(d) Data transfer b/w nodes : 1 GB/sec

Encryption overhead : 5% = 0.05

Effective data transfer rate : 1 Gbps − encryption overhead

= 1 − 0.05

= 0.95 Gbps.

* Unit - 5 :-

① HDFS storage Mechanism :-

* Blocks :- Large files are divided into blocks, which are distributed across Data Nodes

* Replication :- Each block is replicated (typically three times) to ensure fault tolerance.

HDFS Components :-

* Name Node :- Manages metadata (file structure) but does not store actual data.

* Data Node :- stores actual data blocks and handles read/write operations.

Architecture Comparison :-

* peer-to-peer :- Each node has equal responsibilities coordination and consistency are challenging.

* master-slave (HDFS) :- Name Node manages metadata ; Data Node stores data

Simplifies metadata management and fault tolerance through data replication.

* Real-Time Examples :-
  * yahoo Search
  * Facebook

(9) Predict the cloud services used in virtual meeting (Gmeets, Zoom, etc)

1. Compute and virtual machines :-

Function :- provide scalable computing resources to host meeting applications.

Examples :- Google Compute Engine, Amazon EC2, Azure Virtual machines

2. Real-Time Communication Services :-

Function :- Enable audio and video communication between participants in real-time.

Examples :- web RTC, Amazon Chime SDK, Twillo Programmable Video.

3. Content Delivery Networks (CDNs) :-

Function :- Distribute audio and video streams efficiently to reduce latency.

Examples :- Google cloud CDN, Amazon cloud Front, Azure CDN

4. Storage services :-

Function :- store meeting recordings, chat logs and other data.

Examples :- Google cloud storage, Amazon S3, Azure Blob storage

5. Identity and Access Management (IAM) :-

Function :- manage user identities and control access to meeting resources securely

Examples :- Google Identity, AWS, IAM, Azure Active Directory.

(A) Outline about the cloud services used in Google Colab.

Google colab, short for "Colaboratory", is a free cloud service provided by Google Colab that allows users to write and execute python code in a web-based, interactive notebook environment.

1. Google Compute Engine (GCE) :-

* Function :- provides the virtual machines for running colab notebooks.
* Uses :- offers scalable and powerful Compute resources for executing code.

2. Google drive Integration :-

* Function :- Allows seamless access and storage of files.
* Uses :- Users can save and load notebooks and datasets directly from google Drive.

3. Google cloud storage :-

* Function :- provides durable and scalable Object Storage.
* Uses :- Storing large datasets and files accessible from cola

4. Tensor processing Units :-

* Function :- It is a specialized hardware for accelerating machine learning computations.
* Uses :- Running deep learning models with more efficiency in colab.

5. Google cloud AI platform :-

* Function :- provides tools for building, training and deploying machine learning models.
* Uses :- Enhances machine learning workflows with colab, integrating with Other AI services.

(7) Discover the location services used in Google maps.

=> main location services used in Google maps

1. Gps (Google positioning system):-

* Function :- provides precise location data by triangulating signals from multiple satellites.

* Uses :- determining exact geographic coordinates

=> enabling turn-by-turn navigation.

2. Wifi positioning system :-

* Function :- uses the proximity of wi-fi networks to determine a device's location.

* Uses :- enhancing location accuracy indoors or in areas where Gps signals are weak.

=> providing location data in urban environments with dense WiFi networks.

3. Cell Tower Triangulation :-

* Function :- estimates a device's location based on its distance from multiple cell towers.

* Uses :- offering location data when Gps and WiFi signals are unavailable

4. Bluetooth Low Energy Beacons :-

* Function :- Uses Bluetooth signals from beacons for positioning

* Uses :- Indoor navigation and proximity-based services.

5. Crowdsource Data :-

* Function :- Aggregates anonymous location data from users to improve map accuracy and traffic predictions.

* Uses :- updating road conditions and maps

=> providing real-time traffic updates.

6. Offline maps and Local Data Storage :-

* Function :- Stores map data locally on the device for use without an Internet Connection.

* Uses :- providing navigation and location services in areas with poor connectivity.

⑥ Analyze the cloud services used in E commerce applications.

cloud services play a critical role in the infrastructure of E-commerce applications.

1. Infrastructure as a Service (IaaS)

Examples :- Amazon web Services, microsoft Azure, Google Cloud platform.

Uses :- Scalability :- easily scale up or down based on demand

Flexibility :- Customize the infrastructure to meet specific requirements.

Cost management :- pay only for resources used.

2. platform as a Service (Paas)

Examples :- Heroku, Google App Engine, AWS, Elastik Beanstal

Uses :-

Development Efficiency :- provides a platform to develop, run and manage applications without worrying about the underlying infrastructure.

Integration :- Simplifies integration with other services and applications.

Speed to market :- Accelerates the development process by providing preconfigured environments.

3 Software as a Service : (Saas)

Examples :- Shopify, magento, Big Commerce.

Uses :- Ease of use :- Ready to use applications for various e-commerce functions.

maintenance :- The provider manages updates, security, maintenance

Accessibility :- Accessible from any device with internet Connectivity.

(8) Hadoop clusters : Analysis :-

(i) Robustness :- Data Replication :- Ensures system resilience and availability.

(ii) Data Disk failures :-

Replication :- Data is replicated across multiple Data Nodes

Automatic Recovery :- Lost blocks are re-replicated

(iii) Heartbeats and Re-replication :-

Heartbeats :- Monitor Data Node health.

Re-replication :- Replaces lost blocks to maintain replica

(iv) Cluster Rebalancing :-

Purpose :- Distributes data evenly across the cluster t prevent hotspots.

(v) Data integrity :-

Checksums :- Validate and correct data corruption

(vi) metadata Disk Failure :-

Single point of failure :- Addressed with rumened backup and high availa setups.

(vii) Snapshots :-

purpose :- Capture filesystem state for back and recovery.

(4) Data transfer b/w nodes = 1 GB/sec

encryption overhead = 5% = 5/100 = 0.05

Effective data transfer rate = 1 Gbps - encryption overhead

= 1 - 0.05

= 0.95 Gbps.

* Unit - 5 :-

(1) HDFS storage Mechanism :-

* Blocks :- Large files are divided into blocks, which are distributed across Data Nodes

* Replication :- Each block is replicated (typically three) to ensure fault tolerance.

HDFS Components :-

* Name Node :- Manages metadata (file structure) but does not store actual data.

* Data Node :- stores actual data blocks and handles read/write operations.

Architecture Comparison :-

* peer-to-peer :- Each node has equal responsibilities. Coordination and consistency are challenging.

* master-slave (HDFS) :- Name Node manages metadata; Data Nodes stores data

Simplifies metadata management

tolerance through data replication

* Real-Time Examples :-

* yahoo! Search

* Facebook