Contents

[Introduction 1](#_Toc67549676)

[**EC2 Comparison with our Desktop/Laptop** 1](#_Toc67549677)

[EC2 Options 2](#_Toc67549678)

[Instance Type 3](#_Toc67549679)

[**Elastic Block Storage (EBS)** 4](#_Toc67549680)

[**EBS Volume Types:** 6](#_Toc67549681)

[**Amazon Machine Image (AMI)** 7](#_Toc67549682)

[**Key Pairs** 8](#_Toc67549683)

[**Elastic IP and Dynamic IP** 8](#_Toc67549684)

[**Security Groups** 9](#_Toc67549685)

[**Putty & Putty KeyGen (Windows)** 9](#_Toc67549686)

[**Elastic Load Balancers** 9](#_Toc67549687)

[**Application Load Balancers** 10](#_Toc67549688)

[Network Load Balancer 10](#_Toc67549689)

[**Classic Load Balancer** 11](#_Toc67549690)

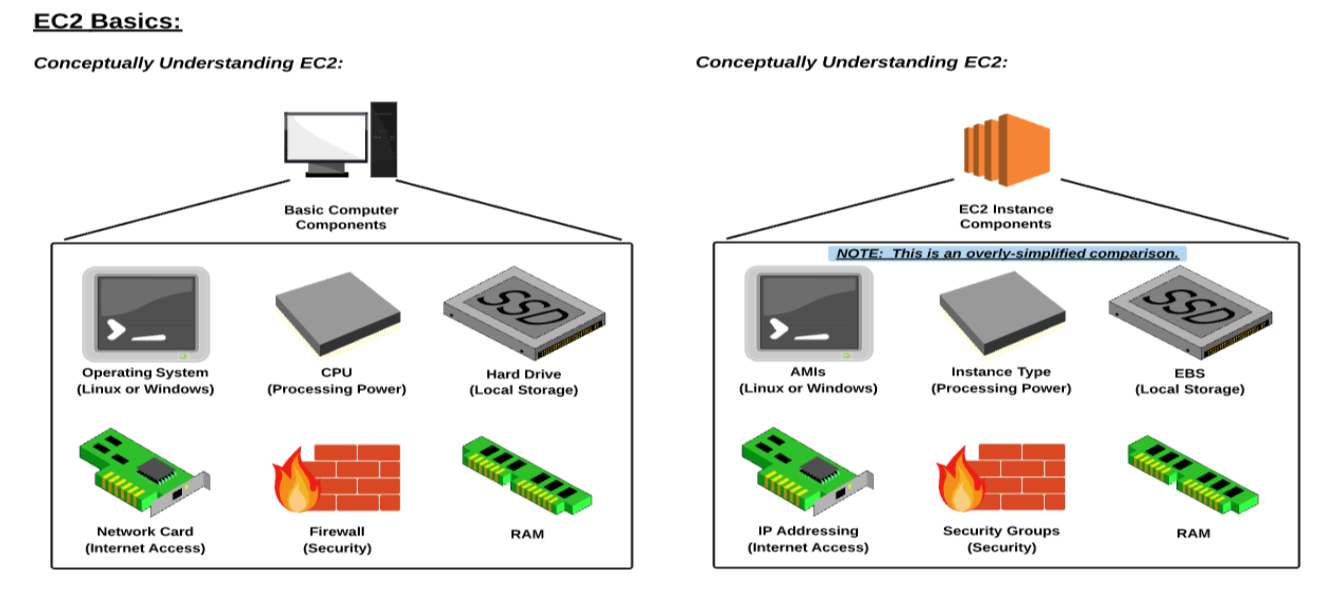
[EFS 11](#_Toc67549691)

[**EC2 Auto Scaling** 11](#_Toc67549692)

# Introduction

* Think of EC2 as your basic desktop computer
* Amazon Elastic Compute Cloud (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.
* Configure Security and Networking and Manage Storage

# **EC2 Comparison with our Desktop/Laptop**

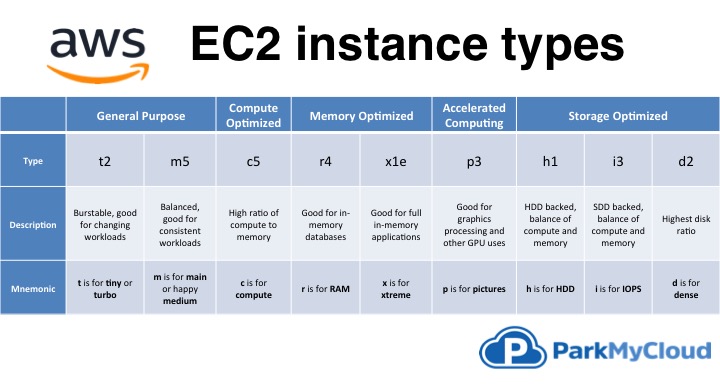


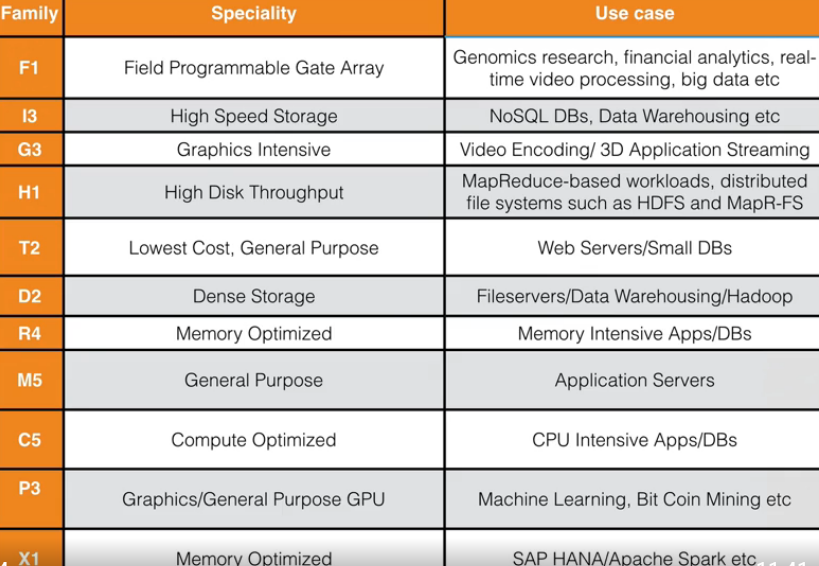
# EC2 Options

* **On Demand** – allows you to pay a fixed rate by the hour (or by the second) with no commitment.
* 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** – provides you with a capacity reservation, and offer a significant discount on the hourly charge for an instance. 1 year or 3 year terms.
* 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 (Upto 75% off on-demand)
  + Convertible Ris (Upto 54% off on-demand)
  + Scheduled Ris are available to launch within the time window you reserve.
* **Spot** – enables you to bid what ever price you want for instance capacity, providing for even greater savings if your applications have flexible start and end times.
* 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
* If a Spot instance is terminated by Amazon EC2, you will not be charged for a partial hour of usage. However, if you terminate the instance yourself, you will be charged for the complete hour in which the instance ran.
* **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.
* 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 upto 70% off the On-Demand price.

# Instance Type

* The instance type is the CPU (Compute Power) of your instance. Each instance type offers different compute, memory and storage capabilities





# **Elastic Block Storage (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.
* Virtual Disk in the Cloud







# **EBS Volume Types:**

* **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 GB and above.
* **Provisioned IOPS SSD**
  + Designed for IO intensive applications such as large relational or NoSQL databases.
  + Use if you need more than 10,000 IOPS
  + Can provision up to 20,000 IOPS per volume.
* **Thoughput Optimized HDD (ST1) (Magnetic)**
  + 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.

# **Amazon Machine Image (AMI)**

* You can select your AMI based on
  + Region
  + 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.



# **Key Pairs**

* A Key pair consists of a public key that AWS stores, and a private key file that you stores. Together they allow you to connect to your instance securily.
* For Linux AMIs, the private key file allows you to securily SSH into your instance.
* For Windows AMIs, the private key file is required to obtain the password used to log into your instance.

# **Elastic IP and Dynamic IP**

* Elastic IP address is a public **static IPv4 address** which is reachable from the Internet.
* Elastic IP addresses are used by AWS to manage its dynamic cloud computing services.
* *Elastic IP to an Instance which doesn't change after you stop / start the instance Whereas Dynamic IP will get changed whenever you start and Stop the EC2 Instance*
* In short Elastic IP is a permanent IP for your instance and Dynamic IP is a temporary for your instance

# **Security Groups**

* A Security Group is a Virtual Firewall
* 1 instance can have multiple Security Groups
* Inbound and Outbound rules
* Stateful, anything you added to inbound rules, outbound rules are added automatically
* 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 cannot block specific IP addresses using Security Groups, instead use Network Access Control Lists.
* You can specify allow rules, but not deny rules.

# **Putty & Putty KeyGen (Windows)**

* Putty won’t support pem files
* PuttyKeyGen is to convert pem files into ppk files
* Download Putty and PuttyKeyGen (<https://www.puttygen.com/download-putty>)
* Open PuttyKeyGen
* Load pem file
* Save Private Key
* Open Putty
* user@Public IP address of EC2 Instance
* SSH 🡪 Auth – Refer Private key
* Click on Open

# **Elastic Load Balancers**

* Elastic Load Balancing automatically distributes incoming application traffic across multiple targets, such as Amazon EC2 instances, containers, IP addresses, and Lambda functions.
* It can handle the varying load of your application traffic in a single Availability Zone or across multiple Availability Zones.
* Elastic Load Balancing offers three types of load balancers that all feature the high availability, automatic scaling, and robust security necessary to make your applications fault tolerant.
  + Application Load Balancer
  + Network Load Balancer
  + Classic Load Balancer – Not using now (It is of Previous Generation)
* A Virtual appliance/Virtual machine balances the load of your web applications, Balances the load across Web servers

## **Application Load Balancers**

* Application Load Balancer is best suited for load balancing of HTTP and HTTPS traffic and provides advanced request routing targeted at the delivery of modern application architectures, including microservices and containers. Operating at the individual request level (Layer 7)
* Application Load Balancer routes traffic to targets within Amazon Virtual Private Cloud (Amazon VPC) based on the content of the request.
* Key features
  + IP addresses as Targets
  + Lambda functions as targets
  + Content Based routing
  + Sticky sessions
  + Delete Protection
  + Logging

## Network Load Balancer

* Network Load Balancer operates at the connection level (Layer 4), routing connections to targets - Amazon EC2 instances, microservices, and containers – within Amazon Virtual Private Cloud (Amazon VPC) based on IP protocol data.
* Ideal for load balancing of both TCP and UDP traffic, Network Load Balancer is capable of handling millions of requests per second while maintaining ultra-low latencies.
* Network Load Balancer is optimized to handle sudden and volatile traffic patterns while using a single static IP address per Availability Zone. It is integrated with other popular AWS services such as Auto Scaling, Amazon EC2 Container Service (ECS), Amazon CloudFormation and AWS Certificate Manager (ACM).
* Use for extreme performance!

## **Classic Load Balancer**

* Are the legacy ELBs. You can load balance HTTP/HTTPS applications and use Layer 7-specific features, such as x-forwarded and sticky sessions. You can also use strict Layer 4 Load Balancing for applications that rely purely on the TCP protocol

# EFS

* EFS is a File storage service for EC2 instances. 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.
* EC2 Auto SCalingSupports the NFS version 4 Protocol
* You only pay for the storage you use (no pre-provisioning required)
* Can scale up to the petabytes
* Can support thousands of concurrent NFS connections
* Data is stored across multiple AZs within a region
* Block based storage
* Read after Write consistency

# **EC2 Auto Scaling**

