# Design of an AWS Architecture for Library e-Portal

Due on August 30, 2021

*Basanta Joshi, Ph.D.*

*Dipak Poudel*


## Ashlesh Pandey

## PUL074BEX007

# Abstract

With the growing field of Information Technology (IT), the entire world is shifting from on-premise services to cloud computing. The benefits of cloud computing over traditional on-premise methodologies is unmatched. As a part of the elective subject, Enterprise Computing, under the Department of Electronics and Communication Engineering, Pulchowk Campus, Institute of Engineering, this report presents a library e-Portal architecture on the cloud using Amazon Web Services (AWS).

# Contents

**Abstract** i

**List of Figures** iv

**List of Abbreviations** v

**1 Problem Statement** 1

**2 Requirement of the Architecture** 1

**3 Background Theory** 3

3.1 Global Infrastructures . . . 3

3.1.1 AWS Cloud . . . 3

3.1.2 AWS Regions . . . 3

3.1.3 Availability Zones . . . 3

3.2 Server related Services . . . 3

3.2.1 Amazon Elastic Compute Cloud (EC2) . . . 3

3.2.2 Elastic Load Balancing . . . 4

3.2.3 Autoscaling . . . 4

3.3 Database and Storage related Services . . . 4

3.3.1 Amazon Aurora . . . 4

3.3.2 Amazon Memcached . . . 5

3.3.3 Amazon Simple Storage Service (S3) . . . 5

3.3.4 Amazon Elastic File System (EFS) . . . 5

3.4 Networking related Services . . . 5

3.4.1 Virtual Private Cloud (VPC) . . . 5

3.4.2 Subnet . . . . . 5
3.4.3 Amazon Route 53 . . . . . 6
3.4.4 Amazon Cloudfront . . . . . 6
3.4.5 Internet Gateway . . . . . 6
3.4.6 Network Address Translation (NAT) Gateway . . . . . 6

**4 Library e-Portal Architecture** **7**

# List of Figures


1 Single server versus divided components . . . . . 2
2 Basic setup with a single EC2 instance . . . . . 7
3 Faster content delivery . . . . . 8
4 Separate database sevice . . . . . 9
5 Redundancy across availability zones . . . . . 10
6 Resource utilization as per demand . . . . . 11
7 Placing instances in private subnets . . . . . 12
8 Caching for faster response time . . . . . 13
9 Shared file system . . . . . 14
10 Final architecture for library e-Portal . . . . . 15

# List of Abbreviations

**AMI** Amazon Machine Image.

**AWS** Amazon Web Services.

**CDN** Content Delivery Network.

**DNS** Domain Name System.

**EC2** Elastic Compute Cloud.

**EFS** Elastic File System.

**ELB** Elastic Load Balancer.

**IaaS** Infrastructure as a Service.

**IP** Internet Protocol.

**IT** Information Technology.

**NAT** Network Address Translation.

**PaaS** Platform as a Service.

**S3** Simple Storage Service.

**SaaS** Software as a Service.

**VPC** Virtual Private Cloud.

# 1 Problem Statement

With the advent and growth of cloud computing services, the on-premise servers and data centers are gradually becoming obsolete. There is no doubt that with growing demand of resources, and the necessity of widespread availability, most services have to switch to the cloud to sustain a positive impression on the users. One such field is the library facility that almost every campus must ensure for a well-rounded education system. Library facilities include book-lending, book-renewal, manual/report/project database and tracking of user's book-transactions. To ensure easy and faster access along with durable and secure results, such library systems are shifting to cloud, which is why this report suggests one such architecture for a library e-Portal on AWS cloud.

# 2 Requirement of the Architecture

For a functional library e-Portal, the overall architecture must be able to perform the following key functions,

1. Host static content.
2. Host dynamic content.
3. Store database content.
4. Store file system content.

If a basic library e-Portal was to be designed with a single instance, the EC2 host would perform all the tasks. However, for a properly implemented scalable, secure, durable and fast architecture, the functions are divided among different components provided as service by AWS.

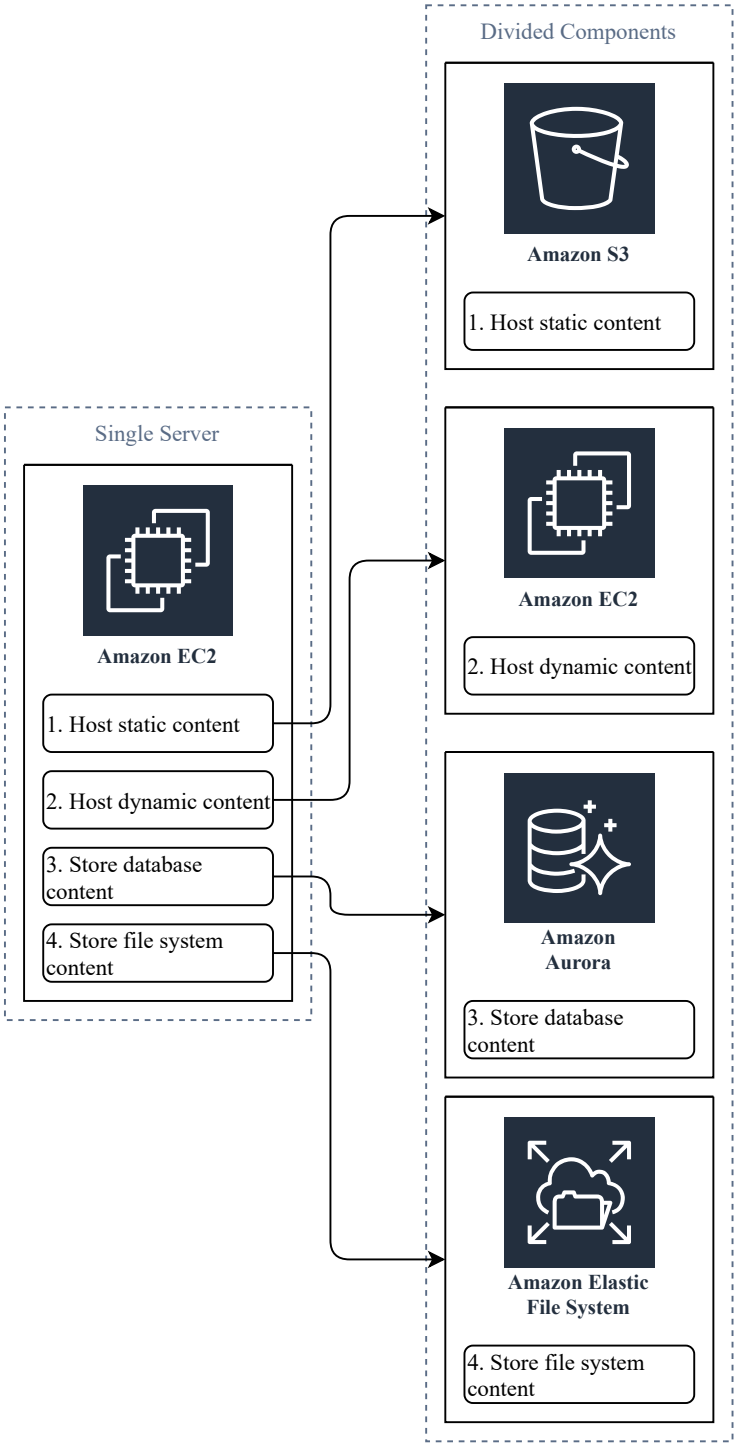


Figure 1: Single server versus divided components

# 3 Background Theory

## 3.1 Global Infrastructures

### 3.1.1 AWS Cloud

AWS cloud is a group of services provided by Amazon that include IaaS, PaaS and SaaS along with different necessary components for an entire cloud architecture. In order to deliver highly reliable, scalable cloud services over a secure connection, Amazon has segregated it's global infrastructures with 81 availability zones in 25 actual global regions around the globe.

### 3.1.2 AWS Regions

These are physical geographical locations where data replication across regions depends on the clients need. Currently there are 25 AWS regions that themselves house one or more availability zones. Each region is completely independent of one another and pricing vary in these regions.

### 3.1.3 Availability Zones

These are isolated AWS infrastructures inside an AWS region that are connected via high bandwidth private networks. Currently there are 81 availability zones that themselves house one or more data centers. The availability zones ensure isolation from natural calamities. 21 new availability zones in 7 AWS regions have been announced recently.

## 3.2 Server related Services

### 3.2.1 Amazon Elastic Compute Cloud (EC2)

Amazon Elastic Compute Cloud (EC2) is a service available on the Amazon console dashboard that allows clients to create virtual machines, generally known as EC2 instances. In terms of

cloud computing types, an EC2 can be regarded as an IaaS since the client can choose the AMI, and capacity of servers. The said EC2 instances allow clients to host virtual machines on which they can deploy applications in somewhat a similar fashion as an on-premise environment. Few use-cases where EC2 instances are used are as application, web, game, media, mail, database servers. EC2 instances can be launched in the different possible availability zones.

### 3.2.2 Elastic Load Balancing

Elastic Load Balancing is a service provided by AWS that is used to distribute the incoming traffic across different target that maybe Amazon EC2, IP addresses, applications or Lambda functions. There are four types of Elastic Load Balancer (ELB), viz. Application Load Balancer, Network Load Balancer, Classic Load Balancer and Gateway Load Balancer.

### 3.2.3 Autoscaling

Auto scaling is a service provided by AWS that allows users to set scaling configurations based on which various resources are scaled based on the demand requirements of the system. It is used to dynamically allocate computational resources with changing demands and helps in cost reduction, proper utilization of resources and scalability.

## 3.3 Database and Storage related Services

### 3.3.1 Amazon Aurora

AWS offers Amazon Aurora as a simple and cost-effective enterprise level relational database service that is also compatible with MySQL and PostgreSQL.

### 3.3.2 Amazon Memcached

Amazon Memcached is an open-source high-performance, in-memory data store used as cache store for applications on AWS cloud.

### 3.3.3 Amazon Simple Storage Service (S3)

Amazon Simple Storage Service (S3) is the object storage service provided by AWS. Amazon promotes S3 as a highly scalable, available, durable (99.999999999%) and secure object storage service. Objects are stored in S3 buckets that are globally unique and located in a region.

### 3.3.4 Amazon Elastic File System (EFS)

Amazon Elastic File System (EFS) is a file storage system for the applications that run on AWS. It is used to avoid performance bottlenecks as it provides flexible storage capacity to multiple connected instances.

## 3.4 Networking related Services

### 3.4.1 Virtual Private Cloud (VPC)

Virtual Private Cloud (VPC) is a service provided by AWS that allows users to create private, isolated virtual networks on the public environment essentially providing much needed isolation to different resources on AWS.

### 3.4.2 Subnet

Subnet, as the name suggests is a smaller (sub) piece of a larger network (net). Logical division of IP addresses create the required ambiguity among different components of a larger system.

Private subnets are those that don't have public IP assigned, hence aren't directly accessible over the internet. Similarly, public subnets can be accessed over the internet.

### 3.4.3 Amazon Route 53

Amazon Route 53 is a Domain Name System (DNS) service provided by AWS that provides domain registration, health checking in addition to DNS routing.

### 3.4.4 Amazon Cloudfront

Amazon Cloudfront is a Content Delivery Network (CDN) service provided by AWS that provides a globally distributed content delivery service using proxy servers to cache contents at various edge locations.

### 3.4.5 Internet Gateway

An internet gateway, as the name suggests is the gateway between a VPC and the internet. It provides a target in a VPC route table for the traffic and also performs Network Address Translation (NAT) for the instances that have been assigned public IP addresses.

### 3.4.6 Network Address Translation (NAT) Gateway

For instances that are placed in a private subnet, a Network Address Translation (NAT) gateway is required to connect to the internet. A NAT gateway provides instances in private subnets access to public internet outbound traffic while blocking incoming connections. It is essential for connecting instances to the internet during software updates or patch installations.

# 4 Library e-Portal Architecture

The architecture design starting from the basic setup to the final design is included as,

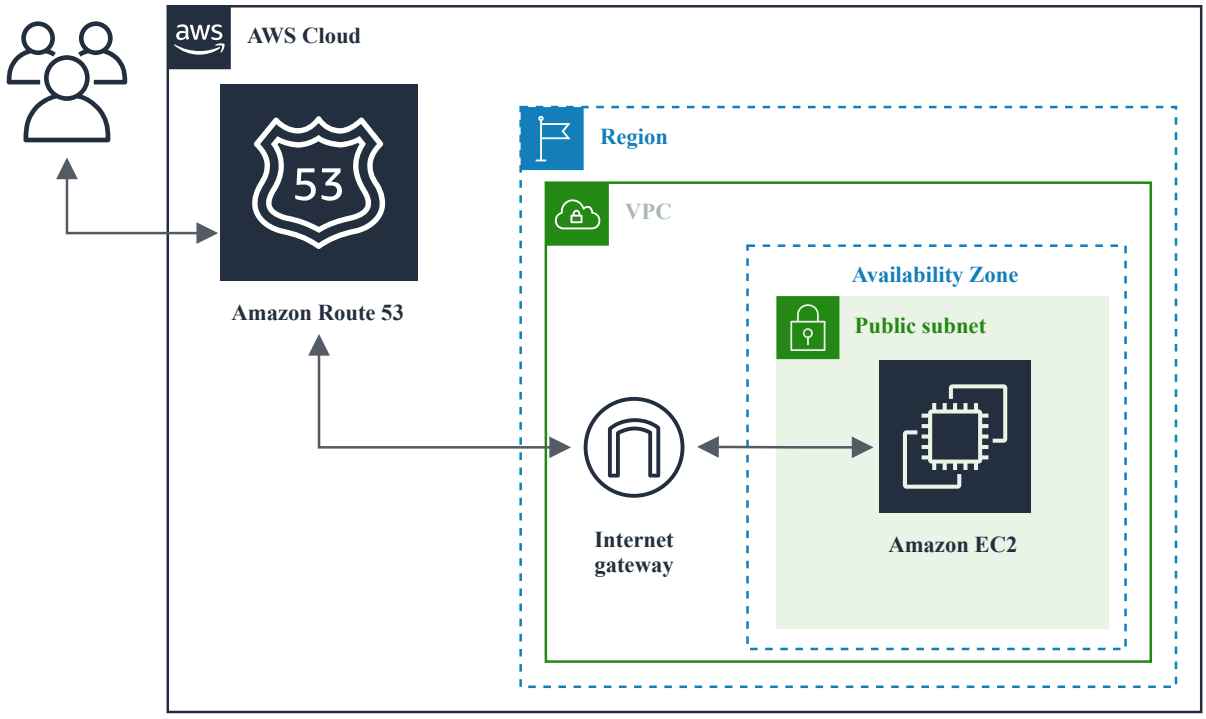


Figure 2: Basic setup with a single EC2 instance

A simple library e-Portal can be run using a single EC2 instance connected to an internet gateway and a working DNS system provided by Amazon Route53. All the work-load would be handled by the single instance and a larger system would be tedious.

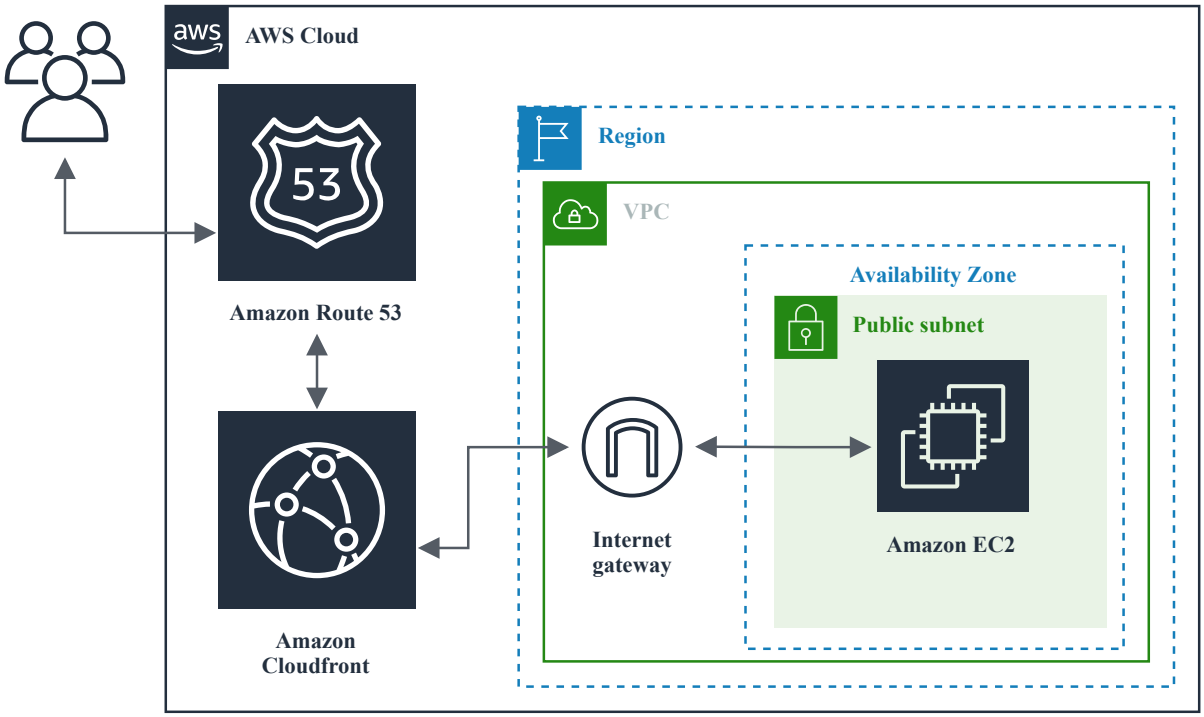


Figure 3: Faster content delivery

As discussed earlier, Amazon Cloudfront is essential if the system requires best content delivery to users in different locations of the world. This not only boosts the access speed but also saves compute capacity from the instance used.

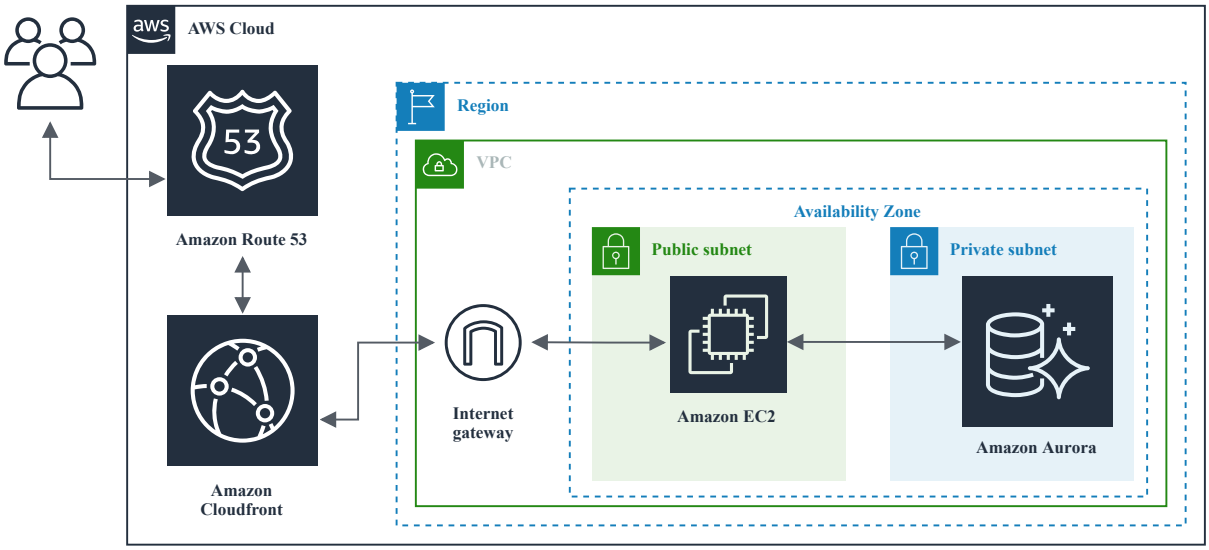


Figure 4: Separate database sevice

Since the system requires a fully-managed database, Amazon Aurora is used in a private subnet to provide additional security to the database. This allows better database management for the e-Portal and also secures the database since it will contain critical data of users.

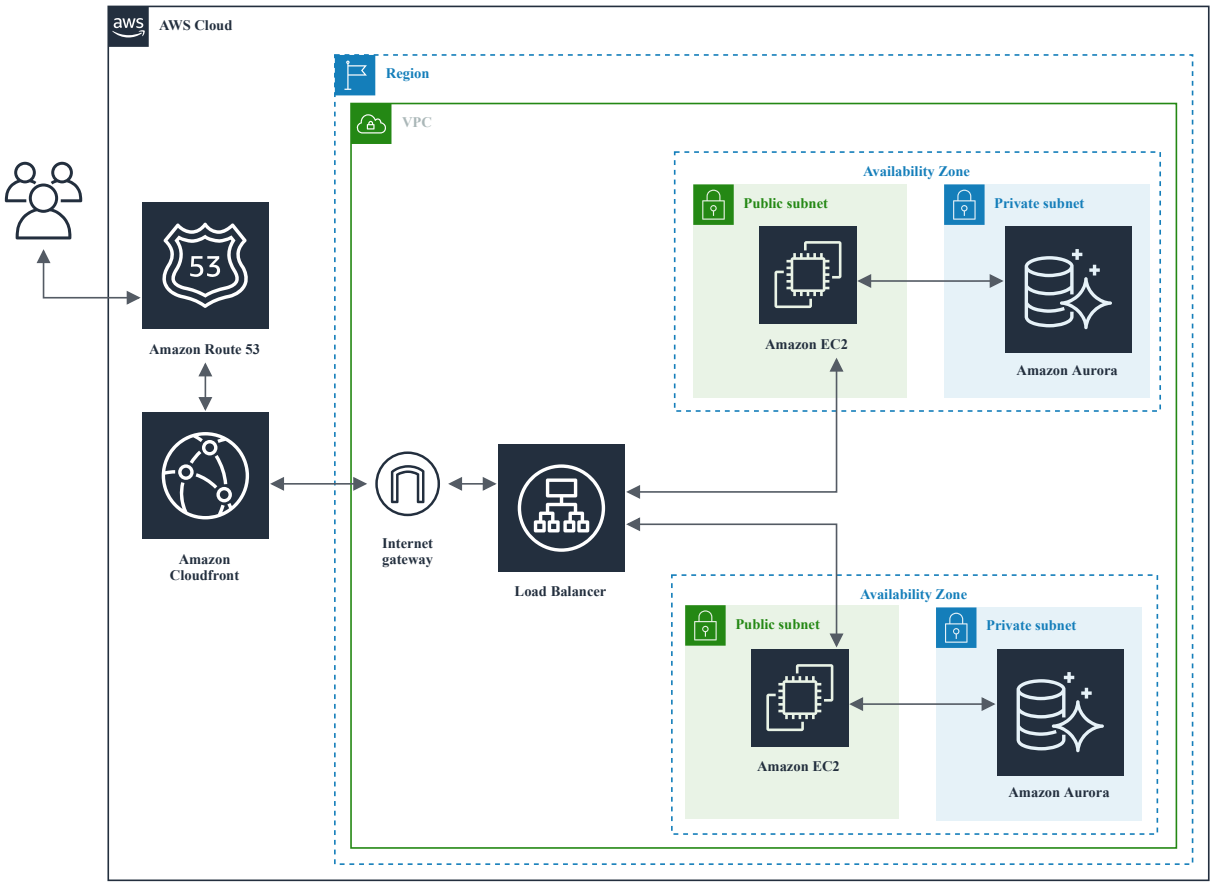


Figure 5: Redundancy across availability zones

To avoid downtime and assure database security, a replication can be added in another availability zone. An ELB is required to distribute traffic to the different instances. This way the e-Portal will have better durability since a rare failure of one instance or a database wouldn't totally damage the system.

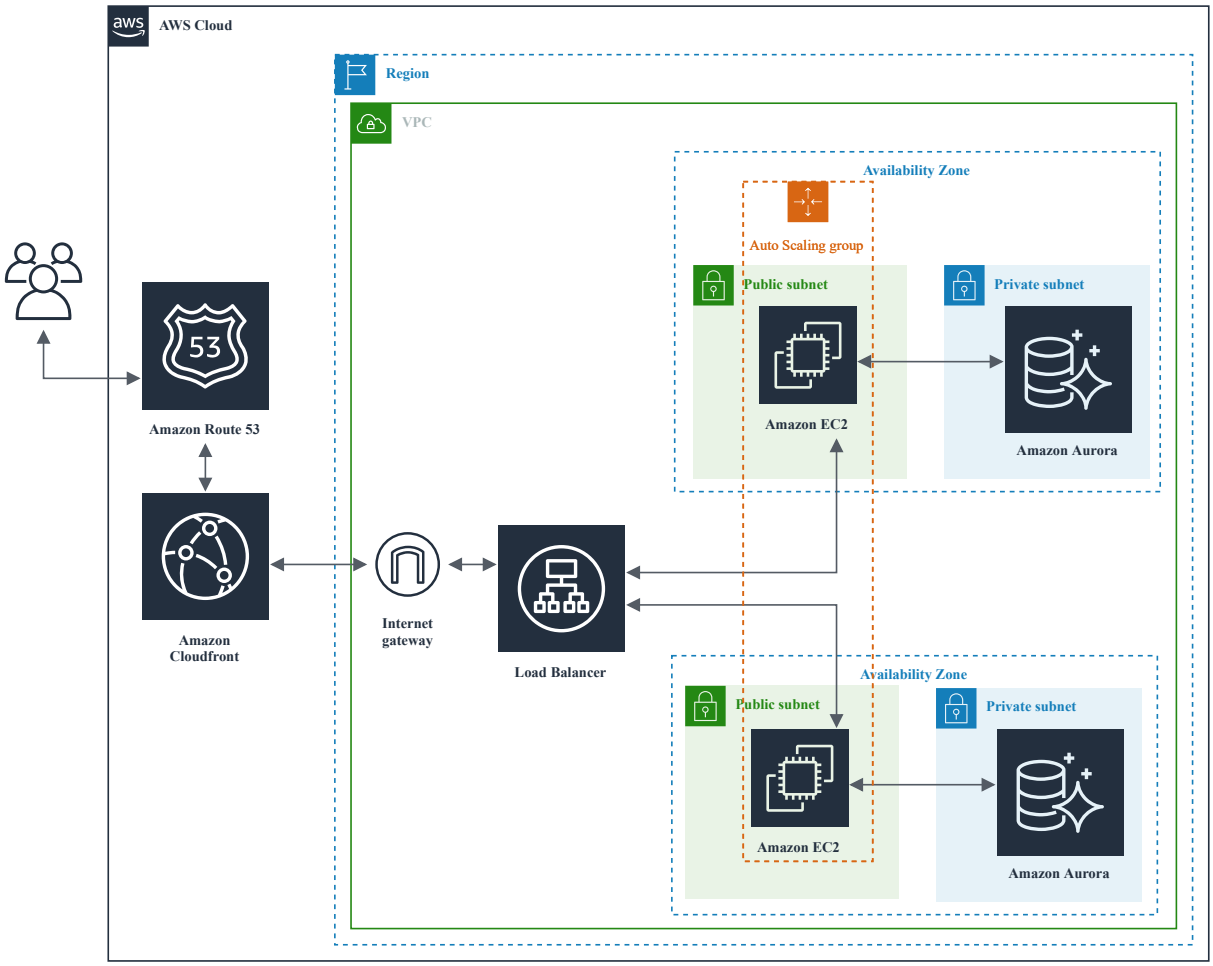


Figure 6: Resource utilization as per demand

As discussed earlier, auto-scaling enables a system to scale according to the requirements hence saving cost since the servers will be scaled down during holidays when the users won't be much active and also ensure that the service don't get overwhelmed during peak traffic such as examination time.

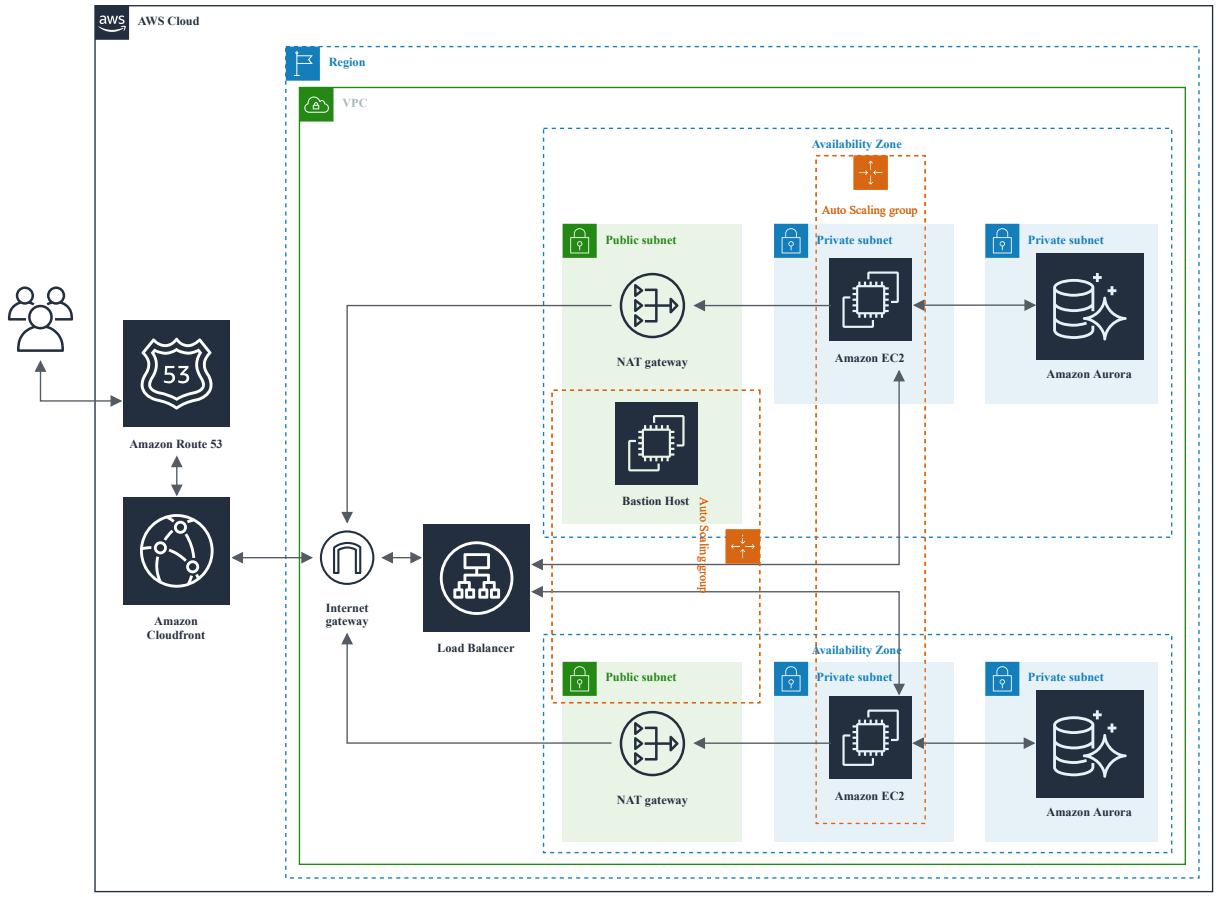


Figure 7: Placing instances in private subnets

Similar to the database servers, the EC2 instances placed in private subnets provide additional security from malicious contents. However, the instance would not have connection to the internet hence disrupting the software updates and patch installations, which is why NAT gateways are used. AWS suggests the use of a Bastion Host, which is a way to securely access EC2 instances located in a private subnet. Placing the Bastion Host inside an auto-scaling group ensures the proper scaling if traffic were to increase.

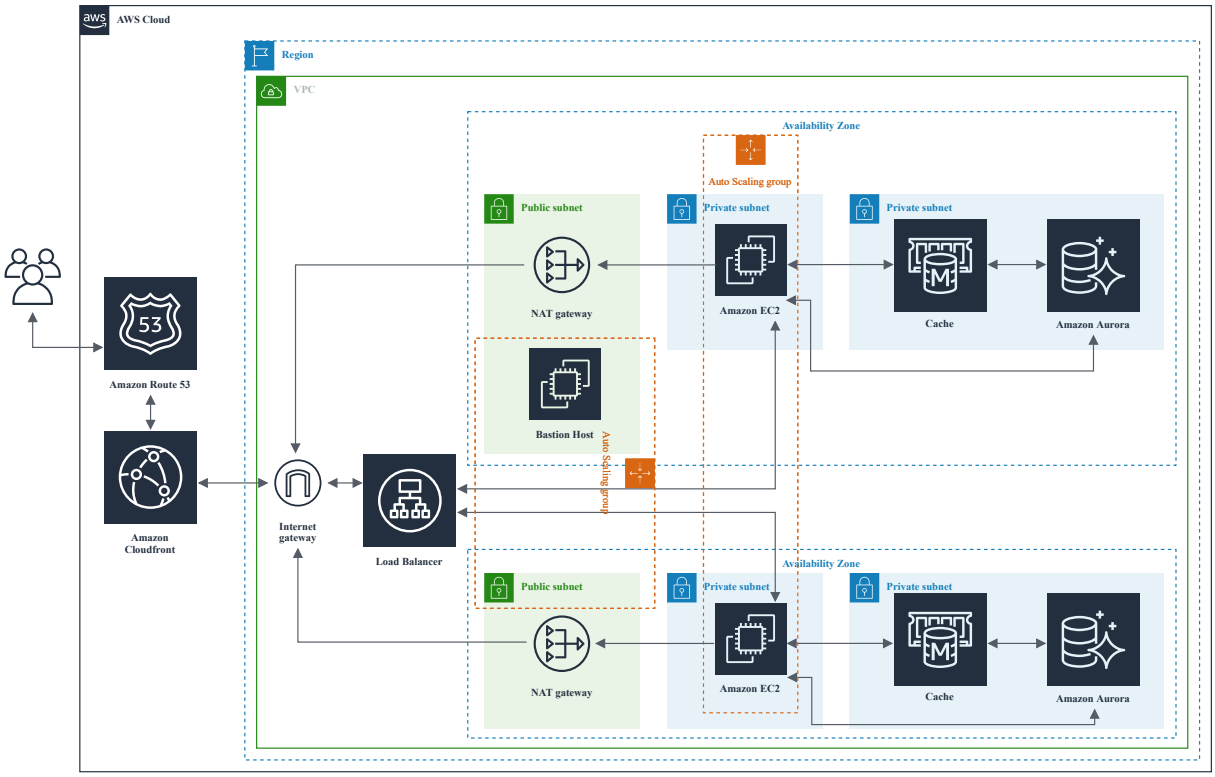


Figure 8: Caching for faster response time

Use of a cache memory system reduces the database burden and also provides faster responses making it quite useful for performance boosts. This way the database requests would have a faster response time and frequently searched queries would be cached to reduce load.

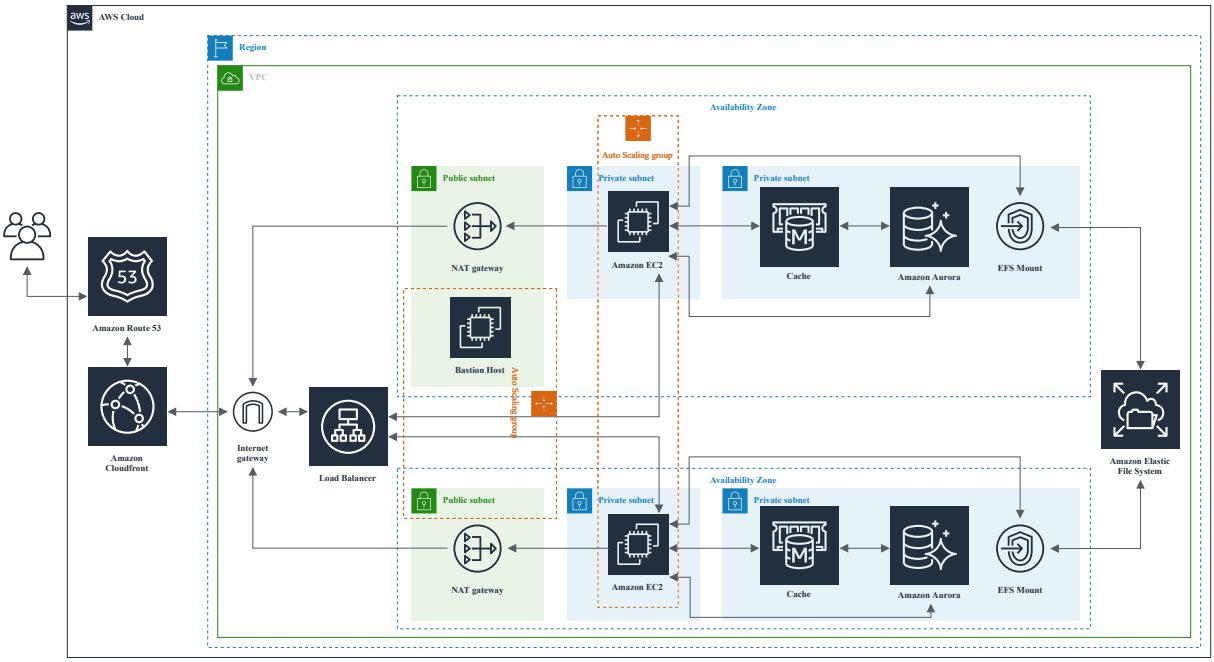


Figure 9: Shared file system

As discussed earlier, Amazon EFS can be used to provide a shared file system that the different EC2 instances can use for a shared disk space.

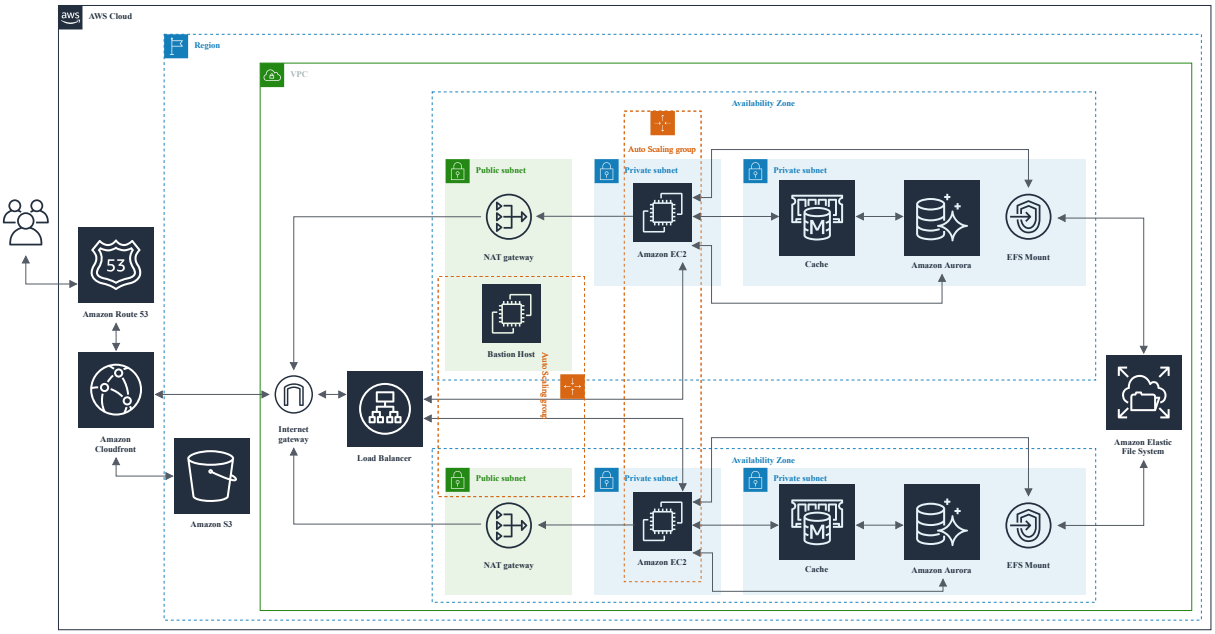


Figure 10: Final architecture for library e-Portal

Lastly the static contents for the portal can be placed in a S3 bucket. The static objects placed in the bucket can be easily access providing an essential storage service for some reports or books, images, lecture videos and such for the portal.