Understanding EC2 & EBS in AWS:

Amazon EC2 (Elastic Compute Cloud) is the core service in AWS that allows users to rent virtual servers (instances) to run applications. Along with it, Amazon EBS (Elastic Block Store) provides persistent block-level storage for EC2 instances.

In this blog, we'll break down:

- The hardware-based types of EC2 instances
- The cost models of EC2
- The types of EBS volumes (SSD and HDD)

🔨 Types of EC2 Instances (Hardware-Based)

1. General Purpose Instances

These instances offer a balance of compute, memory, and networking resources. They're ideal for applications that don't need special hardware like GPUs or high memory.

Example instances: t3.micro, t2.medium, m5.large

Use cases: Web servers, development and testing environments, small databases.

2. Compute Optimized Instances

Designed for compute-heavy workloads, these instances provide high-performance processors.

Example instances: c5.large, c6g.xlarge

Use cases: High-performance web servers, scientific modeling, batch processing, gaming servers.

3. Memory Optimized Instances

These instances are perfect for memory-intensive applications. They provide high RAM to support large datasets in memory.

Example instances: r5.large, x1e.xlarge

Use cases: In-memory databases (like Redis), real-time big data analytics, SAP HANA.

4. Storage Optimized Instances

These are designed for workloads requiring high, sequential read and write access to large data sets on local storage.

Example instances: i3.large, d2.xlarge

Use cases: NoSQL databases like Cassandra, data warehousing, Hadoop clusters.

5. Accelerated Computing Instances

These instances use GPUs or FPGAs for tasks that require massive parallel processing power.

Example instances: p3.2xlarge, inf1.xlarge

Use cases: Machine learning inference/training, video rendering, financial simulations, high-performance computing.

💰 EC2 Pricing Models (Cost-Based)

1. On-Demand Instances

You pay for compute capacity by the hour or second without any long-term commitments.

Best for: Short-term, unpredictable workloads.

Example: Testing a new application.

Cost: Higher than other models due to flexibility.

2. Reserved Instances

You commit to a 1- or 3-year term in exchange for a significant discount (up to 75%) compared to On-Demand.

Best for: Steady-state workloads.

Example: Running a backend server for a production web app continuously.

Cost: Lower if committed upfront.

3. Savings Plans

A flexible pricing model where you commit to a consistent amount of usage (e.g., $10/hour) over 1 or 3 years. Works across instance families.

Best for: Users looking for cost savings with flexibility.

Example: Multiple services in use across different EC2 types.

Cost: Discounted, similar to Reserved Instances.

4. Spot Instances

You bid on unused EC2 capacity and can save up to 90%. Instances can be terminated anytime AWS needs the capacity back.

Best for: Fault-tolerant or flexible applications.

Example: Big data jobs, background image processing.

Cost: Very low, but not reliable for critical systems.

5. Dedicated Instances

These run on hardware that's dedicated to you but may share hardware with other instances in the same account.

Best for: Compliance or licensing needs.

Example: Running legacy applications that need dedicated resources.

Cost: Higher than On-Demand.

6. Dedicated Hosts

You get physical servers fully dedicated to you, with visibility and control over placement.

Best for: Licensing that requires physical separation (e.g., Microsoft or Oracle).

Example: Enterprises with strict compliance and licensing requirements.

Cost: Highest, but gives full hardware control.

💾 Amazon EBS: Elastic Block Store Types

EBS is a persistent storage service used with EC2 instances. You can attach EBS volumes to your instances and keep data even if the instance is stopped or terminated.

EBS volumes are of two main categories: SSD-backed and HDD-backed.

◇ SSD-Backed (Solid State Drive)

What is SSD?

A storage device that uses flash memory. It has no moving parts, making it **faster** and **more durable**.

1. **gp3 (General Purpose SSD):**

Default choice for most workloads. Offers a balance of price and performance.

Use cases: Boot volumes, small to medium-size databases, dev/test.

2. **io2 / io2 Block Express (Provisioned IOPS SSD):**

High-performance SSD for latency-sensitive and IOPS-intensive workloads.

Use cases: High-performance databases like Oracle, SAP HANA, or SQL Server.

◇ HDD-Backed (Hard Disk Drive)

What is HDD?

A traditional spinning disk drive. It's **slower** but **cheaper** than SSD and better for large, sequential data access.

1. **st1 (Throughput Optimized HDD):**

Optimized for large, sequential workloads.

Use cases: Big data, data warehouses, log processing.

2. **sc1 (Cold HDD):**

Lowest-cost EBS volume, used for infrequently accessed data.

Use cases: Archival storage, file backups, rarely used logs.

Conclusion

Amazon EC2 and EBS are core AWS services offering virtual servers and block storage for applications. This article explores various EC2 instance types—General Purpose, Compute Optimized, Memory Optimized, Storage Optimized, and Accelerated Computing—each suited for different workloads. It also covers EC2 pricing models, including On-Demand, Reserved, Savings Plans, Spot, Dedicated Instances, and Dedicated Hosts, each with unique cost implications. Additionally, it examines EBS volume types, SSD-backed (gp3, io2) and HDD-backed (st1, sc1), designed for different performance and cost needs.