Containers on AWS: ECS, Fargate, ECR, and EKS

Uday Manchanda

August 23, 2021

1 ECS

- elastic container service, allows you to launch docker containers on AWS
- must provision and maintain the infrastructure (EC2 instances)
- AWS takes care of starting and stopping containers
- integrates with ALB to expose application to the web
- Launch Types:
  - EC2 launch type for ECS:
    - ECS cluster in a single VPC
    - Create ASG to create EC2 instances aka container instances.
    - Run ECS agent on all the instances, register those instances to the ECS cluster, and start launching some tasks.
    - Each EC2 instance can run one or more tasks
  - Fargate.
    - No EC2 instances to provision. Serverless offering
    - AWS just runs the containers for you based on CPU/RAM needs.
    - Fargate will launch a task
    - To access the task, we will have an ENI (elastic network interface) launched within the VPC to bind this task to a network IP.
    - More tasks = more ENIs. Each ENI is a distinct IP address

1.1 IAM Roles for ECS Tasks

- ECS tasks might need to interact with other AWS services - will need IAM roles for them
- EC2 instance profile - used by ECS agent, makes API calls to ECS service, send container logs
- ECS task role:
  - allows each task to have a specific role
  - use different roles for the different ECS services you run
  - task role is defined in the task definition
  - task roles allow, for example, access to an s3 bucket

1.2 Data Volumes - EFS File Systems

- if you want to share data: create EFS file system and mount it directly on the ECS tasks.
- works for both EC2 and fargate tasks.

1.3 ECS Scaling

- Service CPU usage which is a CW metric, triggers CW alarm, autoscale
- SQS queue, polls for messages, if queue length exceeds limit (CW metric), triggers an alarm, autoscale

1.4 ECS Rolling Updates

- we can control how many tasks can be started and stopped, and in which order
- EX: min 50 percent, max 100 percent, starting number of tasks: 4

2 ECS Services and Tasks

- it is common to spin up an ECS cluster with a running service that connects to an ALB
- when launching a container, it will get a random port assigned to it
- ALB has dynamic port mapping feature - supports finding the right port on your EC2 instances
- You can run multiple instances of the same container onto the same EC2 instance and have them exposed thru the same ALB.
- You must allow on the EC2 instance SG any port from the ALB SG

2.1 Load Balancing for Fargate

- Fargate creates ENI for every task run. Each ENI has a unique IP but port remains the same.
- Must allow on the ENI SG the task port from the ALB SG

2.2 ECS tasks invoked by Event Bridge

- allows you to invoke an ECS task, like by an event bridge or cloud watch event
- EX: Fargate cluster, whenever a user uploads an object into S3 bucket, run a fargate task to process that object and insert some metadata into dynamoDB
- Create AWS Event Bridge event from S3 bucket. The event has a rule as a target which is to run an ECS task. Make sure the task has the proper task role.

3 ECR

- Elastic Container Registry
- Store, manage, and deploy containers on AWS, pay for what you use.

4 EKS

- Elastic kubernetes service
- launch managed kubernetes clusters on AWS
- kubernetes - deployment, scaling, and management of containerized applications
- alternate to ECS. best if you already use kubernetes on-prem and want to migrate to AWS
- EKS pods are like ECS tasks. They run on EKS nodes. The nodes can be managed by an ASG.