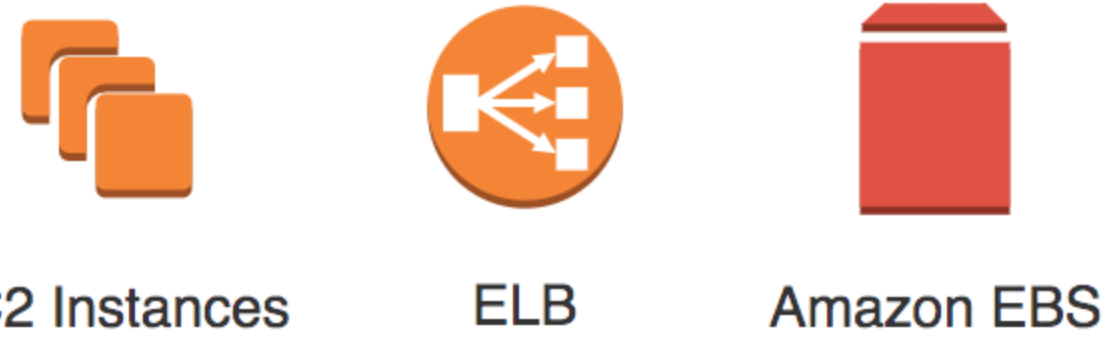
EC2 Fundamentals

EC2(Elastic Compute Cloud)

- In corporate data centers, applications are deployed to physical servers
- Where do you deploy applications in the cloud?
  - Rent virtual servers
  - **EC2 instances** - Virtual servers in AWS (billed by second)
  - **EC2 service** - Provision EC2 instances or virtual servers

EC2 Features



- Create and manage lifecycle of EC2 instances
- **Load balancing** and **auto scaling** for multiple EC2 instances
- **Attach storage** (& network storage) to your EC2 instances
- Manage **network connectivity** for an EC2 instance
- **Our Goal**:
  - Setup EC2 instances as HTTP Server
  - Distribute load with Load Balancers

EC2 Hands-on

- Let's create a few EC2 instances and play with them
- Let's check out the lifecycle of EC2 instances
- Let's use EC2 Instance Connect to SSH into EC2 instances

EC2 Instance Types

- Optimized combination of **compute(CPU, GPU), memory, disk (storage) and networking** for specific workloads
- 270+ instances across 40+ instance types for different workloads
- **t2.micro**:
  - **t** - Instance Family
  - **2** - generation. Improvements with each generation.
  - **micro** - size. *(nano < micro < small < medium < large < xlarge < .....)*
- (Remember) As size increases, compute(CPU, GPU), memory and networking capabilities increase proportionately

EC2 - Instance Metadata Service and Dynamic Data

Instance Metadata Service:

- Get details about EC2 instance **from inside** an EC2 instance:
  - AMI ID, storage devices, DNS hostname, instance id, instance type, security groups, IP addresses etc
- URL: ***http://169.254.169.254/latest/meta-data/***
- URL Paths: network, ami-id, hostname, local-hostname, local-ipv4 , public-hostname, public-ipv4, security-groups, placement/availability-zone

Dynamic Data Service:

- Get dynamic information about EC2 instance:
- URL: ***http://169.254.169.254/latest/dynamic/***
- Example: ***http://169.254.169.254/latest/dynamic/instance-identity/document***

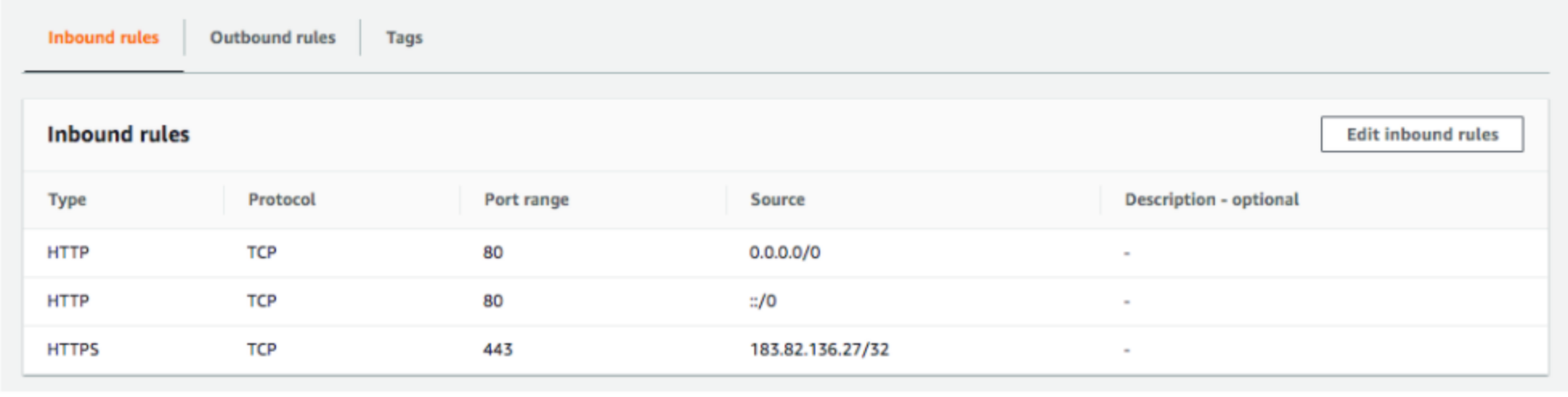
EC2 Hands-on : Setting up a HTTP server

```
sudo su
yum update -y
yum install httpd -y
systemctl start httpd
systemctl enable httpd
echo "Getting started with AWS" > /var/www/html/index.html
echo "Welcome to in28minutes $(whoami)" > /var/www/html/index.html
echo "Welcome to in28minutes $(hostname)" > /var/www/html/index.html
echo "Welcome to in28minutes $(hostname -i)" > /var/www/html/index.html
```

Security Groups

- **Virtual firewall** to control **incoming and outgoing** traffic to/from AWS resources (EC2 instances, databases etc)
- Provides additional layer of security - Defense in Depth

Security Groups Rules



- Security groups are **default deny**
  - If there are no rules configured, no outbound/inbound traffic is allowed
- You can specify **allow rules ONLY**
- You can configure **separate rules** for inbound and outbound traffic
- You can assign multiple (upto five) security groups to your EC2 instances

Security Groups

- You can add and delete security groups to EC2 instances at any time.
  - Changes are immediately effective
- Traffic NOT explicitly allowed by Security Group **will not reach** the EC2 instance
- Security Groups are **stateful**:
  - If an outgoing request is allowed, the incoming response for it is automatically allowed.
  - If an incoming request is allowed, an outgoing response for it is automatically allowed

Security Group - Trivia

- What if there are no security group rules configured for inbound and outbound?
  - Default DENY. No traffic is allowed in and out of EC2 instance.
- Can I change security groups at runtime?
  - Yes. Changes are immediately effective.

EC2 IP Addresses

- Public IP addresses are internet addressable.
- Private IP addresses are **internal** to a corporate network
- You CANNOT have two resources with same public IP address.
- HOWEVER, two different corporate networks CAN have resources with same private IP address
- **All EC2 instances** are assigned private IP addresses
- Creation of public IP addresses **can be enabled** for EC2 instances in public subnet
- (Remember) When you stop an EC2 instance, public IP address is lost
- **DEMO**: EC2 public and private addresses

Elastic IP Addresses

- Scenario : How do you get a **constant public IP address** for a EC2 instance?
  - Quick and dirty way is to **use an Elastic IP**!
- **DEMO**: Using Elastic IP Address with an EC2 instance

Elastic IP Addresses - Remember

- Elastic IP can be switched to another EC2 instance **within the same region**
- Elastic IP **remains attached** even if you stop the instance. You have to manually detach it.
- Remember : You are charged for an Elastic IP when you are NOT using it! Make sure that you explicitly release an Elastic IP when you are not using it
- You will be charged for Elastic IP when:
  - Elastic IP is NOT associated with an EC2 instance OR
  - EC2 instance associated with Elastic IP is stopped

Simplify EC2 HTTP server setup

- How do we reduce the number of steps in creating an EC2 instance and setting up a HTTP Server?
- Let's explore a few options:
  - Userdata
  - Launch Template
  - AMI

Bootstrapping with Userdata

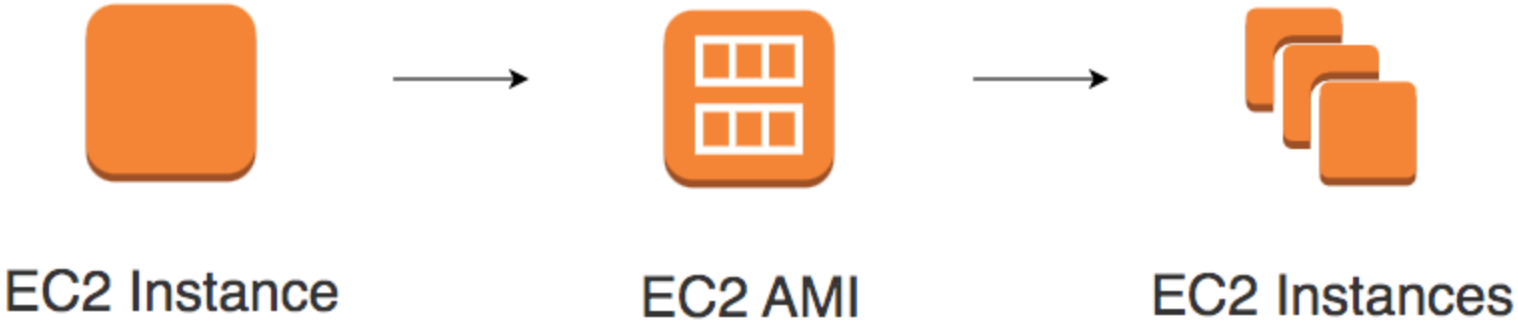
```
#!/bin/bash
yum update -y
yum install -y httpd
systemctl start httpd
systemctl enable httpd
curl -s http://169.254.169.254/latest/dynamic/instance-identity/document > /var/www/html/ir
```

- **Bootstrapping**: Install OS patches or software when an EC2 instance is launched.
- In EC2, you can configure **userdata** to bootstrap
- Lookup user data - ***http://169.254.169.254/latest/user-data/***
- **DEMO** - Using Userdata

Launch Templates

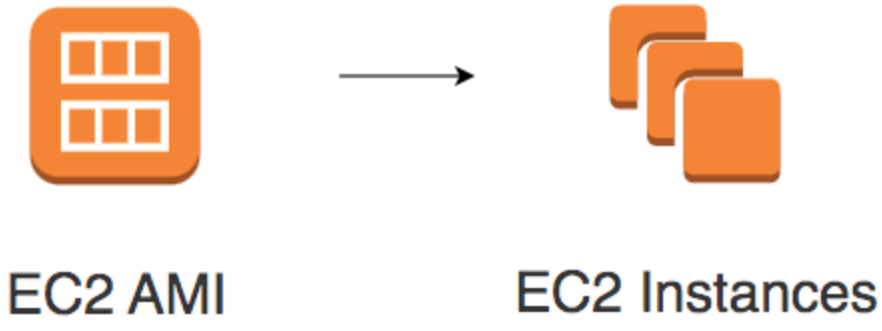
- Why do you need to specify all the EC2 instance details (AMI ID, instance type, and network settings) **every time** you launch an instance?
- How about creating a **Launch Template**?
- Allow you to launch Spot instances and Spot fleets as well
- **DEMO** - Launch EC2 instances using Launch Templates

Reducing Launch Time with Customized AMI



- Installing OS patches and software using userdata at launch of EC2 instances **increases boot up** time
- How about creating customized AMIs with OS patches and software **pre-installed**?
  - **Hardening an Image** - Customize EC2 images to your corporate security standards
- **Prefer** using Customized AMI to userdata
- **DEMO** : Create a Customized AMI and using it in Launch Template

AMI - Amazon Machine Image



- What operating system and what software do you want on the instance?
- Three AMI sources:
  - Provided by AWS
  - **AWS Market Place**: Online store for customized AMIs. Per hour billing
  - **Customized AMIs**: Created by you.

EC2 Amazon Machine Image - AMI - Remember

- AMIs contain:
  - Root volume block storage (OS and applications)
  - Block device mappings for non-root volumes
- You can configure launch permissions on an AMI
  - Who can use the AMI?
  - You can **share your AMIs** with other AWS accounts
- AMIs are stored in Amazon S3 (**region specific**).
- **Best Practice**: Backup upto date AMIs in multiple regions
  - Critical for Disaster Recovery

EC2 Security - Key Pairs

- EC2 uses public key cryptography for protecting login credentials
- Key pair - public key and a private key
  - Public key is stored in EC2 instance
  - Private key is stored by customer

Connecting to EC2 instance(s) - Troubleshooting

- You need to have the **private key** with you
- Change permissions to **0400** (*chmod 400 /path/my-key-pair.pem*)
  - Default permissions on private key - 0777 (**VERY OPEN**)
- (Windows Instances) In addition to private key, you need admin password
  - (At Launch) Random admin password is generated and encrypted using public key
  - Decrypt the password using the private key and use it to login via RDP
- Security Group should allow inbound SSH or RDP access:
  - Port 22 - Linux EC2 instance (SSH)
  - Port 3389 - RDP (Remote Desktop - Windows)
- Connect to your instance using its Public DNS: `ec2-**-**-**-**.compute.amazonaws.com`

Important EC2 Scenarios - Quick Review

Scenario	Solution
You want to identify all instances belonging to a project, to an environment or to a specific billing type	Add Tags. Project - A. Environment - Dev
You want to change instance type	Stop the instance. Use "Change Instance Type" to change and restart.
You don't want an EC2 instance to be automatically terminated	Turn on Termination Protection. (Remember) EC2 Termination Protection is not effective for terminations from a) Auto Scaling Groups (ASG) b) Spot Instances c) OS Shutdown
You want to update the EC2 instance to a new AMI updated with latest patches	Relaunch a new instance with an updated AMI
Create EC2 instances based on on-premise Virtual Machine (VM) images	Use VM Import/Export. You are responsible for licenses.

Important EC2 Scenarios - Quick Review

Scenario	Solution
Change security group on an EC2 instance	Assign at launch or runtime. Security Group changes are immediately effective.
You get a timeout while accessing an EC2 instance	Check your Security Group configuration
You are installing a lot of software using user data slowing down instance launch. How to make it faster?	Create an AMI from the EC2 instance and use it for launching new instances
I've stopped my EC2 instance. Will I be billed for it?	ZERO charge for a stopped instance (If you have storage attached, you have to pay for storage)

Quick Review

AMI

- What operating system and what software do you want on the instance?
- Reduce boot time and improve security by creating customized hardened AMIs.
- Region specific.
- Backup AMIs in multiple regions.
- You can share AMIs with other AWS accounts.

EC2 Instance Types

- Optimized combination of compute(CPU, GPU), memory, disk (storage) and networking for specific workloads.

Quick Review

Security Groups

- Virtual firewall to control incoming and outgoing traffic to/from AWS resources (EC2 instances, databases etc)
- Default deny. Separate allow rules for inbound and outbound traffic
- Stateful and immediately effective

Key Pairs

- Public key cryptography (Key Pairs) used to protect your EC2 instances
- You need private key with right permissions (chmod 400) to connect to your EC2 instance. (Windows EC2 instances only) You need admin password also.
- Security group should allow SSH(22) or RDP(3389)

Quick Review

- **Instance Metadata Service** - Get details about EC2 instance from inside an EC2 instance. ***http://169.254.169.254/latest/meta-data/***
- **Userdata** - Used for bootstrapping. Install OS patches or software when an EC2 instance is launched.
- **Elastic IP Addresses** - Static public IP address for EC2 instance.
- **Launch Templates** - Pre-configured templates (AMI ID, instance type, and network settings) simplifying the creation of EC2 instances.