**Lec 1:**

Cloud computing: cloud computing is the delivery of computing services—including servers, storage, databases, networking, software, analytics, and intelligence—over the Internet (“the cloud”) to offer faster innovation, flexible resources, and scalability. pay only for cloud services you use.

Less cost-- (pay when u only use), eliminates the capital expense of buying hardware, setting datacenters, power, cooling and IT maintainers and software

Speed-- Most cloud computing services are provided self service and on demand, so even vast amounts of computing resources can be provisioned in minutes.

Global scale/broad network access: access cloud services over the network from any geographical location and on any device.

Resource sharing/pooling: multiple customers can share physical resources.

Scalability/Elasticity scale down or up resources as per use.

Security: offers secure policy and certificates to secure client’s data and has excellent disaster recovery

All the services are offered and measured/monitored. They track usage and charge acc to that.

In cloud computing, "the cloud" is used to describe a network of remote servers and data centers that are accessed over the internet

**Lec 2:**

Public clouds

offered by third-party providers over public internet. Either free or pay per use. No need to manage as managed by third party. Shared b/w multiple clients.

Pvt cloud

A private cloud is a cloud computing environment that is dedicated to a single organization, managed and maintained. More security, control

It is a dedicated area of internet for a user or an organization.

Private clouds are often used by organizations that require a high level of control and data security, such as financial institutions, healthcare providers, and government agencies.

Hybrid: combination

Cloud bursting is an application deployment technique in which an application runs in a private cloud or data center and bursts into a public cloud when the demand for computing capacity spikes.

Multi cloud: more than 1 cloud provider

A multi-cloud strategy is beneficial when you want to avoid vendor lock-in, ensure redundancy and high availability, or take advantage of specialized services from multiple cloud providers. It is suitable for organizations that want to balance cost and performance across different cloud providers.

IAAS :Infrastructure as a service (IaaS) is a form of cloud computing that provides virtualized computing resources over the internet. physical and virtual servers, storage and networking - the backend IT infrastructure

PAAS: ready-to-use, cloud-hosted platform for developing, running, maintaining and managing applications.

Saas: n-demand access to ready-to-use, cloud-hosted application software.

A screenshot of a computer

Description automatically generated



Scaling, from an IT resource perspective, represents the ability of the IT

resource to handle increased or decreased usage demands”[1]

• Horizontal Scaling is the allocation or release of IT resources that are of the

same type

• Horizontal allocation is referred to as scaling out

• Horizontal release is referred to as scaling in

• Vertical Scaling is the replacement of an existing IT resource by another with

higher or lower capacity

• Replacing with higher capacity is referred to as scaling up

• Replacing with lower capacity is referred to as scaling down

**Lec 3:**

store data on remote servers. Data centers are placed in different continents to ensure efficient, and reliable service

structured data: has pattern and structure. Has no and values. Ex-normal data, rdbms

unstructured data: has no pattern and hard to analyse. Ex- videos,audio

DynamoDB: nosql db. Does not require to define database structure.

File storage is when all the data is saved together in a single file with a file extension type that’s determined by the application used to create the file

Block storage is when the data is split into fixed blocks of data and then stored separately with unique identifiers. The blocks can be stored in different environments, such as one block in Windows and the rest in Linux. Fastest

Object storage is ideal for storing and managing large volumes of unstructured data, images, videos, document. accompanied by metadata

Object better for searching in unstructured.

Block is raw and fast compared to others

In unmanaged storage, the storage service provider makes storage capacity available to users. Ex- drive or dropbox

In a managed cloud storage system, the user provisions storage on demand and pays for the storage using a pay-as-you-go model. Ex-s3

CSA (Cloud Security Alliance): CSA provides guidelines and best practices related to compliance,audit,security,policies

Cloud Control Matrix (CCM) – provides a security control list and framework that enables detailed understanding of security concepts and principles.

Symmetric encryption uses one key to encrypt and decrypt. Blowfish, AES, DES etc. better speed

Asymmetric encryption is different on each side; the sender and the recipient use two different keys. Asymmetric encryption, also known as public key encryption, uses a public key-private key pairing: data encrypted with the public key can only be decrypted with the private key. Better security

Q1

To make an informed decision, you should consider the specific needs of your car manufacturing plant's applications. If you require real-time data processing, analysis, and high write/read throughput, Google Cloud Bigtable might be the better choice. If you need to improve the performance of frequently accessed data and reduce latency for certain applications, Google Cloud Memorystore for caching could be a valuable addition to your architecture. In some cases, a combination of both services may be appropriate, depending on the different needs of various applications within the manufacturing plant.

Q2

In summary, whether object storage is better than block storage or vice versa depends on your specific needs. Object storage excels at handling large volumes of unstructured data with cost-effective scalability, while block storage is well-suited for high-performance applications that require low-latency and fine-grained control. In many scenarios, a combination of both storage types may be used to meet various application and data storage requirements within an organization.

Lec 4:

Virtualization is technology that you can use to create virtual representations of servers, storage, networks, and other physical machines. Ex- divide a pc into multiple virtual machines. VMs on the same server allows applications to better share the server. Better utilization and easily migrated

A diagram of a computer

Description automatically generated

VM migration steps: a VM is stopped, its state is saved as a file, the file is transported to another server,  
and the VM is restarted.

Hypervisor software that creates and runs virtual machines (VMs). Limited to 1 server only

A type 1 hypervisor, also referred to as a native or bare metal hypervisor, runs directly on the host’s hardware to manage guest operating systems. It takes the place of a host operating system and VM resources are scheduled directly to the hardware by the hypervisor. most common in an enterprise data center. Ex-KVM, Microsoft Hyper-V, and VMware vSphere

A type 2 hypervisor is also known as a hosted hypervisor, and is run on a conventional operating system as a software layer or application. It works by abstracting guest operating systems from the host operating system. A type 2 hypervisor is better for individual users who want to run multiple operating systems

A screenshot of a computer

Description automatically generated

The VIM (virtual infrastructure manager) provides a range of features for administering multiple hypervisors across physical servers

A diagram of a computer network

Description automatically generated

Architectures

A diagram of a cloud

Description automatically generated

A screenshot of a web page

Description automatically generated

A diagram of a cloud computing system

Description automatically generated

Diagram of a diagram of a cloud structure

Description automatically generated

A computer screen shot of a diagram

Description automatically generated

A diagram of a computer

Description automatically generated

A diagram of a cloud

Description automatically generated

A diagram of a cloud structure

Description automatically generated

A close-up of a computer

Description automatically generated

Amazon Elastic Container Service (Amazon ECS) is a fully managed container orchestration service provided by Amazon Web Services (AWS). It simplifies the deployment, management, and scaling of containerized applications using Docker containers. Amazon ECS is commonly used in scenarios where you need to manage and scale containerized applications with ease and flexibility.

Amazon Elastic Container Registry (Amazon ECR) is a fully managed Docker container registry service provided by Amazon Web Services (AWS). ECR provides a private container registry where you can store your Docker images securely.

Amazon SageMaker is a fully managed machine learning (ML) service provided by Amazon Web Services (AWS). It is designed to simplify the process of building, training, and deploying machine learning models at scale.

Amazon Athena is an interactive query service provided by Amazon Web Services (AWS) that allows you to analyze and query data stored in Amazon S3 using standard SQL.

AWS Glue is a fully managed extract, transform, and load (ETL) service provided by Amazon Web Services (AWS). automate the process of preparing and loading data from various sources into data lakes, data warehouses, and data stores. Helps building datawarehouses. It support analytics and machine learning workloads. It also queries against s3. Do transformation and also event based or scheduled etl jobs

Data Catalog: AWS Glue provides a central metadata repository known as the AWS Glue Data Catalog. It stores metadata information about the data sources, transformations, and schemas, making it easier to discover and access data.

Data Crawling: Glue can automatically discover and catalog metadata from a variety of data sources, including databases, data lakes, and Amazon S3, using data crawling. This helps in understanding the structure of the data.

Google Cloud Data Fusion is a fully managed, cloud-native data integration service provided by Google Cloud. It is designed to simplify and accelerate the process of building, orchestrating, and managing ETL (Extract, Transform, Load) pipelines for data processing and analysis. Data Fusion is particularly useful for organizations looking to streamline data integration tasks and make data readily available for analytics and business intelligence.

Steps

1. Create instance of fusion
2. Create bucket
3. Upload file
4. Open wrangler for etl
5. Parsing conditions. Ex- file format and first row as header
6. Transformation selection

A screenshot of a computer

Description automatically generated

**Q How to use AWS Glue to migrate data from Google BigQuery to AWS S3**

To use AWS Glue to migrate data from Google BigQuery to AWS S3, you can follow these general steps:

1. Set Up Your AWS Environment:

Ensure you have an AWS account and access to the AWS Glue service.

Create an S3 bucket in AWS where you want to store the data migrated from BigQuery.

1. Prepare Your Data in Google BigQuery:

Make sure your data in BigQuery is organized and structured correctly for the migration.

1. Create an AWS Glue Crawler:

In the AWS Glue console, create a crawler that is responsible for discovering the schema of your data in BigQuery. The crawler will catalog the data, making it easier to work with in AWS Glue.

1. Create an AWS Glue Job:

Create an AWS Glue job to extract data from Google BigQuery and write it to AWS S3. The job will use a script written in PySpark to define the migration process. In the job script, you'll need to set up the connections to BigQuery and S3 and define the transformation logic for your data if needed.

1. Set Up the Data Target in S3:

Specify the S3 bucket and prefix where you want to write the data from BigQuery. Ensure you have the necessary permissions for writing to S3.

1. Run the AWS Glue Job:

Execute the Glue job to extract data from Google BigQuery and write it to S3. The job will use the configuration and logic defined in your script.

1. Monitoring and Verification

In the context of modernizing a monolithic hospital management system, a serverless approach, such as using Amazon Web Services (AWS) Lambda, can offer several advantages. Here's how you can approach this task:

**Amazon Kinesis** is a suite of real-time data streaming and processing services provided by Amazon Web Services (AWS). It is designed to help you collect, process, and analyze streaming data from various sources, making it useful for a wide range of applications.

Decompose the Monolith:

The first step is to break down the monolithic system into smaller, more manageable services. Identify the various functionalities like infrastructure management, day-to-day operations, logistics, security, and transport, and separate them into microservices.

Leverage AWS Lambda:

AWS Lambda is a serverless computing service that allows you to run code without provisioning or managing servers. Each microservice can be implemented as an AWS Lambda function, which can be triggered in response to various events.

Use API Gateway:

AWS API Gateway can be used to create APIs for your microservices. This enables the web and mobile applications to interact with the serverless backend through RESTful or GraphQL APIs.

Data Storage:

For data storage, you can utilize AWS services like Amazon DynamoDB for NoSQL databases, Amazon RDS for relational databases, and Amazon S3 for file storage. Migrate or redesign the data storage to fit the serverless architecture.

Event-Driven Architecture:

Make use of AWS EventBridge to build an event-driven architecture. Events can be used to trigger functions and enable communication between microservices.

Authentication and Authorization:

Implement security measures by using AWS Cognito for user authentication and authorization. IAM (Identity and Access Management) can control access to AWS resources.

Logging and Monitoring:

AWS CloudWatch can be used for logging and monitoring. It provides insights into the performance and health of your serverless applications.

Scalability and Resilience:

Serverless applications inherently offer scalability and resilience. AWS Lambda automatically scales with the number of incoming requests, ensuring optimal performance.

Deployment and CI/CD:

Set up a continuous integration and continuous deployment (CI/CD) pipeline using AWS CodePipeline and AWS CodeBuild to automate the deployment of your serverless microservices.

Vendor-Agnostic Approach:

While AWS provides a robust serverless platform, it's essential to follow best practices for cloud-agnostic design. This ensures that your application can migrate to other cloud providers if needed, offering long-term flexibility.

In terms of cloud vendor preferences, AWS is a strong choice for serverless solutions, but other cloud providers like Microsoft Azure and Google Cloud Platform also offer similar services. The choice of vendor may depend on factors like existing cloud expertise, cost considerations, and specific services available in each cloud platform.

Tearing down the entire monolithic system and adopting a serverless approach offers greater flexibility, scalability, and cost-efficiency, while also enhancing the ability to adapt to the evolving needs of a hospital management system over time.