additional resources on-demand from external physically virtualized resource pools, in order to meet the QOS requirements of applications.

Development Environment:

This development environment provides a rich set of API's for developing applications that can utilize free resources of the infrastructure. These APIs expose different programming abstractions, such as the Task Model, Thread Model and Map Reduce.

Storage Service:

The storage service provides a temporary repository for application files such as input files that are required for task execution and output files that are the result of execution.

Any output files produced as a result of the execution are uploaded back to the storage servers. From here they are staged-out to the remote storage location.

Scientific Application

Scientific computing involves the construction of mathematical models and numerical solution techniques to solve scientific, social scientific and engineering problems

Classification of Scientific Applications and services in the Cloud

Scientific computing involves the construction of mathematical models and numerical solution techniques to solve scientific, social scientific and engineering problems.

There are three layers consists in the cloud as:

1. Software as a Service (SaaS)
2. Platform as a Service (PaaS)
3. Infrastructure as a Service (IaaS)

Software as a Service:

- It provides ready-to-Run services that are deployed and configured for the users. User has no control over the underlying cloud infrastructure.
- No client-side software required.
- All data manipulated in remote infrastructure.
- It can be used Scientific Portal or a visualization tool. Service Providers: Google Apps, Sales Force.

Example: TeraGrid Science Gateway

Platform as a Service:

- PaaS model provides the capability of developing application using programming tools like Java, Python.

Example: Microsoft Dryad, Google Map Reduce

Infrastructure as a Service:

- Majority of scientific applications rely on IaaS cloud services.
- Only IaaS provides sufficient programmatic control to express decomposition and Dynamic execution modes that are more important for scientific Applications.

Example: Amazon Web Services, Tata Communications

SAGA

SAGA means Simple API for Grid Applications. It is used to develop scientific applications in the cloud. It provides framework for implementing higher level programming for Scientific Applications.

SAGA-Based scientific application tools in cloud environment

SAGA is used to develop scientific application that can utilize the cloud infrastructure from vanilla cloud such as EC2 to open-source cloud such as Eucalyptus. It provides framework for implementing higher level programming for Scientific Applications. It is the ultimate approach to develop scientific applications in cloud platform.

The SAGA-Based Scientific application tools are

1. SAGA MapReduce
2. SAGA Montage

SAGA MapReduce

- SAGA MapReduce provides Application Development and runtime environment for scientific Applications.
- It gives maximum control over the deployment, Distribution and runtime decomposition. It is a prominent tool for PaaS Model.
- MapReduce job usually splits the input data-set into independent chunks which are processed by the map tasks in a completely parallel manner. The framework sorts the outputs of the maps, which are then input to the reduce tasks.
- The framework takes care of scheduling tasks, monitoring them and re-executes the failed tasks.

SAGA Montage

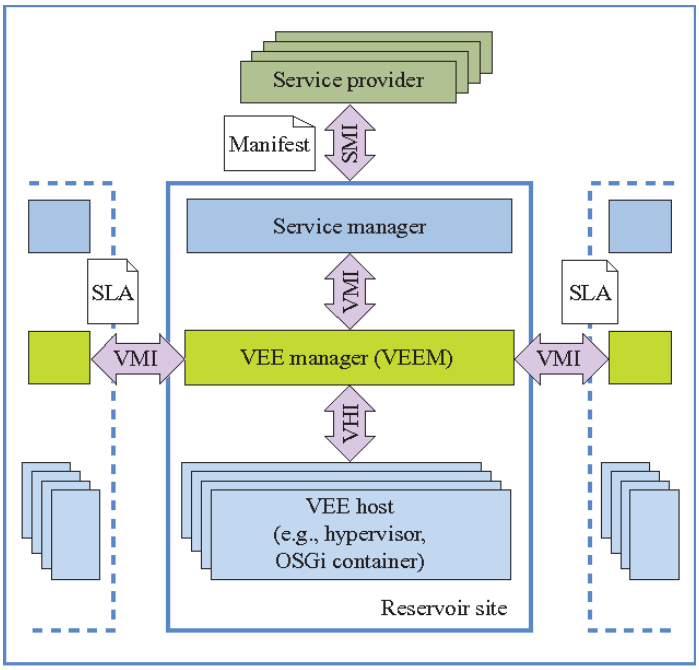
- It is designed to take multiple astronomical images from telescope and other instrument and stitch them together.
- Montage is a set of programming modules or tools, executable program that can run on a single computer, parallel or distributed system.

Federal Cloud Computing

A federated cloud computing is the deployment and management of multiple external and internal cloud computing services to match business needs.

Federated Cloud Model

All cloud computing providers have a finite capacity of resources. To grow beyond this capacity, cloud computing providers should be able to form federation of providers such that they can collaborate and share their resources.
Each provider can buy or sell on-demand resources from other providers.



Federated Cloud Model Components

1. Infrastructure Providers (IPs)
2. Service Providers (SPs)
3. Service Application Software
4. Virtual Execution Environment Manager(VEEM)
5. Virtual Execution Environment Host(VEEH)
6. Service Management Interface
7. VEE Management Interface

Infrastructure Providers(IPs):

Infrastructure Providers provide seemingly infinite pool of computational, Within each IP, optimal resource utilization is achieved by partitioning physical resources through virtualization layer into Virtual Execution Environment.

Service Providers(SPs):

Service Providers lease the resources from Infrastructure Providers and these resources to customers. There is a contract and SLA between SPs and IPs

Service Application Software:

It is a set of software components that collectively to achieve a common goal. Each component of such service applications executes in a dedicated VEE.

Virtual Environment Manager(VEEM):

It is a fully isolated runtime environments that abstract away the physical characteristics of the resources and enable resource sharing. It is responsible for the optimal placement of VEEs into VEE Hosts to optimize the process. It is free to place and move VEEs anywhere even on the remote site as long as the placement satisfies the constraints. It is responsible for the federation of remote sites.

Virtual Execution Environment Host(VEEH) :

It consists of:

- Virtualized Computational Resources
- Virtualization Layer(Hypervisor)
- All Management enablement Components.

It is responsible for the basic control and monitoring of VEEs and their resources such as

- Creating a VEE
- Allocating resources to VEE
- Monitoring VEE
- Migrating a VEE

VEEs belonging to the same application may be placed on multiple VEEHs. VEEHs must support transparent VEE migration to any compatible VEEH within the Federated Cloud, regardless of the site location.

Service Management Interface :

Each layer will be able to interact with each other using SMI with its service manifest.

VEE Management Interface :

VMI's support of VEEM-to-VEEM communication that simplifies cloud federation interoperability. It also implements different and independent IT optimization strategies.

SLO

A service level objective (SLO) is a criteria that is used to evaluate the performance of a business or technology service.

The existing approaches used to measure the Quality Of Service(QOS) in cloud are:

1.Load Balancing Approach

2.Admission Control Approach

Load Balancing Approach:

Cloud load balancing is the process of distributing workloads and computing resources in a cloud computing environment. Load balancing allows enterprises to manage application or workload demands by allocating resources among multiple computers, networks or servers. Load Balancing algorithm is executed on a physical machine (front-end node) that interfaces with clients. This machine receives the incoming request and distribute these requests to different physical machines (Back-end nodes). These backend nodes are responsible for serving the incoming requests.

Load Balancing Algorithm:

Load balancing algorithm is categorized into three:

1. Class-Agnostic
2. Class-Aware
3. Admission Control

Class-Agnostic: The front-end node not aware of the category of the request such as browsing, payment, selling, buying etc.

Class-Aware: This algorithm aware the type of request send from the clients and decide which back-end server should execute.

Admission Control: Admission Control Algorithm play an important role in deciding set of request that should be admitted into the application server which is very heavy load.

1.Request-based Algorithm
2.Session-based algorithm

Request-Bases Algorithm: It rejects new requests if the servers running to their capacity.

Session-Based Algorithm: Once the session is admitted into the server, all future requests belonging to that session are admitted and new sessions are rejected.

SLA

A service-level agreement (SLA) is a contract between an external service provider and its customers or between ISPa and SPs. A service level agreement, or SLA, is a formal set of service commitments made to a customer by a service provider.

Types of SLA

There are two types of SLA:

1. Infrastructure SLA
2. Application SLA

Infrastructure SLA:

Infrastructure Service provider offers guarantees on availability of the infrastructure such as

- Server Machine
- CPU Power
- Network Connectivity

Application SLA:

- Service providers are flexible in allocating and de-allocating computing resources among the application hosted in the server.
- It is agreement between a customer and service providers indicating performance criteria and cost structure.
- Application SLA may not be a single SLA and it may be set of SLAs based on service performance.
- Customer can switch from one SLA to another SLA during run time.

Life Cycle of SLA

Each SLA goes through a sequence of steps. Such sequence of step is called SLA life Cycle and consists of the following five phases:

1. Contract Definition
2. Publishing and Discovery
3. Negotiation
4. Operationalization
5. De-Commissioning

Contract Definition: Generally, Service Providers define a set of service offering in standard templates. These templates may be in the form of on-line catalog. This is called SLA Template.

Publishing and Discovery: Service Providers advertise these base services offering through standard publication media. Customers can search different competitive offering and shortlist a few that fulfill their requirements.

Negotiation: Once the customer has discovered a service provider , before signing the agreement , both parties engages in negotiation to be mutually agreed upon.

Operationalization: It is known as SLA Enforcement or execution of SLA agreement. It includes:

1. Monitoring
2. Performance Metrics
3. Identifying deviation and correcting
4. Detecting SLA violation and provide the penalty paid option.

De-Commissioning: It involves termination of all activities performed under particular SLA when the hosting relationship between service providers and customers end.

SLA Management activities

SLA Management of applications hosted on cloud platform involves five phases

1. Feasibility Study
2. On-Boarding Application
3. Pre-Production
4. Production
5. Termination

Feasibility Study

Feasibility study on hosting application on cloud platforms.
Three kinds of feasibility:
1.Technical Feasibility
2.Infrastrucure Feasibility
3.Financial Feasibility

Technical Feasibility:

- Ability of application to scale out.
- Compatibility of Data Centre.
- Availability of specific hardware and software required.
- Preliminary information on application performance.

Infrastructure Feasibility:

- Determining the availability of infrastructural resources in sufficient quantity to meet the projected demand.

Financial Feasibility:

- It involves determining approximate cost to be incurred by MSP. MSP Means Managed Service Providers

On-Boarding Application

- Once the customer and MSP agree to host the application based on feasibility study, the application is moved from customer server to Cloud platform. This moving activity is called On- Boarding.

Pre-Production

- Once the policies are completed, the application is hosted in a simulated production environment.
- Customer verifies and validated the application performance and other details given in the SLA.
- Once both parties agree on the cost, term and conditions of the SLA, MSP allows the application to go on live.

Production

- In this phase, the application is made accessible to its end users under the agreed SLA.

Termination

- When the customer wishes to withdraw the hosted application and doesn't wish to continue to use the services of the MSP, the termination activity is initiated.

Automated Policy Based Management

There are three types of policy. They are:

1. Business Policy
2. Operational Policy
3. Provisioning Policy

Business Policy:

Business Policy help prioritize access to the resources. It includes

1. Application class- Platinum, Gold, Silver
2. Whether application breaching the SLA
3. Whether application has already breached the SLA
4. Number of applications breached by the same customer
5. Number of applications about to breach by the same customer.
6. Type of action to be taken

Operational Policy:

It specifies functional relationship between the infrastructural attributes and SLA goals. It helps identifying the quantum of resources to be allocated to various parts of applications.

Provisioning Policy:

It helps to identifying a sequence of actions corresponding to user request such as

1.Scale-in
2.Scale-out
3.Start
4.Stop
5.Suspend
6.Resume

Automation of Policies

Automation of these policies are done by various software components such as

1. Prioritization Engine
2. Provisioning Engine
3. Rules Engine
4. Monitoring System
5. Auditing
6. Accounting /Billing System

Prioritization Engine: Identifying user requests based on priority and make it executed. It is a business Policy

Provisioning Engine: It is based on provision policy which makes set of necessary steps to carry out application by providing resources.

Rules Engine: This component works based on operation policy. It defined sequence of action to be taken under different condition based on SLA.

Monitoring System: It collects the defined metrics from SLA and these metrics are used for monitoring resource failure and evaluating operational policy.

Auditing: The predefined SLA should be monitored and recorded. Non-compliance leads to strict penalties.

Accounting and Billing System: It is used to make bill based on Payment Model, Resources Utilized, Fixed Cost and Recurring Cost.

Fundamental Concept of MapReduce

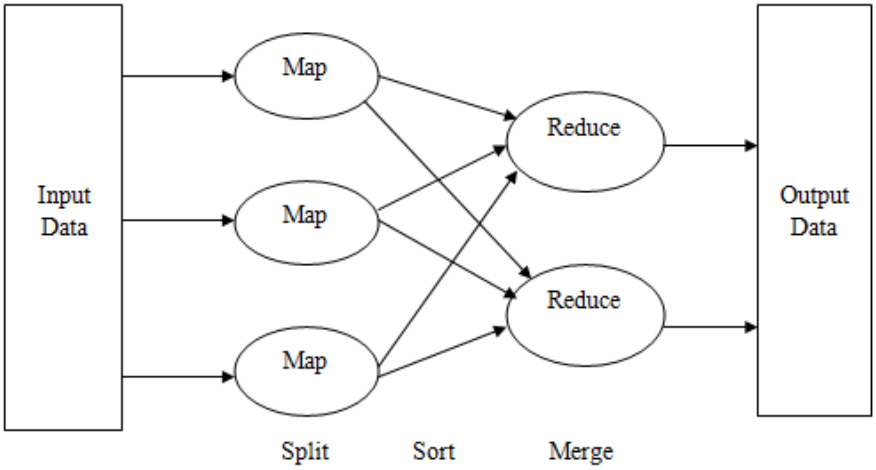
Google facilitated MapReduce as a programming framework in order to analyze huge amount of data. With the help of MapReduce developers can perform complex computations in a simple way and can hide the details of data distribution, parallelization and fault tolerance. In MapReduce programming model every map reduce are independent of other ongoing maps and reduces and the operation runs in parallel on different key and values. The most important aspect of MapReduce processing model is that the map tasks are carried on the nodes where the data lives. This ensures that there is very little or no movement of data between nodes.

Fundamental phases of MapReduce programming model are:

Map phase: the input data is divided into M Map functions called as Mapper. Mappers run in parallel and the output of MapReduce is intermediate key and value pairs.

Shuffle and Sort phase: output from the mappers is partitioned by hashing the output key. Here the number of partitions is equal to the number of reducers. In the shuffle phase, all key and value pairs share the same key that belong to the same division. Each division is stored by a key to merge all values of that key, after partitioning the MapReduce.

Reduce phase: output from the second phase are portioned into R Reduce function called as Reducer. Reducers process different in-between keys and also run in parallel.



Phases of MapReduce

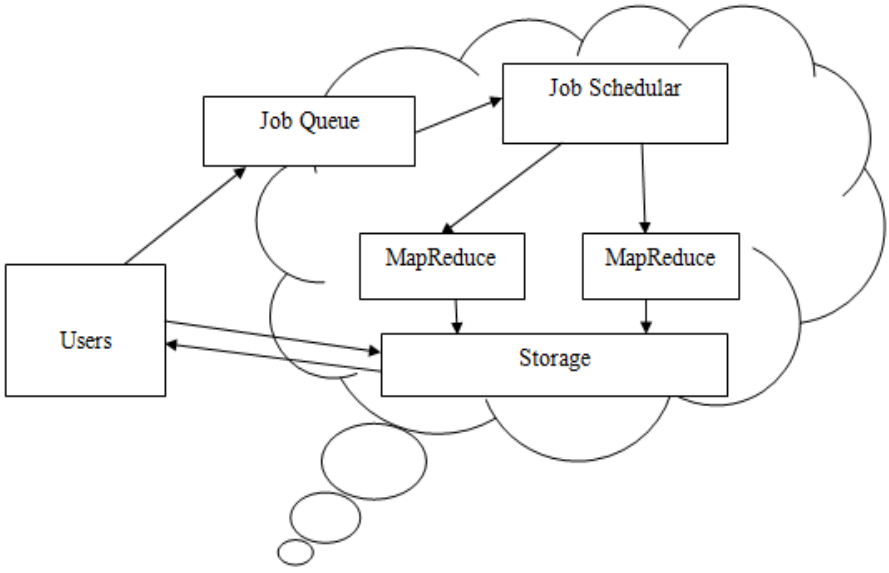
MapReduce In Cloud Computing

Cloud computing enable us to store, collect, share and transfer large amounts of data at very high speeds in a flawless and transparent manner that would out of complete necessity order all data to be "totally virtual". Thus, all data in cloud computing captures the concept of data virtualization through a new programming model which treats all data as a single entity through a process called MapReduce. MapReduce is widely used for big data processing in cloud platforms. Hadoop an open-source implementation of MapReduce hides the complexity of parallel execution across hundreds of servers in a cloud environment. It allows developers to process terabytes of data. How parallel programming work away from the developer is the main reason of using MapReduce with cloud computing.

Amazon cloud is using cloud MapReduce as a MapReduce implementation.

Cloud MapReduce has three main advantages over other MapReduce implementations built on traditional OS:

- Cloud MapReduce implementation is faster than other implementations.
- It has high scalability and failure resistance.
- It has simple line of code.



Cloud Computing with MapReduce

MODULE 5

Grid Computing

Grid computing is a processor architecture that combines computer resources from various domains to reach a main objective.

High Performance Computing (HPC)

High Performance Computing most generally refers to the practice of aggregating computing power the use of parallel processing for running advanced application programs efficiently, reliably and quickly.

Example: GRID

S.No	GRID COMPUTING	CLOUD COMPUTING
1	Distributed resource of computational Power	Centralized resource of Computational power.
2	Grid forms cluster of Physical resources which are faster.	Virtual cluster of resources are economic but slow
3	Parallel processing possible with physical cluster resource.	Cloud doesn't support parallel processing.
4	Grid technology was designed using bottom-up approach.	Cloud technology was designed using top-down approach
5	Grid is difficult to be used , does not give performance guarantee.	Cloud is easy to use , scalable and always gives user what they want.
6	It is used by narrow community of scientist to solve specific problem,	It is used by all communities.

Performance related issues linked to the adoption of cloud In the High Performance Computing (HPC)

There are three important issues occurred related with cloud implementation in HPC. They are:
1. Difference performance evaluation between HPC and Cloud paradigm
2. A Comparison of cloud and HPC approach in terms of advantages and drawbacks.
3. New performance evaluation techniques and tools to support HPC in Cloud System.

Difference performance evaluation between HPC and Cloud paradigm:

The difference between typical HPC and cloud paradigm are given below.

Issues	HPC	CLOUD
Cost	Buy-and –Maintain Paradigm	Pay-per-use Paradigm
Performance Optimization	Tuning of the application to hardware	Joint tuning of application and hardware
System Dimensioning	Performance Under System Administrator Control	Performance under user control.

A Comparison of cloud and HPC approach in terms of advantages and drawbacks

The next issue is performance comparison between classical HPC system and the new Cloud paradigm. It is possible to point out the advantages and disadvantages of the two approaches and will enable us to understand how cloud can be useful for HPC.

Advantages

The advantages of HPC in the cloud are clear. It's scalable, on-demand, fast and inexpensive. Some organizations are concerned about security in the public cloud. And others may worry about the latency effect of moving large amounts of data.

Once the data is in the cloud, it gives HPC customers a unique advantage: it doesn't need to move anymore. Keeping it in a cloud environment makes collaboration easier. Scientific research teams often want to work with other teams in the same interest area, but they may be distributed around the world.

Drawbacks

HPC customers with one big challenge in a cloud computing environment: getting their data into the cloud in the first place. Scientific research applications often deal with data in the Terabyte range. Uploading it to someone else's server can be like pushing a data lake through a straw.

New performance evaluation techniques and tools to support HPC in Cloud System.

Very few performance measuring tools are provided by CSP or third parties. These tools are useful only to measure for the virtualized environment. Ex: CloudWatch offered by Amazon.

Cloud Security

Cloud Security refers to a broad set of policies, technologies, and controls deployed to protect data, applications, and the associated infrastructure of cloud computing.

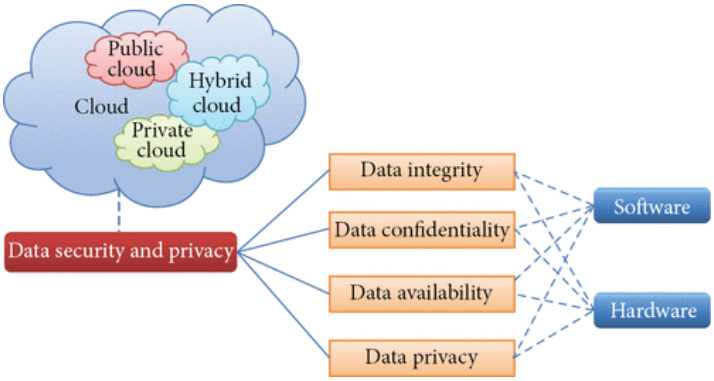
Current State of Data Security in the Cloud

- Cloud security has clearly emerged with both a technological and business case, but from a security perspective, it's still a bit in a state of flux.
- A key challenge that many information security professionals are struggling with is how to classify the cloud and define the appropriate type of controls to secure data entering the cloud.
- The cloud is inherently untrusted since it is not simply an extension of the organization, but it's an entirely separate environment that is out of the organization's control.
- In the cloud computing environment, it becomes particularly serious because the data is located in different places even in all the globe. Data security and privacy protection are the two main factors of user's concerns about the cloud technology.
- Data security and privacy protection issues are relevant to both hardware and software in the cloud architecture.
- This study is to review different security techniques and challenges from both software and hardware aspects for protecting data in the cloud and aims at enhancing the data security and privacy protection for the trustworthy cloud environment.

Homo Sapiens

It means cloud computing is seen as a revolutionary move forward in the use of technology to enhance the modern human communication.

Data security risk in cloud computing



Cloud computing model faces old and new data security risks. Data are uploaded into a cloud and stored in a data center and these data are accessed by users from these data center maintained by Cloud Service Providers such as Google, Amazon, Microsoft and so on. The most obvious risk is associated with the storage of that data. This action has several risks associated with it. There are

Risk 1:

A user uploading or creating cloud-based data that are stored and maintained by a third-party provider and the data may be hijacked on the way into the data base.

Action to be taken: It is necessary to protect the data during upload into the data center to ensure that the data do not get hijacked on the way into the database. It is necessary to the stored data in the data centers to ensure that they are encrypted all times.

Access those data need to be controlled and the control should be applied to
1. Hosting Company
2. Administrator of the Hosting Company.

Risk 2:

Access Control Risk- Access control becomes a much more fundamental issue in cloud –based system. Access control is a security technique that can be used to regulate who or what can view or use resources in a computing environment.

Action to be taken: Information-centric security is an approach to information security paradigm that emphasizes the security of the information itself rather than the security of networks, applications, or even simply data.

Risk 3:

Use of Content after Access: This risk is potentially higher in cloud network. A user printing of a sensitivity document within the office of a company after accessing data from the cloud.

Action to be taken: More privacy techniques to be applied to avoid this kind of risk and ensure that the accessed content should not be faced any risk

Risk 4:

The development of Web 2.0 technologies has created a dynamic method of communicating information through blogs, social networking sites, web conferencing, wikis. Data within these domains are more at risk.

Action to be taken: Cloud computing needs to ensure that protection is applied at the inception of the information, in a content centric manner, ensuring that a security policy becomes as integral part of that data throughout its life cycle.

Encryption is a vital component of the protection policy but further control over the access of that data and the use of that data must be designed.

Digital Identity

A **digital identity** is a number, code, record, object or collection of attributes that are used by information technologies to identify entities such as people, organizations, users and customers.

Example: This includes usernames and passwords, online search activities, birth date, social security

Content Level Security/ Content-Centric Security

Content Level Security is designed to control what content of file or document is permitted to a user through the cloud. It protects the contents as follows:

1. Certain section of the content is visible certain persons
2. Copy protection
3. Sharing of content to only authorized people
4. Only authorized people allowed to access the content.

Advantages:

- **Document Centric:** Complete control over your content down to the document.
- **Remote Wipe:** Eliminate Access to your content at anytime from anywhere
- **Content Tracking:** It enables content owners to understand when, how and for how long a user accessed the content.
- **Authorized Access:** Protect the content of document by assigning access to persons who hold a managed information card which contain certain claims.
- **Individual Content Control:** Content access can be based on an individual identity , individual content control and some users can be given stronger rights restrictions than others.
- **Protection from Administrator:** Content level security does not allow the cloud server administrator to access the content from the cloud server.
- **Simple to use:** Container security is done in a much simpler way of securing data. Content level security can be easily done for database and storage level using existing encryption techniques.

Data protected at the content level has other benefits such as Greater control, More focused on access control, Increased granular protection over content and Assurance within the cloud-hosted system.

Disadvantages:

The implementation of content level security becomes difficult for the clouds which use multi- center storage and replication of data.

Legal issues in cloud computing system

- One of the foremost and fundamental concerns faced by an organization while migrating to cloud services is with respect to the security and privacy of its data.
- However, despite the lack of clarity, most developed countries including EU, UK and the United States are at different stages of creating a legal framework for cloud-based services.
- The UK's Cloud Industry Forum has formulated a code of practice for Cloud service providers. Similarly, New Zealand has a Cloud Computing code of practice.
- In the US there is proposal to enact a Cloud Computing Act. In the EU, a Cloud Computing Information Assurance Framework has been proposed.

Cross border transfer of data:

The global nature of cloud architecture coupled with the diversity of legal mechanisms.

Privacy Shield lays down seven privacy principles which are worth mentioning and which should comprise the yardstick to which any cross border transfer of data should be subjected to:

- **Notice:** Information to an end user/ consumer that their data is being collected and how it will be used;
- **Choice:** Individual's right to opt out of collection and forward transfer of data to third parties;
- **Safety:** Safeguards to prevent loss of collected information;
- **Data Integrity and purpose limitation:** Data must be relevant and reliable for the purpose it was collected;
- **Access:** Individual's right to access information held about him and to correct or delete it, if inaccurate;
- **Enforcement & Liability:** Effective means to enforce these rules.

Encryption and data security

Encryption is one of the key tools employed by an organization to ensure security and privacy of its data in a cloud architecture where the data is frequently in transit and in cases of a multi- tenant environment- where data is stored on a physical hardware that is often shared with third parties.

Liability Issue: The liability assumed when entering into a contract in which either party to the contract fails to perform in accordance with the terms, otherwise known as a breach of contract.

Contract Law: Contract law is a body of law that governs, enforces, and interprets agreements related to an exchange of goods, services, properties, or money. The branch of civil law that deals with interpretation and enforcement of contracts between two or more parties.

Data Portability: Data portability is the ability to move data among different application programs, computing environments or cloud services. Data portability is growing more important as an increasing number of organizations store greater and greater quantities of data in the cloud.

Copyright: Copyright is a legal means of protecting an author's work. It is a type of intellectual property that provides exclusive publication, distribution, and usage rights for the author. This means whatever content the author created cannot be used or published by anyone else without the consent of the author.

Compliance: Contract Compliance is state of acting in conformance with the predefined and agreed rules or guidelines.

Contract

A contract is a legally enforceable agreement between two or more parties where each assumes a legal obligation that must be completed to provide a product or Service.

Types of contracts

Contracts can be classified as follows.

1. Licensing Agreements Versus Service Agreements
2. On-Line Agreements Versus Standard Contracts
3. Importance of Privacy Policies Terms and Conditions
4. Risk Allocation and Limitation of Liability

Licensing Agreements Versus Service Agreements

Licensing Agreement

- A software license agreement is the legal contract between the licensor and/or author and the purchaser of a piece of software which establishes the purchaser's rights.
- A software license agreement details how and when the software can be used, and provides any restrictions that are imposed on the software.
- A software license agreement also defines and protects the rights of the parties involved in a clear and concise manner.
- Most of software license agreements are in digital form and are not presented to the purchaser until the purchase is complete.

Service Agreement

- A service agreement is an agreement between two persons or businesses where one agrees to provide a specified service to the other.
- In cloud computing models, the access to the cloud-based technology is provided as a service contract.
- A service contract provides all the basic terms and conditions that provide adequate protection to the cloud user.
- There are two contracting models under which a cloud provider will grant access to its services- On-Line Agreement and Standard Contract.

On-line Agreement Versus Standard Contract

On-Line Agreement

- It is a click wrap agreement with which a cloud user will be presented before initially accessing the service.
- It is a non-negotiable agreement.
- Example: When user enters into when he/she checks an “I Agree” box.

Standard Contract

- It is a negotiated, signature-based contract will have its place and time with all terms and condition.

Privacy Policy and Terms and Condition

- A privacy policy is a statement or a legal document (in privacy law) that discloses some or all of the ways a party gathers, uses, discloses, and manages a customer or client's data. It fulfills a legal requirement to protect a customer or client's privacy.
- It’s required by law if you collect personal information from users.
- It’s required by third-party services you may use.
- Users are interested in their privacy.
- It’s ubiquitous.

Risk Allocation and Limitation of Liability:

Risk Allocation

- Allocation of risk in commercial contracts represents a key negotiation point. Each party to a commercial contract seeks to minimize its risk and maximize its reward.
- Allocation of risk is central to all commercial contract negotiations.
- Each party to a commercial contract seeks to minimize its risk and maximize its reward, which creates an inherent tension between contracting parties.
- Parties can manage risk by carefully negotiating and drafting many common contractual provisions.
- Examples: Representations and warranties, Indemnification, Limitation of liability, Express contractual remedies, Payment terms.

Limitation of Liability

- A limitation of liability clause stipulates that a party will be obligated to pay to the other in such an event under the terms of an agreement.
- This clause limits the amount as well as the types of damages a party can recover from the other.

ANEKA TOOL

ANEKA Cloud Platform:

- ANEKA is a software platform and a framework for developing distribution application on the cloud. It simplifies the computing resources of a heterogeneous network of workstations and servers or data centers on demand.
- Aneka provides developers with a rich set of API's for using the resources by expressing application logic with a variety of programming abstraction. It is a workflow management tool.
- Aneka is a distributed middleware for deploying platform as a service. Aneka, which is both a development and runtime environment, is available for public use for a cost.
- It can be installed on corporate networks or dedicated clusters or it can be hosted on Infrastructure clouds like Amazon Ec2 to support work flow management.
- Aneka was developed on Microsoft.Net framework 2.0 and is compatible with other implementation. Aneka can run on popular platforms such as Microsoft windows, Linux and Macosx.
- The runtime environment consists of a Aneka containers running on physical or virtualized models. Each of these containers can be configured to play a specific role such as scheduling or execution.
- The Aneka service stack provides services for infrastructure management, application execution management, accounting, licensing and security.

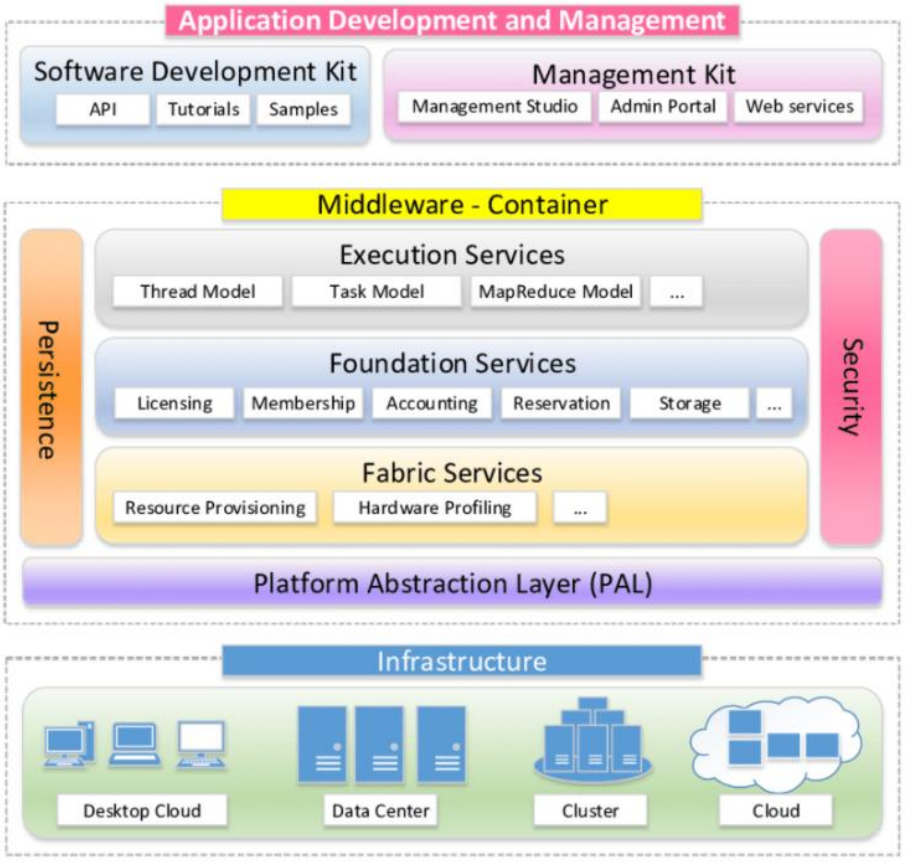
Dynamic resource Provisions:

This service enables horizontal scaling depending on the overall load in the cloud. The platform is elastic in nature and can provision additional resources on-demand from external physically virtualized resource pools, in order to meet the QOS requirements of applications.

Development Environment:

This development environment provides a rich set of API's for developing applications that can utilize free resources of the infrastructure. These APIs expose different programming abstractions, such as the:

1.Task Model
2.Thread Model
3.Map Reduce



Storage Service:

The storage service provides a temporary repository for application files such as input files that are required for task execution and output files that are the result of execution. Any output files produced as a result of the execution are uploaded back to the storage servers. From here they are staged-out to the remote storage location.

Execution Service:

They are responsible for scheduling and executing applications. Each of the programming models supported by Aneka defines specialized implementations of these services for managing the execution of a unit of work defined in the model.

Foundation Service:

These are the core management service of the Aneka container. They are in charge of metering application, allocating resources for execution, managing the collection of available nodes and keeping the services registry updated.

Fabric Services:

They provide access to the resources managed by the cloud. An important service in this layer is Resource Provisioning Service which enables horizontal scaling in the cloud. Resource Provisioning makes Aneka elastic and allows it to grow or to shrink dynamically to meet the Quality-of-Service requirements of Applications.

COMET CLOUD

Comet Cloud is an autonomic computing engine for cloud and grid environments. It supports highly heterogeneous and dynamic cloud and grid infrastructure, integration of public and private clouds and autonomic cloud bursts.

Autonomic Behavior of Comet Cloud

Cloud Bursting: Cloud bursting is an application deployment model in which an application runs in a private cloud or data center and bursts into a public cloud when the demand for computing capacity spikes.

Autonomic Cloud Bridging: Autonomic Cloud Bridging is to connect Comet Cloud and a virtual cloud which consists of Public Cloud, Data Centre and Grid by the dynamic needs of resources for the application.

Autonomic Cloud Bursting: The main goal of autonomic cloud bursting is to seamlessly and securely integrate private enterprises clouds and data centers with public utility clouds on-demand. It helps dynamic deployment of applications components onto to a public cloud to support dynamic workloads.

Comet Cloud supports autonomic cloud bursting and cloud bridging for the real-world cloud application. It also integrates all local computational environment and public cloud services dynamically and provide to manage cloud resources on demand.

Key Features of Autonomic Cloud Bursting

- **Load Dynamic:** The computational environment dynamically grows or shrink in response to dynamic application workload.
- **Accuracy:** The computational environment on comet cloud dynamically adapt to satisfy the accuracy requirements.
- **Collaboration of different groups:** Different groups can run the same application with different dataset policies, as collaboration groups join or leave the work, the computational environment must grow or shrink to satisfy their SLA.

Economics:

Application tasks can have varied heterogeneous and dynamic priorities and must be assigned resources and scheduled accordingly based on budget and economic model.

Failures:

The computation must be able to manage failures without impacting application quality of service.