**Compute Service**

Workloads = Applications, database, container

Graphical user interface, application

Description automatically generated

***Amazon EC2 instance types***

1. General Purpose
2. Compute Optimized
3. Memory Optimized
4. Accelerated Computing
5. Storage Optimized

***Amazon Auto Scaling Group***

Within Amazon EC2 Auto Scaling, you can use two approaches-

* ***Dynamic scaling*** responds to changing demand.
* ***Predictive scaling***automatically schedules the right number of Amazon EC2 instances based on predicted demand.

Diagram

Description automatically generated

Amazon EC2 Instance

If you do not specify the desired number of Amazon EC2 instances in an Auto Scaling group, the desired capacity defaults to your minimum capacity.

*Amazon Elastic Load Balancing*

Automatically distribute traffic across multiple EC2 instances.

**Advantages**

* Increase availability and fault tolerance
* Configure health checks
* Offload encryption and decryption

**Load Balancer Type**

* Application Load Balancer (App layer)
* Network Load Balancer (Network layer)
* Gateway Load Balancer (third party virtual appliances)

ELB is regional construct. (runs on regional level)

*A picture containing diagram

Description automatically generated*

Auto Scaling Group

ELB (Elastic Load Balancing)

|  |  |
| --- | --- |
| Amazon EC2 | Host Traditional Application Full Access to OS |
| AWS Lambda (Serverless) | Host short running function  Service Oriented Application Event Driven Application  No provisioning or managing servers |
| EKS Or ECS | Host Container based workload on AWS  Amazon EC2 that you manage |
| Amazon Fargate (Serverless) | Host Container based workload on AWS  Amazon EC2 that AWS manage for you  It is a serverless compute engine for containers. AWS Fargate is compatible [with both Amazon Elastic Container Service](https://aws.amazon.com/ecs/?pg=ln&sec=hiw) (ECS) and [Amazon Elastic Kubernetes Service](https://aws.amazon.com/eks/?pg=ln&sec=hiw) (EKS). |