**Amazon Elastic Compute Cloud (EC2):**

It allows users to rent virtual computers on which to run their own computer applications.

**AMI – Amazon Machine Images:**

It is the operating system which is installed in you VM.

***Types:***

* Amazon Maintained
* Marketplace Maintained
* Your own machine images

**Compute Instance types:**

1. General purpose – T3, T2, M5a, M5 M4 M5d
2. Compute optimized – 21d, C5, C4, C5d
3. Storage and I/O optimized – I3, H1, D2
4. Memory Optimized – X1, Z1d, R4, R5, R5d, X1c
5. GPU Graphics – G3, EG, G2
6. GPU and FPGA compute – F1, P3, P2, G3

***General Purpose:***

* Ideal for low traffic websites and web applications.
* Hosting Small and mid-size databases.

***Compute Optimized:***

* High Performance front end Fleets
* Video encoding

***Memory Optimized:***

* High performance Databases
* Distributed memory cache

***Storage Optimized:***

* Used for Data warehousing
* Log and data processing apps

***GPU instances:***

* 3D app streaming
* Machine learning

**Instance Type:**

***M5d.xlarge***

M – Instance family

5 – Instance generation

D – Additional capabilities

xlarge – Instance size

***On – Demand Instances:***

* Pay-as-you-go instances
* You can pay whenever you need
* Provides Scalability, Stability and Flexibility
* When we don’t know how long we need the instances, use this
* For short term projects

***Reserved Instances:***

* Same as On-Demand instances but we can reserve this for one or two years.
* Standard – Use instance within one family and one OS.
* Convertible – Use instance from different family and different OS.

***Spot Instances:***

* Borrowing Some Extra capacity.
* Unreliable
* Useful for stateless, containerized apps for short term.