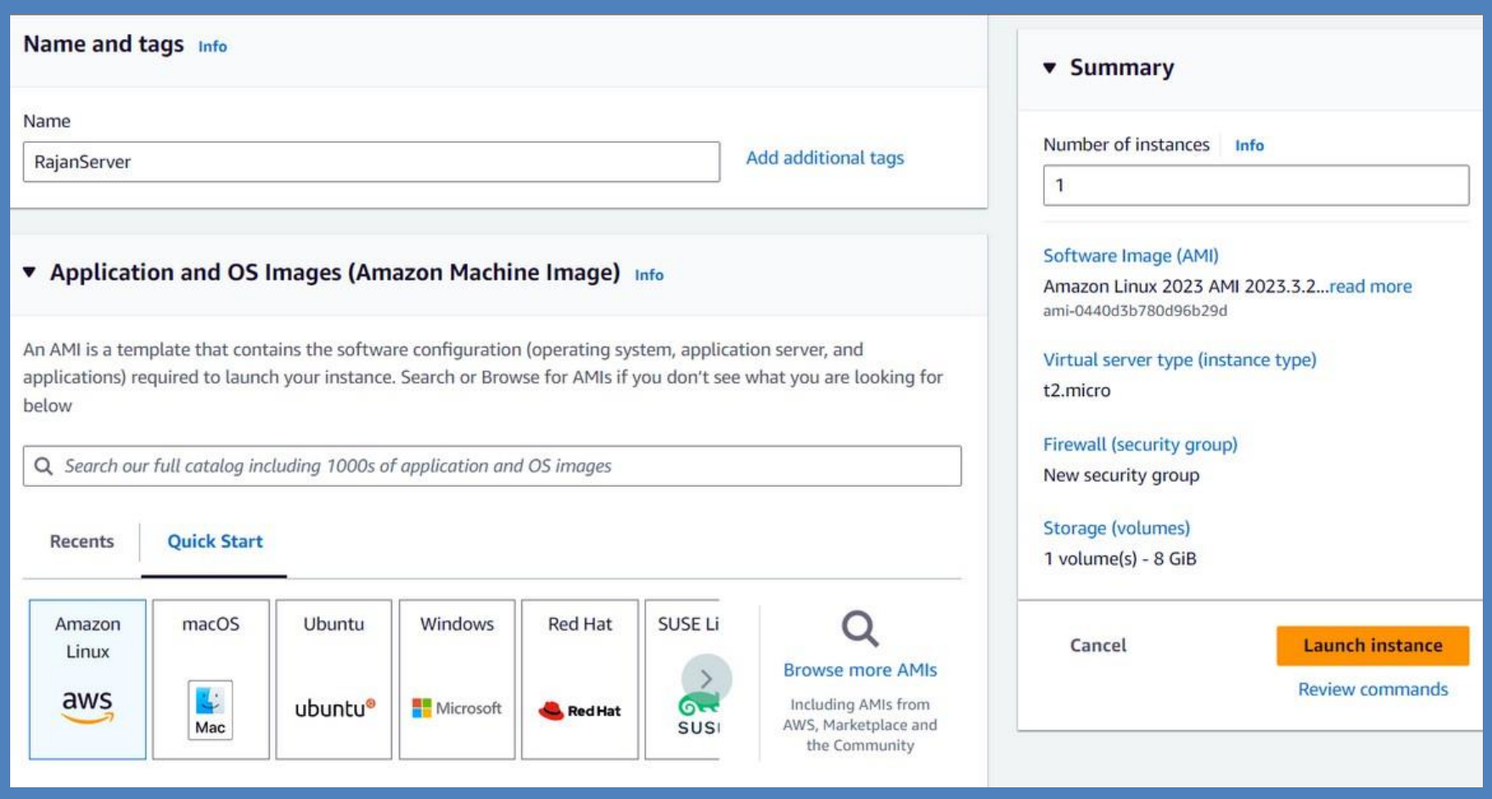
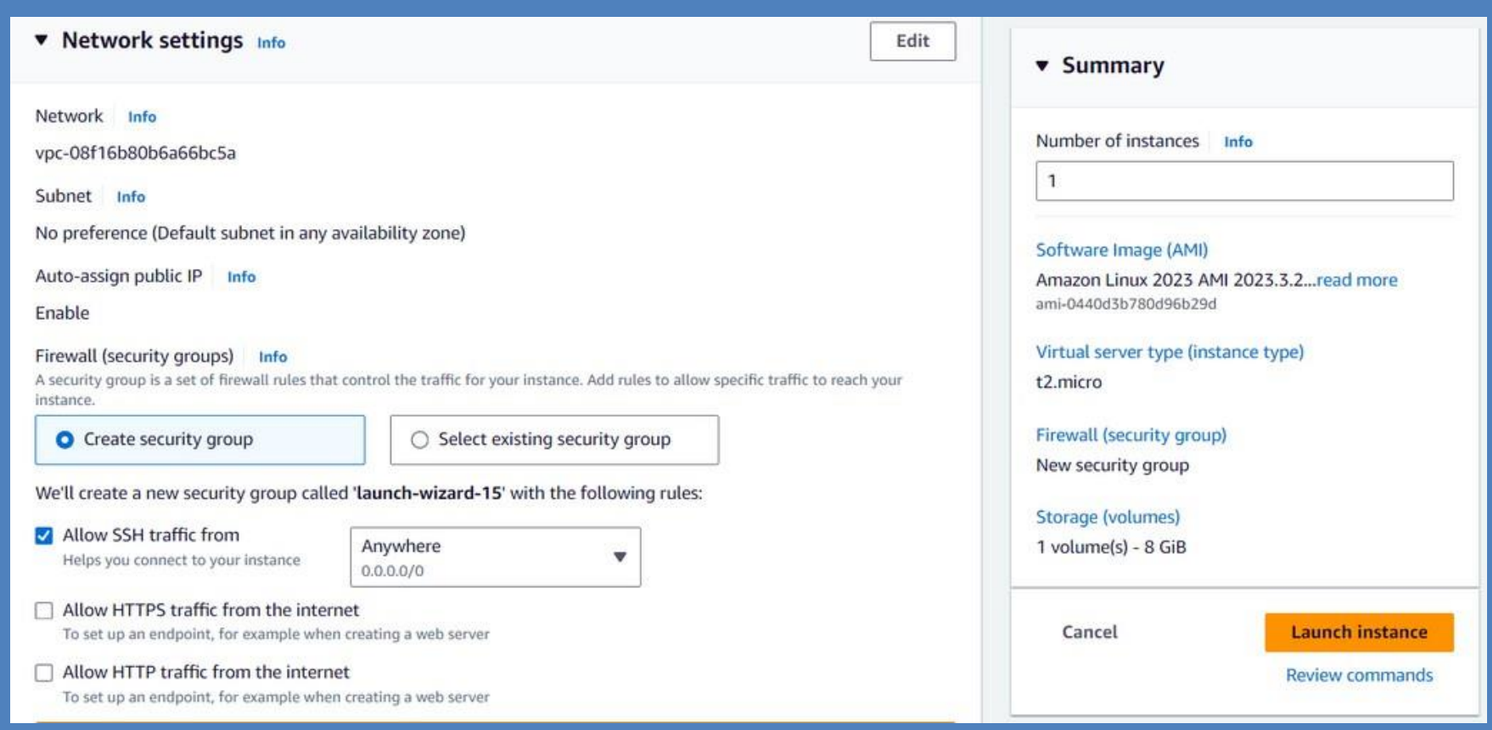
AWS EC2 Full Guide

Amazon Elastic Compute Cloud

1. Getting Started: Amazon EC2

EC2 sizing & configuration options





When configuring EC2 instances, the choice of Operating System is flexible, supporting Linux, Windows, or Mac OS.

For Computing Resources, users can tailor the amount of compute power and CPU cores to meet their specific workload demands.

Random-access memory (RAM) is customizable, allowing users to allocate the appropriate memory for optimal performance.

Storage options include Elastic Block Store (EBS) and Elastic File System (EFS) for network-attached storage, as well as EC2 Instance Store for hardware-based storage.

Network configurations involve specifying the network card speed and determining the need for a Public IP address.

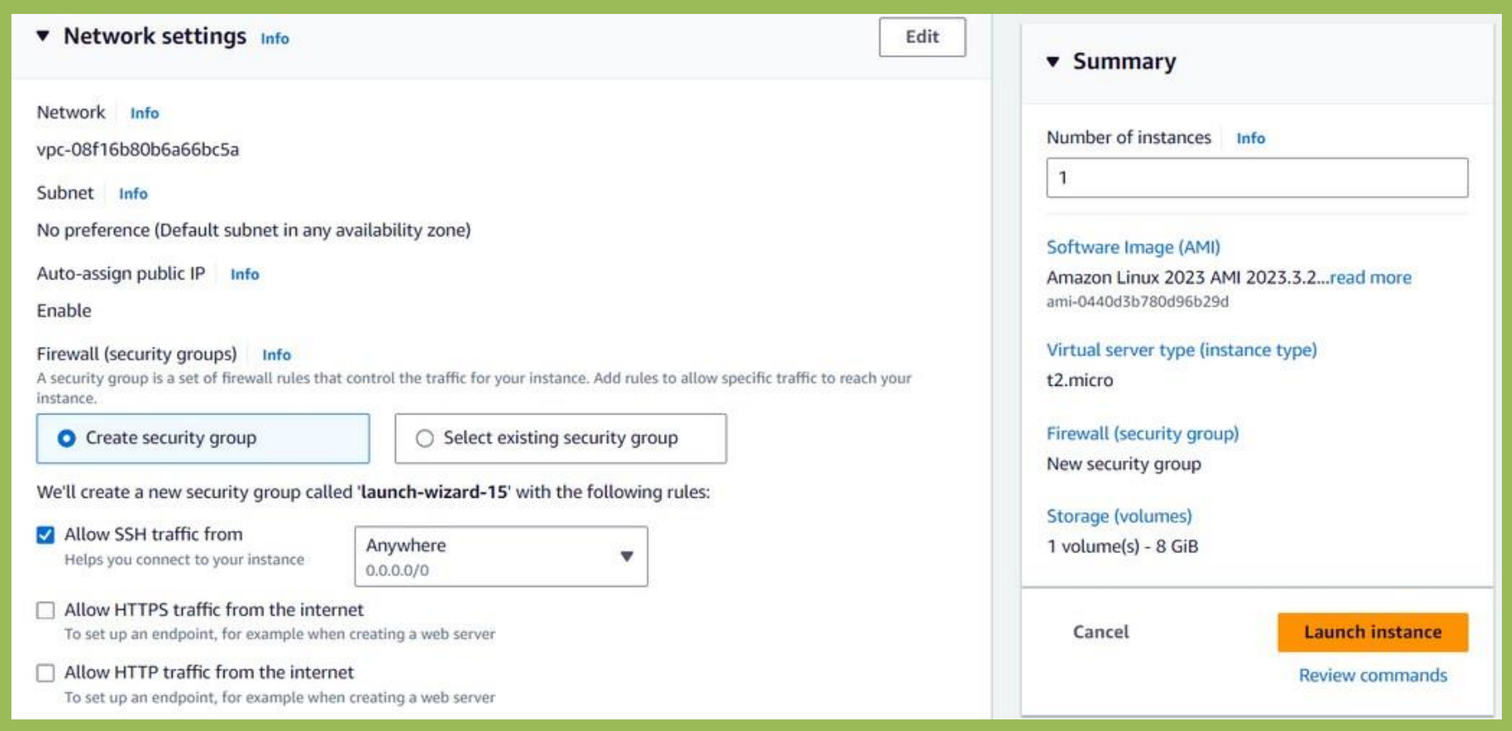
Security is managed through security groups, enabling users to define firewall rules for effective protection of EC2 instances.

What is Aamzon EC2

EC2, part of AWS, is essential for understanding the cloud. It offers virtual machines, storage (EBS), load balancing (ELB), and scalability through auto-scaling groups (ASG). A cornerstone service, EC2 enables flexible and scalable application deployment on the AWS platform.

2. Getting Started: Amazon EC2

EC2 User Data



Overview

It is possible to bootstrap our instances using an EC2 User Data Script.

Bootstrapping means launching commands when a machine starts.

That script is only run once at the instance first start.

Use Cases

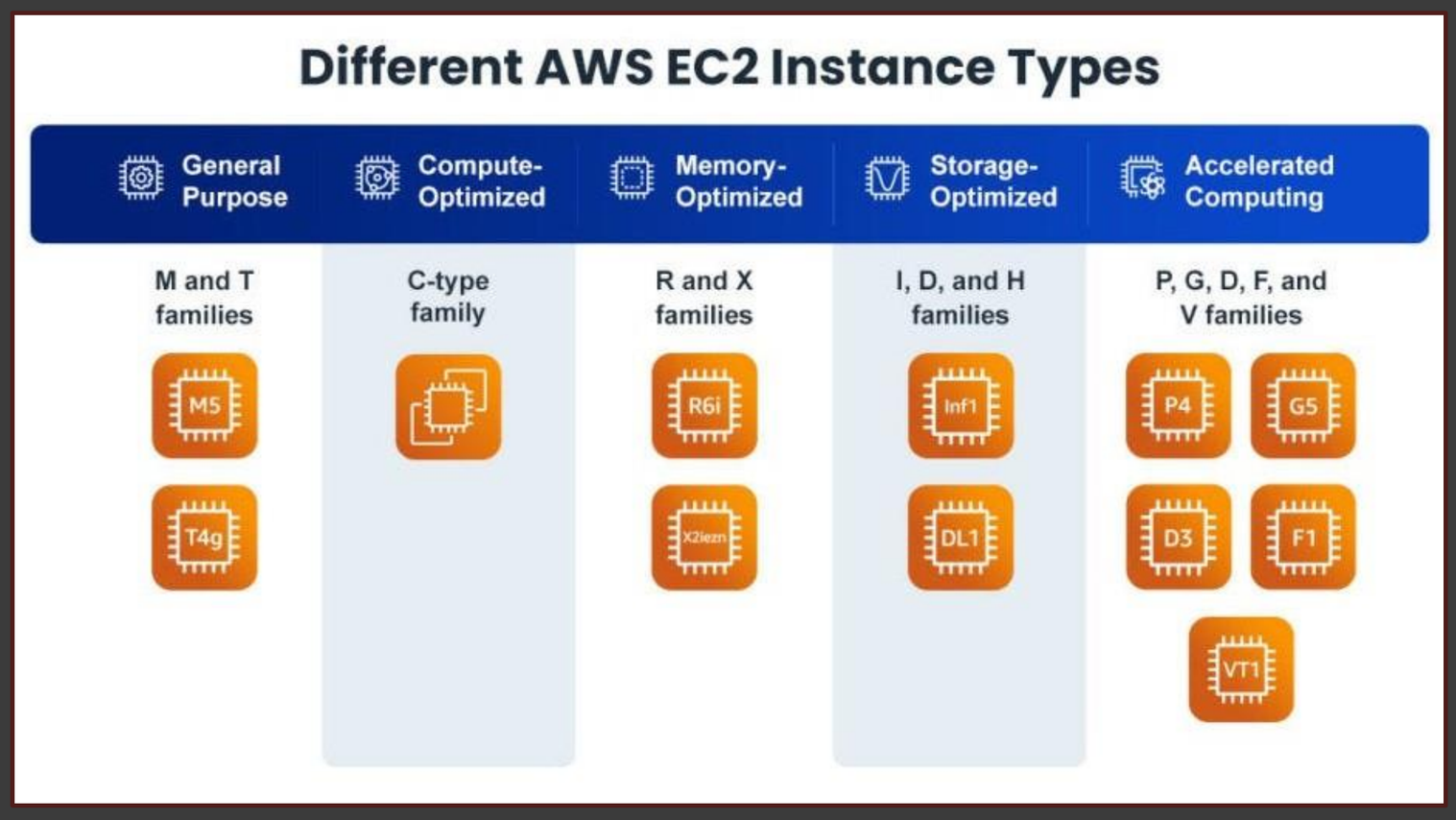
EC2 user data is used to automate boot tasks such as:

- Installing updates
- Installing software
- Downloading common files
- Anything you can think of

The EC2 User Data Script runs with the root user

3. Amazon EC2 Instance Types

EC2 Instance Types Overview



Overview

You can use different types of EC2 instances that are optimized for different use cases

Find More details here:- https://aws.amazon.com/ec2/instance-types/

Naming Convention

AWS has the following naming convention:

m5.4xlarge

m: instance class
5: generation (AWS improves them over time)

4xlarge: size within the instance class

4. Amazon EC2 Instance Types

Understanding Instance types use-cases

General Purpose

General purpose instances provide a balance of compute, memory and networking resources, and can be used for a variety of diverse workloads. These instances are ideal for applications that use these resources in equal proportions such as web servers and code repositories.

M7g | M7i | M7i-flex | M7a | Mac | M6g | M6i | M6in | M6a | M5 | M5n | M5zn | M5a | M4 | T4g | T3 | T3a | T2

Great for a diversity of workloads such as web servers or code repositories
Balance between:
- Compute
- Memory
- Networking

Compute Optimized

Compute Optimized instances are ideal for compute bound applications that benefit from high performance processors. Instances belonging to this category are well suited for batch processing workloads, media transcoding, high performance web servers, high performance computing (HPC), scientific modeling, dedicated gaming servers and ad server engines, machine learning inference and other compute intensive applications.

C7g | C7gn | C7i | C7a | C6g | C6gn | C6i | C6in | C6a | C5 | C5n | C5a | C4

Great for compute-intensive tasks that require high performance processors:
- High performance web servers
- Batch processing workloads
- Scientific modeling & machine learning

Memory Optimized

Memory optimized instances are designed to deliver fast performance for workloads that process large data sets in memory.

R8g | R7g | R7i | R7iz | R7a | R6g | R6i | R6in | R6a | R5 | R5n | R5b | R5a | R4 | X2gd | X2idn | X2iedn | X2iezn | X1 | X1e | High Memory | z1d

Fast performance for workloads that process large data sets in memory.
Use-case:
- High performance, relational/non-relational databases
- Distributed web scale cache stores

Storage Optimized

Storage optimized instances are designed for workloads that require high, sequential read and write access to very large data sets on local storage. They are optimized to deliver tens of thousands of low-latency, random I/O operations per second (IOPS) to applications.

I4g | Im4gn | Is4gen | I4i | I3 | I3en | D2 | D3 | D3en | H1

Great for storage-intensive tasks that require high, sequential read and write access to large data sets on local storage.
Use-Case:
- High frequency online transaction processing (OLTP) systems
- Relational & NoSQL databases

EC2 Instance Types: example

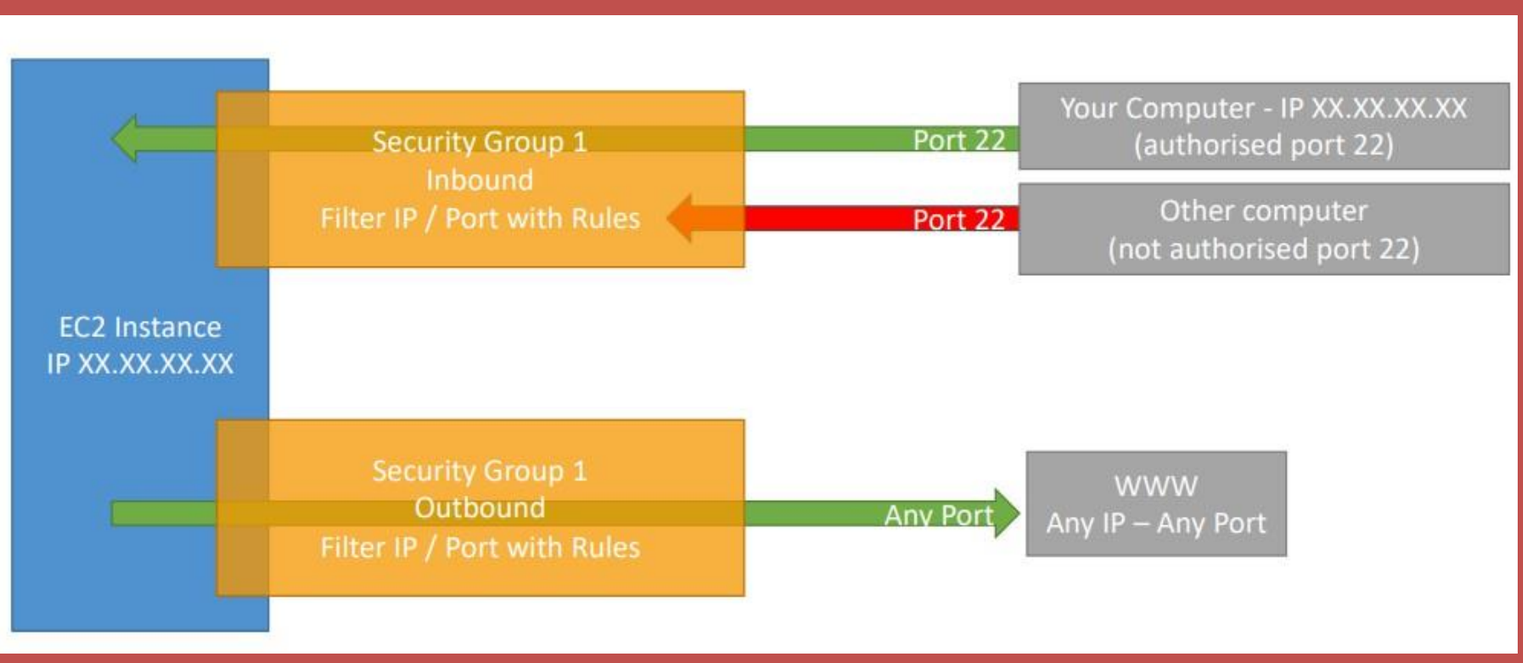
Instance	vCPU	Mem (GiB)	Storage	Network Performance	EBS Bandwidth (Mbps)
t2.micro	1	1	EBS-Only	Low to Moderate	
t2.xlarge	4	16	EBS-Only	Moderate	
c5d.4xlarge	16	32	1 x 400 NVMe SSD	Up to 10 Gbps	4,750
r5.16xlarge	64	512	EBS Only	20 Gbps	13,600
m5.8xlarge	32	128	EBS Only	10 Gbps	6,800

t2.micro is part of the AWS free tier (up to 750 hours per month)

5. Amazon EC2: Security Groups

Security Groups Deeper Dive

Type	Protocol	Port Range	Source	Description
HTTP	TCP	80	0.0.0.0/0	test http page
SSH	TCP	22	122.149.196.85/32	
Custom TCP Rule	TCP	4567	0.0.0.0/0	java app



What are Security Groups

AWS Security Groups are core to network security, managing traffic to and from EC2 instances. They consist of rules that can reference either IP addresses or other Security Groups for flexible access control.

Security Groups are acting as a firewall on EC2 instances.
They regulate:

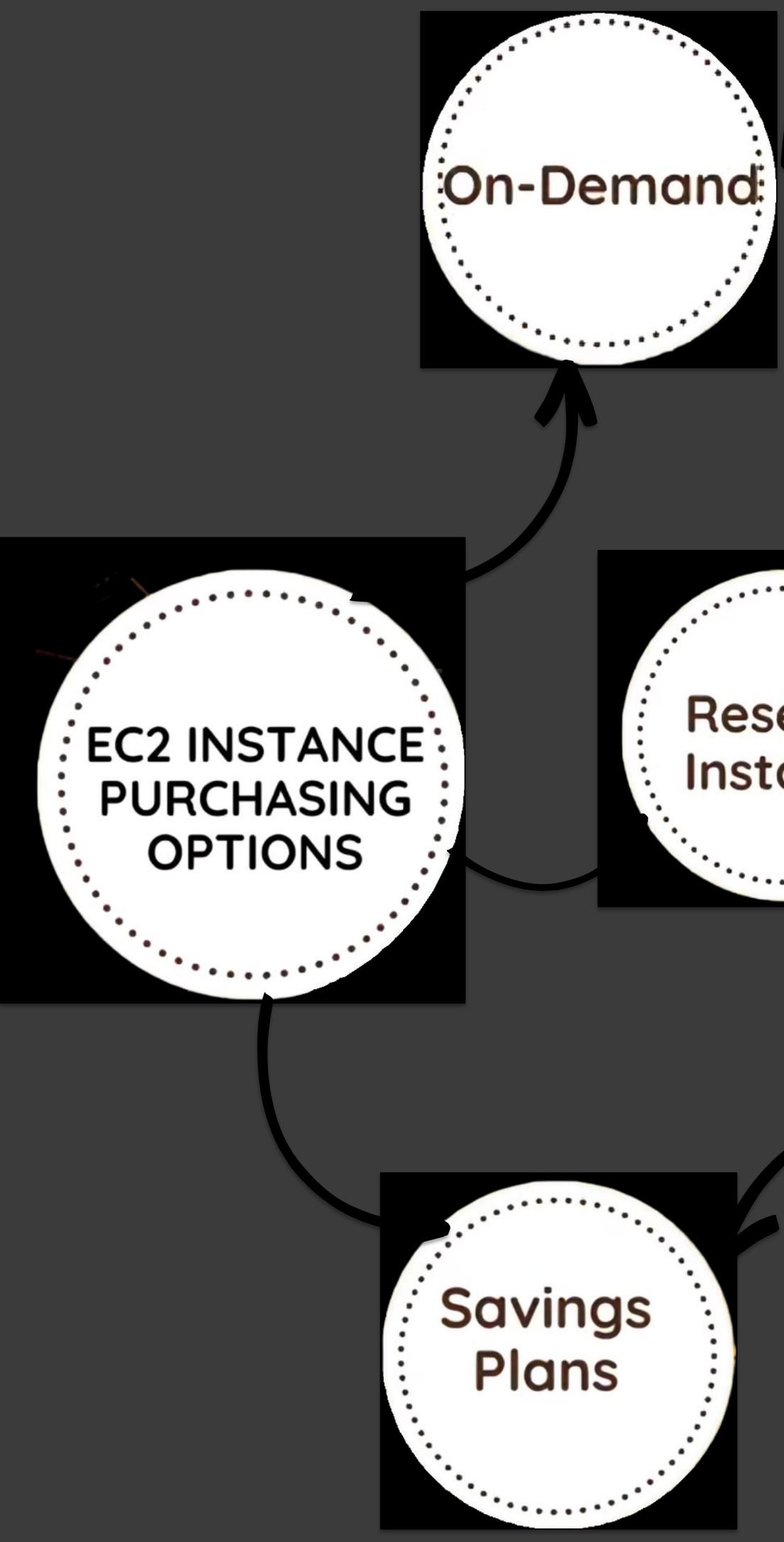
- Access to Ports
- Authorised IP ranges
  IPv4 and IPv6
- Control of inbound network
  (from other to the instance)
- Control of outbound network
  (from the instance to other)

Security Groups Good to know

- Can be attached to multiple instances
- Locked down to a region / VPC combination
- Does live "outside" the EC2 – if traffic is blocked the EC2 instance won't see it
- It's good to maintain one separate security group for SSH access
- If your application is not accessible (time out), then it's a security group issue
- If your application gives a "connection refused" error, then it's an application error or it's not launched
- All inbound traffic is blocked by default
- All outbound traffic is authorised by default

6. Amazon EC2: Purchasing Options

Exploring Different Purchasing Options



EC2 On Demand

You Pay only for what you use:

- Linux or Windows - billing per second, after the first minute
- All other operating systems - billing per hour

Costy than other option but no upfront payment is needed.

Recommended for short-term and un-interrupted workloads, where you can't predict how the application will behave.

EC2 Reserved Instances

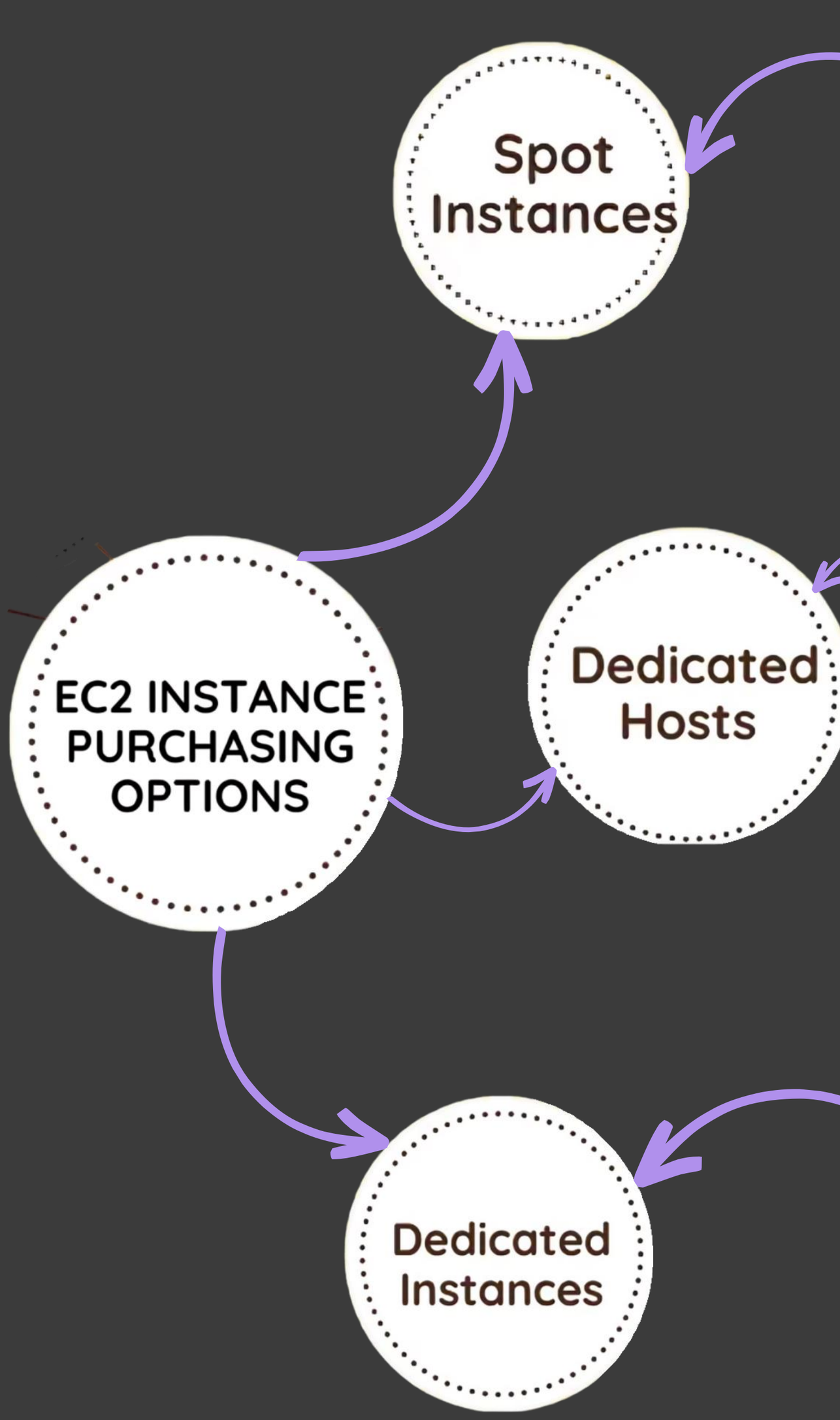
- Up to 72% discount compared to On-demand
- You reserve a specific instance attributes (Instance Type, Region, Tenancy, OS)
- Reservation Period – 1 year or 3 years
- Recommended for steady-state usage applications (think database)
- You can buy and sell in the Reserved Instance Marketplace

EC2 Savings Plans

- Get a discount based on long-term usage (up to 72% - same as RIs)
- Commit to a certain type of usage ($8/hour for 1 or 3 years)
- Usage beyond EC2 Savings Plans is billed at the On-Demand price
- Locked to a specific instance family & AWS region
- Flexible across:
  - Instance Size (e.g., m5.xlarge, m5.2xlarge)
  - OS (e.g., Linux, Windows)
  - Tenancy (Host, Dedicated, Default)

7. Amazon EC2: Purchasing Options

Exploring Different Purchasing Options



EC2 Spot Instances

- Can get a discount of up to 90% compared to On-demand
- Instances that you can "lose" at any point of time if your max price is less than the current spot price
- The MOST cost-efficient instances in AWS
- Useful for workloads that are resilient to failure
- Not suitable for critical jobs or databases

EC2 Dedicated Hosts

- A physical server with EC2 instance capacity fully dedicated to your use
- Allows you address compliance requirements

- Purchasing Options:
  - On-demand – pay per second for active Dedicated Host
  - Reserved - 1 or 3 years (No Upfront, Partial Upfront, All Upfront)

- The most expensive option

EC2 Dedicated Instances

- Instances run on hardware that's dedicated to you
- May share hardware with other instances in same account
- No control over instance placement (can move hardware after Stop / Start)

8. Price Comparision

Example – m4.large – us-east-1

Price Type	Price (per hour)
On-Demand	$0.10
Spot Instance (Spot Price)	$0.038 - $0.039 (up to 61% off)
Reserved Instance (1 year)	$0.062 (No Upfront) - $0.058 (All Upfront)
Reserved Instance (3 years)	$0.043 (No Upfront) - $0.037 (All Upfront)
EC2 Savings Plan (1 year)	$0.062 (No Upfront) - $0.058 (All Upfront)
Reserved **Convertible** Instance (1 year)	$0.071 (No Upfront) - $0.066 (All Upfront)
Dedicated Host	On-Demand Price
Dedicated Host Reservation	Up to 70% off
Capacity Reservations	On-Demand Price

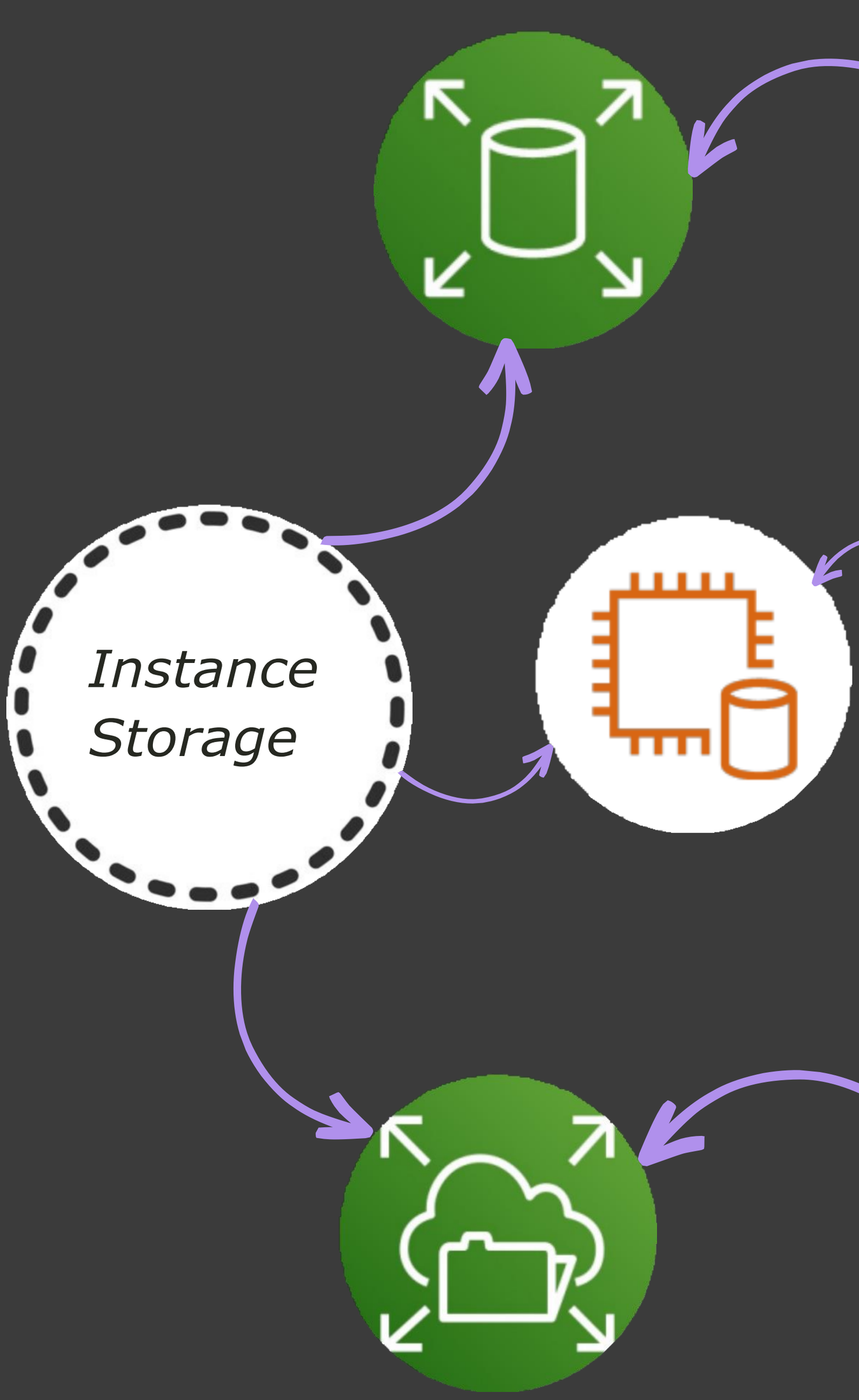
Selecting a purchasing option

Choosing AWS purchasing options involves understanding key models. On-Demand Instances provide flexibility, ideal for variable workloads. Reserved Instances demand upfront payment for stable workloads with cost savings. Spot Instances offer potential cost reduction through bidding, suitable for flexible and fault-tolerant workloads.

Tailor your choice based on workload characteristics, budget, and scalability needs for an efficient and cost-effective AWS deployment.

9. Amazon EC2: Instance Storage

Exploring Different Instance Storage



EBS Volume

- An EBS (Elastic Block Store) Volume is a network drive you can attach to your instances while they run
- It allows your instances to persist data, even after their termination
- They can only be mounted to one instance at a time

EC2 Instance Store

- High-performance hardware disk
- Better I/O performance
- EC2 Instance Store lose their storage if they’re stopped
- Good for buffer / cache / scratch data / temporary content
- Backups and Replication are your responsibility

EC2 Elastic File System

- Managed NFS (network file system) that can be mounted on many EC2
- EFS works with EC2 instances in multi-AZ
- Highly available, scalable, expensive (3x gp2), pay per use
- Use cases: content management, web serving, data sharing, Wordpress
- File system scales automatically, pay-per-use, no capacity planning!

Thanks for Reading

Follow me for more Guides