AWS EC2 Complete Details

Amazon Elastic Compute Cloud (Amazon EC2) is a web service that provides secure, resizable compute capacity in the cloud.

Instance Categories

1. General-Purpose: Ideal for business applications, small and mid-sized databases, web tier applications.
T2 Websites and web apps, development environments, build servers, code repositories, micro services, test and staging environments
T3 Micro-services, low-latency interactive applications, small and medium databases, virtual desktops
M4 Small and mid-size databases, backend servers for SAP and other enterprise applications
M5 Small and mid-size databases, backend servers for SAP and other enterprise applications

** Amazon EC2 allows you to choose between Fixed Performance Instances (e.g. M5, C5, and R5) and Burstable Performance Instances (e.g. T3).
Burstable Performance Instances provide a baseline level of CPU performance with the ability to burst above the baseline.

Type	CPU Count	RAM	Cost
t2.2xlarge	8	32	$0.3712 per Hour
t2.xlarge	4	16	$0.1856 per Hour
t3.xlarge	4	16	$0.1664 per Hour
t3.2xlarge	8	32	$0.3328 per Hour
m5.2xlarge	8	32	$0.384 per Hour
m5.4xlarge	16	64	$0.768 per Hour
m4.2xlarge	8	32	$0.80 per Hour

2. COMPUTE OPTIMIZED: ideal for compute bound applications that benefit from high performance processors. Used in dedicated gaming servers and ad server engines, machine learning inference
C5 c5.4xlarge 16cpu 32ram c4.2xlarge 8cpu 15ram
3. MEMORY OPTIMIZED: Designed to deliver fast performance for workloads that process large data sets in memory. **R5 R4 X1...** high performance databases, distributed web scale in-memory caches, mid-size in-memory databases, real time big data analytics
4. ACCELERATED COMPUTING: Designed to work more efficiently than is possible in software running on CPUs. **P3, G3, G4**
Uses: computational fluid dynamics, computational finance, seismic analysis, speech recognition, autonomous vehicles, drug discovery.
running graphics-intensive applications
5. 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. IOPS is price calculation
I3 instance i3.2xlarge 8CPU 61Ram
Uses: NoSQL databases (e.g. Cassandra, MongoDB, Redis), scale-out transactional databases, data warehousing. Mostly used General, Compute and Storage types
https://www.cloudhealthtech.com/blog/aws-instance-types-and-comparison

	Type	Description	Mnemonic
General Purpose	a1	Good for scale-out workloads, supported by Arm	**a** is for Arm processor – or as light as **A1** steak sauce
	t-family: t3, t3a, t2	Burstable, good for changing workloads	**t** is for **tiny** or **turbo**
	m-family: m6g, m5, m5a, m5n, m4	Balanced, good for consistent workloads	**m** is for **main** or happy **medium**
Compute Optimized	c-family: c5, c5n, c4	High ratio of compute to memory	**c** is for **compute**
Memory Optimized	r-family: r5, r5a, r5n, r4	Good for in-memory databases	**r** is for **RAM**
	x1-family: x1e, x1	Good for full in-memory applications	**x** is for **xtreme**
	High memory	Good for large in-memory databases	High memory is for... high memory.
	z1d	Both high compute and high memory	**z** is for **zippy**
Accelerated Computing	p-family: p3, p2	Good for graphics processing and other GPU uses	**p** is for **pictures**
	Inf1	Support machine learning inference applications	**Inf** is for **inference**
	g-family: g4, g3	Accelerate machine learning inference and graphics-intensive workloads	**g** is for **graphics**
	f1	Customizable hardware acceleration with field programmable gate arrays (FPGAs)	**f** is for **FPGA** or **feel** as in hardware
Storage Optimized	i-family: i3, i3en	SDD-backed, balance of compute and memory	**i** is for **IOPS**
	d2	Highest disk ratio	**d** is for **dense**
	h1	HDD-backed, balance of compute and memory	**H** is for **HDD**

Elastic IP address: Every instance comes with its own private and public address. the public address is associated exclusively with the instance until it is stopped or terminated.
However, this can be replaced by the Elastic IP address, which stays with the instance as long as the user doesn't manually detach it. if you are hosting multiple websites on your EC2 server, in that case you may require more than one Elastic IP address.

AMIs(Amazon Machine Images) are like templates of virtual machines and an instance is derived from an AMI
Security group: is just like a firewall, it controls the traffic in and out of your instance. In AWS terms, the inbound and outbound traffic.
Spot Instances : A Spot Instance is an unused EC2 instance that is available for less than the On-Demand price. Your Spot Instance runs whenever capacity is available and the maximum price per hour for your request exceeds the Spot price. Spot Instances are a cost-effective choice if you can be flexible about when your applications run and if your applications can be interrupted.
Reserved instances : Reserved Instances are not physical instances, but rather a billing discount applied to the use of On-Demand Instances in your account.

Dedicated instances : Dedicated Instances are Amazon EC2 instances that run in a virtual private cloud (VPC) on hardware that's dedicated to a single customer. Dedicated Instances that belong to different AWS accounts are physically isolated at the hardware level.

On-Demand instances : An On-Demand Instance is an instance that you use on demand. You have full control over its lifecycle—you decide when to launch, stop, hibernate, start, reboot, or terminate it. There is no long-term commitment required when you purchase On-Demand Instances.

Spot instance different from an On-Demand instance or Reserved Instance?
Spot instances are just like bidding, the bidding price is called Spot Price.
The Spot Price fluctuates based on supply and demand for instances, but customers will never pay more than the maximum price they have specified.If the Spot Price moves higher than a customer's maximum price, the customer's EC2 instance will be shut down automatically.
But the reverse is not true, if the Spot prices come down again, your EC2 instance will not be launched automatically, one has to do that manually. In Spot and On demand instance, there is no commitment for the duration from the user side, however in reserved instances one has to stick to the time period that he has chosen.

AWS CLI EC2

Create,launch,stop,terminate an Instance from AWS CLI
run-instances:
1. $ ***aws ec2 run-instances --image-id --instance-type --security-group-ids --key-name --count --subnet-id***
Example: $ aws ec2 run-instances --image-id ami-abc1234 --count 1 -instance-type m4.large --key-name keypair --user-data file://my_script.txt --subnet-id subnet-abcd1234 --security-group-ids sg-abcd1234

AWS CLI ami 2018: "InstanceId": "i-0b0dd2a2f57392c00"
RedHat 7.5 : "InstanceId": "i-0ece23056d2c0e663"
Ubuntu 16: "InstanceId": "i-044ea9c22acbc63a5"

1. START-INSTANCES
$ ***aws ec2 start-instances --instance-ids***
Example: $ aws ec2 start-instances --instance-ids i-1234567890abcdef0

2. STOP-INSTANCES
$ ***aws ec2 stop-instances --instance-ids***
Example: $ aws ec2 stop-instances --instance-ids i-1234567890abcdef0

3. TERMINATE-INSTANCES
$ ***aws ec2 terminate-instances --instance-ids***
Example: $ aws ec2 terminate-instances --instance-ids i-1234567890abcdef0

4. CREATE A SECURITY GROUP
$ ***aws ec2 create-security-group --description --group-name --vpc-id***
Example: $ aws ec2 create-security-group --group-name MySecurityGroup --description "My security group" --vpc-id vpc-1a2b3c4d
"GroupId": "sg-00c7cfc8fac66dea4"

5. AUTHORIZE-SECURITY-GROUP-INGRESS(INBOUND)

$ ***aws ec2 authorize-security-group-ingress --group-id --ip-permissions --protocol --port***

Example:$ aws ec2 authorize-security-group-ingress --group-id sg-903004f8 --protocol tcp --port 22 --cidr 203.0.113.0/24

$ aws ec2 authorize-security-group-ingress --group-id sg-00c7cfc8fac66dea4 --port 22 --cidr 0.0.0.0/0 --protocol tcp

6. DELETE-SECURITY-GROUP

$ ***aws ec2 delete-security-group --group-name***

Example: aws ec2 delete-security-group --group-name MySecurityGroup

aws ec2 delete-security-group --group-id sg-903004f8