


AWS EC2 INSTANCES



Instances Types

1. On Demand Instances
2. Reserved Instances
3. Spot Instances
4. Dedicated Instances

Instance Type

Dedicated Instances

If you want to choose dedicated instances then when launching an instances go to the Configure instance
→ Tenancy → Dedicated

Shared Instances

- On-demand Instances
- Spot Instances
- Reserved Instances

Pricing : Reserved Instances have three pricing model all-together.

- full-upfront : → You pay your all money at the time of booking
- Partial-upfront : → You pay your few money at booking and then as you required
- No-upfront : → You pay after using

On-Demand Instances!

if the workload is short-term, spikey,
unpredicted.

You are charged by hour

AWS EC2 INSTANCES



On Demand Instances

1. This Instance is one which we usually create.
2. There is no Time Commitment.
3. No Advance Payment. → No UPFRONT
4. You only need to pay for hours which you used.
5. Highest Cost as there is no commitment.

Benefits →

→ When you are uncertain about time.

By default what you choose is a on-demand instance. If you want to change it then after launching choose instance type and click on Purchasing option.

Suppose you have web of 4 servers. But for a few days there is an additional traffic. For that you need 2 more servers. For this additional servers you can opt for on-demand servers.

For long time → Reserved Instances

For high performance → Spot Instances as we get high-performance servers in cheaper price

Reserved Instances :— Where applications
wave steady-state, predictable usage.

AWS EC2 INSTANCES



Reserved Instances

1. For these, Instance there is a time commitment. You buy it for a time frame such as 3/6 Months or 1/3 Year.
2. You need to purchase it via Console and it will automatically get applied on your On Demand Instances if you are using any of same configuration.
3. Various Payment option like All Upfront, No Upfront or Partial Upfront.
4. You get a good discount like 30% to 70%.



EC2 - Reserved Instances (RI)

Best Long-term

Designed for applications that have a **steady-state, predictable usage**, or require **reserved capacity**.

Reduced Pricing is based on **Term x Class Offering x Payment Option**

Platform	Linux/UNIX	Tenancy	Default	Offering Class	Standard				
Instance Type	t2.micro	Term	12 months - ...	Payment Option	Partial Upfront				
Seller	Term	Effective Rate	Upfront Price	Hourly Rate	Payment Option	Offering Class	Quantity Available	Desired Quantity	Normalized units per hour
AWS	36 months	\$0.005	\$66.00	\$0.002	Partial Upfront	standard	Unlimited	1	0.5

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 eg. once a week for a few hours. 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 org

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

Spot Instances - Big capacity servers cheap
Price is cheapest

AWS EC2 INSTANCES



Spot Instances

1. These Instance are cheapest one and good for small duration like few hours to few days only.
2. Work on bid pricing means you put a bid for instance, and if that instance is available under bid price, it is available for you and if bid price goes higher, AWS will pull the instance automatically.
3. Good for one who want to use high processing power at low price and need for minimum duration.



EC2 - Spot Instances

Biggest Savings

AWS has unused compute capacity that they want to maximize the utility of their idle servers.
It's like when a hotel offers discounts for to fill vacant suites or planes offer discount to fill vacant seats.

Spot Instances provide a discount of **90%** compared to On-Demand Pricing
Spot Instances can be terminated if the computing capacity is needed by on-demand customers.

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

Tell us your application or task need

To help us identify the most appropriate compute capacity for your job, select the closest match for your application or task need.

Load balancing workloads

Launch instances of the same size, in any Availability Zone. Good for running web services.

Flexible workloads

Launch instances of any size, in any Availability Zone. Good for running batch and CI/CD jobs.

Big data workloads

Launch instances of any size, in a single Availability Zone. Good for MapReduce jobs.

Defined duration workloads

Launch instances into a spot block for 1 to 6 hours.

One hour



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 any hour that it ran.

Dedicated Instances :- Performance is Best

AWS EC2 INSTANCES



Dedicated Instances

1. Costlier than default Shared Instances
2. Give better performance as allotted CPU and Memory is not shared with other clients.

Resources will not be shared for dedicated instance with other clients (except CPU & RAM)

If we choose services here
that will be on-demand
services.

The screenshot shows the AWS Cloud Learning environment with the Mumbai region selected. The EC2 Instances page is open, displaying a single instance named "Web Server". The instance details are as follows:

Name	Instance ID	Instance Type	Availability Zone	Instance State	Status Checks	Alarm Status	Public DNS (IPv4)
Web Server	i-044535769b0d97e83	t2.micro	ap-south-1a	stopped	Loading...	-	-

Below the table, the instance details are expanded:

Description	Value	Description	Value
Instance ID	i-044535769b0d97e83	Public DNS (IPv4)	-
Instance state	stopped	IPv4 Public IP	-
Instance type	t2.micro	IPv6 IPs	-
Elastic IPs	-	Private DNS	ip-172-31-25-131.ap-south-1.compute.internal
Availability zone	ap-south-1a	Private IPs	172.31.25.131



EC2 - Dedicated Host Instances

Most Expensive

Designed to meet regulatory requirements. When you have strict **server-bound licensing** that won't support multi-tenancy or cloud deployments.

Multi-Tenant vs Single Tenant

When multiple customers are running workloads on the same hardware. **Virtual Isolation** is what separate customers. (think apartment)



Multi-Tenant

When a single customer has dedicated hardware. **Physical Isolation** is what separates customers (think house)



Single-Tenant



Single-Tenant



Single-Tenant

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



Enterprises and **Large Organizations** may have security concerns or obligations about against sharing the same hardware with other AWS Customers.



EC2 - Pricing Model

On-Demand

Least Commitment

- low cost and flexible
- only pay per hour
- short-term, spiky, unpredictable workloads
- cannot be interrupted
- For first time apps

Spot

upto 90%

Biggest Savings

- request spare computing capacity
- flexible start and end times
- Can handle interruptions (server randomly stopping and starting)
- For non-critical background jobs

Reserved

upto 75% off
Best Long-term

- steady state or predictable usage
- commit to EC2 over a 1 or 3 year term
- Can resell unused reserved instances

Dedicated

Most Expensive

- Dedicated servers
- Can be on-demand or reserved (upto 70% off)
- When you need a guarantee of isolate hardware (enterprise requirements)



EC2 Pricing - *CheatSheet*

- EC2 has 4 pricing models **On-Demand, Spot, Reserved Instances (RI)** and **Dedicated**
- **On-Demand** (least commitment)
 - low cost and flexible
 - only pay per hour
 - **Use case:** short-term, spiky, unpredictable workloads, first time apps
 - Ideal when your workloads cannot be interrupted
- **Reserved Instances** upto 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 year
 - **Payment Options:** All Upfront, Partial Upfront, and 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 or equal in value.
 - **Scheduled** You reserve instances for specific time periods eg. once a week for a few hours. Savings vary



EC2 Pricing - *CheatSheet*

- **Spot Pricing** upto 90% off (Biggest Savings)
 - request spare computing capacity
 - flexible start and end times
 - **Use case:** Can handle interruptions (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.
- **Dedicated Hosting** (Most Expensive)
 - Dedicated servers
 - Can be on-demand or reserved (upto 70% off)
 - **Use case:** When you need a guarantee of isolate hardware (enterprise requirements)