Overview

> **!!** Compute for any workload.
>
> - Instances
> - Containers
> - Serverless computing

Compute: Refers to using a computer to process something.

Instances

---

aka Virtual Machines.

- **Amazon EC2**: Secure and resizable virtual machines in the cloud
- **Amazon EC2 Spot**: Fault tolerant workloads at 90% off.
- **Amazon EC2 Auto Scaling**: Add or remove capacity based on demand.
- **Amazon Lightsail**: Easy to use cloud platform to build applications or websites.

Containers

- **Amazon ECS**: Run secure, reliable and scalable containers.
- **Amazon ECR**: Registry for storing, managing and deploying container images.
- **Amazon EKS**: Fully managed Kubernetes service.

Serverless Computing

- **AWS Lambda**: A compute service to run code without servers.

Edge

- **AWS Outposts**: Run AWS services on-premises
- **AWS Snow Family**: Tools and services to bring migrate data to AWS.
- **AWS WaveLength**: Access AWS services via 5G networks.
- **VMWare Cloud on AWS**: Migrate VMWare workloads.
- **AWS Local Zones**: Run latency sensitive apps closer to end users (edge computing)

EC2

- Rent virtual computers.
- Choose from various types with differing CPU, RAM and storage.
- Different optimizations for different use cases.
- Different payment modes, (ex: by hour or by second).

Benefit of pay as you go

Consider a 24-hour workload. Given the cost of an EC2 instance to be $0.10 per hour, it would cost $2.40 to do it on a single instance, and take 24 hours. But if the load can be broken into 24 1-hour loads, they can be parallelly completed across 24 separate EC2 instances, taking a total of 1 hour and still costing the same amount, $2.40.

Elastic / Resizable

A user can start with a basic configuration, but change specs as the need arises by simply turning off the instance, changing the config and starting the instance again. They would be charged only for the time spent on each configuration.

Containers

Create a package for a program, including the required dependencies and environment.

Serverless

Compute without dedicated servers.

Avoids the need of having a 24/7 server, or complicated power on/off cycles for something like cron jobs, or conditional code execution (like hitting an API endpoint). The code is executed only when needed.

Major provider is AWS Lambda which.

- Provides serverless compute service.
- Runs code in response to events.
- Runs code somewhere (it doesn't matter where).

Example uses

1. Automated image resizing / thumbnail generation on uploading image.

2. Posting tweets every day at 9am.

Pricing

- Charged by the millisecond.
- Price depends on RAM use.