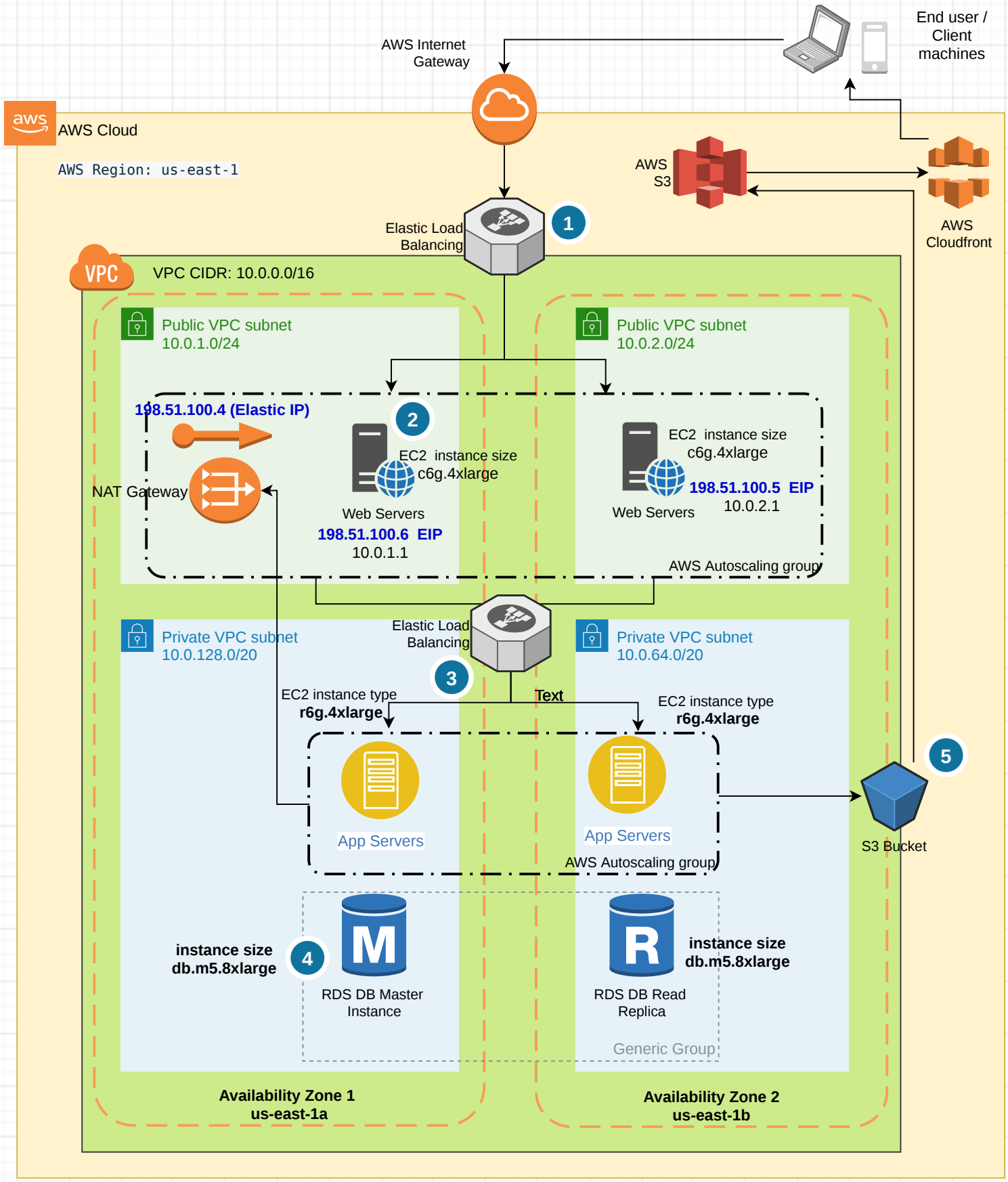
AWS Architecture - New social media app development project

This architecture is intended to support 50,000 single-region users with consideration to be a cost effective solution. Expected workloads are a natural fit for AWS due to unexpected traffic patterns & request rates. We leverage the flexibility of AWS to start small and power up according to the increase on demand. Decision about AWS instance type focus on current generation only to ensure better performance for the price.

For the app server instance the choice of instance type is for **r6g.4xlarge** memory-optimized instances types which provide the lowest cost per GB of RAM. Regarding the high performance web server it will benefit from high compute power and then the choice is the **C6g** compute-optimized instance types which provide the lowest cost for CPU performance. Web servers are not expected to require consistently high levels of CPU but will benefit significantly from having full access to very fast CPUs when they need them.

Author: David de Lucas Salmerón



1. Route users to backend using **elastic load balancing**, that scales automatically for incoming traffic.
2. Web servers on AWS EC2 instances in an **autoscaling group** which spans two AZs. According to the usage estimates c6g.4xlarge instance types with 32 GB Memory and 16 vCPUs
3. Web and App tier separated, use a Elastic Load Balancer in private subnet to allocate incoming traffic to the App layer.
4. AWS RDS with read replica in a different AZ.
5. AWS S3 to store static assets using CloudFront as a globally distributed cache of content