Cloud Computing

Session 7 – IaaS (EC2, EBS, S3, EFS)

BITS
Pilani

AWS compute services

Services	Key Concepts	Characteristics	Ease of Use
• Amazon EC2	• Infrastructure as a service (IaaS) • Instance-based • Virtual machines	• Provision virtual machines that you can manage as you choose	A familiar concept to many IT professionals.
• AWS Lambda	• Serverless computing • Function-based • Low-cost	• Write and deploy code that runs on a schedule or that can be triggered by events • Use when possible (architect for the cloud)	A relatively new concept for many IT staff members, but easy to use after you learn how.
• Amazon ECS • Amazon EKS • AWS Fargate • Amazon ECR	• Container-based computing • Instance-based	• Spin up and run jobs more quickly	AWS Fargate reduces administrative overhead, but you can use options that give you more control.
• AWS Elastic Beanstalk	• Platform as a service (PaaS) • For web applications	• Focus on your code (building your application) • Can easily tie into other services—databases, Domain Name System (DNS), etc.	Fast and easy to get started.

Choosing the optimal compute service

The optimal compute service or services that you use will depend on your use case

Some aspects to consider –

- What is your application design?

- What are your usage patterns?

- Which configuration settings will you want to manage?

Selecting the wrong compute solution for an architecture can lead to lower performance efficiency

- A good starting place—Understand the available compute options

Amazon Elastic Compute Cloud (Amazon EC2)

- Example uses of Amazon EC2 instances

  - Application server

  - Web server

  - Database server

  - Game server

  - Mail server

  - Media server

  - Catalog server

  - File server

  - Computing server

  - Proxy serve
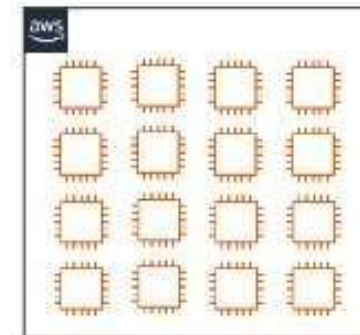
Amazon EC2 instances

Amazon EC2 overview

- Amazon Elastic Compute Cloud (Amazon EC2)

- Provides virtual machines—referred to as EC2 instances—in the cloud.

- Gives you full control over the guest operating system (Windows or Linux) on each instance.

- You can launch instances of any size into an Availability Zone anywhere in the world.

- Launch instances from Amazon Machine Images (AMIs).

- Launch instances with a few clicks or a line of code, and they are ready in minutes.

- Control traffic to and from instances

Launching an Amazon EC2 instance

Choices made using the Launch Instance Wizard:

1. AMI
2. Instance Type
3. Network settings
4. IAM role
5. User data
6. Storage options
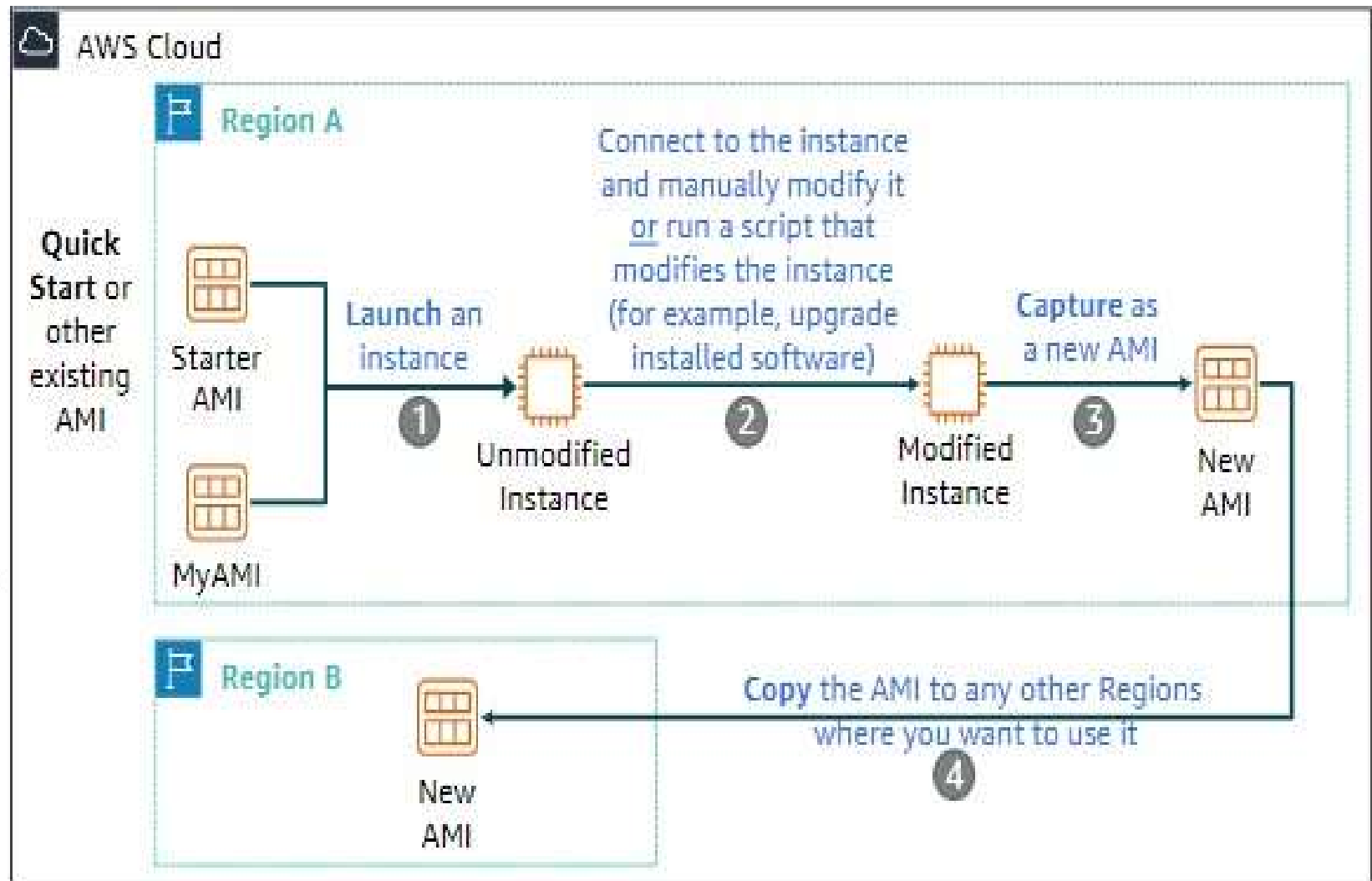7. Tags
8. Security group
9. Key pair

- **Amazon Machine Image (AMI)**
  - Is a template that is used to create an EC2 instance (which is a **virtual machine, or VM**, that runs in the AWS Cloud)
  - Contains a **Windows** or **Linux** operating system
  - Often also has some **software** pre-installed
- **AMI choices:**
  - Quick Start – *Linux and Windows AMIs that are provided by AWS*
  - My AMIs – *Any AMIs that you created*
  - AWS Marketplace – *Pre-configured templates from third parties*
  - Community AMIs – *AMIs shared by others; use at your own risk*

Creating a new AMI: Example

Launching an Amazon EC2 instance

Choices made using the
Launch Instance Wizard:

1. AMI
2. Instance Type
3. Network settings
4. IAM role
5. User data
6. Storage options
7. Tags
8. Security group
9. Key pair

- Consider your use case
  - How will the EC2 instance you create be used?
- The **instance type** that you choose determines –
  - Memory (RAM)
  - Processing power (CPU)
  - Disk space and disk type (Storage)
  - Network performance
- Instance type categories –
  - General purpose
  - Compute optimized
  - Memory optimized
  - Storage optimized
  - Accelerated computing
- Instance types offer *family*, *generation*, and *size*
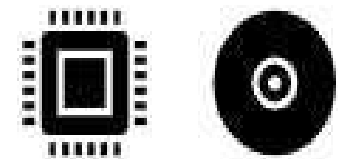
EC2 instance type   naming and sizes

Instance type naming

- Example: **t3.large**
  - T is the family name
  - 3 is the generation number
  - Large is the size

Example instance sizes

Instance Name	vCPU	Memory (GB)	Storage
t3.nano	2	0.5	EBS-Only
t3.micro	2	1	EBS-Only
t3.small	2	2	EBS-Only
t3.medium	2	4	EBS-Only
t3.large	2	8	EBS-Only
t3.xlarge	4	16	EBS-Only
t3.2xlarge	8	32	EBS-Only

Select instance type: Based on use case
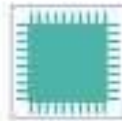
|	General Purpose	Compute Optimized	Memory Optimized	Accelerated Computing	Storage Optimized
Instance Types	a1, m4, m5, t2, t3	c4, c5	r4, r5, x1, z1	f1, g3, g4, p2, p3	d2, h1, i3
Use Case	Broad	High performance	In-memory databases	Machine learning	Distributed file systems

Launching an Amazon EC2 instance

3. Specify network settings

Choices made by using the Launch Instance Wizard:

1. AMI
2. Instance Type
3. Network settings
4. IAM role
5. User data
6. Storage options
7. Tags
8. Security group
9. Key pair

- Where should the instance be deployed?
  - Identify the **VPC** and optionally the **subnet**
- Should a **public IP address** be automatically assigned?
  - To make it internet-accessible

Example: specify to deploy the instance here
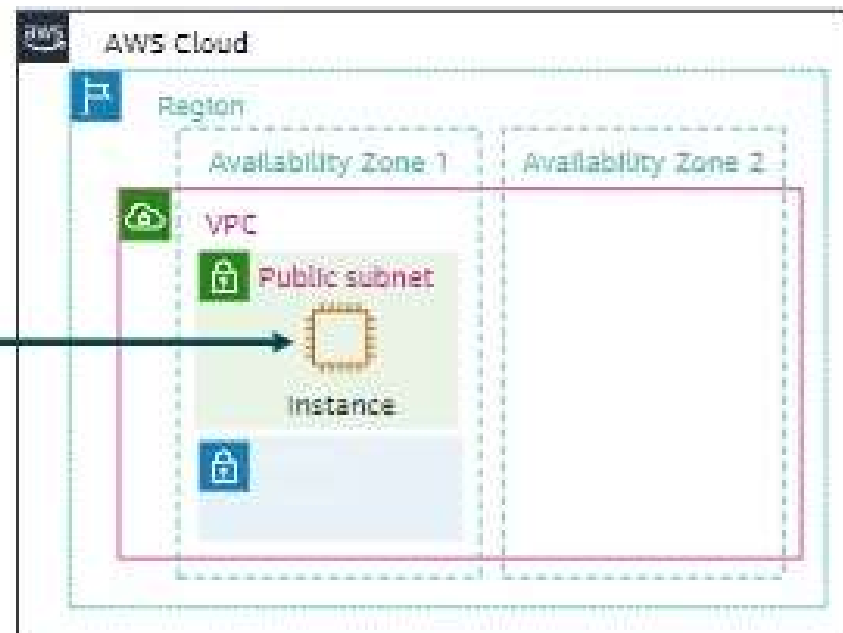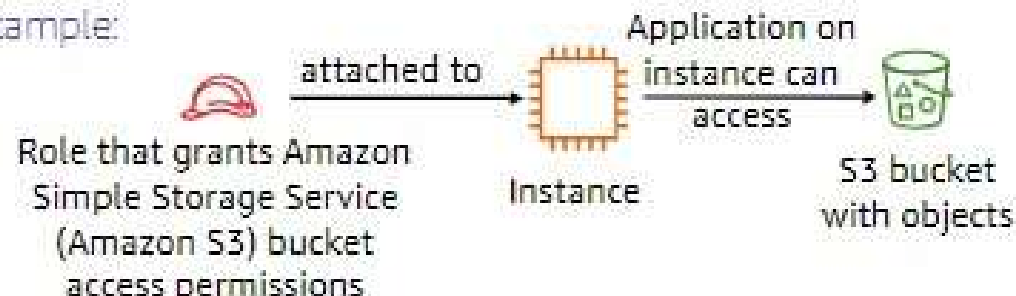
Launching an Amazon EC2 instance

4. Attach IAM role (optional)

Choices made by using the Launch Instance Wizard:

1. AMI
2. Instance Type
3. Network settings
4. IAM role
5. User data
6. Storage options
7. Tags
8. Security group
9. Key pair

- Will software on the EC2 instance need to interact with other AWS services?
  - If yes, attach an appropriate **IAM Role**.
- An AWS Identity and Access Management (IAM) role that is attached to an EC2 instance is kept in an **instance profile**.
- You are *not* restricted to attaching a role only at instance launch.
 - You can also attach a role to an instance that already exists.

Example:

Role that grants Amazon Simple Storage Service (Amazon S3) bucket access permissions — attached to → Instance — Application on instance can access → S3 bucket with objects

Launching an Amazon EC2 instance

5. User data script (optional)

Choices made by using the Launch Instance Wizard:

1. AMI
2. Instance Type
3. Network settings
4. IAM role
5. User data
6. Storage options
7. Tags
8. Security group
9. Key pair

User data

```
#!/bin/bash
yum update -y
yum install -y wget
```

AMI → Running EC2 instance
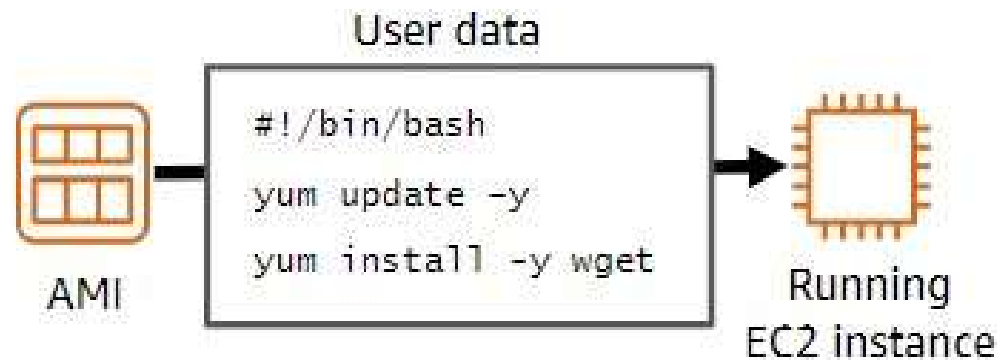
- Optionally specify a user data script at instance launch
- Use **user data** scripts to customize the runtime environment of your instance
  - Script runs the first time the instance starts
- Can be used strategically
  - For example, reduce the number of custom AMIs that you build and maintain
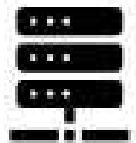
Launching an Amazon EC2 instance

6. Specify storage

Choices made by using the Launch Instance Wizard:

1. AMI
2. Instance Type
3. Network settings
4. IAM role
5. User data
6. Storage options
7. Tags
8. Security group
9. Key pair

- Configure the root volume
  - Where the guest operating system is installed
- Attach additional storage volumes (optional)
  - AMI might already include more than one volume
- For each volume, specify:
  - The **size** of the disk (in GB)
  - The **volume type**
    - Different types of solid state drives (SSDs) and hard disk drives (HDDs) are available
  - If the volume will be deleted when the instance is terminated
  - If **encryption** should be used

Launching an Amazon EC2 instance

7. Add tags

Choices made by using the Launch Instance Wizard:

1. AMI
2. Instance Type
3. Network settings
4. IAM role
5. User data
6. Storage options
7. Tags
8. Security group
9. Key pair

- A tag is a label that you can assign to an AWS resource.
  - Consists of a *key* and an optional *value*.
- Tagging is how you can attach **metadata** to an EC2 instance.
- Potential benefits of tagging—Filtering, automation, cost allocation, and access control.

Example:

Key (128 characters maximum)	Value (256 characters maximum)
Name	WebServer1

Add another tag   (Up to 50 tags maximum)

Launching an Amazon EC2 instance

8. Security group settings

Choices made by using the Launch Instance Wizard:

1. AMI
2. Instance Type
3. Network settings
4. IAM role
5. User data
6. Storage options
7. Tags
8. Security group
9. Key pair

- A security group is a **set of firewall rules** that control traffic to the instance.
  - It exists *outside* of the instance's guest OS.
- Create **rules** that specify the **source** and which **ports** that network communications can use.
  - Specify the **port** number and the **protocol**, such as Transmission Control Protocol (TCP), User Datagram Protocol (UDP), or Internet Control Message Protocol (ICMP).
  - Specify the **source** (for example, an IP address or another security group) that is allowed to use the rule.

Example rule:

Type ⓘ	Protocol ⓘ	Port Range ⓘ	Source ⓘ	
SSH	TCP	22	My IP	72.21.198.67/32

Launching an Amazon EC2 instance

9. Identify or create the key pair

**Choices made by using the Launch Instance Wizard:**

1. AMI
2. Instance Type
3. Network settings
4. IAM role
5. User data
6. Storage options
7. Tags
8. Security group
9. Key pair

- At instance launch, you specify an existing key pair *or* create a new key pair.
- A key pair consists of –
  - A **public key** that AWS stores.
  - A **private key** file that you store.
- It enables secure connections to the instance.
- For **Windows AMIs** –
  - Use the private key to obtain the administrator password that you need to log in to your instance.
- For **Linux AMIs** –
  - Use the private key to use SSH to securely connect to your instance.

mykey.pem

Launching an Amazon EC2 instance

Amazon EC2 console view of a running EC2 instance
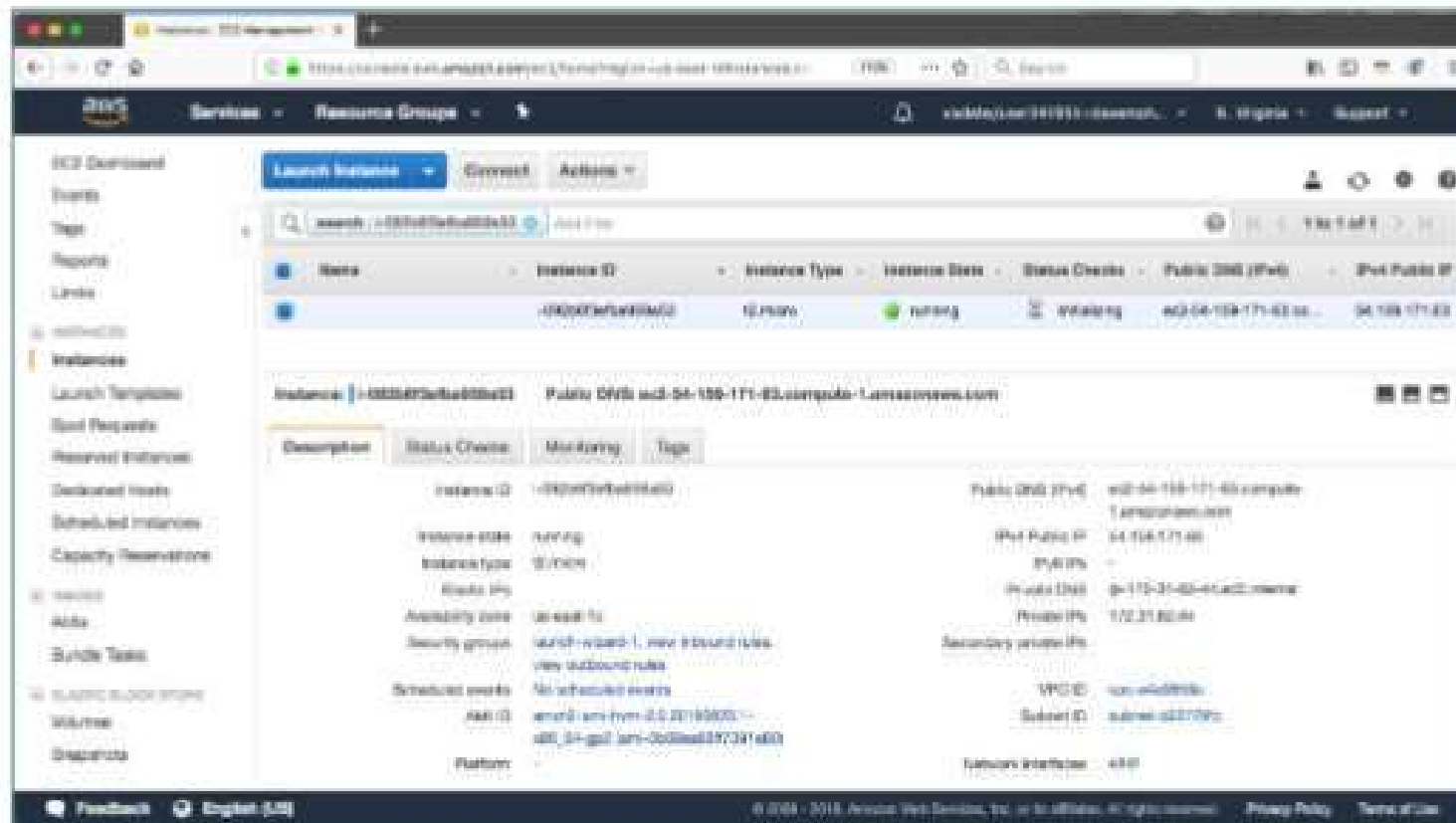
Launching an Amazon EC2 instance

Another option: Launch an EC2 instance with the AWS Command Line Interface

- EC2 instances can also be created programmatically.

AWS Command Line Interface (AWS CLI)

- This example shows how simple the command can be.

  - This command assumes that the key pair and security group already exist.

  - More options could be specified. See the AWS CLI Command Reference for details.

Example command:

```
aws ec2 run-instances \
--image-id ami-1a2b3c4d \
--count 1 \
--instance-type c3.large \
--key-name MyKeyPair \
--security-groups MySecurityGroup \
--region us-east-1
```

Amazon EC2 instance lifecycle
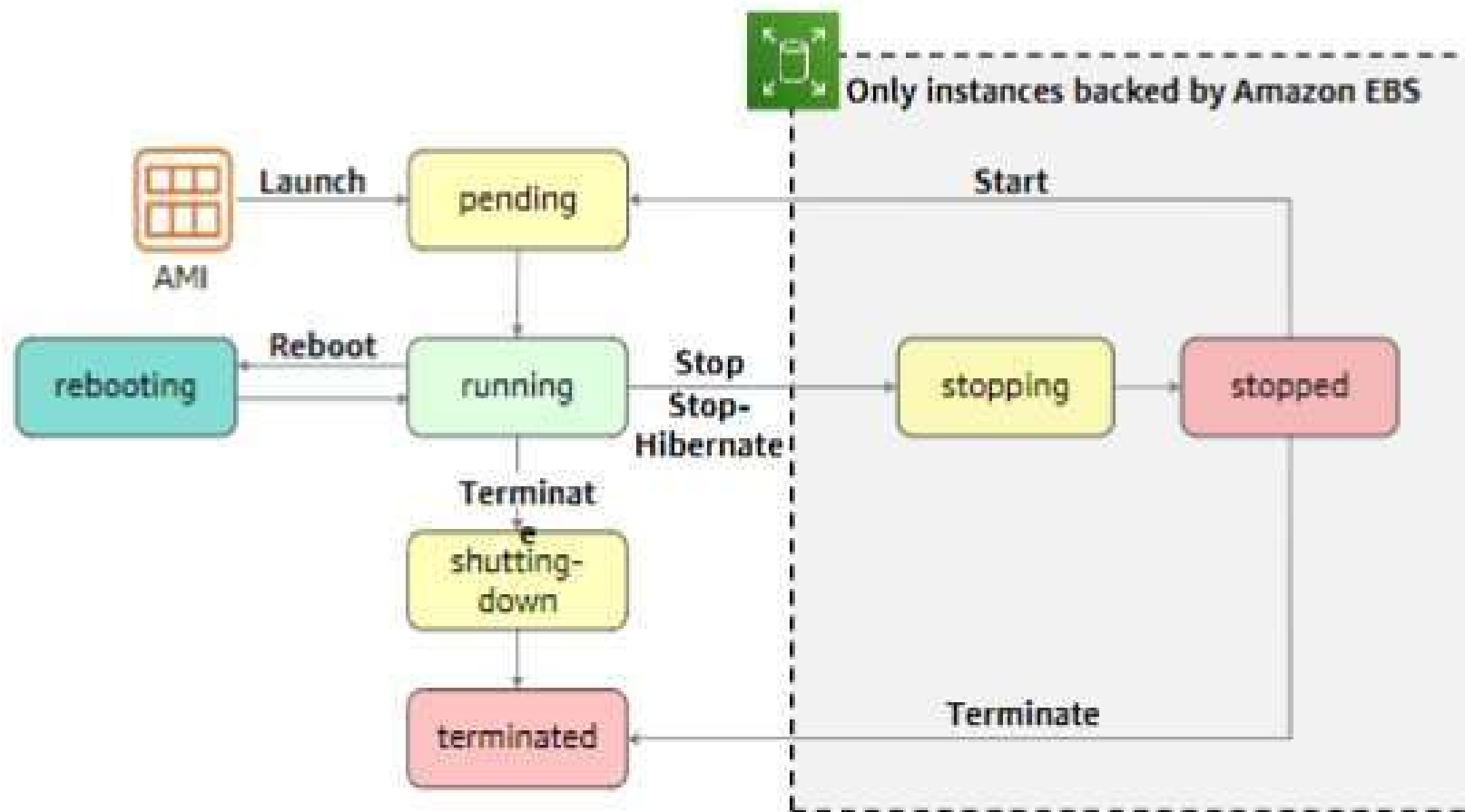
Consider using an Elastic IP address

- Rebooting an instance will not change any IP addresses or DNS hostnames

- When an instance is stopped and then started again –

    - The publicIPv4 address and external DNS hostname will change.

    - The privateIPv4 address and internal DNS hostname do not change

- If you require a persistent public IP address –

    - Associate an Elastic IP address with the instance.

    - Elastic IP address characteristics –

    - Can be associated with instances in the Region as needed.

    - Remains allocated to your account until you choose to release it

EC2 instance metadata

- Instance metadata is data about your instance.

- While you are connected to the instance, you can view it –•In a browser: http://169.254.169.254/latest/meta-data/

- In a terminal window: curl http://169.254.169.254/latest/meta- data/

- Example retrievable values –

- Public IP address, private IP address, public hostname, instance ID, security groups, Region, Availability Zone.

- Any user data specified at instance launch can also be accessed at: http://169.254.169.254/latest/user-data/

- It can be used to configure or manage a running instance.

- For example, author a configuration script that reads the metadata and uses it to configure applications or OS settings

Amazon EC2 pricing models

- On-Demand Instances
  - Pay by the hour
  - No long-term commitments.
  - Eligible for the AWS Free Tier
- Dedicated Hosts
  - A physical server with EC2 instance capacity fully dedicated to your use
- Dedicated Instances
  - Instances that run in a VPC on hardware that is dedicated to a single customer
- Reserved Instances
  - Full, partial, or no upfront payment for instance you reserve.
  - Discount on hourly charge for that instance.
  - 1-year or 3-year term
- Scheduled Reserved Instances
  - Purchase a capacity reservation that is always available on a recurring schedule you specify.
  - 1-year term

Amazon EC2 pricing models

- Spot Instances
  - Instances run as long as they are available and your bid is above the Spot Instance price.
    - They can be interrupted by AWS with a 2-minute notification.
    - Interruption options include terminated, stopped or hibernated.
    - Prices can be significantly less expensive compared to On-Demand Instances
    - Good choice when you have flexibility in when your applications can run
- Per second billing available for On-Demand Instances, Reserved Instances, and Spot Instances that run Amazon Linux or Ubuntu

Amazon EC2 pricing models: Benefits and Usecases

- Benefits

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

Spiky Workloads	Time-Insensitive Workloads	Steady-State Workloads	Highly Sensitive Workloads

- Use cases

| On-Demand Instances	Spot Instances	Reserved Instances	Dedicated Hosts
• Short-term, spiky, or unpredictable workloads • Application development or testing	• Applications with flexible start and end times • Applications only feasible at very low compute prices • Users with urgent computing needs for large amounts of additional capacity	• Steady state or predictable usage workloads • Applications that require reserved capacity, including disaster recovery • Users able to make upfront payments to reduce total computing costs even further	• Bring your own license (BYOL) • Compliance and regulatory restrictions • Usage and licensing tracking • Control instance placement

AWS Storage

- Amazon Elastic Block Store (Amazon EBS)

- Amazon Simple Storage Service (Amazon S3)

- Amazon Elastic File System (Amazon EFS)

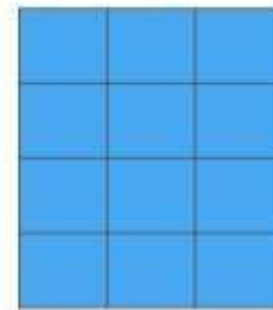- Amazon Simple Storage Service Glacier

Storage

- AWS storage options: Block storage versus object storage

- What if you want to change one character in a 1-GB file?

- Block: Change one block (piece of the file)that contains the character

- Object: Entire file must be updated
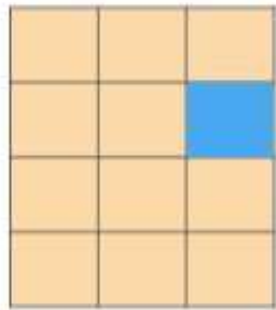
Block storage

Object storage

Amazon EBS

- Amazon EBS enables you to create individual storage volumes and attach them to an Amazon EC2 instance:

- Amazon EBS offers block-level storage.

- Volumes are automatically replicated within its Availability Zone.

- It can be backed up automatically to Amazon S3 through snapshots.

- Uses include –

    - Boot volumes and storage for Amazon Elastic Compute Cloud (Amazon EC2) instances

    - Data storage with a file system

    - Database hosts

    - Enterprise applications

Amazon EBS volume types

	Solid State Drives (SSD)		Hard Disk Drives (HDD)	
	General Purpose	Provisioned IOPS	Throughput-Optimized	Cold
Maximum Volume Size	16 TiB	16 TiB	16 TiB	16 TiB
Maximum IOPS/Volume	16,000	64,000	500	250
Maximum Throughput/Volume	250 MiB/s	1,000 MiB/s	500 MiB/s	250 MiB/s

Amazon EBS volume type use cases

Solid State Drives (SSD)		Hard Disk Drives (HDD)	
General Purpose	Provisioned IOPS	Throughput-Optimized	Cold
• This type is recommended for most workloads	• Critical business applications that require sustained IOPS performance, or more than 16,000 IOPS or 250 MiB/second of throughput per volume	• Streaming workloads that require consistent, fast throughput at a low price	• Throughput-oriented storage for large volumes of data that is infrequently accessed
• System boot volumes	• Large database workloads	• Big data	• Scenarios where the lowest storage cost is important
• Virtual desktops		• Data warehouses	• It cannot be a boot volume
• Low-latency interactive applications		• Log processing	
• Development and test environments		• It cannot be a boot volume	

Amazon EBS features

- Snapshots –
  - Point-in-time snapshots
  - Recreate a new volume at any time
- Encryption –
  - Encrypted Amazon EBS volumes
  - No additional cost
- Elasticity –
  - Increase capacity
  - Change to different types

Amazon EC2 storage options

- Amazon Elastic Block Store (Amazon EBS) –

    - Durable, block-level storage volumes.

    - You can stop the instance and start it again, and the data will still be there.

- Amazon EC2 Instance Store –

    - Ephemeral storage is provided on disks that are attached to the host computer where the EC2 instance is running.

    - If the instance stops, data stored here is deleted.

- Other options for storage (not for the root volume) –

    - Mount an Amazon Elastic File System (Amazon EFS) file system.

    - Connect to Amazon Simple Storage Service (Amazon S3)

Amazon EBS: Volumes, IOPS, and pricing

- Volumes –
  - Amazon EBS volumes persist independently from the instance.
  - All volume types are charged by the amount that is provisioned per month.

- IOPS –
  - General Purpose SSD:
    - Charged by the amount that you provision in GB per month until storage is released.

  - Magnetic:
    - Charged by the number of requests to the volume.
  - Provisioned IOPS SSD:
    - Charged by the amount that you provision in IOPS (multiplied by the percentage of days that you provision for the month).

Amazon EBS: Snapshots and data transfer

- Snapshots –
    - Added cost of Amazon EBS snapshots to Amazon S3 is per GB-month of data stored.

- Data transfer –
    - Inbound data transfer is free.
    - Outbound data transfer across Regions incurs charges.

Amazon Simple Storage Service (Amazon S3) - overview

- Data is stored as objects in buckets

- Virtually unlimited storage

- Single object is limited to 5 TB

- Designed for 11 9s of durability

- Granular access to bucket and objects

Amazon S3 bucket

- To upload your data:

  - Create a bucket in an AWS Region.

  - Upload almost any number of objects to the bucket
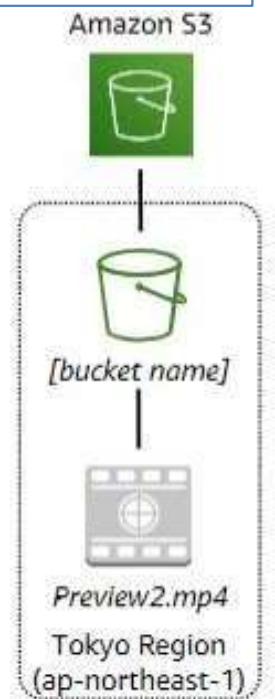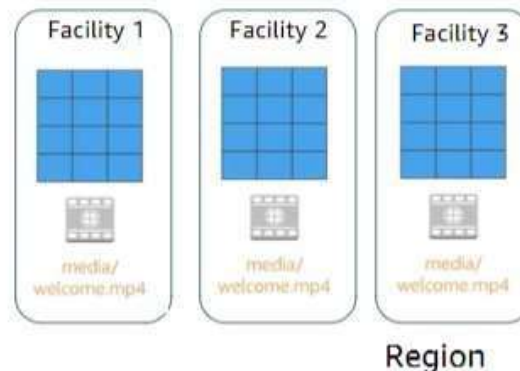
Bucket path-style URL endpoint:

https://s3.ap-northeast-1.amazonaws.com/bucket-name

Region code   Bucket name

Bucket virtual hosted-style URL endpoint:

https:// bucket-name.s3-ap-northeast-1.amazonaws.com

Bucket name   Region code

Amazon S3

- Data is redundantly stored in the Region

media/welcome.mp4

my-bucket-name

Facility 1   Facility 2   Facility 3

media/ welcome.mp4   media/ welcome.mp4   media/ welcome.mp4

Region

[bucket name]

Preview2.mp4

Tokyo Region
(ap-northeast-1)

Amazon S3 bucket

- Access the data anywhere

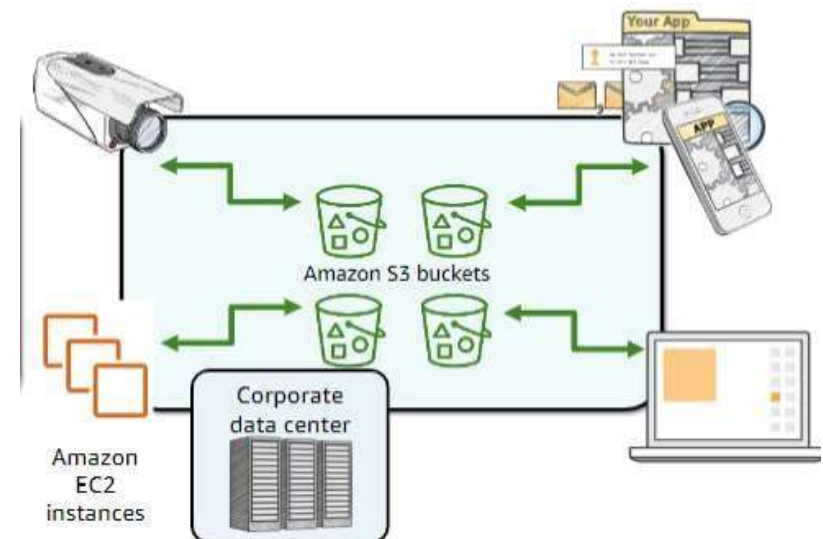AWS Management Console

AWS Command Line Interface

SDK

- Common use cases
  - Storing application assets
  - Static web hosting
  - Backup and disaster recovery(DR)
  - Staging area for bigdata
  - Many more..

Your App

Amazon S3 buckets

Corporate data center

Amazon EC2 instances

Amazon S3 pricing

- Pay only for what you use, including–

  - GBs per month

  - Transfer OUT to other Regions

  - PUT, COPY, POST, LIST, and GET requests

- You do not pay for –

  - Transfers IN to Amazon S3

  - Transfers OUT from Amazon S3 to Amazon CloudFront or Amazon EC2 in the same Region

Amazon Elastic File System (Amazon EFS)

- Amazon EFS features
  - File storage in the AWS Cloud
  - Works well for big data and analytics, media processing workflows, content management, web serving, and home directories
  - Petabyte-scale, low-latency file system
  - Shared storage
  - Elastic capacity
  - Supports Network File System (NFS) versions 4.0 and 4.1 (NFSv4)
  - Compatible with all Linux-based AMIs for Amazon EC2

Amazon S3 Glacier

- Amazon S3 Glacier is a data archiving service that is designed for security, durability, and an extremely low cost.

- Amazon S3 Glacier is designed to provide 11 9s of durability for objects.

- It supports the encryption of data in transit and at rest through Secure Sockets Layer (SSL) or Transport Layer Security (TLS).

- The Vault Lock feature enforces compliance through a policy.

- Extremely low-cost design works well for long-term archiving.

- Provides three options for access to archives—expedited, standard, and bulk—retrieval times range from a few minutes to several hours.

Amazon S3 Glacier use cases

- Media asset archiving
- Healthcare information archiving
- Regulatory and compliance archiving
- Scientific data archiving
- Digital preservation
- Magnetic tape replacement

Amazon S3 bucket → Archive after 30 days → Amazon S3 Glacier Archive → Delete after 5 years → (trash)

IaaS for you

*Thanks, I feel so "**Clouded**" now*