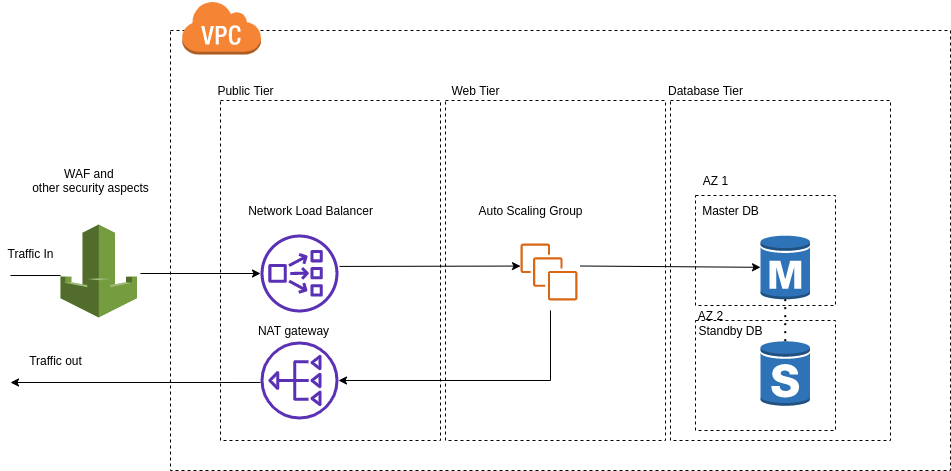
**Diagram**:



**Traffic flow:**

When web page is accessed, traffic is passed from the load balancer port 80 to other ports used by the auto scaling group to access the web page. If user is required to save or access data, (e.g. logging into system), the database is accessed for data retrieval.

**Justifications:**

**For high availability DB setup**

* DB should be using an Active and standby DB so that if the Master DB fails, traffic can automatically failover to standby DB
* Active and Standby DB are located in different AZ.
* AWS Aurora is not used as it is more costly if Database traffic is less

**For high availability Web**

* Scaling group should auto recover if instance is found to be unhealthy (kill and create new)
* Scaling group should scale out once CPU of all instances in scaling group reaches 60% for 5mins or more
* Scaling group should scale in once CPU of all instances in scaling group is less than 60% for 5mins or more
* Instances created should be with a minimum requirement of 4 c5.xlarge instances at all times

**Network Load Balancer**

* Selected over Application load balancer and options to do checking and routing for layer 4 traffic and layer 7 traffic
* AWS managed, least administrative effort

**NAT gateway**

* NAT gateway to route traffic out of web tier as autoscaling group instances are of private network and require a NAT gateway to return traffic to users.

**WAF**

* WAF and other security aspects to filter bad traffic and prevent security attacks
* Stringent security rules to be added for WAF
* Can also consider to add https certification on load balancer.

**Assumptions**

* Assume that for auto scaling, instance creation takes around 1 min or less
* Assume that for DB servers, traffic is not intensive as there is no requirement for loadbalancer on db tier and also no read replicas are required
* Assume that calculations for amount of traffic is completed and would require minimum of 4 c5.xlarge instances running at all times
* Assume that cloudfront is not required as majority of projected traffic is from singapore, and hosted area is also in singapore
* Assume that security