EC2

Rahul Daware

April 15, 2019

Contents

1 EC2 101

Amazon Elastic Compute Cloud (Amazon EC2) is a web service that provides resizable compute capacity in the cloud. Amazon EC2 reduces the time required to obtain and boot new server instances to minutes, allowing you to quickly scale capacity, both up and down, as your computing requirements change.

EC2 Pricing Models:

- On Demand - Allows you to pay a fixed rate by the hour (or by the second) with no commitment

- Reserved - Provides you with a capacity reservation, and offe a significant discount on the hourly charge for an instance. Contract Terms are 1 Year or 3 Year Terms.

- Spot - Enables you to bid whatever price you want for instance capacity, providing for even great savings if your applications have flexible start and end times.

- Dedicated Hosts - Physical EC2 Server dedicated for your suse. Dedicated Hosts can help you reduce costs by allowing you to use your existing server-bound software licenses.

1.1 On Demand Pricing

- Users that want the low cost and flexiblility of Amazon EC2 without any up-front payment or long term commitment

- Applications with short term, spiky, or unpredictable workloads that cannot be interrupted.

- Applications being developed or tested on Amazon EC2 for the first time.

1.2 Reserved Pricing

- Applications with steady state or predictable usage

- Applications that require reserved capacity

- Users able to make upfront payments to reduce their total computing costs even further

- **Standard Reserved Instances** - These offer up to 75% off on demand instances. The more you pay up front and the longer the contract, the greater the discount.

- **Convertible Reserved Instances** - These offer up to 54% off on demand capability to change the attrbutes of the RI as long as the exchange results in the creation of Reserved Instances of equal or greater value.

- **Scheduled Reserved Instances** - These are available to launch within the time windows tou reserve. This option allows you to match your capacity reservation to a predictable recurring schedule that only requires a fraction of a day, a week, or a month.

1.3 Spot Pricing

- Applications that have flexible start and end times

- Applications that are only feasible at very low compute prices

- Users with urgent computing needs for large amounts of additional capacity

1.4 Dedicated Hosts Pricing

- Useful for regulatory requirements that may not support multi-tenant virtualization

- Great for licensing which does not support multi-tenancy or cloud deployments

- Can be purchased On-Demand (hourly)

- Can be purchased as a reservation for up to 70% off the on-demand price

1.5 EC2 Instance Types

Family	Specialty	Use Case
F1	Field Programmable Gate Array	Genomics research. financial analytics, real-time video processing, big data etc
I3	High Speed Storage	NoSQL DBs, Data Warehousing etc
G3	Graphics Intensive	Video Encoding, 3D Application Streaming
H1	High Disk Throughput	MapReduce based workloads, distributed file systems such as HDFS and MapR-FS
T3	Lowest Cost, General Purpose	Web Servers/Small DBs
D2	Dense Storage	Fileservers/Data Warehousing/ Hadoop
R5	Memory Optimized	Memory Intensive Apps/ DBs
M5	General Purpose	Application Servers
C5	Compute Optimized	CPU Intensive Apps/ DBs
P3	Graphics/ General Purpose GPU	Machine Learning , Bit Coing Mining etc
X1	Memory Optimized	SAP HANA/Apache Spark etc
Z1D	High compute capacity and a high memory footprint	Ideal for electronic design automation (EDA) and certain relational database workloads with high per-core licensing costs.
A1	Arm-based workloads	Scale out workloads such as web servers
U-6tb1	Bare Metal	Bare Metal Capabilities that eliminate virtualization overhead

1.6 Exam Tips

- Termination Protection is turned off by default, you must turn it on.

- On an EBS-back instance, the default action is for the root EBS volume to be deleted when the instance is terminated.

- EBS Root Volumes of your DEFFAULT AMIs cannot be encrypted. You can use a third party tol (such as bit locker etc) to encrypt the root volume, or this can be done when creating AMIs in the AWS console or using the API.

- Additional volumes can be encrypted.

2 Security Groups Basics

- All inbound traffic is blocked by default.

- All outbound traffic is allowed.

- Changes to security groups take effect immediately.

- You can have any number of EC2 instances within a security group.

- You can have multiple security groups attached to EC2 instances.

- Security Groups are STATEFUL.

- If you create an inbound rule allowing traffic in, that traffic is automatically allowed back out again.

- You cannot block specific IP addresses using security groups, instead use network access control lists.

- You can specify allow rules, but no deny rules

3   EBS

Amazon Elastic Block Store (EBS) provides persistent block storage volumes for use with Amazon EC2 instances in the AWS cloud. Each Amazon EBS volume is automatically replicated within its availability zone to protect you from component failure, offering high availability and durability

Types of EBS Storage:

- General purpose(SSD)

- Provisioned IOPS (SSD)

- Throughput optimised hard disk drive

- Cold hard disk drive

- Magnetic

Solid State Drives:

Volume Type	Description	Use Cases	API Name	Volume Size	Max IOPS
General Purpose SSD	General Purpose SSD volume that balances price and performance for wide variety of transactional workloads	Most Work Loads	gp2	1GiB-16TiB	16,000
Provisioned IOPS SSED	highest-performance SSD volume designed for mission critical applications	Databases	io1	4GiB-16TiB	64,000

Hard Disk Drives:

Volume Type	Description	Use Cases	API Name	Volume Size	Max IOPS
Throughput Optimized HDD	Low cost HDD volume designed for frequently accessed, throughput intensive workloads	Big Data, Data Warehouses	st1	500GiB-16TiB	500
Cold HDD	Lowest cost HDD volume designed for less frequently accessed workloads	File servers	sc1	500GiB-16TiB	250
EBS Magnetic	Previous Generation HDD	Workloads where data is infrequently accessed	Standard	1GiB-1TiB	40-200