What is Cloud?

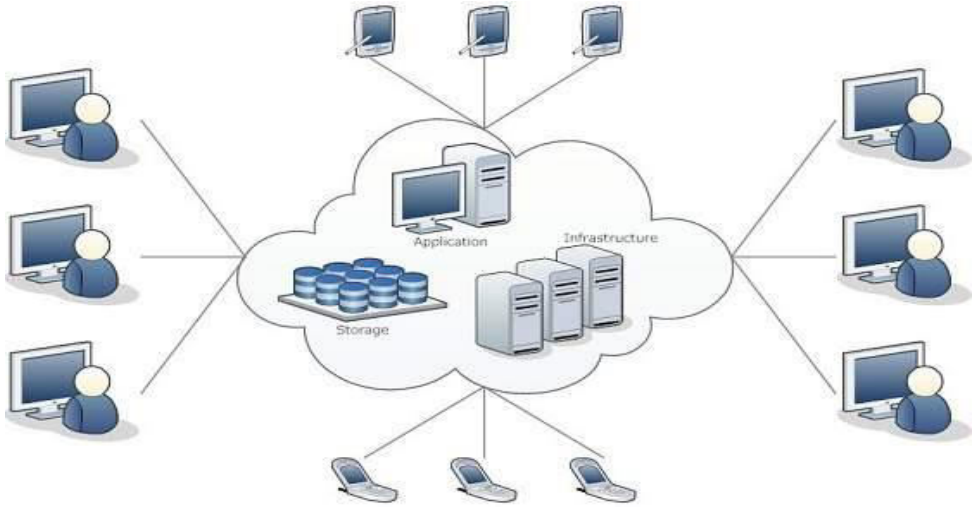
The term **Cloud** refers to a **Network** or **Internet.**

In other words, we can say that Cloud is something, which is present at remote location. Cloud can provide services over public and private networks, i.e., WAN, LAN or VPN.

Applications such as e-mail, web conferencing, customer relationship management (CRM) execute on cloud.

What is Cloud Computing?

Cloud Computing refers to **manipulating, configuring,** and **accessing** the hardware and software resources remotely. It offers online data storage, infrastructure, and application.



Cloud computing offers **platform independency,** as the software is not required to be installed locally on the PC. Hence, the Cloud Computing is making our business applications **mobile** and **collaborative.**

Cloud Computing provides us resources of accessing the applications as utilities over the Internet. It allows us to create, configure, and customize the applications online.

Basic Concepts

There are certain services and models working behind the scene making the cloud computing feasible and accessible to end users.

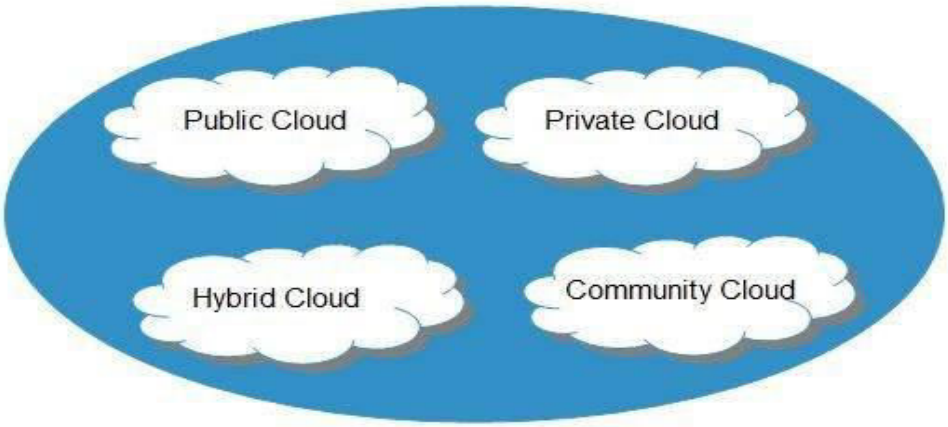
Following are the working models for cloud computing:

- Deployment Models
- Service Models

Deployment Models:

Deployment models define the type of access to the cloud, i.e., how the cloud is located?

Cloud can have any of the four types of access: Public, Private, Hybrid, and Community.



Public Cloud:

The **public cloud** allows systems and services to be easily accessible to the general public. Public cloud may be less secure because of its openness.

Private Cloud:

The **private cloud** allows systems and services to be accessible within an organization. It is more secured because of its private nature.

Community Cloud:

The **community cloud** allows systems and services to be accessible by a group of organizations.

Hybrid Cloud:

The **hybrid cloud** is a mixture of public and private cloud, in which the critical activities are performed using private cloud while the non-critical activities are performed using public cloud.

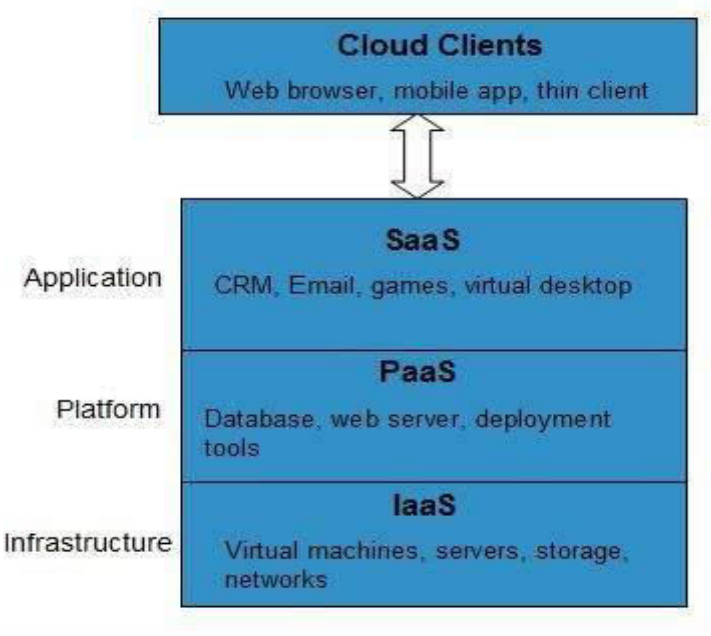
Service Models:

Cloud computing is based on service models. These are categorized into three basic service models:-

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

Anything-as-a-Service (XaaS) is yet another service model, which includes Network-as-a-Service, Business-as-a-Service, Identity-as-a-Service, Database-as-a-Service or Strategy-as-a-Service.

The **Infrastructure-as-a-Service (IaaS)** is the most basic level of service. Each of the service models inherit the security and management mechanism from the underlying model, as shown in the following diagram:



Infrastructure-as-a-Service (IaaS):

IaaS provides access to fundamental resources such as physical machines, virtual machines, virtual storage, etc.

Platform-as-a-Service (PaaS):

PaaS provides the runtime environment for applications, development and deployment tools, etc.

Software-as-a-Service (SaaS):

SaaS model allows to use software applications as a service to end-users.

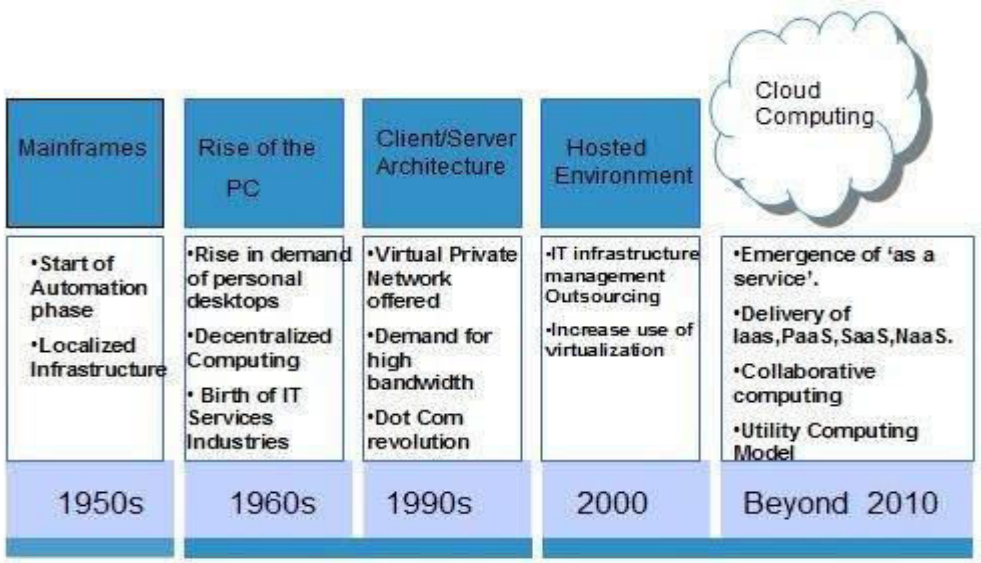
History of Cloud Computing

The concept of **Cloud Computing** came into existence in the year 1950 with implementation of mainframe computers, accessible via **thin/static clients.**

Since then, cloud computing has been evolved from static clients to dynamic ones and from software to services.

[A **thin client** is a computer that runs from resources stored on a central server instead of a localized hard drive. **Thin clients** work by connecting remotely to a server-based computing environment where most applications, sensitive data, and memory, are stored.]

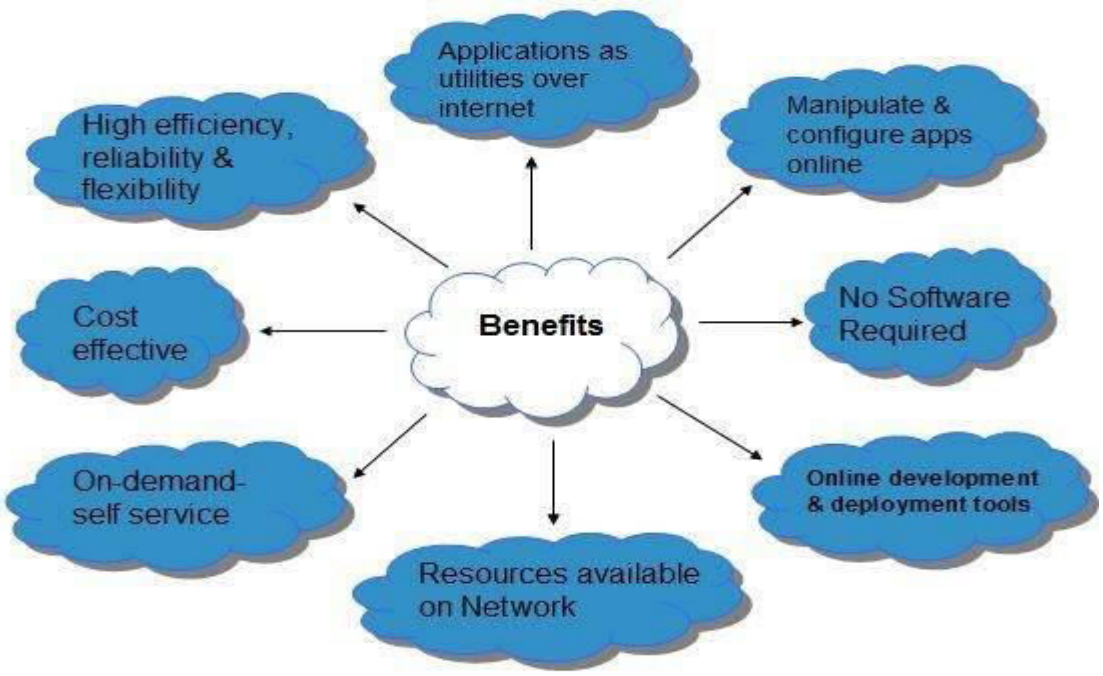
The following diagram explains the evolution of cloud computing:



Benefits

Cloud Computing has numerous advantages. Some of them are listed below -

- One can access applications as utilities, over the Internet.
- One can manipulate and configure the applications online at any time.
- It does not require to install a software to access or manipulate cloud application.
- Cloud Computing offers online development and deployment tools, programming runtime environment through **PaaS model.**
- Cloud resources are available over the network in a manner that provide platform independent access to any type of clients.
- Cloud Computing offers **on-demand self-service.** The resources can be used without interaction with cloud service provider.
- Cloud Computing is highly cost effective because it operates at high efficiency with optimum utilization. It just requires an Internet connection
- Cloud Computing offers load balancing that makes it more reliable.



Risks related to Cloud Computing:

Although cloud Computing is a promising innovation with various benefits in the world of computing, it comes with risks as below:

Security and Privacy:

It is the biggest concern about cloud computing. Since data management and infrastructure management in cloud is provided by third-party, it is always a risk to handover the sensitive information to cloud service providers.

Although the cloud computing vendors ensure highly secured password protected accounts, any sign of security breach may result in loss of customers and businesses.

Lock In:

It is very difficult for the customers to switch from one **Cloud Service Provider (CSP)** to another. It results in dependency on a particular CSP for service.

Isolation Failure:

This risk involves the failure of isolation mechanism that separates storage, memory, and routing between the different occupiers.

Management Interface Compromise:

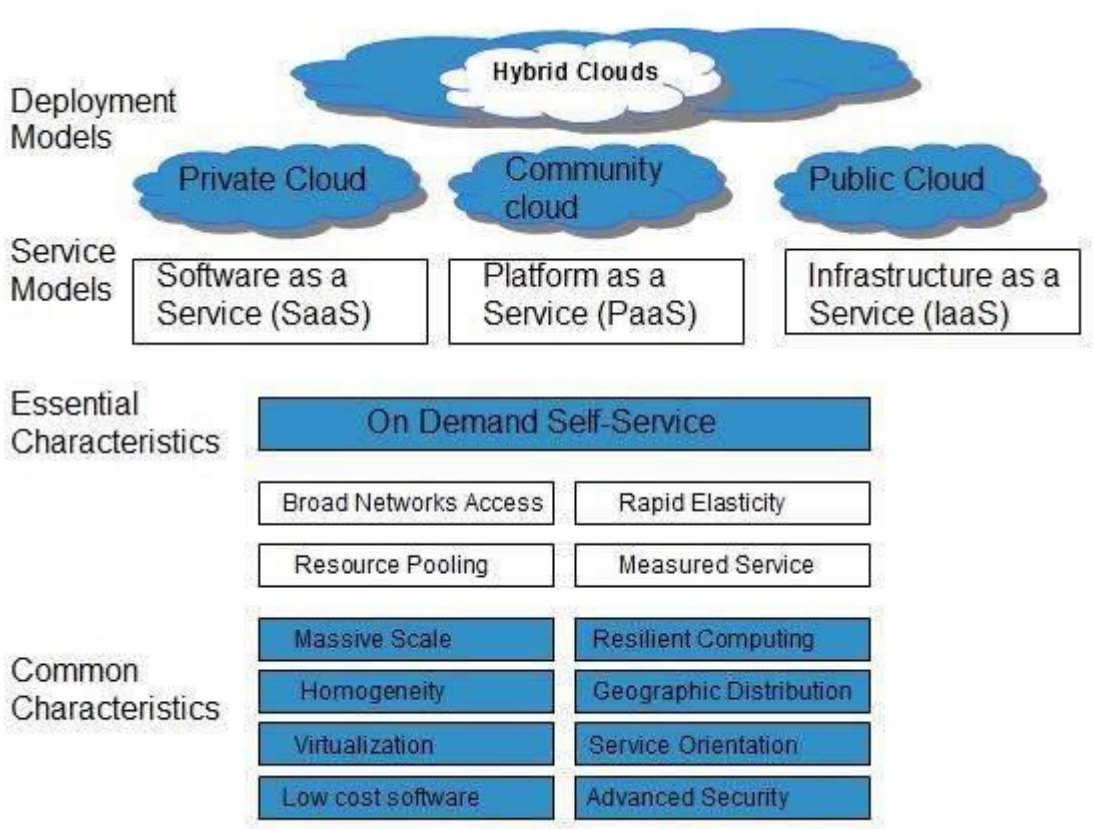
In case of public cloud provider, the customer management interfaces are accessible through the Internet.

Insecure or Incomplete Data Deletion:

It is possible that the data requested for deletion may not get deleted. It happens because either of the following reasons

- Extra copies of data are stored but are not available at the time of deletion
- Disk that stores data of multiple occupiers is destroyed.

Characteristics of Cloud Computing:



Following are the essential characteristics of Cloud Computing:

On Demand Self Service:

Cloud Computing allows the users to use web services and resources on demand. One can logon to a website at any time and use them.

The Cloud computing services does not require any human administrators, user themselves are able to provision, monitor and manage computing resources as needed.

Broad Network Access:

Since cloud computing is completely web based, it can be accessed from anywhere and at any time.

The Computing services are generally provided over standard networks and heterogeneous devices.

Resource Pooling:

Cloud computing allows multiple residents to share a pool of resources. One can share single physical instance of hardware, database and basic infrastructure.

The IT resource (e.g., networks, servers, storage, applications, and services) present are shared across multiple applications and occupant in an uncommitted manner. Multiple clients are provided service from a same physical resource.

Rapid Elasticity:

It is very easy to scale the resources vertically or horizontally at any time. Scaling of resources means the ability of resources to deal with increasing or decreasing demand.

The resources being used by customers at any given point of time are automatically monitored.

The Computing services should have IT resources that are able to scale out and in quickly and on as needed basis.

Whenever the user require services it is provided to him and it is scale out as soon as its requirement gets over

Measured Service:

In this service cloud provider controls and monitors all the aspects of cloud service. Resource optimization, billing, and capacity planning etc. depend on it.

The resource utilization is tracked for each application and occupant, it will provide both the user and the resource provider with an account of what has been used. This is done for various reasons like monitoring billing and effective use of resource

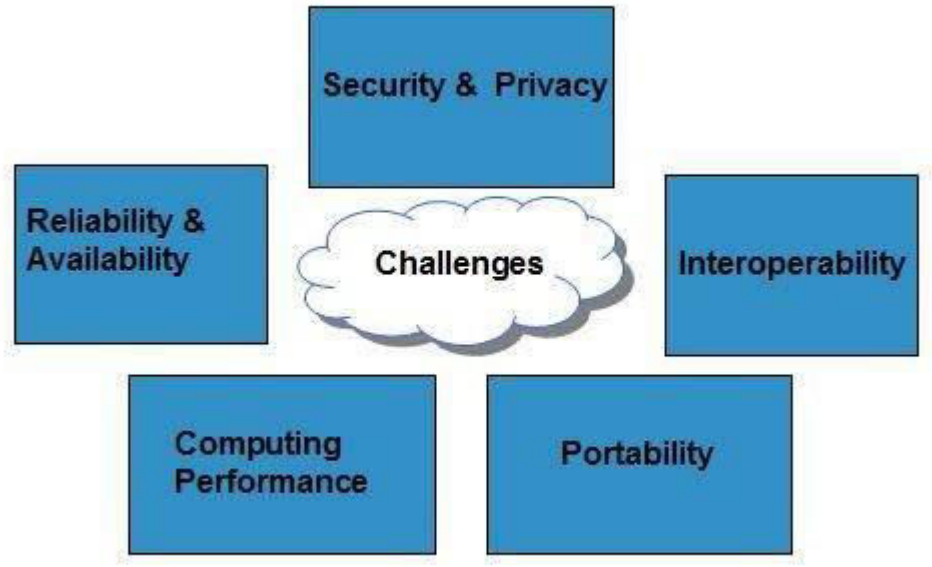
Cloud Vulnerability

> **Vulnerability** is a cyber-security term that refers to a flaw in a system that can leave it open to attack. A **vulnerability** may also refer to any type of weakness in a computer system itself, in a set of procedures, or in anything that leaves information security exposed to a threat.

Denial of Service (DoS) and Distributed Denial of Service (DDoS) attacks are **cloud** security **vulnerabilities** that make servers inaccessible for users by flooding the network's traffic. Data Loss.

Cloud Computing Challenges

Cloud computing, an emergent technology, has placed many challenges in different aspects of data and information handling. Some of these are shown in the following diagram:



Security and Privacy:

Security and Privacy of information is the biggest challenge to cloud computing. Security and privacy issues can be overcome by employing encryption, security hardware and security applications.

Portability:

This is another challenge to cloud computing that applications should easily be migrated from one cloud provider to another. There must not be vendor lock-in. However, it is not yet made possible because each of the cloud provider uses different standard languages for their platforms.

Interoperability:

It means the application on one platform should be able to incorporate services from the other platforms. It is made possible via web services, but developing such web services is very complex.

Computing Performance:

Data intensive applications on cloud requires high network bandwidth, which results in high cost. Low bandwidth does not meet the desired computing performance of cloud application.

Reliability and Availability:

It is necessary for cloud systems to be reliable and robust because most of the businesses are now becoming dependent on services provided by third-party.

Cloud Migration

Cloud migration is the process of moving data, applications or other business elements to a cloud computing environment. ... However, a cloud migration could also involve moving data and applications from one cloud platform or provider to another -- a model known as cloud-to-cloud migration.

There are various types of cloud migrations an enterprise can perform. One common model is the transfer of data and applications from a local, on-premises data center to the public cloud.

The most important part of any cloud migration is making sure the migration gets your company where it needs to be. The four types of cloud migration are called lift and shift, shift to Software-as-a-Service (SaaS), application refactoring, and replatforming.

Lift and Shift:

The lift and shift approach is for organizations looking to get out of the data center and stop managing hardware.

Lift and shift provides the same software that your company used in the data center, but now in the cloud. There isn't any learning curve for the cloud applications, since they work exactly the same as before. This is the fastest method for migrating applications to the cloud, and the one that causes the least disruption. It only requires the involvement of the infrastructure and security teams, leaving everyone else free to pursue their work uninterrupted.

It's also the option with the least upfront cost. Moving the application to the cloud allows it to handle peak performance, without your company having to pay for it.

Lift and shift comes with its drawbacks, however. This cloud migration can't take full advantage of the speed and versatility the cloud can provide.

Since the process doesn't change the application – it just moves the code to a new location – the shifted version of the application doesn't usually have better performance than the original.

It's also unlikely to lead to long-term savings. This model is best suited to companies with a regular peak schedule and slow, predictable changes in the market.

Food delivery companies with their regular peaks on Monday, Saturday, and Sunday are a good example of the former.

Tax companies are a good example of the latter, since tax rules change on a yearly basis and new rules are released a year in advance, giving the companies plenty of time to prepare.

Lift and shift cloud migration involves copying, bit by bit, the workloads, tasks, and applications housed within an organization's virtual machines, and storing the replicated version in a cloud-based location. Because the data is simply a replication, no code is modified and no costly, time-consuming redesign is required during migration.

The method is certainly less complex than other migration techniques, IT teams and their cloud provider will want to be sure data sets are properly matched with handling systems in the new environment. Additionally, they need to ensure that all applications have the resources needed to operate effectively and efficiently.

Shift to SaaS: Save Time and Trouble:

Companies that want to stop allocating time and resources to applications outside their core business should think about shifting to SaaS.

Shifting to SaaS means outsourcing one or more applications to a cloud services company that specializes in managing those applications. Companies do this on an application by application basis and only shift the applications they need to.Static applications can remain on-premises.

Shifting to SaaS frees employees up to focus on core competencies and the things that make a business unique and competitive.

It's extremely important when shifting an application to SaaS that you pick the right service.

The main drawback of shifting to SaaS, while you can personalize it, is that customizing it can lead to problems.

Shifting to SaaS should only be used for routine functions – not for anything that needs to be unique.

Email is a good example of a routine business function that can be shifted to SaaS. For example, we worked with a manufacturing company that built engine components. The company was sick of managing its own email. While it needed to have email for all its employees, the customers didn't care what the email service looked like – so long as it worked. We helped them find a hosting service that freed the company of needing to worry about their email so it could focus on manufacturing.

Application Refactoring: App Modernization:

App modernization is a preferred approach for organizations that have specific applications which could benefit from the cloud.

With refactoring, organizations can copy their inheritance applications whole and unbroken onto a cloud platform.

What makes it low risk is that inheritance applications can run in parallel while new applications are constructed, with the immediate benefits of quickness and speed to market. This approach focuses on the applications that benefit the most from a cloud platform.

Refactoring is about prioritization; it provides lots of opportunities to save over time by minimizing spending on things you won't need once you're in the cloud.

You can save money on the platform itself as well by switching to cloud native services that cost less than the ones you use on-premises.

Refactoring is not just about cutting costs. It also allows you to make changes to your enterprise very quickly, which means that you can keep up with your customers.

Refactoring lets you respond faster and prioritize updates. One big box store we worked with started out with its applications so hard coded that it took months to do simple things like changing the font or background color. Refactoring got them moving.

Refactoring usually requires outside help. We worked with a company that wanted to refactor its entire application suite into modern technology. The project would have taken four years for the company working on their own, but we were able to cut that time down to eighteen months. We provided the company making the migration with its unique model for retraining.

When introducing the new technology, we rotate through everyone who needs training. Slowly, over time, we dial back how many consultants they have helping until the local IT people are able to handle everything on their own. This means that the IT staff learns by doing, so there is no awkward learning curve. At every point in the process, there is someone trained to handle difficulties on-site.

Replatforming: Develop Applications in the Cloud:

Replatforming is for companies looking to hold benefits of the cloud, enterprise-wide.

These companies want their core capabilities to be scalable, flexible, robust, redundant, and available.

This is the hardest option to implement, requires the most planning for the future, and comes with the most upfront cost, but it's the only option that lets you utilize the full strength and flexibility of the cloud.

Replatforming is replacing the application at the code level to make it cloud native. This is a complete reimagination of the application and usually requires a complete rewrite.

When considering replatforming, think about how fast your company can change. Then think about how fast it needs to change to keep up with customers and the market.

By making applications truly cloud native, they can be updated and those updates pushed out at the speed of the cloud. This boosts the speed to change across the board for all aspects of the business.

Replatformed applications can also be designed to be more modular and thus easier to maintain.

Refactoring the application can save development time, since modules of code from your first refactored application can be used to extend the capabilities of new applications.

Unlike refactored applications, refactored platforms can work across multiple cloud providers. This makes it easier to port from one mobile platform to another, which positions replatforming as the ideal strategy for those looking to develop mobile applications.

We worked with a healthcare company building a net-new mobile application for patient management, which included sending push notifications on appointments and subscriptions. The company wanted to build this application as a cloud-native company. However, they already had internal patient managing applications for the nurses and doctors that complemented the new applications. The company used replatforming to get their inheritance patient management application to the cloud, so the application would be ready and waiting while it developed its net-new mobile application.

Determine your goals before you select your cloud migration model:

Each of the four primary methods we identified comes with its share of advantages, but also with its disadvantages. Finding the method that matches your organizations goals and needs is the first step to a successful migration.

Cloud Service Provider

A cloud provider is a company that delivers cloud computing based services and solutions to businesses and/or individuals. This service organization may provide rented and provider-managed virtual hardware, software, infrastructure and other related services.

Amazon Web Services:

AWS is Amazon's cloud web hosting platform which offers fast, flexible, reliable and cost-effective solutions. It offers a service in the form of building block which can be used to create and deploy any kind of application in the cloud. It is the most popular as it was the first to enter the cloud computing space.

Cloudways:

Cloudways provides managed cloud hosting to agencies, stores, etc.

The platform has partnered with top cloud providers including AWS, Google Cloud, DigitalOcean,etc.

Experience the freedom to build, deploy and manage applications including PHP, WordPress, etc without requiring any knowledge of cloud server management.

Cloudways users can focus on business growth without worrying about the technical complexities of server management, security, and maintenance.

DigitalOcean:

Digitalocean's droplet is a scalable computer service. It is more than just virtual machines.

This cloud platform offers add-on storage, security, and monitoring capabilities to run production applications easily.

Rackspace"

Rackspace is another useful cloud computer service tool. It offers services like hosting web applications, cloud files, cloud backup, database, and cloud server, etc.

Alibaba Cloud:

Alibaba is the largest Chinese cloud computing company. It is a new platform which created a global footprint with over 1500 Nodes worldwide of 19 regions and 56 availability zones across more than 200 countries.

[A content delivery network (**CDN**) is a system of distributed servers (network) that deliver pages and other web content to a user, based on the geographic **locations** of the user, the origin of the webpage and the content delivery server. ... CDNs also provide protection from large flows in traffic.]

Microsoft Azure:

Azure is a cloud computing platform which is launched by Microsoft in February 2010. This open source and flexible cloud platform which helps in development, data storage, service management & hosting solutions.

Google Cloud Platform:

Google Cloud is a set of solution and products which includes GCP & G suite. It helps you to solve all kind of business challenges with ease.

Oracle Cloud:

Oracle Cloud offers innovative and integrated cloud services. It helps you to build, deploy, and manage workloads in the cloud or on premises. Oracle Cloud also helps companies to transform their business and reduce complexity.

IBM Cloud:

IBM cloud is a full stack cloud platform which spans public, private and hybrid environments. It is built with a robust suite of advanced and AI tools.

Roles and Responsibilities:

There is a important growth of cloud adoption across small as well as large enterprises. This has resulted in a large spectrum of cloud offerings including cloud delivery models and a variety of cloud computing services that are being provided by cloud hosting companies.

Improved accessibility and security:

Cloud adoption not only helps improve business processes and enhances the efficiency of IT infrastructures but also brings down costs of running, upgrading, and maintaining on-site IT facilities.

Your business-critical data is armed with added security in the cloud environment. In reality, the data is not actually being placed up in the cloud but is distributed to a number of remote data center facilities that are owned and operated by third-party service providers. These establishments consist of climate-controlled rooms to house enterprise-grade servers for seamless protection and easy accessibility for maintaining business continuity in spite of any tragic event that may impact the main office of your enterprise.

The cloud data centers are designed to house a multitude of servers for storing data under strict security controls. The arrangement is aimed at enabling uninterrupted connectivity among vast networks comprising of millions of machines. Cloud computing is controled by end users as well as cloud hosting companies for the enhancement of their services.

Understanding the cloud's role in businesses:

In order to understand the precise reasons for increased cloud adoption in enterprise setups, we should have in-depth knowledge about of cloud's attributes that boost business processes.

Cloud services are designed to set your IT staff free from boring(ordinary) and time-consuming tasks of maintaining, repairing, and upgrading hardware equipment such as servers. On-site IT infrastructure in enterprises will be thinner after moving workloads to cloud data center. In the majority of cases, there will be no need to allocate separate space for housing servers and other IT equipment.

The direct benefit of cloud computing is associated with reduced capital expenditure as companies need not invest funds in purchasing costly hardware equipment. improvement of hardware costs is also backed by freedom from maintenance and repair costs of web servers. There is a definite reduction in upfront costs of ownership of cost-intensive software as well as hardware.

Performance with a promise of security:

In comparison with a physical server, a cloud hosting delivers better performance. This is because established web hosting service providers are in a better position to afford enterprise-grade cloud servers as against small or medium-sized enterprises.

Cloud hosting providers attach great importance to the security of customers' digital assets by spending a significant amount of financial and manpower resources. These providers harden the defences by the implementation of strict measures such as firewalls, anti-malware and anti-virus deployments. In addition to this, the host data centers are armed with barrier-like security for safeguarding physical as well as networking assets.

Greater affordability:

By provisioning top of the line hardware and software resources to customers at affordable prices, cloud hosting service providers help business enterprises reduce their capital as well as operating costs without impacting performance.

Cloud services go all out by investing huge sums of money to offer world-class resources to customers at economical prices. Their efficient staffs are well equipped to look after the routine tasks as well as technical problems irrespective of the time of the day for all weekdays.

Demand-oriented resource provisioning:

Users of cloud services are allowed to access the optimum amount of resources in response to resource requirements. This not only assures guaranteed resource availability but also helps businesses achieve resource optimization for reduction of operating costs.

Cloud-based infrastructure also enables users to access a variety of resources such as applications or platforms via any internet enabled device, from any location. These services are always available on round the clock basis for improved efficiency of enterprises. Employees can use a number of devices including smart-phones, tablets, and laptops to get their hands on a huge number of files and folders without the need to make a trip to the office. Cloud-based solutions are inherently flexible and accessible and businesses can easily keep their employees well-connected with each other for greater efficiency.

Freedom from maintenance:

On-site IT infrastructures are resource intensive and need to be regularly upgraded and maintained. In contrast, cloud service providers shoulder the entire responsibility of looking after the performance of servers, bandwidth, network, and software applications. This also includes periodic upgrades and security patching of operating systems and other business-critical applications.

This kind of infrastructure management requires large teams of software professionals to be available for 24 hours a day for 365 days in a year. Majority of companies that adopt cloud are

driven by the need to have consistently available, flexible, secure, and well managed IT infrastructure in the absence of any on-premise facility.

These are some of the valuable benefits of cloud computing that signify the role of cloud service providers. Therefore, the future of faultless data management is secure in the hands of established cloud service providers.

Cloud Service Consumer

A cloud consumer represents a person or organization that maintains a business relationship with, and uses the service from a cloud provider.

A cloud consumer browses the service catalog from a cloud provider, requests the appropriate service, sets up service contracts with the cloud provider, and uses the service.

The main expectation of cloud service consumer is to have a reliable service.

To satisfy consumer's expectation several Data centres are established all over the world and each Data centre contain thousands of servers.

The idle servers and resources in data center wastes huge amount of energy.

An appealing customer experience is essential for consumer business. If you can't manage expectations for consumer fast enough, they too may go elsewhere.

1. Flexibility:

Change is the only constant in the digital economy. A hybrid cloud model gives you the flexibility to adjust and grow.

This flexibility allows you to go after the big opportunities knowing that the location of your company's data won't be an obstacle.

The value is in more than just cost and speed, though these remain key principle for the public cloud.

The hybrid model also enables you to realize new insights across your entire ecosystem and quickly move priorities and resources to meet opportunities.

2. Freedom to choose:

Never forget you have options. You need to be able to easily change where an application runs based on your business needs. Which cloud helps you realize the most value: public, private?

What if that changes? If you're locked in with a public vendor, for example, moving data without disturbance can be very costly and time-consuming.

Look for solutions that give you the freedom to choose, with easy application portability regardless of your architectural environment across any cloud.

Solutions like IBM WebSphere Application Server Version 9 are built to put clients in control, not cloud providers.

3. Cognitive insights:

We're living in the cognitive era, inseparable from cloud innovation.

Cognitive is the way to outthink the competition and make sense of information. What can your data do? Bring new customer experiences, new applications and even new business models, for starters.

IBM Cloud offers a host of accessible cognitive capabilities which you can build into your applications.

You can rapidly infuse apps with cognitive capabilities to gain operational insights and dazzle your customers. Use these cognitive capabilities to breathe new life into your existing investments and extend their value while still putting the customer first.

Service level agreements in Cloud computing

A **Service Level Agreement (SLA)** is the bond for performance negotiated between the cloud services provider and the client.

Earlier, in cloud computing all Service Level Agreements were negotiated between a client and the service consumer. Nowadays, with the initiation of large utility-like cloud computing providers, most Service Level Agreements are standardized until a client becomes a large consumer of cloud services.

Service level agreements are also defined at **different levels** which are mentioned below:

- Customer-based SLA
- Service-based SLA
- Multilevel SLA

Few Service Level Agreements are enforceable as contracts, but mostly are agreements or contracts which are more along the lines of an Operating Level Agreement (OLA) and may not have the restriction of law.

It is fine to have an advocate review the documents before making a major agreement to the cloud service provider.

Service Level Agreements usually specify **some parameters** which are mentioned below:

1. Availability of the Service
2. Latency or the response time
3. Service components reliability
4. Each party accountability
5. Warranties

In any case, if a cloud service provider fails to meet the stated targets of minimums then the provider has to pay the penalty to the cloud service consumer as per the agreement. So, Service Level Agreements are like insurance policies in which the corporation has to pay as per the agreements if any casualty occurs.

Microsoft publishes the Service Level Agreements linked with the Windows Azure Platform components, which is demonstrative of industry practice for cloud service vendors.

Each individual component has its own Service Level Agreements.

Below are two **major Service Level Agreements (SLA)** described:

1. **Windows Azure SLA –**

   Window Azure has different SLA's for compute and storage. For compute, there is a guarantee that when a client deploys two or more role instances in separate fault and upgrade domains, client's internet facing roles will have external connectivity minimum 99.95% of the time. Moreover, all of the role instances of the client are monitored and there is guarantee of detection 99.9% of the time when a role instance's process is not runs and initiates properly.
2. **SQL Azure SLA –**

   SQL Azure clients will have connectivity between the database and internet gateway of SQL Azure. SQL Azure will handle a "Monthly Availability" of 99.9% within a month. Monthly Availability Proportion for a particular tenant database is the ratio of the time the database was available to customers to the total time in a month. Time is measured in some intervals of minutes in a 30-day monthly cycle. Availability is always remunerated for a complete month. A portion of time is marked as unavailable if the customer's attempts to connect to a database are denied by the SQL Azure gateway.

Service Level Agreements are based on the usage model. Frequently, cloud providers charge their pay-as-per-use resources at a premium and deploy standards Service Level Agreements only for that purpose.

Clients can also subscribe at different levels that guarantees access to a particular amount of purchased resources.

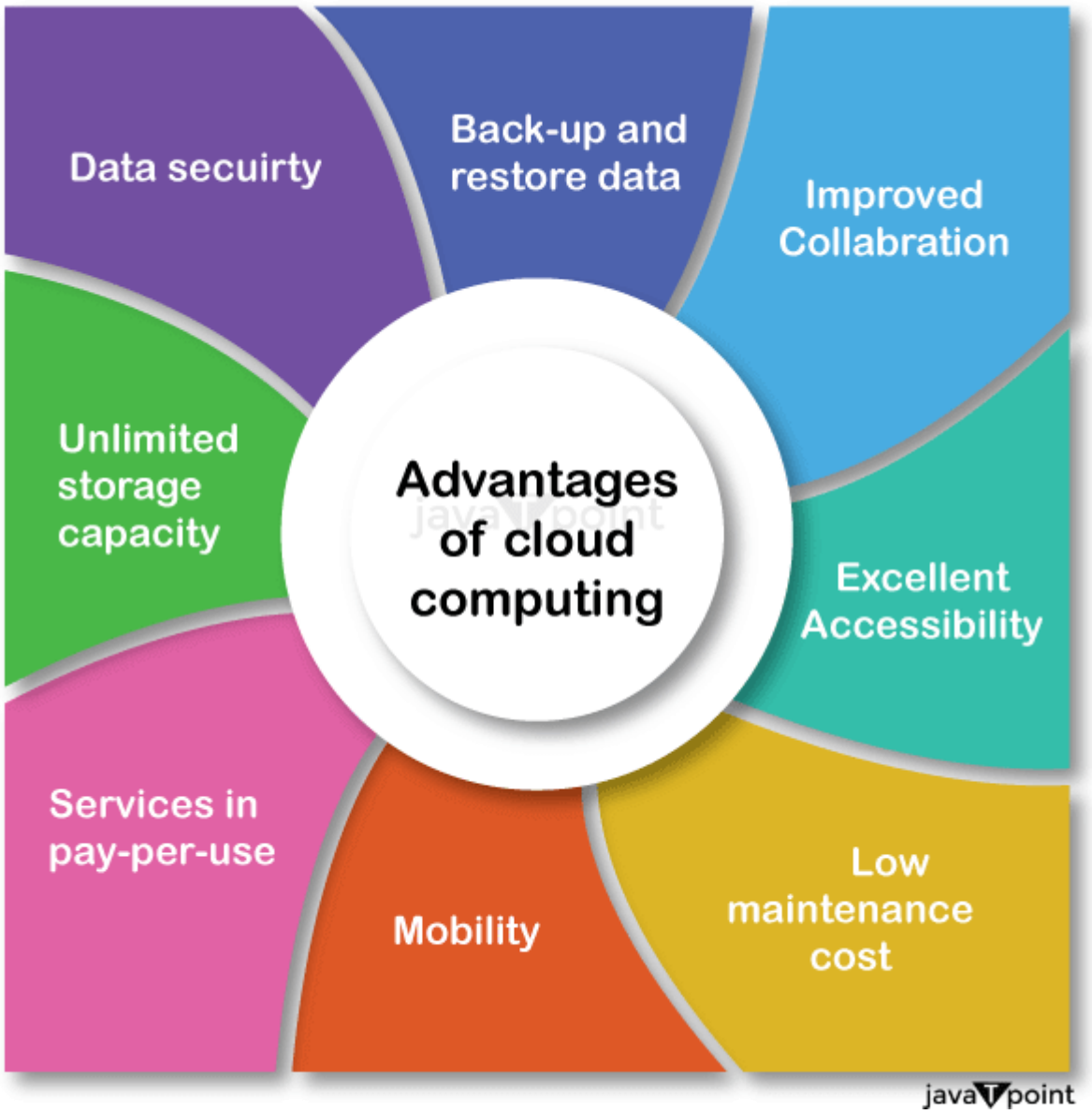
The Service Level Agreements (SLAs) attached to a subscription many times offer various terms and conditions. If client requires access to a particular level of resources, then the client need to subscribe to a service.

Advantages and Disadvantages of Cloud Computing

Here, we are going to discuss some important advantages of Cloud Computing-

Advantages:

Business operations are being transformed by the transformational technology known as cloud computing. With its extensive advantages and possibilities, cloud computing has emerged as a crucial strategic tool for businesses in a range of sectors. Businesses can take advantage of various benefits that promote development, innovation, and operational effectiveness by leveraging the power of the cloud.



Let's explore the advantages that cloud computing has to offer:

Data Backup and Restoration:

Cloud computing offers a quick and easy method for data backup and restoration. Businesses may simply access and restore their data in the event of any data loss or system failure by keeping it in the cloud.

Improved Collaboration:

Collaboration is improved because cloud technologies make it possible for teams to share information easily. Multiple users may work together on documents, projects, and data thanks to shared storage in the cloud, enhancing productivity and teamwork.

Excellent Accessibility:

Access to information stored in the cloud is made possible. Users can access their data from anywhere in the world with an internet connection, making remote work, flexibility, and effective operations possible.

Cost-effective Maintenance:

Organizations using cloud computing can save money on both hardware and software upkeep. Because cloud service providers manage the maintenance and updates, businesses no longer need to make costly infrastructure investments or set aside resources for continuous maintenance.

Upkeep and Updates:

Cloud service providers take care of infrastructure upkeep, security patches, and updates, freeing organizations from having to handle these duties themselves.

This frees up IT teams' time and resources to work on higher-value projects like application development, data analysis, or strategic initiatives rather than wasting them on rote upkeep and updates.

Mobility:

Cloud computing makes it simple for mobile devices to access data. Utilizing smartphones and tablets, users can easily access and control their cloud-based applications and data, increasing their mobility and productivity.

Pay-per-use Model:

Cloud computing uses a pay-per-use business model that enables companies to only pay for the services they really utilize. This method is affordable, eliminates the need for up-front investments, and offers budget management flexibility for IT.

Scalable Storage Capacity:

Businesses can virtually store and manage a limitless amount of data in the cloud. The cloud offers a scalable and centralized storage option for all types of data, including documents, photos, audio, video, and other kinds of files.

Enhanced Data Security:

Cloud computing places a high focus on data security. To guarantee that data is handled and stored safely, cloud service providers offer cutting-edge security features like encryption, access limits, and regular security audits. Businesses can rest easy knowing that their important data is secure.

Disaster Recovery and Business Continuity:

Cloud computing provides reliable options for these two issues. Businesses can quickly bounce back from any unforeseen disasters or disruptions thanks to data redundancy, backup systems, and geographically dispersed data centers.

Agility and Innovation:

Businesses can continue to be innovative and nimble thanks to cloud computing. Organizations may quickly embrace new solutions, test out emerging trends, and promote corporate growth with access to a variety of cloud-based tools, services, and technology.

Green Computing:

By maximizing the use of computer resources, lowering energy use, and minimizing e-waste, cloud computing may support environmental sustainability.

By utilizing technologies like virtualization and load balancing to maximize the use of computer resources, cloud providers can operate large-scale data centers built for energy efficiency, resulting in lower energy usage and a smaller carbon footprint.

These benefits of cloud computing give companies the ability to use cutting-edge technology offered by cloud service providers while maximizing productivity, cost savings, scalability, and data security. They also enable them to concentrate on their core capabilities.

Disadvantages of Cloud Computing

When we talk about the "disadvantages of cloud computing," we're talking about any potential drawbacks or difficulties that businesses might have when utilizing cloud computing services. These drawbacks draw attention to some restrictions or risks related to cloud computing that businesses should take into account before making a choice.

Some of the Disadvantages of Cloud Computing are as follows:

- **Vendor Reliability and Downtime:**

Because of technological difficulties, maintenance needs, or even cyberattacks, cloud service providers can face outages or downtime. Users may not be able to access their data or applications during these times, which can interfere with business operations and productivity.

- **Internet Dependency:**

A dependable and fast internet connection is essential for cloud computing. Business operations may be delayed or interrupted if there are connectivity problems or interruptions in the internet service that affect access to cloud services and data.

- **Limited Control and Customization:**

Using standardized services and platforms offered by the cloud service provider is a common part of cloud computing. As a result, organizations may have less ability to customize and control their infrastructure, applications, and security measures. It may be difficult for some organizations to modify cloud services to precisely match their needs if they have special requirements or compliance requirements.

- **Data Security and Concerns about Privacy:**

Concerns about data security and privacy arise when sensitive data is stored on the cloud. Businesses must have faith in the cloud service provider's security procedures, data encryption, access controls, and regulatory compliance. Unauthorized access to data or data breaches can have serious repercussions, including financial loss, reputational harm, and legal obligations.

- **Hidden Costs and Pricing Models:**

Although pay-as-you-go models and lower upfront costs make cloud computing more affordable, businesses should be wary of hidden charges. Data transfer fees, additional storage costs, fees for specialized support or technical assistance, and expenses related to regulatory compliance are a few examples.

- **Dependency on Service Provider:**

When an organization depends on a cloud service provider, it is dependent on that provider's dependability, financial security, and longevity. Users may have disruptions and difficulties switching to alternate options if the provider runs into financial difficulties, changes their pricing policy, or even closes down their services.

- **Data Location and Compliance:**

When data is stored in the cloud, it frequently sits in numerous data centers around the globe that may be governed by multiple legal systems and data protection laws. This may pose compliance issues, especially if some sectors of the economy or nations have stringent data sovereignty laws.

Organizations should carry out a comprehensive risk assessment, thoroughly examine the dependability and security procedures of possible cloud service providers, and build backup and disaster recovery strategies to counteract these drawbacks.

What is NaaS (Network-as-a-Service) ?

Last Updated : 30 Apr, 2024

-
-
-

In the ever-evolving scene of networking administration and cloud computing, organizations are progressively turning to innovative solution for smooth out their tasks, improve productivity, and reduce costs. Network as a Service (NaaS) is one such solution that has gained popularity recently.

NaaS addresses a change in perspective in the way businesses manage and use their networking infrastructure. By utilizing cloud-based technologies, NaaS offers organizations the capacity to access, provision, and manage network assets on-request, all through an incorporated and effectively open platform.

In this complete guide, we dig into the complexities of Network as a Service, exploring its core ideas, primary terminologies, functional cycles, and practical applications, from understanding the basic standards of NaaS to examining real-world and use cases, this guide expects to furnish you with the information expected to successfully explore and harness the power of NaaS.

In order to assist you in comprehending how Network as a Service (NaaS) functions and its potential advantages for businesses, this guide will cover key terminologies, operational procedures, and practical examples.

What is NaaS ?

Network as a Service (NaaS) is a cloud computing model that provides organizations with on-demand access to networking resources over the Internet. In NaaS, traditional systems networking parts like routers, switches, firewalls, and different devices are virtualized and given as services through cloud service providers. This makes it possible for businesses to access and manage their network infrastructure without having to purchase any physical hardware. As a result, there is a reduction in operational complexity, costs associated with maintenance, and initial investment.

NaaS works on the standards of virtualization and software-defined networking (SDN), empowering associations to progressively arrange, configure, and manage their network resources through software-based interfaces. Users can undoubtedly increase their network infrastructure up or down based on changing prerequisites, change configurations, and monitor execution progressively.

Pay-per-use billing, in which organizations are charged based on how much they actually use networking services, and centralized management, where users can monitor and manage their network resources from a single dashboard, are two important aspects of NaaS, this utilization-based evaluating model offers cost proficiency and adaptability, as associations just compensate for the assets they consume.

How Does Network as a Service (NaaS) Work?

Through a cloud-based platform, Network as a Service (NaaS) provides organizations with internet access to networking resources, this is the way NaaS typically works:

- **Cloud-Based Platform: NaaS is conveyed through a cloud-based platform managed by a service provider. Organizations buy into the NaaS platform and access networking services through online portals, APIs, or command-line interfaces (CLIs),**
- **Centralized Management: Organizations can monitor and manage their network infrastructure from a single dashboard to centralized management interfaces provided by NaaS platforms. Administrators can see network execution measurements, break down traffic design, troubleshoot issues, and make configuration changes as needs be.**
- **Security and Compliance: Network infrastructure and data are protected from unauthorized access, data breaches, and other security threats by NaaS platforms' extensive security measures. They additionally offer consistence confirmations and adherence to industry principles to ensure data privacy, respectability, and accessibility.**
- **On-Demand Provisioning: With NaaS, organizations can provision networking resources on-request as per their particular prerequisites, through the self-service interfaces of the NaaS platform, they can set up virtual networks, define routing policies, set security setting, and deploy other networking components as needed.**
- **Pay-Per-Use Billing: Pay-per-use billing is typically used by NaaS providers, with organizations being charged based on how much they actually use networking services, this utilization based estimating model offers cost effectiveness and adaptability, as associations just compensation for the resources they consume, as opposed to putting resources into excess limit.**
- **Scalability and Flexibility: Organizations can dynamically scale their network infrastructure in response to shifting demand with NaaS. They can add or remove virtual networking**

components, adjust configurations, and allocate resources depending on the situation to support business development, seasonal fluctuations, or new initiative.

- **Virtualization of Networking Components: NaaS use virtualization technologies to abstract conventional networking components parts like routers, switches, firewalls, and load balancers from the basic actual equipment. These parts are virtualized and provisioned as software-based instances in the cloud.**

What are the types of Network as a Service resources?

Network as a Service (NaaS) offers an extensive variety of networking resources that organizations can access and use over the internet. Some of the key types of NaaS resources include:

Virtual Networks

- **Virtual Networks allows organizations to establish isolated network conditions in the cloud, complete with virtual subnets, IP addresses, and routing tables. These virtual networks empower secure communication between various applications, users, and services while maintaining logical separation.**

Routers and Switches

- **NaaS platforms give virtual routers and switches that work with the routing and switching of network traffic inside virtualized conditions, these product based routing and switching parts empower associations to define and manage network traffic streams, carry out routing policies, and ensure proficient data transmission.**

Load Balancers

- **NaaS platform offer virtual load-balancers that appropriate approaching organization traffic across numerous servers or resources to enhance execution, unwavering quality, and accessibility, these load balancers help equitably convey responsibilities, prevent over-load of individual resources, and ensure a smooth and steady user experience.**

VPN (Virtual Private Network)

- **Virtual private networks given by NaaS platform empower secure correspondence and data exchange between remote users, branch offices, and cloud-based resources, VPNs encrypt network traffic, verify users, and lay out private and secure associations over open network like the internet.**

Firewalls

- **Virtual firewalls presented as a component of NaaS arrangements allow organizations to support security policies and control access to network resources. These firewalls review incoming and active traffic, block unapproved access endeavors, and provide interruption detection and prevention abilities to protect against cyber threats.**

SD-WAN (Software-Defined Wide Area Network)

- **SD-WAN arrangements presented as a component of NaaS empower associations to improve and manage network across distributed areas. SD-WAN advances use software defined networking standards to powerfully route traffic over different network paths,**

further develop application execution, and decrease latency and data transfer capacity costs.

Advantages of Network as a Service (NaaS)

- **Scalability: Businesses can easily up or down their network infrastructure using NaaS, depending on their requirements, businesses that are experiencing rapid growth or have varying demands especially benefit from this flexibility.**
- **Cost-Efficiency: Businesses can save a lot of money upfront by using NaaS, which typically operates on a subscription or pay-per-use model, because businesses only pay for the services they use, this also makes it easier to predict costs.**
- **Accessibility: NaaS makes network resources accessible from anywhere with an internet connection, making it possible to work remotely and for teams with different locations to work together.**
- **Reliability: Service level agreements (SLAs) that ensures a certain level of uptime and performance are provided by many NaaS providers. Businesses that heavily rely on their network infrastructure for critical operations may particularly benefit from this.**
- **Rapid deployment: Businesses can quickly adapt to new requirements or market conditions adapt to the rapid deployment of NaaS solutions. In today's fast-paced business environment, this agility is essential.**
- **Global Reach: NaaS providers often have a global presence, businesses can easily expand their network's reach to new locations without making significant infrastructure investment.**
- **Improved efficiency: High-performance connectivity is typically provided by NaaS providers through the use of cutting-edge networking infrastructure and technologies, ensuring the best possible user experience for essential applications and services**
- **Compliance: When it comes to data protection and privacy, NaaS providers assist businesses by adhering to industry standards and regulations.**

Disadvantages of Network as a Service (NaaS)

- **Vendor Lock-In: Adopting NaaS solutions from a single service provider may result in vendor lock-in, making it difficult or costly for businesses to switch providers or return to in-house network management. Organizations that lack flexibility may be unable to take advantage of emerging technologies or adapt to shifting business requirements.**
- **Reliability and Availability: NaaS solutions may still experience downtime due to network issues, maintenance activities, or failures in the service provider's infrastructure, despite the promises of high availability and redundancy. To reduce the likelihood of disruptions to their operations, businesses must evaluate the provider's reliability and availability history.**
- **Dependency on Service Providers: Businesses are dependent on the provider's dependability, uptime, and quality of service when they rely on a third-party service provider for their network infrastructure, operations could be disrupted if the service provider has any issues.**

- **Data Security Concerns: Concerns about data privacy and security arise when sensitive data are entrusted to a third-party service providers. To ensure the safety of their data, businesses must carefully evaluate the security measures taken by NaaS providers.**
- **Potential Performance Issues: Limitations on bandwidth, latency, and congestion can have an impact on network performance, which can vary depending on the service provider and the underlying infrastructure, the NaaS solution's performance should be evaluated to see if it meets the needs of requirements**
- **Limited Control: With NaaS, organizations relinquish a level of control over their network infrastructure to the service provider. Concerns regarding customization, configuration, and troubleshooting may result from this lack of control, particularly for businesses with particular network requirements**
- **Integration Challenges: If the NaaS solution does not seamlessly integrate with other technologies utilized by the business, it can be difficult to integrate NaaS with existing IT systems, applications, and workflows. Compatibility issues may arise, necessitating additional integration efforts and resources.**

Naas Providers who are providing NaaS cloud computing platform

A Some of the top Network as a Service (NaaS) providers in the cloud computing industry include:

PROVIDER	DESCRIPTION
Amazon Web Services (AWS)	AWS offers various networking services like Virtual Private Cloud (VPC), AWS Transit Gateway, and AWS Global Accelerator to provide networking solutions
Google Cloud Platform (GCP)	GCP offers various networking services like Cloud VPN, Cloud Load Balancing and Virtual Private Cloud (VPC) to create global networks
Microsoft Azure	Azure provides networking services like Azure Virtual Network, Azure VPN Gateway, Azure Express Route and Azure Load Balancer to build secure and High-permanence networks in cloud
IBM Cloud	IBM Cloud provides networking services like VPC, IBM Cloud Load Balancer and IBM Cloud VPN offering to secure and reliable networks options in cloud
Oracle Cloud Infrastructure (OCI)	OCI offers networking services like VCN that is Virtual Cloud Network, VPN Connect, FastConnect and load balancing these are enabling high available and scalable networks in the cloud.

PROVIDER	DESCRIPTION
Cisco SD-WAN	**Cisco offers SD-WAN solutions this was provide secure connection.**
Juniper Contrail SD-WAN	**Juniper Network offers contrail SD-WAN this provide automated networking and security to simply network infrastructure.**

Conclusion on NaaS

Network as a Service (NaaS) provides businesses with a compelling set of benefits, which include simplified management, scalability, accessibility, and cost efficiency. businesses can build agile, secure, and high-performance networks in the cloud by utilizing NaaS solutions from leading providers like Amazon Web Services (AWS), Microsoft Azure, Google Cloud Platform (GCP), and others.

However, it is essential to acknowledge that NaaS has its own set of difficulties and considerations. Data security concerns, reliance on service providers, limited control, potential performance issues, and the need to comply with regulatory requirements are just a few examples, when adopting NaaS solutions, businesses should also carefully consider the risks, integration difficulties, and long-term costs of vendor lock-in.

In this complete guide, we dig into the complexities of Network as a Service, exploring its core ideas, primary terminologies, functional cycles, and practical applications, from understanding the basic standards of NaaS to examining real-world and use cases, this guide expects to furnish you with the information expected to successfully explore and harness the power of NaaS.

What is Database as a Service (DBaaS)?

Like SaaS, PaaS, and IaaS of cloud computing, we can consider DBaaS (also known as Managed Database Service) as a cloud computing service. It allows users associated with database activities to access and use a cloud database system without purchasing it.

1. DBaaS and cloud databases come under Software as a Service (SaaS) whose demand is growing so fast
2. Database as a Service (DBaaS) is self-service/on-demand database consumption coupled with automation of operations.
3. As we know cloud computing services are like pay-per-use so DBaaS is also based on the same payment structure like how much you will use just pay for your usage.
4. This DBaaS provides the same function as standard traditional and relational database models.
5. Using DBaaS, organizations can avoid database configuration, management, degradation, and security.
6. DBaaS consists of an info manager element, that controls all underlying info instances via API.

7. This API is accessible to the user through a management console, typically an online application, that the user might use to manage and assemble the info and even provision or deprovision info instances.

Features of DBaaS

1. **Scalability:** DBaaS platforms supports scaling resources automatically up and down based on the demand and the amount of storage and processing can be easily adjusted by the users without downtime.
2. **High Availability:** In DBaaS, data is replicated over multiple servers to ensure continuous availability.
3. **Automated Failover Mechanism:** Automated failover mechanism in DBaaS ensures minimal disruption in case of network or hardware failure.
4. **Caching:** Frequently accessed data is cache to speed up the query response time.
5. **Regular Updates:** The service provider handles regular updates, security fixes, and patches.
6. **Shared Infrastructure:** Multiple tenants share the same infrastructure, thus reducing the costs.
7. **Support for Multiple Database Engines:** DBaaS provides support for multiple database engines like MySQL, PostgreSQL, MongoDB, etc.
8. **Cost-efficiency:** The charges are based on the actual usage, thus this is more cost-effective than the traditional database solutions.

Who Uses DBaaS?

DBaaS is utilized by a wide range of organizations and users. Below are some of the user groups that utilize DBaaS:

1. **SaaS Providers:** They can leverage DBaaS efficiently to manage multi-tenant architectures, ensuring high availability and reliability.
2. **Data Analysts and Scientists:** They can utilize the features of DBaaS like performance and scalability while working with large datasets for analyzing and machine learning.
3. **Startups:** Startups have limited IT resources and can benefit from cost efficiency of DBaaS.
4. **Small and Medium-Sized Enterprises (SMEs):** SMEs can use DBaaS to avoid cost and maintenance expenses of the traditional database infrastructure.
5. **Large Enterprises:** Large enterprises can benefit from the performance optimization and scalability of DBaaS with complex and diverse workloads.

How does DBaaS Work?

DBaaS providers host data and database infrastructure while enabling access through API endpoints. They offer features like alerts, notifications, monitoring, constant support, and geo-replication for backups and availability.

1. DBaaS is delivered over the internet and users can access it through a web-based interface or API.

2. It can help to reduce the complexity and cost of managing databases.
3. It allows the development team to deploy and access a database without worrying about hardware purchase, hardware setup, database installation, and configuration, database maintenance and administration.
4. The organizations can free up the resources to focus on strategic initiatives by offloading tasks like data backup, recovery, etc. to a service provider.
5. It provides disaster recovery capability and helps to improve the availability and performance of databases.

DBaaS vs IaaS vs PaaS

Below are the differences between DBaaS, IaaS, and PaaS:

Parameters	**DBaaS**	**IaaS**	**PaaS**
Full Form	Database as a Service	Infrastructure as a Service	Platform as a Service
Definition	It is delivered as managed database services where the provider takes care of the patching, upgrading, and backing up the database.	In this access to cloud infrastructure resources is rented as individual services from a cloud service provider.	It is a complete development and deployment environment in the cloud.
User Responsibility	DBaaS requires minimal setup and maintenance.	In IaaS, users must secure their own data, software stacks, and operating systems that run their applications.	User is responsible for developing, maintaining, and managing data and user access within their applications.
Provider Responsibility	Provider responsibilities in DBaaS includes database setup, maintenance, backups.	Provider is responsible for maintaining the physical infrastructure.	Provider is responsible for physical security, power, cooling, and network connectivity.
Scalability	Automatic or easy scaling.	This requires manual scaling.	Automatic or easy scaling.

Parameters	DBaaS	IaaS	PaaS
Setup Complexity	Setup complexity is low in DBaaS.	Setup complexityis high in IaaS.	Setup complexity is medium in PaaS.
Security	Security is managed by provider and it is focussed on DB security.	Security is managed by user and have high control over security.	Security is managed by provider and is focussed mainly on application security.
Examples	Azure SQL, Amazon RDS, Amazon Aurora, IBM Db2.	Amazon Web Services, Microsoft Azure, and Google Compute Engine.	AWS Lambda, Microsoft Azure, Heroku, Cloud Foundry.

How to Choose a DBaaS?

Below are some of the key factors to consider when choosing a DBaaS:

1. **Database Type:** Determine the type of database needed. If you need a relational database like MySQL or a non-relational database like MongoDB.
2. **Performance:** Evaluate the performance capabilities such as latency, throughput before deciding the DBaaS.
3. **Uptime:** Check provider's Service Level Agreement to check for uptime guarantees.
4. **Security:** Check to ensure that DBaaS offers robust security features like data encryption, Identity and Access Management, and network security.
5. **Cost:** Consider the cost related to storage, data transfer, understand the pricing model and ensure that it aligns with your budget.
6. **Documentation:** Look for comprehensive and detailed documentation to help with setup and troubleshooting.

DBaaS Tools and Vendors

There are numerous DBaaS tools and vendors available, each catering to different needs. Here are some of prominent DBaaS tools and vendors:

1. **Amazon RDS:** This offers managed relational databases with support for multiple database engines, including MySQL, PostgreSQL, Oracle, MariaDB, and Microsoft SQL server. It offers automated backups, scaling, and patching.
2. **Google Cloud SQL:** It is fully managed relational relational database service for MySQL, PostgreSQL, and SQL server. It is best for businesses looking for seamless integration with Google Cloud Sevices. It offers services like automated backups, failover, scaling, and maintenance.

3. **MongoDB Atlas:** It is a fully managed cloud database service for MongoDB, offering flexible NoSQL database capabilities. It is suited for applications requiring flexible schema designs, high performance, and scalability, specially for unstructured data.
4. **Oracle Autonomous Database:** It offers self-driving, self-securing, and self-repairing database capabilities for both data warehousing and transaction processing. It is suited for enterprises needing high-performance, self-managing database.
5. **Microsoft Azure SQL Database:** It is a fully managed relational database service with AI-powered features for optimization and performance tuning. It is well-suited for enterprises seeking a robust, scalable SQL database solution.

Benefits of DBaaS

DBaaS offers numerous benefits to the organizations, enhancing efficiency, cost-efficiency, and many more. Below are some of the benefits of using DBaaS:

1. **Scalability:** DBaaS can automatically scale resources up and down based on the demand, ensuring cost-efficiency and optimal performance.
2. **High Availability:** DBaaS services ensures high availability as they often include built-in redundancy and failover mechanisms.
3. **Reliability:** DBaaS services ensures high reliability and business continuity with automated backup and disaster recovery solutions to help protect data against loss.
4. **Cost-efficiency:** DBaaS services offer pay-as-you-go pricing that allow organizations to pay only for the resources they use.
5. **Performance Optimization:** DBaaS providers ensures consistent performance by managing and optimizing the underlying infrastructure.
6. **Flexibility:** Many DBaaS providers support a variety of database engines like SQL, NoSQL, thus giving organizations the flexibility to choose the best tool for their needs.

Limitations of DBaaS

Database as a Service (DBaaS) offers many benefits but it also hs several limitations. Here are some of the limitations of DBaaS:

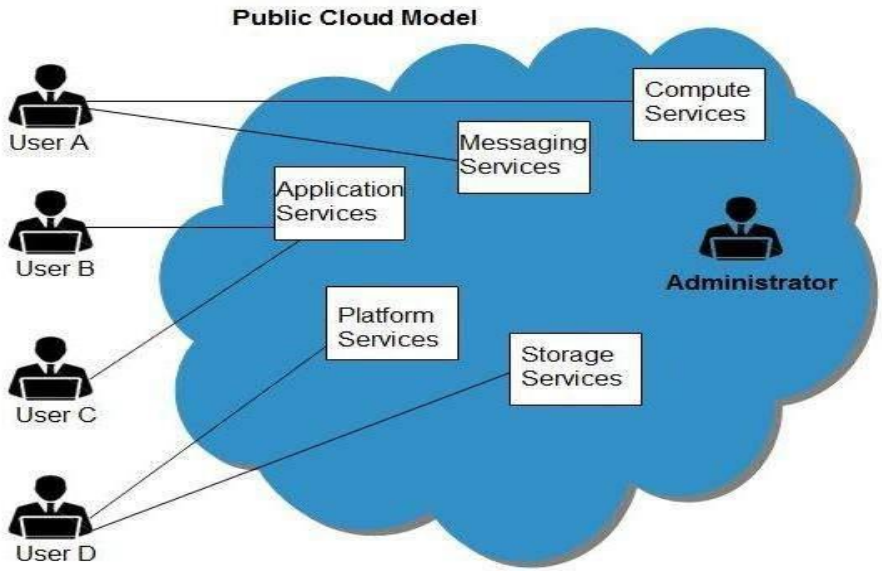
1. **Performance Variability:** DBaaS operates on shared resources due to which sometimes performance can vary due to resource contention with other tenants.
2. **Complex Customizations:** Sometimes it can be difficult to optimize the database for specific needs as users have limited control over the underlying hardware and software.
3. **Hidden Costs:** Additional storage charges and charges for premium features can lead to unexpected expenses.
4. **Complex Data Portability:** Transferring huge volumes of data between providers or back to on-premises systems can be complex and costly.
5. **Time-consuming Data Transfer:** Migrating large datasets to a DBaaS platform can be time-consuming and may require downtime.

Types of Cloud

Public Cloud:

Public Cloud allows systems and services to be easily accessible to general public. The IT giants such as Google, Amazon and Microsoft offer cloud services via Internet.

The Public Cloud Model is shown in the diagram below.



Benefits:

There are many benefits of deploying cloud as public cloud model. The following diagram shows some of those benefits:

Cost Effective

Since public cloud shares same resources with large number of customers it turns out inexpensive.

Reliability

The public cloud employs large number of resources from different locations. If any of the resources fails, public cloud can employ another one.

Flexibility

The public cloud can smoothly integrate with private cloud, which gives customers a flexible approach.

Location Independence

Public cloud services are delivered through Internet, ensuring location independence.

Utility Style Costing

Public cloud is also based on pay-per-use model and resources are accessible whenever customer needs them.

High Scalability

Cloud resources are made available on demand from a pool of resources, i.e., they can be scaled up or down according the requirement.

Disadvantages:

Here are some disadvantages of public cloud model:

Low Security

In public cloud model, data is hosted off-site and resources are shared publicly, therefore does not ensure higher level of security.

Less Customizable

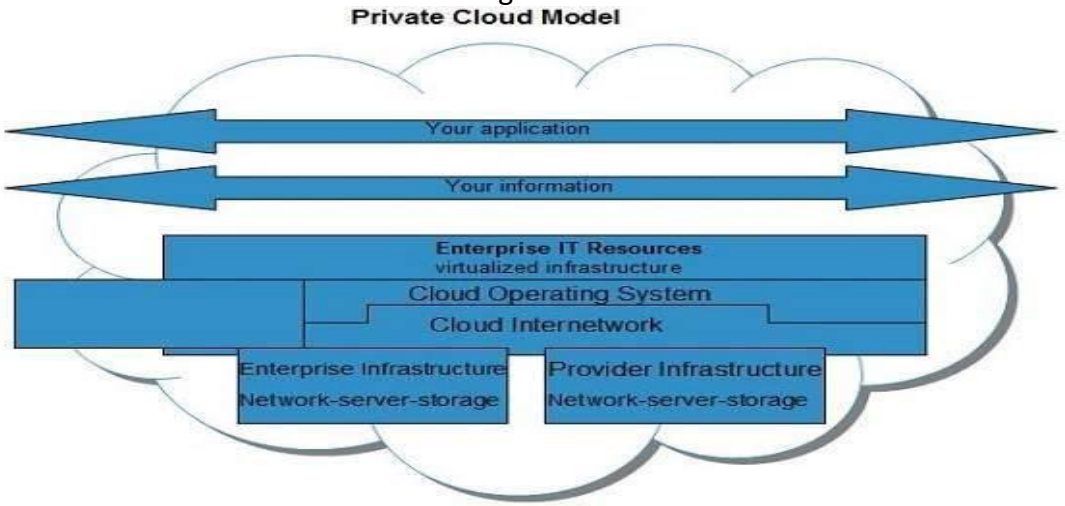
It is comparatively less customizable than private cloud.

Private Cloud:

Private Cloud allows systems and services to be accessible within an organization.

The Private Cloud is operated only within a single organization. However, it may be managed internally by the organization itself or by third-party.

The private cloud model is shown in the diagram below.



Benefits:

There are many benefits of deploying cloud as private cloud model. The following diagram shows some of those benefits:

High Security and Privacy

Private cloud operations are not available to general public and resources are shared from distinct pool of resources. Therefore, it ensures high security and privacy.

More Control

The private cloud has more control on its resources and hardware than public cloud because it is accessed only within an organization.

Cost and Energy Efficiency

The private cloud resources are not as cost effective as resources in public clouds but they offer more efficiency than public cloud resources.

Disadvantages:

Here are some disadvantages of using private cloud model:

Restricted Area of Operation

The private cloud is only accessible locally and is very difficult to deploy globally.

High Priced

Purchasing new hardware in order to fulfill the demand is a costly transaction.

Limited Scalability

The private cloud can be scaled only within capacity of internal hosted resources.

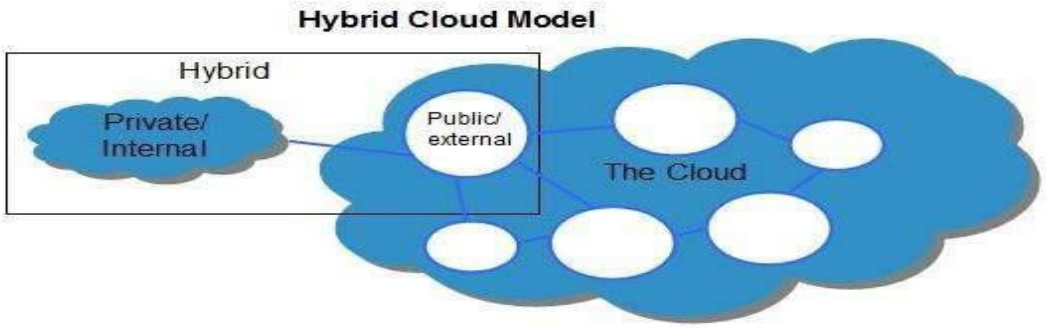
Additional Skills

In order to maintain cloud deployment, organization requires skilled expertise.

Hybrid Cloud:

Hybrid Cloud is a mixture of public and private cloud.

Non-critical activities are performed using public cloud while the critical activities are performed using private cloud. The Hybrid Cloud Model is shown in the diagram below.



Benefits:

There are many benefits of deploying cloud as hybrid cloud model. The following diagram shows some of those benefits:

Scalability

It offers features of both, the public cloud scalability and the private cloud scalability.

Flexibility

It offers secure resources and scalable public resources.

Cost Efficiency

Public clouds are more cost effective than private ones. Therefore, hybrid clouds can be cost saving.

Security

The private cloud in hybrid cloud ensures higher degree of security.

Disadvantages:

Networking Issues

Networking becomes complex due to presence of private and public cloud.

Security Compliance

It is necessary to ensure that cloud services are compliant with security policies of the organization.

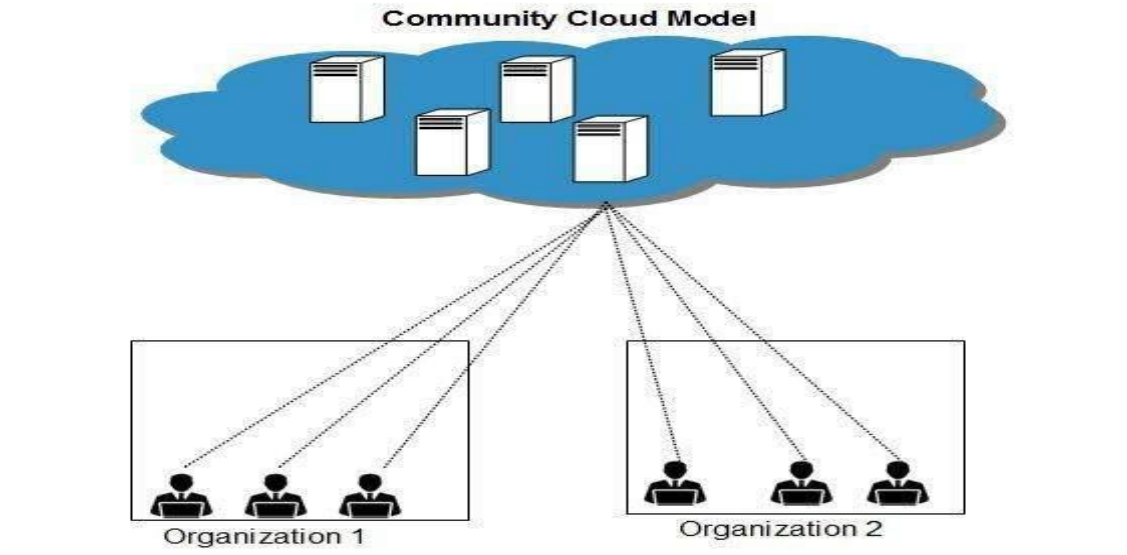
Infrastructure Dependency

The hybrid cloud model is dependent on internal IT infrastructure, therefore it is necessary to ensure redundancy across data centers.

Community Cloud:

Community Cloud allows system and services to be accessible by group of organizations. It shares the infrastructure between several organizations from a specific community.

It may be managed internally by organizations or by the third-party. The Community Cloud Model is shown in the diagram below.



Benefits:

There are many benefits of deploying cloud as community cloud model.

Cost Effective

Community cloud offers same advantages as that of private cloud at low cost.

Sharing Among Organizations

Community cloud provides an infrastructure to share cloud resources and capabilities among several organizations.

Security

The community cloud is comparatively more secure than the public cloud but less secured than the private cloud.

Issues:

- Since all data is located at one place, one must be careful in storing data in community cloud because it might be accessible to others.
- It is also challenging to allocate responsibilities of governance, security and cost among organizations.

Cloud computing architecture

Cloud computing architecture refers to the components and subcomponents required for cloud computing. These components typically consist of a front end platform (fat client, thin client, mobile device), back end platforms (servers, storage), a cloud based delivery, and a network (Internet, Intranet, Intercloud).

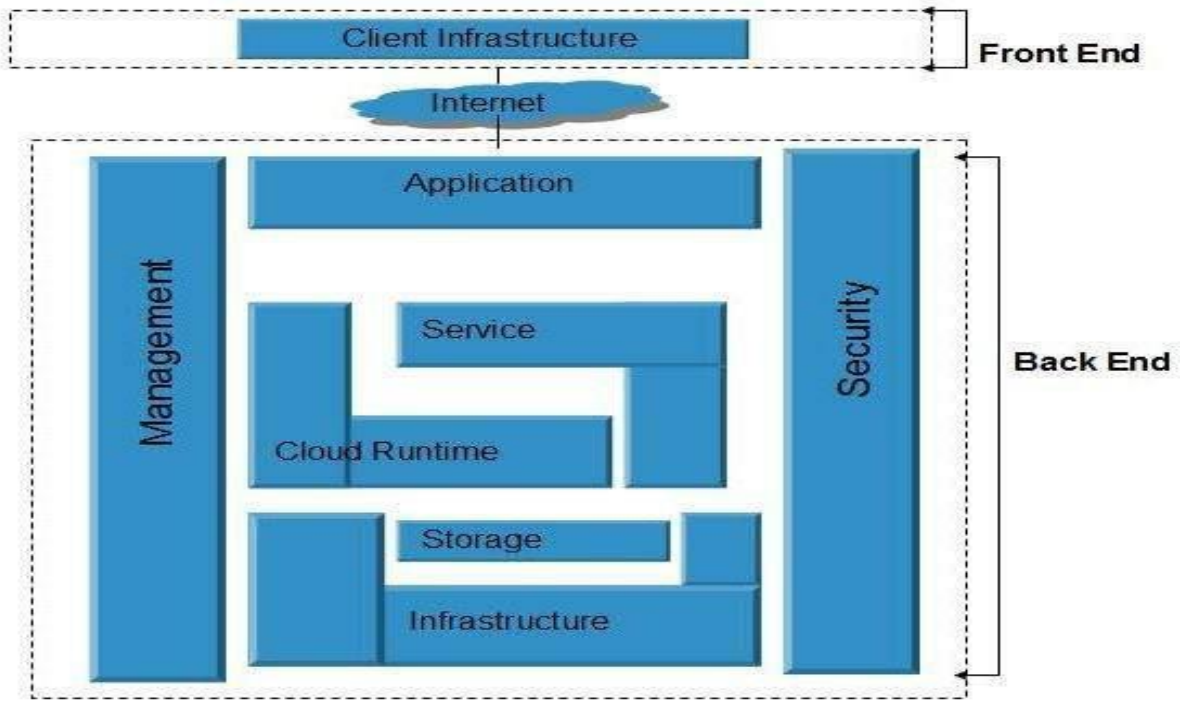
[A **fat client** (also called heavy, **rich** or **thick client**) is a computer (**clients**), in **client**–server architecture or networks, that typically provides **rich** functionality independent of the central server. It is Originally known as just a "client" or "thick client," the name is contrasted to thin client, which describes a computer heavily dependent on a server's applications.]

Cloud Computing architecture comprises of many cloud components, which are loosely coupled. We can broadly divide the cloud architecture into two parts:

- Front End
- Back End

Each of the ends is connected through a network, usually Internet.

The following diagram shows the graphical view of cloud computing architecture:



Front End:

The front end refers to the client part of cloud computing system. It consists of interfaces and applications that are required to access the cloud computing platforms, Example - Web Browser.

Back End:

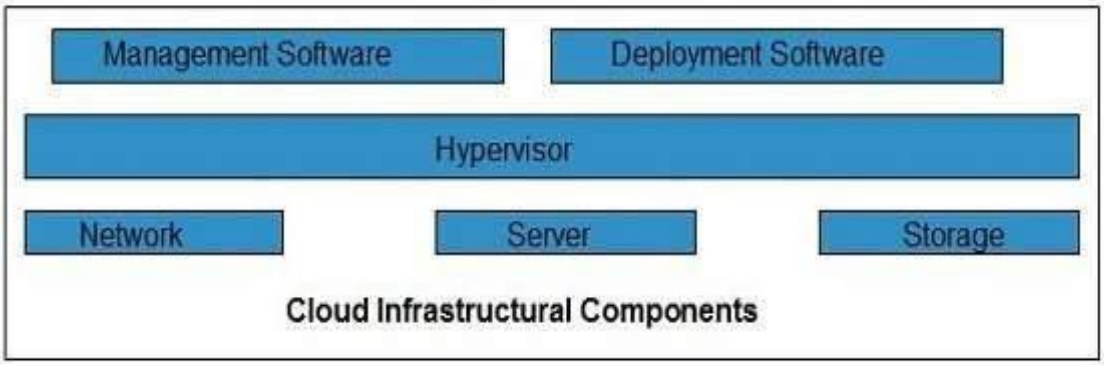
The back End refers to the cloud itself. It consists of all the resources required to provide cloud computing services. It comprises of huge data storage, virtual machines, security mechanism, services, deployment models, servers, etc.

Note:

- It is the responsibility of the back end to provide built-in security mechanism, traffic control and protocols.
- The server employs certain protocols known as middleware, which help the connected devices to communicate with each other.

Cloud Computing Infrastructure

Cloud infrastructure consists of servers, storage devices, network, cloud management software, deployment software, and platform virtualization.



Hypervisor:

Hypervisor is a firmware or low-level program that acts as a Virtual Machine Manager. It allows to share the single physical instance of cloud resources between several occupants.

Management Software:

It helps to maintain and configure the infrastructure.

Deployment Software:

It helps to deploy and integrate the application on the cloud.

Network:

It is the key component of cloud infrastructure. It allows to connect cloud services over the Internet. It is also possible to deliver network as a utility over the Internet, which means, the customer can customize the network route and protocol.

Server:

The **server** helps to compute the resource sharing and offers other services such as resource allocation and de-allocation, monitoring the resources, providing security etc.

Storage:

Cloud keeps multiple replicas of storage. If one of the storage resources fails, then it can be extracted from another one, which makes cloud computing more reliable.

Infrastructural Constraints

Fundamental constraints that cloud infrastructure should implement are shown in the following diagram:

Transparency:

Virtualization is the key to share resources in cloud environment. But it is not possible to satisfy the demand with single resource or server. Therefore, there must be transparency in resources, load balancing and application, so that we can scale them on demand.

Scalability:

Scaling up an application delivery solution is not that easy as scaling up an application because it involves configuration overhead or even re-architecting the network. So, application delivery solution is need to be scalable which will require the virtual infrastructure such that resource can be provisioned and de-provisioned easily.

Intelligent Monitoring:

To achieve transparency and scalability, application solution delivery will need to be capable of intelligent monitoring.

Security:

The mega data center in the cloud should be securely architected. Also the control node, an entry point in mega data center, also needs to be secure.

Virtualization

Virtualization in Cloud Computing is a process in which the user of cloud shares the data present in the cloud which can be application software etc. It provides a virtual environment in the cloud which can be software hardware or any other thing.

Virtualization is a technique how to separate a service from the underlying physical delivery of that service. It is the process of creating a virtual version of something like computer hardware.

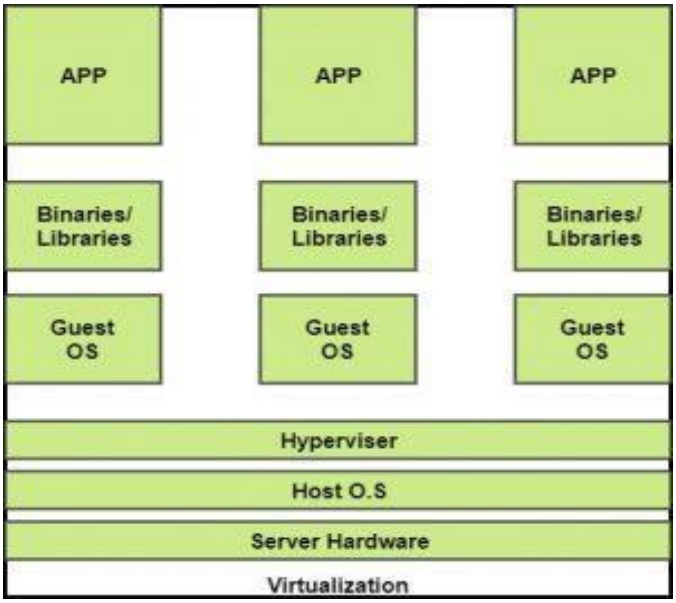
It involves using specialized software to create a virtual or software-created version of a computing resource rather than the actual version of the same resource.

With the help of Virtualization multiple operating systems and applications can run on same Machine and its same hardware at the same time increasing the utilization and flexibility of hardware.

In other words, One of the main cost effective, hardware reducing, energy saving techniques used by cloud providers is virtualization.

Virtualization is a technique, which allows to share a single physical instance of a resource or an application among multiple customers and organizations at one time. It does by assigning a logical name to a physical storage and providing a pointer to that physical resource on demand.

The term virtualization is often synonymous with hardware virtualization, which plays a fundamental role in efficiently delivering Infrastructure-as-a-Service (IaaS) solutions for cloud computing. Moreover, virtualization technologies provide a virtual environment for not only executing applications but also for storage, memory, and networking.



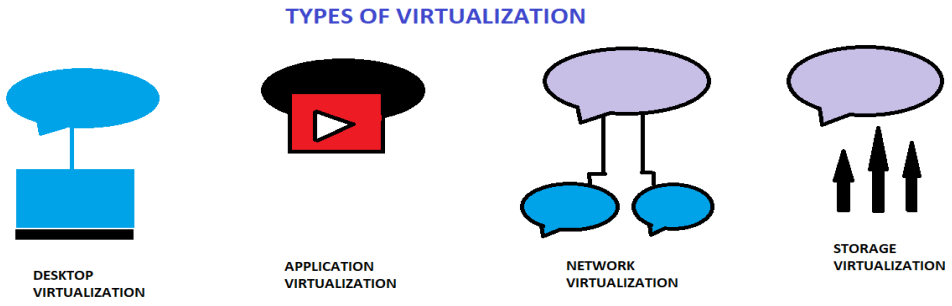
The machine on which the virtual machine is going to be build is known as Host Machine and that virtual machine is referred as a Guest Machine.

BENEFITS OF VIRTUALIZATION:

1. More flexible and efficient allocation of resources.
2. Enhance development productivity.
3. It lowers the cost of IT infrastructure.
4. Remote access and rapid scalibility.
5. High availability and disaster recovery.
6. Pay per use of the IT infrastructure on demand.
7. Enables running multiple operating system.

Types of Virtualization:

1. Application Virtualization.
2. Network Virtualization.
3. Desktop Virtualization.
4. Storage Virtualization.



1. Application Virtualization:

Application virtualization helps user to have a remote access of an application from a server. The server stores all personal information and other characteristics of the application, but can still run on a local workstation through internet. Example of this would be a user who needs to run two different versions of the same software. Technologies that use application virtualization are hosted applications and packaged applications.

2. Network Virtualization:

Network virtualization, provides a facility to create and provision virtual networks—logical switches, routers, firewalls, Virtual Private Network (VPN), and workload security within days or even in weeks.

3. Desktop Virtualization:

Desktop virtualization allows the users' OS to be remotely stored on a server in the data center. It allows the user to access their desktop virtually, from any location by different machine. Users who wants specific operating systems other than Windows Server will need to have a virtual desktop. Main benefits of desktop virtualization are user mobility, portability, easy management of software installation, updates and patches.

4. Storage Virtualization:

Storage virtualization is an array of servers that are managed by a virtual storage system. The servers aren't aware of exactly where their data is stored. It makes managing storage from multiple sources to be managed and utilized as a single repository. storage virtualization software maintains smooth operations, consistent performance and a continuous suite of advanced functions despite changes, break down and differences in the underlying equipment.

Hypervisor

Hypervisor is a form of virtualization software used in Cloud hosting to divide and allocate the resources on various pieces of hardware.

The program which provide partitioning, isolation or abstraction is called virtualization hypervisor.

Hypervisor is a hardware virtualization technique that allows multiple guest operating systems (OS) to run on a single host system at the same time.

A hypervisor is sometimes also called a virtual machine manager(VMM).

Types of Hypervisor:-

TYPE-1 Hypervisor:

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

TYPE-2 Hypervisor:

A Host operating system runs on undrlying host system. It is also known as 'Hosted Hypervisor".

Basically a software installed on an operating system. Hypervisor asks operating system to make hardware calls. Hosted hypervisors are often found on endpoints like PCs.

Choosing the right hypervisor:

Type 1 hypervisors offer much better performance than Type 2 because there's no middle layer, making them the logical choice for mission-critical applications and workloads.

But that's not to say that hosted hypervisors don't have their place – they're much simpler to set up, so they're a good.

One of the best ways to determine which hypervisor meets your needs is to compare their performance metrics. These include CPU overhead, amount of maximum host and guest memory, and support for virtual processors.

The following factors should be examined before choosing a suitable hypervisor:

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

Needs for a virtualization hypervisor are:

a. Flexibility
b. Scalability
c. Usability
d. Availability
e. Reliability
f. Efficiency
g. Reliable support

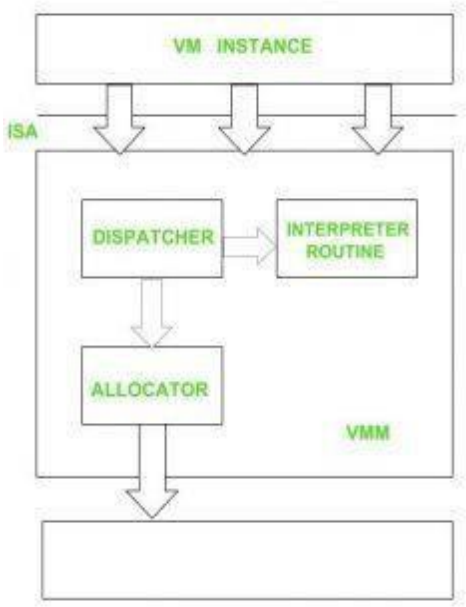
2. The cost of a hypervisor: For many buyers, the toughest part of choosing a hypervisor is striking the right balance between cost and functionality. While a number of entry-level solutions are free, or practically free, the prices at the opposite end of the market can be overwhelming. Licensing frameworks also vary, so it's important to be aware of exactly what you're getting for your money.

3. Virtual machine performance: Virtual systems should meet or exceed the performance of their physical counterparts, at least in relation to the applications within each server. Everything beyond meeting this benchmark is profit.

4. Ecosystem: It's tempting to overlook the role of a hypervisor's ecosystem – that is, the availability of documentation, support, training, third-party developers and consultancies, and so on – in determining whether or not a solution is cost-effective in the long term.

5. Test for yourself: You can gain basic experience from your existing desktop or laptop. You can run both Hypervisor to create a nice virtual learning and testing environment.

HYPERVISOR REFERENCE MODEL



There are 3 main modules coordinate in order to follow the fundamental hardware:

1. Dispatcher
2. Allocator
3. Interpreter

DISPATCHER:
The dispatcher behaves like the entry point of the monitor and reroutes the instructions of the virtual machine instance to one of the other two modules.

ALLOCATOR:
The allocator is responsible for deciding the system resources to be provided to the virtual machine instance. It means whenever virtual machine tries to execute an instruction that results in changing the machine resources associated with the virtual machine, the allocator is invoked by the dispatcher.

INTERPRETER:
The interpreter module consists of interpreter routines. These are executed, whenever virtual machine executes a privileged instruction.

CPU Virtualization

Virtualization of the hardware is known as CPU Virtualization. This is where any hardware platform that can be controlled by the user or a guest software over a virtual machine on a platform, virtually. This is not limited to guest software but also several operating systems.

A VM is a duplicate of an existing computer system in which a majority of the VM instructions are executed on the host processor in native mode. Thus, unprivileged instructions of VMs run directly on the host machine for higher efficiency. Other critical instructions should be handled carefully for correctness and stability.

The critical instructions are divided into three categories: privileged instructions, control-sensitive instructions, and behavior-sensitive instructions.

Privileged instructions execute in a privileged mode and will be attentive if executed outside this mode. Control-sensitive instructions attempt to change the configuration of resources used. Behavior-sensitive instructions have different behaviors depending on the configuration of resources, including the load and store operations over the virtual memory.

Memory Virtualization

Virtual memory virtualization is similar to the virtual memory support provided by modern operating systems.

In a traditional execution environment, the operating system maintains mappings of virtual memory to machine memory using page tables, which is a one-stage mapping from virtual memory to machine memory.

All modern x86 CPUs include a memory management unit (MMU) and a translation look aside buffer (TLB) to optimize virtual memory performance. However, in a virtual execution environment, virtual memory virtualization involves sharing the physical system memory in RAM and dynamically allocating it to the physical memory of the VMs.

That means a two-stage mapping process should be maintained by the guest OS and the VMM, respectively: virtual memory to physical memory and physical memory to machine memory.

Furthermore, MMU virtualization should be supported, which is transparent to the guest OS. The guest OS continues to control the mapping of virtual addresses to the physical memory addresses of VMs. But the guest OS cannot directly access the actual machine memory.

The VMM is responsible for mapping the guest physical memory to the actual machine memory.

Since each page table of the guest OS has a separate page table in the VMM corresponding to it, the VMM page table is called the shadow page table.

Nested page tables add another layer of indirection to virtual memory. The MMU already handles virtual-to-physical translations as defined by the OS. Then the physical memory addresses are translated to machine addresses using another set of page tables defined by the hypervisor.

Since modern operating systems maintain a set of page tables for every process, the shadow page tables will get flooded. Consequently, the performance overhead and cost of memory will be very high.

When the guest OS changes the virtual memory to a physical memory mapping, the VMM updates the shadow page tables to enable a direct lookup.

I/O Virtualization

I/O virtualization involves managing the routing of I/O requests between virtual devices and the shared physical hardware.

There are three ways to implement I/O virtualization: full device emulation, para-virtualization, and direct I/O.

Full device emulation is the first approach for I/O virtualization. Generally, this approach emulates well-known, real-world devices.

All the functions of a device or bus infrastructure, such as device enumeration, identification, interrupts, are replicated in software. This software is located in the VMM and acts as a virtual device.

The I/O access requests of the guest OS are trapped in the VMM which interacts with the I/O devices.

A single hardware device can be shared by multiple VMs that run concurrently. However, software emulation runs much slower than the hardware.

The para-virtualization method of I/O virtualization is typically used in Xen. [Xen Project is a type-1 hypervisor, providing services that allow multiple computer operating systems to execute on the same computer hardware concurrently.]

It is also known as the split driver model consisting of a front end driver and a backend driver. The frontend driver is running in Domain U and the backend driver is running in Domain 0. They interact with each other via a block of shared memory.

The frontend driver manages the I/O requests of the guest OS and the backend driver is responsible for managing the real I/O devices and multiplexing the I/O data of different VMs.
Although para-I/O-virtualization achieves better device performance than full device emulation, it comes with a higher CPU overhead.

Direct I/O virtualization lets the VM access devices directly. It can achieve close-to-native performance without high CPU costs. However, current direct I/O virtualization implementations focus on networking for mainframes.

There are a lot of challenges for commodity hardware devices. For example, when a physical device is reclaimed (required by workload migration) for later reassignment, it may have been set to an arbitrary state that can function incorrectly or even crash the whole system.

Since software-based I/O virtualization requires a very high overhead of device emulation, hardware-assisted I/O virtualization is critical.

VIRTUAL CLUSTERS AND RESOURCE MANAGEMENT

A physical cluster is a collection of servers (physical machines) interconnected by a physical network such as a LAN.

When a traditional VM is initialized, the administrator needs to manually write configuration information or specify the configuration sources. When more VMs join a network, an inefficient configuration always causes problems with overloading or underutilization.

Amazon's Elastic Compute Cloud (EC2) is a good example of a web service that provides elastic computing power in a cloud. EC2 permits customers to create VMs and to manage user accounts over the time of their use.

Most virtualization platforms, including XenServer and VMware ESX Server, support a bridging mode which allows all domains to appear on the network as individual hosts. By using this mode, VMs can communicate with one another freely through the virtual network interface card and configure the network automatically.

Physical versus Virtual Clusters:-

Virtual clusters are built with VMs installed at distributed servers from one or more physical clusters.

The VMs in a virtual cluster are interconnected logically by a virtual network across several physical networks.

Each virtual cluster is formed with physical machines or a VM hosted by multiple physical clusters. The virtual cluster boundaries are shown as distinct boundaries.

The provisioning of VMs to a virtual cluster is done dynamically to have the following interesting properties:

- The virtual cluster nodes can be either physical or virtual machines. Multiple VMs running with different OS can be deployed on the same physical node.
- A VM runs with a guest OS, which is often different from the host OS, that manages the resources in the physical machine, where the VM is implemented.
- The purpose of using VMs is to consolidate multiple functionalities on the same server. This will greatly enhance server utilization and application flexibility.

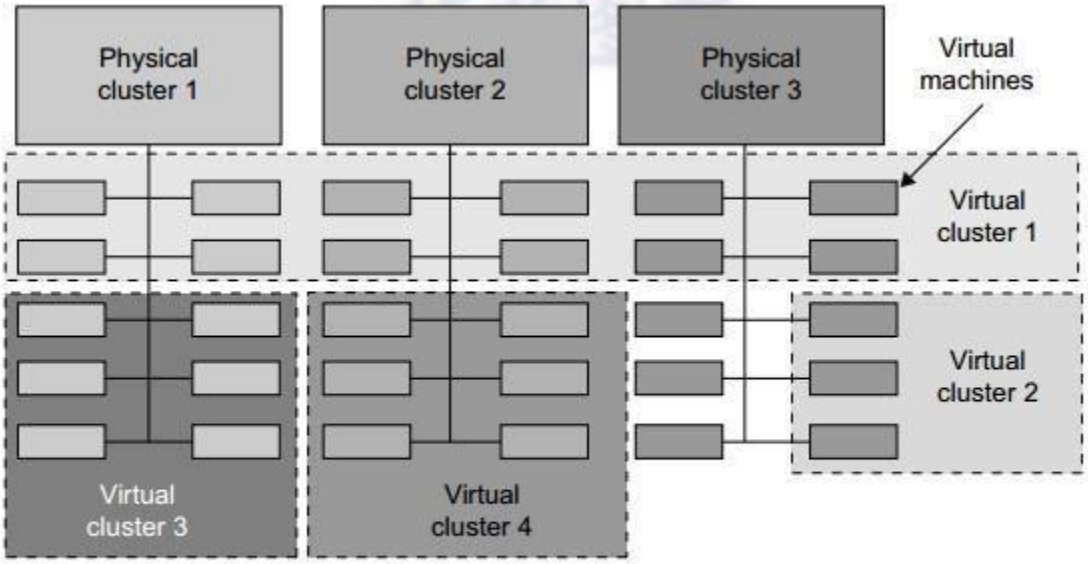


FIGURE 3.18

A cloud platform with four virtual clusters over three physical clusters shaded differently.

- VMs can be colonized (replicated) in multiple servers for the purpose of promoting distributed parallelism, fault tolerance, and disaster recovery.

- The size (number of nodes) of a virtual cluster can grow or shrink dynamically, similar to the way an overlay network varies in size in a peer-to-peer (P2P) network.

- The failure of any physical nodes may disable some VMs installed on the failing nodes. But the failure of VMs will not pull down the host system.

Fast Deployment and Effective Scheduling:

The system should have the capability of fast deployment. Here, deployment means two things: to construct and distribute software stacks (OS, libraries, applications) to a physical node inside clusters as fast as possible, and to quickly switch runtime environments from one user's virtual cluster to another user's virtual cluster.

If one user finishes using his system, the corresponding virtual cluster should shut down or suspend quickly to save the resources to run other VMs for other users.

The concept of "green computing" has attracted much attention recently. However, previous approaches have focused on saving the energy cost of components in a single workstation without a global vision. Consequently, they do not necessarily reduce the power consumption of the whole cluster.

The live migration of VMs allows workloads of one node to transfer to another node. However, it does not guarantee that VMs can randomly migrate among themselves. In fact, the probable overhead caused by live migrations of VMs cannot be ignored.

The overhead may have serious negative effects on cluster utilization and throughput issues. Therefore, the challenge is to determine how to design migration strategies to implement green computing without influencing the performance of clusters.

Another advantage of virtualization is load balancing of applications in a virtual cluster. Load balancing can be achieved using the load index and frequency of user logins. The automatic scale-up and scale-down mechanism of a virtual cluster can be implemented based on this model. Consequently, we can increase the resource utilization of nodes and shorten the response time of systems.

Mapping VMs onto the most appropriate physical node should promote performance. Dynamically adjusting loads among nodes by live migration of VMs is desired, when the loads on cluster nodes become quite unbalanced.

High-Performance Virtual Storage:

The template VM can be distributed to several physical hosts in the cluster to customize the VMs. In addition, existing software packages reduce the time for customization as well as switching virtual environments.

It is important to efficiently manage the disk spaces occupied by template software packages. Some storage architecture design can be applied to reduce duplicated blocks in a distributed file system of virtual clusters.

Basically, there are four steps to deploy a group of VMs onto a target cluster: preparing the disk image, configuring the VMs, choosing the destination nodes, and executing the VM deployment command on every host.

Many systems use templates to simplify the disk image preparation process. A template is a disk image that includes a preinstalled operating system with or without certain application software.

Users choose a proper template according to their requirements and make a duplicate of it as their own disk image. Templates could implement the COW (Copy on Write) format. A new COW backup file is very small and easy to create and transfer. Therefore, it definitely

reduces disk space consumption. In addition, VM deployment time is much shorter than that of copying the whole raw image file.
Every VM is configured with a name, disk image, network setting, and allocated CPU and memory. One needs to record each VM configuration into a file. However, this method is inefficient when managing a large group of VMs.

Live VM Migration Steps and Performance Effects:-

In a cluster built with mixed nodes of host and guest systems, the normal method of operation is to run everything on the physical machine. When a VM fails, its role could be replaced by another VM on a different node, as long as they both run with the same guest OS.

In other words, a physical node can fail over to a VM on another host. This is different from physical-to-physical failover in a traditional physical cluster. The advantage is enhanced failover flexibility.

The potential drawback is that a VM must stop playing its role if its residing host node fails. However, this problem can be mitigated with VM life migration. The migration copies the VM state file from the storage area to the host machine.

There are four ways to manage a virtual cluster:

First, you can use a guest-based manager, by which the cluster manager resides on a guest system. In this case, multiple VMs form a virtual cluster. For example, openMosix is an open source Linux cluster running different guest systems on top of the Xen hypervisor. Another example is Sun's cluster Oasis, an experimental Solaris cluster of VMs supported by a VMware VMM.

Second, you can build a cluster manager on the host systems. The host-based manager supervises the guest systems and can restart the guest system on another physical machine. A good example is the VMware HA system that can restart a guest system after failure. These two cluster management systems are either guest-only or host-only, but they do not mix.

A **third way** to manage a virtual cluster is to use an independent cluster manager on both the host and guest systems. This will make infrastructure management more complex, however.

Finally, you can use an integrated cluster on the guest and host systems. This means the manager must be designed to distinguish between virtualized resources and physical resources. Various cluster management schemes can be greatly enhanced when VM life migration is enabled with minimal overhead.

VMs can be live-migrated from one physical machine to another; in case of failure, one VM can be replaced by another VM. Virtual clusters can be applied in computational grids, cloud platforms, and high-performance computing (HPC) systems.

The major attraction of this scenario is that virtual clustering provides dynamic resources that can be quickly put together upon user demand or after a node failure. In particular, virtual clustering plays a key role in cloud computing.

When a VM runs a live service, it is necessary to make an exchange to ensure that the migration occurs in a manner that minimizes all three metrics.

Furthermore, we should ensure that the migration will not disrupt other active services residing in the same host through resource contention (e.g., CPU, network bandwidth).

A VM can be in one of the following four states:

An **inactive state** is defined by the virtualization platform, under which the VM is not enabled.

An **active state** refers to a VM that has been instantiated at the virtualization platform to perform a real task.

A **paused state** corresponds to a VM that has been instantiated but disabled to process a task or paused in a waiting state.

A VM enters the **suspended state** if its machine file and virtual resources are stored back to the disk.

Llive migration of a VM consists of the following six steps:

Steps 0 and 1: Start migration.

This step makes preparations for the migration, including determining the migrating VM and the destination host. Although users could manually make a VM migrate to an appointed host, in most circumstances, the migration is automatically started by strategies such as load balancing and server consolidation.

Steps 2: Transfer memory.

Since the whole execution state of the VM is stored in memory, sending the VM's memory to the destination node ensures continuity of the service provided by the VM. All of the memory data is transferred in the first round, and then the migration controller recopies the memory data which is changed in the last round. These steps keep iterating until the dirty portion of the memory is small enough to handle the final copy.

Step 3: Suspend the VM and copy the last portion of the data.

The migrating VM's execution is suspended when the last round's memory data is transferred. Other non memory data such as CPU and network states should be sent as well. During this step, the VM is stopped and its applications will no longer run. This "service unavailable" time is called the "downtime" of migration, which should be as short as possible so that it can be negligible to users.

Steps 4 and 5: Commit and activate the new host.

After all the needed data is copied, on the destination host, the VM reloads the states and recovers the execution of programs in it, and the service provided by this VM continues. Then the network connection is redirected to the new VM and the dependency to the source host is cleared. The whole migration process finishes by removing the original VM from the source host.

Migration of Memory, Files, and Network Resources:-

Since clusters have a high initial cost of ownership, including space, power conditioning, and cooling equipment, leasing or sharing access to a common cluster is an attractive solution when demands vary over time.

Shared clusters offer economies of scale and more effective utilization of resources by multiplexing. When one system migrates to another physical node, we should consider the following issues.

Memory Migration:

This is one of the most important aspects of VM migration. Moving the memory instance of a VM from one physical host to another can be approached in any number of ways. But traditionally, the concepts behind the techniques tend to share common implementation paradigms. The techniques employed for this purpose depend upon the characteristics of application/workloads supported by the guest OS.

Memory migration can be in a range of hundreds of megabytes to a few gigabytes in a typical system today, and it needs to be done in an efficient manner.

File System Migration:

To support VM migration, a system must provide each VM with a consistent, location-independent view of the file system that is available on all hosts. A simple way to achieve this is to provide each VM with its own virtual disk which the file system is mapped to and transport the contents of this virtual disk along with the other states of the VM. However, due to the current trend of high-capacity disks, migration of the contents of an entire disk over a network is not a viable solution.

Another way is to have a global file system across all machines where a VM could be located. This way removes the need to copy files from one machine to another because all files are network-accessible.

The relevant VM files are explicitly copied into the local file system for a resume operation and taken out of the local file system for a suspend operation. This approach relieves developers from the complexities of implementing several different file system calls for different distributed file systems.

Network Migration:

A migrating VM should maintain all open network connections without relying on forwarding mechanisms on the original host or on support from mobility or redirection mechanisms.

To enable remote systems to locate and communicate with a VM, each VM must be assigned a virtual IP address known to other entities. This address can be distinct from the IP address of the host machine where the VM is currently located.

Each VM can also have its own distinct virtual MAC address. The VMM maintains a mapping of the virtual IP and MAC addresses to their corresponding VMs.

In general, a migrating VM includes all the protocol states and carries its IP address with it.

Live migration means moving a VM from one physical node to another while keeping its OS environment and applications unbroken. This capability is being increasingly utilized in today's enter-prise environments to provide efficient online system maintenance, reconfiguration, load balancing, and proactive fault tolerance.

It provides desirable features to satisfy requirements for computing resources in modern computing systems, including server consolidation, performance isolation, and ease of management.

Live migration is a key feature of system virtualization technologies. Only memory and CPU status needs to be transferred from the source node to the target node.

Infrastructure as a Service (IaaS)

Iaas is also known as **Hardware as a Service (HaaS)**. It is one of the layers of the cloud computing platform. It allows customers to outsource their IT infrastructures such as servers, networking, processing, storage, virtual machines, and other resources. Customers access these resources on the Internet using a pay-as-per use model.

In traditional hosting services, IT infrastructure was rented out for a specific period of time, with pre-determined hardware configuration. The client paid for the configuration and time, regardless of the actual use. With the help of the IaaS cloud computing platform layer, clients can dynamically scale the configuration to meet changing requirements and are billed only for the services actually used.

IaaS cloud computing platform layer eliminates the need for every organization to maintain the IT infrastructure.

IaaS is offered in three models: public, private, and hybrid cloud:

The private cloud implies that the infrastructure resides at the customer-premise.

In the case of public cloud, it is located at the cloud computing platform vendor's data center.

The hybrid cloud is a combination of the two in which the customer selects the best of both public cloud or private cloud.

IaaS provider provides the following services -

Compute: Computing as a Service includes virtual central processing units and virtual main memory for the Vms that is provisioned to the end- users.
Storage: IaaS provider provides back-end storage for storing files.
Network: Network as a Service (NaaS) provides networking components such as routers, switches, and bridges for the Vms.
Load balancers: It provides load balancing capability at the infrastructure layer.

Advantages of IaaS cloud computing layer:

There are the following advantages of IaaS computing layer -

Shared infrastructure:

IaaS allows multiple users to share the same physical infrastructure.

Web access to the resources:

Iaas allows IT users to access resources over the internet.

Pay-as-per-use model:

IaaS providers provide services based on the pay-as-per-use basis. The users are required to pay for what they have used.

Focus on the core business:

IaaS providers focus on the organization's core business rather than on IT infrastructure.

On-demand scalability:

On-demand scalability is one of the biggest advantages of IaaS. Using IaaS, users do not worry about to upgrade software and troubleshoot the issues related to hardware components.

Disadvantages of IaaS cloud computing layer:

Security:

Security is one of the biggest issues in IaaS. Most of the IaaS providers are not able to provide 100% security.

Maintenance & Upgrade:

Although IaaS service providers maintain the software, but they do not upgrade the software for some organizations.

Interoperability issues:

It is difficult to migrate VM from one IaaS provider to the other, so the customers might face problem related to vendor lock-in.

Some important point about IaaS cloud computing layer:

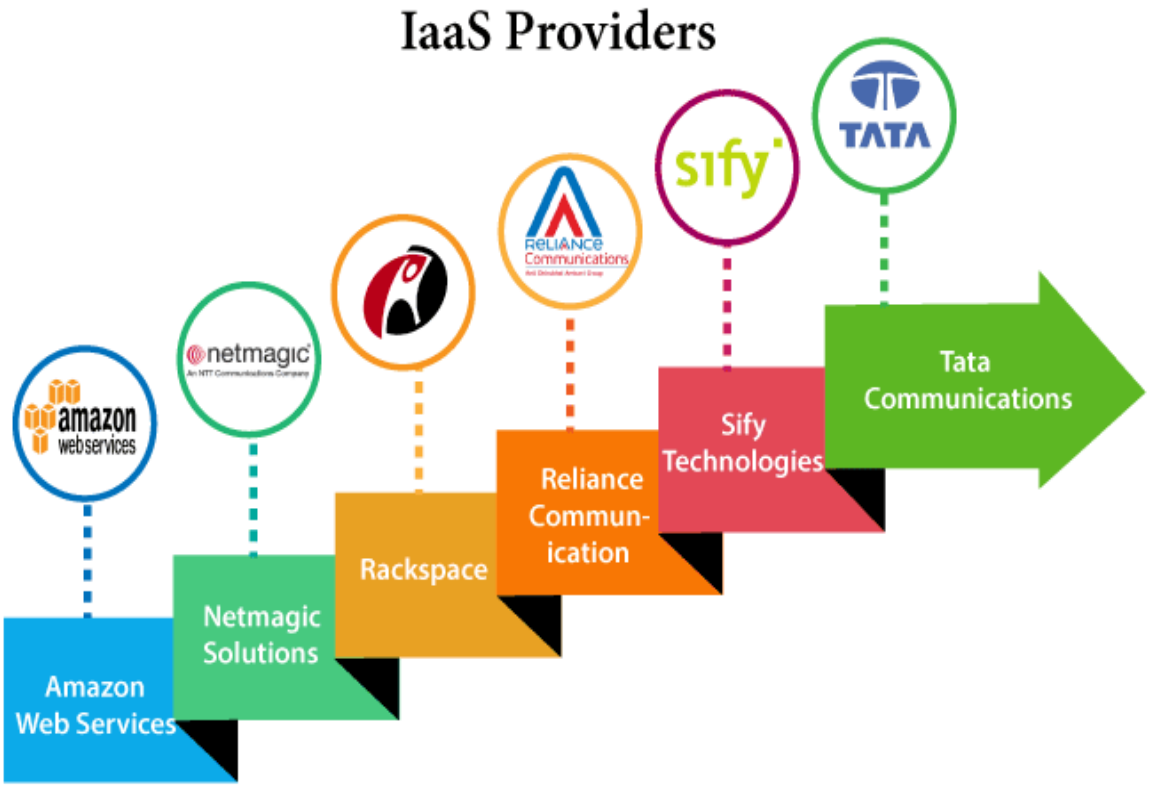
IaaS cloud computing platform cannot replace the traditional hosting method, but it provides more than that, and each resource which are used are predictable as per the usage.

IaaS cloud computing platform may not eliminate the need for an in-house IT department. It will be needed to monitor or control the IaaS setup. IT salary expenditure might not reduce significantly, but other IT expenses can be reduced.

Breakdowns at the IaaS cloud computing platform vendor's can bring your business to the halt stage. Assess the IaaS cloud computing platform vendor's stability and finances. Make sure that SLAs (i.e., Service Level Agreement) provide backups for data, hardware, network, and application failures. Image portability and third-party support is a plus point.

The IaaS cloud computing platform vendor can get access to your sensitive data. So, engage with credible companies or organizations. Study their security policies and precautions.

Top Iaas Providers who are providing IaaS cloud computing platform:



Platform as a Service (PaaS)

Platform as a Service (PaaS) provides a runtime environment. It allows programmers to easily create, test, run, and deploy web applications. You can purchase these applications from a cloud service provider on a pay-as-per use basis and access them using the Internet connection. In PaaS, back end scalability is managed by the cloud service provider, so end-users do not need to worry about managing the infrastructure.

PaaS includes infrastructure (servers, storage, and networking) and platform (middleware, development tools, database management systems, business intelligence, and more) to support the web application life cycle.

Example: Google App Engine, Force.com, Azure.

PaaS providers provide the Programming languages, Application frameworks, Databases, and Other tools:

Programming languages:

PaaS providers provide various programming languages for the developers to develop the applications. Some popular programming languages provided by PaaS providers are Java, PHP, Perl, etc.

Application frameworks:

PaaS providers provide application frameworks to easily understand the application development. Some popular application frameworks provided by PaaS providers are Node.js, Joomla, WordPress, Spring, etc.

Databases:

PaaS providers provide various databases such as ClearDB, PostgreSQL, MongoDB, and Redis to communicate with the applications.

Other tools:

PaaS providers provide various other tools that are required to develop, test, and deploy the applications.

Advantages of PaaS:

There are the following advantages of PaaS -

Simplified Development:

PaaS allows developers to focus on development and innovation without worrying about infrastructure management.

Lower risk:

No need for up-front investment in hardware and software. Developers only need a PC and an internet connection to start building applications.

Prebuilt business functionality:

Some PaaS vendors also provide already defined business functionality so that users can avoid building everything from very scratch and hence can directly start the projects only.

Instant community:

PaaS vendors frequently provide online communities where the developer can get the ideas to share experiences and seek advice from others.

Scalability:

Applications deployed can scale from one to thousands of users without any changes to the applications.

Disadvantages of PaaS cloud computing layer:

Vendor lock-in:

One has to write the applications according to the platform provided by the PaaS vendor, so the migration of an application to another PaaS vendor would be a problem.

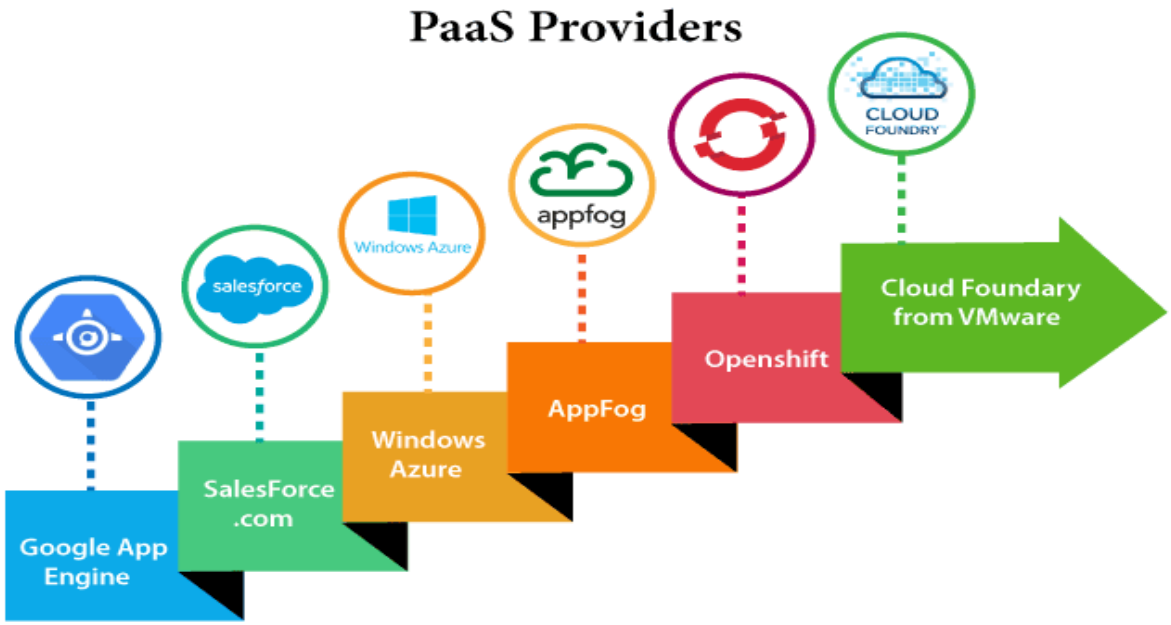
Data Privacy:

Corporate data, whether it can be critical or not, will be private, so if it is not located within the walls of the company, there can be a risk in terms of privacy of data.

Integration with the rest of the systems applications:

It may happen that some applications are local, and some are in the cloud. So there will be chances of increased complexity when we want to use data which in the cloud with the local data.

Popular PaaS Providers:



The below table shows some popular PaaS providers and services that are provided by them :-

Providers	Services
Google App Engine (GAE)	App Identity, URL Fetch, Cloud storage client library, Logservice
Salesforce.com	Faster implementation, Rapid scalability, CRM Services, Sales cloud, Mobile connectivity, Chatter.
Windows Azure	Compute, security, IoT, Data Storage.
AppFog	Justcloud.com, SkyDrive, GoogleDocs
Openshift	RedHat, Microsoft Azure.
Cloud Foundry from VMware	Data, Messaging, and other services.

Software as a Service (SaaS)

SaaS is also known as "**On-Demand Software**". It is a software distribution model in which services are hosted by a cloud service provider. These services are available to end-users over the internet so, the end-users do not need to install any software on their devices to access these services.

There are following services provided by SaaS providers -

Business Services - SaaS Provider provides various business services to start-up the business. The SaaS business services include **ERP** (Enterprise Resource Planning), **CRM** (Customer Relationship Management), **billing**, and **sales**.

Document Management - SaaS document management is a software application offered by a third party (SaaS providers) to create, manage, and track electronic documents.

Example: Slack.

Social Networks - As we all know, social networking sites are used by the general public, so social networking service providers use SaaS for their convenience and handle the general public's information.

Mail Services - To handle the unpredictable number of users and load on e-mail services, many e-mail providers offering their services using SaaS.

Advantages of SaaS cloud computing layer:

SaaS is easy to buy:

SaaS pricing is based on a monthly fee or annual fee subscription, so it allows organizations to access business functionality at a low cost, which is less than licensed applications.

Unlike traditional software, which is sold as a licensed based with an up-front cost (and often an optional ongoing support fee), SaaS providers are generally pricing the applications using a subscription fee, most commonly a monthly or annually fee.

One to Many:

SaaS services are offered as a one-to-many model means a single instance of the application is shared by multiple users.

Less hardware required for SaaS:

The software is hosted remotely, so organizations do not need to invest in additional hardware.

Low maintenance required for SaaS:

Software as a service removes the need for installation, set-up, and daily maintenance for the organizations. The initial set-up cost for SaaS is typically less than the enterprise software. SaaS vendors are pricing their applications based on some usage parameters, such as a number of users using the application. So SaaS does easy to monitor and automatic updates.

No special software or hardware versions required:

All users will have the same version of the software and typically access it through the web browser. SaaS reduces IT support costs by outsourcing hardware and software maintenance and support to the IaaS provider.

Multidevice support:

SaaS services can be accessed from any device such as desktops, laptops, tablets, phones, and thin clients.

API Integration:

SaaS services easily integrate with other software or services through standard APIs.

No client-side installation:

SaaS services are accessed directly from the service provider using the internet connection, so do not need to require any software installation.

Disadvantages of SaaS cloud computing layer:

Security:

Actually, data is stored in the cloud, so security may be an issue for some users. However, cloud computing is not more secure than in-house deployment.

Latency issue:

Since data and applications are stored in the cloud at a variable distance from the end-user, there is a possibility that there may be greater latency when interacting with the application compared to local deployment. Therefore, the SaaS model is not suitable for applications whose demand response time is in milliseconds.

Total Dependency on Internet:

Without an internet connection, most SaaS applications are not usable.

Switching between SaaS vendors is difficult:

Switching SaaS vendors involves the difficult and slow task of transferring the very large data files over the internet and then converting and importing them into another SaaS also.

Popular SaaS Providers:

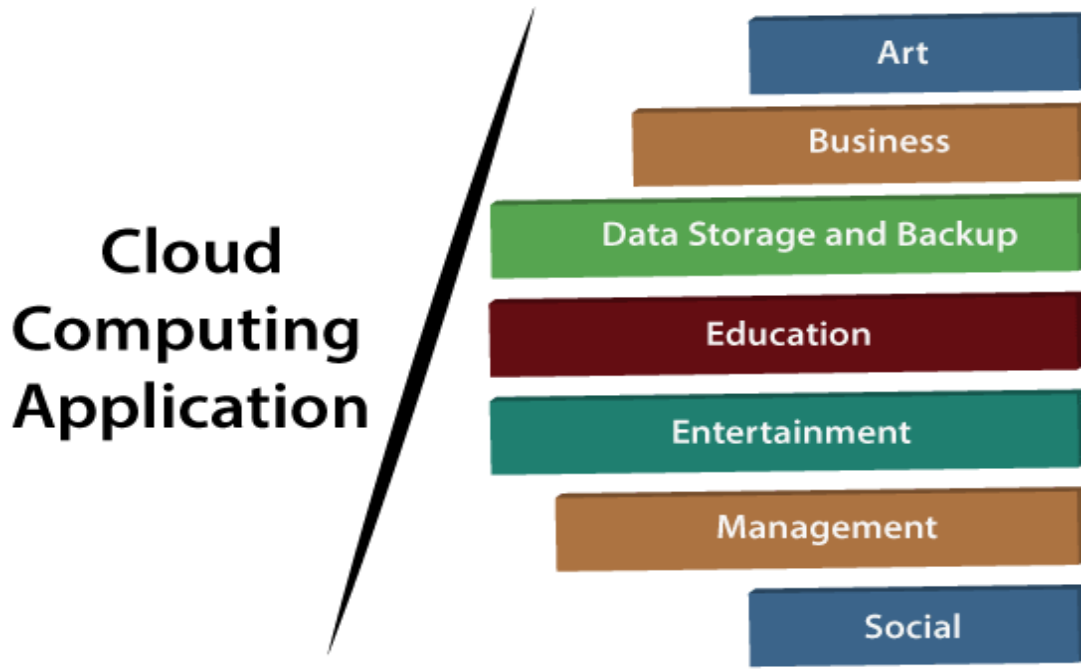
The below table shows some popular SaaS providers and services that are provided by them -

Provider	Services
Salseforce.com	On-demand CRM solutions
Microsoft Office 365	Online office suite
Google Apps	Gmail, Google Calendar, Docs, and sites
NetSuite	ERP, accounting, order management, CRM, Professionals Services Automation (PSA), and e-commerce applications.
GoToMeeting	Online meeting and video-conferencing software
Constant Contact	E-mail marketing, online survey, and event marketing
Oracle CRM	CRM applications
Workday, Inc	Human capital management, payroll, and financial management.

Cloud Computing Applications

Cloud service providers provide various applications in the field of art, business, data storage and backup services, education, entertainment, management, social networking, etc.

The most widely used cloud computing applications are given below -



Art Applications:

Cloud computing offers various art applications for quickly and easily design **attractive cards, booklets,** and **images**. Some most commonly used cloud art applications are given below:

i Moo

Moo is one of the best cloud art applications. It is used for designing and printing business cards, postcards, and mini cards.

ii. Vistaprint

Vistaprint allows us to easily design various printed marketing products such as business cards, Postcards, Booklets, and wedding invitations cards.

iii. Adobe Creative Cloud

Adobe creative cloud is made for designers, artists, filmmakers, and other creative professionals. It is a suite of apps which includes PhotoShop image editing programming, Illustrator, InDesign, TypeKit, Dreamweaver, XD, and Audition.

Business Applications:

Business applications are based on cloud service providers. Today, every organization requires the cloud business application to grow their business. It also ensures that business applications are 24*7 available to users.

There are the following business applications of cloud computing -

i. MailChimp

MailChimp is an **email publishing platform** which provides various options to **design, send,** and **save** templates for emails.

ii. Salesforce

Salesforce platform provides tools for sales, service, marketing, e-commerce, and more. It also provides a cloud development platform.

iii. Chatter

Chatter helps us to **share important information** about the organization in real time.

iv. Paypal

Paypal offers the simplest and easiest **online payment** mode using a secure internet account. Paypal accepts the payment through debit cards, credit cards, and also from Paypal account holders.

v. Slack

Slack stands for **Searchable Log of all Conversation and Knowledge**. It provides a **user-friendly** interface that helps us to create public and private channels for communication.

vi. Quickbooks

Quickbooks works on the terminology "**Run Enterprise anytime, anywhere, on any device**." It provides online accounting solutions for the business. It allows more than 20 users to work simultaneously on the same system.

Data Storage and Backup Applications:

Cloud computing allows us to store information (data, files, images, audios, and videos) on the cloud and access this information using an internet connection. As the cloud provider is responsible for providing security, so they offer various backup recovery application for retrieving the lost data.

A list of data storage and backup applications in the cloud are given below -

i. Box.com

Box provides an online environment for **secure content management, workflow,** and **collaboration**. It allows us to store different files such as Excel, Word, PDF, and images on the cloud. The main advantage of using box is that it provides drag & drop service for files and easily integrates with Office 365, G Suite, Salesforce, and more than 1400 tools.

ii. Mozy

Mozy provides powerful **online backup solutions** for our personal and business data. It schedules automatically back up for each day at a specific time.

iii. Google G Suite

Google G Suite is one of the best **cloud storage** and **backup** application. It includes Google Calendar, Docs, Forms, Google+, Hangouts, as well as cloud storage and tools for managing cloud apps. The most popular app in the Google G Suite is Gmail. Gmail offers free email services to users.

Education Applications:

Cloud computing in the education sector becomes very popular. It offers various **online distance learning platforms** and **student information portals** to the students. The advantage of using cloud in the field of education is that it offers strong virtual classroom environments, Ease of accessibility, secure data storage, scalability, greater reach for the students, and minimal hardware requirements for the applications.

There are the following education applications offered by the cloud -

i. Google Apps for Education

Google Apps for Education is the most widely used platform for free web-based email, calendar, documents, and collaborative study.

ii. Chromebooks for Education

Chromebook for Education is one of the most important Google's projects. It is designed for the purpose that it enhances education innovation.

iii. Tablets with Google Play for Education

It allows educators to quickly implement the latest technology solutions into the classroom and make it available to their students.

iv. AWS in Education

AWS cloud provides an education-friendly environment to universities, community colleges, and schools.

Entertainment Applications:

Entertainment industries use a **multi-cloud strategy** to interact with the target audience. Cloud computing offers various entertainment applications such as online games and video conferencing.

i. Online games

Today, cloud gaming becomes one of the most important entertainment media. It offers various online games that run remotely from the cloud. The best cloud gaming services are Shaow, GeForce Now, Vortex, Project xCloud, and PlayStation Now.

ii. Video Conferencing Apps

Video conferencing apps provides a simple and instant connected experience. It allows us to communicate with our business partners, friends, and relatives using a

cloud-based video conferencing. The benefits of using video conferencing are that it reduces cost, increases efficiency, and removes interoperability.

Management Applications

Cloud computing offers various cloud management tools which help admins to manage all types of cloud activities, such as resource deployment, data integration, and disaster recovery. These management tools also provide administrative control over the platforms, applications, and infrastructure.

Some important management applications are -

i. Toggl

Toggl helps users to track allocated time period for a particular project.

ii. Evernote

Evernote allows you to sync and save your recorded notes, typed notes, and other notes in one convenient place. It is available for both free as well as a paid version.

It uses platforms like Windows, macOS, Android, iOS, Browser, and Unix.

iii. Outright

Outright is used by management users for the purpose of accounts. It helps to track income, expenses, profits, and losses in real-time environment.

iv. GoToMeeting

GoToMeeting provides **Video Conferencing** and **online meeting apps**, which allows you to start a meeting with your business partners from anytime, anywhere using mobile phones or tablets. Using GoToMeeting app, you can perform the tasks related to the management such as join meetings in seconds, view presentations on the shared screen, get alerts for upcoming meetings, etc.

Social Applications:

Social cloud applications allow a large number of users to connect with each other using social networking applications such as **Facebook, Twitter, Linkedln,** etc.

There are the following cloud based social applications -

i. Facebook

Facebook is a **social networking website** which allows active users to share files, photos, videos, status, more to their friends, relatives, and business partners using

the cloud storage system. On Facebook, we will always get notifications when our friends like and comment on the posts.

ii. Twitter

Twitter is a **social networking** site. It is a **microblogging** system. It allows users to follow high profile celebrities, friends, relatives, and receive news. It sends and receives short posts called tweets.

iii. LinkedIn

LinkedIn is a **social network** for students, freshers, and professionals.

Service Oriented Architecture

SOA (**Service Oriented Architecture**) is built on computer engineering approaches that offer an architectural advancement towards enterprise system. **SOA** provides a translation and management layer within the **cloud architecture** that removes the barrier for **cloud** clients obtaining desired services.

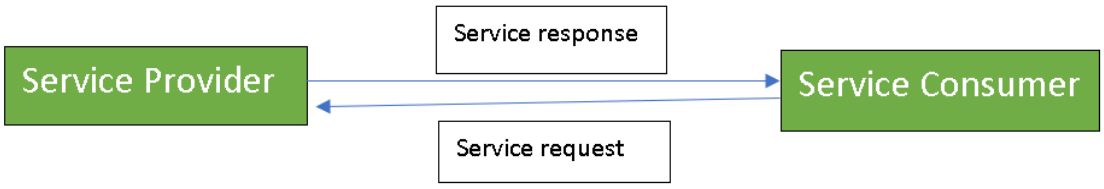
Service-Oriented Architecture (SOA) is an architectural approach in which applications make use of services available in the network. In this architecture, services are provided to form applications, through a communication call over the internet.

- SOA allows users to combine a large number of facilities from existing services to form applications.
- SOA encompasses a set of design principles that structure system development and provide means for integrating components into a coherent and decentralized system.
- SOA based computing packages functionalities into a set of interoperable services, which can be integrated into different software systems belonging to separate business domains.

There are two major roles within Service-oriented Architecture:

1. **Service provider:** The service provider is the maintainer of the service and the organization that makes available one or more services for others to use. To advertise services, the provider can publish them in a registry, together with a service contract that specifies the nature of the service, how to use it, the requirements for the service, and the fees charged.

2. **Service consumer:** The service consumer can locate the service metadata in the registry and develop the required client components to bind and use the service.



Guiding Principles of SOA:

1. **Standardized service contract:** Specified through one or more service description documents.
2. **Loose coupling:** Services are designed as self-contained components; maintain relationships that minimize dependencies on other services.

3. **Abstraction:** A service is completely defined by service contracts and description documents. They hide their logic, which is encapsulated within their implementation.
4. **Reusability:** Designed as components, services can be reused more effectively, thus reducing development time and the associated costs.
5. **Autonomy:** Services have control over the logic they encapsulate and, from a service consumer point of view, there is no need to know about their implementation.
6. **Discoverability:** Services are defined by description documents that constitute supplemental metadata through which they can be effectively discovered. Service discovery provides an effective means for utilizing third-party resources.
7. **Composability:** Using services as building blocks, sophisticated and complex operations can be implemented.

Advantages of SOA:

- **Service reusability:** In SOA, applications are made from existing services. Thus, services can be reused to make many applications.
- **Easy maintenance:** As services are independent of each other they can be updated and modified easily without affecting other services.
- **Platform independent:** SOA allows making a complex application by combining services picked from different sources, independent of the platform.
- **Availability:** SOA facilities are easily available to anyone on request.
- **Reliability:** SOA applications are more reliable because it is easy to debug small services rather than huge codes
- **Scalability:** Services can run on different servers within an environment, this increases scalability

Disadvantages of SOA:

- **High overhead:** A validation of input parameters of services is done whenever services interact this decreases performance as it increases load and response time.
- **High investment:** A huge initial investment is required for SOA.
- **Complex service management:** When services interact they exchange messages to tasks. The number of messages may go in millions. It becomes a cumbersome task to handle a large number of messages.

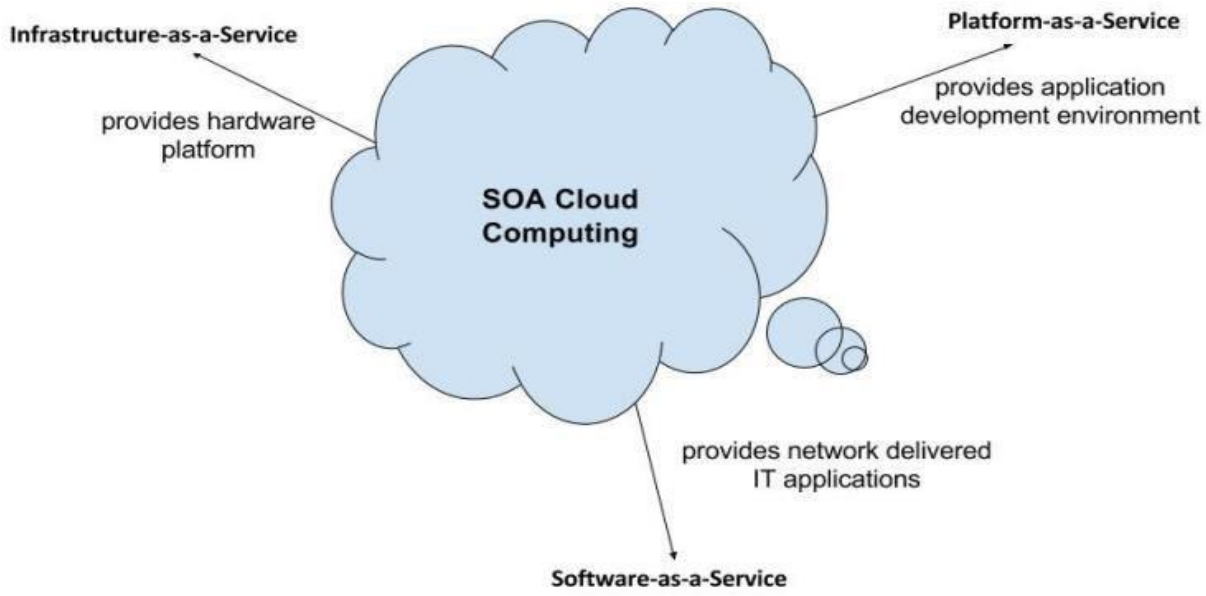
Concepts in Cloud Computing

Cloud computing is a model used for enabling convenient and usage-based network access to a configurable computing resources (eg. networks, servers etc) that can be provided and used rapidly.

- It provides a chance to business users to implement services with usage-based billing that is changed according to their requirements without need of consulting with IT department.

- It provides an abstraction layer between computing resources and its technical implementation details and sequentially enables computational resources to be used while avoiding efforts in infrastructure management.

The below figure shows the SOA cloud computing along with the models:



Below are the models that are differentiated on the horizontal scaling basis in cloud computing:

- **Infrastructure-as-a-Service (IaaS)**: It provides a hardware platform as a service.
- **Platform-as-a-Service (PaaS)**: It provides end-users an application development environment delivered over the internet.
- **Software-as-a-Service (SaaS)**: It provides end-users a standardized, network-delivered IT applications.

The distinctions are made according to availability and the location of installation in the deployment models. Private clouds are internal company services whereas public clouds are the services that are available to the public on internet.

In the large companies where IT plays an important role, internal company cloud solutions are often built in their own data centers. Small and medium companies often use public cloud services.

Cloud Computing provides a very flexible and scalable platform through processing external services and also has the ability to connect with customers, suppliers etc.

Diversified Services

Diversified is a full-service information technology service provider delivering data. deploying a data centre on virtualization and cloud computing technologies.

Virtualization:

Data centre virtualization is the process of designing, developing and deploying a data centre on virtualization and cloud computing technologies. It primarily enables virtualizing physical servers in a data centre facility along with storage, networking and other infrastructure devices and equipment. Diversified's expertise in virtualization helps data centres achieve the following benefits:

- Reduce hardware vendor lock-in
- Improve disaster recovery
- Smooth migration to cloud
- Reduce data center footprint
- Faster server provisioning

Premise, Hybrid & Cloud Computing:

The overwhelming number of vendors, products and hybrid options make it challenging to develop your future IT roadmap. Diversified aides clients in navigating the pros and cons of hosting locally, in a private/public cloud, or a combination thereof to create a strategy that works.

Storage:

A proper strategy for storage is essential for quality and cost efficiency. Diversified assists clients in selecting the proper storage to fit their needs and budget.

- Design for migrating inheritance storage solutions from drives, tapes, and arrays.
- Deployment of complete storage solutions consisting of internal and external storage, backup and storage software, storage networking, hyperconverged, cloud, and policies.

Hyperconverged Infrastructure:

The drive to reduce complexity while increasing scalability is also becoming increasingly popular in the data centre. Diversified's hyperconverged infrastructure solutions provide a software-centric architecture that tightly integrates computing, storage and virtualization resources in a single system that runs on off the shelf servers.

Performance Issues in Cloud Computing Services

Application that fits the cloud:

Not all the applications are suitable for cloud. It is extremely important to identify the most suitable applications for migration and identify any potential problems. Create a checklist to ensure a complete and successful migration.

Dealing with the performance issues:

If you are managing application performance in the cloud, you need a topological map of service delivery across all tiers. Although cloud computing offers numerous benefits, performance issues can complicate or reduce the benefits.

Addressing the topological dependencies:

While moving to cloud, various businesses need to face the impact of moving from a primarily static to dynamic network architecture. Firewall, load balancing and security services are still required for the network architecture.

Monitoring consumption for every service:

While transitioning from a resource focused cost center to a business service focused profit center, it demands assessing the resource consumption. Unfortunately, the traditional chargeback and AMP (Advanced Malware Protection) tools lack the ability to enable the business aligned costing and chargeback paradigms. This means that you need to come up with a solution to monitor consumption for every service across multiple applications and tiers.

Have a clear picture of resource consumption:

In order to make it sure that SLA's in cloud are met, you need to prioritize the allocation of resources based on the measurement of the end user performance. It demands a clear picture of the resource consumption at the transaction level.

Lack of infrastructure configuration for service deployment:

The lack of knowledge and infrastructure configuration for the service deployment has limited the ability of researchers to study the impact of resource management inside the cloud infrastructures on the service performance through measurement-based evaluations. This makes it difficult for a service customer to use a measurement-based method to get insight about the performance behaviours of new Cloud services.

Depending upon the size and the type of business you are running, the cloud offers you various benefits. For a startup, cloud computing provides essential differentiators to keep the business up and running quickly with minimal up-front costs. Larger businesses often

face complex challenges to ensure the availability and performance of the high traffic websites.

You will face various risks along the way along that can be complex to manage. Transitioning to the cloud is a non-trivial decision that demands a proper evaluation of both the data and services. It is important to perform a thorough evaluation on the cloud service performance as this is what will be beneficial for both service providers as well as customers.

Overall, cloud performance issues are considerable and demands innovation to overcome the challenges.

Data Centre

Data centres are simply centralized locations where computing and networking equipment is concentrated for the purpose of collecting, storing, processing, distributing or allowing access to large amounts of data. They have existed in one form or another since the introduction of computers.

Data centre refers to on-premise hardware while the cloud refers to off-premise computing. The cloud stores your data in the public cloud, while a data centre stores your data on your own hardware.

They are also responsible for data backup and recovery, as well as networking. These centres also host websites, manage e-mails, and instant messaging services. They support cloud storage applications and e-commerce transactions. Even online gaming communities require data centres to manage their online activities.

Data centres connect communication networks so end-users can access information remotely. These vast numbers of clustered servers and related equipment can be found in a room or even in an entire compound.

Evolution of Data Centres:

Decades ago, early computers were massive machines that can occupy whole rooms. But as technology evolved, equipment shrunk and became cheaper than before. However, with this progress, data processing demands have also begun to increase exponentially.

Unlike before, where data centres are just one big supercomputer, modern aged data centres functions using multiple servers to optimize further and boost its processing power. Now data centres consist of thousands of potent and tiny servers that run non-stop around the clock.

Importance of Data Centres:

Almost every modern business and government offices need their very own data centre, or they may decide to lease data centr. Big corporations and government institutions may choose to build and manage them in-house if they have the resources. While others choose to rent servers at 'colos' or colocation facilities. Some business owners also have the choice to use public cloud-based services.

Corporations that handle education, finance, telecommunication, retailers, and social networking services process a lot of information every day. This business that produces and utilizes data requires data centres in running their operations. Without these centres, they will suffer the absence of speedy and secure access to data. This failure in delivering services will ultimately lead to the loss of clients and profits.

Now, we must remember that all of this information needs to be housed somewhere. The idea of running or storing our data and resources at home or work computers is getting replaced by faraway storing mentality. Many firms are also migrating their professional applications to data centre services to minimize the cost of running an in-house server.

That is why data centres are an essential resource for a business that wants to run their operation without worries. The importance of data centres in the modern world has increased ten times due to the rising demand of information trading.

How Do Data Centres Work?

The data that is stored on a data centre server is distributed into packets before transmission and is sent via routers that decide the most suitable path for that data to progress.

It then uses a series of wired and wireless networks to reach the user's Internet service provider and finally arrive at the end user's computer. Every time a Web address is enters into a browser, it automatically requests information from a server. If the end-user wants to upload information, then the process will be reversed.

Types of Data Centres:

With how data centres are essential in running big corporations and with even small-medium enterprises joining the trend, choosing one to fit a business model is essential. There are different types of data centres and service models.

Here are four main types of data centres:

Colocation Data Centers:

Colocation data centers or most commonly known as "colo" is a company that rents space within a data centre that they do not own and is housed outside the company's premises. The colocation data centre provides the infrastructure like the building itself, cooling,

bandwidth, and security, among others. While the company produces and maintains the components, which include the servers, storage system, and security firewalls.

Enterprise Data Centres:

Enterprise data centres are established, owned, and managed by companies. These data canters are operated under a single purpose and that this optimized service for their end-user clients. Enterprise data canters are often located inside corporate compounds.

Managed Services Data Canters:

These data canters are operated by a third-party entity or a managed services provider instead of the company. The company rents the equipment and infrastructure to cut costs.

Cloud Data Canters:

Cloud data canters are an off-premises form of a data centre. The most common cloud hosting services are Amazon Web Services (AWS), Microsoft (Azure), and IBM Cloud.

How Reliable Is a Data Centre Facility?

Business owners are in constant need of reliability when in terms of maintaining a smooth operation. Good thing that a Data centre is built to withstand a 24/7 service easily. However, the components require a significant amount of infrastructure support in both hardware and software areas.

These include power subsystems, stable and uninterruptible power supplies, proper ventilation, high-quality cooling systems, fire control, reliable backup generators, and connections to external networks.

The business world is moving at an incredibly fast pace that matches the overwhelming demand for information. With the ever-changing requirements of the modern business model, many companies place their confidence in data canters, as these facilities play a crucial role in reaching their IT specifications.

Data centre service providers are capable of handling higher volumes of traffic without making compromises on security and storage capacity of data. Generally, a typical data centre carries the responsibility of managing significant characteristics like data workloads, operating conditions, data protection and security fulfilment.

Data canters are more than just a safe and secure facility with space that is equipped with reliable power, and network. They are becoming a valuable addition to many businesses as they prove to be a dependable extension of their IT team. That is why data canters in modern business set up are increasingly becoming an essential factor for success.

Legal issues in cloud computing service provision

Cloud computing, being one of such recent advancements, have raised a number of legal issues including privacy and data security, contracting issues, issues relating to the location of the data, and business considerations. Issues relating to contractual relation between the cloud service provider and the customer.

Cloud computing is bringing amazing advantages and benefits companies. But it also brings some challenges. There are several legal issues that must be taken into consideration when moving into the cloud. Let's see which are the most challenging legal issues around cloud computing.

Security procedures:

The majority of companies which implemented cloud solutions and services do not have security procedures in place. Also, they lack measures to approve or evaluate cloud applications. When adopting the BYOD trend for example, organizations needed these security procedures more than ever. General data security trainings, multiple levels of security, rigorous procedures to use one's own device and to transfer or copy data are some of the options available to protect data in organizations. The bottom line is that security procedures must be established according to every company's objectives and work flow.

[Bring your own device (**BYOD**) refers to the **trend** of employees using personal devices to connect to their organizational networks and access work-related systems and potentially sensitive or confidential data. Personal devices could include smartphones, personal computers, tablets, or USB drives.]

Third party access issues:

Third-party involvement could be a risk. All third parties using a multi-tenant shared cloud are using the same administration interface, so make sure multi-factor authentication and enhanced security is present. Also, look for HIPAA (Health Insurance Portability and Accountability Act) compliant providers – a business associate agreement (BAA) with third-party vendor who access Protected Health Information (PHI) is necessary to ensure privacy and security requirements. A partnership with a HIPAA solutions provider that signs a BAA is an efficient method to make sure this goes smoothly and everything is secure. And don't forget to read carefully the terms and conditions before signing up for a cloud based services.

Intellectual Property Rights:

Intellectual Property Rights differ from one country to another, so it is not very clear what intellectual property laws will apply in the cloud computing environment. Make sure you are aware of the regulations and rights from the country you store your intellectual work. The provider you choose should know how to protect intellectual property it stores and how to avoid potential infringement pitfalls.

Confidential data theft attacks:

Data stored in the cloud might be compromised or breached. Therefore, most cloud computing providers also offer the customer different levels of security protection, which allows for more enhanced security. Encryption might seem to have failed in protecting data from theft attacks, but other methods have been discovered and implemented, including monitoring data access in the cloud to detect abnormal data access patterns. The customer has to understand the cloud provider's disclosure policy and how quickly the breach would be disclosed to them. Most of the U.S. states have security breach disclosure laws requiring the provider to inform the customers when their data has been compromised.

Many of these legal issues and the methods to inform about them or to solve them should be mentioned in the Service Level Agreement. It is essential to understand all the terms of the cloud's provider and to consider the needs and objectives of the enterprise before signing an agreement.

What is Big Data? Introduction, Types, Characteristics, Examples

What is Data?

The quantities, characters, or symbols on which operations are performed by a computer, which may be stored and transmitted in the form of electrical signals and recorded on magnetic, optical, or mechanical recording media.

What is Big Data?

Big Data is a collection of data that is huge in volume, yet growing exponentially with time. It is a data with so large size and complexity that none of traditional data management tools can store it or process it efficiently. Big data is also a data but with huge size.

What is an Example of Big Data?

Following are some of the Big Data examples-

The **New York Stock Exchange** is an example of Big Data that generates about **one terabyte** of new trade data per day.

Social Media

The statistic shows that **500+terabytes** of new data get ingested into the databases of social media site **Facebook**, every day. This data is mainly generated in terms of photo and video uploads, message exchanges, putting comments etc.

A single **Jet engine** can generate **10+terabytes** of data in **30 minutes** of flight time. With many thousand flights per day, generation of data reaches up to many **Petabytes.**

Types Of Big Data

Following are the types of Big Data:

1. **Structured**
2. **Unstructured**
3. **Semi-structured**

Structured

Any data that can be stored, accessed and processed in the form of fixed format is termed as a 'structured' data. Over the period of time, talent in computer science has achieved greater success in developing techniques for working with such kind of data (where the format is well known in advance) and also deriving value out of it. However, nowadays, we are foreseeing issues when a size of such data grows to a huge extent, typical sizes are being in the rage of multiple zettabytes.

Do you know? 10^{21} bytes equal to **1 zettabyte** or **one billion terabytes** forms **a zettabyte**.

Looking at these figures one can easily understand why the name Big Data is given and imagine the challenges involved in its storage and processing.

Do you know? Data stored in a relational database management system is one example of a **'structured'** data.

Examples Of Structured Data

An ‘Employee’ table in a database is an example of Structured Data

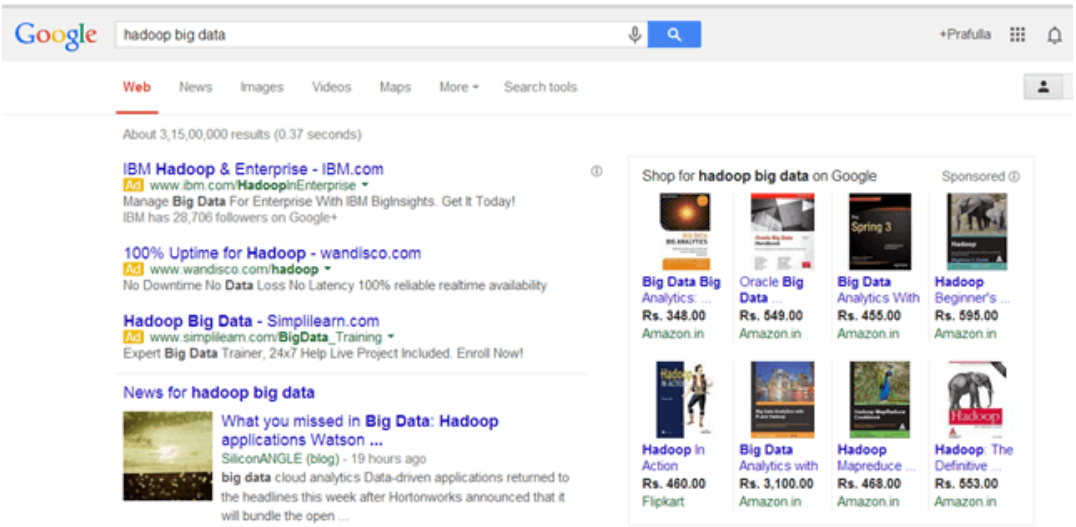
Employee_ID	Employee_Name	Gender	Department	Salary_In_lacs
2365	Rajesh Kulkarni	Male	Finance	650000
3398	Pratibha Joshi	Female	Admin	650000
7465	Shushil Roy	Male	Admin	500000
7500	Shubhojit Das	Male	Finance	500000
7699	Priya Sane	Female	Finance	550000

Unstructured

Any data with unknown form or the structure is classified as unstructured data. In addition to the size being huge, un-structured data poses multiple challenges in terms of its processing for deriving value out of it. A typical example of unstructured data is a heterogeneous data source containing a combination of simple text files, images, videos etc. Now day organizations have wealth of data available with them but unfortunately, they don’t know how to derive value out of it since this data is in its raw form or unstructured format.

Examples Of Un-structured Data

The output returned by ‘Google Search’



Example Of Un-structured Data

Semi-structured

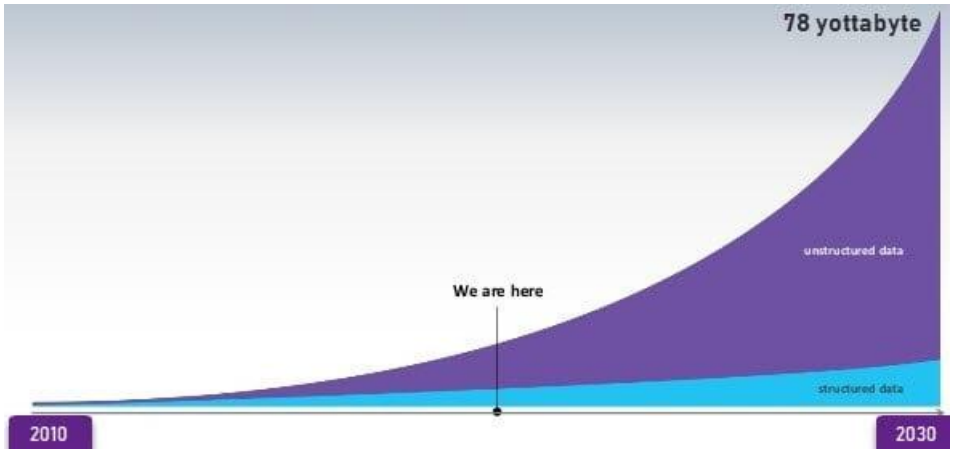
Semi-structured data can contain both the forms of data. We can see semi-structured data as a structured in form but it is actually not defined with e.g. a table definition in relational DBMS. Example of semi-structured data is a data represented in an XML file.

Examples Of Semi-structured Data

Personal data stored in an XML file-

```
<rec><name>Prashant Rao</name><sex>Male</sex><age>35</age></rec>
<rec><name>Seema R.</name><sex>Female</sex><age>41</age></rec>
<rec><name>Satish Mane</name><sex>Male</sex><age>29</age></rec>
<rec><name>Subrato Roy</name><sex>Male</sex><age>26</age></rec>
<rec><name>Jeremiah J.</name><sex>Male</sex><age>35</age></rec>
```

Data Growth over the years



Data Growth over the years

Please note that web application data, which is unstructured, consists of log files, transaction history files etc. OLTP systems are built to work with structured data wherein data is stored in relations (tables).

Characteristics Of Big Data

Big data can be described by the following characteristics:

- Volume
- Variety
- Velocity

- Variability

(i) Volume – The name Big Data itself is related to a size which is enormous. Size of data plays a very crucial role in determining value out of data. Also, whether a particular data can actually be considered as a Big Data or not, is dependent upon the volume of data.
Hence, **'Volume'** is one characteristic which needs to be considered while dealing with Big Data solutions.

(ii) Variety – The next aspect of Big Data is its **variety**.

Variety refers to heterogeneous sources and the nature of data, both structured and unstructured. During earlier days, spreadsheets and databases were the only sources of data considered by most of the applications. Nowadays, data in the form of emails, photos, videos, monitoring devices, PDFs, audio, etc. are also being considered in the analysis applications. This variety of unstructured data poses certain issues for storage, mining and analyzing data.

(iii) Velocity – The term **'velocity'** refers to the speed of generation of data. How fast the data is generated and processed to meet the demands, determines real potential in the data.

Big Data Velocity deals with the speed at which data flows in from sources like business processes, application logs, networks, and social media sites, sensors, Mobile devices, etc. The flow of data is massive and continuous.

(iv) Variability – This refers to the inconsistency which can be shown by the data at times, thus hampering the process of being able to handle and manage the data effectively.

Advantages Of Big Data Processing

Ability to process Big Data in DBMS brings in multiple benefits, such as-

- Businesses can utilize outside intelligence while taking decisions

Access to social data from search engines and sites like Facebook, Twitter are enabling organizations to fine tune their business strategies.

- Improved customer service

Traditional customer feedback systems are getting replaced by new systems designed with Big Data technologies. In these new systems, Big Data and natural language processing technologies are being used to read and evaluate consumer responses.

- Early identification of risk to the product/services, if any
- Better operational efficiency

Big Data technologies can be used for creating a staging area or landing zone for new data before identifying what data should be moved to the data warehouse. In addition, such integration of Big Data technologies and data warehouse helps an organization to offload infrequently accessed data.

Summary

- Big Data definition : Big Data meaning a data that is huge in size. Bigdata is a term used to describe a collection of data that is huge in size and yet growing exponentially with time.
- Big Data analytics examples includes stock exchanges, social media sites, jet engines, etc.
- Big Data could be 1) Structured, 2) Unstructured, 3) Semi-structured
- Volume, Variety, Velocity, and Variability are few Big Data characteristics
- Improved customer service, better operational efficiency, Better Decision Making are few advantages of Bigdata

Difference between Traditional data and Big data

1. Traditional data: Traditional data is the structured data that is being majorly maintained by all types of businesses starting from very small to big organizations. In a traditional database system, a centralized database architecture used to store and maintain the data in a fixed format or fields in a file. For managing and accessing the data Structured Query Language (SQL) is used.
Traditional data is characterized by its high level of organization and structure, which makes it easy to store, manage, and analyze. Traditional data analysis techniques involve using statistical methods and visualizations to identify patterns and trends in the data.

Traditional data is often collected and managed by enterprise resource planning (ERP) systems and other enterprise-level applications. This data is critical for businesses to make informed decisions and drive performance improvements.

2. Big data: We can consider big data an upper version of traditional data. Big data deal with too large or complex data sets which is difficult to manage in traditional data-processing application software. It deals with large volume of both structured, semi structured and unstructured data. Volume, Velocity and Variety, Veracity and Value refer to the 5'V characteristics of big data. Big data not only refers to large amount of data it refers to extracting meaningful data by analyzing the huge amount of complex data sets. semi-structured
Big data is characterized by the three Vs: volume, velocity, and variety. Volume refers to the vast amount of data that is generated and collected; velocity refers to the speed at which data is generated and must be processed; and variety refers to the many different types and formats of data that must be analyzed, including structured, semi-structured, and unstructured data.

Due to the size and complexity of big data sets, traditional data management tools and techniques are often inadequate for processing and analyzing the data. Big data technologies, such as Hadoop, Spark, and NoSQL databases, have emerged to help organizations store, manage, and analyze large volumes of data.

The main differences between traditional data and big data as follows:

- Volume: Traditional data typically refers to small to medium-sized datasets that can be easily stored and analyzed using traditional data processing technologies. In contrast, big data refers to extremely large datasets that cannot be easily managed or processed using traditional technologies.
- Variety: Traditional data is typically structured, meaning it is organized in a predefined manner such as tables, columns, and rows. Big data, on the other hand, can be structured, unstructured, or semi-structured, meaning it may contain text, images, videos, or other types of data.

- Velocity: Traditional data is usually static and updated on a periodic basis. In contrast, big data is constantly changing and updated in real-time or near real-time.
- Complexity: Traditional data is relatively simple to manage and analyze. Big data, on the other hand, is complex and requires specialized tools and techniques to manage, process, and analyze.
- Value: Traditional data typically has a lower potential value than big data because it is limited in scope and size. Big data, on the other hand, can provide valuable insights into customer behavior, market trends, and other business-critical information.

Some similarities between them, including:

- Data Quality: The quality of data is essential in both traditional and big data environments. Accurate and reliable data is necessary for making informed business decisions.
- Data Analysis: Both traditional and big data require some form of analysis to derive insights and knowledge from the data. Traditional data analysis methods typically involve statistical techniques and visualizations, while big data analysis may require machine learning and other advanced techniques.
- Data Storage: In both traditional and big data environments, data needs to be stored and managed effectively. Traditional data is typically stored in relational databases, while big data may require specialized technologies such as Hadoop, NoSQL, or cloud-based storage systems.
- Data Security: Data security is a critical consideration in both traditional and big data environments. Protecting sensitive information from unauthorized access, theft, or misuse is essential in both contexts.
- Business Value: Both traditional and big data can provide significant value to organizations. Traditional data can provide insights into historical trends and patterns, while big data can uncover new opportunities and help organizations make more informed decisions.

The difference between Traditional data and Big data are as follows:

Traditional Data	Big Data
Traditional data is generated in enterprise level.	Big data is generated outside the enterprise level.
Its volume ranges from Gigabytes to Terabytes.	Its volume ranges from Petabytes to Zettabytes or Exabytes.
Traditional database system deals with structured data.	Big data system deals with structured, semi-structured,database, and unstructured data.

Traditional Data	Big Data
Traditional data is generated per hour or per day or more.	But big data is generated more frequently mainly per seconds.
Traditional data source is centralized and it is managed in centralized form.	Big data source is distributed and it is managed in distributed form.
Data integration is very easy.	Data integration is very difficult.
Normal system configuration is capable to process traditional data.	High system configuration is required to process big data.
The size of the data is very small.	The size is more than the traditional data size.
Traditional data base tools are required to perform any data base operation.	Special kind of data base tools are required to perform any databaseschema-based operation.
Normal functions can manipulate data.	Special kind of functions can manipulate data.
Its data model is strict schema based and it is static.	Its data model is a flat schema based and it is dynamic.
Traditional data is stable and inter relationship.	Big data is not stable and unknown relationship.
Traditional data is in manageable volume.	Big data is in huge volume which becomes unmanageable.
It is easy to manage and manipulate the data.	It is difficult to manage and manipulate the data.
Its data sources includes ERP transaction data, CRM transaction data, financial data, organizational data, web transaction data etc.	Its data sources includes social media, device data, sensor data, video, images, audio etc.

Conclusion :
The key differences between traditional data and big data are related to the volume, variety, velocity, complexity, and potential value of the data. Traditional data is typically small in size, structured, and static, while big data is large, complex, and constantly changing. As a result, big data requires specialized tools and techniques to manage and analize effectively.

Data Warehousing

A Database Management System (DBMS) stores data in the form of tables and uses an ER model and the goal is ACID properties. For example, a DBMS of a college has tables for students, faculty, etc.
A **Data Warehouse** is separate from DBMS, it stores a huge amount of data, which is typically collected from multiple heterogeneous sources like files, DBMS, etc. The goal is to produce statistical results that may help in decision-making. For example, a college might want to see quick different results, like how the placement of CS students has improved over the last 10 years, in terms of salaries, counts, etc.

Need for Data Warehouse

An ordinary Database can store MBs to GBs of data and that too for a specific purpose. For storing data of TB size, the storage shifted to the Data Warehouse. Besides this, a transactional database doesn't offer itself to analytics. To effectively perform analytics, an organization keeps a central Data Warehouse to closely study its business by organizing, understanding, and using its historical data for making strategic decisions and analyzing trends.

Benefits of Data Warehouse

- **Better business analytics:** Data warehouse plays an important role in every business to store and analysis of all the past data and records of the company. which can further increase the understanding or analysis of data for the company.
- **Faster Queries: The data** warehouse is designed to handle large queries that's why it runs queries faster than the database.
- **Improved data Quality:** In the data warehouse the data you gathered from different sources is being stored and analyzed it does not interfere with or add data by itself so your quality of data is maintained and if you get any issue regarding data quality then the data warehouse team will solve this.
- **Historical Insight:** The warehouse stores all your historical data which contains details about the business so that one can analyze it at any time and extract insights from it.

Data Warehouse vs DBMS

Database	Data Warehouse
A common Database is based on operational or transactional processing. Each operation is an indivisible transaction.	A data Warehouse is based on analytical processing.
Generally, a Database stores current and up-to-date data which is used for daily operations.	A Data Warehouse maintains historical data over time. Historical data is the data kept over years and can used for trend analysis, make future predictions and decision support.
A database is generally application specific. **Example –** A database stores related data, such as the student details in a school.	A Data Warehouse is integrated generally at the organization level, by combining data from different databases. **Example –** A data warehouse integrates the data from one or more databases , so that analysis can be done to get results , such as the best performing school in a city.
Constructing a Database is not so expensive.	Constructing a Data Warehouse can be expensive.

Example Applications of Data Warehousing

Data Warehousing can be applied anywhere where we have a huge amount of data and we want to see statistical results that help in decision making.

- **Social Media Websites:** The social networking websites like Facebook, Twitter, Linkedin, etc. are based on analyzing large data sets. These sites gather data related to members, groups, locations, etc., and store it in a single central repository. Being a large amount of data, Data Warehouse is needed for implementing the same.
- **Banking:** Most of the banks these days use warehouses to see the spending patterns of account/cardholders. They use this to provide them with special offers, deals, etc.
- **Government:** Government uses a data warehouse to store and analyze tax payments which are used to detect tax thefts.

Features of Data Warehousing

Data warehousing is essential for modern data management, providing a strong foundation for organizations to consolidate and analyze data strategically. Its distinguishing features empower businesses with the tools to make informed decisions and extract valuable insights from their data.

- **Centralized Data Repository:** Data warehousing provides a centralized repository for all enterprise data from various sources, such as transactional databases, operational systems, and external sources. This enables organizations to have a comprehensive view of their data, which can help in making informed business decisions.
- **Data Integration:** Data warehousing integrates data from different sources into a single, unified view, which can help in eliminating data silos and reducing data inconsistencies.
- **Historical Data Storage:** Data warehousing stores historical data, which enables organizations to analyze data trends over time. This can help in identifying patterns and anomalies in the data, which can be used to improve business performance.
- **Query and Analysis:** Data warehousing provides powerful query and analysis capabilities that enable users to explore and analyze data in different ways. This can help in identifying patterns and trends, and can also help in making informed business decisions.
- **Data Transformation:** Data warehousing includes a process of data transformation, which involves cleaning, filtering, and formatting data from various sources to make it consistent and usable. This can help in improving data quality and reducing data inconsistencies.
- **Data Mining:** Data warehousing provides data mining capabilities, which enable organizations to discover hidden patterns and relationships in their data. This can help in identifying new opportunities, predicting future trends, and mitigating risks.
- **Data Security:** Data warehousing provides robust data security features, such as access controls, data encryption, and data backups, which ensure that the data is secure and protected from unauthorized access.

Advantages of Data Warehousing

- **Intelligent Decision-Making:** With centralized data in warehouses, decisions may be made more quickly and intelligently.
- **Business Intelligence:** Provides strong operational insights through business intelligence.
- **Historical Analysis:** Predictions and trend analysis are made easier by storing past data.
- **Data Quality:** Guarantees data quality and consistency for trustworthy reporting.
- **Scalability:** Capable of managing massive data volumes and expanding to meet changing requirements.
- **Effective Queries:** Fast and effective data retrieval is made possible by an optimized structure.
- **Cost reductions:** Data warehousing can result in cost savings over time by reducing data management procedures and increasing overall efficiency, even when there are setup costs initially.

- **Data security:** Data warehouses employ security protocols to safeguard confidential information, guaranteeing that only authorized personnel are granted access to certain data.

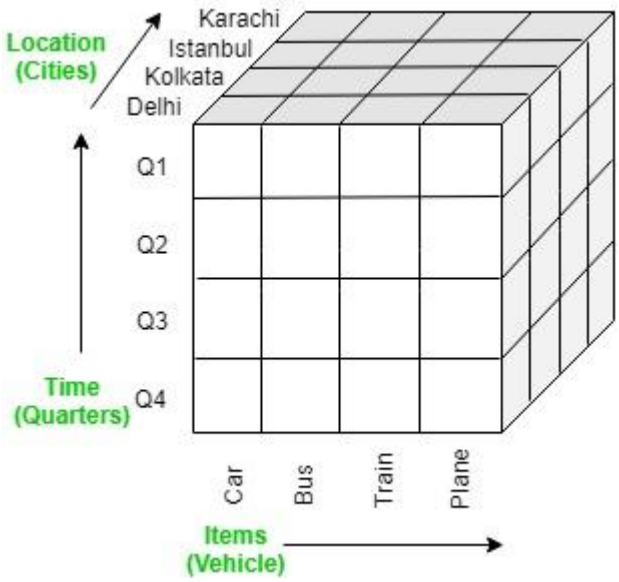
Disadvantages of Data Warehousing

- **Cost:** Building a data warehouse can be expensive, requiring significant investments in hardware, software, and personnel.
- **Complexity:** Data warehousing can be complex, and businesses may need to hire specialized personnel to manage the system.
- **Time-consuming:** Building a data warehouse can take a significant amount of time, requiring businesses to be patient and committed to the process.
- **Data integration challenges:** Data from different sources can be challenging to integrate, requiring significant effort to ensure consistency and accuracy.
- **Data security:** Data warehousing can pose data security risks, and businesses must take measures to protect sensitive data from unauthorized access or breaches.

There can be many more applications in different sectors like E-Commerce, telecommunications, Transportation Services, Marketing and Distribution, Healthcare, and Retail.

OLAP Operations in DBMS

OLAP stands for ***Online Analytical Processing*** Server. It is a software technology that allows users to analyze information from multiple database systems at the same time. It is based on multidimensional data model and allows the user to query on multi-dimensional data (eg. Delhi -> 2018 -> Sales data). OLAP databases are divided into one or more cubes and these cubes are known as *Hyper-cubes*.

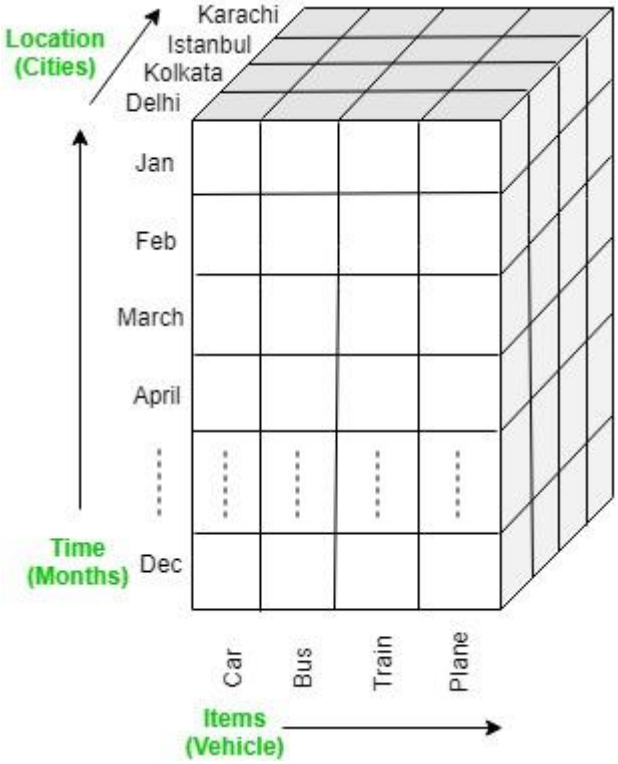


OLAP operations:

There are five basic analytical operations that can be performed on an OLAP cube:

1. **Drill down:** In drill-down operation, the less detailed data is converted into highly detailed data. It can be done by:
    - Moving down in the concept hierarchy
    - Adding a new dimension

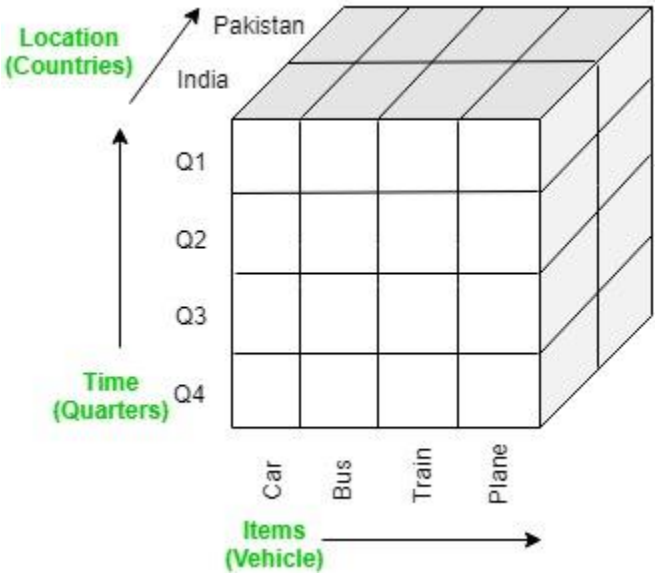
    In the cube given in overview section, the drill down operation is performed by moving down in the concept hierarchy of *Time* dimension (Quarter -> Month).



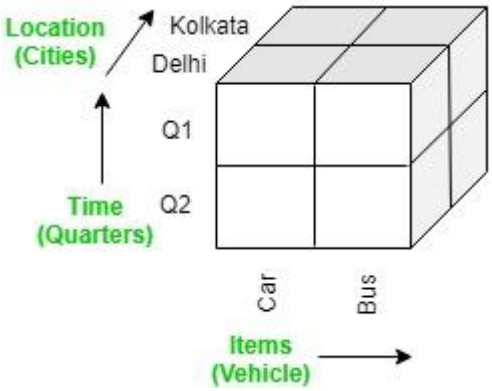
2. **Roll up:** It is just opposite of the drill-down operation. It performs aggregation on the OLAP cube. It can be done by:
    - Climbing up in the concept hierarchy
    - Reducing the dimensions

    In the cube given in the overview section, the roll-up operation is performed by climbing up in the concept hierarchy

of *Location* dimension (City -> Country).

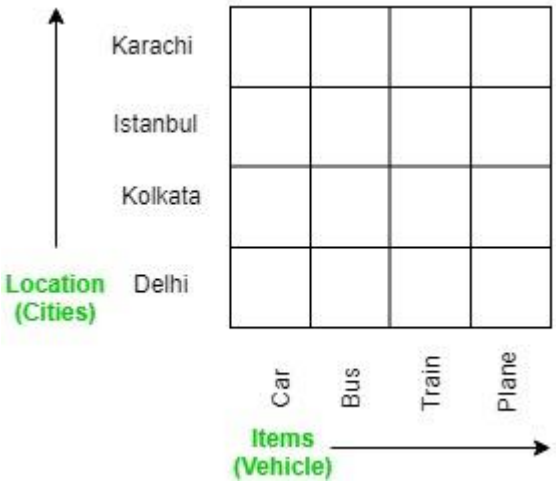


3. **Dice:** It selects a sub-cube from the OLAP cube by selecting two or more dimensions. In the cube given in the overview section, a sub-cube is selected by selecting following dimensions with criteria:
    - Location = “Delhi” or “Kolkata”
    - Time = “Q1” or “Q2”
    - Item = “Car” or “Bus”

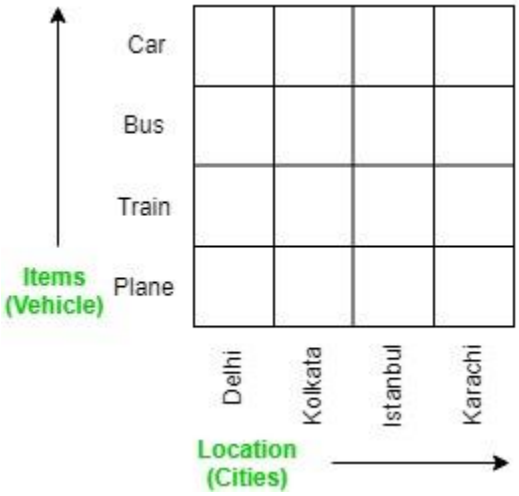


4. **Slice:** It selects a single dimension from the OLAP cube which results in a new sub-cube creation. In the cube given in the overview section,

Slice is performed on the dimension Time = “Q1”.



5. **Pivot:** It is also known as *rotation* operation as it rotates the current view to get a new view of the representation. In the sub-cube obtained after the slice operation, performing pivot operation gives a new view of it.



What is OLTP?

OLTP is an operational system that supports transaction-oriented applications in a 3-tier architecture. It administers the day to day transaction of an organization. OLTP is basically focused on query processing, maintaining data integrity in multi-access environments as well as effectiveness that is measured by the total number of transactions per second.

The full form of OLTP is Online Transaction Processing.

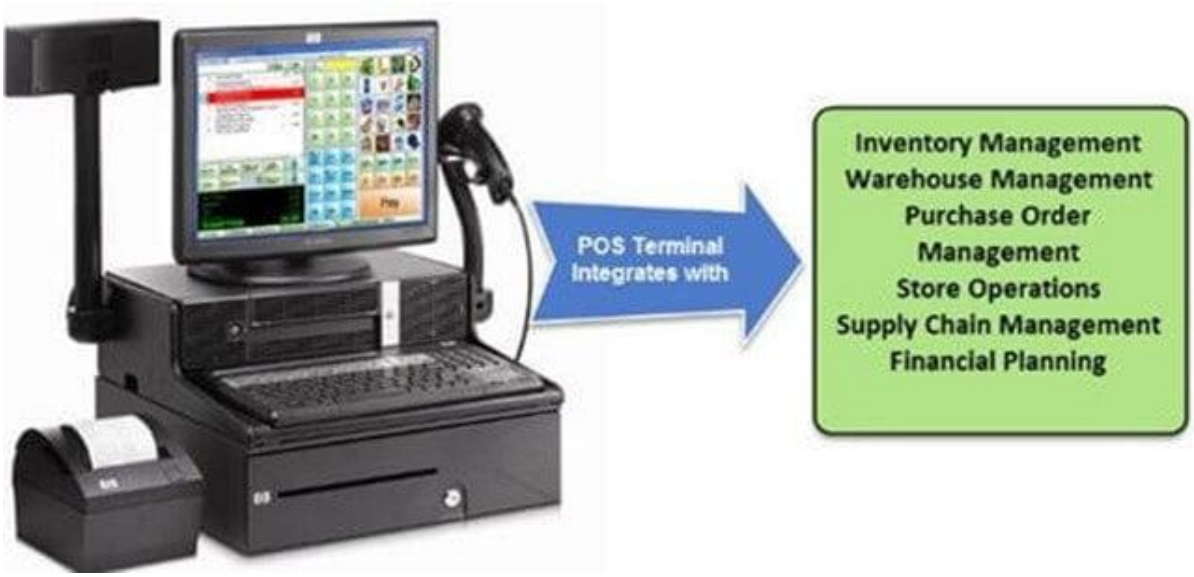
Characteristics of OLTP

Following are important characteristics of OLTP:

- OLTP uses transactions that include small amounts of data.
- Indexed data in the database can be accessed easily.
- OLTP has a large number of users.
- It has fast response times
- Databases are directly accessible to end-users
- OLTP uses a fully normalized schema for database consistency.
- The response time of OLTP system is short.
- It strictly performs only the predefined operations on a small number of records.
- OLTP stores the records of the last few days or a week.

Type of queries that an OLTP system can Process

OLTP system is an online database changing system. Therefore, it supports database query such as insert, update, and delete information from the database.



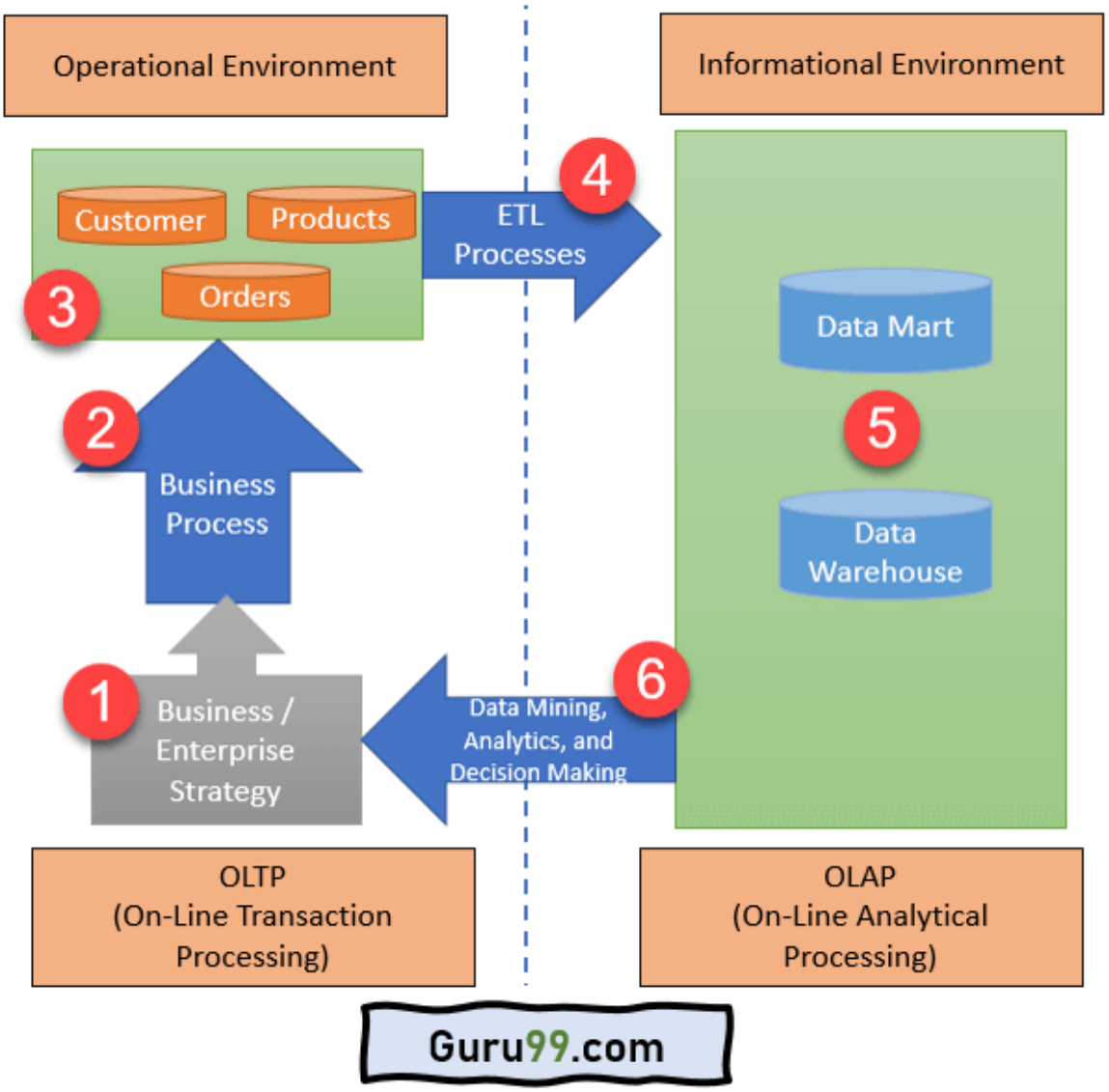
POS system for OLTP

Consider a point of sale system of a supermarket, following are the sample queries that this system can process:

- Retrieving the description of a particular product.
- Filtering all products related to the supplier.
- Searching the record of the customer.
- Listing products having a price less than the expected amount.

Architecture of OLTP

Here is the architecture of OLTP:



OLTP Architecture

1. **Business / Enterprise Strategy:** Enterprise strategy deals with the issues that affect the organization as a whole. In OLTP, it is typically developed at a high level within the firm, by the board of directors or the top management

2. **Business Process:** OLTP business process is a set of activities and tasks that, once completed, will accomplish an organizational goal.
3. **Customers, Orders, and Products:** OLTP database store information about products, orders (transactions), customers (buyers), suppliers (sellers), and employees.
4. **ETL Processes:** It separates the data from various RDBMS source systems, then transforms the data (like applying concatenations, calculations, etc.) and loads the processed data into the Data Warehouse system.
5. **Data Mart and Data warehouse:** A Data Mart is a structure/access pattern specific to data warehouse environments. It is used by OLAP to store processed data.
6. **Data Mining, Analytics, and Decision Making:** Data stored in the data mart and data warehouse can be used for data mining, analytics, and decision making. This data helps you to discover data patterns, analyze raw data, and make analytical decisions for your organization's growth.

Example of OLTP Transaction

An example of the OLTP system is the ATM center. Assume that a couple has a joint account with a bank. One day both simultaneously reach different ATM centers at precisely the same time and want to withdraw the total amount present in their bank account.

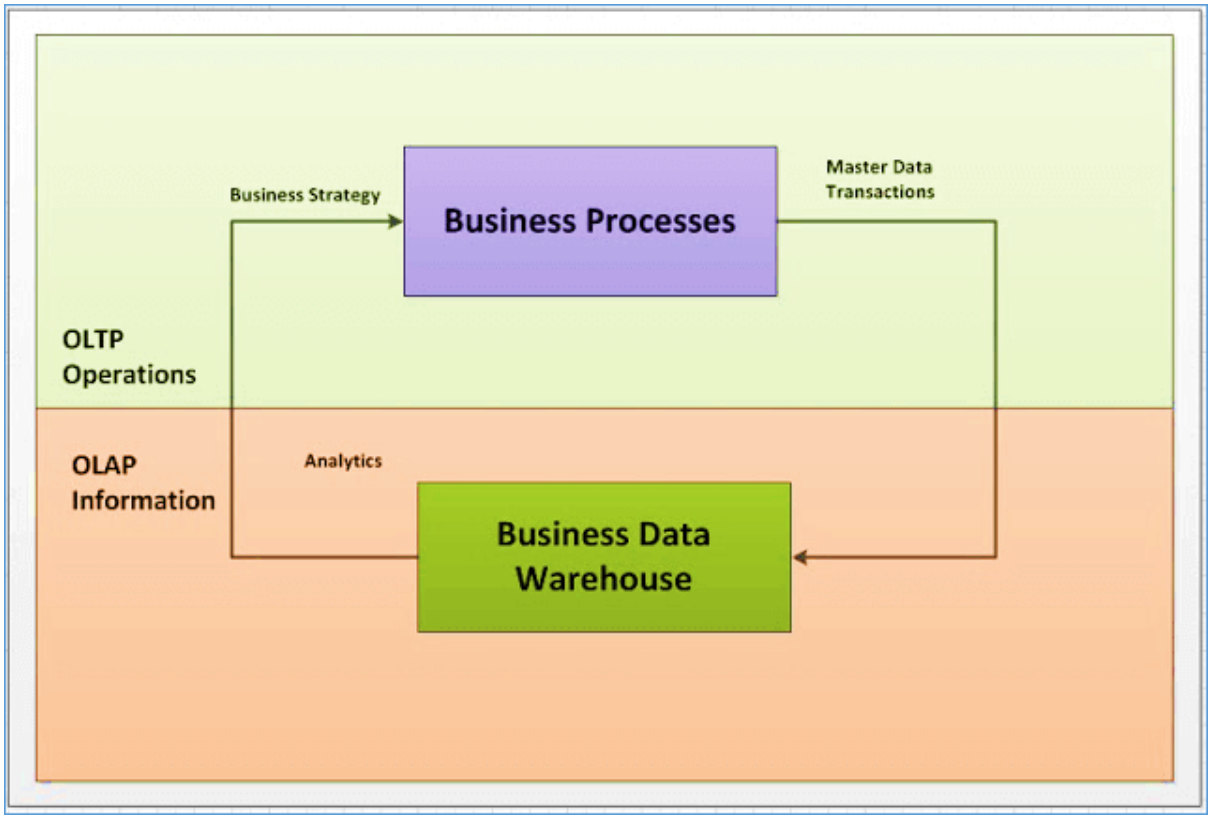
OLTP for ATM image

However, the person that completes the authentication process first will be able to get money. In this case, the OLTP system makes sure that the withdrawn amount will be never more than the amount present in the bank. The key to note here is that OLTP systems are optimized for transactional superiority instead of data analysis.

Other examples of OLTP system are:

- Online banking
- Online airline ticket booking
- Sending a text message
- Order entry
- Add a book to shopping cart

OLTP vs. OLAP



Here is the important difference between OLTP and OLAP:

OLTP	OLAP
OLTP is an online transactional system.	OLAP is an online analysis and data retrieving process.
It is characterized by large numbers of short online transactions.	It is characterized by a large volume of data.
OLTP is an online database modifying system.	OLAP is an online database query management system.
OLTP uses traditional DBMS.	OLAP uses the data warehouse.
Insert, Update, and Delete information from the database.	Mostly select operations

OLTP	OLAP
OLTP and its transactions are the sources of data.	Different OLTP databases become the source of data for OLAP.
OLTP database must maintain data integrity constraints.	OLAP database does not get frequently modified. Hence, data integrity is not an issue.
It's response time is in a millisecond.	Response time in seconds to minutes.
The data in the OLTP database is always detailed and organized.	The data in the OLAP process might not be organized.
Allow read/write operations.	Only read and rarely write.
It is a market-orientated process.	It is a customer orientated process.
Queries in this process are standardized and simple.	Complex queries involving aggregations.
Complete backup of the data combined with incremental backups.	OLAP only need a backup from time to time. Backup is not important compared to OLTP
DB design is an application-oriented example: Database design changes with the industry like retail, airline, banking, etc.	DB design is subject-oriented. Example: Database design changes with subjects like sales, marketing, purchasing, etc.
It is used by Data critical users like clerk, DBA & Data Base professionals.	It is used by Data knowledge users like workers, managers, and CEO.
It is designed for real time business operations.	It is designed for analysis of business measures by category and attributes.
Transaction throughput is the performance metric	Query throughput is the performance metric.
This kind of Database user allows thousands of users.	This kind of Database allows only hundreds of users.
It helps to Increase user's self-service and productivity	Help to Increase the productivity of business analysts.
Data Warehouses historically have been a development project which may prove costly to build.	An OLAP cube is not an open SQL server data warehouse. Therefore, technical knowledge and experience are essential to managing the OLAP server.
It provides a fast result for daily used data.	It ensures that response to the query is quicker consistently.
It is easy to create and maintain.	It lets the user create a view with the help of a spreadsheet.
OLTP is designed to have fast response time, low data redundancy, and is normalized.	A data warehouse is created uniquely so that it can integrate different data sources for building a consolidated database

Advantages of OLTP

Following are the pros/benefits of OLTP system:

- OLTP offers accurate forecast for revenue and expense.
- It provides a solid foundation for a stable business /organization due to timely modification of all transactions.

- OLTP makes transactions much easier on behalf of the customers.
- It broadens the client base for an organization by speeding up and simplifying individual processes.
- OLTP provides support for bigger databases.
- Partition of data for data manipulation is easy.
- We need OLTP to use the tasks which are frequently performed by the system.
- When we need only a small number of records.
- The tasks that include insertion, updation, or deletion of data.
- It is used when you need consistency and concurrency in order to perform tasks that ensure its greater availability.

Disadvantages of OLTP

Here are cons/drawbacks of OLTP system:

- If the OLTP system faces hardware failures, then online transactions get severely affected.
- OLTP systems allow multiple users to access and change the same data at the same time, which many times created an unprecedented situation.
- If the server hangs for seconds, it can affect to a large number of transactions.
- OLTP required a lot of staff working in groups in order to maintain inventory.
- Online Transaction Processing Systems do not have proper methods of transferring products to buyers by themselves.
- OLTP makes the database much more susceptible to hackers and intruders.
- In B2B transactions, there are chances that both buyers and suppliers miss out efficiency advantages that the system offers.
- Server failure may lead to wiping out large amounts of data from the database.
- You can perform a limited number of queries and updates.

Challenges of an OLTP System

- It allows more than one user to access and change the same data simultaneously. Therefore, it requires concurrency control and recovery technique in order to avoid any unprecedented situations
- OLTP system data are not suitable for decision making. You have to use data of OLAP systems for “what if” analysis or the decision making.

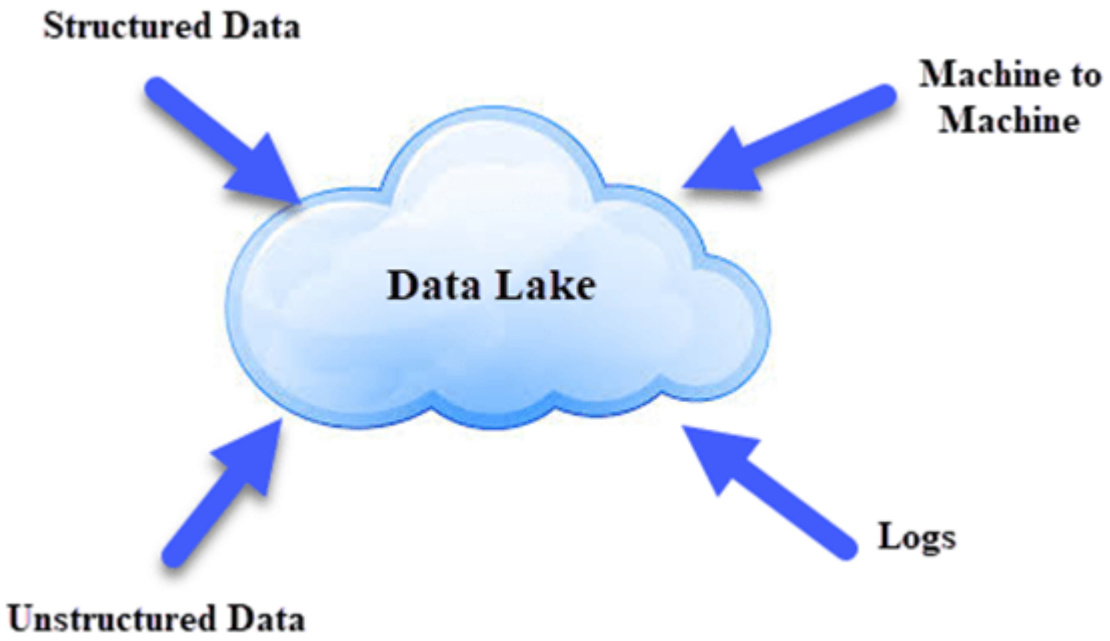
Summary

- OLTP is defined as an operational system that supports transaction-oriented applications in a 3-tier architecture.
- OLTP uses transactions that include small amounts of data.
- OLTP system is an online database changing system.
- The architecture of OLTP contains 1) Business / Enterprise Strategy, 2) Business Process, 3) Customers, Orders, and Products, 4) ETL Processes, 5) Data Mart and Data warehouse, and 6) Data Mining, Analytics, and Decision Making.
- OLTP is an online transactional system, whereas OLAP is an online analysis and data retrieving process.
- OLTP provides a solid foundation for a stable business /organization due to timely modification of all transactions.
- OLTP systems allow multiple users to access and change the same data at the same time, which many times created an unprecedented situation.

What is Data Lake?

A Data Lake is a storage repository that can store large amount of structured, semi-structured, and unstructured data. It is a place to store every type of data in its native format with no fixed limits on account size or file. It offers high data quantity to increase analytic performance and native integration.

Data Lake is like a large container which is very similar to real lake and rivers. Just like in a lake you have multiple branches coming in, a data lake has structured data, unstructured data, machine to machine, logs flowing through in real-time.



Data Lake

The Data Lake democratizes data and is a cost-effective way to store all data of an organization for later processing. Research Analyst can focus on finding meaning patterns in data and not data itself.

Unlike a hierarchal Data Warehouse where data is stored in Files and Folder, Data lake has a flat architecture. Every data elements in a Data Lake is given a unique identifier and tagged with a set of metadata information.

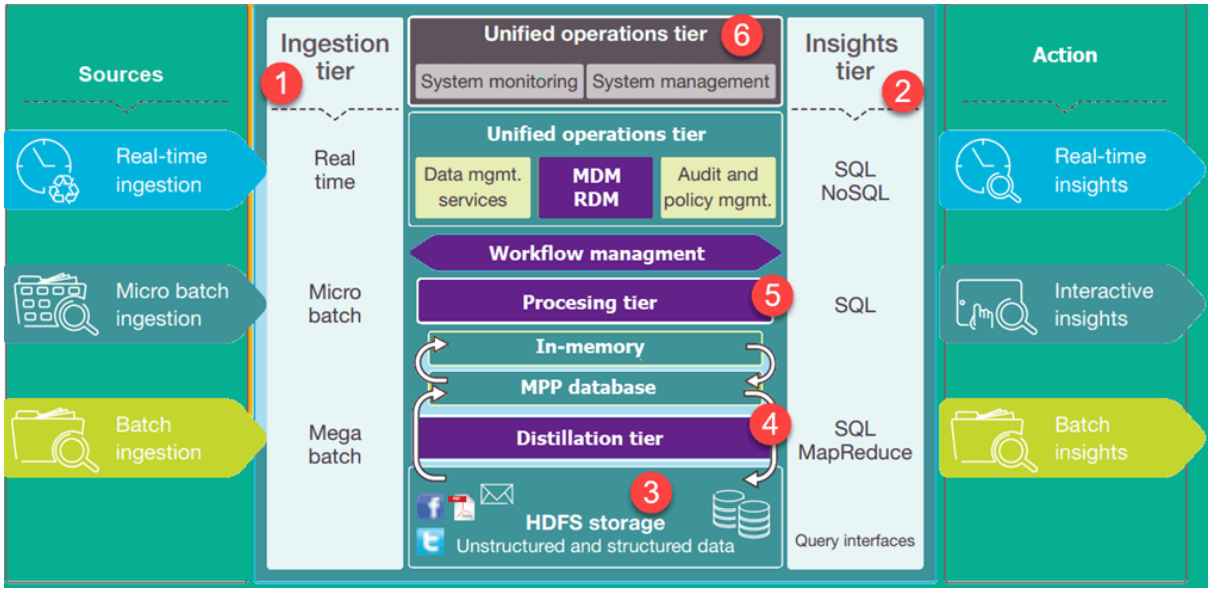
Why Data Lake?

The main objective of building a data lake is to offer an unrefined view of data to data scientists.

Reasons for using Data Lake are:

- With the onset of storage engines like Hadoop storing disparate information has become easy. There is no need to model data into an enterprise-wide schema with a Data Lake.
- With the increase in data volume, data quality, and metadata, the quality of analyses also increases.
- Data Lake offers business Liveliness
- Machine Learning and Artificial Intelligence can be used to make profitable predictions.
- It offers a competitive advantage to the implementing organization.
- There is no data silo structure. Data Lake gives 360 degrees view of customers and makes analysis more robust.

Data Lake Architecture



Data Lake Architecture

The figure shows the architecture of a Business Data Lake. The lower levels represent data that is mostly at rest while the upper levels show

real-time transactional data. This data flow through the system with no or little latency. Following are important tiers in Data Lake Architecture:

1. **Ingestion Tier**: The tiers on the left side depict the data sources. The data could be loaded into the data lake in batches or in real-time
2. **Insights Tier:** The tiers on the right represent the research side where insights from the system are used. SQL, NoSQL queries, or even excel could be used for data analysis.
3. **HDFS** is a cost-effective solution for both structured and unstructured data. It is a landing zone for all data that is at rest in the system.
4. **Distillation tier** takes data from the storage tire and converts it to structured data for easier analysis.
5. **Processing tier** run analytical algorithms and users queries with varying real time, interactive, batch to generate structured data for easier analysis.
6. **Unified operations tier** governs system management and monitoring. It includes auditing and proficiency management, data management, workflow management.

Key Data Lake Concepts

Following are Key Data Lake concepts that one needs to understand to completely understand the Data Lake Architecture

Key

Data Ingestion

Data Ingestion allows connectors to get data from a different data sources and load into the Data lake.

Data Ingestion supports:

- All types of Structured, Semi-Structured, and Unstructured data.
- Multiple ingestions like Batch, Real-Time, One-time load.
- Many types of data sources like Databases, Webservers, Emails, IoT, and FTP.

Data Storage

Data storage should be scalable, offers cost-effective storage and allow fast access to data exploration. It should support various data formats.

Data Governance

Data governance is a process of managing availability, usability, security, and integrity of data used in an organization.

Security

Security needs to be implemented in every layer of the Data lake. It starts with Storage, Unearthing, and Consumption. The basic need is to stop access for unauthorized users. It should support different tools to access data with easy to navigate GUI and Dashboards.

Authentication, Accounting, Authorization and Data Protection are some important features of data lake security.

Data Quality

Data quality is an essential component of Data Lake architecture. Data is used to exact business value. Extracting insights from poor quality data will lead to poor quality insights.

Data Discovery

Data Discovery is another important stage before you can begin preparing data or analysis. In this stage, tagging technique is used to express the data understanding, by organizing and interpreting the data ingested in the Data lake.

Data Auditing

Two major Data auditing tasks are tracking changes to the key dataset.

1. Tracking changes to important dataset elements
2. Captures how/ when/ and who changes to these elements.

Data auditing helps to evaluate risk and compliance.

Data Lineage

This component deals with data's origins. It mainly deals with where it movers over time and what happens to it. It eases errors corrections in a data analytics process from origin to destination.

Data Exploration

It is the beginning stage of data analysis. It helps to identify right dataset is vital before starting Data Exploration.

All given components need to work together to play an important part in Data lake building easily evolve and explore the environment.

Maturity stages of Data Lake

The Definition of Data Lake Maturity stages differs from textbook to other. Though the crux remains the same. Following maturity, stage definition is from a layman point of view.

Maturity stages of Data Lake

Stage 1: Handle and ingest data at scale

This first stage of Data Maturity Involves improving the ability to transform and analyze data. Here, business owners need to find the tools according to their skillset for obtaining more data and build analytical applications.

Stage 2: Building the analytical muscle

This is a second stage which involves improving the ability to transform and analyze data. In this stage, companies use the tool which is most appropriate to their skillset. They start acquiring more data and building applications. Here, capabilities of the enterprise data warehouse and data lake are used together.

Stage 3: EDW and Data Lake work in unison

This step involves getting data and analytics into the hands of as many people as possible. In this stage, the data lake and the enterprise data warehouse start to work in a union. Both playing their part in analytics

Stage 4: Enterprise capability in the lake

In this maturity stage of the data lake, enterprise capabilities are added to the Data Lake. Adoption of information governance, information lifecycle management capabilities, and Metadata management. However, very few organizations can reach this level of maturity, but this tally will increase in the future.

Best practices for Data Lake Implementation

- Architectural components, their interaction and identified products should support native data types
- Design of Data Lake should be driven by what is available instead of what is required. The schema and data requirement is not defined until it is queried
- Design should be guided by disposable components integrated with service API.
- Data discovery, ingestion, storage, administration, quality, transformation, and visualization should be managed independently.
- The Data Lake architecture should be tailored to a specific industry. It should ensure that capabilities necessary for that domain are an inherent part of the design
- Faster on-boarding of newly discovered data sources is important
- Data Lake helps customized management to extract maximum value
- The Data Lake should support existing enterprise data management techniques and methods

Challenges of building a data lake:

- In Data Lake, Data volume is higher, so the process must be more reliant on programmatic administration
- It is difficult to deal with sparse, incomplete, volatile data
- Wider scope of dataset and source needs larger data governance & support

What is Data Lake?

- A **Data Lake** is a storage repository that can store a large amount of structured, semi-structured, and unstructured data. It is a place to store every type of data in its native format with no fixed limits on account size or file. It offers a large amount of data quantity for increased analytical performance and native integration.
- Data Lake is like a large container which is very similar to real lake and rivers. Just like in a lake, you have multiple tributaries coming in; similarly, a data lake has structured data, unstructured data, machine to machine, logs flowing through in real-time.

What is Data Warehouse?

- **Data Warehouse** is a blend of technologies and components for the strategic use of data. It collects and manages data from varied sources to provide meaningful business insights. It is the electronic storage of a large amount of information designed for query and analysis instead of transaction processing. It is a process of transforming data into information.
- Next, we will learn the key difference between data warehouse vs data lake.

Difference between Data Lake and Data Warehouse

- Here are the key differences between data lake versus data warehouse:

Parameters	Data Lake	Data Warehouse
Storage	In the data lake, all data is kept irrespective of the source and its structure. Data is kept in its	A data warehouse will consist of data that is extracted from transactional systems or

Parameters	Data Lake	Data Warehouse
	raw form. It is only transformed when it is ready to be used.	data which consists of quantitative metrics with their attributes. The data is cleaned and transformed
History	Big data technologies used in data lakes is relatively new.	Data warehouse concept, unlike big data, had been used for decades.
Data Capturing	Captures all kinds of data and structures, semi-structured and unstructured in their original form from source systems.	Captures structured information and organizes them in schemas as defined for data warehouse purposes
Data Timeline	Data lakes can retain all data. This includes not only the data that is in use but also data that it might use in the future. Also, data is kept for all time, to go back in time and do an analysis.	In the data warehouse development process, significant time is spent on analyzing various data sources.
Users	Data lake is ideal for the users who indulge in deep analysis. Such users include data scientists who need advanced analytical tools with capabilities such as predictive modeling and statistical analysis.	The data warehouse is ideal for operational users because of being well structured, easy to use and understand.
Storage Costs	Data storing in big data technologies are relatively inexpensive then storing data in a data warehouse.	Storing data in Data warehouse is costlier and time-consuming.
Task	Data lakes can contain all data and data types; it empowers users to access data prior the process of transformed, cleansed and structured.	Data warehouses can provide insights into pre-defined questions for pre-defined data types.

Parameters	Data Lake	Data Warehouse
Processing time	Data lakes empower users to access data before it has been transformed, cleansed and structured. Thus, it allows users to get to their result more quickly compares to the traditional data warehouse.	Data warehouses offer insights into pre-defined questions for pre-defined data types. So, any changes to the data warehouse needed more time.
Position of Schema	Typically, the schema is defined after data is stored. This offers high agility and ease of data capture but requires work at the end of the process	Typically schema is defined before data is stored. Requires work at the start of the process, but offers performance, security, and integration.
Data processing	Data Lakes use of the ELT (Extract Load Transform) process.	Data warehouse uses a traditional ETL (Extract Transform Load) process.
Complain	Data is kept in its raw form. It is only transformed when it is ready to be used.	The chief complaint against data warehouses is the inability, or the problem faced when trying to make change in in them.
Key Benefits	They integrate different types of data to come up with entirely new questions as these users not likely to use data warehouses because they may need to go beyond its capabilities.	Most users in an organization are operational. These type of users only care about reports and key performance metrics.

Benefits and Risks of using Data Lake

Here are some major benefits in using a Data Lake:

- Helps fully with product ionizing & advanced analytics
- Offers cost-effective scalability and flexibility
- Offers value from unlimited data types
- Reduces long-term cost of ownership
- Allows economic storage of files
- Quickly adaptable to changes
- The main advantage of data lake is the **centralization** of different content sources
- Users, from various departments, may be scattered around the globe can have **flexible access** to the data

Risk of Using Data Lake:

- After some time, Data Lake may lose significance and momentum
- There is larger amount risk involved while designing Data Lake
- Unstructured Data may lead to Ungoverned Chao, Unusable Data, Disparate & Complex Tools, Enterprise-Wide Collaboration, United, Consistent, and Common
- It also increases storage & computes costs
- There is no way to get insights from others who have worked with the data because there is no account of the lineage of findings by previous analysts
- The biggest risk of data lakes is security and access control. Sometimes data can be placed into a lake without any oversight, as some of the data may have privacy and regulatory need

Summary

- A Data Lake is a storage repository that can store large amount of structured, semi-structured, and unstructured data.
- The main objective of building a data lake is to offer an unrefined view of data to data scientists.

- Unified operations tier, Processing tier, Distillation tier and HDFS are important layers of Data Lake Architecture
- Data Ingestion, Data storage, Data quality, Data Auditing, Data exploration, Data discover are some important components of Data Lake Architecture
- Design of Data Lake should be driven by what is available instead of what is required.
- Data Lake reduces long-term cost of ownership and allows economic storage of files
- The biggest risk of data lakes is security and access control. Sometimes data can be placed into a lake without any oversight, as some of the data may have privacy and regulatory need.