FINAL PROJECT PAPER

ITC 6480: AWS CLOUD ARCHITECTING, FALL 2022

December 17th, 2022

Professor Carmen Taglienti

Submitted by -

Shreya Karakata

Richa Umesh Rambhia

Keith Medas

Introduction

A medical company which is a start-up software-as-a-service (SaaS) company has started an online medical social networking and diagnosis assistance application for its users. User mainly resides in the United States, Europe, and APAC. The application establishes connections between patients and doctors via online consultations, remote appointments, diagnostics, and prescription services. It also has a feature for text extraction from documents and images. The company wants to use cloud services for its new application once it is launched, as the current environment uses a traditional on-premise server approach for test and development, proves time-consuming and costly when upgrading the infrastructure.

This report highlights using different AWS services to help them control the anticipated growth after launching their new application and host the development, test, and production environments through AWS cloud infrastructure. It also talks about the overall requirements and further details the cloud solution, including a list of all AWS services used, and explains it in depth. Additionally covered are the features and solutions of topics including user authentication, network security, web and application layers, business continuity, and auditing.

Executive Summary

The Medical Startup is a startup that offers software as a service (SaaS); the provider currently employs an antiquated way of leveraging physical resources such as servers and a hosting company to facilitate infrastructure development and testing of the production and development environments. If the resources demand an infrastructure upgrade, this strategy can become costly, time-consuming to manage, and challenging to perform modifications and upgrades to essential applications. The current architecture is divided by three tiers: the web layer as the client tier, the application tier as the server tier, and the database tier. The medical company has created a new application that has yet to be made public; once made public, the company forecasts a tremendous increase in website traffic and application use. The requirement for cloud resources to improve efficiency, scalability, and security will be critical to the product's success and reputation.

The company is committed to constructing cloud infrastructure for clients; the team will employ Amazon Web Services (AWS) to transition from the on-premise model to the cloud environment to meet the customers' needs.

Customer Requirements:

1. Configuring access permissions to conform with AWS best practices.
2. Building networks that conform to AWS best practices while providing all the necessary network services to the application in different environments.
3. Building an architecture that matches the current server hosting company's existing architecture can handle double the number of servers.
4. Securing all medical information, as medical information usually contains highly-sensitive personally identifiable information (PII).
5. Utilizing load balancers for the web and application tiers that must support HTTP, HTTPS, and TCP protocols plans to move their application into AWS.

Overall Requirements

The customer requirements include the following concerning the medical company.

1. High Availability: One requirement used in building the cloud architecture was High Availability. It simply means that the systems are reliable and effective, meaning that these systems continue to operate even if some of their critical components fail. Thus, the major requirement of building this architecture was to make it highly available such that the system or the web application can function even if the components of the system fails. This architecture provides a secured network for a web application that uses the firewall. The new features or components are deployed without causing any problems to the present users.

2. Scalability: Scalability refers to the expansion of every application or component of the infrastructure that can handle the increased load. The main reason for scaling is to increase the performance and efficiency of the application and the infrastructure. Hence, the next requirement of the architecture for the medical company web application was scalability to increase the application's performance each time it is in use such that the increased load on the application can be handled.

3. Security: The security pillar in AWS can protect data, systems and applications, and assets to take advantage of the cloud technologies to improve the security of the architecture. Thus, this is one of the requirement of the architecture in order to provide security to the highly sensitive medical information.

4. Utilization of load balancers: Load balancers are required for the web and application tier; hence, classic load balancers are used in the architecture as it supports the HTTP, HTTPS, and TCP protocols as per the customer requirements.

5. Supporting multiple locations: The application should be available for APAC, US, and Europe users.

Solution

1. **Identify AWS Services:**

   The AWS services used for the project are as follows:

   - **AWS WAF (Web Application Firewall)** - It enables tracking HTTP(S) requests that are routed to the resources of secure web applications.

   - **AWS CloudTrail -** It monitors user behavior and API usage, auditing, security monitoring, and operational troubleshooting.

   - **AWS Cognito -** It enables adding user sign-up and authentication to web applications quickly. Also used to authenticate users through a third-party identity provider and receive temporary security credentials to access the AWS resources. (Amazon Web Services, Inc., 2022)

   - **AWS Route 53** - It enables the connection of user requests to web applications that are running on-premises or on AWS.

   - **AWS CloudFront** - It expedites the delivery of customers' access to web application material, such as HTML, CSS, JS, and picture files.

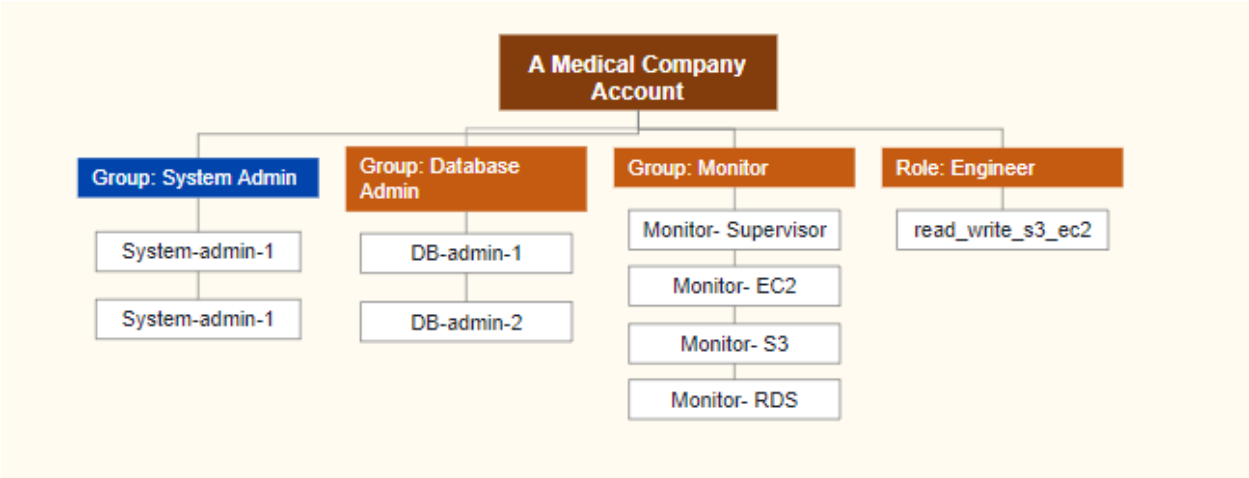
   - **AWS CloudWatch -** It provides real-time monitoring for applications operating on the Amazon infrastructure.

- **AWS IAM (Identity and Access Management)** - It provides to securely manage a user's access to AWS resources and services. It facilitates the creation and management of AWS users and groups, as well as the use of permissions to grant and restrict users' access to AWS resources. (IAM Identities, 2022)
- **AWS S3 (Simple Storage Service) -** It facilitates in storing and retrieving any quantity of data whenever and wherever needed.
- **AWS Classic Load Balancer** - It functions at both the request and connection levels and offers fundamental load balancing across several Amazon EC2 instances.
- **AWS VPC (Virtual Private Cloud)** - It offers a private area within the AWS cloud. Implementing a VPC gives full management of the virtual network, including setting up network gateways and route tables and choosing an IP range. (Virtual private clouds (VPC), 2022)
- **AWS EC2 (Elastic Compute Cloud)** - It gives access to scalable, secure computational resources in the cloud. It is intended to facilitate web-scale cloud computing.
- **AWS NAT Gateway** - It makes it simple for instances inside an Amazon Virtual Private Cloud(VPC) to connect to the Internet.
- **AWS Auto Scaling Group** - It includes a set of EC2 instances used for automated scaling and administration as a logical grouping.
- **AWS Application Load Balancer** - It enables path-based routing and makes routing decisions at the application layer. (Neal, 2022)
- **AWS RDS (Relational Database System)** - It is aimed at making relational database setup, operation, and scalability for application usage.
- **AWS Availability Zones** - It is designed to protect against failures in other Availability Zones.

2. User Authentication:

The users are divided into the following IAM groups:

The system and database administrator group have two users each. The monitor group has four users. The read_write_c3_ec2 role is described to develop, test, and maintain the system mostly to deal with auto-scaling.



Group/Role #	Group/Role Name	Permissions
Group	System Administrator	AWS Console Management Access Programmatic Access
Group	Database Administrator	AWS Console Management Access Programmatic Access
Group	Monitor	AWS Console Management Access
Role	read_write_s3_ec2	AWS Console Management Access

The administrator are given higher access with both Programmatic and AWS Console Management Access. Virtual MFA is enabled for administrator groups that requires them to input a randomly generated password they will get on their mobile devices each time they log in.

The medical company's password policy requires the following:

- Past three passwords cannot be used again
- Passwords must be changed every 90 days

- Should include at least one number
- One special character
- One capital letter
- One lowercase letter.

The requirements will be implemented through an IAM password policy setup.

Requirement	Solution
Should be at least 8 characters and have 1 uppercase, 1 lowercase, 1 special character, and a number.	Checking the following in IAM Password Policy: → Enforce password minimum length : 8 characters → Require at least one uppercase letter → Require at least one lowercase letter → Require at least one number → Require at least one non-alphanumeric character
Change passwords every 90 days and ensure that the previous three passwords can't be re-used.	Checking the following in IAM Password Policy: → Enable password expiration : 90 days → Prevent password reuse : 3
All administrators require programmatic access	Give administrator groups programmatic access through IAM groups
Administrator sign-in to the AWS Management Console requires the use of Virtual MFA.	Enable virtual MFA for administrator groups

3. Network and Security

Network and Security are the critical components in the AWS architecture. The architecture built for the medical company must satisfy the customer requirements where the application is easily accessible to the users. High availability is achieved in the tiers for the architecture, and the application and database tier are in the private subnet; hence no external access to these tiers is given, whereas the web tier is in the public subnet in order to make the application accessible to the users in the available regions. The development and production environment of the company is maintained in separate networks; each of the tiers can receive requests from the respective tiers and load balancers.

VPC	Region	Purpose	Subnets (Each AZs)	AZs	CIDR Range
1	us-east-1	Development	1 public (web tier) 2 private(app and db tier)	use1-az1 use1-az2	10.0.0.0/8
2	us-east-1	Production	1 public (web tier) 2 private(app and db tier)	use1-az1 use1-az2	11.0.0.0./8

The Development and Production VPC is present in region us-east-1(N.Virginia) region having CIDR ranges 10.0.0.0/8 and 11.0.0.0/8, respectively, with two availability zones. Each availability zone has three subnets: web tier subnet is public, and the application and database tier subnets are private.

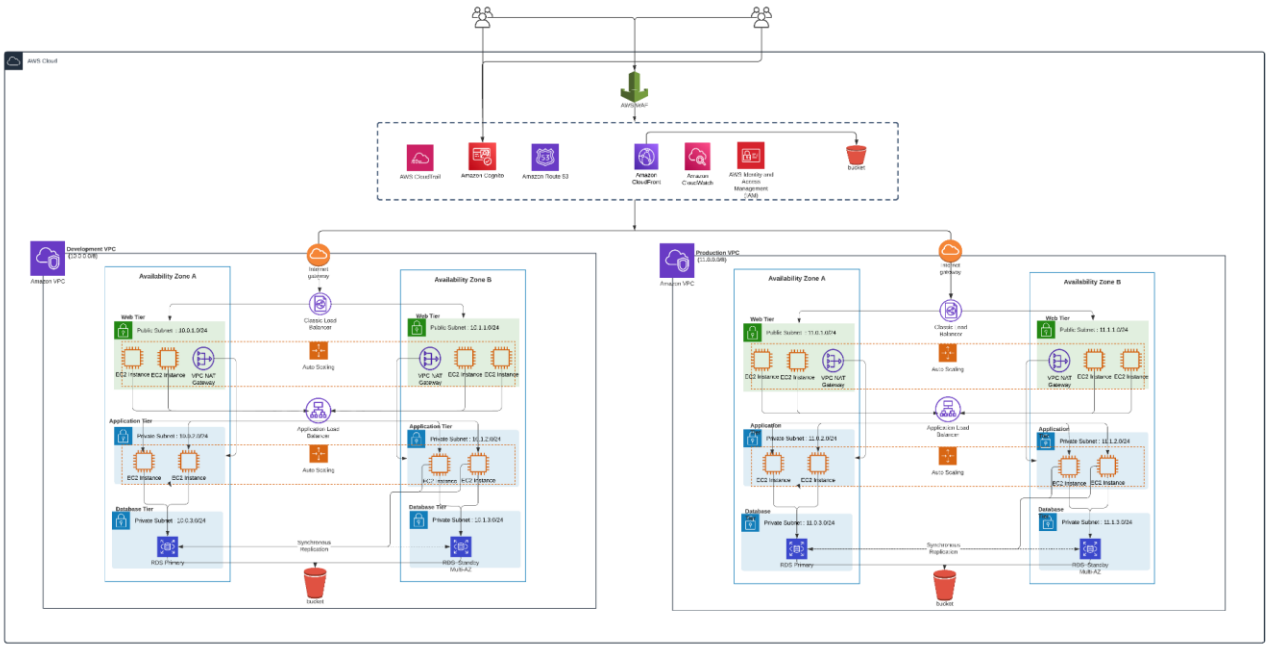
Development VPC

Subnet Name	VPC	Subnet Type (Public/private)	AZ	Subnet Address
dev_web_pub_1	#1	Public	1	10.0.1.0/24
dev_web_pub_2	#1	Public	2	10.1.1.0/24
dev_app_priv_1	#1	Private	1	10.0.2.0/24
dev_app_priv_2	#1	Private	2	10.1.2.0/24
dev_db_priv_1	#1	Private	1	10.0.3.0/24
dev_db_priv_2	#1	Private	2	10.1.3.0/24

Production VPC

Subnet Name	VPC	Subnet Type (Public/private)	AZ	Subnet Address
prod_web_pub_1	#2	Public	1	11.0.1.0/24
prod_web_pub_2	#2	Public	2	11.1.1.0/24
prod_app_priv_1	#2	Private	1	11.0.2.0/24
prod_app_priv_2	#2	Private	2	11.1.2.0/24
prod_db_priv_1	#2	Private	1	11.0.3.0/24
prod_db_priv_2	#2	Private	2	11.1.3.0/24

4. Architecture Diagram (AWS Project : Lucidchart)



5. Web and Application Tier

The current environment that the medical company used was based on Microsoft Windows servers, which hosted their web and application tiers using the traditional methodology. The new architecture built for the company considers the various requirements and AWS services implemented in different availability zones connected to the classic load balancer. In order to be able to access the application such that it could work on port 80, classic load balancers and application load balancers are used for the web and application tiers, as these load balancers support HTTP, HTTPS, and TCP protocols for the inbound port. (Neal, 2022)

Microsoft 2016 does not have IIS installed and therefore cannot use port 80 by default. To overcome this, we use user data for the web and application tier that deploys IIS and can access port 80.

Tier	Tag*	OS	Type	Size	Justification	# of instances	User Data?
Web	Key = Name Value = web-tier	MS Windows 2016	t3 medium	4 GB Memory	For the size and it is required for a high network performance	2	Yes
App	Key = Name Value = app-tier	MS Windows 2016	t3 large	16 GB Memory	For the size and it is required for a high network performance and less interference	2	Yes
DB	Key = Name Value = db-tier	MS Windows with SQL server SE	db.t3 3x large	32 GB Memory, 5TB Storage	For the size and to support all on-demand services	1	No

Load Balancer	Name*	External/ Internal	Subnets	SG Name*	Rule	Source
For Web Tier	web-elb	External	prod_web_pub_1 prod_web_pub_2	web-elb-sg	Inbound port 80 and 443	80 (Internet)
For App Tier	app-elb	Internal	prod_app_priv_1 prod_app_priv_2	app-elb-sg	Inbound port 8080	8080 (Web Tier)

Instance Tier	SG Name*	Rule	Source
Web Tier	web-tier-sg	Inbound port 80 Receives requests from web tier load balancer	web-elb
App Tier	app-tier-sg	Inbound port 80 Receives requests from application tier load balancer	app-elb
Database Tier	db-tier-sg	Inbound port 1433 Receives requests from application tier	App Tier

6. Business Continuity

The database tier contains a master-slave database architecture across availability zones. The backup or slave server acts as a primary server in the event of a breakdown. This ensures continuity in operation, fault tolerance, and resiliency.

Auto-scaling of EC2 instances in the web and application tier across availability zones provides the capacity to enhance traffic and accommodate the rapid expansion of the firm. The auto-scaling helps increase or decrease the instances with respect to business requirements.

Tier	OS	Type	Size	Configuration Name*	Role	Security Group
Web	Microsoft Windows	(t3) medium	4 CPU , 8 GB	WebTier	read_write_s3_ec2	System admin
App	Microsoft Windows	(t3) xlarge	6 CPU , 32 GB	AppTier	read_write_s3_ec2	System admin

Tier	Launch Configuration*	Group Name*	Group Size	VPC	Subnets	ELB	Tags
Web	WebTier	WebTier	Min : 2 Max 4	Production	prod_web_pub_1 prod_web_pub_2	web-elb	Key =Name Value =web-tier
App	AppTier	AppTier	Min : 2 Max 4	Production	prod_app_priv_1 prod_app_priv_2	app-elb	Key =Name Value =app-tier

7. Auditing

Auditing is done with the help of the following services:

- AWS CloudTrail provides operational auditing, governance, compliance, and risk auditing of AWS accounts. This proves as an advantage for constant monitoring of the AWS system and account behaviors, logging history of every AWS account operation history via the AWS Control Console, AWS SDKs, command line programs, and other AWS utilities.
- VPC flow logging captures and stores data related to network traffic. This is in the form of raw data it collects about IP traffic to and from specified network interfaces in Amazon CloudWatch, where it may be downloaded and inspected.
- AWS Config provides an inventory of AWS resources, and a history of configurations. It also alerts of configuration changes to facilitate security

and control. Therefore, it is possible to audit compliance, analyze security, track resource change, and troubleshoot.

- Amazon Cognito helps you to quickly and easily add user sign-up, user sign-in, and access control to your online and mobile apps. (Amazon Web Services, Inc., 2022)
- AWS CloudFront monitors AWS resources such as Amazon Elastic Compute Cloud (EC2) instances, Amazon Elastic Block Store (EBS) volumes, Elastic Load Balancing, and Amazon Relational Database Service (RDS) instances.

Conclusion

The architecture built follows the best practices of AWS and is in accordance with the company's requirements. It is intended to assist the medical firm in transitioning to cloud architecture from its current setup. High availability, scalability, security, utilization of load balancers, and supporting multiple locations are provided with the AWS solution.

IAM groups, users, and roles, along with the permissions, are outlined in the solution. Additionally, it implements other necessary requirements, including virtual MFA and a password policy. The use of AWS VPC, subnets, and usage of multiple zone server deployment also ensures that the application is fault-tolerant and resilient. Security groups are provided to enhance the security for each tier. Auto-Scaling is provided for the EC2 instances in the web and application tier to ensure fault tolerance and better availability. Auditing is performed through services like CloudTrail, VPC logs and AWS Config to track and audit AWS-related operations on the infrastructure.

References

Amazon Web Services. (n.d.-a). *FAQs | Amazon Cognito | Amazon Web Services (AWS)*. Amazon Web Services, Inc. https://aws.amazon.com/cognito/faqs/

Amazon Web Services. (n.d.-b). *How Amazon VPC works - Amazon Virtual Private Cloud*. Docs.aws.amazon.com. https://docs.aws.amazon.com/vpc/latest/userguide/how-it-works.html

Amazon Web Services. (n.d.-c). *Virtual private clouds (VPC) - Amazon Virtual Private Cloud*. Docs.aws.amazon.com. https://docs.aws.amazon.com/vpc/latest/userguide/configure-your-vpc.html

Amazon Web Services. (2019a). *Identities (Users, Groups, and Roles) - AWS Identity and Access Management*. Amazon.com. https://docs.aws.amazon.com/IAM/latest/UserGuide/id.html

Amazon Web Services. (2019b). *Regions, Availability Zones, and Local Zones - Amazon Elastic Compute Cloud*. Amazon.com. https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/using-regions-availability-zones.html

Amazon Web Services. (2020, December 23). *How to upgrade Windows Server 2008R2 using CloudEndure and AWS Managed Services | Microsoft Workloads on AWS*. Aws.amazon.com. https://aws.amazon.com/blogs/modernizing-with-aws/how-to-upgrade-windows-server-2008r2-using-cloudendure-and-aws-managed-services/

Neal. (2022, January 5). *AWS Elastic Load Balancing (AWS ELB)*. Digital Cloud Training. https://digitalcloud.training/aws-elastic-load-balancing-aws-elb/

Parms, J. (2021, March 8). *Port 80 (HTTP) vs. Port 443 (HTTPS): Everything You Need to Know*. SSL2BUY Wiki - Get Solution for SSL Certificate Queries. https://www.ssl2buy.com/wiki/port-80-http-vs-port-443-https

Port 8080 (tcp/udp). (n.d.). SpeedGuide. https://www.speedguide.net/port.php?port=8080