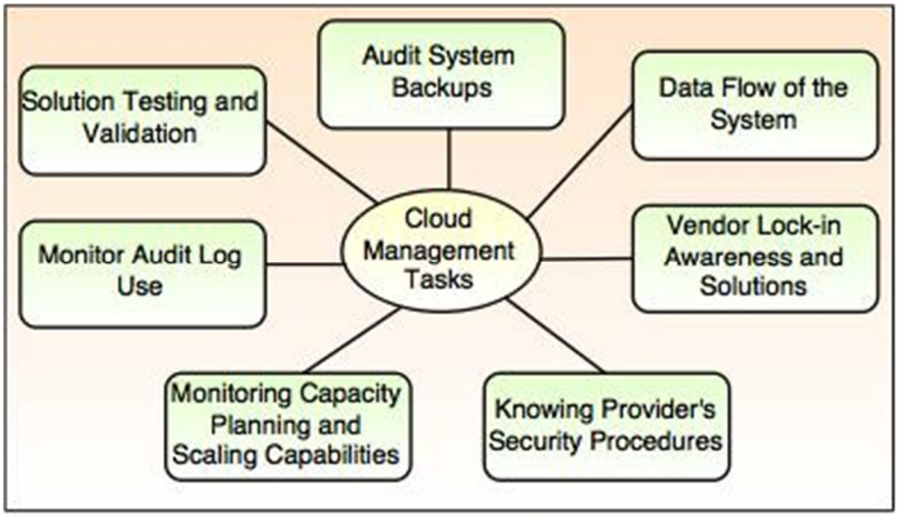
UNIT – III

SERVICE MANAGEMENT IN CLOUD COMPUTING

Cloud service management and operations refer to all the activities that an organization does to plan, design, deliver, operate, and control the IT and cloud services that it offers to customers. Service management includes the operational aspects of your applications and services. Cloud management tools help to ensure that cloud computing-based resources are working optimally and properly interacting with users and other services.

Service management includes the operational aspects of your applications and services. After an application is pushed to production, it must be managed. Applications are monitored to ensure availability and performance according to service level agreements (SLAs) or service level objectives (SLOs).

Some of the cloud management tasks executed by cloud provider are:



1) Audit System Backups

It is needed to audit backups timely to verify restoring of randomly selected files of different users.

Following are the two ways to perform backups:

i) Backing up files by the company, from on-site computers to disks which occupy in the cloud.

ii) Backing up files by the cloud provider.

2) Data Flow of the System

The managers are responsible to develop a diagram to describe a detailed process flow. This process flow explains the movement of data belonging to an organization throughout the cloud solution.

3) Vendor Lock-In Awareness and Solutions

The procedure to exit from services of a particular cloud provider must be known to the managers.

The procedures are defined to allow the cloud mangers to export data of an organization from their system to another cloud provider.

4) Knowing Provider's Security Procedures

It is important to the security plans of the provider for the services like Multiuser use, E-commerce processing, Employee screening, Encryption policy.

5) Monitoring Capacity Planning and Scaling Capabilities

The manager should know the capacity planning for checking whether the cloud provider would be able to meet the future requirement of business or not.

The managers should manage the scaling capabilities for ensuring services that can be scaled up or down, according to the need.

6) Monitor Audit Log Use

To recognize the errors in the system, managers should audit the log on a regular basis.

7) Solution Testing and Validation

When a cloud provider offers a solution, it is necessary to test that solution and verify the results and for error-free solution. It is mandatory for a system to be robust and reliable.

SERVICE LEVEL AGREEMENTS (SLA's)

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.

As an example, Internet service providers and telco's (telephone companies) will commonly include service level agreements within the terms of their contracts with customers to define the level(s) of service. In this case the SLA will typically have a technical definition in mean time between failures (MTBF), mean time to repair or mean time to recovery (MTTR); identifying which party is responsible for reporting faults or paying fees; responsibility for various data rates; throughput; jitter; or similar measurable details.

In order to develop a well-organized service level agreement, there are six key components should be included:

1. **Agreement Overview:** The agreement overview includes details such as the individuals involved, effective/expiry date as well as a general statement on what other details the particular SLA will cover.

2. **Goals and Objectives:** Here, the purpose of the agreement, including the ability to obtain a mutual agreement, will be outlined.
3. **Stakeholders:** It defines the parties involved in the agreement. For example, an IT service provider and an IT customer.
4. **Periodic Review:** There should be mention of a periodic review, which will outline the effective/expiry date, as well as the parameters regarding review timelines of a particular SLA.
5. **Service Agreement:** It features many key components for which the service provider takes responsibility such as Service scope, Customer requirements, details on payments.
6. **Service Management:** It involves both service availability and service requests. A concise SLA will feature information on the availability of telephone support, response time for service requests, as well as options regarding remote assistance.

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

- Customer-based SLA
- Service-based SLA
- Multilevel SLA or Hierarchical SLA

- **Customer-based SLA**: An agreement with an individual customer group, covering all the services they use. For example, an SLA between a supplier (IT service provider) and the finance department of a large organization for the services such as finance system, payroll system, billing system, procurement/purchase system, etc.
- **Service-based SLA**: An agreement for all customers using the services being delivered by the service provider**.** For example:
  - A mobile service provider offers a routine service to all the customers and offers certain maintenance as a part of an offer with the universal charging.
  - An email system for the entire organization. There are chances of difficulties arising in this type of SLA as level of the services being offered may vary for different customers (for example, head office staff may use high-speed LAN connections while local offices may have to use a lower speed leased line).
- **Multilevel SLA**: The SLA is split into the different levels, each addressing different set of customers for the same services, in the same SLA.

  Corporate level: All of the general issues relevant to the organization are covered, and they are the same throughout the entire organization. For example, with security SLA at the organization level, every employee needs to create passwords of 8 characters and must change it every thirty days—or every employee needs to have an access card with an imprinted photograph.

  Customer level: Those issues specific to a customer can be dealt. Security requirements of one or more departments within the organization are higher. For example, the financial department needs more top security measures by virtue of its crucial role and handling of financial resources.

Service Level: All issues relevant to a specific service (in relation to the customer) can be covered. Applies to all customers that contract the same service — for example, contracting IT support services for everyone who uses a particular IP telephony provider.

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

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

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.

BILLING AND ACCOUNTING IN CLOUD COMPUTING

Cloud billing is the process of generating bills from the resource usage data using a set of predefined billing policies. The bill can be for real money or it can refer to a more abstract notion of exchange, depending on your individual cloud computing general policies.
The cloud billing service module is enabled for a service oriented architecture, covering both functional requirements — a quote service, conversion functions and policies, payment schemes, and user identification — and the non-functional, but essential, requirements such as security, scalability, standards, and fault tolerance.

Parts of a cloud billing service module

1. **Working module**
2. **Billing policies, Billing syntax, Language**
3. **The rules/inference engine and its role**

The general cloud billing service module should provide port types to support all of your functional requirements:

- The billing service generates bills according to policies that it retrieves from a Billing Policies Repository.
- The bills are stored in the Bills Repository for later use by the manager.

The bills generation process is not automated and needs to be initiated by the manager. All the functionalities of the service are implemented in a web service. The billing service establishes the connection to the web server through a PHP interface and gets the record document from the Usage Record (UR) Repository.

However simple or complex you make your billing policies, there will be a time when the resource manager will want to modify them. Capturing every case is simply impossible and changing the source code each time you add new policies is, of course, not feasible.

The billing policies are not part of the service; they constitute an external element that the billing service makes use of. They are written by the resource manager and stored in a repository. The billing service uses this repository to extract the necessary guidelines to generate the final bill.

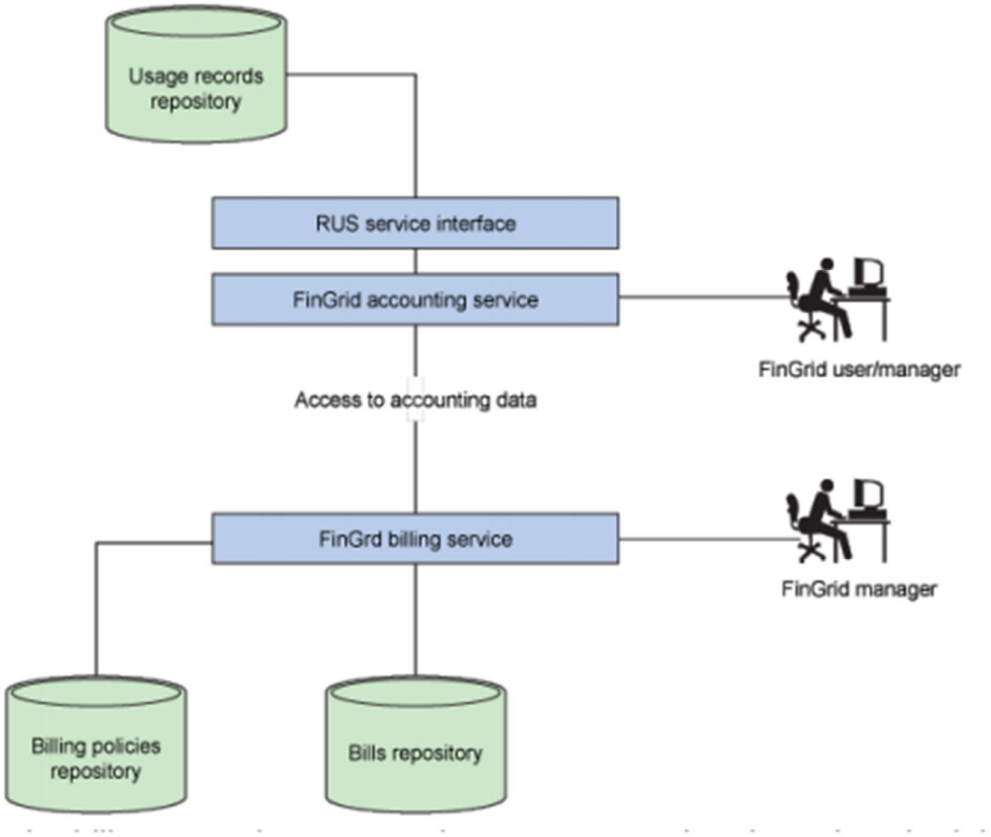


Figure: Architecture of Service Module

There are also different types of languages used to describe the policies; each has its own benefits and drawbacks. A natural language permits the expression of rules in plain English. The natural language is easy to use and has a flat learning curve.

The same policy description in a domain specific language (which uses the same syntax and semantic used in the specification document). The domain specific language is more structured and is easier to make quick edits and changes in than the natural language.

The process of matching the new or existing facts against production rules is called pattern matching. There are a number of algorithms used for pattern matching by rules engines. The rules are stored in the production memory and the facts that the inference engine matches against the working memory. Facts are inserted into the working memory where they are modified or retracted. A system with a large number of rules and facts may result in many rules being true for the same fact assertion; these rules are said to be in conflict.

Cloud accounting is the practice of using an **accounting** system that's accessed through the internet. Cloud accounting software is similar to traditional, on-premises, or self-install accounting software, only the accounting software is hosted on remote servers, similar to the SaaS (Software as a Service) business model. Data is sent into "the cloud," where it is processed and returned to the user.

For one, cloud accounting is more flexible. Accounting data can be accessed from anywhere on any device with an Internet connection, rather than on a few select on-premises computers. Secondly, unlike traditional accounting software, cloud accounting software updates financial information automatically and provides financial reporting in real-time. This means account balances are always accurate and fewer errors take place due to manual data entry. They are also better able to handle multi-currency and multi-company transactions more efficiently.

With cloud accounting, it's also easier to get real-time reporting and visibility throughout your organization, with greater mobile capabilities and collaboration. Subscription-based models are popular among cloud accounting providers, and in most cases these subscriptions are usage-based.

Cloud vs. Traditional Accounting

In the on-premises world, every time a firm grows, they encounter greater software license and maintenance costs as well as new licenses and fees for database, systems management and other software. The firm might also have to make expensive capital purchases of new hardware, such as servers. With cloud solutions, businesses don't get stuck with permanent, expensive equipment and licenses when your business contracts are up and, likewise, there are no big spikes in costs when it expands a little.

Cloud accounting solutions provide an equally secure method of storing financial information than traditional accounting software. For instance, a company computer or laptop with critical financial information could be lost or stolen, which could lead to an information breach. Cloud accounting, however, leaves no trace of financial data on company computers, and access to that data in the cloud is encrypted and password protected.

Sharing data is also less worrisome. With cloud accounting, two people simply need access rights to the same system with their unique passwords. Traditional methods often require flash drives to transport data, which could be lost or stolen.

Also, cloud accounting requires far less maintenance than its traditional counterpart. The cloud provider completes the backups, updates occur automatically and nothing needs to be downloaded or installed on a company computer.

Lastly, cloud providers usually have backup servers in two or more locations. If one server network gets down, you still have access to your data. Information kept just

on-premises could be destroyed or damaged in a fire or natural disaster, and may never be recovered.

There are two types of Cloud Billing accounts:

- **Self-serve (or Online) account:** Payment instrument is a credit or debit card or ACH direct debit, depending on availability in each country or region. Costs are charged automatically to the payment instrument connected to Cloud Billing account.
- **Invoiced (or Offline) account:** Payment instrument can be check or wire transfer. Invoices are sent by mail or electronically. Invoices are also accessible in the Cloud Console, as are payment receipts.

There are two types of payments profiles:

- **Individual:** You're using your account for your own personal payments.
- **Business:** You're paying on behalf of a business, organization, partnership, or educational institution.

Benefits of Cloud Accounting

Cloud-based accounting software now offers all the functionality and reliability of your tried and trusted desktop accounting system, but with a excess of additional benefits that only online technology can deliver. If your business is looking for a more effective way to manage its financial affairs, here are six reasons for seriously considering a move to cloud accounting.

1. Mobile access at any time

With cloud accounting, you can access your accounts and key financial figures at any time, from anywhere.

When you use an old-fashioned, desktop-based system, you are effectively tied to the office. Your software, your data and your accounts are all sat on a local drive. And that limits the access you can have to your financial information.

Cloud-based accounting frees you up from this restriction. Your data and records are all safely encrypted and stored on a cloud server, and there is no software application for you to download – you log in and work from your web browser, wherever you have Wi-Fi and an Internet connection.

So, wherever you are, you can always check on the status of your business.

2. A cost and time-effective solution

Working online reduces your IT costs and saves you time by keeping you constantly connected to the business.

Desktop-based systems require an investment in IT hardware, plus the maintenance of that hardware. You require a server to house the application software and the related data. And you will need to pay an IT expert to maintain both the server and the office network – that can be an expensive overhead.

Online accounting is carried out entirely from the cloud. There is no costly IT infrastructure for you to maintain, and you can access the software whether you are in the office or not, you can immediately approve payments, or send out invoices to customers, saving you time and making your financial processes far more effective.

3. High security and no time-consuming back-ups

When you are cloud-based, your accounts and records are all saved and backed up with high levels of encryption. If you have used desktop accounting, you will be aware of the need to back-up your work at the end of each day. And you will also know about the need for updates each time your provider brings out a new version of the software.

Security is another area where cloud accounting overcomes a desktop system. Your data is no longer sat on a physical server in the office, or on the hard drive of your laptop. All your accounting information is encrypted at source and saved to the cloud. So the only person who can access your confidential information is you, plus selected members of your team and advisers.

4. Share and collaborate with ease

Working with colleagues, and sharing data with your advisers, is an extremely straightforward process when you're based in the cloud.

Using the old, desktop approach, you had limited access to your accounts – and that made collaboration with colleagues and advisers difficult. With a system like online accounting, you, your colleagues, your management team and your advisers can all access the same numbers – instantly, from any geographical location.

5. Reduces paperwork and is more sustainable

Using cloud accounting can deliver the dream of having a paperless office. With traditional accounting, dealing with paperwork, everything must be printed out and dealt with in hard copy, and this is slow, ineffective and bad for the environment.

With an online accounting system, you can significantly reduce your reliance on paperwork. Invoices can be emailed out directly to clients, removing the costs of printing and postage – and speeding up the payment process. And because your documents are all digitized and stored in the cloud, there is no need to keep the paper originals.

6. Better control of your financial processes

The efficiencies of online accounting software give you greatly improved control of your core financial processes. The online invoicing function streamlines the whole invoice process, giving you a better view of expected income, an overview of outstanding debts and a clear breakdown of what each customer owes your business.

COMPARING SCALING HARDWARE: TRADITIONAL Vs. CLOUD

Scaling is defined as the ability for an IT resource to handle growing or decreasing demands in a capable manner. It is one of the most popular and beneficial features of cloud computing, as businesses can scale up or down to meet demands based on season, projects, growth and more.

Traditional IT infrastructures consist of various kinds of hardware's, such as a desktop computer, which are connected to a network via a remote server. Businesses with this IT model must purchase additional hardware and upgrades in order to scale up their data storage and services to support more users. For many businesses with IT data centers, an in-house IT department is needed to install and maintain the hardware.

Cloud computing is far more abstract as a virtual hosting solution. Instead of being accessible via physical hardware or investing money into purchasing physical servers and data storage space in-house, you can rent them from cloud computing providers on a more cost effective pay-per-use basis. It's a real-time virtual environment hosted between several different servers at the same time.

Choosing an IT model for your business is a very important decision. Every company needs a safe and secure storage space, where data and applications can be easily accessed and running costs are kept to a minimum. If you're thinking of migrating your data from traditional IT infrastructure to cloud based platforms, read on to explore the differences between the two, to better understand the benefits of such a move.

There are many reasons why your business should to move to the cloud and away from traditional IT. Getting out of the data center business makes you more efficient, more secure, and provides greater flexibility. The cloud also offers other benefits, such as automation, increased collaboration, efficient disaster recovery, and faster time to market.

The core difference between traditional IT and the cloud is location. With a traditional approach to IT, the technology capabilities are housed on-site. Whereas, Cloud services are hosted off-site. Employees at a company access the programs that they need via an internet connection. This core difference creates many additional differences between the two approaches.

Resilience and Elasticity

The information and applications hosted in the cloud are evenly distributed across all the servers, which are connected to work as one. Therefore, if one server fails, no data is lost and downtime is avoided. The cloud also offers more storage space and server resources, including better computing power. This means your software and applications will perform faster.

Traditional IT systems are not so resilient and cannot guarantee a consistently high level of server performance. They have limited capacity and are susceptible to downtime, which can greatly hinder workplace productivity.

Flexibility and Scalability

Cloud hosting offers an enhanced level of flexibility and scalability in comparison to traditional data centers. The on-demand virtual space of cloud computing has unlimited storage space and more server resources. Cloud servers can scale up or down depending on the level of traffic your website receives, and you will have full control to install any software as and when you need to. This provides more flexibility for your business to grow.

With traditional IT infrastructure, you can only use the resources that are already available to you. If you run out of storage space, the only solution is to purchase or rent another server. If you hire more employees, you will need to pay for additional software licenses and have these manually uploaded on your office hardware. This can be a costly venture, especially if your business is growing quite rapidly.

Automation

A key difference between cloud computing and traditional IT infrastructure is how they are managed. Cloud hosting is managed by the storage provider who takes care of all the necessary hardware, ensures security measures are in place, and keeps it running smoothly. Traditional data centers require heavy administration in-house, which can be costly and time consuming for your business. Fully trained IT personnel may be needed to ensure regular monitoring and maintenance of your servers – such as upgrades, configuration problems, threat protection and installations.

Running Costs

Cloud computing is more cost effective than traditional IT infrastructure due to methods of payment for the data storage services. With cloud based services, you only pay for what is used – similarly to how you pay for utilities such as electricity.

With traditional IT infrastructure, you will need to purchase equipment and additional server space upfront to adapt to business growth. If this slows, you will end up paying for resources you don't use. Furthermore, the value of physical servers decreases year on year, so the return on investment of investing money in traditional IT infrastructure is quite low.

Deployment of Services

When you deploy a new service, traditional IT requires that you purchase a new program for that service. Your company pays an upfront cost, and hired IT personnel to load the program onto the computers in the system, which can take some time depending on the nature of the service to be deployed as well as the size and schedule of your IT team.

Whereas deployment of cloud services is taken care of for you off-site, so there is no need for IT experts at your company to spend valuable time on deployment. Typically with cloud services you pay a small monthly fee rather than paying a large upfront cost. Compared to traditional IT deployment of services, cloud deployment is faster and easier on the budget.

Security

Cloud computing is an external form of data storage and software delivery, which can make it seem less secure than local data hosting. Anyone with access to the server can view and use the stored data and applications in the cloud, wherever internet connection is available.

With traditional IT infrastructure, you are responsible for the protection of your data, and it is easier to ensure that only approved personnel can access stored applications and data.

What kind of infrastructure is a good choice for your business? It depends on what your company does and what are your needs. Nevertheless, more and more organizations today prefer cloud infrastructure.

ECONOMIES OF SCALING: BENEFITTING ENORMOUSLY

Economies of Scale refer to the cost advantage experienced by a firm/organization when it increases its level of output or when production becomes efficient. Companies can achieve economies of scale by increasing the production and lowering the per-unit cost. This happens because costs are spread over a larger number of goods. Costs can be both fixed and variable.

Understanding Economies of Scale

Economies of scale are an important concept for any business in any industry and represent the cost-savings and competitive advantages larger businesses have over smaller ones.

Most consumers don't understand why a smaller business charges more for a similar product sold by a larger company. That's because the cost per unit depends on how much the company produces. Larger companies are able to produce more by spreading the cost of production over a larger amount of goods. An industry may also be able to dictate the cost of a product if there are a number of different companies producing similar goods within that industry.

There are several reasons why economies of scale give rise to lower per-unit costs.

- First, specialization of labor and more integrated technology boost production volumes.
- Second, lower per-unit costs can come from bulk orders from suppliers, larger advertising buys, or lower cost of capital.
- Third, spreading internal function costs across more units produced and sold helps to reduce costs.

For example, if an industry is producing the video games, there is the one-time cost of actually creating the game. As we create more copies of the game, the cost per game decreases as the one-time cost is distributed. We may get lower the costs on CDs and other raw materials, higher the manufacturing efficiency. So, once a company can establish itself, expand the customer base for a specific game, and raise demand for that game. This extra output required to meet that demand lowers overall cost in the long run.

Types of Economies of Scale

Economies of scale can be both internal as well as external. Internal economies of scale are borne from within the company (based on management decisions), while external ones are based on outside factors.

Internal economies: Internal economies of scale happen when a company cuts costs internally, as they're unique to that particular firm. This may be the result of the limited size of a company or because of decisions from the firm's management. Larger companies may be able to achieve internal economies of scale—lowering their costs and raising their production levels—because they can buy resources in bulk, have special technology, or can access more capital.

Examples of internal economies of scale;

Specialization through the division of labor

Workers in larger-scale factories and other such production operations can do more precise, specific jobs. This situation increases economic efficiency as relatively limited training can allow workers to become excellent at their assigned tasks.

Technology

A larger firm may be able to adopt production technologies of production that a smaller firm just can't. That's because large-scale businesses can afford to invest in expensive, specialized capital in the form of specialized machinery and other manufacturing equipment.

Bulk buying power

With greater buying power, a large firm can purchase its factor inputs in bulk at discounted prices. They can buy more from suppliers at a lower price.

Financial

Larger firms tend to be more creditworthy so that they have access to credit with especially favorable rates of borrowing. They have better financing options and lower interest rates. This means they access capital more cheaply. Additionally, large firms listed on stock markets can raise money by selling equity more easily.

Risk-bearing

Investments that might end up being extremely worthwhile may also be quite costly and high-risk. Generally, only large companies can afford to accept such risks and make large investments.

External economies: External economies of scale originate from outside the firm. External economies of scale, on the other hand, are achieved because of external factors, or factors that affect an entire industry. These factors benefit the entire industry, and no single firm has control over these costs. That means no one company controls costs on its own. These occur when there is a highly-skilled labor pool, subsidies and/or tax reductions, and partnerships and joint ventures—anything that can cut down on costs to many companies in a specific industry.

Examples of external economies of scale

Transportation

The presence of larger firms may create better transportation networks. This reduces costs for companies that use these networks.

Skilled labor

High concentrations of skilled labor often appear as workers receive training and education to serve particular firms and industries. One of the best-known examples of this concentration is the Silicon Valley, where there are huge concentrations of programmers drawn to the presence of massive tech firms like Apple and Google.

Benefits: Internal economies of scale are beneficial to the business because decreased costs mean they are able to decrease prices to gain a competitive advantage, or increase profit margins. Technical economies describe a larger company's ability to invest in the latest technology in order to increase efficiency. Managerial economies are when a large company attracts a highly skilled workforce because of the benefits they can offer an employee. Marketing economies are seen in a larger company's ability to gain bulk buying discounts and financial economies occur when a company can attract investors and are able to borrow at lower interest rates. Therefore economies of scale are hugely beneficial to an expanding business due to the competitive advantage that can be achieved.

MANAGING DATA- LOOKING AT DATA

IT infrastructure is becoming increasingly complex and enterprises should look out for scalable data management solutions in order to stay flexible. Companies are extensively adopting the cloud for it offers cost savings, data availability, flexibility, scalability, and more. However, any organization/company wants to leverage the fullest potential of infrastructure and have a possibility to work both in the cloud and on premise storage transitioning easily between the two.

Businesses used to have centralized on-premise data warehouses where information was safe. Yet, as time went by, it became harder to maintain them; highly skilled manpower and greater maintenance fees are required for them. Management methods applied to data that is stored in the cloud differ from the traditional ones since cloud data analytics has to meet the requirements of enhanced cloud data security and integrity. Data migration to the cloud is the real deal that requires a holistic approach.

Data management is an administrative process that includes acquiring, validating, storing, protecting, and processing required data to ensure the accessibility, reliability, and timeliness of the data for its users. It is a broad term that can refer to a role (a data manager), while also referring to an organizational responsibility.

Within the parameters of data management exist responsibility for the entire data lifecycle, from collection to consumption. This includes its point of origin (data provenance) and transformative journey from origination to current point of reference or observation (data lineage). These attributes are particularly useful in managing data.

Data stored in the cloud has its own rules for data integrity and security. Traditional data management methods may not apply to the cloud, so having management in place designed for the particular requirements of the cloud is vital. As with cloud storage, the functions of a cloud data management system will vary between vendors and the conditions outlined in the service-level agreement between the vendor and user.

Best practices for data management in cloud computing

If there are core principles that lay the foundation for the strategic management of data in the cloud and certain pitfalls to avoid, then there must be methods and techniques that are, if compared with the traditional ones, aimed at the operational excellence and overall improvement of your experience.

Figure: Strategic Management of data in the cloud

1. Ensure a sophisticated infrastructure. Everything will work smoothly and efficiently if there is a possibility to choose whether you want to move data to on-premises storages, to the cloud, or across different clouds. The cloud is not the only destination of a mass data migration. The structure has to be sophisticated yet this whole system should have centralized management.

2. Choose your cloud data management platform. Platforms are used for control, monitoring and other relevant cloud activities. Modern enterprises tend to constantly change their IT environments by making them larger and more complex. If you do provide such an infrastructure managing different types of data across various cloud computing services and local servers, then selecting a single platform is highly recommended. This platform approach will help you maintain a certain level of consistency and reduce bottlenecks. Besides you can opt for a platform that is native, cloud provider-specific, or available from a third-party vendor.

3. Leverage the Cloud Data Management Interface. It is a generally accepted standard of interface's functioning which allows enterprises to manage data elements increasing the system's interoperability. Accommodation of requirements from multiple vendors instead of using the storage system with a unique interface might be challenging, so the deployment of CDMI compatible systems is the right thing to do.

4. Create a framework for cloud management. Before moving data to the cloud, make sure there is a solid framework. Upon having one established, it will be easier for an enterprise to say how to best manage its cloud resources. Migration of systems to more capable platforms is a natural process, but it has to be a conscious and informed decision.

Challenges and opportunities

The majority of cloud data management challenges are the same drawbacks cited for cloud technologies in general, particularly security and costs. As long as security is a concern with cloud technologies, it will be considered a barrier to cloud data management adoption. However, as more strides are made in cloud-based technologies and it becomes more secure, that reluctance could fade.

The benefits of using cloud data management include consolidation of processes such as backup, disaster recovery, archiving and analytics, and cost savings. Some cloud data management companies also offer ransom-ware protection, keeping data and applications native to the platform in a secure, immutable format.

Cloud data management companies and products

Rubrik, which brands itself as "the Cloud Data Management Company," is considered a major data management player in the cloud computing space. The vendor's Cloud Data Management scale-out platform uses a single interface to manage data across public and private clouds. In addition to offering ransom-ware recovery, Rubrik's platform is the first one to support hybrid cloud environments.

Other vendors in the cloud data management space include Commvault, which has a cloud data management platform that supports multiple clouds and on-premises data; and VMware and its vRealize Suite that delivers cloud data management for hybrid cloud environments. Komprise and Red Hat's CloudForms platform also focus on cloud-based data management.

You can have the best products and employees in the world, but without data they are powerless, so take steps to ensure it flows freely and safely. Smart data management will empower your staff to leverage the latest cloud technologies, innovate new products and services and differentiate your organization from the competition.

SCALABILITY & CLOUD SERVICES

A scalable system is able to rise or shrink its performance, resources and functionalities according to user's needs. Scalability is one of the key features of Cloud Computing solutions, one of the reasons why Cloud has been so successful on the market and it will keep increasing.

Scalability in cloud computing is the ability to quickly and easily increase or decrease the size or power of an IT solution. This is one of the most valuable and predominant feature of cloud computing. Through scalability you can scale up your data storage capacity or scale it down to meet the demands of your growing business.

Scaling in the cloud provides you the best experience of flexibility of time and money for your business. When business demands are increasing, you can easily add nodes to increase your storage space, or you can increase the number of servers currently used.

When the increased demand is reduced then you can move back to your original configuration.

To effectively leverage scalability you need to understand the complexity and the types of scalability. Virtualization is what makes scalability in cloud computing possible.

Types of scalability

Scale Vertically – Scale Up:

Vertical Scaling or Scaling up is easy, it can be done by moving the application to bigger virtual machines deployed in the cloud or you can scale up by adding expansion units as well with your current infrastructure. This ability to add resources to accommodate increasing workload volumes is vertical scaling. It can resize your server with no change in your code.

The downside to scaling up is that it increases storage capacity but the performance is reduced because the compute capacity remains the same. Workloads requiring higher throughput demand reduced latency and this can only by fulfilled by Horizontal Scaling / Scaling out.

Scale Horizontally – Scale out:

Horizontal Scaling or Scaling out is the addition of nodes to the existing infrastructure to accommodate additional workload volumes. The total workload volume is aggregated over the total number of nodes and latency is effectively reduced. This scaling is ideal for workloads that require reduced latency and optimized throughput.

Contrary to Vertical Scaling, Horizontal Scaling also delivers performance along with storage capacity.

Scale Diagonally:

Diagonal scaling helps you combine the scaling up and scaling down. As the term suggests, scaling down is the removal of storage resources as requirements decrease. Diagonal scaling delivers flexibility for workload that requires additional storage resources for specific instances of time. For instance, a website sets up diagonal scaling; as the traffic increases, the compute requirements are accommodated. As the traffic decreases, the computation capacity is restored to its original size.

This type of scaling introduces enhanced budgeting and cost effectiveness for environments and businesses dealing with variable workload volumes.

A practical example: Containers

Containers are a great example of Cloud infrastructure with maximum scalability: they enable both horizontal and vertical scaling and represent the most innovative and performing Cloud solution currently on the market. Containerized systems recognize automatically when increasing of computing power it's needed, as well as is requested a decrease, so that the system can be maximum efficient without waste.

Scalable Cloud Based Services

- Infrastructure-as-a-Service (IaaS)
- Platform-as-a-Service (PaaS)
- Storage-as-a-Service (STaaS)
- Data-as-a-Service (DaaS)
- Database-as-a-Service (DBaaS)

Benefits of cloud scalability

Scalability is the key concept for anyone in the **DevOps** world. Let's briefly discuss the benefits of scalability.

Performance: One core benefit of scalability in the cloud is that it facilitates performance. Scalable architecture has the ability to handle the bursts of traffic and heavy workloads that will come with business growth.

Cost-efficient: You can allow your business to grow without making any expensive changes in the current setup. This reduces the cost implications of storage growth making scalability in the cloud very cost effective.

Easy and Quick: Scaling up or scaling out in the cloud is simpler; you can commission additional VMs with a few clicks and after the payment is processed, the additional resources are available without any delay.

Capacity: Scalability ensures that with the continuous growth of your business the storage space in cloud grows as well. Scalable cloud computing systems accommodate your data growth requirements. With scalability, you don't have to worry about additional capacity needs.

DATABASE & DATA STORES IN CLOUD

A cloud database is a collection of informational content, either structured or unstructured that resides on a private, public or hybrid cloud computing infrastructure platform. From a structural and design perspective, a cloud database is no different than one that operates on a business's own on-premises servers. The critical difference lies in where the database resides. Where an on-premises database is connected to local users through a corporation's internal local area network (LAN), a cloud database resides on servers and storage furnished by a cloud or database as a service (DBaaS) provider and it is accessed solely through the internet. Database services take care of scalability and high availability of the database.

Types of Cloud Databases

Two cloud database environment models exist:

- **Traditional cloud model**
- **Database as a service (DBaaS) model**

In a **traditional cloud model**, a database runs on an on-site infrastructure with a virtual machine. Cloud platforms allow users to purchase virtual-machine instances for a limited time, and one can run a database on such virtual machines.

The **DBaaS model** is a fee-based subscription service in which the database runs on the service provider's physical infrastructure. With a database as a service model, application owners do not have to install and maintain the database themselves. Instead, the database service provider takes responsibility for installing and maintaining the database, and application owners are charged according to their usage of the service.

SQL databases are one type of database which can run in the cloud, either in a virtual machine or as a service, depending on the vendor. While SQL databases are easily vertically scalable, horizontal scalability poses a challenge.
NoSQL databases are another type of database which can run in the cloud. NoSQL databases are built to service heavy read/write loads and can scale up and down easily and therefore they are more natively suited to running in the cloud. The NoSQL databases have proven to provide efficient horizontal scalability, good performance, and ease of assembly into cloud applications.

Datastore:

Datastore is a NoSQL document database built for automatic scaling, high performance, and ease of application development. Cloud Datastore automatically handles sharding* and replication*, providing you with a highly available and durable database that scales automatically to handle your applications load. Cloud Datastore provides a numerous of capabilities such as ACID transactions, SQL-like queries, indexes, and much more.

Data store features include

***Atomic transactions**:* Ensure the integrity of your data by executing multiple datastore operations in a single transaction with ACID characteristics, so all the grouped operations succeed or all fail.

Massive scalability with high performance: Datastore uses a distributed architecture to automatically manage scaling. Data store uses a mix of indexes and query constraints so your queries scale with the size of your result set, not the size of your data set.

Flexible storage and querying of data: Datastore maps naturally to object-oriented and scripting languages, and is exposed to applications through multiple clients. It also provides a SQL-like query language.

Balance of strong and eventual consistency: Datastore ensures that entity lookups by key and ancestor queries always receive strongly consistent data. All other queries are eventually consistent.

Encryption at rest: Datastore automatically encrypts all data before it is written to disk and automatically decrypts the data when read by an authorized user.

Multiple access methods: Access your data via our JSON API, open source clients, or community maintained ORMs (Objectify, NDB).

Fully manage: Cloud Datastore is fully managed, which means Google automatically handles sharding and replication in order to provide you with a highly available and consistent database.

Note: Firestore, the newest version of Datastore, makes all queries strongly consistent and introduces several improvements over Datastore.

*Sharding is partitioning where the database is split across multiple smaller databases to improve performance because it can improve both read and write performance.

*In replication, we basically copy the database across multiple databases to provide a quicker look and less response time. It can greatly improve read performance but does little for applications that have a lot of writes.

LARGE SCALE DATA PROCESSING IN CLOUD COMPUTING

The computer industry is being challenged to develop methods and techniques for affordable data processing on large datasets at optimum response times. The technical challenges in dealing with the increasing demand to handle vast quantities of data is daunting and on the rise. One application area for cloud computing is large-scale data processing. Today, the need to process large amount of data has been enhanced in the area of Engineering, Science, Commerce and the Economics of the world.

Many organizations face difficulties when dealing with a large amount of data. They are unable to manage, manipulate, process, share and retrieve large amounts of data by traditional software tools due to them being costly and time-consuming for data processing. The term large-scale processing is focused on how to handle the applications with massive datasets. Such applications devote the largest fraction of execution time to movement of data from data storage to the computing node in a computing environment. The main challenges behind such applications are data storage capacity and processor computing power constrains. Developers need hundreds or thousands of processing nodes and large volume of storage devices to process complex applications with large datasets, such applications process multi-terabyte to petabyte-sized datasets and using traditional data processing methods like sequential processing and centralized data processing are not effective to solve these kinds of application's problems.

A lot of researchers have been interested in scale-out (horizontal scaling) technology and have started research and development (R&D) activities as a way to deal with large-volume data analysis by adding blade/rack servers to increase processing speed and expand data capacity. NTT (Nippon Telegraph and Telephone Corporation is a Japanese telecommunications company) Information Sharing Platform Laboratories has also been developing a large-scale distributed data processing platform called CBoC Type 2 (CBoC: Common IT Bases over Cloud Computing; IT: information technology).

Analysis of such huge data (known as big data) in a conventional manner becomes exponentially costly even when the data is collected by the system, so the data has been either stored wastefully or discarded. The advent of scale-out technology, however, has reduced the cost of constructing systems for processing large-scale data, and new advanced services such as personalization based on analytical results are now possible.
Large-scale data processing enables the use of diverse types of big data in a cloud environment in order to create mash-up services. A large-scale distributed data processing platform collects and stores the big data produced by IT systems or the Internet. By analyzing such large volumes of data, one can acquire new knowledge and expertise and create new mash-up services. A large-scale distributed data processing platform is expected to serve as a platform for creating knowledge on which to base advanced services for customers.

One of the recent processing models with a more efficient and intuitive solution to rapidly process large amount of data in parallel is called MapReduce. It is a framework defining a template approach of programming to perform large-scale data computation on clusters of machines in a cloud computing environment. MapReduce provides automatic parallelization and distribution of computation based on several processors. It hides the complexity of writing parallel and distributed programming code.

MapReduce:

One of the best ways for huge data processing is to perform parallel and distributed computing in a cloud computing environment. Cloud computing as a distributed computing paradigm aims at large datasets to be processed on available computer nodes by using a MapReduce framework. MapReduce is a software framework introduced to the world by Google in 2004; it runs on a large cluster of machines and is highly scalable. It is a high-performance processing technique to solve large-scale dataset problems. MapReduce computation processes petabyte to terabyte of unit data on thousands of processors. Google uses MapReduce for indexing web pages. Its main aim is to process large amount of data in parallel stored on a distributed cluster of computers. It is a basis to take advantage of cloud computing paradigm as a new realistic computation industry standard.
MapReduce program works in two phases, namely, Map and Reduce. Map tasks deal with splitting and mapping of data while Reduce tasks shuffle and reduce the data.

Map Function:

This function accepts a chunk of input data and produces an intermediate key and value pairs. Map (in_key, in_value) → list (out_key, intermediate_value) All output intermediate key and value pairs from the Map function become integrated according to a common and known intermediate key K between them and then they are passed into the Reduce function. The Map function can be executed in parallel on non-overlapping portions of the input data.

Reduce Function:

The Reduce function takes an intermediate key K and all the set of values for that intermediate key; after that it combines these values to produce a small set of values.

Reduce (out_key, list (intermediate_value)) → list (out_value) the result of Reduce function is a zero or a single output value per function call. The Reduce function can be executed in parallel on each set of intermediate pairs with the same key.

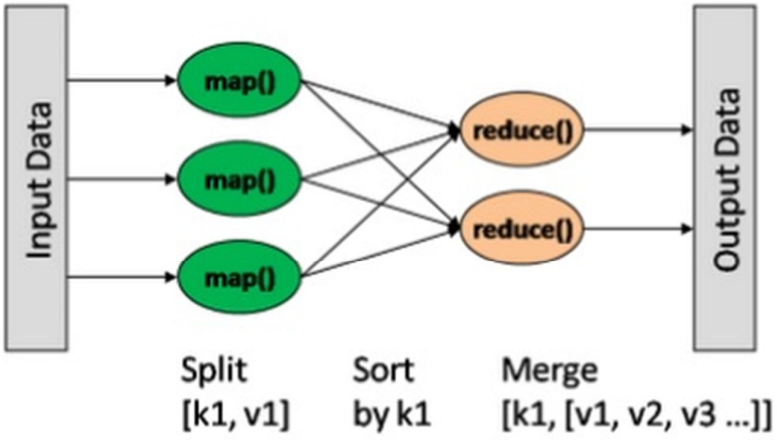


Fig. The MapReduce programming model

Consider you have following input data for your Map Reduce Program;

Welcome to Hadoop Class
Hadoop is good
Hadoop is bad

Hadoop:

Hadoop is an open-source, java based Apache Software foundation project. It is one of the most interesting third-party implementations of MapReduce for distributed computing. Hadoop provides a framework to support large dataset distributed computing on a cluster of servers. Hadoop was inspired in 2006 by Doug Cutting (named by his son's stuffed elephant) now being used by major companies, including Amazon, IMB, Yahoo, Facebook and a growing number of other companies. Hadoop provides its own distributed file system called Hadoop Distributed File System (HDFS), HDFS is a distributed file system designed to store multiple copies of data block on a cluster of compute nodes, to enable reliable and rapid computations. Hadoop runs in a distributed environment. Hadoop uses HDFS to store a dataset and applies the MapReduce's power to distribute and parallelize the processing of this dataset. Stream data is first fragmented into typical HDFS-sized block (64MB) and then smaller packets (64KB) by the user. It means a file into HDFS is divided into 64 MB chunk and each chunk is resided on a different working machine.

A Major web company, Amazon web services platform offers a service called Amazon Elastic MapReduce to store and process massive datasets by running MapReduce system on Amazon cloud. It utilizes a hosted Hadoop framework running on the web-scale infrastructure of Amazon Elastic Compute Cloud (Amazon EC2) and Amazon Simple Storage Service (Amazon S3). EC2is a web service platform that provides resizable compute capacity in a cloud. Amazon S3 provides a simple web services interface that can be used to store and retrieve any amount of data, at any time, from anywhere on the web.

CASE STUDIES:

Eucalyptus: (Elastic **U**tility **C**omputing **A**rchitecture for **L**inking **Y**our **P**rograms **T**o **U**seful **S**ystems.)

EUCALYPTUS is an open-source software framework for cloud computing that implements Infrastructure as a Service (IaaS) in a private or hybrid cloud computing environment; systems that give users the ability to run and control entire virtual machine instances deployed across a variety physical resources. EUCALYPTUS enables users familiar with existing Grid and HPC systems to explore new cloud computing functionality while maintaining access to existing, familiar application development software and Grid middle-ware.

It is a popular way of integrating public and private clouds using Eucalyptus. Eucalyptus is an open source software platform that implements IaaS-style cloud computing using the Linux-based infrastructure found in many modern data centers. While it can be deployed solely for private clouds, because it is interface-compatible with Amazon's AWS, it is possible to move workloads between AWS and the data center without code modification.

Eucalyptus was founded out of a research project in the Computer Science Department at the University of California, Santa Barbara, and became a for-profit business called Eucalyptus Systems in 2009. Eucalyptus Systems announced a formal agreement with Amazon Web Services (AWS) in March 2012, allowing administrators to move instances between a Eucalyptus private cloud and the Amazon Elastic Compute Cloud (EC2) to create a hybrid cloud. The partnership also allows Eucalyptus to work with Amazon's product teams to develop unique AWS-compatible features.

Eucalyptus features include:

- Supports both Linux and Windows virtual machines (VMs).
- Application program interface- (API) compatible with Amazon EC2 platform.
- Compatible with Amazon Web Services (AWS) and Simple Storage Service (S3).
- Works with multiple hypervisors including VMware, Xen and KVM.
- Can be installed and deployed from source code or DEB and RPM packages.
- Internal processes communications are secured through SOAP and WS-Security.
- Multiple clusters can be virtualized as a single cloud.
- Administrative features such as user and group management and reports.

Version 3.3, which became generally available in June 2013, adds the following features:

- **Auto Scaling:** Allows application developers to scale Eucalyptus resources up or down based on policies defined using Amazon EC2-compatible APIs and tools
- **Elastic Load Balancing:** AWS-compatible service that provides greater fault tolerance for applications
- **Cloud Watch:** An AWS-compatible service that allows users to collect metrics, set alarms, identify trends, and take action to ensure applications run smoothly
- **Resource Tagging:** Fine-grained reporting for show back and chargeback scenarios; allows IT/ DevOps to build reports that show cloud utilization by application, department or user
- **Expanded Instance Types:** Expanded set of instance types to more closely align to those available in Amazon EC2. It was 5 before, now up to 15 instance types.
- **Maintenance Mode:** Allows for replication of a virtual machine's hard drive, evacuation of the server node and provides a maintenance window.

Microsoft Azure:

Microsoft Azure, formerly known as Windows Azure, is Microsoft's public cloud computing platform. Windows Azure is software *cum* services platform in the cloud providing a range of cloud services including compute, analytics, networking, hosting, management, and scalable storage, as well as it provides a management infrastructure for provisioning and geo-distribution of cloud-based services, and a development platform for the Azure Services layer. **Azure** provides stronger and faster PaaS capabilities which nowadays are more important part of Cloud infrastructure.

Windows Azure is an open platform that will support both Microsoft and non-Microsoft languages and environments. Windows Azure is not grid computing, packaged software, or a standard hosting service. It is an integrated development, service hosting and management environment maintained at Microsoft datacenters.

The environment includes a robust and efficient core of compute and simple storage capabilities and support for a rich variety of development tools and protocols.

Components of MS Azure;

Compute:

- Component that runs running applications.
- Support applications that have a very large number of simultaneous users and that can scale out. So, it can scale up scale down things.

Storage:

- Storing and accessing data.
- Applications request storage as a simple as simple blobs, a more structure way to store information or, a way to exchange data between different parts of the applications.

Fabric:

- Manage the underlying resources.
- Provide a common way to manage, monitor applications that use the cloud platform.

Types of roles you can implement in Windows Azure are:

Web and Worker Roles: A Windows Azure application consists of one or more hosted roles running within the Azure data centers. Typically there will be at least one Web role that is exposed for access by users of the application. The application may contain additional roles, including Worker roles that are typically used to perform background processing and support tasks for Web roles.

Data management services and features available for storing and managing data:

Azure Storage: This provides four core services for persistent and durable data storage in the cloud. The services support a REST interface that can be accessed from within Azure-hosted or on-premises (remote) applications.

- **Azure Table Service** provides a table-structured storage mechanism based on the familiar rows and columns format.
- **Binary Large Object (BLOB) Service** provides a series of containers aimed at storing text or binary data. It provides both Block BLOB containers for streaming data, and Page BLOB containers for random read/write operations.
- **Queue Service** provides a mechanism for reliable, persistent messaging between role instances (such as between a Web role and a Worker role).
- **Windows Azure Drives** provide a mechanism for applications to mount a single volume NTFS VHD (Virtual Hard Disk) as a Page BLOB, and upload and download VHDs via the BLOB.

SQL Azure Database: This is a highly available and scalable cloud database service built on SQL Server technologies, and supports the familiar T-SQL based relational database model. It can be used with applications hosted in Windows Azure, and with other applications running on-premises or hosted elsewhere.

Networking Services provided by Azure:

Windows Azure provides several networking services that you can take advantage of to maximize performance, implement authentication, and improve manageability of your hosted applications. These services include the following:

- Content Delivery Network (CDN)
- Virtual Network Connect
- Virtual Network Traffic Manager
- Access Control
- Service Bus

AMAZON EC2: (Amazon Elastic Compute Cloud)

Amazon Elastic Compute Cloud (Amazon EC2) is a web service that provides secure, resizable compute capacity in the cloud. It is designed to make web-scale cloud computing easier for developers. Amazon EC2's simple web service interface allows you to obtain and configure capacity with minimal friction.

EC2 is a pay-as-you-go cloud platform that includes compute power, storage and database services as its core components. Users select the CPU, memory, storage, networking capacity and access controls, OS, security, and additional software needed to run a virtual computing environment, or instance.

AWS Elastic Compute Service or EC2 is IaaS(Infrastructure as a Service). This is because Amazon takes the responsibility of networking, storage, server and virtualization and the user is responsible for managing the Operating System, middleware, runtime, data and application.

Key components of AWS EC2 service:

1. INSTANCES: Instances are basically servers that are hosted in AWS cloud using EC2 service. For e.g., Launch Templates, On-demand instance, Spot Instances, Reserved Instances, Dedicated Instance.

2. IMAGES: AMI provides you template with OS and application pre-configured to reduce the instance deployment. You can bundle a Windows instance to create a Windows instance store-backed AMI.

3. ELASTIC BLOCK STORE: EBS volume is a durable, block-level storage device that you can attach to a single EC2 instance. You can back up the data on your Amazon EBS volumes to Amazon S3 by taking point-in-time snapshots. Snapshots are incremental backups, which mean that only the blocks on the device that have changed after your most recent snapshot are saved.

4. NETWORK & SECURITY: A security group acts as a virtual firewall that controls the traffic for one or more instances. An Elastic IP address is a static IPv4 address designed for dynamic cloud computing. You can connect to the internet using elastic IP. Amazon EC2 uses public–key cryptography to encrypt and decrypt login information. Public–key cryptography uses a public key to encrypt a piece of data, such as a password, then the recipient uses the private key to decrypt the data. The public and private keys are known as a key pair. An elastic network interface is a logical networking component in a VPC that represents a virtual network card. Identity & Access Management or IAM role is used for access management of AWS. Virtual Private Cloud or VPC enables you to launch AWS resources into a virtual network that you've defined. This virtual network closely resembles a traditional network that you'd operate in your own data center, with the benefits of using the scalable infrastructure of AWS.

5. **LOAD BALANCING:** Load Balancing distributes the incoming application or network traffic across multiple targets, such as Amazon EC2 instances, containers, and IP addresses, in multiple Availability Zones. Load Balancing scales your load balancer as traffic to your application changes over time and can scale to the vast majority of workloads automatically. Target groups are used to direct traffic from a load balancer to a specific destination.

6. **AUTO SCALING:** A launch configuration is a template that an Auto Scaling group uses to launch EC2 instances. When you create a launch configuration, you specify information for the instances such as the ID of the Amazon Machine Image or AMI, the instance type, a key pair, one or more security groups, and a block device mapping. An Auto Scaling group contains a collection of EC2 instances that share similar characteristics and are treated as a logical grouping for the purposes of instance scaling and management.

7. **MONITORING:** Amazon Cloud Watch monitors your AWS resources and the applications you run on AWS in real time. You can use Cloud Watch to collect and track metrics, which are variables you can measure for your resources and applications. Cloud Watch alarms send notifications or automatically make changes to the resources you are monitoring based on rules that you define. Amazon EC2 performs automated checks on every running EC2 instance to identify hardware and software issues. You can view the results of these status checks to identify specific and detectable problems.

8. **SYSTEM MANAGER:** AWS Systems Manager is a collection of capabilities for configuring and managing your Amazon EC2 instances, on-premises servers and virtual machines, and other AWS resources at scale. Systems Manager includes a unified interface that allows you to easily centralize operational data and automate tasks across your AWS resources.

9. **Region & Availability Zone:** Amazon EC2 is hosted in multiple locations worldwide. These locations are composed of regions and Availability Zones. Each region is a separate geographic area. Each region has multiple, isolated locations known as Availability Zones.

Amazon EC2 provides you with the ability to place resources, such as instances, and data in multiple locations. Resources aren't replicated across regions unless you do so specifically.