

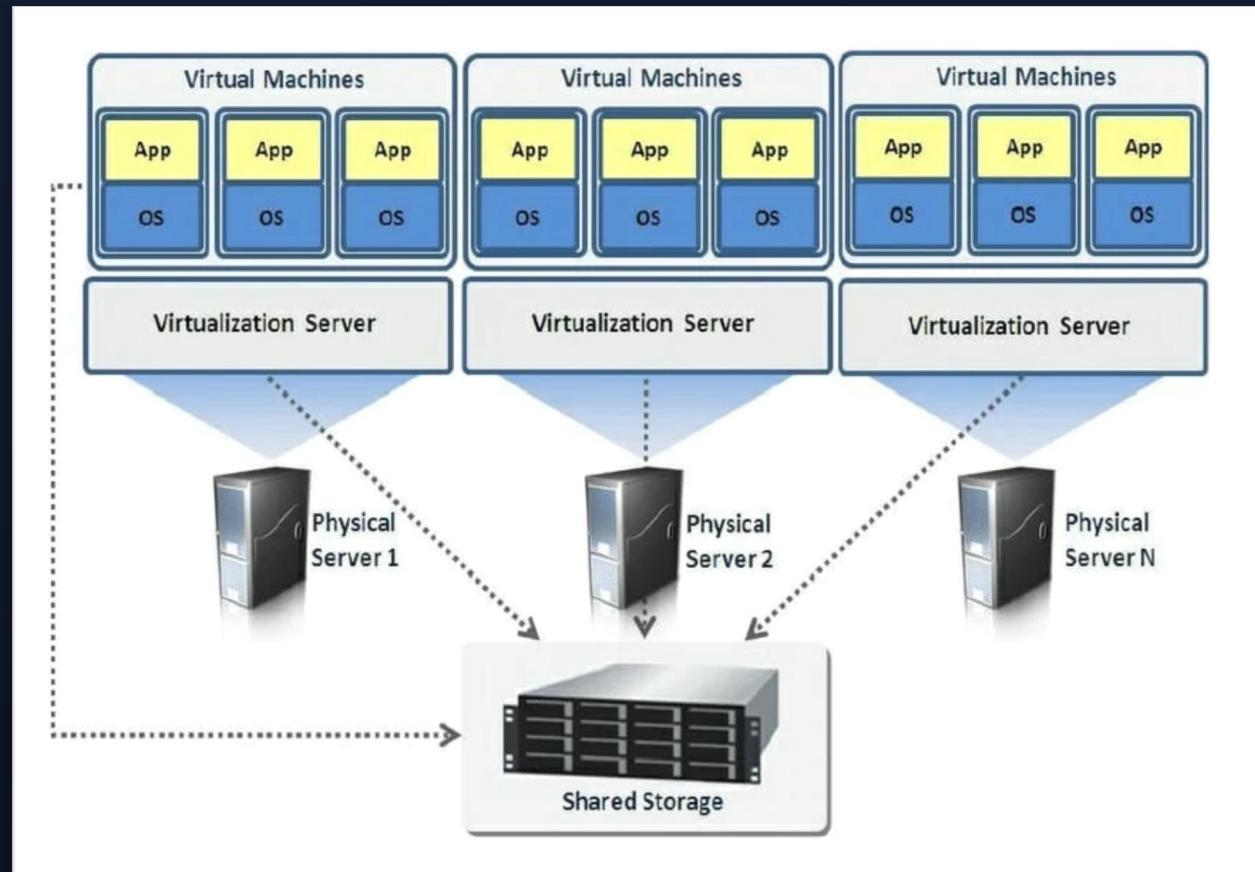
Agenda

- ① Understanding Virtual Machine Concepts
- ② Compute services on Aws
- ③ Aws offering of VMs (Ec2)
- ④ Types of Instances and the Components
- ⑤ Processors available
- ⑥ Purchasing Options
- ⑦ Security groups and NACL

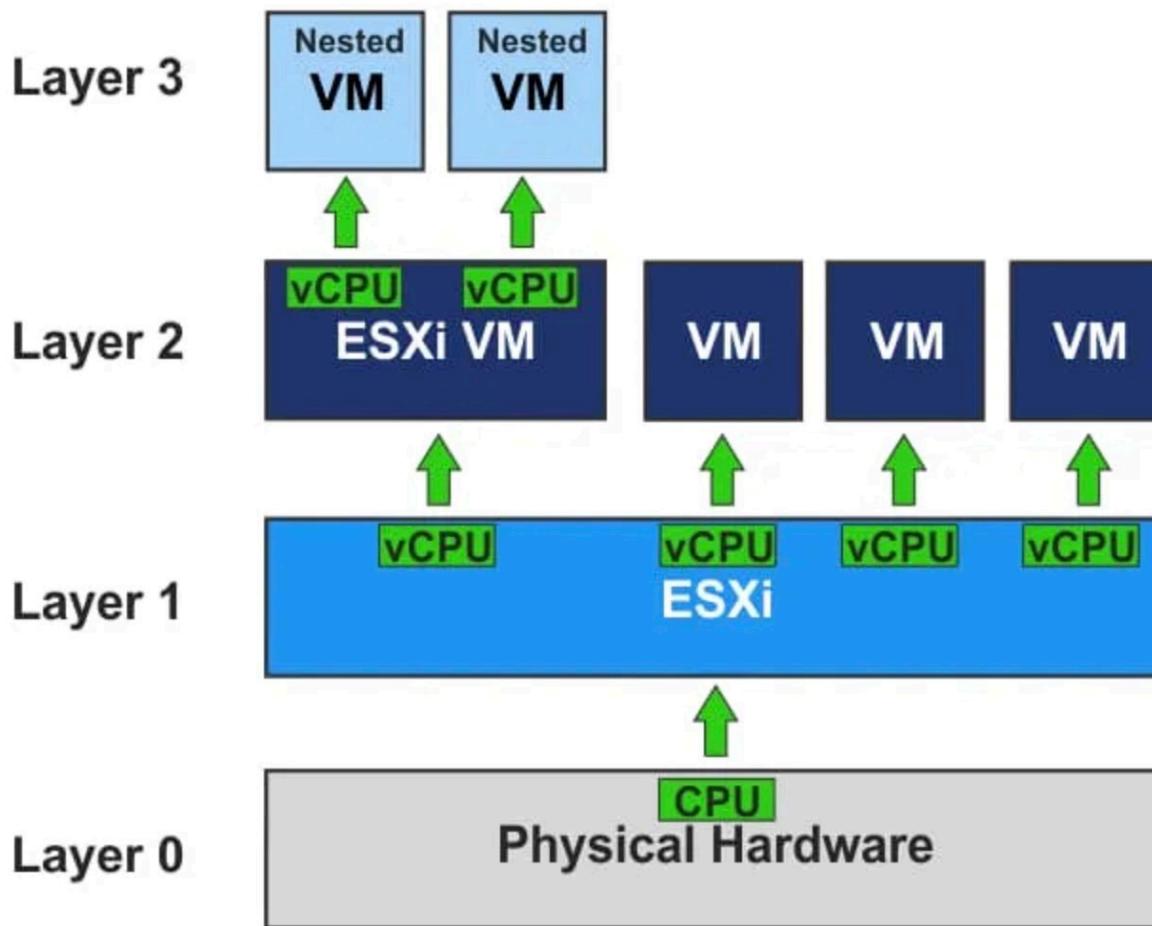
Quick Recap of S3 and Identity Centre

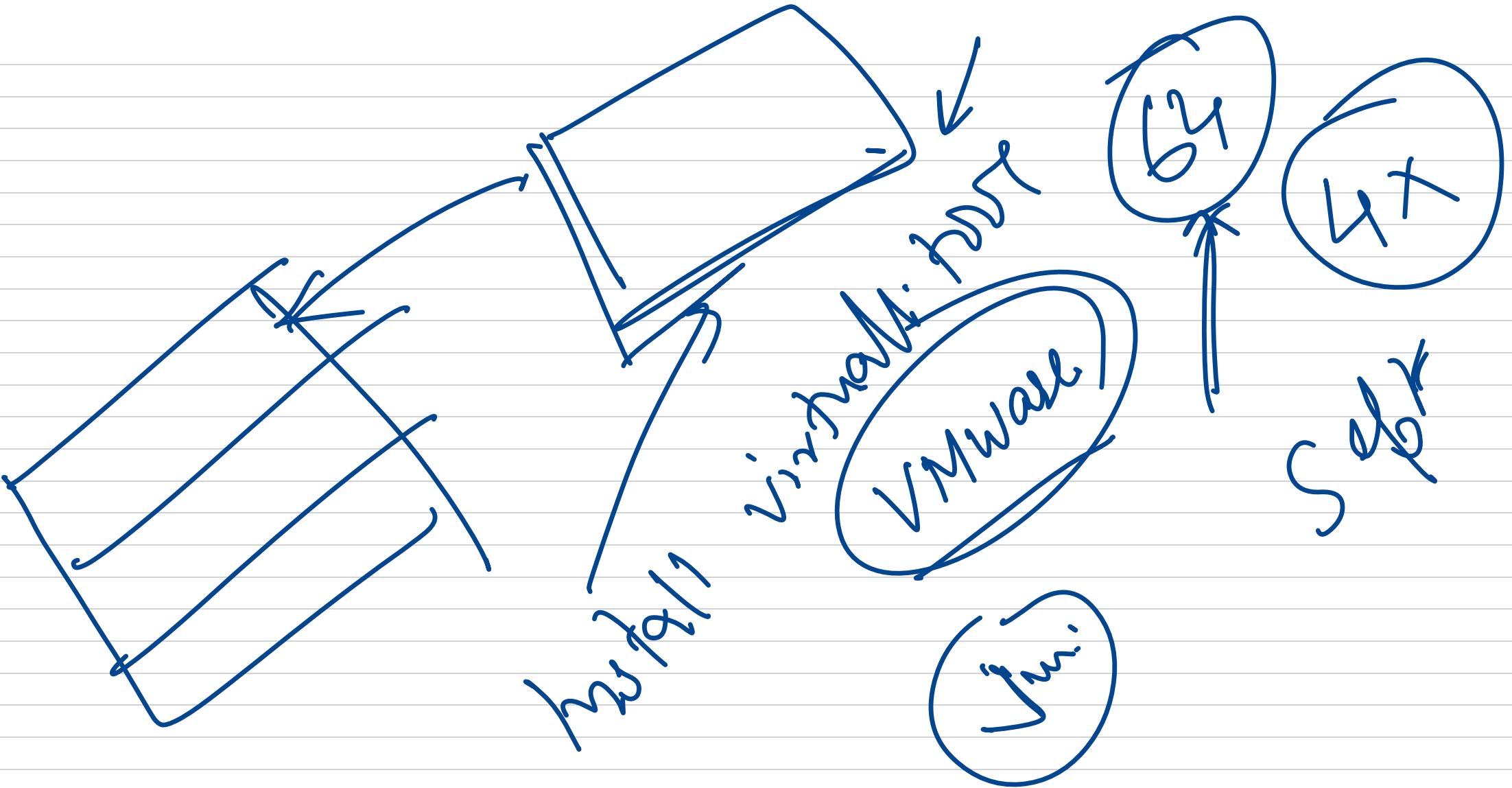
Virtual Machines

A virtual machine (VM) is essentially a **software-based computer** that runs on a physical host machine. It uses a program called a **hypervisor** to share the host's physical resources—like CPU, memory, and storage—allowing it to function as a self-contained, independent computer with its own operating system and applications. This creates an **isolated environment**, which is highly efficient for running multiple operating systems on a single piece of hardware, testing software safely, and improving security, as an issue within one VM doesn't affect the host or other VMs.



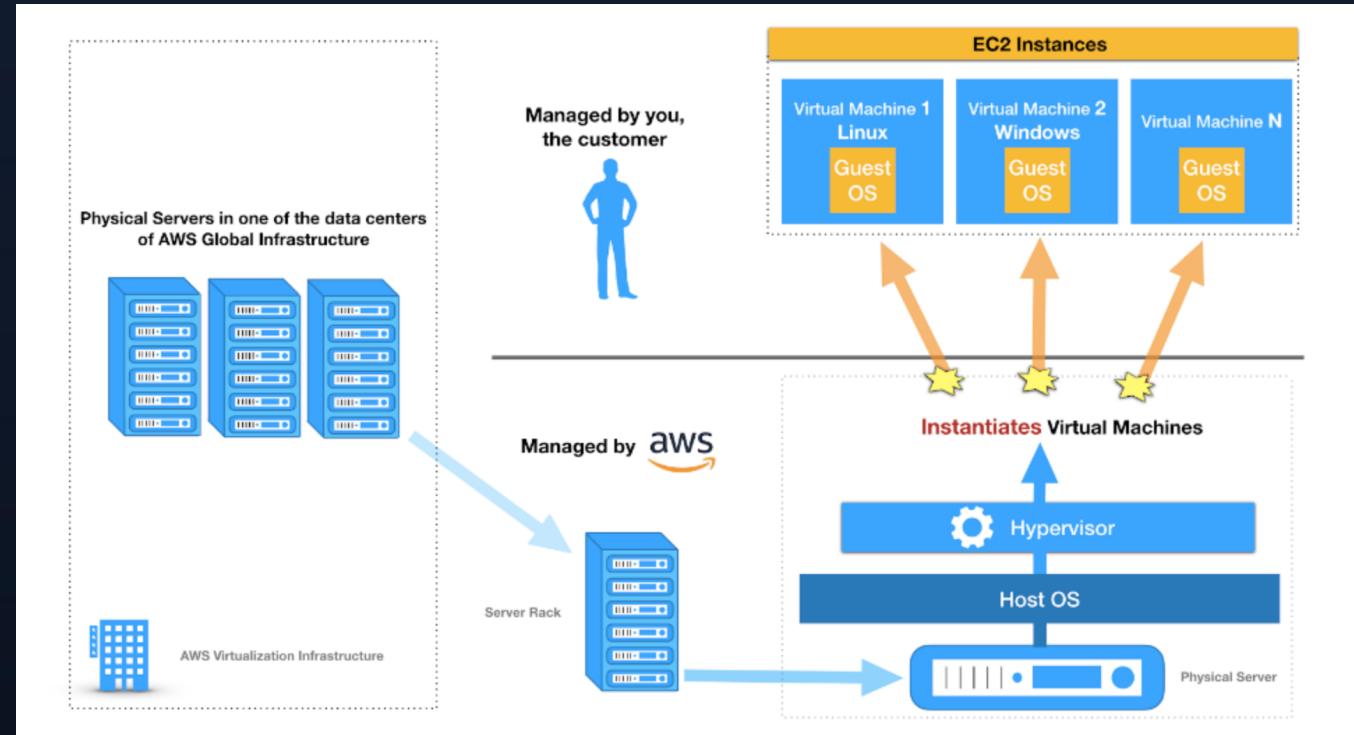
VMware Nested Virtualization (ESXi)



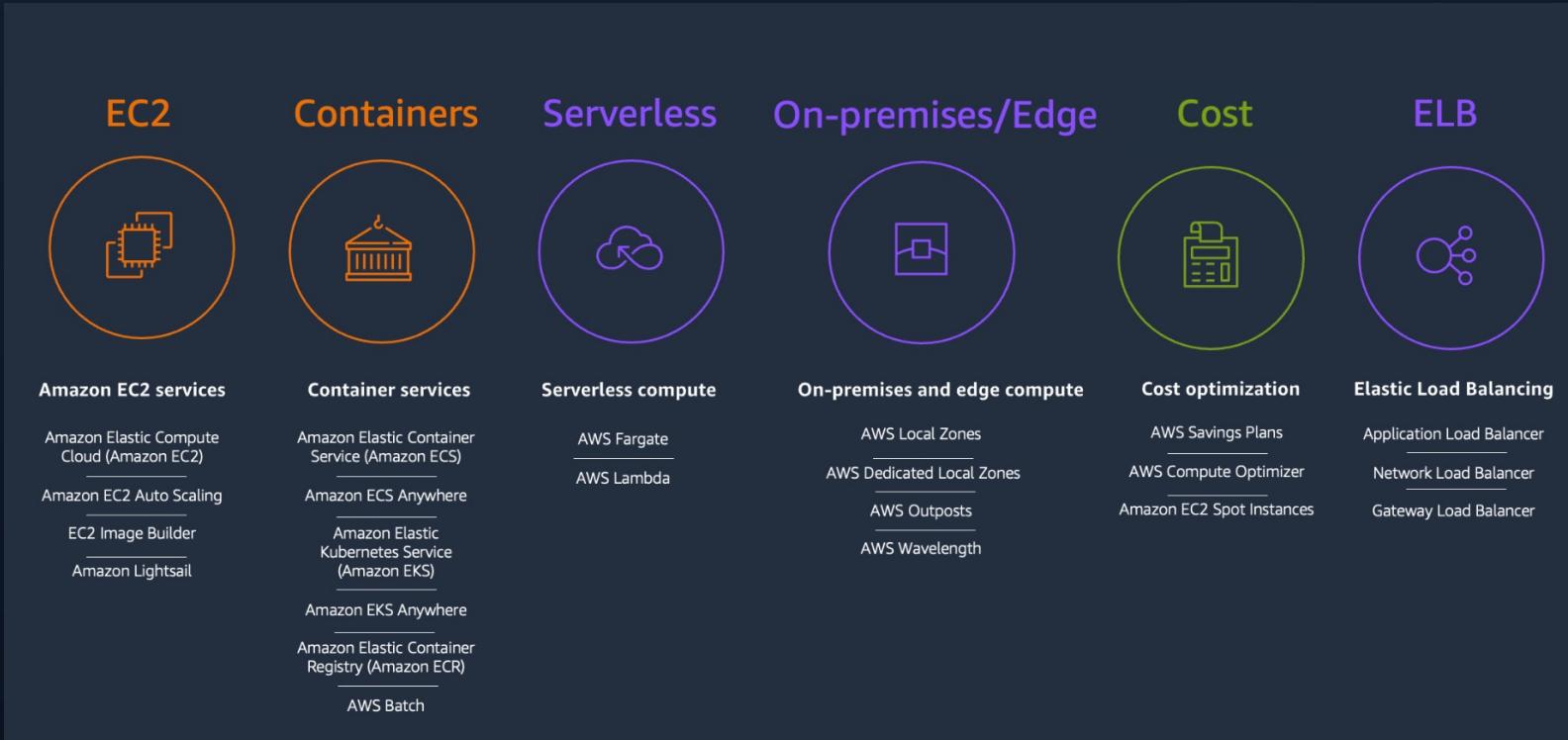


VMs in AWS

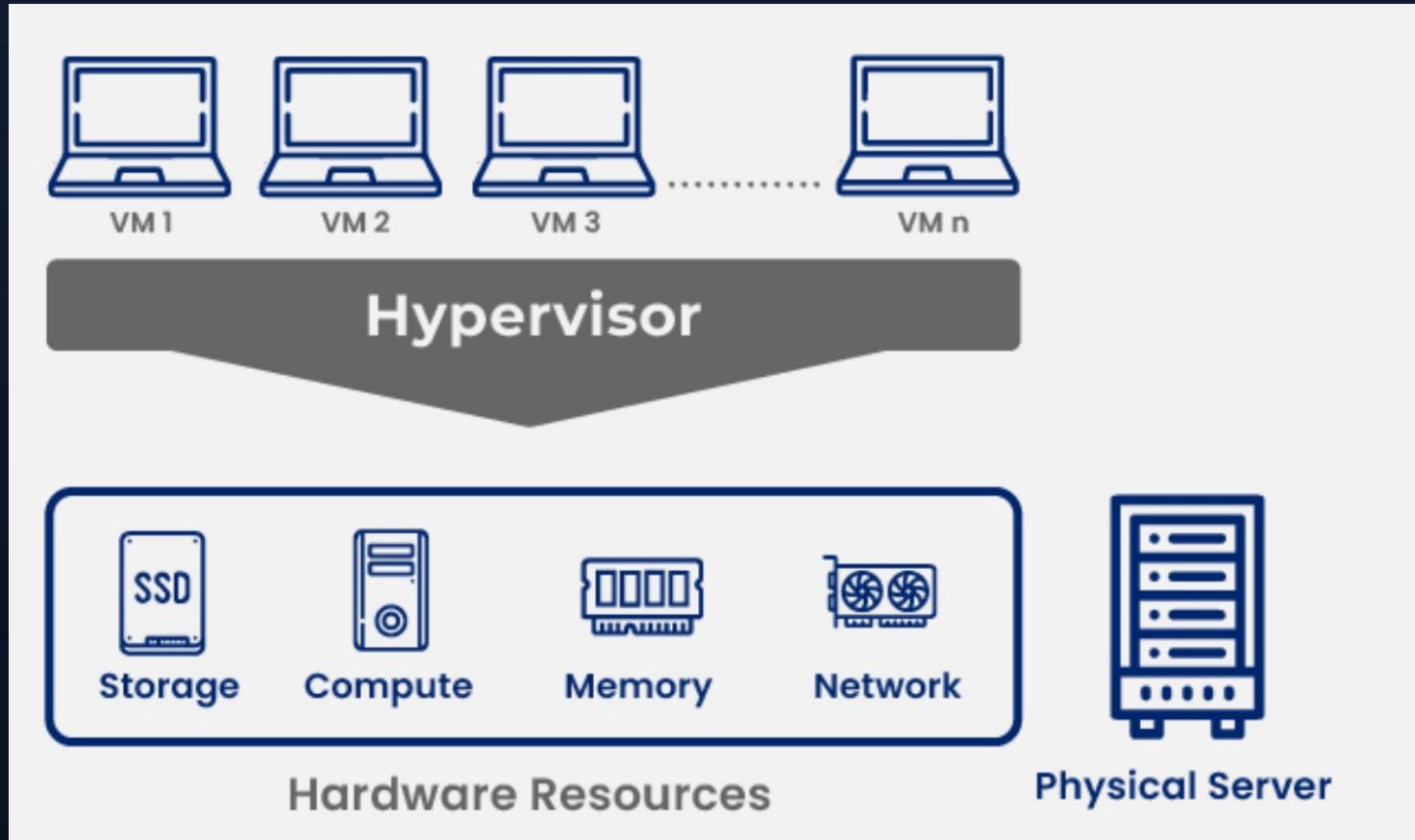
AWS achieves virtualization to host its virtual machines (VMs), known as EC2 instances, primarily through its advanced **Nitro System**. This system isn't just a single hypervisor; it's a combination of dedicated hardware and a lightweight hypervisor that offloads many traditional virtualization functions, like networking and storage management, onto specialized hardware cards. By doing this, AWS minimizes the performance overhead, or "virtualization tax," allowing the host server's CPU and memory to be almost entirely dedicated to the customer's VM. This approach provides performance that is nearly indistinguishable from bare metal, while still offering the security, isolation, and flexibility of traditional virtualization.



Compute Services in AWS

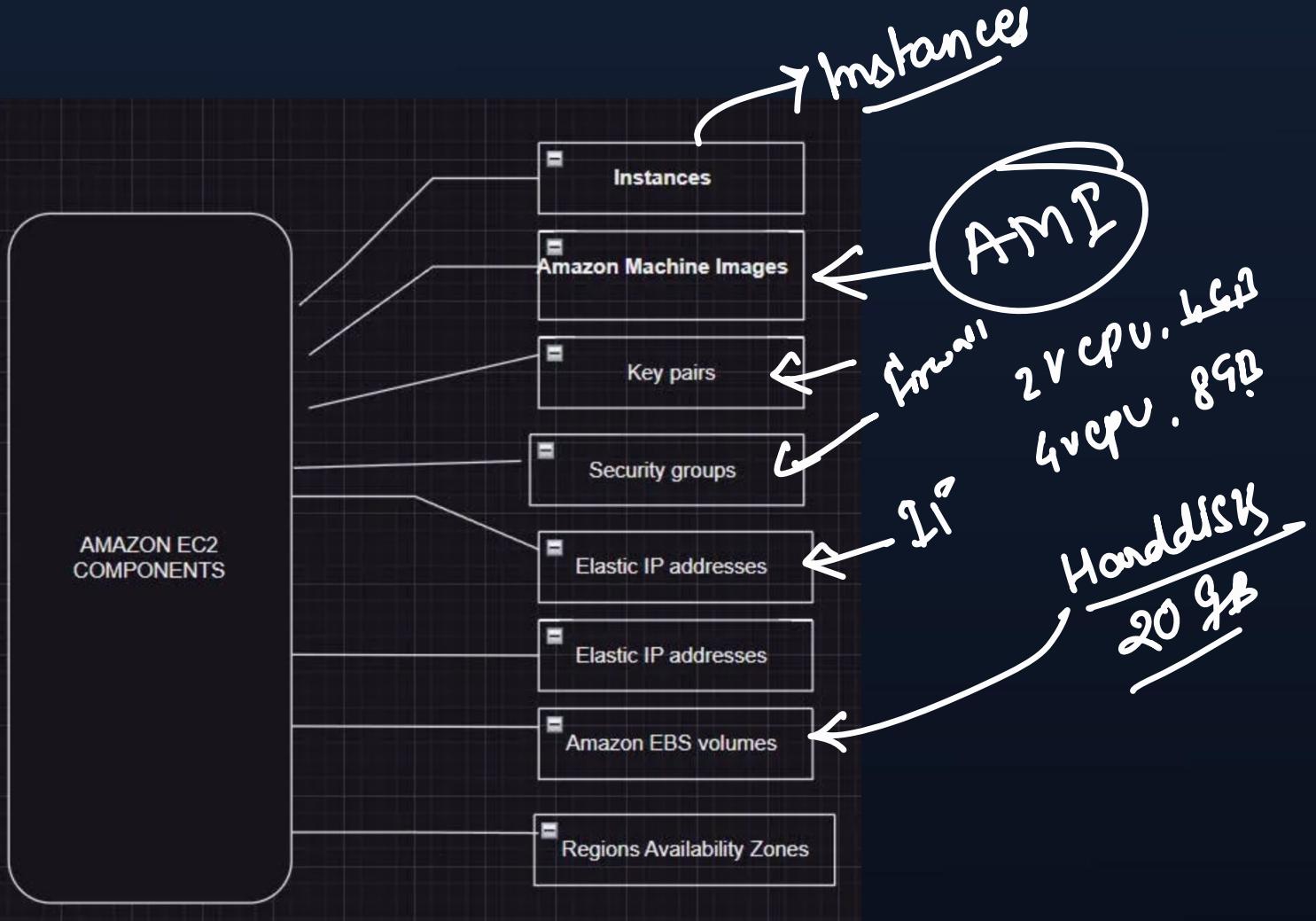


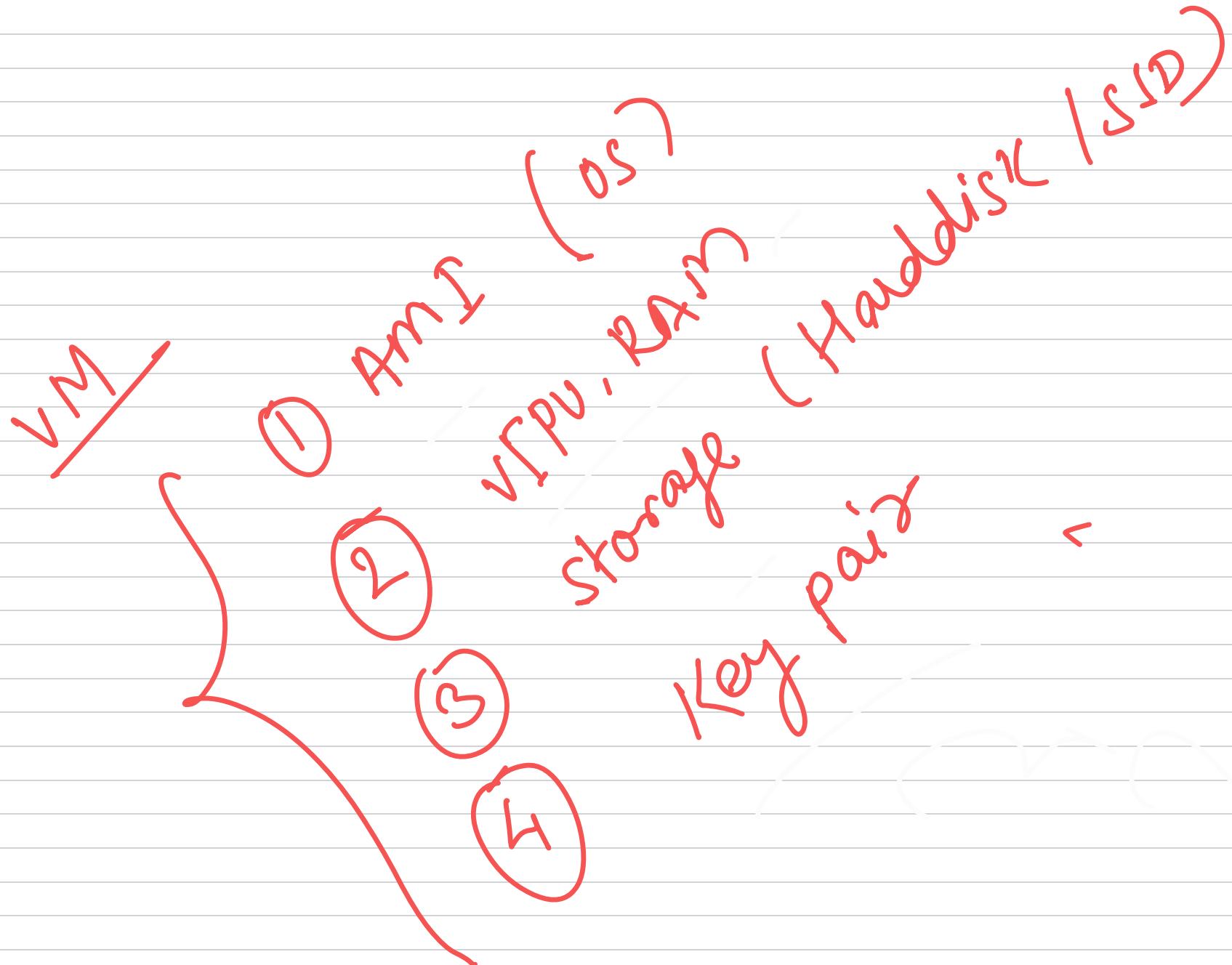
Components we need to Launch a Server



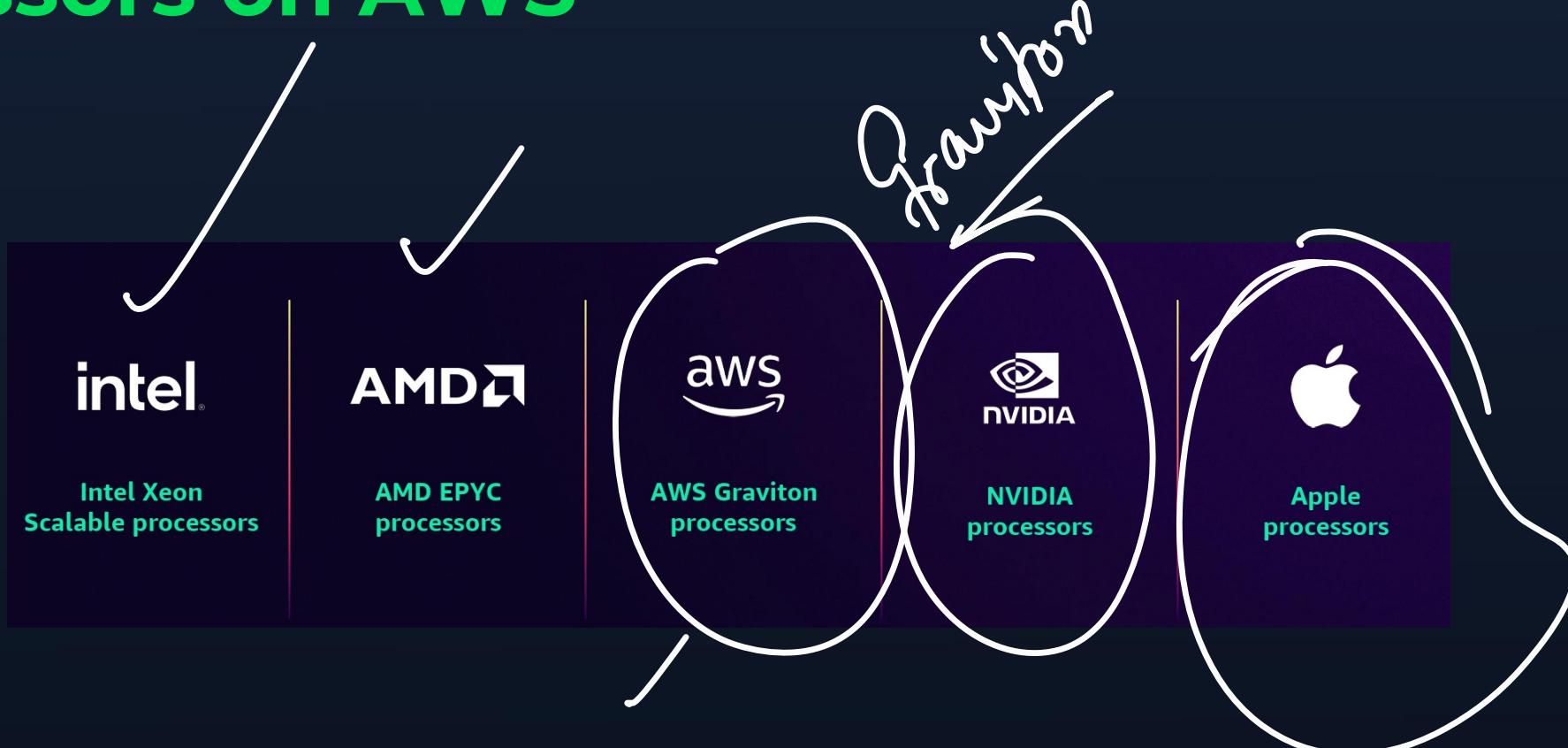
156 File

Introduction to EC2

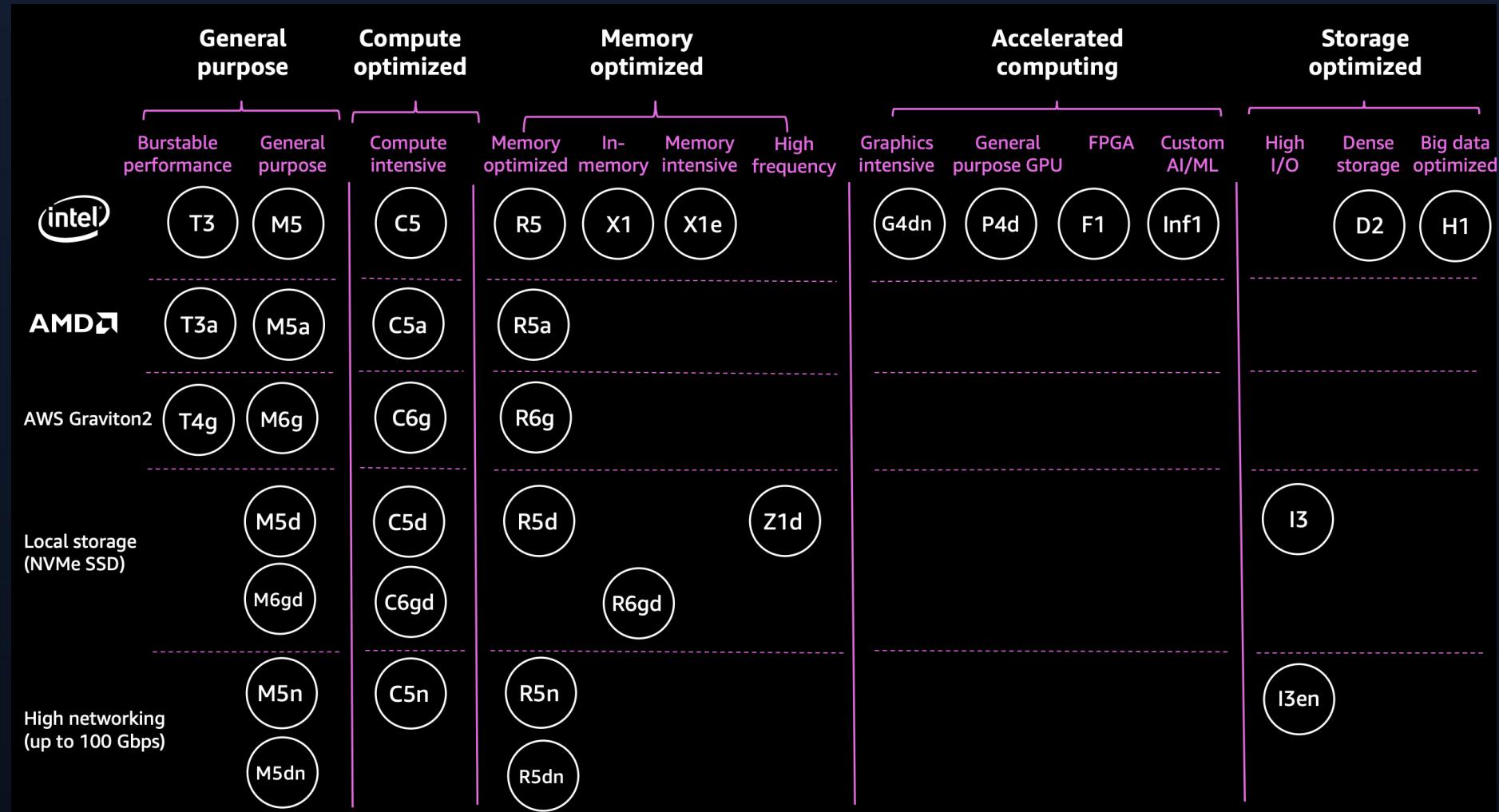




Processors on AWS

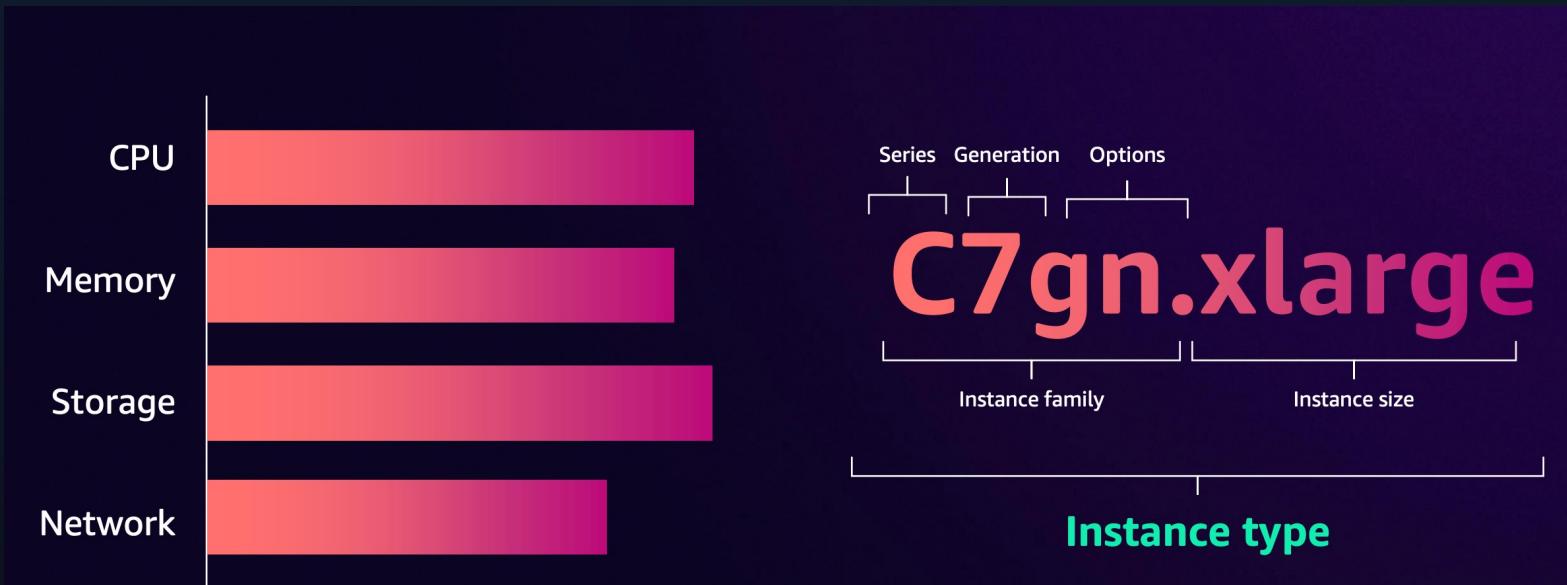


Types of EC2 Instance



Instance Type Characteristics

How to pick the right Instance type and family?



Purchasing Options for EC2

Purchasing Method	Commitment Structure	Exchanging Options	% Savings over On-Demand
On-Demand	Charged on a \$/second basis for instances that are deployed	Not Applicable	Not Applicable
Compute Savings Plans	Commitments are made on \$/hour basis for term lengths of 1 or 3 years	Not Applicable	Up to 66%
EC2 Savings Plans	Commitments are made on \$/hour basis for term lengths of 1 or 3 years	Not Applicable	Up to 72%
Standard RI	Commitments are made to specific instance configurations for term lengths of 1 or 3 years	Can be bought and sold in the AWS RI Marketplace	Up to 72%*
Convertible RI	Commitments are made to specific instance configurations for term lengths of 1 or 3 years	Can be exchanged for other instances of same or different properties	Up to 66%*

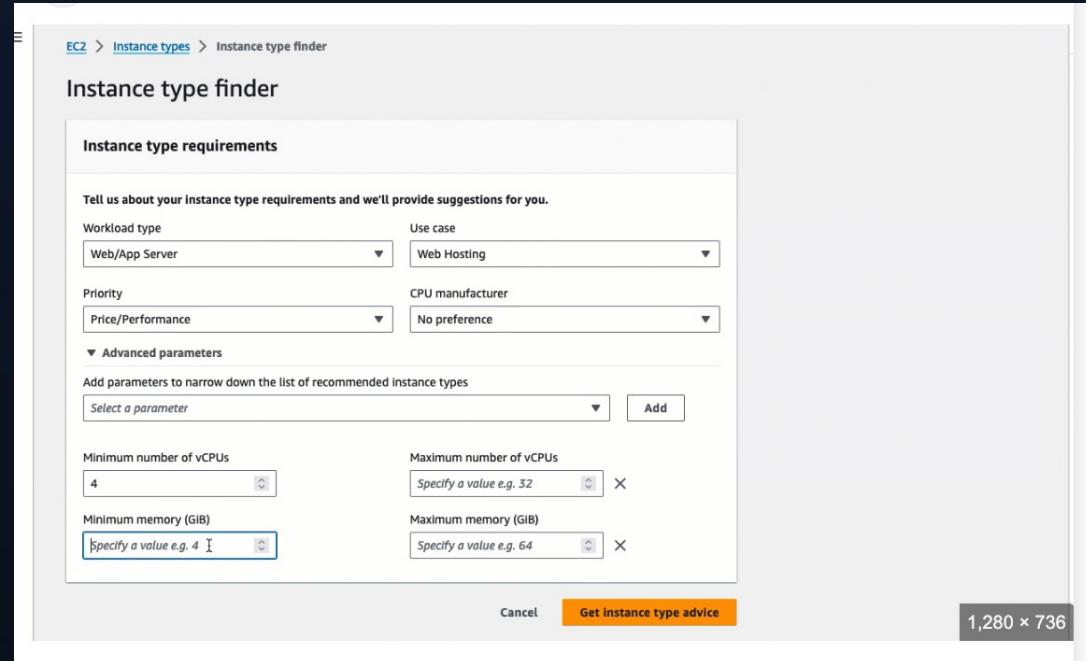
* RIs save 5-10% over their Savings Plan counterparts



Instance Type Identification

How to pick the right Instance type and family?

Have you heard of Instance Type Finder feature that AWS has Launched recently?



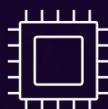
EC2 Instance Selection Assistant

AWS EC2 Instance Selection Assistant

USE NATURAL LANGUAGE TO DESCRIBE YOUR WORKLOAD



Generative AI analyzes workload requirements



Provides personalized instance suggestions

Use GenAI to get the Instance Type

The image displays two screenshots side-by-side. On the left is the 'Launch an instance' wizard in the AWS Management Console, showing fields for 'Name and tags', 'Application and OS Images (Amazon Machine Image)', and a detailed view of the 'Amazon Linux 2023 AMI'. On the right is the Amazon Q AI interface, which has recommended a high-performance EC2 instance type for training a machine learning model.

Left Side: AWS EC2 - Launch an instance

- Name and tags:** A text input field containing "e.g. My Web Server" and a button to "Add additional tags".
- Application and OS Images (Amazon Machine Image):** A dropdown menu showing "Amazon Linux", "macOS", "Ubuntu", "Windows", "Red Hat", and "SUSE Linux". Below this is a search bar with placeholder text "Search our full catalog including 1000s of application and OS images".
- Amazon Machine Image (AMI):** A detailed view of the "Amazon Linux 2023 AMI", including its ID, architecture, virtualization type, and root device type.

Right Side: Amazon Q

I need to train a machine learning model, could you recommend a suitable high-performance EC2 instance type for this task?

Recommended Instance Families: C8g, G5g, Trn1, C7gn, C7gd

Additional Context:

1. The recommended instances C8g, G5g, Trn1, C7gn, and C7gd are optimized for high-performance Machine Learning workloads in Data Science & Analytics use cases, where performance is the preferred criteria.
2. C8g instances, powered by the latest AWS Graviton4 processors, provide the best price-performance for compute-intensive workloads, offering up to 30% better performance than the previous generation C7g instances.
3. G5g instances, featuring NVIDIA T4G Tensor Core GPUs and AWS Graviton2 processors, are designed for graphics workloads like Android game streaming and cost-effective ML inference.
4. Trn1 instances, powered by the AWS Trainium ML chip, deliver the best price-performance for training deep learning models in the cloud.
5. C7gn instances offer up to 200 Gbps of network bandwidth and up to 3x higher packet-processing performance per vCPU compared to current generation x86-based network-optimized instances.
6. C7gd instances are compute-optimized instances with local NVMe-based SSDs, providing up to 45% better real-time NVMe storage performance than comparable Graviton2-based instances.
7. For workloads requiring high network bandwidth, the C7gn instances are available. For workloads requiring local storage, the C7gd instances are available.

You can view the suggested instance families in the [Instance types table](#).

Ask me anything about AWS

Max 1000 characters

Amazon Q Developer uses generative AI. You may need to verify responses. See the [AWS Responsible AI Policy](#).