PROG 8870

CLOUD ARCHITECTURES AND INFRASTRUCTURE AS CODE

ASSIGNMENT 1

PROFF. DARSHVIR SINGH

BY, MARTIN JOHNY

8945124

Cloud Services Selection

XYZ Retail is migrating from an on-premise infrastructure to the cloud to ensure **high availability, scalability, security, and analytics capabilities**.

Compute Resources

AWS Services:

- **Amazon EC2 (Elastic Compute Cloud)**: Provides **scalable and resizable virtual servers**. EC2 is suitable for hosting applications, handling peak traffic, and ensuring availability during high-demand seasons like Christmas.
- **AWS Lambda**: A **serverless computing service** that automatically scales based on incoming requests. It's useful for microservices and event-driven functions.

GCP Services:

- **Google Compute Engine (GCE)**: Provides **customizable virtual machines** similar to EC2, allowing on-demand and preemptible instances for cost optimization.
- **Google Cloud Functions**: A **serverless execution environment** that triggers functions in response to events, reducing infrastructure overhead.

Compute resources must be **scalable** and **cost-effective**. Serverless solutions like AWS Lambda and Cloud Functions reduce operational workloads, while EC2 and GCE provide flexibility for handling high workloads.

2. Storage

AWS Services:

- **Amazon S3 (Simple Storage Service)**: Highly **durable, scalable, and secure** object storage for hosting images, videos, and backups.
- **Amazon EFS (Elastic File System)**: A **fully managed shared file storage** service that scales automatically and supports distributed access.

GCP Services:

- **Google Cloud Storage**: Equivalent to S3, providing **multi-regional object storage** with built-in redundancy and access control policies.

- **Google Filestore**: A **fully managed file storage solution** that supports shared access and integrates with GCP compute services.

Storage solutions must be **highly durable, scalable, and support lifecycle management** to optimize costs. Object storage (S3/Cloud Storage) is great for static content, while file storage (EFS/Filestore) supports real-time workloads.

3. Network Security

AWS Services:

- **AWS WAF (Web Application Firewall)**: Protects applications from **common web threats** such as SQL injection and cross-site scripting.
- **AWS Shield**: Provides **automated DDoS protection** for applications, ensuring uptime during traffic spikes.

GCP Services:

- **Google Cloud Armor**: Offers **real-time application-level security** to block unauthorized traffic and mitigate web-based attacks.
- **Google Cloud DDoS Protection**: Helps prevent **large-scale DDoS attacks**, ensuring uninterrupted service availability.

A retail business like XYZ Retail faces **high traffic volumes and cyber threats**. Using WAF and DDoS protection ensures website security and uptime.

4. Database

AWS Services:

- **Amazon RDS (Relational Database Service)**: Fully managed SQL database with **automated backups, replication, and high availability**. Suitable for storing customer transactions and inventory.
- **Amazon DynamoDB**: A **serverless NoSQL database** designed for high-speed, scalable applications such as product catalogs and user sessions.

GCP Services:

- **Google Cloud SQL**: A **managed relational database service** that supports MySQL, PostgreSQL, and SQL Server. It ensures automatic replication and scaling.

- **Google Firestore (NoSQL)**: A NoSQL database optimized for **real-time, scalable applications** such as recommendation engines and chat applications.

Databases must be **highly available, scalable, and secure**. Relational databases (RDS/Cloud SQL) support structured data, while NoSQL solutions (DynamoDB/Firestore) handle dynamic, high-speed data.

5. Backup & Disaster Recovery

AWS Services:

- **AWS Backup**: Provides **automated backups for EC2, RDS, and S3**, ensuring data protection.
- **Amazon S3 Glacier**: A **low-cost, long-term storage** solution for archival data that's not frequently accessed.

GCP Services:

- **Google Cloud Backup and DR**: Ensures **automated backups and disaster recovery** for databases and virtual machines.
- **Google Archive Storage**: Provides **cost-effective archival storage** for compliance and regulatory requirements.

Backups must be **automated, cost-efficient, and compliant with data retention policies**. Using solutions like AWS Backup and Cloud Backup & DR ensures business continuity.

6. Scalability & Load Balancing

AWS Services:

- **AWS Elastic Load Balancer (ELB)**: Distributes traffic across multiple EC2 instances, ensuring **fault tolerance and improved performance**.
- **AWS Auto Scaling**: Automatically **adjusts compute resources** based on traffic demand, reducing costs.

GCP Services:

- **Google Cloud Load Balancing**: Provides **global traffic distribution**, ensuring minimal latency and failover capabilities.

- **Google Autoscaler**: Dynamically **adjusts VM instances** based on usage metrics.

Load balancing ensures **consistent performance and uptime** during peak seasons. Auto-scaling prevents resource wastage and optimizes costs.

7. Security & Compliance

AWS Services:

- **AWS IAM (Identity and Access Management)**: Controls user access and enforces **least privilege principles**.
- **AWS KMS (Key Management Service)**: Manages **encryption keys** to protect sensitive data.

GCP Services:

- **Google Cloud IAM**: Provides **role-based access control** to enforce security policies.
- **Google Cloud KMS**: Ensures **secure encryption key management** for stored data.

Access control and encryption are **critical for compliance with regulations like GDPR and PCI-DSS**. IAM and KMS services provide enterprise-level security.

8. Monitoring & Logging

AWS Services:

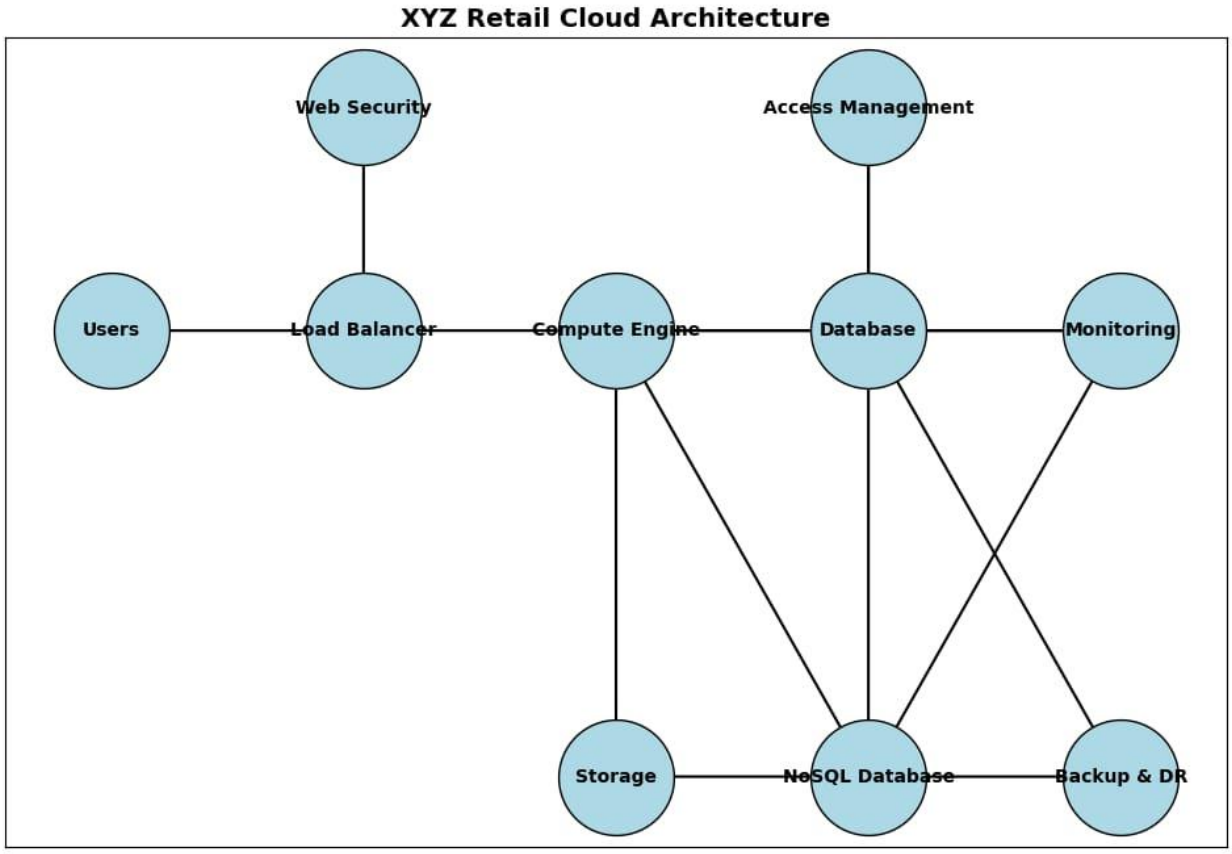
- **Amazon CloudWatch**: Monitors infrastructure, logs, and generates alerts for resource health.
- **AWS CloudTrail**: Tracks API activity and user actions for security auditing.

GCP Services:

- **Google Cloud Operations (formerly Stackdriver)**: Provides **centralized monitoring, logging, and tracing** for applications.
- **Google Cloud Audit Logs**: Captures **security-related events** and administrative actions.

Monitoring ensures **proactive incident response** and system reliability. Audit logs help maintain compliance and security transparency.

Task 2: Cloud Architecture Diagram



This diagram represents the cloud infrastructure for XYZ Retail, designed for high availability, scalability, security, and efficient data management.

1. Users

- Represents customers accessing the XYZ Retail platform.

2. Load Balancer

- Distributes incoming traffic across multiple compute resources to ensure high availability and fault tolerance. Prevents server overload and ensures smooth handling of peak traffic.

3. Compute Engine

- Hosts the application microservices that process user requests.
- Scales dynamically based on traffic demand.

- Can be deployed using AWS EC2 (Amazon Elastic Compute Cloud) or Google Compute Engine (GCE).

4. Storage

- Stores static assets such as images, videos, and logs.
- Uses AWS S3 (Simple Storage Service) or Google Cloud Storage for highly durable and scalable object storage.

5. Database

- Stores structured business-critical data such as transactions, user details, and inventory.
- Uses AWS RDS (Relational Database Service) or Google Cloud SQL to provide a managed SQL database with automatic backup and scaling.

6. NoSQL Database

- Stores fast, scalable, and unstructured data, such as session data, catalogs, and user interactions.
- Uses AWS DynamoDB or Google Firestore, optimized for real-time applications.

7. Web Security

- Protects the system from cyber threats, including DDoS attacks, SQL injections, and cross-site scripting.
- Uses AWS WAF & Shield or Google Cloud Armor to enhance security at the application layer.

8. Access Management

- Ensures secure user authentication and role-based access control.
- Uses AWS IAM (Identity and Access Management) or Google Cloud IAM to enforce security policies and manage permissions.

9. Monitoring

- Provides real-time monitoring, logging, and security auditing.
- Uses AWS CloudWatch & CloudTrail or Google Cloud Operations (formerly Stackdriver) to track system health, API activity, and performance metrics.

10. Backup & DR (Disaster Recovery)

- Ensures data resilience and recovery in case of failure or accidental data loss and Uses AWS Backup & S3 Glacier or Google Cloud Backup & Archive Storage to store backups cost-effectively with lifecycle management.