APPROACH FOR HIGH AVAILABILITY

- 1. We will spread our application across multiple Availability zones, to ensure that our application remains available even if one Availability zone goes down.
- 2. We will create multiple worker nodes in our cluster, to ensure that our application remains available even if one node goes down.
- 3. We will apply health check on our pods such as liveness and readiness probes to detect and automatically recover from failures.
- 4. We will use a managed Kubernetes service, such as Amazon EKS(Elastic Kubernetes Service) which provides us many features like automatic node replacement and cluster scaling, which can help improve the deployment and availability of our application.
- 4. We will use a load balancer to distribute our traffic across multiple worker nodes and ensure that our application remains available even if one worker goes down.