**What is Elastic Container Service?**

***Elastic Container Service (ECS)*** is a container management service that makes it easy to run, stop, and manage Docker containers on a cluster. You can host your tasks on a cluster of EC2 instances or on a serverless infrastructure using the Fargate.

Features:

* launch and stop container-based applications with simple API calls
* schedule the placement of containers across your cluster based on your resource needs
* create a consistent deployment and build experience, manage, and scale batch workloads

**What is Fargate?**

***Fargate*** is a technology that is used with ECS to run containers without having to manage servers or clusters of Amazon EC2 instances.

You don’t have to provision, configure, or scale clusters to run containers. This removes the need to choose server types, decide when to scale your clusters.

Guide:<https://docs.aws.amazon.com/AmazonECS/latest/developerguide/getting-started-fargate.html>

**Clusters**

***Cluster*** is a Region-specific:

* logical grouping of tasks or services
* grouping of container instances (if you are running tasks or services that use the EC2)
* logical grouping of capacity providers (if you are using capacity providers)

States:

* ACTIVE - cluster is ready to accept tasks
* PROVISIONING - resources needed for the capacity provider are being created
* DEPROVISIONING - resources needed for the capacity provider are being deleted
* FAILED - resources needed for the capacity provider have failed to create
* INACTIVE - cluster has been deleted

Guide:<https://docs.aws.amazon.com/AmazonECS/latest/developerguide/create_cluster.html>

**Task Definitions**

***Task definition*** is required to run Docker containers in ECS.

Parameters you can specify:

* The Docker image to use with each container in your task
* How much CPU and memory to use with each task or each container
* The launch type to use
* The Docker networking mode to use for the containers in your task
* The logging configuration to use for your tasks
* Whether the task should continue to run if the container finishes or fails
* The command the container should run when it is started
* Any data volumes that should be used with the containers in the task
* The IAM role that your tasks should use

Guide:<https://docs.aws.amazon.com/AmazonECS/latest/developerguide/create-task-definition.html>

**Services**

***Service*** is an ability to run and maintain a specified number of instances of a task definition simultaneously in ECS cluster.

If any of your tasks should fail or stop for any reason, the ECS service scheduler launches another instance of your task definition to replace it and maintain the desired count of tasks.

Load Balancer

Elastic Load Balancing is a tool to distribute traffic across the tasks in your service.

ECS services support the Application Load Balancer (route HTTP/HTTPS traffic), Network Load Balancer and Classic Load Balancer (route TCP traffic) load balancer types.

Auto Scaling

Auto Scaling is the ability to increase or decrease the desired count of tasks in your Amazon ECS service automatically.

**CloudWatch**

CloudWatch is a toll which collects and processes raw data from ECS into readable, near real-time metrics.

These statistics are recorded for a period of two weeks, so that you can access historical information and gain a better perspective on how your clusters or services are performing. ECS metric data is automatically sent to CloudWatch in 1-minute periods.