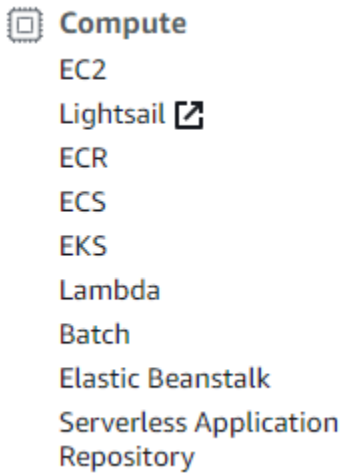
Compute services @ AWS

- Potureddi Gowtham

Compute

- A cloud infrastructure compute will let you have compute hosts, known as instance.
- You can launch as many instances as you need based on user use-case and requirements.
- Once after you launch an instance you can securely login from your computer, restart it, attach and detach with volumes or data and terminate once work is done.
- But there won't be any data saved once after you terminate.



Amazon EC2 (Elastic compute cloud)

- EC2 is a webservice that provide secure, reliable and resizable compute capacity in cloud.
- This is specifically made to make web scaling easy for users.
- This is infrastructure as a service, because we were left with complete control over instance and its computing resources.
- This takes a very less time to boot up and get the server up and running (maybe in few mins).
- Because of this flexibility and speed, it's easy to scale both up and down.

Pricing models

1. On Demand
    a. This allows you to pay a fixed rate by the hour or by even seconds with no commitment.
    b. Awesome for developers because you can go launch an instance, work on it and play around then terminate.
    c. Awesome for startups.

2. Reserved
    a. Provides you with a capacity reservation and offers a significant discount on the hourly charges for an instance.
    b. Contract term is 1 year to 3 years.
    c. More you pay upfront more u can save.
    d. Different Reserved
        i. Standard Reserved instances
            - They offer up to 75% off on demand instances.
        ii. Convertible Reserved instances
            - They offer 54% off, but you can change the OS or any services you want on the go.
        iii. Scheduled Reserved instances
            - This is to launch with-in the time windows you reserve.
            - Ex: like when you know website traffic going to increase at a particular time frame.

3. Spot
    a. Enables you to bid whatever price you are comfortable with to pay for instance capacity.
    b. It's like when instances are available for the price you want to pay, you will get those.
    c. If instance price increase, your instances will be terminated instantly.
    d. It's like stock market.

4. Dedicated Hosts
    a. Physical servers dedicated for your use.

EC2 Instance types:

F – FPGA I – IOPS G – Graphics H – High Disk Throughput

T – Cheap General purpose (t2) D – Density R – Ram

M – Main Choice for general purpose apps C – Compute

P – Graphics (pics) X – Extreme Memory U – Bare Metal

Z – Extreme memory and CPU A – Arm-based workloads

Family	Speciality	Use case
F1	Field Programmable Gate Array	Genomics research, financial analytics, real-time video processing, big data etc
I3	High Speed Storage	NoSQL DBs, Data Warehousing etc
G3	Graphics Intensive	Video Encoding/ 3D Application Streaming
H1	High Disk Throughput	MapReduce-based workloads, distributed file systems such as HDFS and MapR-FS
T3	Lowest Cost, General Purpose	Web Servers/Small DBs
D2	Dense Storage	Fileservers/Data Warehousing/Hadoop
R5	Memory Optimized	Memory Intensive Apps/DBs
M5	General Purpose	Application Servers
C5	Compute Optimized	CPU Intensive Apps/DBs
P3	Graphics/General Purpose GPU	Machine Learning, Bit Coin Mining etc
X1	Memory Optimized	SAP HANA/Apache Spark etc
Z1D	High compute capacity and a high memory footprint.	Ideal for electronic design automation (EDA) and certain relational database workloads with high per-core licensing costs.
A1	Arm-based workloads	Scale-out workloads such as web servers
U-6tb1	Bare Metal	Bare metal capabilities that eliminate virtualization overhead

remember - '**Fight dr. M. cpxz AU** '

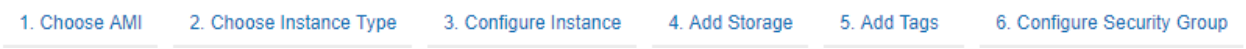
FYI :

- 'ec2-t2 micro' instance is free to use for 1 year.
- Feel free to play around.

Launch EC2 instance

Steps

- Choose AMI
- Instance type
- Configure instance
- Add storage
- Add Tags
- Configure Security group
- Launch



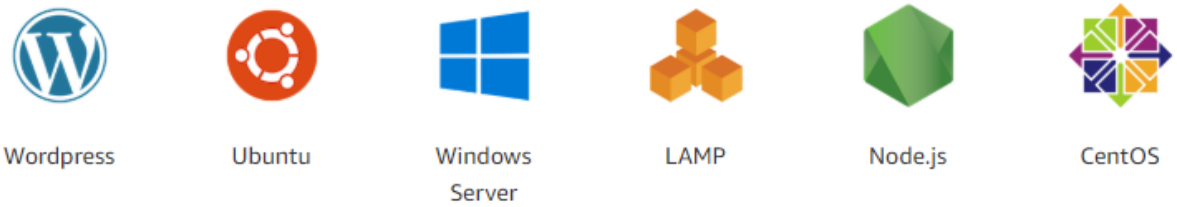
LightSail

- Amazon LightSail is designed to be the easiest way to launch and manage a virtual private server.
- LightSail includes

    b. virtual machine
    c. SSD-based storage
    d. data transfer
    e. DNS management
    f. and a static IP address – for a low price start $ 3.5 current price.

Uses :

1. Simple web application
    - Easily deploy a web application with LAMP, Nginx, MEAN and Node.js.
2. Websites
    - pre-configured applications like WordPress, Magento, Plesk, and Joomla.
3. Business software
    - quickly launch your line-of-business software, like file storage and sharing, backups, financial and accounting software.
4. Dev / Test environments
    - Spin up a developer or test environment in seconds complete needed test or development. And, once you're done, you can quickly destroy them.

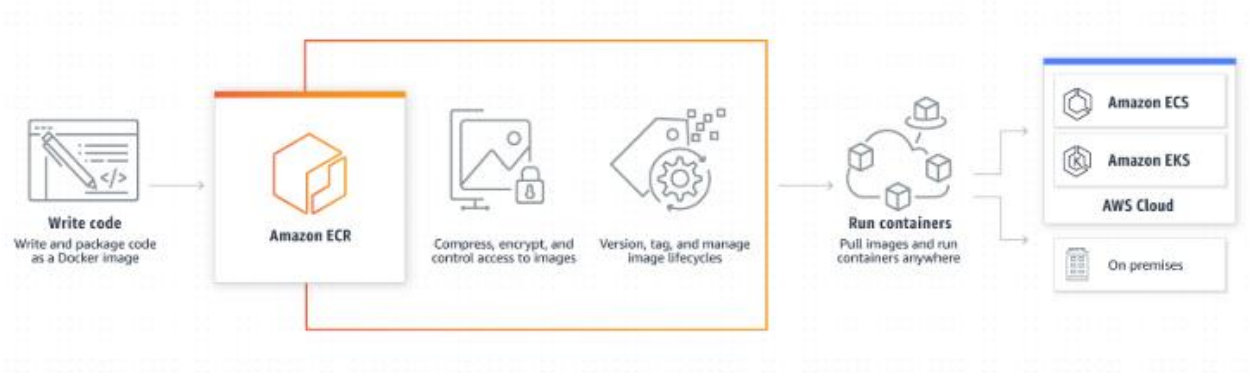


ECR (Amazon Elastic Container Registry)

- This fully-managed Docker container registry that makes it easy for developers to store, manage, and deploy Docker container images.
- Using this you need not operate your own containers and worry about scaling the underlying infrastructure.
- Use IAM to provides resource-level control of each repository.
- And pay only for the memory and data you are using.

Benefits

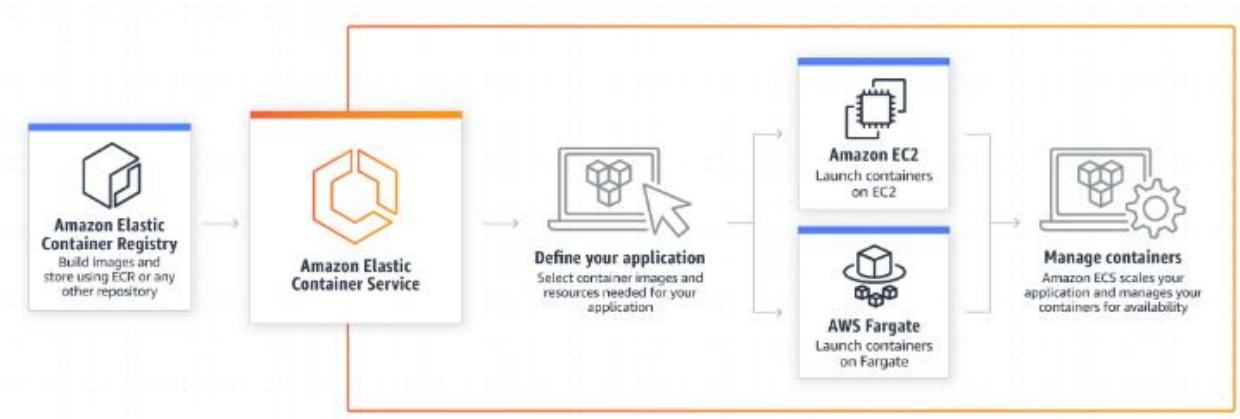
- Fully Managed
- Secure
- Highly available
- Simplified workflow



ECS (Elastic Container Service)

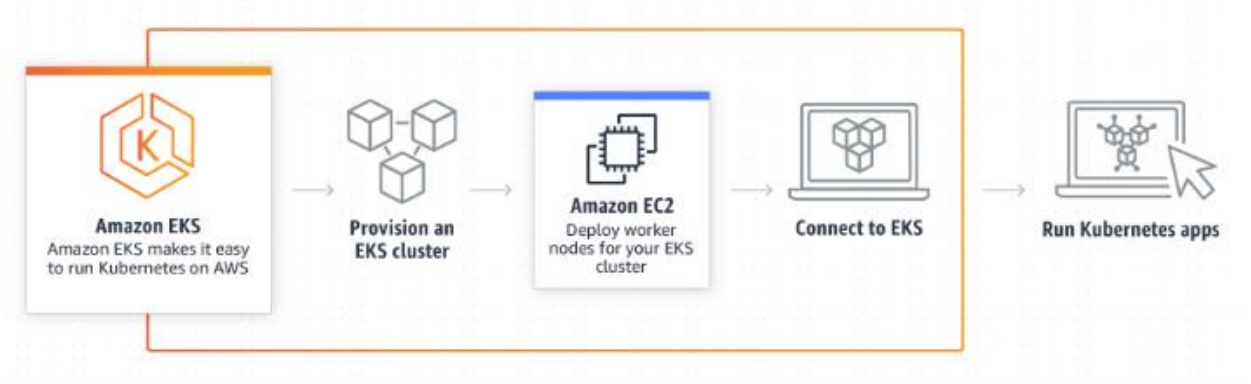
- This is a Container Orchestration tool.
- We need not worry about installing our own tool like kubernetes.
- This can take care of manage, scale a cluster of virtual machines, schedule containers on these virtual machines.
- With simple API calls, you can launch, stop Docker enabled applications.
- You can use

- IAM roles, security groups
- load balancers
- Amazon CloudWatch Events
- AWS CloudFormation templates
- AWS CloudTrail logs



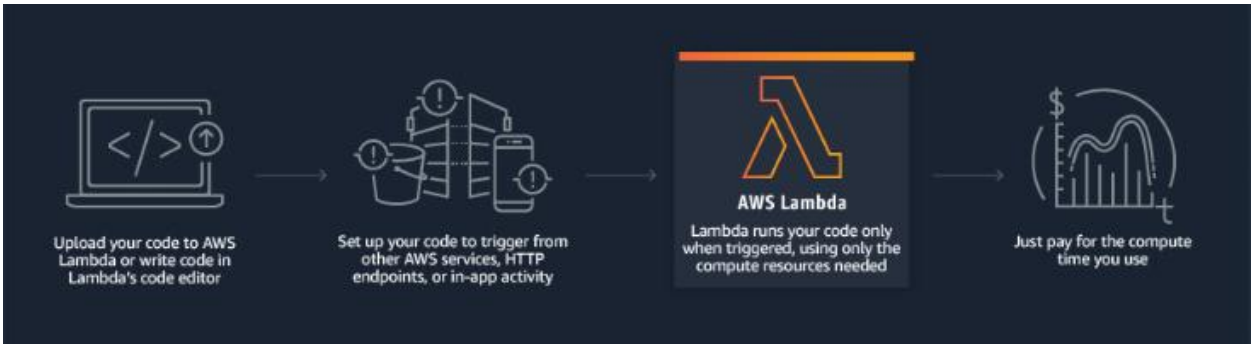
EKS (Elastic Container Service for Kubernetes)

- You can easily deploy, manage, and scale containerized application using Kubernetes.
- Using this we can run Kubernetes and use all the plugins from partners and kube community.



Lambda

- You can let Lambda run your code with needed infrastructure.
- You can use it to run any applications, backend services and many more.
- Just you need to upload your code, and lambda takes care of everything it requires to run and scale you code with high availability.
- You can set up your code to automatically be called from you web application or mobile application etc.
- You need not manage any servers.

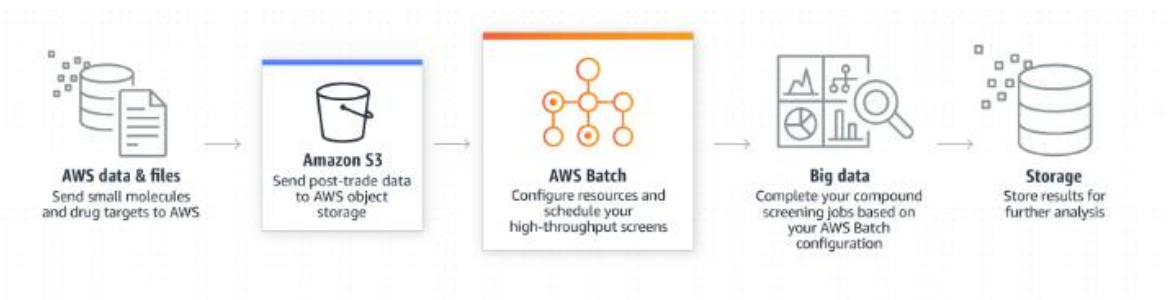


Batch

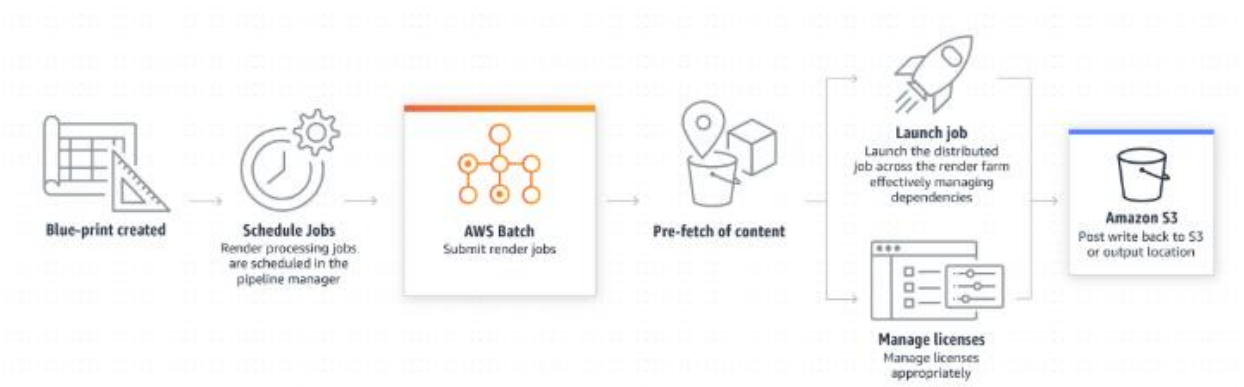
- You can run hundreds and thousands of batch computing jobs.
- You need not bother to set needed computational power or resources, AWS will do that for you.
- No need to install any resources.
- You should only pay for aws resources you used.

Ex:

Huge data for Life Sciences jobs



Digital media jobs



Elastic Beanstalk

- This is Platform as a service.
- You just need to upload you code and give some details.
- Pay only for the services you are using.
- Mostly used to build websites, blogs etc.
- You can change the configurations and make it work as you like.
- Internally we get EC2, auto scaling, notifications sender

You can use services developed on

- Java
- .net
- Php
- node.js
- python
- ruby
- go
- Docker

Servers

- Apache
- Nginx
- Passenger
- IIS

Serverless Application Repository

- Here teams, organizations, and individual developers can store and share reusable applications, and easily assemble and deploy serverless architectures.
- You can use the prebuild applications in the serverless applications repository.
- So, you or your team need not do duplicated work.
- You can publicly share the application or privately send to other teams with in your organization.
- Each application is packaged with a Serverless application model (SAM) template that defines the aws resources used.
- It free to use the application SAM templates in our projects.
- You only need to pay for the resources you are using.