1. Define the term cloud computing and explain various goals of cloud computing and advantages and disadvantages.

   Cloud computing: cloud Computing is a model for enabling ubiquitous, Convenient on demand network access to a shared pool of configurable computing resources, such as networks, servers, storage applications and services with minimal management efforts or service provider interaction.

   Goals of cloud computing:

   - Quick time value: Cloud computing allows you to quickly consume the IT resources and focus on key areas of your business without having to Worry about managing an underlying infrastructure.
   - Reduce cost: Without having to put capital investment in procuring the hardware, Cloud computing enables you to pay for what you use.
   - Infinite scaling: Cloud providers typically have massive amount of computing resources. These resources can be dynamically consumed based on your needs.
   - Maximum availability: Cloud provider maximum availability of computing resources. The resources are aggregated in large resource pool. If any of resources is temporarily out of service, it is easily replaceable. The cloud services providers typically provide service level agreements SLA with respect to service that they provide.
   - Rapid innovation: Cloud providers are rapidly innovating to brief new services and to deliver optimum performance at the cheapest possible cost. You directly benefit from any key technology breakthrough or any cost savings arising out of such innovation.

   Advantages:

   I. Cost effective
   II. Massive resources to consume.
   III. No maintenance required.
   IV. Lesser liabilities and commitments
   V. Quick time the volume with increased focus on business.

   Disadvantages:

   I. Limited flexibility and customization
   II. Vendor lock-in.
   III. Lower control on day to day infrastructure operations
   IV. Increased burden of security and compliance.
   V. Requires high speed network connectivity.

2. List the cloud computing characteristics and explain them briefly.

   Characteristics:

   I. On Demand self-service:
      This is perhaps the most important characteristic of cloud computing.
      it means whenever you desire to acquire information technology resources such as computers, storage, network, applications, and several others, these resources are available for immediate consumption.

You don’t require procuring hardware, Provision the servers, installing and deploying applications or Carrying out any placement or maintenance planning Before consuming the resources for your requirements.

II. Broad network access:
Capabilities are available over the network and access through standard mechanisms that promote use by heterogeneous thin or thick client platforms Such as mobile phones, tablets, laptops and workstations.

III. Resource pooling:
The cloud computing service provider has a wider resource pool of IT resources such as CPU, memory, disc and network. These resource pool is used to serve multiple users.
In cloud computing world. A user or a company consuming the cloud resources is called tenant.
There is a sense of location independence in that the consumer generally has no control or knowledge over the exact location of The provided resources but may be able to specify Location at a higher level of abstraction.

IV. Rapid elasticity:
Rapid elasticity of cloud allows you to dynamically change and accessed your cloud resources consumption based on your needs. You are not stuck on choosing one set of resource subscription and then facing inflexibility as you are dim and grows or comes down.

V. Measured service:
Any service model is billed on actual consumption of resources.
Cloud services are built according to what you have subscribed to and consumed. You are not directly or additionally for any hardware or any operational costs associated with providing the services to you. If you consumed 5 hours of computing, you are billed for 5 hours, no more and no less.
The cloud providers have a mechanism to measure the services provided to you and then charge you based on that measurement. Different types of resources are billed differently. For example, CPU is billed according to its clock.

3. Explain cloud delivery models with examples.

4. Comparison between software as a service platform, as a service and information as a service.

Basis Of	IAAS	PAAS	SAAS
Stands for	Infrastructure as a services.	Platform as a services.	Software as a services.
Uses	IAAS is used by network architects.	PAAS is used by developer.	SAAS is used by end user.

Access	IAAS give access to the resources like virtual machines and virtual storage.	PAAS give access to run time environment to deployment and development tools for application.	SAAS give access to the end user.
Model	It is service model that provide visualized computing resources over internet.	It is a cloud computing model that delivers tools that is used for development of application.	It is a service model in cloud computing that host software make available for client.
Technical understanding.	It required technical knowledge.	In which you required knowledge of subject to understand basic setup.	There is no requirement about technicalities company handle everything.
Popularity.	It is popular between developer and researchers.	It popular between developer who focus on the development of apps and scripts.	It is popular between consumer and company. such as file sharing, email and networking.
Cloud services.	Amazon web services, sun, vcloud express.	Facebook, and google search engine.	M.S office web, Facebook and google apps.
Enterprise services.	AWS virtual private cloud.	Microsoft azure.	IBM cloud analysis.
Outsourced cloud services.	Salesforced.	Force.com, Gigaspaces.	AWS, terremark

5. Explain deployment model of cloud computing.

   The cloud deployment model identifies the specific type of cloud environment based on ownership, scale, and access, as well as the cloud's nature and purpose. The location of the servers you're utilizing and who controls them are defined by a cloud deployment model.

   It specifies how your cloud infrastructure will look, what you can change, and whether you will be given services or will have to create everything yourself. Relationships between the infrastructure and your users are also defined by cloud deployment types.

Different types of cloud computing deployment models are:

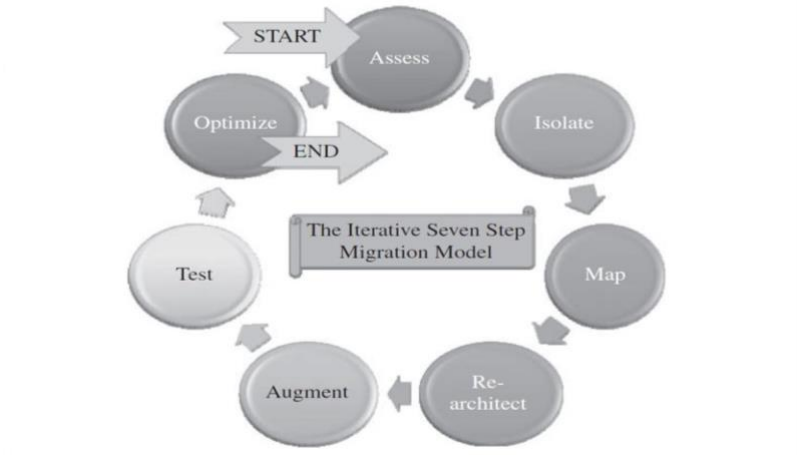
I. Private cloud:
   - Private Cloud is created for self-use of any organization. The organization itself creates a private cloud or it can let a third party vendor set it up and manage on behalf of it.
   - The resources in the private cloud are for exclusive use of the organization for which The private cloud is created.
   - The organization can choose to have its private cloud deployed within its physical parameter or outside of the company boundaries.

II. Public cloud:
   - Public Cloud is the cloud deployment environment that is open to public use.
   - It is not tight to particular organization's usage exclusively.
   - This is the most prevalent form of cloud deployment options today. Public cloud providers serve multiple tenants by sharing an isolating the computing resources.

III. Hybrid cloud:
   - Hybrid cloud is necessarily an interconnection between any of the three types of deployment model discussed as private, public or community.
   - Hybrid cloud can also be created by extending your datacentre and connecting to any of the Cloud deployment models.
   - Organization usually do this to get benefit from that respective characteristics and to overcome any limitations or challenges.

IV. Community cloud:
   - Community Cloud is similar to private cloud, except one key difference that instead of single organization for which the private cloud is created, Community Cloud is created by and for a group of organizations.
   - These organizations are similar in nature in terms of their mission, business, market requirements, policies, legal implications, policies, legal implications, compliance and customers.

6. Describe seven step model or migration into a cloud.



I. Assess: The cloud migration journey starts with an assessment of issues relating to migrations at application, code, design and architecture level. Moreover, assessments are also required for tool being used, functionality test cases and configuration of the application.
II. Isolate: the second step is the isolation of all the environmental and systematic dependencies of the enterprise applications within the Data center. These include library, application and architectural dependencies. This step results in a better understanding of the complexity of the migration.
III. Maps: a mapping construct is a generated to separate the components that should reside in the data center from ones That will go into the cloud.
IV. Re-architect: It is likely that a substantial part of the application has to be re architecture and implemented in the cloud. This can affect the functionality of the application and some of these might be lost. You could also return mine if there is another migration strategy such as we host or re platform that would be better suited.
V. Augment: The features of cloud computing services are used to augment the application.
VI. Best: Once the augmentation is done The application needs to be validated and tested. This is to be done using test suit for the applications on the cloud. New test cases due to augmentation and proof of concept are also tasted at this stage.
VII. Optimise: the test results from the last step can be mixed and so require iteration and optimization. It may take several optimizing iterations for the migration to be successful. It is the best to iterate through the seven step model as this will ensure the migration to be reversed and comprehensive.

7. Write a short note on data storage and explain types of data storage system.
Data storage:

- There are two types of digital information: input and output data. Users provide the input data. Computer provides output data. But the computer CPU can't compute anything or produce output data without the users input. User can enter the data directly into the computer. However, we have found early on in the computer era that continually entering data manually in time and energy consuming.
- Big data storage space user can save data onto a device and it is saved even if the device is powered down. Also, instead of manually entering data into a computer, you can instruct the computer to pull data from storage devices.
- Computers can read input data from various sources as needed, and it can then create and save the output to the same source or other storage locations. You can also share data storage with others.

Types of Data Storage Systems:

I. The storage device which is permanently attached to a desktop computer. DAS is for a single user (Hard drive attached to a computer). DAS is well suited for a small-to-medium sized business where sufficient amounts of storage can be configured at a low startup cost. The DAS enclosure will be a separate adjacent cabinet that contains the additional disk drives.

Component of Directly Attached Storage (DAS) –

1. Storage devices
2. Cables
3. Disk Array
4. Protocol
5. Storage protocols: ATA, SCSI, SAS, SASA, FC

II. **Network Attached Storage (NAS):**

This Storage Device is attached on the Local Area Network and used for sharing of data among different users attached to the Local Area Network. Instead of accessing data at the sector level, users can access information on file level over the network. This NAS system is having its own file system, which is once set with proper configuration of NAS and is not dependent upon the operating system of computers from which it is connected. This type of network requires a medium for attaching with several computers. File sharing protocols like NFS, AFP, or CIFS provide access to files in a network.

Components of Network Attached Storage (NAS) –

1. Head unit: CPU, Memory
2. Network Interface Card (NIC)
3. Optimized operating system
4. Protocols
5. Storage protocols: ATA, SCSI, FC

8. Explain general architecture of cloud file system.

Typically, architecture of cloud file system falls into two categories:

1. Client server architecture:
   - In client-server architecture, the file server hosts the file system that can be mounted by the clients. One file server can host multiple file shares and each file share can be mounted and operated by multiple clients. All file operations are then synchronized back to the file server so that the other clients that have mounted the same file share can get updates as well.
   - Example is NFS (network file system)

then synchronised back to the file server be

ates as well.

ows block diagram that depicts the client-server architecture at a high-level.

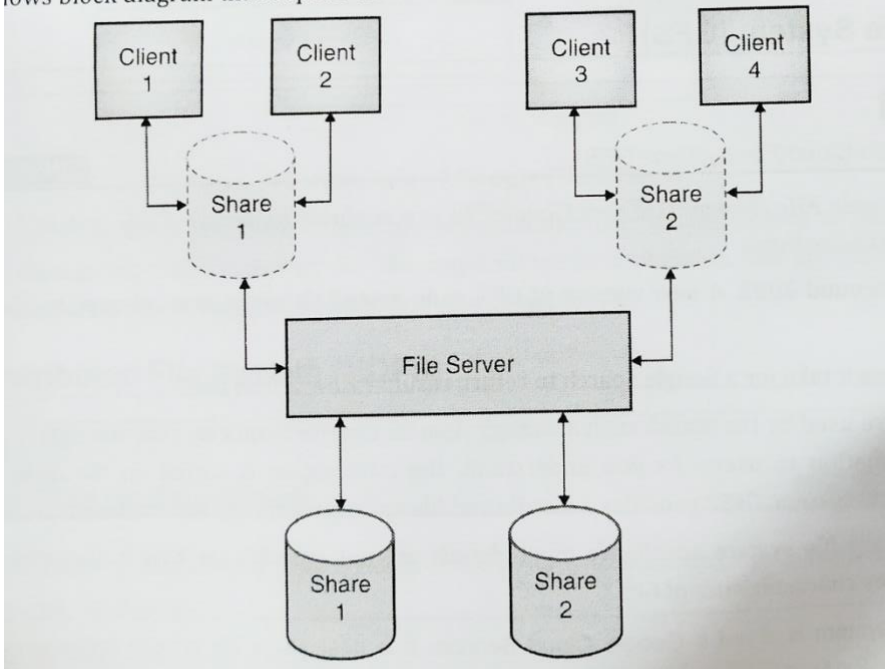


Fig. 2.2.1 : Client-Server Architecture

6/2019)

2. Clustering based architecture.
    - In cluster based architecture the file is broken into smaller parts called chunks and each chunk is stored on the storage server. The chunks are redundantly stored on the several servers to, with stand any faults and have high availability. This architecture does not depend upon a single server for hosting the file system. The file system is distributed and provides parallelism That significantly improves the scale and performance. This architecture is commonly used today in the cloud environment.
    - Examples are Google File system, Amazon S3, etc.

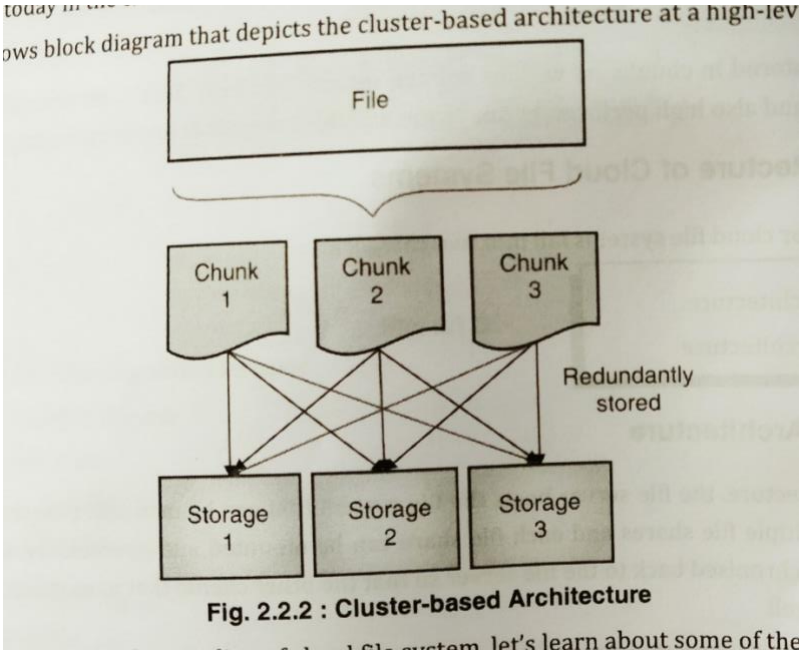


Fig. 2.2.2 : Cluster-based Architecture

9. Draw and explain cloud file system HDFS.

How do distributed file system(HDFS):

- The hard up distributed file system is distributed file system designed to run on general and low cost hardware.
- It is mainly designed for processing large data sets.

Architecture of HDFS:

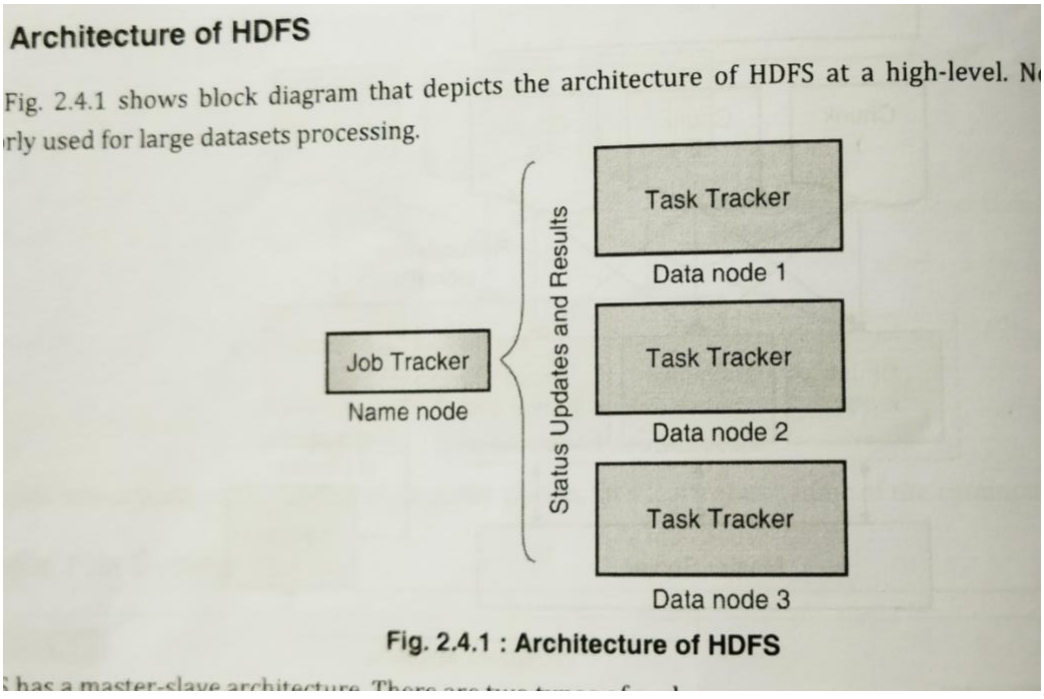


Fig. 2.4.1 : Architecture of HDFS

- HDFS has master slave architecture. There are two types of nodes:

i. Name notes: HDFS cluster consist of a single NameNode. The NameNode runs a master server that manage is the file system and controls the access to the files by clients.

ii. Data notes: There can be multiple data nodes. data nodes are processing units in HDFS cluster. Each data node manage is the storage attached to it.

- HDFS exposes a file system grouping and allows user data to be stored in files. Internally a file it split into one or more blocks and then these blocks are stored in the set of data nodes.
- The name node executes file system Namespace operations like opening, closing and renaming files and directories. It also determines the mapping of blocks to data node.
- The data nodes are responsible for serving read and write requests from the file systems clients. The data node also performs block creation, deletions, and replication upon instruction from the name node.

10. Explained the architecture of H-base.

Architecture of the H-base:

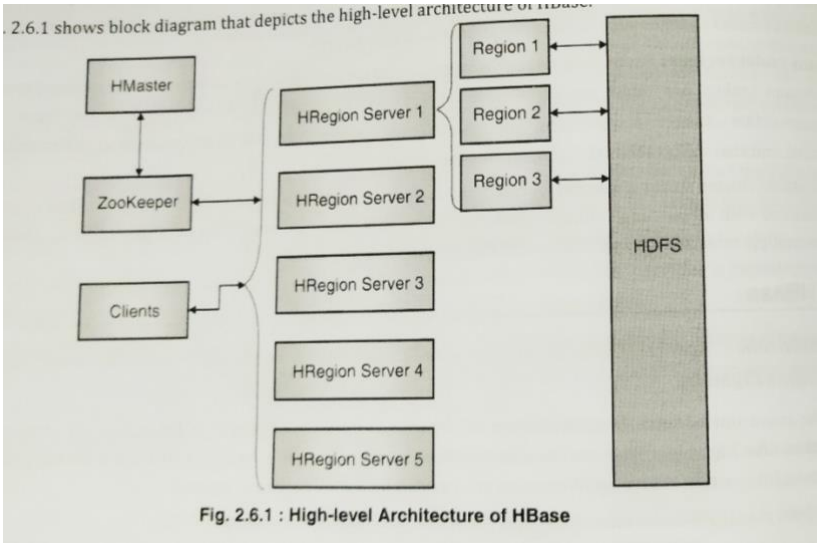


Fig. 2.6.1 : High-level Architecture of HBase

i. HDFS: all H base data is stored on HDFS.

ii. Regions: table teenage base are divided horizontally by rookie range into regions. Region contains all rows in the table between the regions start key and end key. Regions are assigned to the nodes in the cluster called region servers and these serves data for reads and writes to the clients. Region server can serve around 1000 regions.

iii. Master server: the master server coordinates the Cluster server and performs. Administrative operations. Such as assigning regions to the region servers and balancing the load. It also performs other administrative operations, such as creating undeleting tables.

iv. Region server: The region servers perform data processing. Each region server stores a subset of the data of each table. Clients talk to region servers to access the data in Hbase.

v. Zoo keeper: Zoo Keeper is a centralized service for maintaining configuration information, naming and providing distributed synchronization and providing group services. Hbase uses zoo keeper as a distributed coordination service to maintain server state in the cluster. Zoo Keeper maintains which region servers are alive and available and providing servers failure notification. To the master server to coordinate administrative tasks such as region assignment. Zoo Keeper uses consensus to guarantee common shared state.

vi. Clients: clients work with the HBS data as appropriate. Clients directly connect to the region servers for accessing the data or performing any operations on data.

11. Explain the characteristic of Dynamo.

Characteristics of the Dynamo:

i. NoSQL database: Dynamic provides a key value storage system. It does not provide relational database schemes. It provides simple read and write operations on data items that are uniquely identified by keys.

ii. ACID properties: ACID is a set of properties that guarantee that database transactions are processed Reliably. Usually, data stores that provide ACID guarantees tend to have poor availability. Dynamo targets applications that operate with weaker consistency, and this results in high availability. Dynamo does not provide any isolation guarantees and permits only single key updates at a time.

iii. High performance: Dynamo can run on commodity hardware and provides single digit latency.

iv. Incremental scalability: Dynamo can scale out one storage host at a time with minimal impact on both operators of the system and the system itself.

v. Symmetrical: every node in Dynamo have the same set of responsibilities at its peer. There is no distinguished node or nodes that can take special role or extra set of responsibilities. This symmetry specifies the process of system provisioning and maintenance.

12. Explain the architecture of the big table.

Architecture of the big table:

- The big table implementation has three major components.
  i. A library that is linked to every client: The library is the interface between your application code and big table. Clients communicate directly with the table servers from Read and writes. big Table clients do not rely on the master for table location information and hence most clients never communicate with the master. As a result, the master is lightly loaded in practice.
  ii. One master server: The master server is responsible for assigning. Cables and tablet servers. Detecting the addition and expiration of table servers balancing tablet server load and garbage collection of files in GFS. In addition, it handles schema changes such as table and column family creations.
  iii. Menu tablet servers: Each table server manage is a set of tables. Tablet servers can be dynamically added or removed from the cluster to accommodate

changes in workloads. The tablet server handles read and write a request to tablets that is has loaded and also split tablets that have grown too large.

- A big data cluster stores a number of tables. Each table consists of set and tablets, and each tablet contains all data associate with a row range. Initially each table consist of just one tablet. As the table grows, it automatically splits into multiple tablets, each approximately 100 to 200 MB insights by default.