EC2 Pricing Study Notes

Andrew Brown

On Demand

When you launch an EC2 instance it is by default <u>On-Demand</u> Pricing

Key Points

- On-Demand has **no up-front payment** and **no long-term commitment**

- You are charged by the **hour** or by the **minute** (varies based on EC2 Instance Types

- On-Demand is for applications where the workload is for **short-term, spikey, or unpredictable.**

- Used for when you have a **new app** for development or you want to run an experiment

EC2- Reserved Instances (RI)

<u>Reserved Instance Pricing:</u> Designed for applications that have a **steady-state**, **predictable usage**, or require **reserved capacity**.

Key Points

- Reduced pricing is based on **Term x Class Offering x Payment Option**

- **Types of payment options:**

- **Standard:** Up to **75%** reduced pricing compared to on-demand. Cannot change RI Attributes
- **Convertible:** Up to **54%** reduced pricing compared to on-demand. Allows you to change RI Attributes if greater or equal in value.
- **Scheduled:** You reserve instances for specific time periods. **Savings vary.**

- **Terms**- You commit to a 1 year or 3 year contract. The longer the term the greater savings

- **Payment Options**- All Upfront, Partial Upfront, and No Upfront. The greater upfront the greater the savings.

- **RIs** can be shared between multiple accounts within an organization

- **Unused RIs** can be sold in the **Reserved Instance Marketplace**.

Spot Instances

Amazon EC2 Spot Instances: Run fault-tolerant workloads for up to **90% off**

Key Points

- Amazon EC2 Spot Instances let you take advantage of **unused EC2 capacity** in the AWS cloud.

- Designed for applications that have flexible start and end times or applications that are only feasible at **very low** compute costs.

- AWS Batch is an easy and convenient way to use Spot Pricing

- **Termination Conditions:**

  - Instances can be terminated by AWS **at anytime.**
  - If your instance is terminated by AWS, **you don't get charged** for a partial hour of usage.

■ If you terminate an instance **you will still be charged** for an hour that it ran.

Dedicated Host Instances

<u>Dedicated Host Instances</u>: dedicated hardware to support existing software licenses and improve compliance.

Key Points

■ Most expensive

■ Designed to meet regulatory requirements.

■ Use when you have strict **server-bound licensing** that won't support multi-tenancy or cloud deployments

■ **Multi-tenant:** When multiple customers are running workloads on the same hardware. **Virtual Isolation** is what separate customers (Like a multi-tenant apartment)

■ **Single Tenant:** When a single customer has dedicated hardware. **Physical Isolation** is what separates customers(Like a single-family home)

■ Offered in both **On-demand** and **Reserved** (70% of on-demand pricing)

■ **Enterprises** and **Large Organizations** my have security concerns or obligations against sharing the same hardware with other AWS Customers

Savings Plans

<u>AWS Savings Plans</u> was introduced as an alternative to Reserved Instance (RI) Pricing to simplify purchasing. They sound similar but there are some key differences.

With **Savings Plans** you have to commit up-front to:

- 1 year contract
- 3 years contract

and any usage over the committed usage is charged at On-Demand rate

With **Reserved Instances**, you can commit to full, partial or up-front. You can commit to

- 1 year contract
- 3 years contract

With **Savings Plans** you can only save with EC2 and Fargate

With **Reserved Instances** you can only save with EC2, RDS, Redshift and ElastiCache

With *Savings Plans* **you commit to a minimum dollar per hour spend With** *Reserved Instances* you commit to a specific amount of utilization

With **Convertible Reserved Instance** you can resell your utilized resources in the AWS Marketplace With **Savings Plans** you cannot resell in the AWS Marketplace.

Cheat Sheet

- **EC2 has 4 pricing models**
    1. On-Demand
    2. Spot
    3. Reserved Instances (RI)
    4. Dedicated

- **<u>On-Demand</u>** (Least commitment)
  - Low cost and flexible
  - Only pay per hour
  - Use Case: short-term, spiky, unpredictable workloads, and first time apps
  - Ideal when your workloads cannot be interrupted.
- **<u>Reserved Instances</u>**- up to 75% off (Best long-term value)
  - **Use case:** steady state or predictable usage
  - Can resell unused reserved instances (Reserved Instance Marketplace)
  - Reduced Pricing is based on **Term x Class Offering x Payment Option**
  - **Payment Terms:** 1 year or 3 years
  - **Payment Options:** All Upfront, Partial Upfront, No Upfront
  - **Class Offerings:**
    - **Standard:** Up to 75% reduced pricing compared to on-demand. Cannot change RI attributes
    - **Convertible:** Up to 54% reduced pricing compared to on-demand. Allows you to change RI Attributes if greater than or equal in value.
    - **Scheduled:** You reserve instances for specific times periods. Savings vary.
- **<u>Spot Pricing</u>**- up to 90% off (Biggest Savings)
  - Request spare computing capacity
  - Flexible start and end times
  - **Use case:** Can handle interruptions like a server randomly stopping and starting
  - **Use case:** For non-critical background jobs
  - Instances can be terminated by AWS **at anytime**
  - If your instance is **terminated by AWS**, you don't get charged for a partial hour of usage.
  - If **you terminate** an instance you will still be charged for any hour that it ran.
- **<u>Dedicated Hosting</u>** (Most Expensive)
  - Dedicated servers
  - Can be on-demand or reserved (up to 70% off)
  - **Use case**: When you need a guarantee of isolated hardware (enterprise requirements)