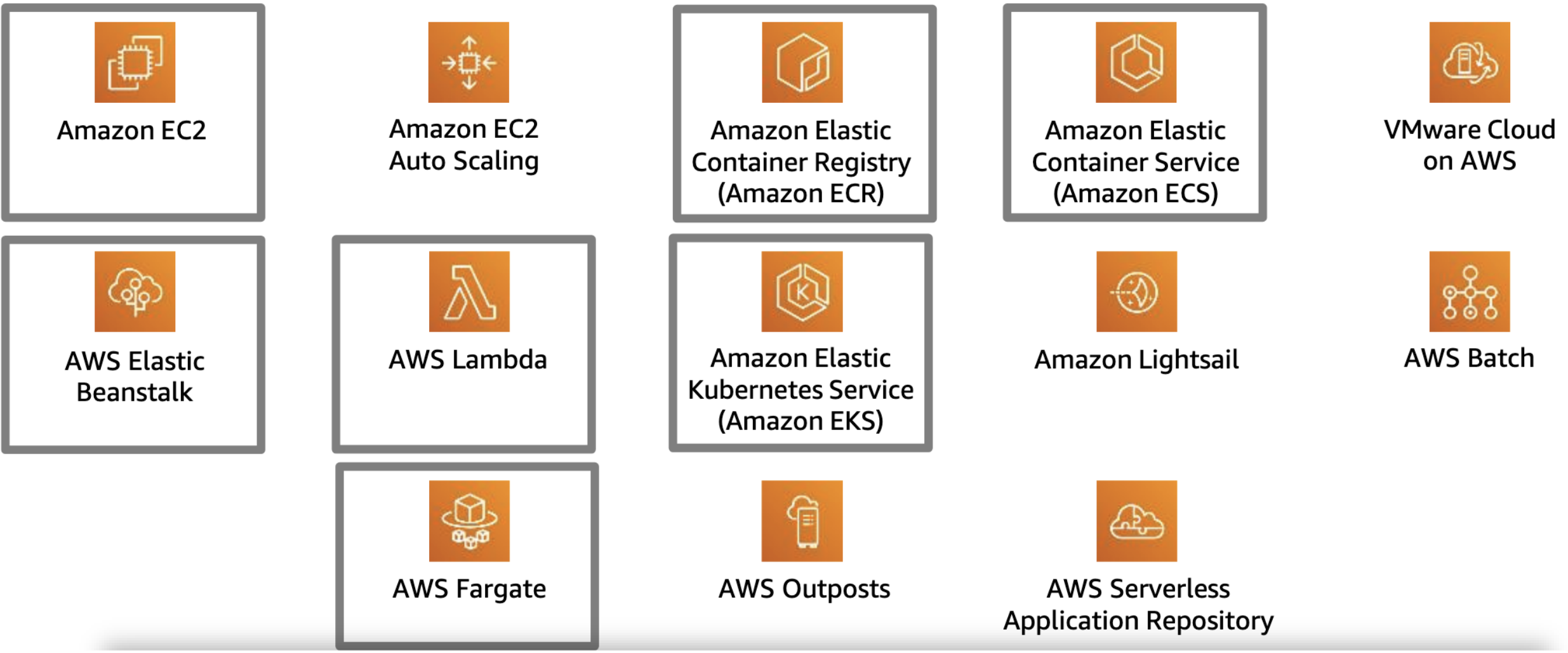
Computing

Alayna Lee

Amazon Web Services (AWS) offers many compute services. This module will discuss the highlighted services.



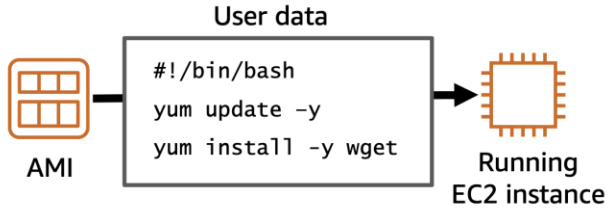
Computing services

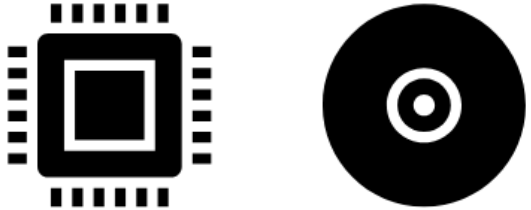
Amazon Elastic Compute Cloud (EC2)

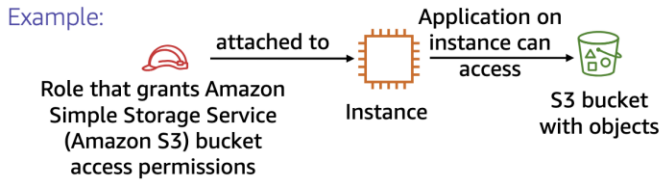
- Provides virtual cloud machines so that you don't have to pay for things on-prem
- Instances can support a variety of resources and is more efficient
- EC2 allows users to have full control over the operating system
- You can launch instances from AMIs (Amazon Machine Images)
- You can control traffic to and from instances
- Elastic refers to how easy it is to increase and decrease the amount of servers
- Compute refers to hosting running apps or processing data
- Cloud refers to the EC2 instances that you run are on the cloud

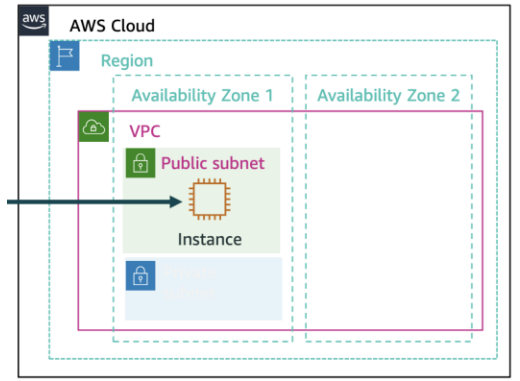
Launching a EC2 instance

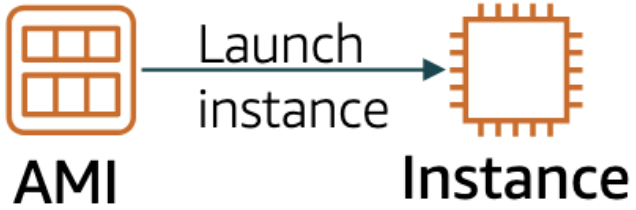
- There are 9 steps when launching a EC2 instance:
  - Select an AMI:
    - This is a template that is use when creating a EC2 instance
    - Often has software already installed
    - You can launch multiple instances from a single AMI
  - Select an instance type:
    - Determines how the instance will be used
    - Determines RAM, CPU, storage, and network performance
  - Specify network settings:
    - Specifies where the instance should be deployed (Region)
    - Identifies the VPC and subnets
  - Attach IAM role (optional):
    - If the Instance needs to interact with different AWS services, then attach a role to the instance
  - User Data Script (optional):
    - Specify user data script to customize runtime environment
  - Specify storage:
    - For each volume specify the size, volume type, and if encryption should be used



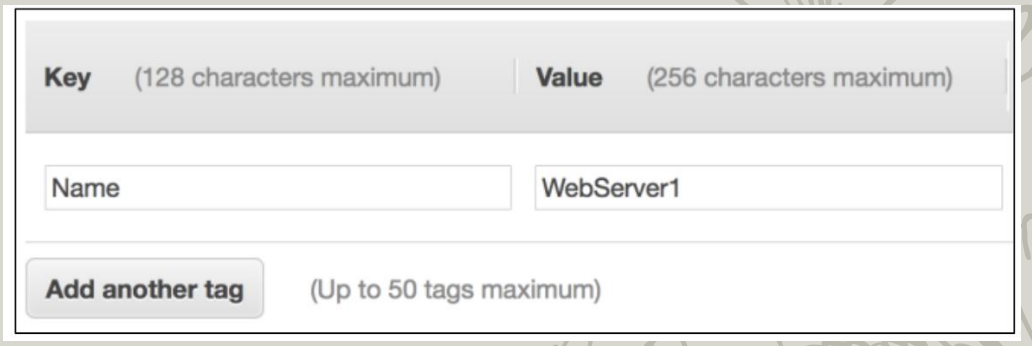




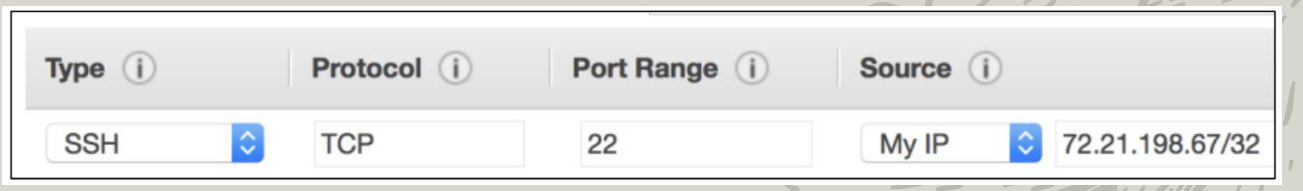




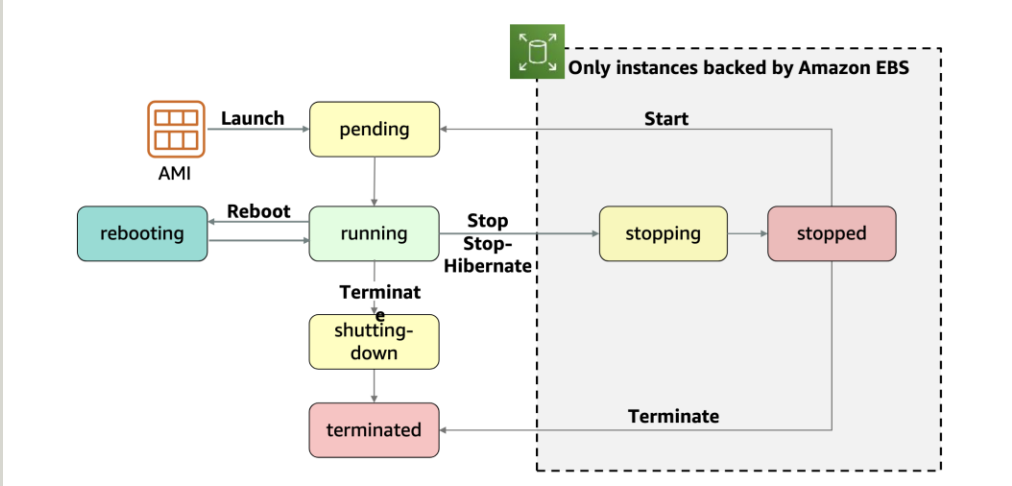
Launching a EC2 instance (Continued)



- Add tags:
  - This is a label you can assign to a AWS resource
  - You can attach metadata to a instance
  - Allows you to have more access control
- Security Group Settings:
  - A set of firewall rules
  - Must specify the port number, protocol, and source that is allowed to use the rule
- Create/Identify key pair:
  - Key pair has a public and private key
  - Users own private key and AWS owns the public key
  - Enables secure connections to instances



EC2 instance lifetime



- Pending:
  - When you first launch an instance
  - When you start a stopped instance
- Running:
  - When the instance is fully boosted and ready to run
- Shutting down:
  - The state between running and terminated
- Terminated:
  - Terminated instance remains visible for a while before getting deleted
- Stopping:
  - The state before the instance obtains the stopped state
- Stopped:
  - Will not incur the same cost as a running instance, and moves the instance to a new host

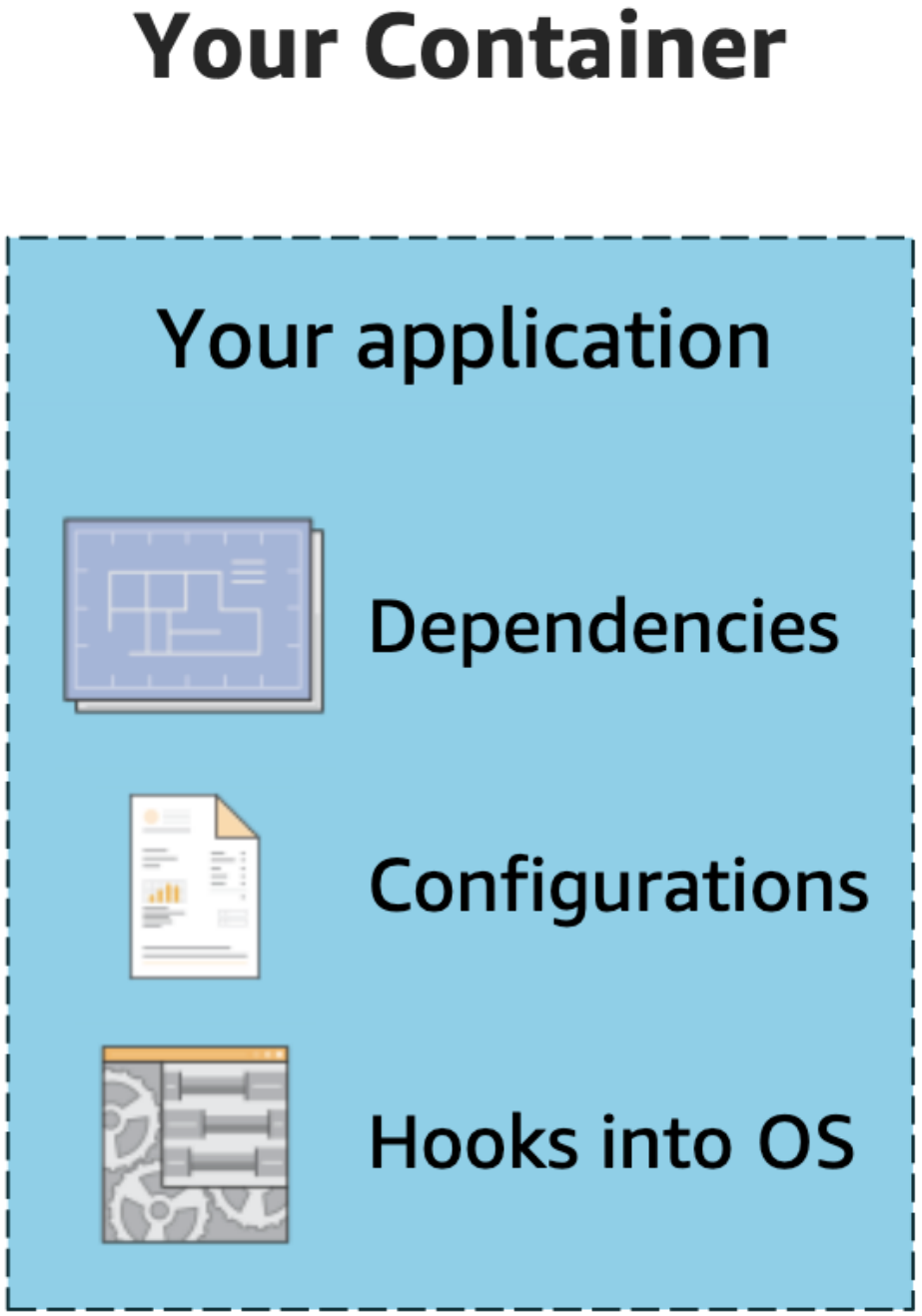
EC2 instance metadata

- This is information about the instance
- It can be used to manage or configure the instance
- You can use Amazon CloudFront to monitor the instance
- Maintains 15 months of historical data

Amazon EC2 pricing model

- Per-second-billing
  - Only available for on-demand, reserved, and spot instances
- On-demand
  - Lowest up front cost
- Dedicated host
  - Physical servers with instance capacity
- Dedicated instance
  - Instances that are on a VPC that runs on hardware that's dedicated to a single customer
- Reserved instance
  - Allows you to reserve instance capacity for 1-3 years
- Scheduled Reserved Instance
  - Allows you to purchase the reserved capacity
- Spot instances
  - Allows you to bid on unused EC2 instances

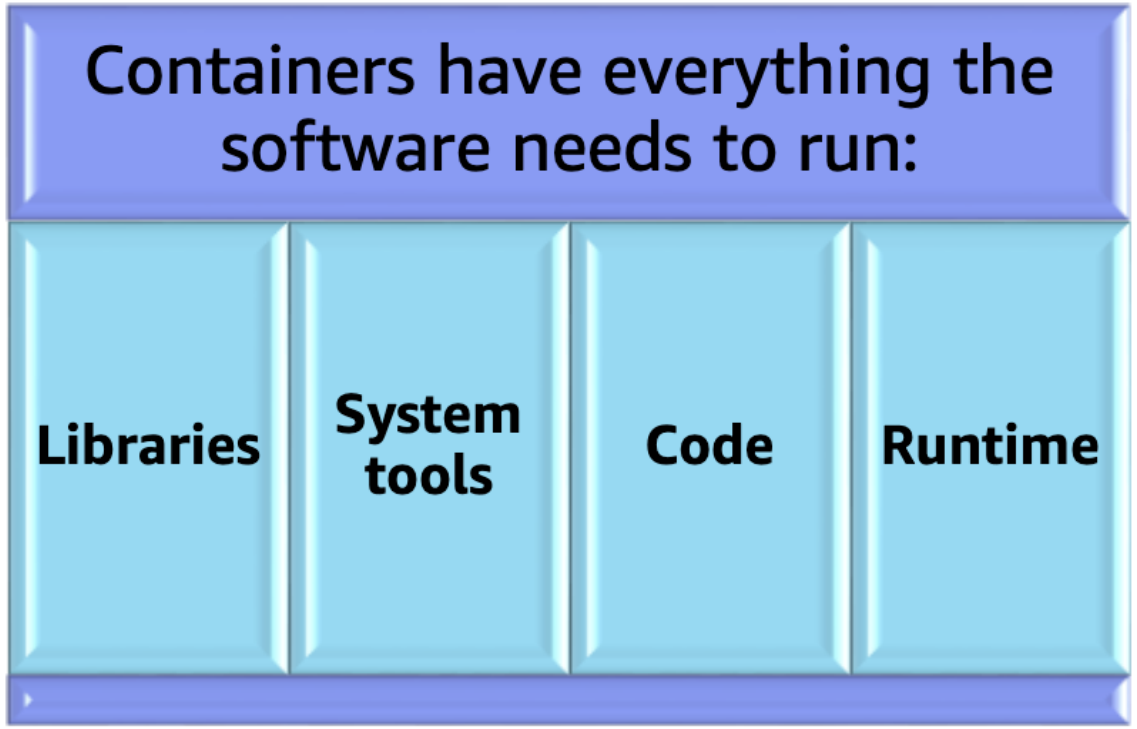
On-Demand Instances	Spot Instances	Reserved Instances	Dedicated Hosts
• Low cost and flexibility	• Large scale, dynamic workload	• Predictability ensures compute capacity is available when needed	• Save money on licensing costs • Help meet compliance and regulatory requirements



Container basics

- A method of OS virtualization
- Allows you to run applications with their dependencies in resource-isolated processes
- Delivers environmental consistency because everything is packed on to one object
- Containers are smaller than VMs

Docker



- Software application that allows you to deploy, build, and test applications quickly
- Containers are ran on dockers
- A container has everything a software app needs to run

Amazon ECS, Kubernetes, EKS, and ECR

- Amazon ECS:
  - Amazon Elastic container service
  - Easy to manage container service
- Kubernetes:
  - Open-source software for container orchestration
- Amazon EKS:
  - Amazon Elastic Kubernetes services
  - Allows the user to use Kubernetes on AWS
  - Helps manage EC2 instance clusters
- Amazon ECR:
  - Elastic container registry
  - Makes it easy to deploy and manage Docker container images

AWS lambda

- Serverless computing service
- Allows you to run code without having to manage servers
- You have to create a lambda function, and then upload code
- You only have to pay for the compute time you use
- It supports multiple coding languages
- Built in fault tolerance
- Completely automated administration
- Event source is a event that triggers a lambda function

AWS Elastic Beanstalk

- Easy way to get web applications up and running
- No additional charge for Elastic Beanstalk
- Automatically handles the deployment you just have to upload your code
- Allows you to develop web applications without focusing on configuring servers