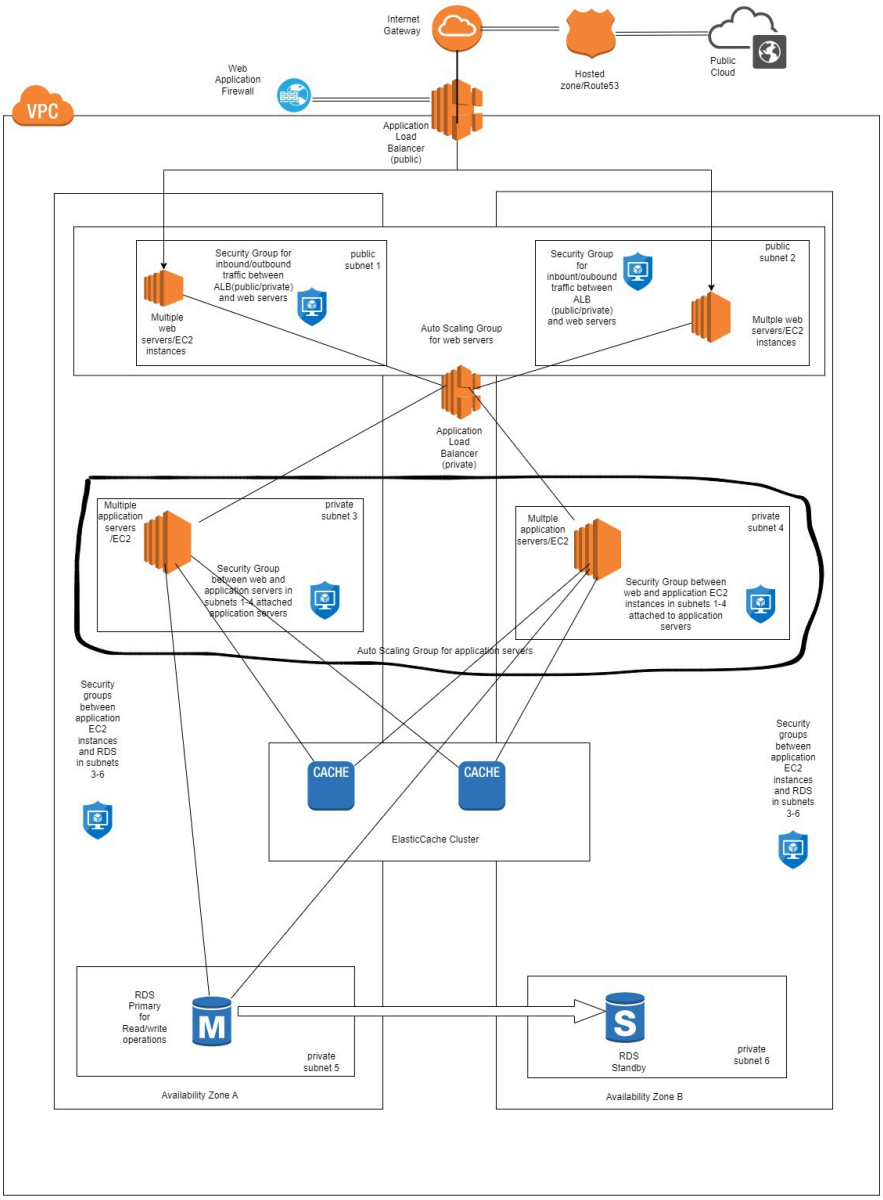
Stefan Rosenberg



1 Use existing or create new VPC with dedicated CIDR block. Create the subnets (CIDR's) for the web and db instances. Configure each subnet with the associated Availability Zone (AZ) - subnets 1,3,5 to AZa, and subnet 2,4,6 to AZb.

2 Security Groups between web and application servers in subnets 1-4. Security groups between application EC2 instances and RDS in subnets 3-6

3 Create the Launch template and include the AMI's for the web application server, along with InstanceType, min-size 1, max-size 2 and deesired-capacity of 2 EC2 instances. Specify the Amazon Machine Image (AMI), instance type, key pair, security groups to allow access to RDS in both AZ's, add custom tags.

4 Create the Auto Scaling Group (ASG), referencing the launch template created in step 3, the VPC and appropriate subnets from Step 1. Attach ASG with the ELB created in step2. Registering the Amazon EC2 instances with a load balancer can be configured here, so I chose the ALB that will be created in Step 6. Turn on Elastic Load Balancing health checks. Under Additional settings, Monitoring, choose whether to enable CloudWatch group metrics collection. For Enable default instance warmup, select this option and choose the warm-up time for your application. If you are creating an Auto Scaling group that has a scaling policy, the default instance warmup feature improves the Amazon CloudWatch metrics used for dynamic scaling

5 Created Amazon RDS master node in AZ a with a Standby node in AZb for failover. All RDS logs, backups, and snapshots are encrypted using AWS KMS key to encrypt these resources.

6 Create an internet-facing Application Load Balancer (ALB). With Application Load Balancers, cross-zone load balancing is always enabled at the load balancer level. When cross-zone load balancing is enabled, each load balancer node distributes traffic across the registered targets in all enabled Availability Zones. Default routing is done via round robin.

7 On the Application Load Balancer, create a target group, which is used to route requests to one or more registered targets (i.e. EC2 instances). When the listener rule is created on the ALB, a target group and conditions are specified. When a rule condition is met, traffic is forwarded to the relevant target group. Again, default routing is done via round robin.

8 I chose an Application Load Balancer over a Network Load Balancer because the ALB examines the application layer protocol data from the request header. Though this takes more time than network load balancing, it allows the balancer to make a more informed decision of where to direct the request. Other advantages: ALB's can listen in on HTTP and HTTPS requests, target groups can used by the ALB to route traffic to different urls (if needed), integration with AWS Certificate Manager to run TLS (HTTPS) connections, make routing decisions based on HTTPS headers, and HTTP routing rules can be based on host header value or URL path pattern.

9 I created two separate Application load balancers - internet-facing and internal. The web servers/EC2 instances attached to the internet-facing application load balancer have public IP addresses, and therefore reside in a public subnet.
The application servers/EC2 intances of the internal application load balancer have only private IP addresses, and therefore reside in private subnets.

10 I created an Internet Gateway and attached it to the VPC. I then created a Route Table and added the IGW as a target to the route table. Therefore, traffic will be routed from the Internet Gateway to the ALB to the web servers in subnets 1 & 2.

11 The following will be created in Cloudwatch by default:

EC2: By default, Amazon EC2 sends metric data to CloudWatch in 5-minute periods

RDS: By default, Amazon RDS automatically sends metric data to CloudWatch in 1-minute periods.

ALB: Elastic Load Balancing publishes data points to Amazon CloudWatch for the application load balancers(internet and internal) and associated targets.

The Flow logs will need to be enabled:
VPC: I'll create a flow log for a VPC, a subnet, or a network interface. For theflow log for a subnet or VPC, each network interface in that subnet or VPC will be monitored.

12 I created a Web Application Firewall (WAF) with an ACL that references a rule group with two rule actions:

SizeRestrictions_QUERYSTRING: Inspects for URI query strings that are over 2,048 bytes

NoUserAgent_HEADER rule: Inspects for requests that are missing the HTTP User-Agent header

13 Created an ElasticCache Cluster in AZ's a & b

14 Application servers in both AZ's will reach the ElasticCache first for cached SQL queries; if not available, the app servers will reach the Primary RDS db in AZ A. If RDS Primary not available, app servers will connect to RDS standby in AZ b