

EC2

Rahul Daware

April 16, 2019

Contents

1	EC2 101	3
1.1	On Demand Pricing	3
1.2	Reserved Pricing	3
1.3	Spot Pricing	3
1.4	Dedicated Hosts Pricing	3
1.5	EC2 Instance Types	4
1.6	Exam Tips	4
2	Security Groups Basics	4
3	EBS	5
3.1	Exam Tips	5
4	AMI Types (EBS vs Instance Store)	6
4.1	Exam Tips	6
5	Volumes and Snapshots	6
5.1	Exam Tips	7
6	CloudWatch 101	7
6.1	CloudWatch and EC2	7
6.2	What is AWS CloudTrail	7
6.3	Exam Tips	7
7	IAM Roles	8
8	Instance MetaData	8

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 offers 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 greater savings if your applications have flexible start and end times.
- Dedicated Hosts - Physical EC2 Server dedicated for your use. 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 flexibility 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 attributes 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 you 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

3.1 Exam Tips

- Volumes exist on EBS. Think of EBS as a virtual hard disk.
- Snapshots exist on S3. Think of snapshots as a photograph of the disk.
- Snapshots are point in time copies of volumes.
- Snapshots are incremental - this means that only the blocks that have changed since your last snapshot are moved to S3.
- If it is your first snapshot, it may take some time.
- To create a snapshot for Amazon EBS volumes that serve as root devices, you should stop the instance before taking the snapshot.
- However you can take a snap while the instance is running.
- You can create AMIs from both volumes and snapshots.

- You can change EBS volume sizes on the fly, including changing the size and storage type.
- Volumes will ALWAYS be in the same availability zone as the EC2 instance.
- To move an EC2 volumes from one AZ to another, take a snapshot of it, create an AMI from the snapshot and then use the AMI to launch the EC2 instance in a new AZ.
- To move an EC2 volume from one region to another, take a snapshot of it, create an AMI from the snapshot, and then copy the AMI from one region to the other. Then use the copied AMI to launch the new EC2 instance in the new region.

4 AMI Types (EBS vs Instance Store)

You can select your AMI based on:

- Region (see Regions and AZ)
- Operating System
- Architecture (32 bit or 64 bit)
- Launch Permissions
- Storage for the root device (Root Device Volume)
 - Instance Store (EPHEMERAL STORAGE)
 - EBS Backed Volumes

All AMIs are categorized as either backed by Amazon EBS or backed by instance store. For EBS volumes, the root device for an instance launched from the AMI is an Amazon EBS volume created from an Amazon EBS snapshot. For instance store volumes, the root device for an instance launched from the AMI is an instance store volume created from a remplate stored in Amazon S3.

4.1 Exam Tips

- Instance Store Volumes are sometimes called Ephemeral Storage.
- Instance store volumes cannot be stopped. If the underlying host fails, you will lose your data.
- EBS backed instances can be stopped. You will not lose the data on this instance if it is stopped.
- You can reboot both, you will not lose your data.
- By default, both ROOT volumes will be deleted on termination. However, with EBS volumes, you can tell AWS to keep the root device volume.

5 Volumes and Snapshots

Steps to create encrypted volumes:

- Create a snapshot of the unencrypted root device volume
- Create a copy of the snapshot and select the encrypt option
- Create an AMI from the encrypted snapshot
- Use that AMI to launch new encrypted instances

5.1 Exam Tips

- Snapshots of encrypted volumes are encrypted automatically.
- Volumes restored from encrypted snapshots are encrypted automatically.
- You can share snapshots, but only if they are unencrypted.
- These snapshots can be shared with other AWS accounts or made public.

6 CloudWatch 101

Amazon CloudWatch is a monitoring service to monitor your AWS resources, as well as the applications that you run on AWS. CloudWatch can monitor things like compute (EC2 instances, autoscaling groups, elastic load balancers, Route53 health checks), Storage and Content Delivery (EBS volumes, Storage Gateways, CloudFront).

6.1 CloudWatch and EC2

Host level metrics consist of:

- CPU
- Network
- Disk
- Status Check

6.2 What is AWS CloudTrail

AWS CloudTrail increased visibility into your user and resource activity by recording AWS management console actions and API calls. You can identify which users and accounts called AWS, the source IP address from which the calls were made, and when the calls occurred.

6.3 Exam Tips

- CloudWatch is used for monitoring performance.
- CloudWatch can monitor most of AWS as well as your applications that run on AWS.
- CloudWatch with EC2 will monitor events every 5 minutes by default.
- You can create 1 minute intervals by turning on detailed monitoring.
- You can create CloudWatch alarms which trigger notifications.
- CloudWatch is all about performance. CloudTrail is all about auditing.
- Standard Monitoring is 5 minutes
- Detailed Monitoring is 1 minute
- You can create dashboards to see what is happening with your AWS environment
- You can set alarms that notify you when particular thresholds are hit.
- CloudWatch events help you to respond to state changes in your AWS resources.
- CloudWatch logs help you to aggregate, monitor and store logs.

7 IAM Roles

- Roles are more secure than storing your access key and secret access key on individual EC2 instances
- Roles are easier to manage
- Roles can be assigned to an EC2 instance after it is created using both the console and command line
- Roles are universal - you can use them in any region

8 Instance MetaData

- Used to get information about an instance (such as public ip)
- Meta Data: `curl http://169.254.169.254/latest/meta-data/`
- User Data: `curl http://169.254.169.254/latest/meta-data/`

9 Elastic File System (EFS)

Amazon Elastic File System (Amazon EFS) is a file storage service for Amazon Elastic Compute Cloud (EC2) instances. Amazon EFS is easy to use and provides a simple interface that allows you to create and configure file systems quickly and easily. With Amazon EFS, storage capacity is elastic, growing and shrinking automatically as you add and remove files, so your applications have the storage they need, when they need it.

9.1 Exam Tips

- Supports the network file system version 4 (NFSv4) protocol
- You only pay for the storage you use (no pre-provisioning required)
- Can scale up to petabytes
- Can support thousands of concurrent NFS connections
- Data is stored across multiple AZs within a region
- Read after write consistency

10 EC2 Placement Groups

10.1 Clustered Placement Group

A clustered placement group is a group of instances within a single Availability Zone. Placement groups are recommended for applications that need low network latency, high network throughput, or both. Only a certain instance can be launched in to a clustered placement group. they cannot span multiple AZs.

10.2 Spread Placement Group

A spread placement group is a group of instances that are each placed on distinct underlying hardware. Spread placement groups are recommended for applications that have a small number of critical instances that should be kept separate from each other. They can span on multiple AZs

10.3 Exam Tips

- The name you specify for a placement group must be unique within your AWS account.
- Only certain types of instances can be launched in a placement group (Compute Optimized, GPU, Memory Optimized, Storage Optimized)
- AWS recommend homogeneous instances within placement groups.
- You cant merge placement groups
- You cant move an existing instance into a placement group. You can create an AMI from your existing instance and then launch a new instance from the AMI into a placement group.