EC2

Basics

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.

Amazon EC2 changes the economics of computing by allowing you to pay only for capacity that you actually use. Amazon EC2 provides developers the tools to build failure resilient applications and isolate themselves from common failure scenarios.

Types of instances

- **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 offer a significant discount on the hourly charge for an instance.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.

On Demand Instances

- Perfect for 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

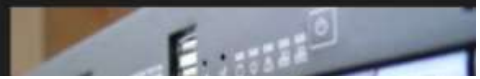
Reserved Instances

- Applications with steady state or predictable usage
- Applications that require reserved capacity
- Users can make up-front payments to reduce their total computing costs even further
  - Standard RIs (Up to 75% off on-demand)
  - Convertible RIs (Up to 54% off on-demand) feature the 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 RIs are available to launch within the time window 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.

Spot and Dedicated

- Applications that have flexible start and end times
- Applications that are only feasible at very low compute prices
- Users with an urgent need for large amounts of additional computing capacity

- 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.



Instance Families

Family	Speciality	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
T2	Lowest Cost, General Purpose	Web Servers/Small DBs
D2	Dense Storage	Fileservers/Data Warehousing/Hadoop
R4	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 Coin Mining etc
X1	Memory Optimized	SAP HANA/Apache Spark etc

Instance Family

- **F** for FPGA
- **I** for IOPS
- **G** - Graphics
- **H** - High Disk Throughput
- **T** cheap general purpose (think T2 Micro)
- **D** for Density
- **R** for RAM
- **M** - main choice for general purpose apps
- **C** for Compute
- **P** - Graphics (think Pics)
- **X** - Extreme Memory

EBS

Amazon EBS allows you to create storage volumes and attach them to Amazon EC2 instances. Once attached, you can create a file system on top of these volumes, run a database, or use them in any other way you would use a block device. Amazon EBS volumes are placed in a specific Availability Zone, where they are automatically replicated to protect you from the failure of a single component.

- General Purpose SSD (GP2)
  - General purpose, balances both price and performance.
  - Ratio of 3 IOPS per GB with up to 10,000 IOPS and the ability to burst up to 3000 IOPS for extended periods of time for volumes at 3334 GiB and above.
- Provisioned IOPS SSD (IO1)
  - Designed for I/O intensive applications such as large relational or NoSQL databases.
  - Use if you need more than 10,000 IOPS.

EBS Volume Types

- Throughput Optimized HDD (ST1)
  - Big data
  - Data warehouses
  - Log processing
  - Cannot be a boot volume
- Cold HDD (SC1)
  - Lowest Cost Storage for infrequently accessed workloads
  - File Server
  - Cannot be a boot volume.

- Magnetic (Standard)
  - Lowest cost per gigabyte of all EBS volume types that is bootable. Magnetic volumes are ideal for workloads where data is accessed infrequently, and applications where the lowest storage cost is important.

Security Group

- 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 not deny rules.

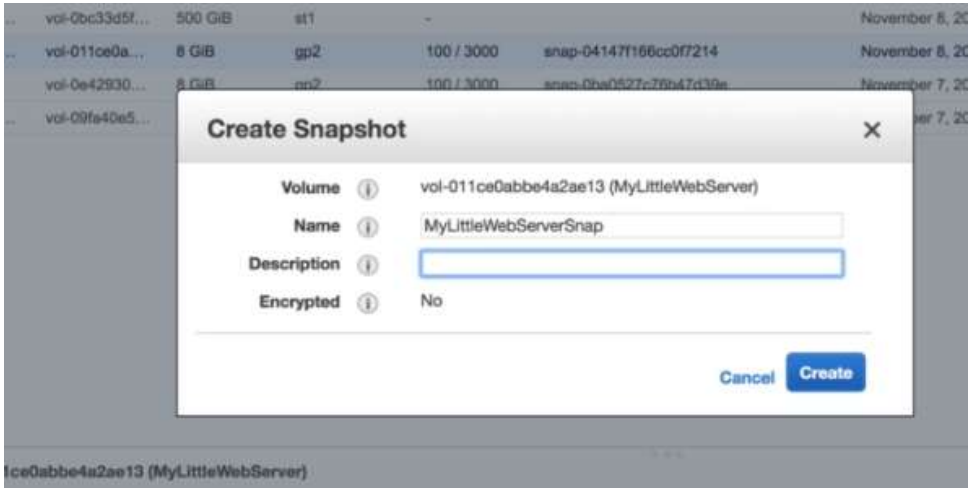
Upgrading EBS Volume

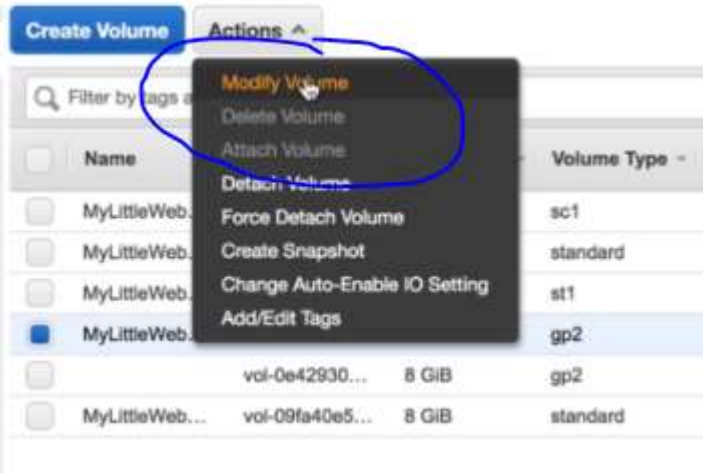
storage options in Amazon EC2.

Volume Type	Device	Snapshot	Size (GiB)	Volume Type	IOPS	Throughput (MB/s)	Delete on Termination	Encrypted	
Root	/dev/xvda	snap-04147f166cc0f7214	8	General Purpose SSD (GP2)	100 / 3000	N/A	☑	Not Encrypted	
EBS	/dev/sdb	Search (case-insensit	8	Magnetic	N/A	N/A	☐	☐	⊗
EBS	/dev/sdc	Search (case-insensit	500	Throughput Optimized HDD (ST1)	N/A	20 / 123	☐	☐	⊗
EBS	/dev/sdd	Search (case-insensit	500	Cold HDD (SC1)	N/A	6 / 40	☐	☐	⊗

Add New Volume

Free tier eligible customers can get up to 30 GB of EBS General Purpose (SSD) or Magnetic storage. Learn more about free usage tier eligibility and usage restrictions.





Snapshot

- Volumes exist on EBS:
  - Virtual Hard Disk
- Snapshots exist on S3.
- 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 this is your first snapshot, it may take some time to create.

- 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 AMI's from EBS-backed Instances 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 volume from one AZ/Region to another, take a snap or an image of it, then copy it to the new AZ/Region

Snapshot

- 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.