ENTERPRISE COMPUTING

ELECTIVE I
CT72507

(IV/I)

Evaluation Scheme:

The questions will cover all the chapters of the syllabus. The evaluation scheme will be as indicated in the table below:

Chapters	Hours	Marks Distribution*
1	4	7
2	4	7
3	4	7
4	6	11
5	5	9
6	6	11
7	6	11
8	3	5
9	3	5
10	4	7
Total	45	80

*There may be minor deviation in marks distribution

Contents

1. Enterprise Computing: A Retrospective

1.1. Introduction

1.1.1. Enterprise Computing

- business-oriented information technology that is critical to a company's daily operations.
- information technology (IT) tools that businesses use for efficient production operations and back-office support.
- Can be a single computing system or as an integrated suite of IT tools
- Computing resource for data processing in large organizations that encompass a variety of operating systems, protocols, and network architectures

1.1.2. Retrospective (History of Enterprise computing)

- The use of computers for enterprise data processing began in the 60s with the early mainframe computers
- computing paradigms have changed dramatically with the emergence of new technology
- In the 80s mainframe is replaced by the Client-server due to the availability of "cheap" computing resources
- 90's Rise of the internet led to Web-based Enterprise application
- improved in terms of scale and ubiquity of access at expense of complexity & cost.
- Now cloud computing offers cheaper and easier access to enterprise hardware and software with minimum complexity.

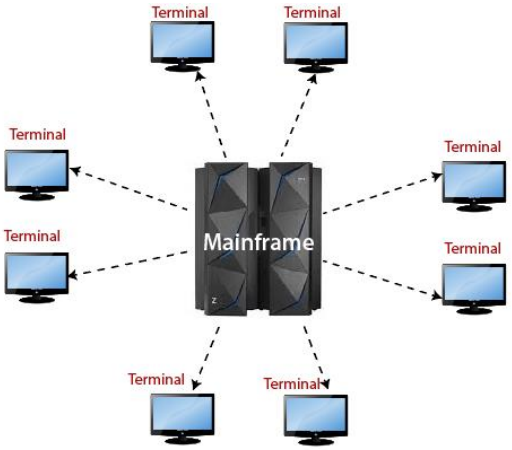
1.1.3. Terms related to Enterprise and Enterprise Computing

- **Enterprise Information** is all the information gathered during day to day operation in Enterprise
- **IS (Information System)** is a set of hardware, software, data, people, and procedures that work together to produce information
- **Information Architecture** is the overall technology strategy for an enterprise
- **Transaction processing systems (TPS)** collect and process data from day to day transactions in an enterprise.
- **Management information system (MIS)** Generates accurate, timely, and organized information so managers can track progress and make decisions
- **Decision support system (DSS)** output from MIS is fed as input and DSS helps in taking decisions.
- **Executive information system (EIS)** supports the needs of executive management for policy and decision-making.
- **The expert system** captures and stores the knowledge of human experts and imitates human reasoning and decision making. Eg: AI, Knowledge Management.
  Two parts:
  - Knowledge base: subject knowledge collected from human experts
  - Inference rules: Logical judgment applied to the knowledge base.
- **Enterprise resource planning (ERP)** provides software applications to coordinate activities of functional units.
- **Customer relationship management (CRM)** manages information about customers, interactions with customers, past purchases, and interests. Used by Sales, Marketing and CSD.
- **A content management system (CMS)** is a combination of databases, software and procedure that organise and allow access to documents and other files
- The technology used in Enterprise computing:
  - Portals, Extranet
  - VPN, Internet, Web services
  - Data warehouse, workflow
- **Virtualization** is the practice of sharing or pooling computing resources

1.1.4. Computing Technologies

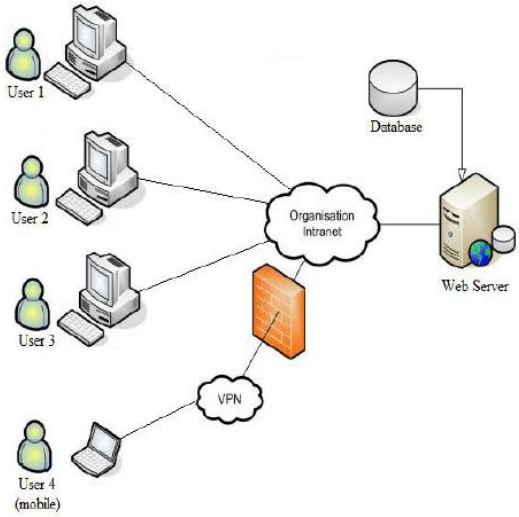
- **Distributed Computing** is networked computers that communicate and coordinate action by passing messages thus sharing computing tasks assigned to the system.
- **Grid Computing** is a group of networked computers that work as a virtual supercomputer to perform large tasks.
- **Edge Computing** is the deployment of computing and storage resources of data

1.2. Mainframe Architecture



- Consist of a mainframe **computer** capable of handling numerous applications and I/O devices simultaneously.
- And **terminals** having limited hardware like display and i/p and o/p only. It is controlled and all processing is handled by the central mainframe. Use **virtual telecommunications access method (VTAM)** to communicate with the mainframe.
- Limited CPU power & I/O bandwidth compared to modern standards
- The application was built using a batch architecture
- Application data was stored either in structured files or in database systems
- The storage subsystem is called '**virtual storage access mechanism' (VSAM)**
- Mainframe systems could run many independent 'guest' operating systems

1.3. Client-Server Architecture

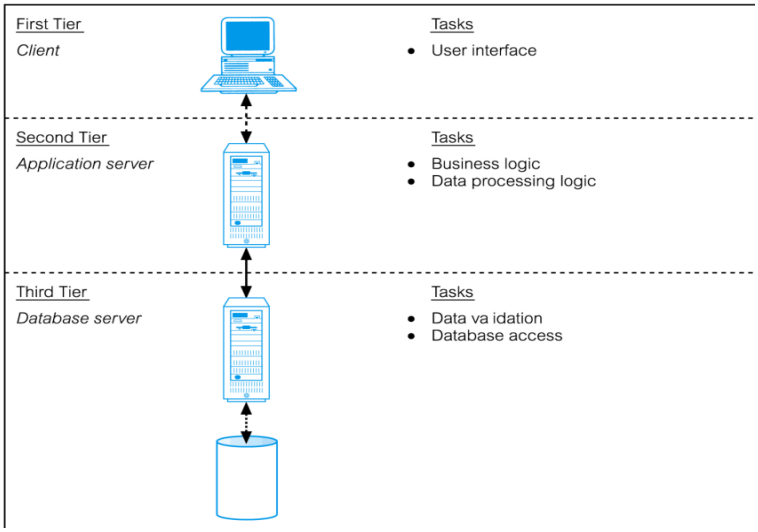


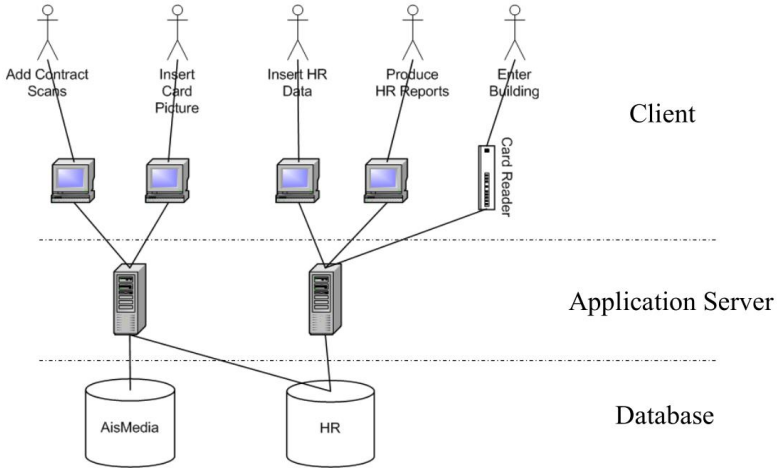
- Microprocessor Revolution in the 80s bring PCs and the development of minicomputers
- has one or more client computers connected to a central server over a network or internet connection.
- A cheaper and easier alternative to the mainframe.
- Here client computers have some level of processing capability hence capable to perform validation and logical functions. And ultimately reduce processing overhead to central computers.
- Client and server have routine and distinct tasks to perform.
- Client and server could be using different OS and may belong to different vendors.
- Applications interact directly with the transport layer protocol to communicate.
- The server houses and provides high-end, computing-intensive services to the client on demand.
- Client/server architecture works when the client computer sends a resource or process request to the server over the network connection, which is then processed and delivered to the client
- A server computer can manage several clients simultaneously, whereas one client can be connected to several servers at a time, each providing a different set of services.
- Different from P2P in a sense that in P2P any system can have the role of Client and server at any time. But in the client-server model, the roles are fixed.
- **Advantages:**
  - Improved data sharing
  - Integration of services
  - Shared resources
  - Security
- **Limitation:**
  - Cant handle numerous traffic at one, Overload
  - Since centralized architecture, Failure of Server would disrupt the whole It infrastructure.

- With the development and growth in the size of Enterprise, the Client-server model became more and more costly and complicated.
- With deployment in WAN, enterprises require regional servers to replicate the main server application and functionalities. Create issues during maintenance and upgrading too.

1.4. 3-Tier Architectures with TP Monitors

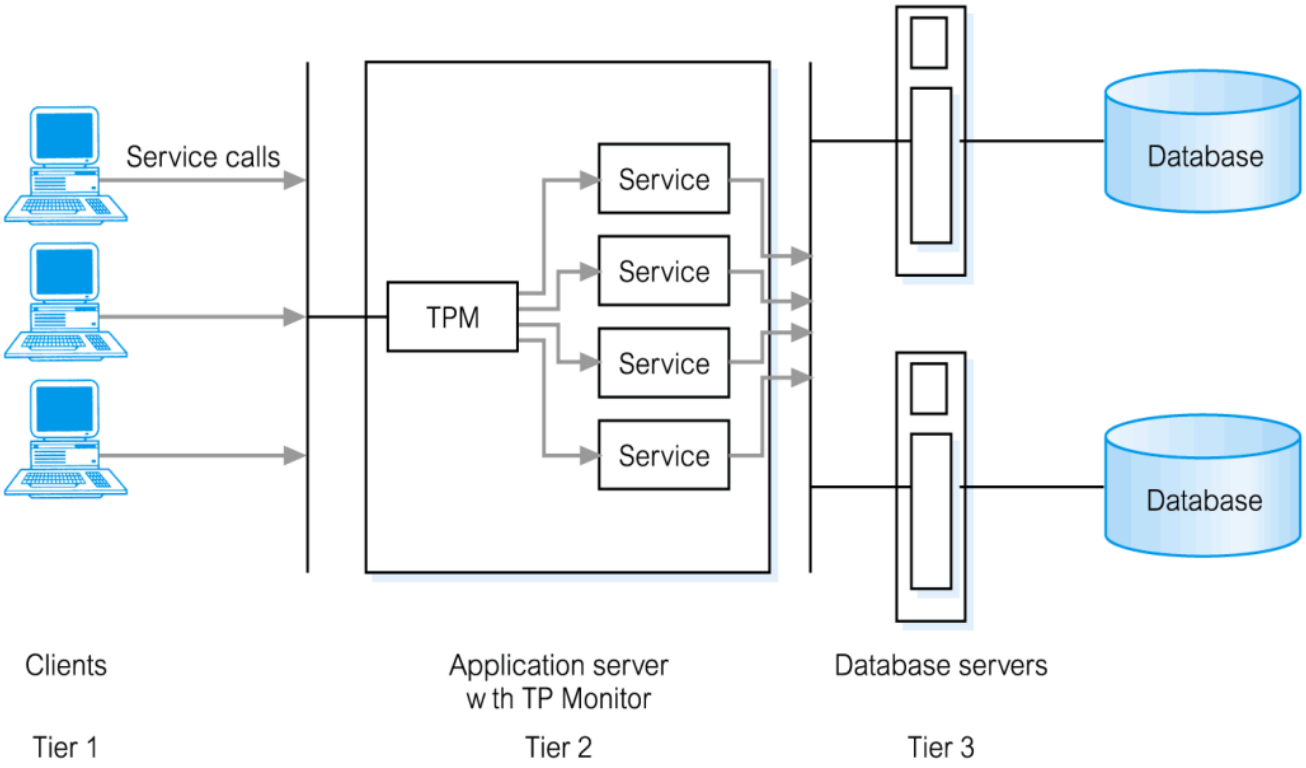
- In the late 90s concepts of 3 tier evolved.
- Modular Client-server architecture having :
  - Presentation Tier (User)
  - Application Tier (Server)
  - Data-tier (database)





- **The presentation tier** is a GUI that communicates with the application and data tier. display information to and collect information from the user. Html, JS, CSS
- **The application tier** handles business logic and can also add, delete or modify data in the data tier. communicates with the data tier using API calls. Python, Java, Ruby
- **Data-tier** stores and manage information, relational and non-relational databases like MySQL, MongoDB
- Each tier has its infrastructure and can be independently maintained and updated.
- **Advantages**:
  - Faster development and deployment
  - Improved scalability, security and reliability
- **Disadvantages**:
  - Increased Complexity
  - Performance depends on database and application server
  - Lack of proper standard
- **Middleware**:
  - Integral Services that must be present in the system but does not necessarily implement business logic
  - Load balancing
  - System integration
  - Availability
  - Logging
  - Threading
  - Pooling
  - Caching
- **Transaction processing Monitors**
  - A program that monitors transactions from one stage to the next, ensuring that each one completes successfully
  - A program that controls data transfer between clients and servers to provide a consistent environment

- purpose/objective is to allow resource sharing and assure optimal use of the resources by applications.
  - Coordinating resources
  - Balancing loads
  - Creating new processes as/when needed
  - Providing secure access to services
  - Routing services
  - Wrapping data messages into messages
  - Unwrapping messages into data packets/structures
  - Monitoring operations/transactions
  - Managing queues
  - Handling errors through such actions as process restarting
  - Hiding interprocess communications details from programmers
- The client connects to TP Monitors instead of the DB server
- The transaction is accepted by the monitor which queues it and takes responsibility to complete it by freeing up the client

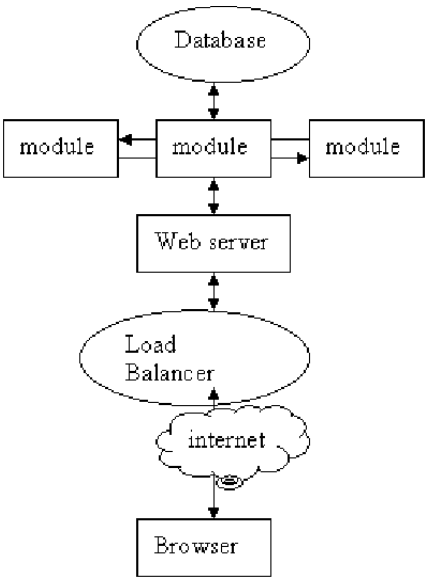


2. Evolution of Computing

2.1. Internet Technology and Web-Enabled Applications

2.1.1. Internet Technology

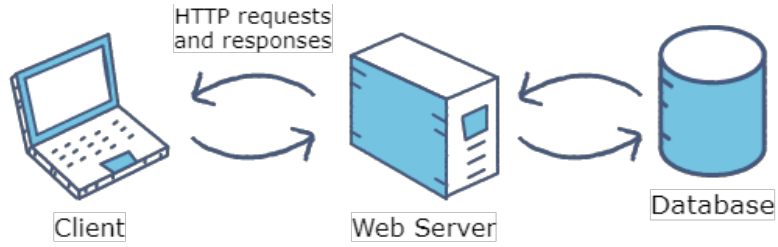
- **The Internet** is a system architecture that allows various computer networks around the world to interconnect.
- Have its seed in ARPANET by DOD, USA.
- Officially born in 1983 after TCP/IP was standardized.
- And available to the public in the early 90s and boomed due to the availability of PCs
- Support communication, access to digital information www, E-business and services over the internet.

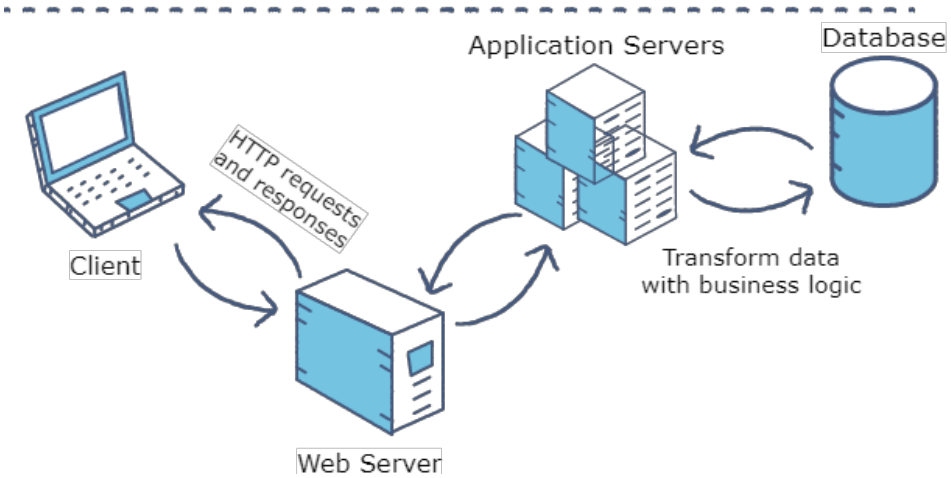


2.1.2. Web-Enabled Application

- Product or service that can be used through, or in conjunction with, the World Wide Web
- A program that does some (or all) of its processing on the user's computer before uploading the output to the Web.
- Enables seamless integration of information exchanges
  - Within the business
  - Between business and customers/suppliers
- It flourished only after the World Wide Web project and HTML by Tim Berners lee
- Its user interface was very limited as only HTML was used.

2.2. Web Application Servers



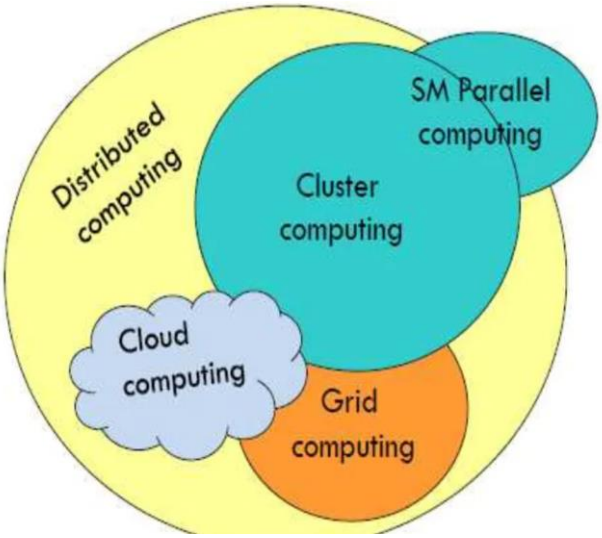


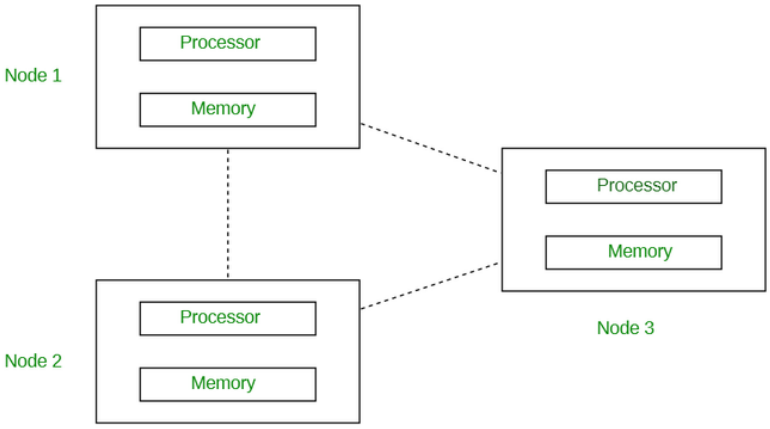
- A **web server** delivers static web content—e.g., HTML pages, files, images, video—primarily in response to (HTTP) requests from a web browser
- An **application server** typically can deliver web content too, but its primary job is to enable interaction between end-user clients and server-side application code to generate and deliver dynamic content, such as transaction results, decision support, or real-time analytics.
- **Web application server** requires more resources and is best suitable for Enterprise applications.

- **Web application server** supports multithreading and supports additional RPC/RMI protocols. (Remote Procedure Call and Remote Method Invocation)
- The application server is working in conjunction with the webserver, where one displays and the other one interacts.

2.3. Overview of Computing Paradigm: Grid Computing, Cluster Computing, Distributed Computing, Utility Computing, Cloud Computing

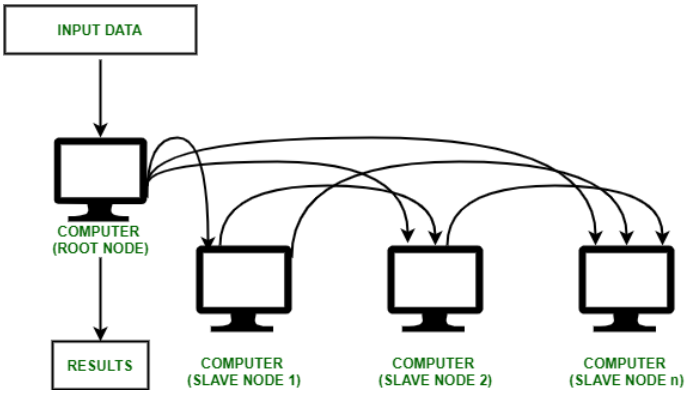
2.3.1. Distributed computing





- Multiple computer systems working on a single problem.
- A single problem is divided into many parts, and each part is solved by different computers thus performing as a single entity.
- Ensure maximum performance by connecting users and It resources and also ensure the system is Fault-tolerant.
- Scalability and Redundancy are two major features of Distributed system
- **Working :**
  - Client machines with very lightweight software agents installed
  - When the agent detects the machine is idle, it requests an application package
  - When received agent run application software (when free)
  - Return the result to the management server.
  - Then the agent frees the resources.

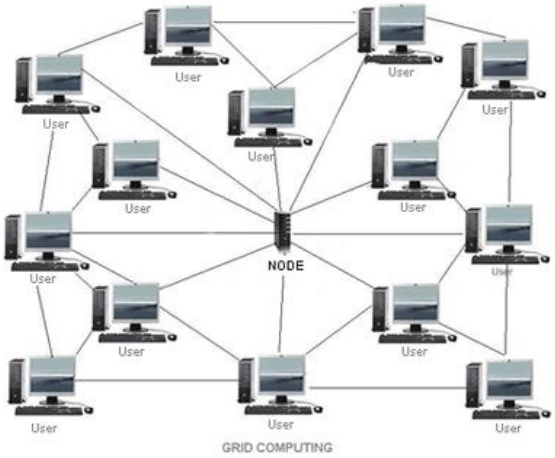
2.3.2. Cluster computing



- consists of two or more independent **Homogeneous** computers referred to as nodes, that work together to execute tasks as a single machine.
- *Homogeneous computers have processors of the same kind.*
- increase the cost efficiency, availability, performance, scalability and simplicity of the system.
- type of parallel and distributed system
- ensures that computational power is **always available .ie Dedicated** and Fault-tolerant.
- expensive and difficult to set up and maintain
- Application:

- Solve the complex computational problem
- Earthquake simulation
- Time-critical system
- Weather forecasting
- Image Rendering

2.3.3. Grid Computing



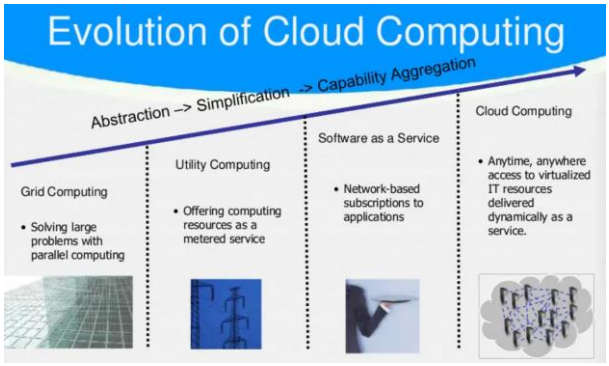
- processor architecture that combines computer resources from various domains to reach the main objective
- the (**homogeneous and heterogeneous**) computers on the network can work on a task together, thus functioning as a virtual supercomputer.
- every resource is on every computer is shared
- Every authorized computer would have access to enormous processing power and storage capacity.
- Use to solve large scale resource-intensive problems in science, engineering, and commerce
- They use computers that are part of the grid only when idle and operators can perform tasks unrelated to the grid at any time.

2.3.4. Utility computing

- Providing computing service through an on-demand, pay-per-use billing method.
- Provider owns, operates and manages the computing infrastructure and resources, and the subscribers access it as and when required on a rental or metered basis.
- The consumer has access to a virtually unlimited supply of computing solutions
- Efficient use of resources while minimizing the cost.
- Limited to “computing as a service “

2.3.5. Cloud computing

- Started with the concept of “Computing as a Service” or Utility Computing
- Cloud computing = (part of) Grid computing + (most of) utility computing + more
- Cloud computing is the delivery of on-demand computing services -- from applications to storage and processing power -- typically over the internet and on a pay-as-you-go basis
- **Advantages:**
  - Efficient use of resources while minimizing the cost
  - Virtually unlimited supply of computing resources
  - Remove upfront cost and complexity of maintaining IT infrastructure.
  - Wide range of services like NLP, AI, ML, office application.
  - Access to vast resources in a fraction of seconds.
  - Global scale.
  - Reliable due to fault-tolerant
  - Availability.



- Public, private, and hybrid are three types of cloud computing.
- AWS, Google Cloud, Azure, IBM cloud are some leading Cloud computing providers.

Key	Cluster computing	Grid computing
Processor and OS at node	Homogeneous	Supports both Hetro and Homogeneous
Computers Dedication	Dedicated to particular work only	Only provide unused resources to the grid
Location	Generally close to each other	Maybe at a huge distance
Connection	High-speed local bus	Low-speed bus or internet
Network topology	Centralized	Distributed
Availability	Whole system functions as a single system	Every node is autonomous and can opt out anytime.
Ownership	Single	Multiple

Key	Cluster computing	Cloud computing
Resource sharing	Shareable	Not shareable
Resource type	Has to be homogeneous	Support Heterogeneous too
Virtualization	Not needed	h/w and s/w resources are virtualized.
Initial cost	high	low
Maintenance	More maintenance	Less maintenance
Scalability	limited	Virtually unlimited resources

Key	Grid computing	Cloud computing
Application type	Batch application	Interactive application
Access	Via grid middleware	Via web protocol (TCP/IP)
Virtualization	Only in the beginning stages	h/w and s/w resources are virtualized.
Business model	Sharing	Pay-as-you go (Utility)
Application development	Locally	On the cloud
Switching cost	Low	Higher

2.4. Internet of Services

- A concept where everything that is needed to use software applications is available as a service on the Internet including software and the underlying platform.
- Generally, IoS provides all the IT resources over the internet.

2.5. Adopting Cloud Computing in Business

2.5.1. Benefits

- **Cost-efficient**:- no upfront cost, pay as you go, model
- **Scalability**:- Virtually unlimited amount of resources available
- **Upgrade and maintenance**:- No downtime during upgrade and maintenance
- **Better security**:- Physical security is handled by the cloud service providers and they support Multi-factor authentication. Additionally, all data is encrypted.
- **Multiple options** :- IaaS, PaaS, SaaS and XaaS

2.5.2. Risk

- Reduced control and visibility
- Resource shared with others
- Incomplete data deletion
- Increase complexity strains in IT staff
- The flaw in one of the APIs could compromise thousands of clients

3. Enterprise Architecture: Role and Evolution

Enterprise Architecture

- The fundamental organization of a system, embodied in its components, their relationship to each other and the environment and the principles governing its design and evolution
- The conceptual blueprint defines the structure and operation of an organization.
- Characterised by a framework that supports the alignment of business and IT strategy
- As enterprises are growing and becoming complex, the Architecture managing the Enterprise is evolving.

3.1. Enterprise Data, Processes and Components

3.1.1. Enterprise data

- All the data that is generated and shared between and among users of the organization
- Could be internal or external data.
- Should have single consistent version sharing and modelling
- Must follow standards for internal and external data to maintain quality of data
- Enterprise must deploy means to secure data from unauthorized access.

3.1.2. Enterprise Processes

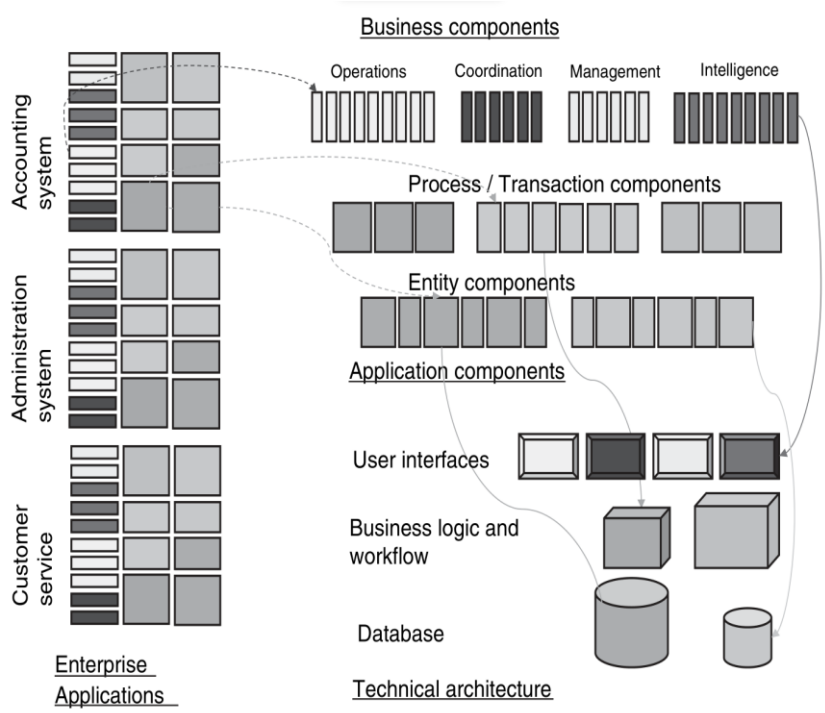
- Vertical Processes:
  - Operate within same functional units
  - Eg: process involving Sales only
- Horizontal Processes:
  - Operate across the functional unit
  - Eg: process involving sales, IT and production

3.1.3. Enterprise Components

- Business Component
  >>collections of high-level enterprise processes

- Application Component
  >> group of smaller processes derived from High-level processes

  - Software Components
    >>unit from Decomposition of Application Component
    >> deals with manipulation of Enterprise data



3.2. Enterprise Architecture Evolution and Frameworks

3.2.1. Enterprise Architecture Evolution

- Zachman framework in 1987
- The Open Group Architecture Framework (TOGAF) in 1995. Different versions released to date
- Federal Enterprise Architecture (FEA) in 2002
- Nepal GEA in 2020

3.2.2. Enterprise Architecture Framework

-
-
-
-

3.3. Monolithic, Application Integration, SOA and Microservices

3.3.1. Monolithic

- The primitive way of building application (a single indivisible unit)
- All application functions are managed and served in a single place.
- Large code and lack modularity.
- Easier to implement and useful for smaller projects where the regular update is not necessary.
- Updates are tedious and can break the entire system.

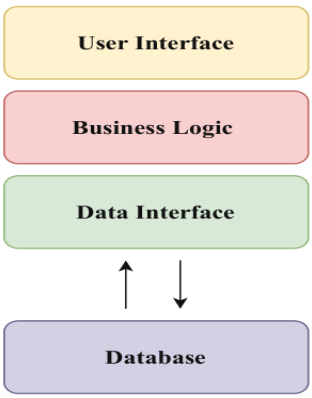


Figure 1: Monolithic architecture

3.3.2. Application Integration

- Application integration enables applications and systems that were built separately to work together
- May result in new capabilities and efficiencies ultimately reducing cost and uncover insights.
- Bridge the gap between existing on-premises systems and fast-evolving cloud-based enterprise applications.

- Seamlessly interconnected processes and data exchange help businesses to operate effectively and efficiently.
- Use Application Programming Interface (API) extensively to interact with software components
- Database triggers are used for data-level Integration.

3.3.3. SOA (Service Oriented Architecture)

- Architecture that idealizes applications containing discrete software agents that are loosely coupled to perform the required function
- Multiple software agents interact with a single central system called the **Enterprise Service Bus.**
- Also called centralized oriented architecture
- Each service is a black box (abstract) so if a single service is faulty, the entire application isn't down.
- Provides service reusability
- **Enterprise Service Bus** becomes the single point of failure and may cause the entire application to collapse.

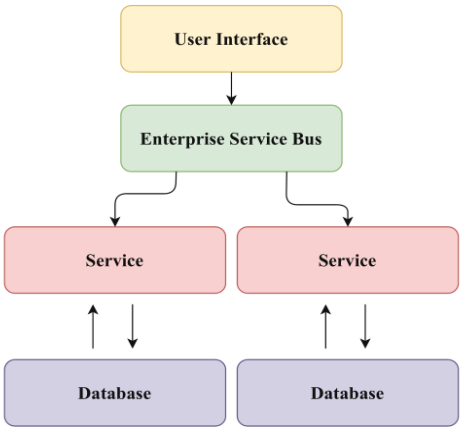
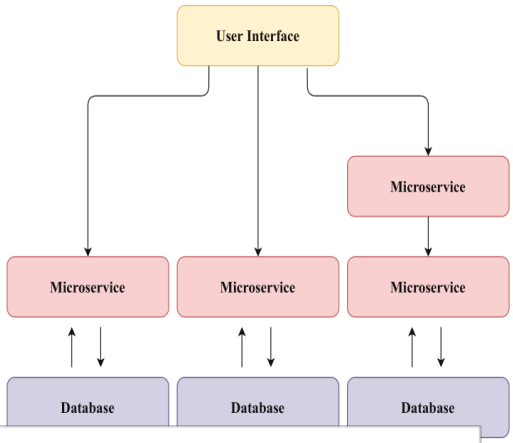


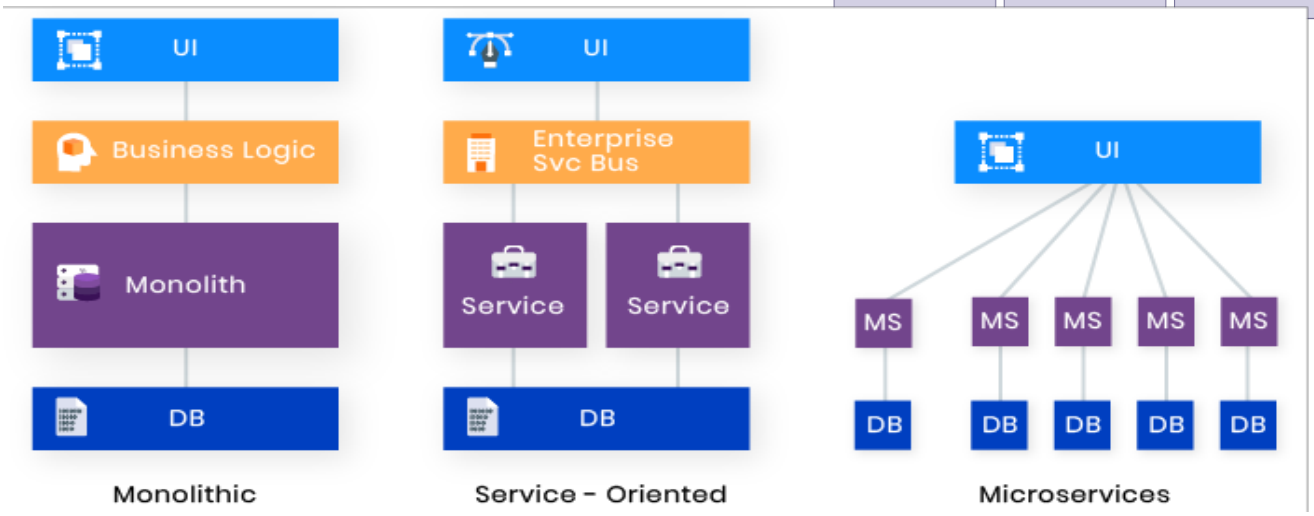
Figure 2: Service-oriented architecture

3.3.4. Microservices

- Type of SOA that idealizes autonomous components that make an application.
- Contains multiple independent components (Microservices) that work together using APIs.
- **Main difference from SOA:** Doesn't rely on a central service, rather all components are self-contained and autonomous.
- Fault tolerance is highest among all discussed.
- A modular code base improves bug detection and troubleshooting.
- Complexity and difficulty in development are major drawbacks.



3.3.5. Comparison



Category	Monolithic	Service-oriented	Microservices
Type of architecture	Single entity	Centralized hub	Decentralized
Developmental complexity	Least	Midway	Most complex
Maintenance complexity	Most complex	Easier to maintain	Special expertise required
Performance	Most	Depends on enterprise service bus	Least (can be increased with hot services)
Inter-function communication	Uses single code base	Uses enterprise service bus	Uses APIs
Modularity	None	Services are modular	Maximum
Use cases	Simple applications with time constraint	Enterprise level applications	Enterprise level applications

Table 1: Comparison between monolithic, service-oriented and microservices based architectures

Key	Monolithic	Service-oriented	Microservices
Architecture	Single Entity	Centralized	Decentralized
Suitable for	Small teams and Small projects	Enterprise Level Application	Enterprise Level Application
Complexity	Easy to deploy but Difficult to maintain when it grows. And have to adapt to new technology	Less Flexible in Deployment. Easier to maintain	Quick & Easy Deployment, Development and testing. Data duplication adds little complexity.
Performance	Higher performance as less or no API calls & and modules are part of a single entity and share the same resources.	Dependent on Enterprise Service Bus (ESB).	Performance is slow when lots of microservices are called at once but if hot services are scaled accordingly then have better performance.
Fault tolerance	Deployed as a single unit so no Fault tolerance	The whole system is divided into services so the failure of one does not affect the whole functionality	Highly scalable and Fault tolerance as each microservice is independent.

Agility	Deployed as a single unit so no modularity and have no agility and have to test the complete application for any minor changes.	Each service can be of any desired technology and can be maintained and updated independently.	Provide a high level of agility and even support Continuous development.
Reusability	No reusability	Since services are independent they can be reused to other applications too.	Highly Reusable but affects agility if overused.
Communication	Single code base so fastest communication.	Use Enterprise Service Bus (ESB) to communicate between and among services	Use API layer to communicate among microservices
Storage	Single storage for the whole application	Services share the same data storage	Each microservices has independent data storage

Each Architecture has its pros and cons and all are still in trend. The selection of one hugely depends on the intention of the end product. If a team has small apps to develop with limited resources then Monolithic will be the best choice. But if the enterprise has to deploy a banking application then centralized SOA will be the best choice in terms of fault tolerance and security. Microservices is best suitable for complex large-scale systems consisting of multiple teams to handle specific areas of application.

3.4. Enterprise Technical Architecture

- Refers to a set of standards or guidelines for an IT infrastructure.
- Definition and management of standards defining the technical architecture, tools and technical components used in an enterprise.
- Covers efforts and technologies aimed at designing an architecture for the IT setups that a business uses.
- Why use ETA
  - **Uniformity**
    - Reduce integration and maintenance effort
    - Simple and saves cost of multiple licenses
  - **Network and Data Security**
    - Use VPN
    - Use server-side encryption
    - Test and prevent against common attacks like SQL injection and cross-site scripting
  - **Implementation architecture and Quick -win**
    - Reduce transaction cost by minimizing the disruption due to transition

3.5. Data Center Infrastructure: Coping with Complexity

- As businesses evolve the data and datacenter's complexity increases
- New business requirements may not be supported by legacy systems, so the system has to be developed for that particular need. In the meantime, the new system should support the data from the legacy system.

3.5.1. Coping with complexity

- **Emphasize Standardization**
  - Long term support for equipment

    - Cheaper, faster and easier --> upgrade and maintenance
    - Simplify the IT infrastructure
- **Streamline Vendors**
    - Consistent support
    - Reduce complexity
    - Simple integration of components
- **Find multipurpose product**
    - So can meet future business needs
- **Aim for Seamless Scaling and Migration**

4. Cloud Concepts

4.1. Cloud Computing (NIST Model), Properties, Characteristics, Benefits

4.1.1. Cloud Computing (NIST Model)

- **National Institute of Standards and Technology (NIST)**
  Cloud Computing: Benefits and Risks of Moving Federal IT into the Cloud
- NIST:- "Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction"

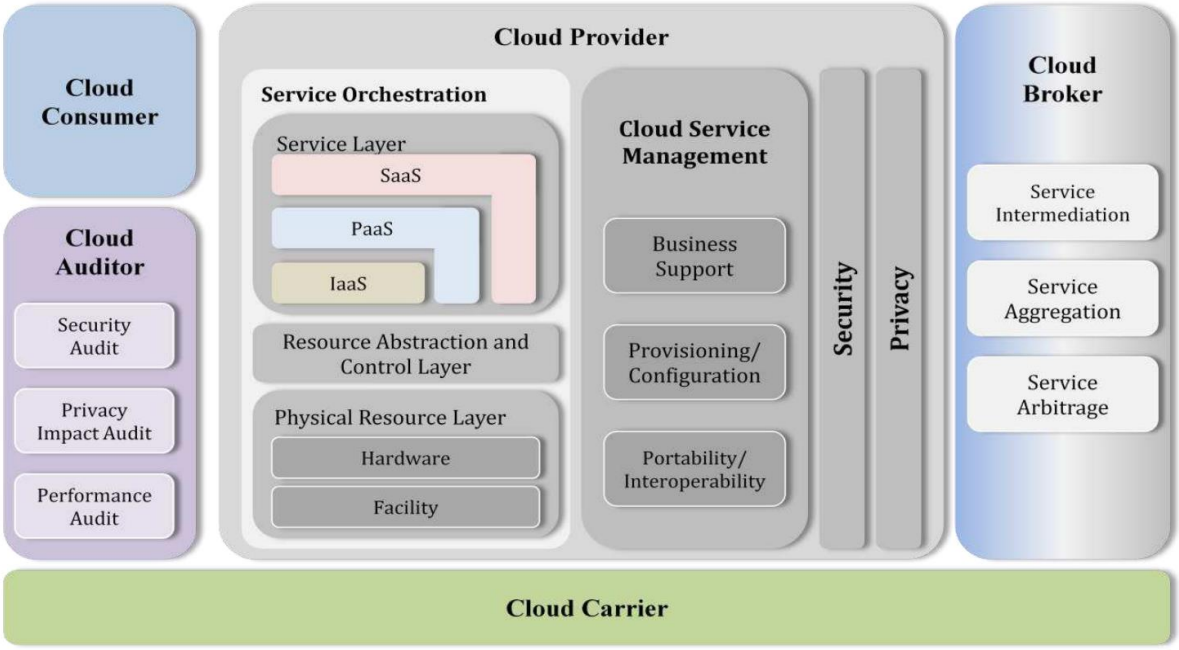


Figure 1: The Conceptual Reference Model

Actor	Definition
Cloud Consumer	A person or organization that maintains a business relationship with, and uses service from, *Cloud Providers*.
Cloud Provider	A person, organization, or entity responsible for making a service available to interested parties.
Cloud Auditor	A party that can conduct independent assessment of cloud services, information system operations, performance and security of the cloud implementation.
Cloud Broker	An entity that manages the use, performance and delivery of cloud services, and negotiates relationships between *Cloud Providers* and *Cloud Consumers*.
Cloud Carrier	An intermediary that provides connectivity and transport of cloud services from *Cloud Providers* to *Cloud Consumers*.

4.1.2. Properties (NIST model)

- Has below-mentioned Characteristics (NIST model)
- List three service model (IaaS, PaaS, SaaS)

- Four Deployment models (private, community, public and hybrid)

4.1.3. Characteristics (NIST model)

- **O**n-demand self-service
- **B**road network access
- **R**esource pooling
- **R**apid elasticity or expansion
- **M**easured service

4.1.4. Benefits (NIST model)

- Cost-saving
- Energy saving
- Rapid deployment
- Customer empowerment
- Characteristics (NIST model)

4.2. Cloud Types; Private, Public and Hybrid Cloud. (Deployment model)

4.2.1. Private Cloud

- Set upon the premise for the exclusive use of an organization and its customer
- Operate Solely for an organization
- **Advantages**
  - High security, privacy and reliability
  - Storage and network components are customizable
  - High control over the corporate information
  - Meet compliance requirement
- **Disadvantages**
  - Costlier form of Cloud in terms of hardware software and IT staff.
  -

4.2.2. Public Cloud

- Hosted on the premise of the service provider
- Accessible to the general public
- **Advantage**
  - Easy access to data and services
  - Flexibility and scalability
  - Cost-effective
  - Continuous operation
  - High guarantee Up-time
  - The service provider handles maintenance and upgrades
- **Disadvantages**
  - Data security and privacy
  - Lack of granular control
  - Mobile difficulty
  - Limited scalability

4.2.3. Hybrid Cloud

- Combination of Private and Public model
- Can enjoy the benefits of public cloud-like scalability and reliability whereas keeping the sensitive data and services Private and controlling their access
- **Advantages**
  - Improved security and privacy
  - Meet compliance requirement

- Reasonable price
- Flexibility as additional resources is available in the public cloud.

➢ **Disadvantages**
- Toggling between public and private could be hard to track causing resource wastage.
- Added complexity

4.2.4. Community cloud

➢ Cloud infrastructure that allows systems and services to be accessible by a group of several organizations to share the information

➢ Shared by a group of organizations of similar industries and backgrounds with similar requirements i.e. mission, security, compliance and IT policies.

➢ **Advantages:**
- Cost reduction
- Improved security, privacy and reliability
- Ease of data sharing and collaboration

➢ **Disadvantages:**
- Higher cost than that of a public one
- Sharing of fixed storage and bandwidth capacity

KEY	Private	Public	Hybrid
Tenancy	Single tenancy: there's only the data of a single organization stored in the cloud.	Multi-tenancy: the data of multiple organizations is stored in a shared environment.	The data stored in the public cloud is usually multi-tenant, which means the data from multiple organizations is stored in a shared environment. The data stored in the private cloud is kept private by the organization.
Exposed to the Public	No: only the organization itself can use the private cloud services.	Yes: anyone can use public cloud services.	The services running on a private cloud can be accessed only by the organization's users, while the services running on the public cloud can be accessed by anyone.
Data Center Location	Inside the organization's network.	Anywhere on the Internet where the cloud service provider's services are located.	Inside the organization's network for private cloud services as well as anywhere on the Internet for public cloud services.
Cloud Service Management	The organization must have its administrators managing its private cloud services.	The cloud service provider manages the services, where the organization merely uses them.	The organization itself must manage the private cloud, while the public cloud is managed by the CSP.

Hardware Components	Must be provided by the organization itself, which has to buy physical servers to build the private cloud on.	The CSP provides all the hardware and ensures it's working at all times.	The organization must provide hardware for the private cloud, while the hardware of CSP is used for public cloud services.
Expenses	Can be quite expensive, since the hardware, applications and network have to be provided and managed by the organization itself.	The CSP has to provide the hardware, set up the application and provide the network accessibility according to the SLA.	The private cloud services must be provided by the organization, including the hardware, applications and network, while the CSP manages the public cloud services.

4.3. Service Models: IaaS, PaaS, SaaS.

4.3.1. IaaS (Infrastructure as Service)

- All the needed infrastructure is from CSP(Cloud Service Provider) and Pay for only the resources rented and returned when not in use.
- Infrastructure is scalable depending on processing and storage needs
- Saves enterprises the costs of buying and maintaining their hardware.
- Enables the virtualization of administrative tasks, freeing up time for other work
- Cost varies depending on consumption
- Here CSP and user share responsibilities like networking, storage, servers, and virtualisation managed by CSP, the rest is the client's responsibility.
- **Suitable for:**
  - Startups and small companies
  - Large companies seeking control over infrastructure
  - Companies experiencing Rapid growth
- **Advantages**
  - most flexible cloud computing model
  - Clients retain complete control of their infrastructure
  - Resources can be purchased as-needed
  - Highly Scalable
- **Disadvantages**
  - Security:- Vulnerability in VM technology compromise client privacy and security
  - Additional resource and training to IT staff

4.3.2. PaaS(Platform-as-a-Service)

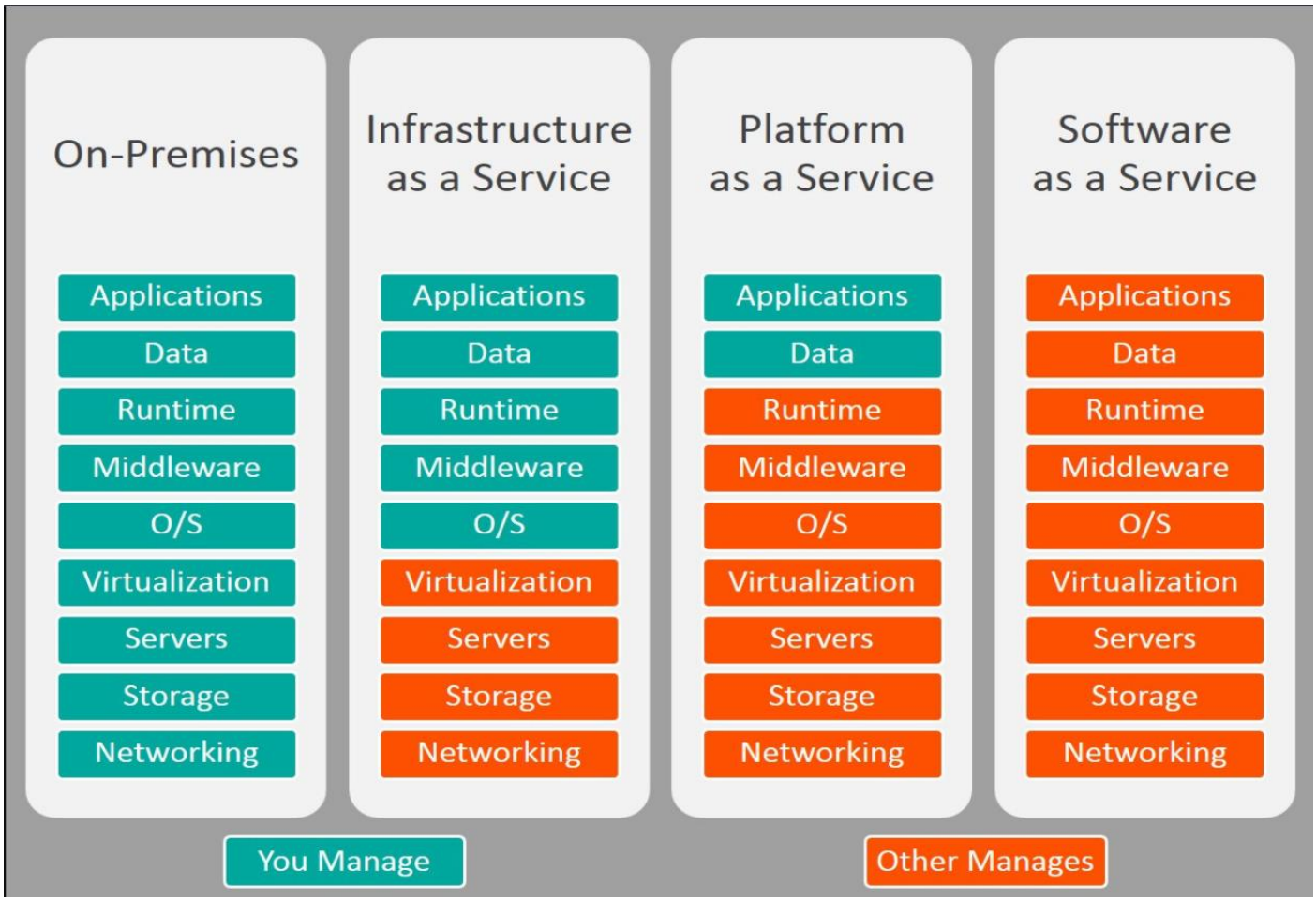
- In addition to IaaS, PaaS CSP(Cloud Service Provider) is also responsible for runtime, middleware and OS.
- Developers have to only deal with applications and Data.
- Delivers a framework for developers that they can build upon and use to create customized applications
- Containerization or Docker technology is generally used in PaaS.
- Provides a variety of services to assist with the development, testing, and deployment of apps
- **Advantages**
  - Simple, cost-effective development and deployment of apps
  - Scalable
  - Highly available
  - Developers can customize apps without the headache of maintaining the software
  - Significant reduction in the amount of coding needed

- **Disadvantages**
  - Since data resides in CSP, poses a security risk
  - Integration:- a complex procedure to connect onsite data centre
  - No support for legacy system
  - Limited framework and language supported

4.3.3. SaaS(Software as a Service)

- In addition to PaaS, SaaS data and applications are also managed by CSP.
- Utilizes the internet to deliver applications, which are managed by a third-party vendor, to its users.
- **Advantages**
  - Reducing the time and money spent on tedious tasks such as installing, managing, and upgrading software.
  - Technical staff could attend to other pressing matters.
- **Disadvantages**
  - Lack of interoperability as it may not support existing apps
  - Data security risk as large volume of data has to be exchanged
  - Minimum customization
  - Minimum control and features

Platform Type	Common Examples
SaaS	Google Workspace, Dropbox, Salesforce, Cisco WebEx, Concur, GoToMeeting
PaaS	AWS Elastic Beanstalk, Windows Azure, Heroku, Force.com, Google App Engine, Apache Stratos, OpenShift
IaaS	DigitalOcean, Linode, Rackspace, Amazon Web Services (AWS), Cisco Metapod, Microsoft Azure, Google Compute Engine (GCE)



4.4. Role of Virtualization in Enabling the Cloud.

- **Virtualization** software allows multiple operating systems and applications to run on the same server at the same time.
- Virtualization is the process of creating a virtual server for providing service devices, various infrastructure, and computing resources
- With virtualization, it performs all of the functions that a piece of hardware performs without the hardware in place.
- Virtualization assigns a logical name to a physical resource and provides a pointer to that physical resource when demanded.
- Cloud uses virtualization technology to provide services that allow end-users to access virtualized servers, applications, etc., without having to purchase that hardware
- Clouds are environments that abstract, aggregate, and share virtual resources over a network
- **Advantages of Virtualization in Cloud**
  - Use the **single system** separately by multiple users.
  - **Flexible** data transfer and retrieval (no limit)
  - **Economical** as no physical hardware is needed.
  - **Efficient** resource management
  - **SLA** service level agreement/guarantee
  - **Disaster recovery** is possible due to seamless backup and the same work can be performed simultaneously on different machines.
  - **Reduced workload** as Service provider takes responsibility for update and maintenance
  - **Multisharing** as the same physical server is virtually divided into various parts and assigned to multiple parties.
  - **Security**:- uses a firewall and different protocols to restrict unauthorized access.

- **Disadvantages of Virtualization in Cloud**
  - May not be always available for service
  - Cloud faces scalability issues if CSP lacks the resources.
  - Security Issues as in Public cloud same hardware and resources are shared by multiple users.
- **Types of virtualization in Cloud**
  - **Hardware Virtualization**:- With the help of Hypervisor one can divide processor, memory and other hardware resources. VMM (Virtual Machine Manager) installed on the hardware
  - **OS Virtualization**:- Multiple guest OS can be run over a single host OS. VMM installed on Host OS. Quicker than Server Virtualization
  - **Server Virtualization**:- One physical server divided into multiple servers. VMM installed on the server system.
  - **Storage Virtualization**:- Multiple physical storage is grouped to act as single large storage. One server sees multiple storages as one (due to virtualization) and storages see the server they are attached to.
  - **Network Virtualization**:- Combine the available resources in a network by splitting up the bandwidth into separate channels.
  - **Software Virtualization**:- Under the SaaS model instances of the same software installed are provided to multiple users.

4.5. Application Availability, Performance, Security and Disaster

4.5.1. Application Availability

- A measure used to evaluate if an application is functioning properly and meets the business requirements
- Cloud provides high availability with a high degree of performance.
- Available 24/7 with nodes in major parts of the world
- Determined based on:
  - Application uptime/downtime
  - Number of complete transactions
  - Timeouts, errors, missing resources
- Uses different techniques to guarantee high availability:-
  - Load balancer
  - Concept of availability zones
  - Health checks

4.5.2. Performance

- A measure used to monitor and ensure effective operations
- Guarantee a certain degree of performance under SLA
- Determined based on:
  - Input/Output Operations per Second (IOPS)
  - File system performance
  - Caching
- Uses different techniques to guarantee Performance:
  - Data backup
  - Disaster recovery
  - Fault tolerance
  - Load balancer
  - Autoscaling
  - Easy maintenance
  - Using appropriate infrastructure
- Types of performance testing:
  - Stress testing
  - Load testing
  - Latency testing
  - Failover testing

4.5.3. Security

- Consists of a set of policies, controls, procedures that work together to protect the cloud-based system
- Configured to protect cloud data and customer privacy under SLA
- CSP is responsible for the physical security of the server and provides various measures to clients to implement on their end. like Firewall, Virtual Private Cloud (VPC), Multi-Factor Authentication (MFA), Access Control List (ACL), Identity Access Management (IAM), Server and client-side Encryption
- Data backup

4.5.4. Disaster

- An unpredictable event that causes an interruption in services
- Cloud provide an effective and cheap Disaster recovery plan (DRP) facility
- If the current functioning server fails then it is the responsibility of CSP to spin another server within a certain time frame and continue its operation.
- CSP employs features like :
    - Data backup
    - Automation in server deployment and recovery
    - Rapid elasticity to minimize downtime
    - Availability zone to redirect traffic to other AZ in the region in case of a disaster in one.

Type of Backup	Advantages	Disadvantages
Full (Save all files)	➢ Fastest recovery method	➢ Longest backup time. ➢ Require more space to back up.
Differential (Save files that are different)	➢ Faster backup method. ➢ Require less space to back up.	➢ Time-consuming recovery as the last full backup and the differential backup is needed.
Incremental (Save a base and incremental recent changes)	➢ Fastest backup method. ➢ Requires less space to back up.	➢ Most time-consuming recovery, as the last full backup (base) and all the incremental backup, are needed.
Selective (Save only required files)	➢ Fast backup method. ➢ More flexibility.	➢ Difficult to manage individual file backups.

5. Cloud Architecture & Framework

5.1. Cloud Global Infrastructure

- Two key components—physical infrastructure and connective network components.
- Physical data centres are arranged into regions and linked by one of the largest interconnected networks.
- **Data centres** are unique physical buildings—located all over the globe—that house a group of networked computer servers for housing the actual data.
- Each data centre provides high availability, low latency, and scalability of data on the cloud.
- **Availability Zones (AZs)** are clusters of data centres spread across the globe designed to provide robust and reliable service to its customer
- AZs are characterized by Redundant Power supply and high bandwidth network connectivity.
- AZs ensure isolation and disaster recovery during calamities.
- Collections of (AZs) are termed as **Regions**. Lies within a latency-defined perimeter and is connected through a dedicated regional low-latency network.
- Regions are completely independent of each other and pricing varies according to region.
- **Edge locations** are end-points near user locations that provide a high-speed connection to cached resources.
- **Features**:
  - Trusted global presence
  - Secure, efficient data centres
  - Fast/low latency, reliable global network
  - Economies of Scale
  - Flexibility and scalability
  - Easier and faster deployment
  - Compliance and data residency

5.2. Deploying And Operating in the Cloud

5.2.1. Deploying

- Provision infrastructure from code **(Infrastructure as a Code (IaaC) and AWS CloudFormation)**
- Deploy artifacts automatically from version control **(CloudDeploy)**
- Configuration managed from code and applied automatically **(CI/CD)**
- **IaaS** (Amazon Elastic Container Service)
- **PaaS** (Elastic Beanstalk, Google App Engine, Heroku)

5.2.2. Operating in the cloud

- Scale your infrastructure automatically **(Autoscaling)**
- Monitor every aspect of the pipeline and the infrastructure **(CloudWatch)**
- Logging for every action **(CloudWatch Logs and CloudTrail)**
- Instance profiles for embedding IAM roles in instances automatically
- Use variables, don't hard code values
- Tagging can be used with automation to provide more insights into what has been provisioned.

5.3. Well-Architected Framework Design Principles

5.3.1. Operational Excellence

- Perform operations as code
- Make frequent, small, reversible changes
- Refine operations procedures frequently
- Anticipate failure
- Learn from all operational failures

5.3.2. Security

- Implement a strong identity foundation

- Enable traceability
- Apply security at all layers
- Automate security best practices
- Protect data in transit and at rest
- Keep people away from data
- Prepare for security events

5.3.3. Reliability

- Automatically recover from failure
- Test recovery procedures
- Scale horizontally to increase aggregate workload availability
- Stop guessing capacity
- Manage change in automation

5.3.4. Performance Efficiency

- Democratize advanced technologies
- Go global in minutes
- Use serverless architectures
- Experiment more often
- Consider mechanical sympathy

5.3.5. Cost Optimization

- Implement cloud financial management
- Adopt a consumption model
- Measure overall efficiency
- Stop spending money on undifferentiated heavy lifting
- Analyze and attribute expenditure

5.4. Software as a Service, SaaS Architectures, Dev 2.0 in the Cloud for Enterprises

For SaaS, refer to SaaS(Software as a Service)

5.4.1. SaaS Architectures

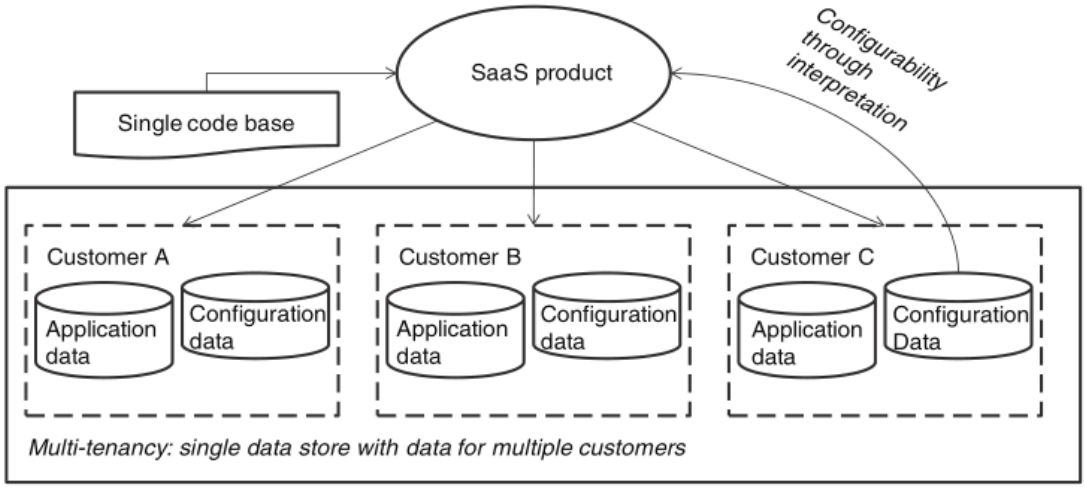


FIGURE 3.1. SaaS architecture

5.4.2. Dev 2.0 in the Cloud for Enterprises

- Aim to bring end-user participation into application development.
- Dev 2.0 aims to replace application-specific code with meta-data that is rendered and edited in web applications.
- Early performance testing is possible using replicas of the production environment.
- Still a relatively new concept.
- **IaaS**=> More IT staff so bottleneck occurs,
- **PaaS**=> Platform provided but code by the business.
- **SaaS**=> Entire software provided as a service.
- **Dev 2.0** uses a database owned by the Dev 2.0 platform (provider) but uses the meta-data of their customer for the specific application.

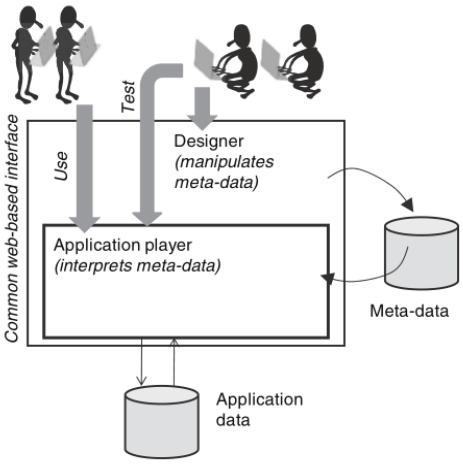


FIGURE 3.2. Dev 2.0 architecture

5.5. Infrastructure as a Service

IaaS (Infrastructure as Service)

5.6. Platform as a Service

PaaS(Platform-as-a-Service)

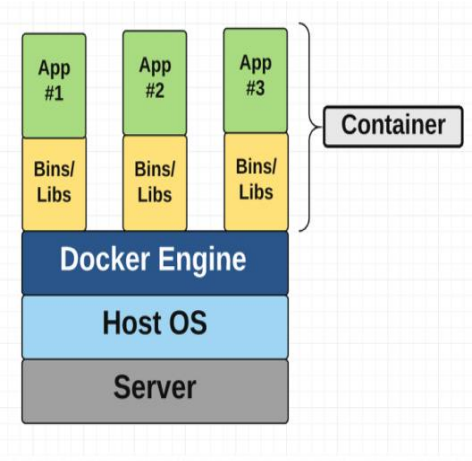
6. Computing and Storage Services

6.1. Compute Services – Virtual Server, Cloud Containers, Cloud Server Less Service, Auto Scaling

6.1.1. Virtual Server

- The backbone of cloud computing.
- Virtualizing servers involves converting one physical server into multiple virtual machines (VMs).
- A virtual server is configured so that multiple users can share its processing power.
- Virtual servers play a key role in building applications, tools, or environments. Like software development and testing.
- **Benefits:**
  - Reduced costs
  - Redundancy
  - Lower power and energy cost
  - Security
  - Remote access
- **AWS**=> Elastic Compute Cloud (EC2)
- **Azure**=>Virtual machines
- **GCP**=> Google Compute Engine

6.1.2. Cloud Containers



- Run multiple instances (OS) on the single host machine
- Virtualize OS
- Lightweight and designed to be created, scaled out, and stopped dynamically
- Docker as service and Portable
- Run containerized apps quickly since no boot-up process is required.
- Containers contain everything the apps needs to run in a "container image"
- **AWS**=>AWS Fargate,
- **Azure**=>Azure Container Instances
- **GCP**=> Cloud Run
- **Kubernetes:**- automating deployment, scaling, and management of containerized applications
  - **AWS**=> Amazon Kubernetes Service
  - **Azure** => Azure Kubernetes Service (AKS)
  - **GCP**=>Google Kubernetes Engine

6.1.3. Cloud Server Less Service

- Also known as Function as a Service
- **AWS**=> AWS Lambda
- **Azure** => Azure Functions and Logic Apps
- **GCP**=>Google Cloud Functions
- Run code without thinking about servers or clusters. Only pay for what you use.
- Event trigger functions that automatically implement the application functions.
- **Benefits:**
  - No servers to manage
  - Continuous scaling
  - Cost-optimized with millisecond metering
  - Consistent performance at any scale
  - Quick deployments and updates are possible

6.1.4. Auto Scaling

- Automatic scaling of resources during spikes and termination during the drop
- Automatically adjust the number of instances as per the estimated traffic, memory and computation required or as per the predefined rules.
- **Vertical Scaling** (Scale UP/DOWN)
  - Here the existing server is upgraded to the higher specification of memory, CPU, Storage etc.
- **Horizontal Scaling** (Scale IN/OUT)
  - Here multiple servers or instances are created having the exact specification as the existing one.
  - It is a more popular type of scaling for applications or services in the deployment phase.
  - It is also created to distribute load among the multiple servers using Load Balancer.
- **AWS**=>Auto Scaling
- **Azure**=> Azure Autoscale
- **GCP**=> Autoscaling

6.2. Storage Services – Object Storage, File Storage, and Block Storage, Archive Storage

6.2.1. Object Storage

- In object storage, the data is broken into discrete units called objects and is kept in a single repository, instead of being kept as files in folders or as blocks on servers.
- The objects stored have an ID, metadata, attributes, and the actual data.
- Abstraction of lower layers of storage.
- Suitable for highly durable and available application
- Storing backups and archives for compliance with regulatory requirements.
- Faster data retrieval, infinite scalability, optimization of resources.
- **AWS**=> Amazon Simple Storage Service (S3)
- **Azure**=> Azure Blob Storage
- **GCP**=> Google Cloud Storage

6.2.2. File Storage

- A cloud-based shared file system for unstructured data.
- Stored in files and may be organized in a folder
- Hierarchically arranged into directories and subdirectories.
- Low latency and provide concurrent access to tens of thousands of clients
- Scalable and predictable performance up to hundreds of thousands of IOPS, tens of GB/s of throughput, and hundreds of TBs.
- An easily set access as well as editing permissions across files and trees such that security and version control are far easier to manage
- **AWS**=> Elastic File System (EFS)
- **Azure**=> Azure Files
- **GCP**=> Google Cloud Firestore

6.2.3. Block Storage

- Chops data into blocks and stores them as separate pieces
- Each block of data is given a unique identifier, which allows a storage system to place the smaller(equal) pieces of data wherever is most convenient.
- Retrieve data quickly
- The more data you need to store, the better off you'll be with block storage.
- **AWS**=> Elastic Block Storage (EBS)
- **Azure**=> Azure Disk Storage
- **GCP**=> Google Persistent Disks

6.2.4. Archive Storage

- Storage service for long-term data retention.
- Data that is infrequently accessed
- Two requirements: Cost must be low and data recovery must be guaranteed.
- Traditionally stored in cheaper magnetic storage but retrieval may not be guaranteed due to storage corruption.
- Benefits:
  - Low cost, high value
  - Long Term Backup Retention
  - Helps to meet compliance requirements
- **AWS**=> S3 Glacier Deep Archive, S3 Infrequent Access
- **Azure**=> Azure Archive Storage, Azure Cool Blob Storage
- **GCP**=> Google Cloud Storage Nearline, Coldline and Archive

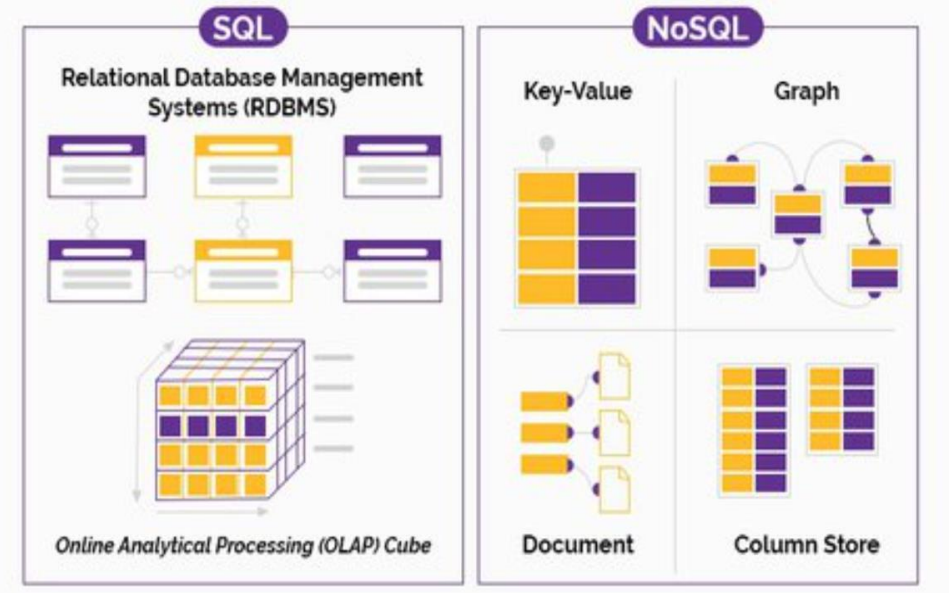
6.3. Databases – SQL, NoSQL and Graph Database and Database Migration

6.3.1. SQL

- Structured Query Language
- Access and manipulate databases
- Intelligent, scalable, relational database
- In the cloud, fully managed service
- **AWS**=> Amazon RDS, Amazon Aurora
- **Azure**=> Azure SQL Database
- **GCP**=> Cloud SQL
- Uses SQL queries to SELECT, DELETE, UPDATE, WHERE, INSERT entries in the database.

6.3.2. NoSQL (Not Only SQL)

- Not only SQL
- Non-relational
- Types: Key-value, column store, document, Graph
- Suitable for applications that require large data volume, low latency, and flexible data models
- Ease of development, functionality, and performance at scale.
- **AWS**=> Amazon DynamoDB
- **Azure**=> Azure CosmosDB
- **GCP**=> Cloud Bigtable



6.3.3. Graph Database

- Purpose-built to store and navigate relationships
- Uses graph structure for queries with node, edge and properties that represent stored data.
- Type of NoSQL database.
- With unstructured real-time data in huge amounts in the day and age of IoT, Graph database allows flexible, dynamic and lower cost integration of new data sources.

6.3.4. Database Migration

- Migrate data from one source DB to one or more target DB using database migration services.
- Requirements of good database migration service:
  - Minimal downtime
  - Support widely used databases
  - Simple to use, low cost and reliable

7. Networking & Security

7.1. Networking and Content Delivery – Cloud Network, CDN, DNS Services, Cloud Load Balancing.

7.1.1. Cloud network

- A computer network exists within or is part of a cloud computing infrastructure.
- Network interconnectivity between cloud-based or cloud-enabled application, services and solutions
- Similar to a standard computer network but its components/devices/operations are centered on cloud computing.
- **Benefits:**
  - Most secure
  - Highest network availability
  - Broadest global coverage
  - Consistent high performance
  - Easier connectivity to scaled infrastructure

7.1.2. Content Delivery Network (CDN)

- Geographically distributed group of servers that work together to provide fast delivery of Internet content.
- Doesn't host content rather caches the content for faster access time.
- Makes use of edge locations to perform regional caching.
- **Benefits:**
  - Reduce load times
  - Save bandwidth
  - Speed responsiveness
  - Global coverage and scalability
  - Encryption and HTTPS support
- **AWS**=> AWS CloudFront
- **Azure**=> Azure CDN
- **GCP**=> Cloud CDN

7.1.3. Domain Name System (DNS) Service

- Phonebook of the internet.
- The naming system for the computers, services or other resources on the internet or a private network.
- Hierarchical distributed database that allows storing IP addresses and other data, and looking them up to user names.
- Managing the mapping between names and numbers.
- **AWS**=> Amazon Route53
- **Azure**=> Azure DNS
- **GCP**=> Cloud DNS

7.1.4. Cloud Load Balancing

- Distribute your load-balanced compute resources in single or multiple regions—close to your users
- Put your resources behind a single anycast IP and scale your resources up or down with intelligent autoscaling
- Make sure no resource is overloaded, underloaded or idle.
- Use parameters like HTTP/ HTTPS, TCP, UDP, location and others to balance and distribute the load.
- **Benefits**:
  - Seamless autoscaling
  - Cloud logging

- Health check
- Increased availability and reliability
- Easy integration with DNS services

➢ **AWS**=> Elastic Load Balancing (ELB)
➢ **Azure**=> Azure Load Balancer
➢ **GCP**=> Cloud Load Balancer

7.2. Cloud Security and Compliance Concepts

7.2.1. Cloud Security

➢ Protection of data and services on the cloud (during transit and at rest).
➢ Protect your information, identities, applications, and devices.
➢ Implement a layered defence-in-depth strategy across identity, data, hosts, and networks
➢ Help meet compliance requirements, such as data locality
➢ Allows them to scale and innovate, while maintaining a secure environment and paying only for the services one uses.
➢ Tools and features to help you to meet security objectives
➢ Encryption by default, at rest and in transit
➢ **Distributed Denial of Service (DDoS) Protection** Google Cloud Armor, AWS Shield, Azure DDoS Protection
➢ **Virtual Private Cloud (VPC)** is the logical division of a CSP's public cloud to support private cloud computing. It provides network isolation with a range of IP addresses called subnets. It controls the network traffic to cloud infrastructure in the VPC so protects from unauthorized access.
➢ **Access Control Lists (ACLs)** control access settings for resources on the cloud. Permissions for access control include read/write access and the user/group of users who can access the resource.
➢ **Network Security Groups** are available for different services and infrastructure that specify the protocol of access, IP address of source/destination and open ports for access.
➢ **Firewall:-** Centrally configure and manage firewall rules
➢ **Multi-Factor Authentication (MFA)** allows additional authentication criteria besides basic username and password login. This way enterprises employing cloud concepts can use MFA for their employees to protect their services and infrastructures from unauthorized access.
➢ **Identity and Access Management (IAM)** allows access management using policies that ensure that the right users or user groups have access to the appropriate resources.
  - Provides fine grain security to services (who, which and how)
  - Includes:
    - **IAM User:** Single user in the IAM hierarchy
    - **IAM Group:** Users having identical authorization
    - **IAM Policy:** Defines which resource can be accessed and the level of access.
    - **IAM Role:** Used to communicate and control resources. "DENY" has the highest priority.
  - **The root user** has all permissions.
  - **Admin** has only permissions issued by the root user.

7.2.2. Compliance Concepts

➢ Cloud Compliance refers to different industry standards and regulations that cloud customers need to comply with.
➢ CSP provides tools to view the current progress towards compliance.
➢ Compliance is a Shared Responsibility
➢
➢ **Examples**
  - **HIPAA** for Healthcare Industry
  - Industry-specific **ISO******
  - **EU model** clauses
  - NIST (National Institute of Standards and Technology)'s **CSF**(Cyber Security Framework)
  - **CSA** (Cloud Security Alliance)
  - **CJIS** (Criminal Justice Information System)

7.3. Shared Responsibility Model

- CSP manages the security **of the cloud**, Customers are responsible for security **in the cloud**.
- Workload responsibility varies whether the Service model is IaaS, PaaS or SaaS (FIGURE ABOVE)
- The customer is always responsible for Data, endpoints, Account and access management.

7.4. Cloud Watch, Cloud Formation, Cloud Logs, Personal Health Dashboard.

7.4.1. Cloud Watch

- It is a metrics repository
- Monitors cloud resources and applications in real-time.
- Used to collect and track variable metrics
- Custom selections of the services we want to know the metrics about.
- Alarms can be set for certain metric criteria to trigger a notification or even make changes to the resources if a threshold is crossed.

7.4.2. Cloud Formation

- Infrastructure as Code
- Developers can deploy, update the resources in simple abstract ways to reduce complexity.
- CloudFormation templates are like JSON or YAML text files.
- Template parameters are used to customize templates during runtime.

7.4.3. Cloud Logs

- CSP provides a platform to securely store, search, analyze, and alert all of the customers log data and events
- Customers can analyze logs in Real-time as they are generated.
- Classified as: **Infrastructure Logs** and **Service Logs**
- **Infrastructure Logs:** Monitoring of infrastructure access, network activity, application, file access and parameters and errors related to the cloud infrastructures.
- **Service Logs:** Monitoring of services provided by various CSPs. May include logs of object storage, load balancers or CDNs.

7.4.4. Personal Health Dashboard

- Provides alerts and guidance for CSP events that might affect your environment
- Configure customizable cloud alerts for active and upcoming service issues
- Analyze health issues,
- Monitor the impact on your cloud resources,
- Get guidance and support, and share details and updates.
- **AWS**=> Personal Health Dashboard
- **Azure**=> Azure Service Health

7.5. Cloud Messaging and Notification Service

- Fully managed messaging service for both application-to-application (A2A) and application-to-person (A2P) communication.
- Cross-platform messaging service provided by CSPs to enable businesses to send messages to their customers.

- Businesses may use cloud messaging services to send event-triggered messages, like maintenance timings, promotional messages or event messages to their customers.
- Reliably deliver messages
- Businesses may also set up notification services to themselves or their employees triggered by some metric threshold.

8. Cloud Computing Economics

8.1. Introduction

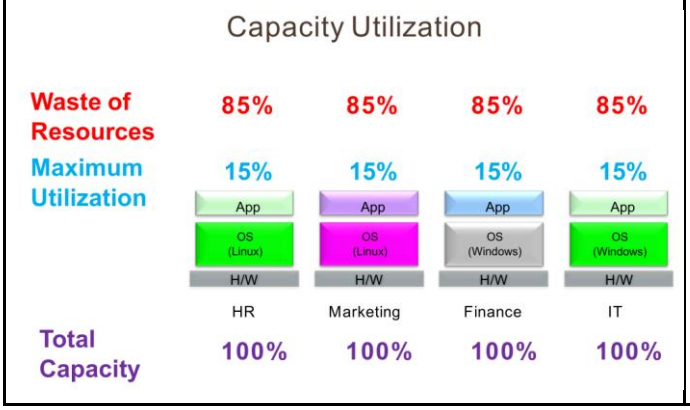
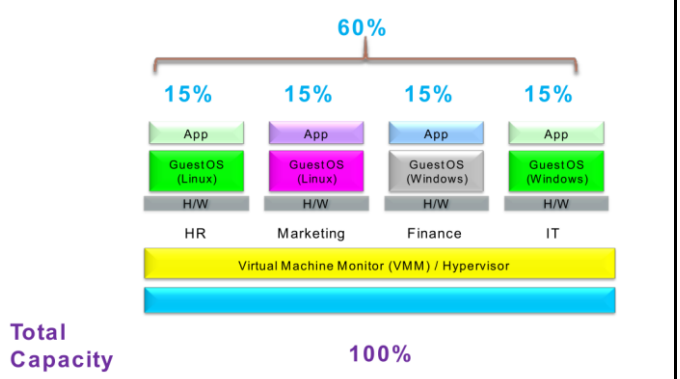
- Cloud Computing Economics is based on the pay-as-you-go method
- Only for the usage of cloud services.
- Involves two primary principles: **economies of scale** and **global reach**.
- From an end-user perspective cloud computing gives the illusion of a potentially infinite capacity with the ability to rapidly leverage additional capacity when needed, and pay only for what one consumes.
- When exploring cloud economics, a company can follow the procedure that includes,
  - **Benchmarking:** Calculate cost of operating current data centre including capital cost.
  - **Cloud costs:** Estimate the cost of cloud infrastructure (private, public or hybrid). Receive quotations from different CSPs and compare the integration cost, security and compliance points.
  - **Migration costs:** Determine the cost to migrate IT operations to the cloud.
- Based on these costs, ROI and TOC are calculated that are used to make decisions.
- **Cost benefits of Cloud Computing**
  - Converts fixed costs into variable costs
  - Reduces the capital costs of infrastructure.
  - Removes the maintenance cost.
  - Removes the administrative cost.

	In-house server	Cloud server
Purchase cost	$9600	
Cost/hr (over 3 yrs)	$0.36	$0.68
Price: Cloud/In-house	1.88	
Efficiency	40%	80%
Cost/Effective-hr	$0.90	$0.85
Power and cooling	$0.36	
Management cost	$0.10	$0.01
Total Cost/Effective-hr	$1.36	$0.86
Cost ratio: In-house/Cloud	**1.58**	

>> For low or variable-volume web-based services PaaS is a cheaper alternative.
>> For heavier, back-end applications (such as those behind web front ends, or for batch processing), IaaS is better suited.

8.2. Economics of Private Clouds

- Through Private cloud inefficiencies of in house data, centres would disappear and the benefits of cloud computing could be achieved on-premises.
- Through Virtualization effective user of resources can be achieved in private cloud

In the house data centre	Private cloud with virtualization
	

8.3. Software Productivity in the Cloud

- Development and testing servers require a different environment than a production environment. Also, these development and testing servers become obsolete after release and waste resources.
- Virtualization can help in this case to meet the growing demand of servers, but the time for provisioning and configuring such servers may bottleneck projects with faster development cycles. For this reason, the public cloud is a better option to provision and release such infrastructure on demand.
- Stress testing during the initial stages is also not possible due to the lack of a proper environment, which is solved by the public cloud.
- The public cloud also enables globally distributed teams to work on a project, which is also known to boost team morale, including skills from different parts of the world. So, it is advantageous to use a public cloud that is centrally located build servers to provide low latency connection to the globally distributed team.
- Likewise, PaaS provides faster and easier deployment for software and provides better scalability as well.
- Dev 2.0 is another benefit of the growing use of software products in the cloud.

8.4. Economies of Scale: Public Vs. Private Clouds

- Public cloud providers enjoy purchasing hardware**,** storage and network are cheaper on large scales in the case of a public cloud than in a private cloud.
- Public cloud providers can gradually pay off the debt of server administration over a large number of servers by employing automation.
- Public cloud providers have their data centres at locations where power cost is less or where power is produced.
- Public clouds enjoy a high degree of utilization compared to private cloud
- Most popular public cloud vendors have pre-established data centres and employ cloud services using those resources at a high level. (Eg. Google, Amazon, Microsoft)

Q) An enterprise plans to host its MIS in the cloud.

a) Make an estimation of Monthly/Yearly cost if the following on demand services with utilization 20 hours/day are used from the cloud service provider.

Service	Charge Unit	On Demand cost (NPR)
Virtual server (2 vCPUs, 8 GiB Memory)	**Hourly**	**12**

30 GB SSD block storage	**Monthly**	**12**
SQL server instances installed on virtual server 2 vCPUs and 8 GiB memory	**Hourly**	**60**
500 GB storage for Database	**Monthly**	**28**

b) If the pricing model of the virtual server is changed to a full year service plan with the commitment of NPR 60000 and with full payment upfront. What will be the percentage change in the cost?

Solution: *Only an approximation calculation daily to monthly conversion is done by multiplying by 30, and monthly to yearly conversion is done by multiplying by 12*.

a) On Demand

For virtual server (per month) = 20 x 12 x 30= Rs. 7200

For 30 GB block storage (per month) = 30 x 12 = Rs. 360

For SQL server instance (per month) = 20 x 60 x 30 = Rs. 36000

For 500 GB database storage (per month) = 500 x 28 = Rs. 14000

Overall total cost (per month) = Rs. 57560

Total cost per year = Rs. 690720

b) Full year service plan

For virtual server (per month) = Rs. 60000/12= Rs. 5000

For 30 GB block (per month) = 30 x 12 = Rs. 360

For SQL server (per month) = 20 x 60 x 30 = Rs. 36000

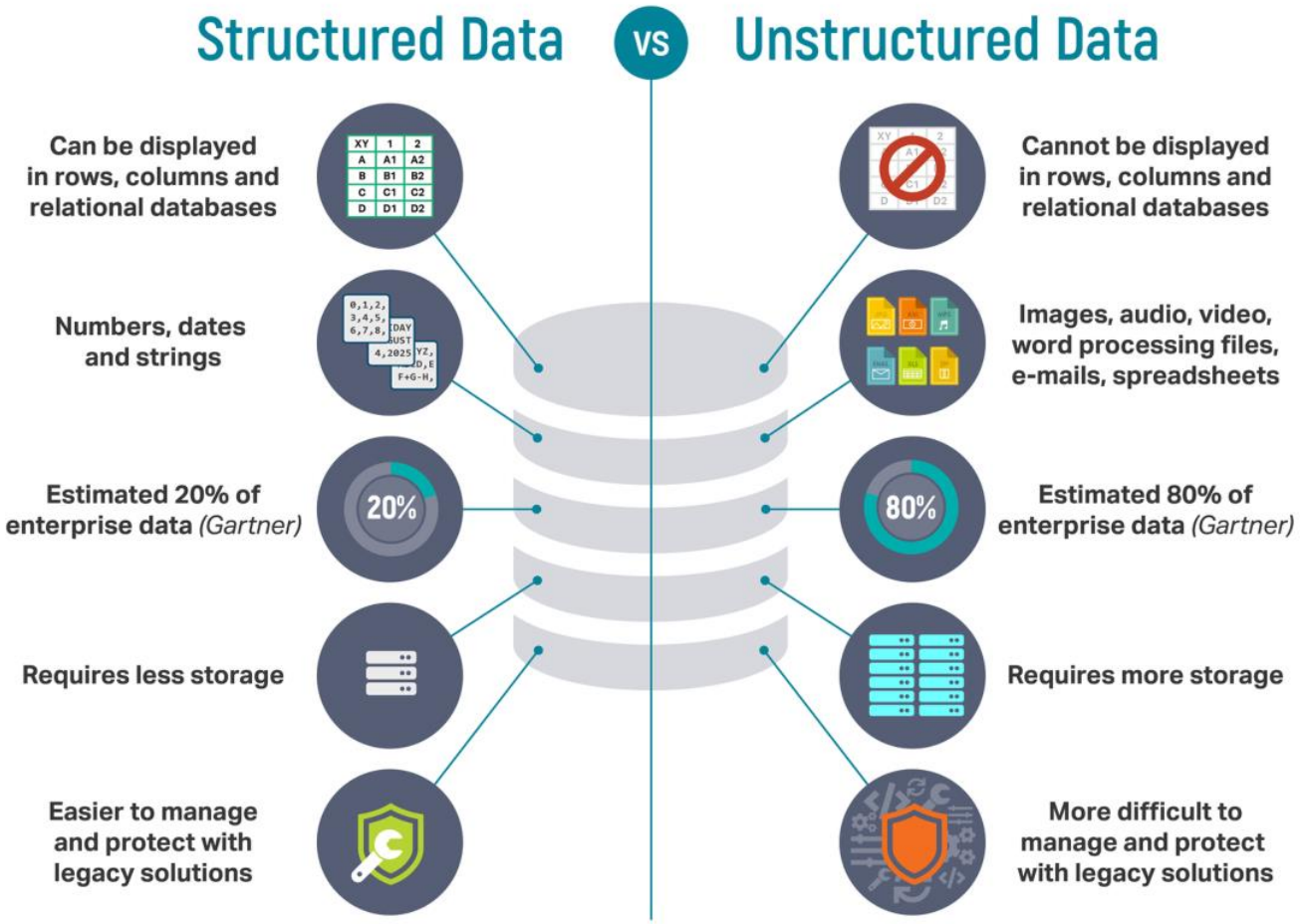
For 500 GB database storage (per month) = 500 x 28 = Rs. 14000

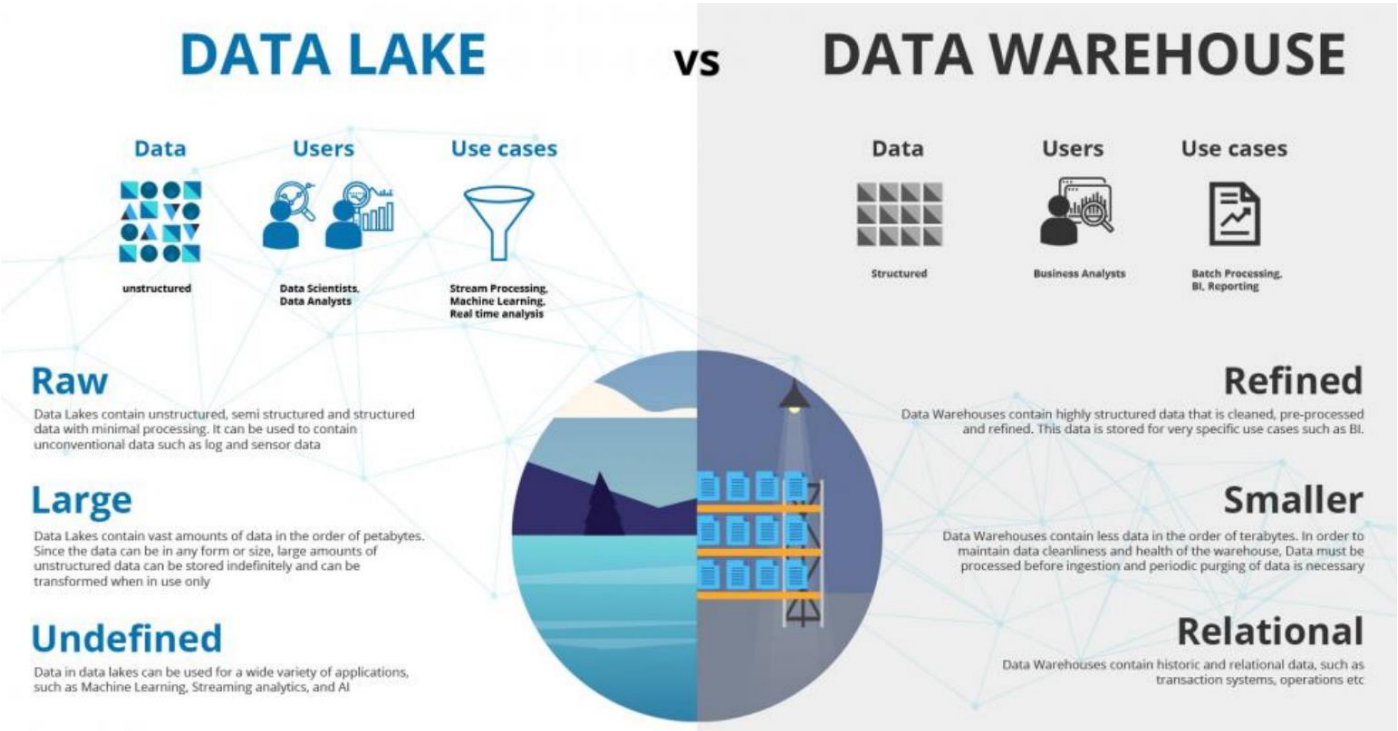
Overall total cost (per month) = Rs. 55360

Total cost per year = Rs. 664320

% change in cost = [(690720-664320)/690720] x 100 % = 3.82 %

9. Enterprise Analytics and Search





9.1. Enterprise Knowledge: Goals and Approaches

- **S**egmenting customers, suppliers, markets or employees based on behaviour.
- **T**argeting advertising and marketing campaigns.
- **D**etecting anomalies.
- **M**ultiple sources of data for a high-level explanation of ground-level evidence.
- **C**ustomer feedback, blogs and emails.

9.2. Business Intelligence

- Helps organizations analyze historical and current data, so they can quickly uncover actionable insights for making strategic decisions.
- Processing large data sets across multiple sources and presenting findings in visual formats that are easy to understand and share.
- **Steps** :
  - **Data collection**:- Collect structured and unstructured from multiple sources then remodelled so that analysis and query could be performed in it.
  - **Analysis/ Mining**:- Analyze data to find patterns and outliers which provide insight into the current state of business. (exploratory, descriptive, statistical, and predictive)
  - **Visualization**:- Make findings (results of analysis) easier to understand and share
  - **Decision/Action**:- quickly move from insights to action.
- **Benefits:**
  - Speed up information analysis and performance evaluation,
  - Reduce inefficiencies
  - Flag potential problems, detect anomalies
  - Find new revenue streams
  - Identify areas of future growth.
  - Tracking of sales, marketing, and financial performance.
  - Insight into customer behavior and shopping patterns.

9.3. Text And Data Mining

- TDM refers to a process of deriving high-quality information from text materials and databases using the software.
- Automated process of selecting and analyzing large amounts of text or data resources
- It is the process of transforming unstructured text into a structured format to identify meaningful patterns and new insights
- Text preprocessing is performed before text mining techniques. It involves the use of techniques such as language identification, tokenization, part-of-speech tagging, chunking, and syntax parsing to format data appropriately for analysis.
- **Applications of Text Mining:-**
  - Customer service
  - Risk management
  - Maintenance
  - Spam filtering
- **Limitation:-**
  - Incomplete information
  - Copyright

9.4. Text And Database Search

9.4.1. Web Search vs Enterprise Search

KEY	WEB	ENTERPRISE
Ranking	Most popular first. So the general assumption is that the most popular is best.	Enterprise needs “Correct“ information

Structure	Contains hyperlinks, so explicit information of inter-relation is available.	Hyperlinks are absent, due to which relationship between data has to be derived. In some cases, foreign links may be present, but often enterprises lack such linked databases.
Information restriction	Information is public on the web.	Information can't be public due to restrictions imposed and sensitivity of data involved
Format of data	Mostly textual and uniformly located via URLs.	A mixed bag of texts, documents, database formats with some foreign links available.

9.4.2. Why search structured data using text search instead of SQL?

- **SQL has a limited extent** to search using keywords and multiple screens need to be searched to access all occurrences of a keyword.
- **Enterprises have a large number of applications** with interrelated but independent databases, so text searching is cheaper and user friendly.
- **Multiple existences of the same keywords on multiple systems** are well discovered by automatically searching using common foreign key values.
- **The linkage between the data items** that are not explicitly maintained as joins.
- **Cloud databases lack support for some integral SQL functions** like ‘join’, so text search is the best way.
- **Unstructured and structured data are augmented** together so there may be a need to search both at once.

10. Enterprise Cloud Computing Ecosystem and Roadmap

10.1. Public Cloud Providers

10.1.1. Amazon Web Services (AWS)

- AWS, a subsidiary of Amazon, is the oldest cloud service, provider.
- Currently has 81 availability zones, which also makes it the cloud service provider among the discussed group with the most available global locations.
- High profile companies such as Netflix, Unilever, and Airbnb use AWS.
- It offers Platform as a Service (PaaS) with the use of AWS Elastic Beanstalk.
- It offers over 200 services, which is the most among the three under consideration.
- AWS has its primary focus on the public cloud rather than private or hybrid cloud models.

10.1.2. Microsoft Azure

- Microsoft Azure, a cloud service provider under the Microsoft group of companies has been providing its services since 2010.
- It falls just behind AWS in terms of global location spread.
- The reputed customer base of companies like Apple, Honeywell, HP and many more.
- It offers PaaS under the alias App Service and Cloud Services.
- Microsoft Azure's revenue has recently been reported to be better than both AWS and Google Cloud combined.
- Microsoft Azure's focus is divided among public and private cloud with enterprise customers being most attracted to the services.

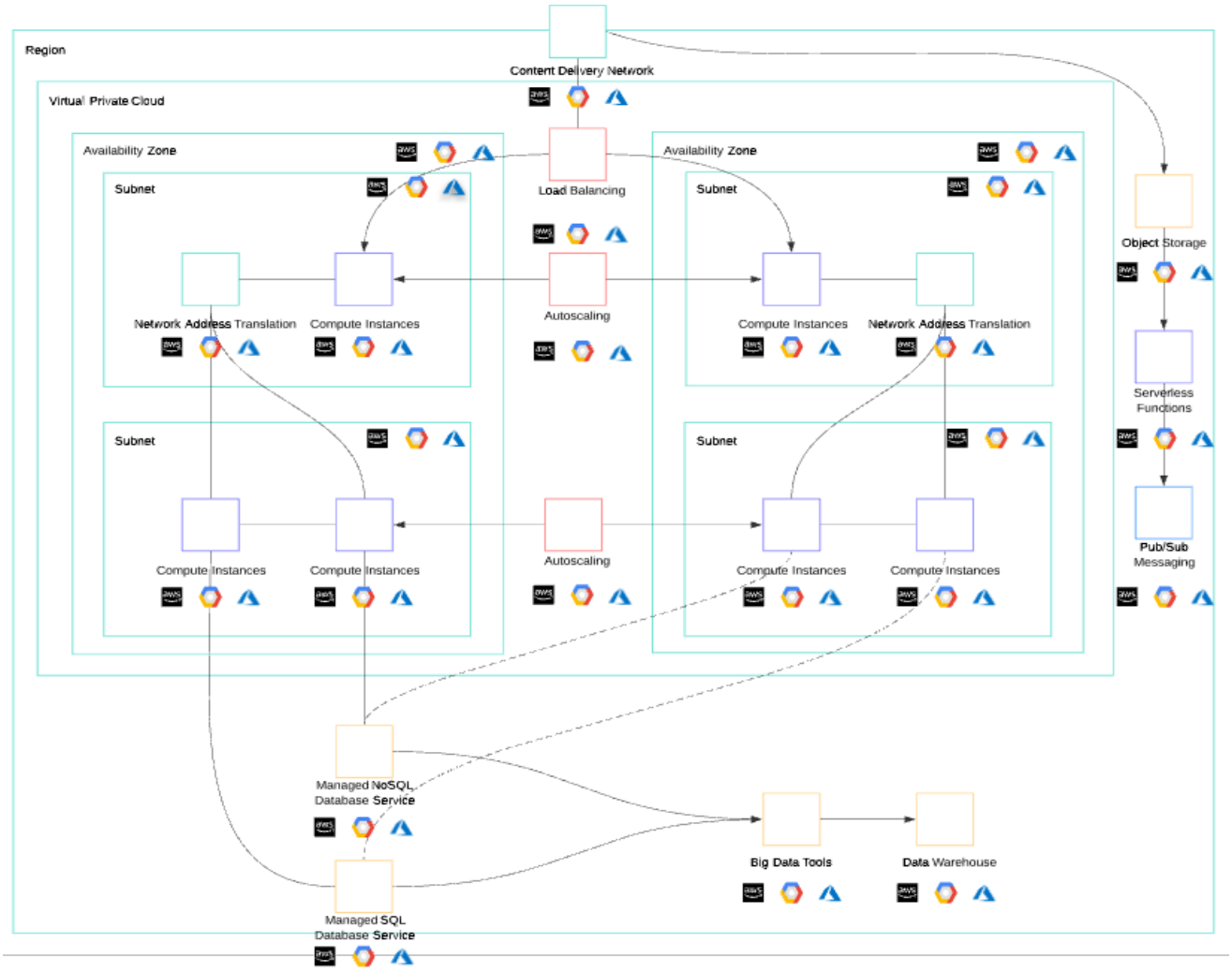
10.1.3. Google Cloud Platform (GCP)

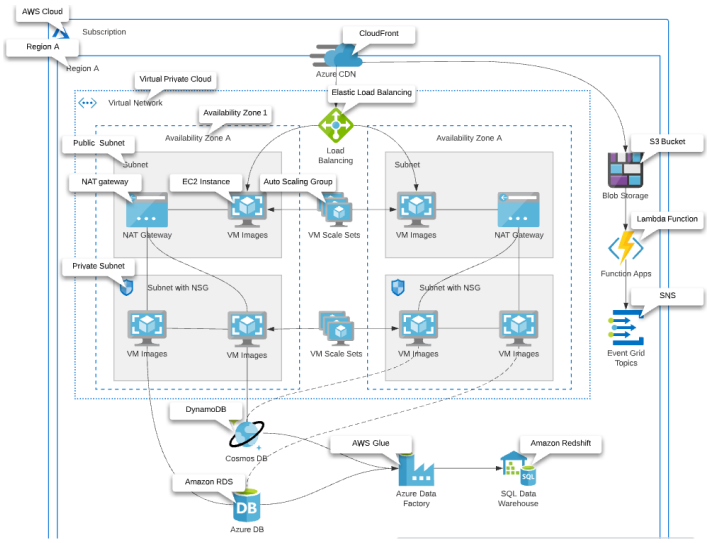
- GCP, offered by Google, is a bunch of cloud services that internally use the same resources used by YouTube, Google Search Engine and other Google products.
- It has the least global locations spread and offers over 60 services which are the least among the discussed options.
- GCP's client base includes companies like PayPal, Dominos, 20th Century Fox.
- It offers PaaS under the alias Google App Engine.
- Both AWS and Microsoft Azure provide pay-per-minute billing, however, GCP allows the customer to opt for pay-per-second billing which means customers save more with GCP than they do with AWS or Microsoft Azure.

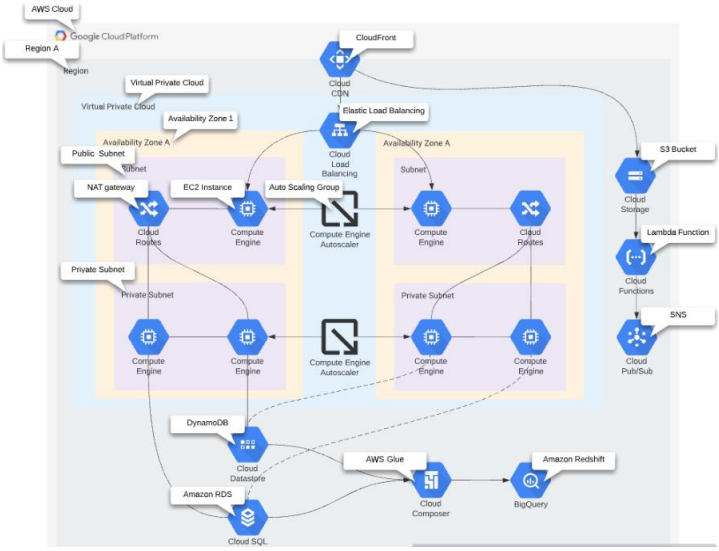
10.1.4. Selection Considerations

- Considering the establishment of the three service providers, AWS is the oldest and the most experienced one in the domain, leading it to occupy one-third of market shares.
- However, GCP has the best growth rate amongst the three.
- With over 200 services, AWS offers the most and with over 60 services GCP offers the least number of services.
- The pricing model offered by GCP is the most flexible.
- When it comes to open-source integration and on-premise systems, Microsoft Azure has the most advantage.
- Considering the brands that already use the services, all platforms are considered equal.

	AWS	GCP	Azure	Generic	Description
Networking	Virtual Private Cloud (VPC)	Virtual Private Cloud (VPC)	Virtual Network (VNet)	Virtual Private Cloud (VPC)	Logically isolated section of the cloud where you can launch resources
	Region	Region	Region	Region	Worldwide locations where resources are hosted
	Availability Zone	Availability Zone	Availability Zone	Availability Zone	Distinct locations (within regions) where resources are housed - designed to be isolated from failures in other AZs
	Subnet	Subnet	Subnet	Subnet	Logical subdivision of an IP network. Can be private or public
	Cloudfront	Cloud CDN and/or CDN Interconnect	Azure CDN	Content Delivery Network (CDN)	Global network that delivers content based on the geographic locations of the user
	Route 53	Cloud DNS	Azure DNS and/or Traffic Manager	Domain Name System (DNS)	Naming database in which internet domain names are located and translated into IP addresses
	IGW	-	-	Internet Gateway	VPC component that allows communication between instances in VPCs and the internet
	Route Tables	Routes	Azure Virtual Network Routing	Network Routes/Routing	A set of rules that are used to determine where network traffic from subnets and/or gateways are directed
	NAT Gateways	Cloud NAT	NAT Gateways	Netwrok Address Translation (NAT)	Enable outbound internet traffic from instances in a private subnet
	Virtual Private Gateway	Cloud VPN	Azure VPN Gateway	VPN Gateway	Private connection to VPCs
	VPC Peering Connections	VPC Network Peering	Virtual Network Peering	Network Peering	A networking connection between two VPCs that enables the routing traffic between them using private IP addresses
	VPC Endpoints	Private services, Private Google Access, and/or Shared VPC	Virtual Network Service Endpoint	VPC Endpoints	Privately connects VPC to other cloud and endpoint services
Security	Security Group	Compute Engine firewall rules	Network Security Group	Security Groups	Contains a list of security rules that allow or deny network traffic to resources
	Identity Access Management (IAM)	Cloud IAM	Azure Role-Based ACL (RBAC) and/or Azure AD	Identity and Access Management (IAM)	Manage access to cloud services and resources securely
	Cross-account Roles	Cloud IAM - Service Account	RBAC - Guests	Third Party and/or Guest Access	Grant access to resources in your cloud account, another cloud account you own, or a third-party account
Load Balance	Auto-scaling group	Compute Engine Autoscaler	VM Scale Sets	Autoscaling	Scales resources (of a group) up or down based on usage and rules
	ELB	Cloud Load Balancing	Load Balancer and/or Application Gateway	Load Balancing	Pushes traffic (in a balanced way) across multiple targets in multiple AZs
Compute	Elastic Compute Cloud (EC2)	Compute Engine	Azure VM	Compute Instances	Scalable compute compute instances (servers)
	Elastic Kubernetes Service (EKS)	Google Kubernetes Engine (GKE)	Azure Kubernetes Service (AKS)	Managed Kubernetes (K8s) Service	Allows you to run K8s without needing to install and operate your own K8s cluster
	Fargate	Cloud Run (+GKE)	Azure Container Instances (ACI)	Serverless Container Service	Run containers without managing servers
	Lambda Functions	Cloud Functions	Azure Functions	Serverless Functions	Event-driven, serverless computing service that runs code in response to events and automatically manages the computing resources required by that code
Data	S3 Buckets	Cloud Storage	Blob Storage	Object Storage	Scalable, highly available storage for objects
	Relational Database Service (RDS)	Cloud SQL	SQL Database, Azure Database for MySQL/PostgreSQL	Managed SQL Database Service	Distributed relational database service capable of running multiple sql engines
	Aurora	Cloud Spanner	Azure SQL Database (Managed Instance)	Managed SQL Database Engine	High performance managed sql engine that runs on a cloud's managed sql service
	DynamoDB	Cloud Firestore, Cloud Bigtable, and/or Cloud Datastore	Cosmos DB	Managed NoSQL Database Service	Distributed non-relational database service
	Redshift	BigQuery	SQL Data Warehouse	Data Warehouse	Central repositories of integrated data from one or more disparate sources
	Kinesis	Dataflow (processing) and/or Pub/Sub (ingest)	Stream and/or Data Lake Analytics	Managed Data Stream Processing Service	
Misc	SNS Topics	Cloud Pub/Sub	Event Grid	Pub/Sub Messaging	Fully managed pub/sub messaging service that enables you to decouple microservices, distributed systems, and serverless applications
	SQS Queues	Cloud Pub/Sub	Azure Queue Storage and/or Service Bus	Messaging Queue Service	Fully managed message queuing service that enables you to decouple and scale microservices, distributed systems, and serverless applications.
	Step Functions	-	Logic Apps	Serverless Orchestration	Task, processes, and workflow orchestration
	API Gateway	Cloud Endpoints	API Management	API Endpoints	Single entryway that allows multiple APIs or microservices to act cohesively


Region
Content Delivery Network
Virtual Private Cloud
Availability Zone
Availability Zone
Subnet
Subnet
Load Balancing
Object Storage
Network Address Translation
Compute Instances
Autoscaling
Compute Instances
Network Address Translation
Serverless Functions
Subnet
Subnet
Pub/Sub Messaging
Compute Instances
Compute Instances
Autoscaling
Compute Instances
Compute Instances
Managed NoSQL Database Service
Big Data Tools
Data Warehouse
Managed SQL Database Service


AWS Cloud
Region A
Subscription
CloudFront
Azure CDN
Virtual Private Cloud
Virtual Network
Elastic Load Balancing
Availability Zone 1
Availability Zone A
Public Subnet
Subnet
Load Balancing
Availability Zone A
S3 Bucket
NAT gateway
EC2 Instance
Auto Scaling Group
Blob Storage
NAT Gateway
VM Images
VM Scale Sets
VM Images
NAT Gateway
Lambda Function
Private Subnet
Function Apps
Subnet with NSG
Subnet with NSG
SNS
VM Images
VM Images
VM Scale Sets
VM Images
VM Images
Event Grid Topics
DynamoDB
Cosmos DB
AWS Glue
Amazon Redshift
Amazon RDS
Azure Data Factory
SQL Data Warehouse
Azure DB


AWS Cloud
Google Cloud Platform
Region A
CloudFront
Cloud CDN
Virtual Private Cloud
Virtual Private Cloud
Availability Zone 1
Elastic Load Balancing
Availability Zone A
Availability Zone A
Public Subnet
Subnet
Cloud Load Balancing
Subnet
S3 Bucket
NAT gateway
EC2 Instance
Auto Scaling Group
Cloud Storage
Cloud Routes
Compute Engine
Compute Engine Autoscaler
Compute Engine
Cloud Routes
Lambda Function
Private Subnet
Private Subnet
Cloud Functions
Compute Engine
Compute Engine
Compute Engine Autoscaler
Compute Engine
Compute Engine
SNS
Cloud Pub/Sub
DynamoDB
Cloud Datastore
AWS Glue
Amazon Redshift
Amazon RDS
Cloud Composer
BigQuery
Cloud SQL

10.2. Cloud Management Platforms and Tools

- Cloud management platforms provide web-based graphical tools to configure and manage complex configurations of servers deployed in the cloud.
- 3tera, RightScale, Kaavo, EnStratus, and Elastic
- Some tools may only work with selected vendors
- These platforms are themselves deployed on the cloud, either by agreement with partner service providers or some smaller hosting providers.

- Cloud management tools also offer dynamic monitoring and load balancing.
- Nowadays these tools are deeply integrated within CSP architecture for example:- Amazon has Elastic load balancing, CloudWatch and Autoscaling
- Hosted technical software for email security.
- rPath, AppZero, MessageLabs, Niche.

10.3. Tools for Building Private Clouds

- Private clouds use virtualization technologies to improve efficiency by improving server utilization.
- However, self-service infrastructure provisioning isn't possible.
- There are emerging tools provided by the vendors that distribute virtualization that provide self-service infrastructure provisioning along with dynamic monitoring and load balancing.
- VMware, GigaSpaces, Elastra, Anomaly, 3tera.

10.4. Future of Enterprise Cloud Computing

- There is a trend digital enterprise depends on technology to deliver competitive services and strong customer experiences
- Is and will be center for emerging technology like AI, ML, IOT, Edge computing, Quantum computing
- Adoption or migration of public cloud from inhouse data center is and will be rising
- Enterprise are embracing multi cloud in combination with hybrid cloud
- Desktop as a Service (DaaS) is a trend as work from home and remote login are becoming a necessity.
- Cloud-Native Applications that are completely powered by cloud will rise in coming year
- Disaster Recovery as a Service (DRaaS) has not reached its maturity and will still evolve in future.
- XaaS is any IT infrastructure as a service
- Some other trending topics in cloud computing will be AI, Block chain, Application Mobility, Serverless computing etc.