For Further Questions?

What is Cloud Computing

Ethan's
TRAINING | RECRUITMENT | CONSULTANCY

Wikipedia defines cloud computing as:

"Cloud computing is internet-based computing in which large groups of remote servers are networked to allow the centralized data storage, and online access to computer services or resources.“

The **National Institute of Standards and Technology** (**NIST**) gives the following definition of cloud computing:

"Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction."

What is Cloud Computing

Ethan's
TRAINING | RECRUITMENT | CONSULTANCY

- Users should be able to provision and release resources on-demand
- The resources can be scaled up or down automatically, depending on the load
- The provisioned resources should be accessible over a network
- Cloud service providers should enable a pay-as-you-go model, where customers are charged based on the type and quantum of resources they consume

Cloud Service Models

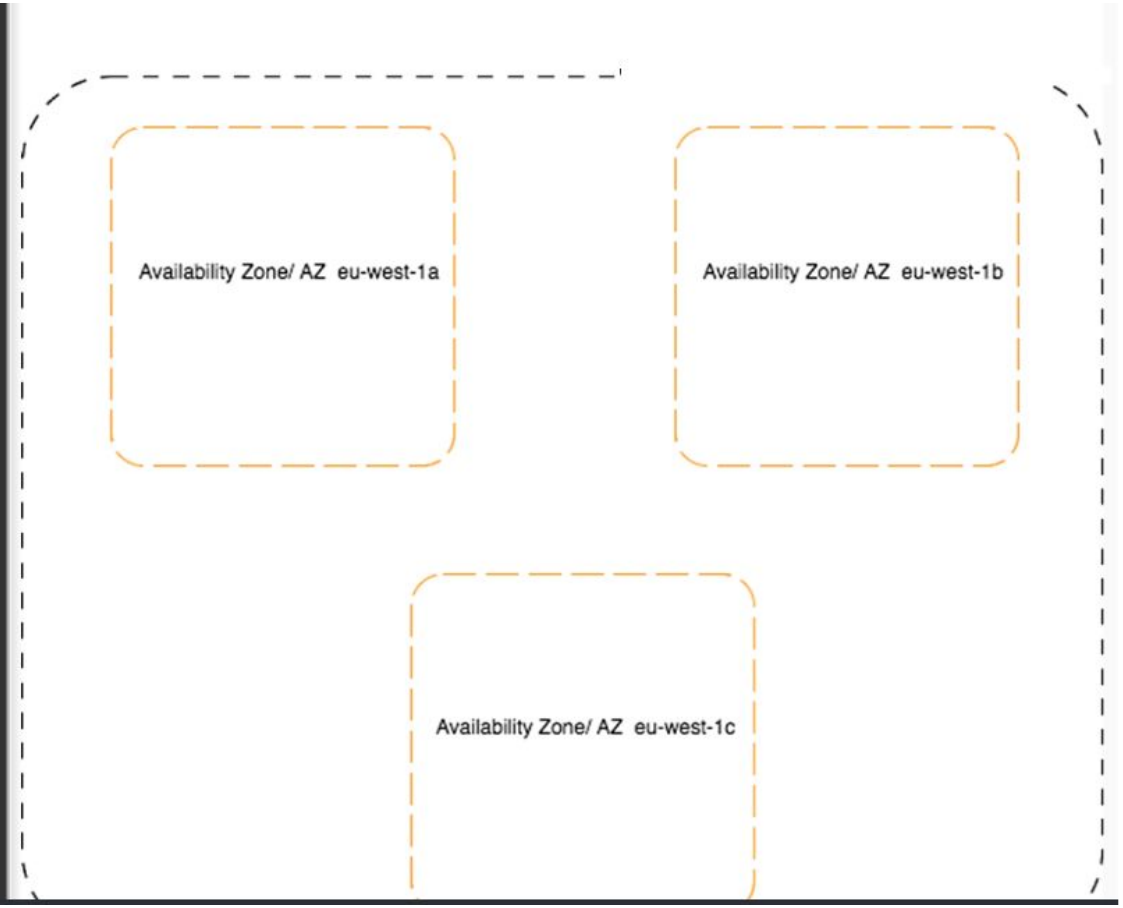
- **Infrastructure as a Service (IaaS)** provides users the capability to provision processing, storage, and network resources on demand.
- In **Platform as a Service(PaaS),** the service provider makes certain core components, such as databases, queues, workflow engines, e-mails, and so on, which are available as services to the customer. The customer then leverages these components for building their own applications.
- In the **Software as a Service(SaaS)** model, typically, third-party providers using a subscription model provide end-user applications to their customers.
- Example Iaas ,PaaS & Saas

Cloud Service Models

Ethan's
TRAINING | RECRUITMENT | CONSULTANCY

On premise	IaaS	PaaS	FaaS	SaaS
Functions	Functions	Functions	Functions	Functions
Application	Application	Application	Application	Application
Runtime	Runtime	Runtime	Runtime	Runtime
Operating system	Operating system	Operating system	Operating system	Operating system
Virtualization	Virtualization	Virtualization	Virtualization	Virtualization
Networking	Networking	Networking	Networking	Networking
Storage	Storage	Storage	Storage	Storage
Hardware	Hardware	Hardware	Hardware	Hardware

- Region:
  - Is a geographical area
- Availability Zone(AZ):
  - Is a data centre
  - 2 or more within a region



- Edge Location
  - Is a CDN endpoint for Cloudfront
  - Used to cache files closer to a user's physical location

Networking and Content Delivery

-  VPC - Virtual data centre in the cloud
-  CloudFront - CDN, used for caching
-  DirectConnect - Dedicated network connection to AWS
-  Route53 - DNS

Compute

-  EC2 - Virtual machine in the cloud
-  ECS - Docker containers
-  Lambda - Serverless
-  Elastic Beanstalk - Web apps without infrastructure code

Storage

-  S3 - Object storage
-  Glacier - Archive
-  Storage Gateway - Hybrid storage
-  EFS - File system

Migration

-  Snowball - Petabyte scale data solution
-  DMS - Database migration <-> AWS, AWS <-> AWS
-  Server Migration (SMS) - Migrate on premise servers to AWS

Database

-  RDS - relational
-  DynamoDB - NoSQL
-  Elasticache - in-memory cache
-  Redshift - data warehouse

Management Tools

-  Cloudwatch - Monitoring & alerting
-  Cloudformation - Infrastructure as code
-  CloudTrail - Audit all API calls
-  OpsWorks - Chef configuration management

Management Tools continued

-  Config - Configuration resources rules
-  Trusted Advisor - Reduce costs and improve security
-  Service Catalog - Organization catalogue

Security, Identity, and Compliance

-  IAM - access control
-  Inspector - agents, security
-  Certificate Manager - SSL/TLS certificates
-  Directory Service - directory store (Active Directory)

Developer Tools

-  CodeCommit - Managed Git source control
-  Code Build - compiles source code, runs tests
-  CodeDeploy - automated deployments on premise & AWS
-  Code Pipeline - CI/CD

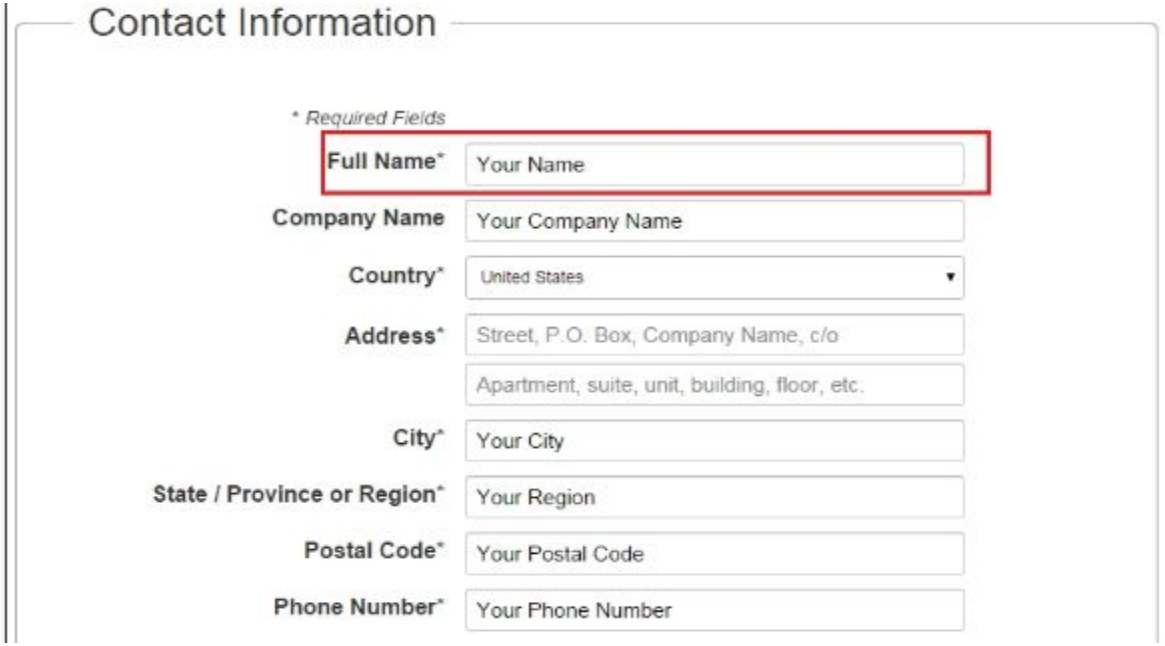
Artificial Intelligence

-  Lex - speech to text, natural language understanding
-  Polly - text to speech
-  Rekognition - image recognition
-  Machine Learning - apply complex algorithms, predictions

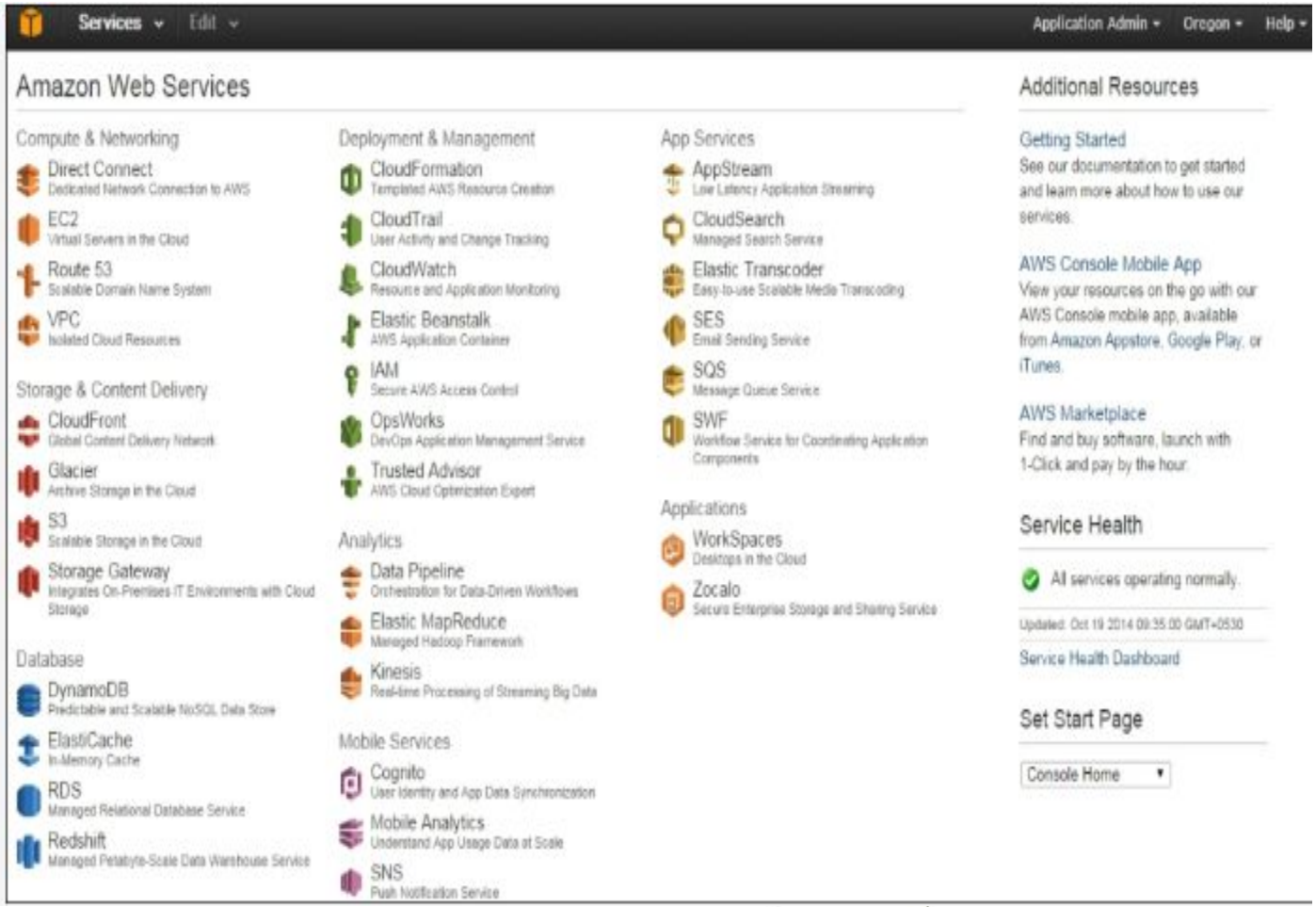
Setting up AWS Account

Point your browser to http://aws.amazon.com/ and click on **Create a Free Account**.

Setting up AWS Account

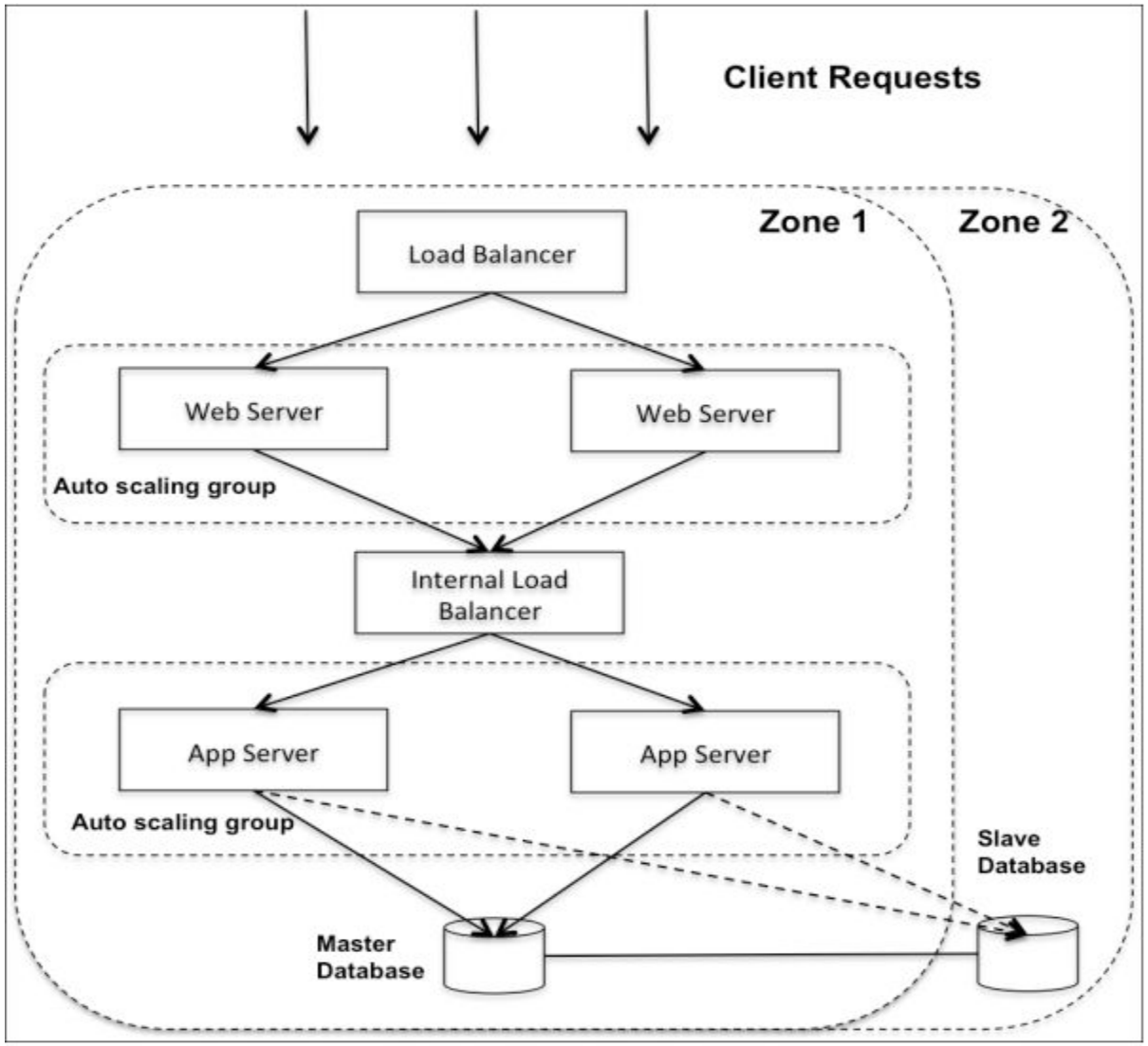


AWS Management Console



Designing Cloud Applications

Multi-tier Architecture



AWS Components

Ethan's
TRAINING | RECRUITMENT | CONSULTANCY

Amazon Elastic Compute Cloud (EC2) :Amazon EC2 is a web service that provides compute capacity in the AWS cloud.

Amazon Elastic Block Storage (EBS): Amazon EBS is highly available and durable persistent block level storage volumes for use with Amazon EC2 instances.

Elastic IP addresses allow you to allocate a static IP address, and programmatically assign it to an instance.

Amazon CloudWatch: We can enable monitoring on EC2 instance using it.

We can also distribute incoming traffic by using the **Elastic Load Balancer** (**ELB**) service.

We can create auto scaling groups using the auto scaling feature to automatically scale your capacity based on CloudWatch.

AWS Components

Ethan's
TRAINING | RECRUITMENT | CONSULTANCY

Amazon S3: Amazon S3 is a highly durable and distributed data store.

Amazon Glacier: Amazon Glacier is low-cost storage service that is typically used for archiving and backups.

Amazon RDS: It provides an easy way to setup, operate, and scale a relational database in the cloud. Database options available from AWS include MySQL, Oracle, SQL Server, PostgreSQL, and Amazon Aurora (in preview at this time)

Amazon DynamoDB : It is a NoSQL database service offered by AWS. It supports both document and key-value pairs, data models, and has a flexible schema.

Amazon ElastiCache: f your application is read-intensive, then you can use the AWS ElastiCache service to significantly boost the performance of your applications. ElastiCache supports Memcached and Redis in-memory caching solutions.

AWS Components

Ethan's
TRAINING | RECRUITMENT | CONSULTANCY

Amazon Simple Queue Service (**Amazon SQS**) is a reliable, highly-scalable, hosted, and distributed queue for storing messages as they travel between computers and application components.

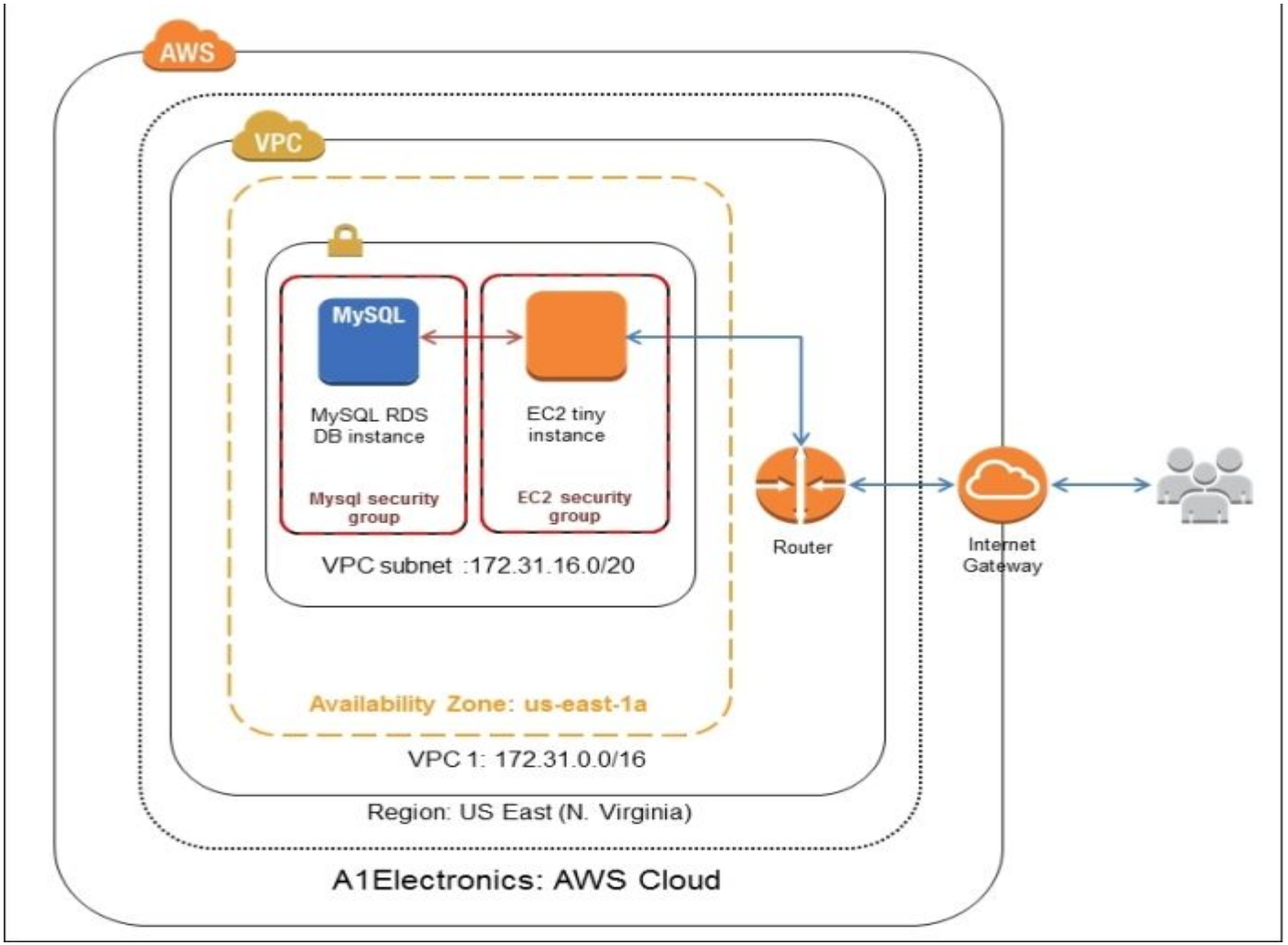
Amazon Virtual Private Cloud (**Amazon VPC**) allows you to extend your corporate network into a private cloud contained within AWS. Amazon VPC uses the IPSec tunnel mode that enables you to create a secure connection between a gateway in your data center and a gateway in AWS.

Amazon Route 53 is a highly-scalable DNS service that allows you to manage your DNS records by creating a hosted zone for every domain you would like to manage.

AWS **Identity and Access Management** (**IAM**) enables you to you to control access to AWS services and resources.

Amazon CloudWatch is a monitoring service for your AWS resources. It enables you to retrieve monitoring data, set alarms, troubleshoot problems, and take actions based on the issues arising in your cloud environment.

Setting up AWS Infrastructure



Setting up AWS Infrastructure

Ethan's
TRAINING | RECRUITMENT | CONSULTANCY

Region: AWS services are hosted in multiple locations around the word and these are known as regions. The regions are connected through the public internet.

Availability Zone: Availability zones (AZ) can be treated as traditional data centers within a region. AZs in the same region are designed to provide infrastructure redundancy in the event of a catastrophic outage, such as earthquakes, snowstorms, Godzilla attacks, and so on. The number of AZs in a region is region specific. In our example, we select the **us-east-1a** AZ.

EC2 Instance: This is a virtual server on which you run your applications. These come in various flavors to meet your computing demand. A high compute EC2 instance also has high network I/O memory associated with it. You cannot have a low compute EC2 instance with high memory and network I/O. EC2 instances have fixed CPU to memory ratios. It is best to select a micro instance for development, since it is free. More on EC2 instance types is available at http://aws.amazon.com/ec2/instance-types/.

Setting up AWS Infrastructure

Ethan's
TRAINING | RECRUITMENT | CONSULTANCY

Security Groups: A security group acts as a virtual firewall for your instance to control inbound and outbound traffic. The security group can be configured by a set of rules for inbound and outbound traffic. The rules define the network protocol, port, and source and destination IP address ranges to accept or send your data to.

Virtual Private Cloud (VPC): VPC lets you provision a private, isolated section of the AWS cloud where you can launch AWS resources in a virtual network, using custom-defined IP address ranges. It is like your own private data centre.

Subnets: Subnets are logical segments of a VPC's address range that allow you to designate to a group of your resources based on security and operational needs

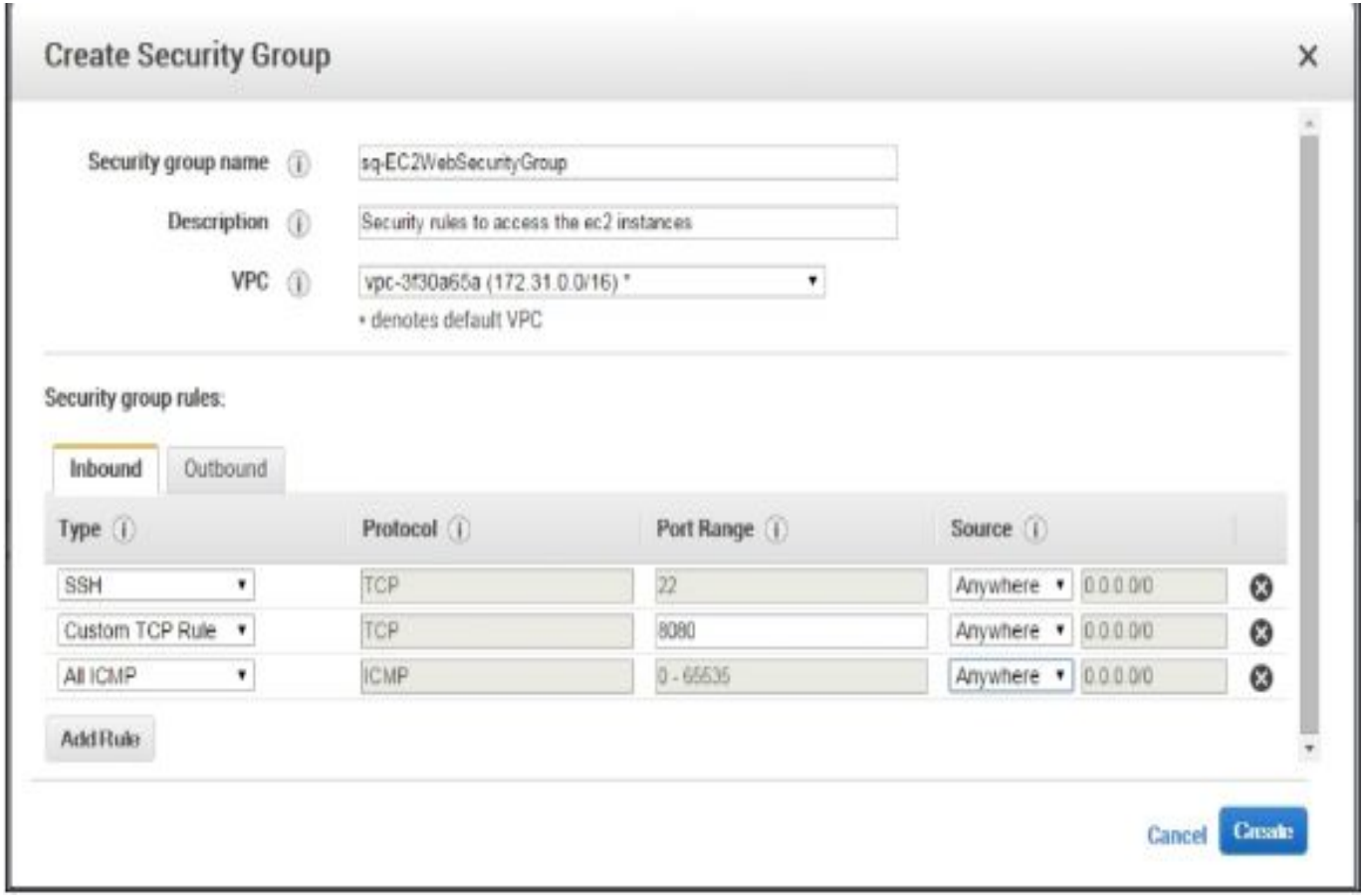
Router: Each VPC comes with a default router in order to communicate with resources outside the VPC. For example, connecting to a database server in other VPCs.

Setting up AWS Infrastructure

Ethan's
TRAINING | RECRUITMENT | CONSULTANCY

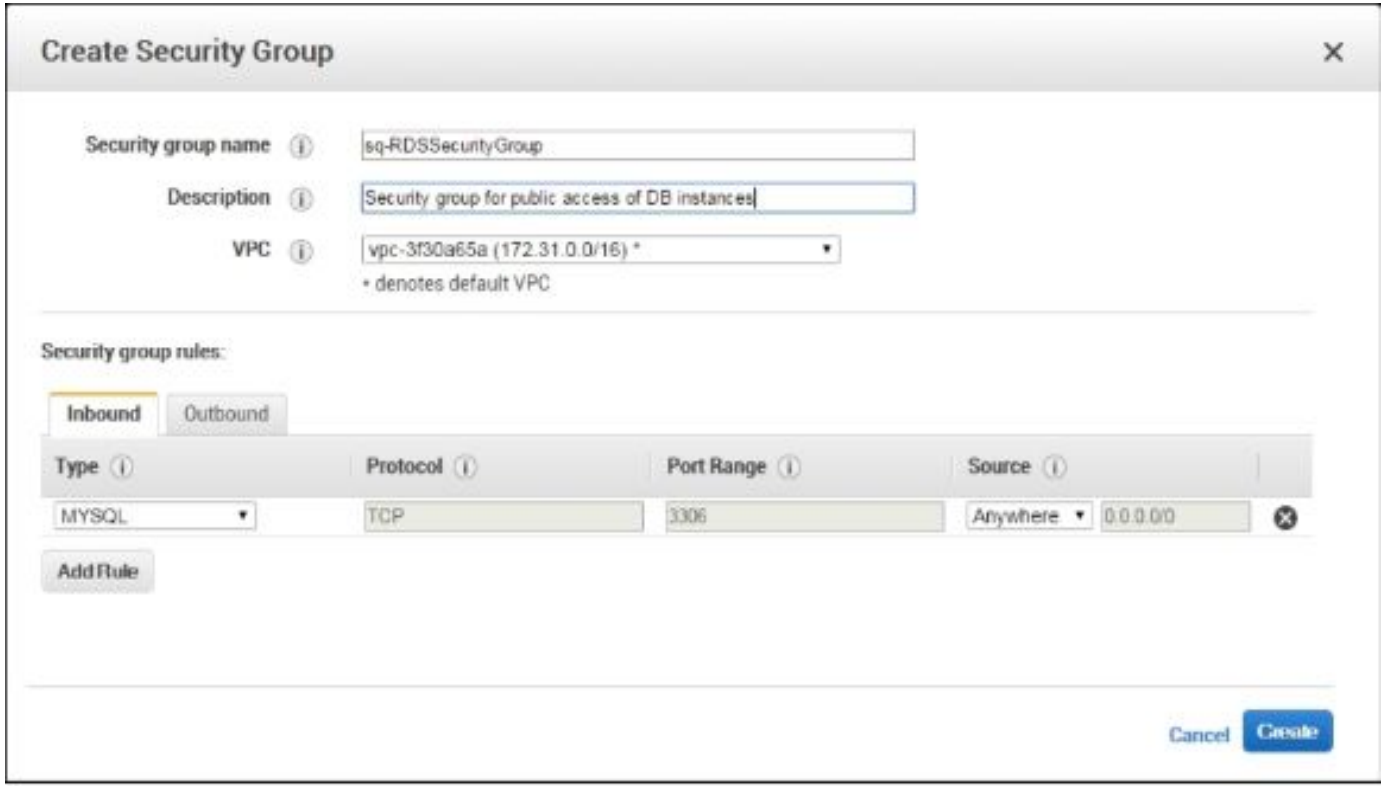
1. From the EC2 dashboard, click on **Security Groups** from the navigation pane and then on the**Create Security Group** button.
2. Create a security group for EC2 instances to allow the following
    1. Web traffic from any IP address on port 8080 (default Tomcat server port)
    2. SSH traffic for remote login from any IP address
    3. ICMP traffic to ping the EC2 instance from a public Internet

Setting up AWS Infrastructure



Setting up AWS Infrastructure

Create a security group for MySQL RDS instances to allow access from the Internet.



Setting up AWS Infrastructure

Ethan's
TRAINING | RECRUITMENT | CONSULTANCY

Creating EC2 instance key pairs

1. From the EC2 dashboard, click on **Key Pairs** from the navigation pane and then on the **Create Key Pair** button.
2. Enter e2accesskey when prompted with a dialog box asking to enter the key pair name. This key pair name will be used while configuring the EC2 instances.

Note - Make sure you select the correct AWS region from the EC2 dashboard to create the keys because key pairs can't be shared across regions

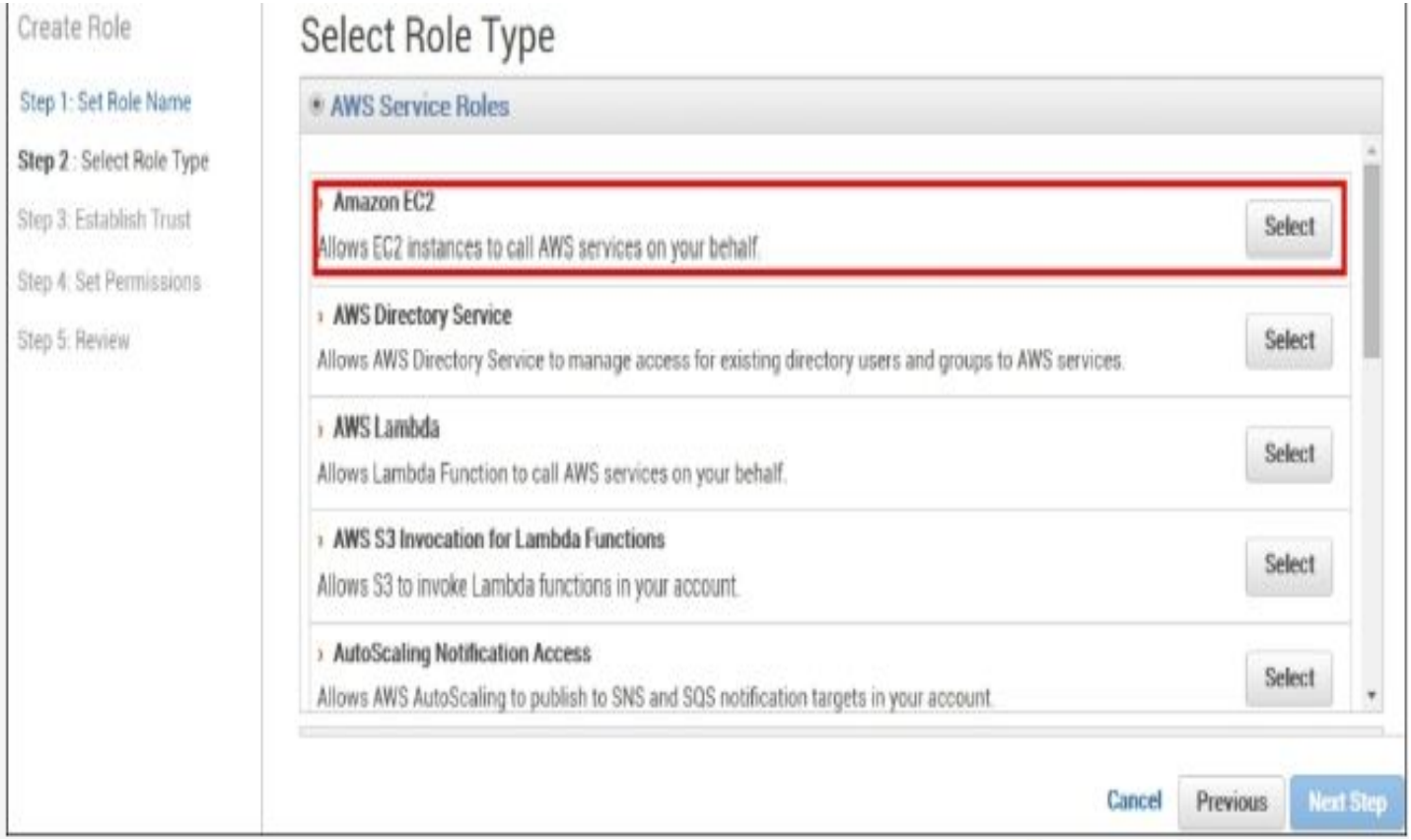
As soon as you create the key pair, your private key will be immediately downloaded to your computer. Secure this private key. This private key file can be only downloaded once during the creation of the keys. You cannot change access keys in your EC2 instances once they have been assigned.

Setting up AWS Infrastructure

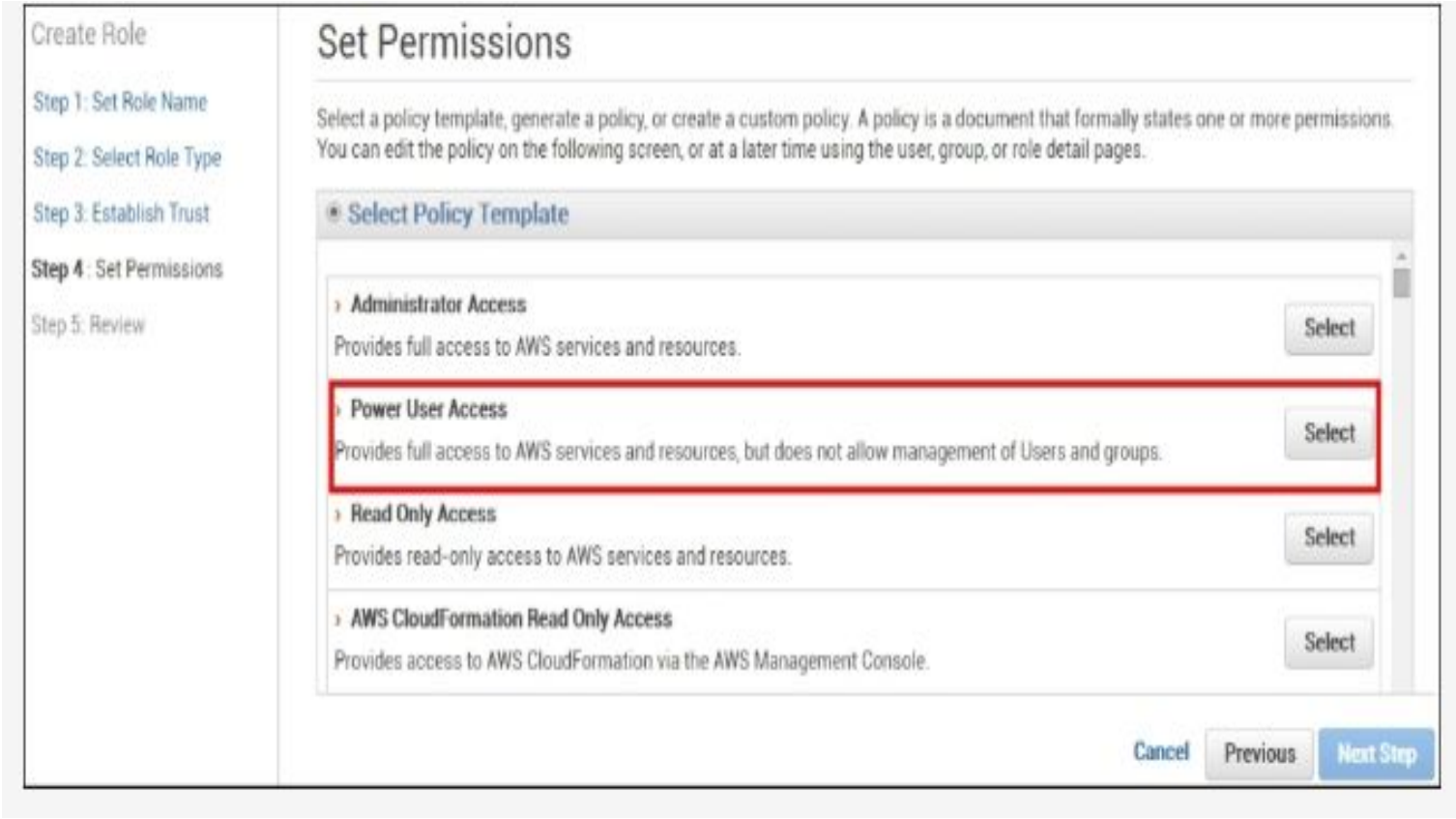
Creating Roles

1. From the IAM dashboard, click on **Roles** in the navigation pane and then on the **Create New Role** button
2. Create a role named ec2Instance for our EC2 instances that have access to all the AWS provided services, as shown in the following screenshot:

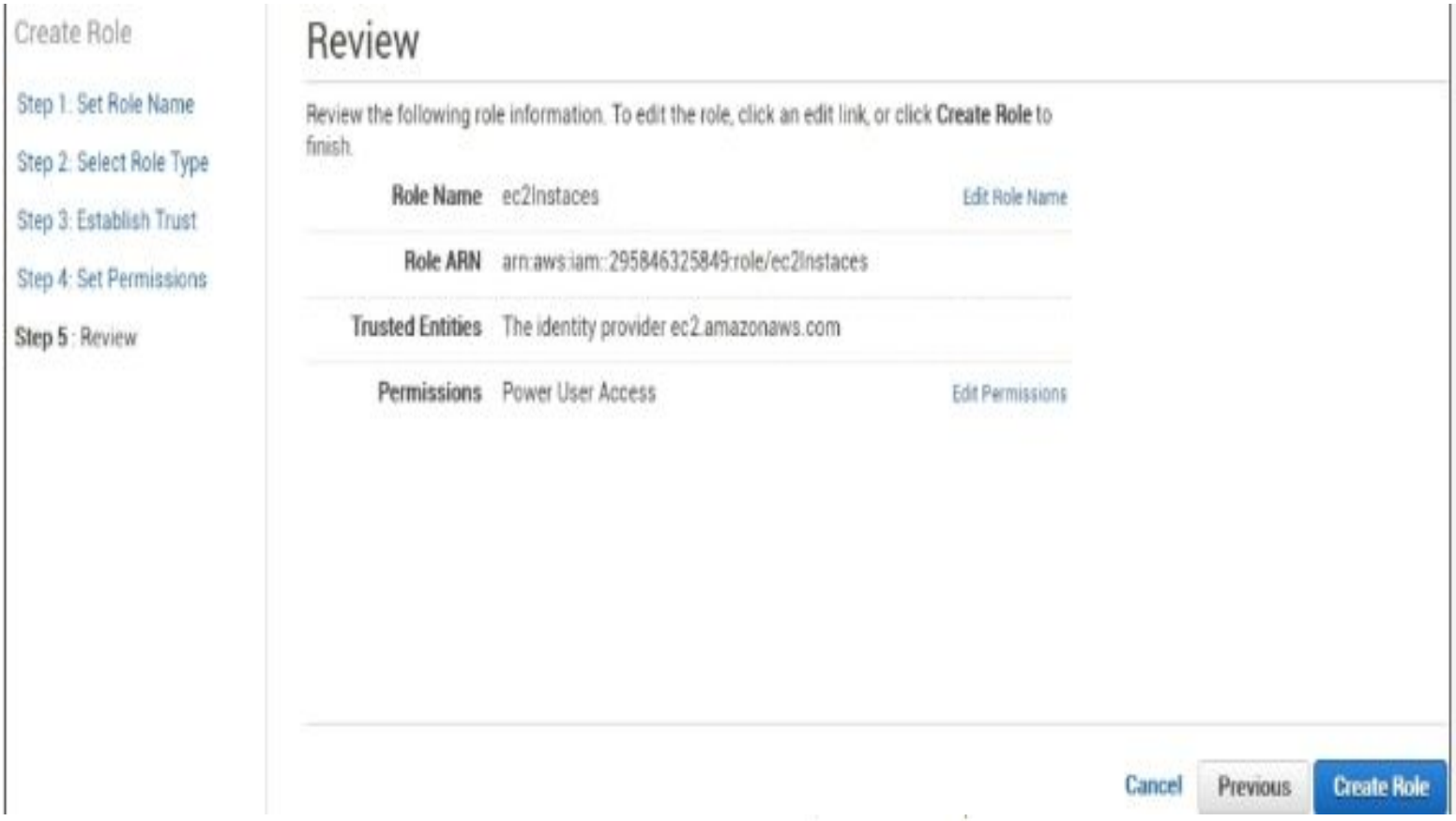
Setting up AWS Infrastructure



Setting up AWS Infrastructure



Setting up AWS Infrastructure

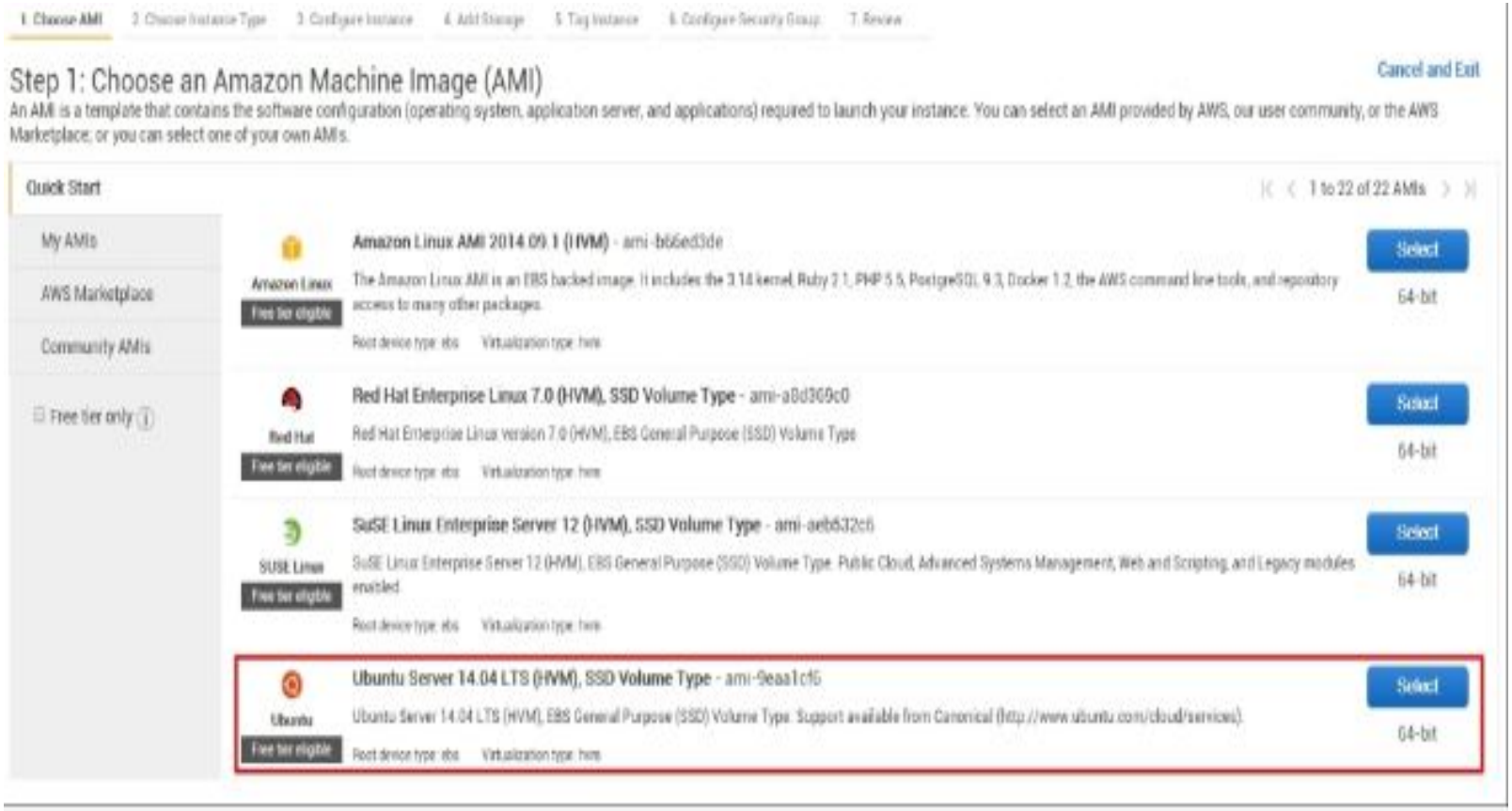


Setting up AWS Infrastructure

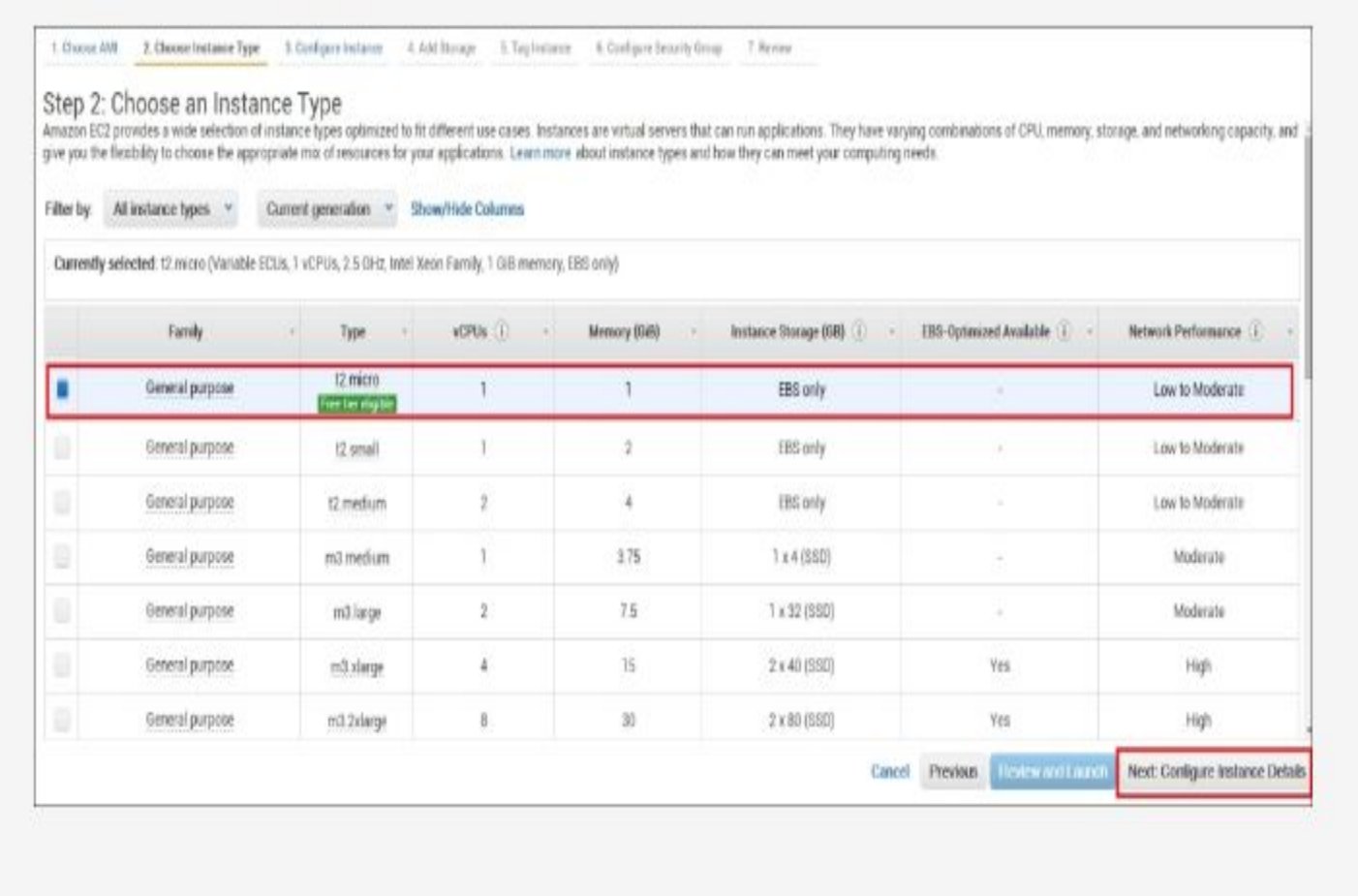
Creating EC2 Instance

1. From the EC2 dashboard, click on **Instances** in the navigation pane and on the **Launch** instance. This will start a process of provisioning an EC2 instance.
2. The next step is to choose an operating system for the EC2 instance; this is done by choosing the correct **Amazon Machine Image** (**AMI**) as per our requirements. Select the**Ubuntu Server 14.04 LTS (HVM) SSD Volume Type** AMI, as shown in the following screenshot:

Setting up AWS Infrastructure



Setting up AWS Infrastructure



1. Choose AMI 2. Choose Instance Type 3. Configure Instance 4. Add Storage 5. Tag Instance 6. Configure Security Group 7. Review

Step 2: Choose an Instance Type

Amazon EC2 provides a wide selection of instance types optimized to fit different use cases. Instances are virtual servers that can run applications. They have varying combinations of CPU, memory, storage, and networking capacity, and give you the flexibility to choose the appropriate mix of resources for your applications. Learn more about instance types and how they can meet your computing needs.

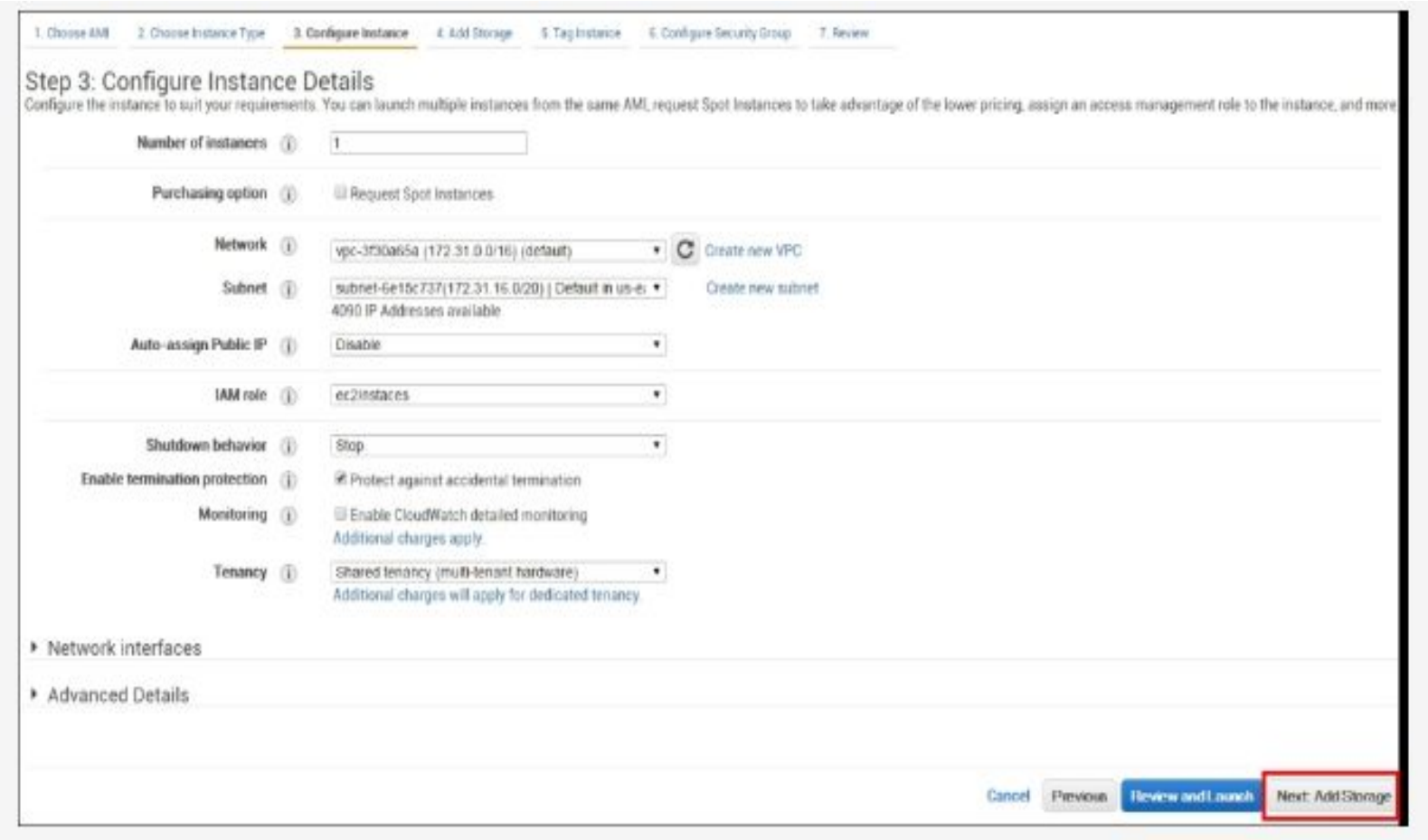
Filter by: All instance types | Current generation | Show/Hide Columns

Currently selected: t2.micro (Variable ECUs, 1 vCPUs, 2.5 GHz, Intel Xeon Family, 1 GiB memory, EBS only)

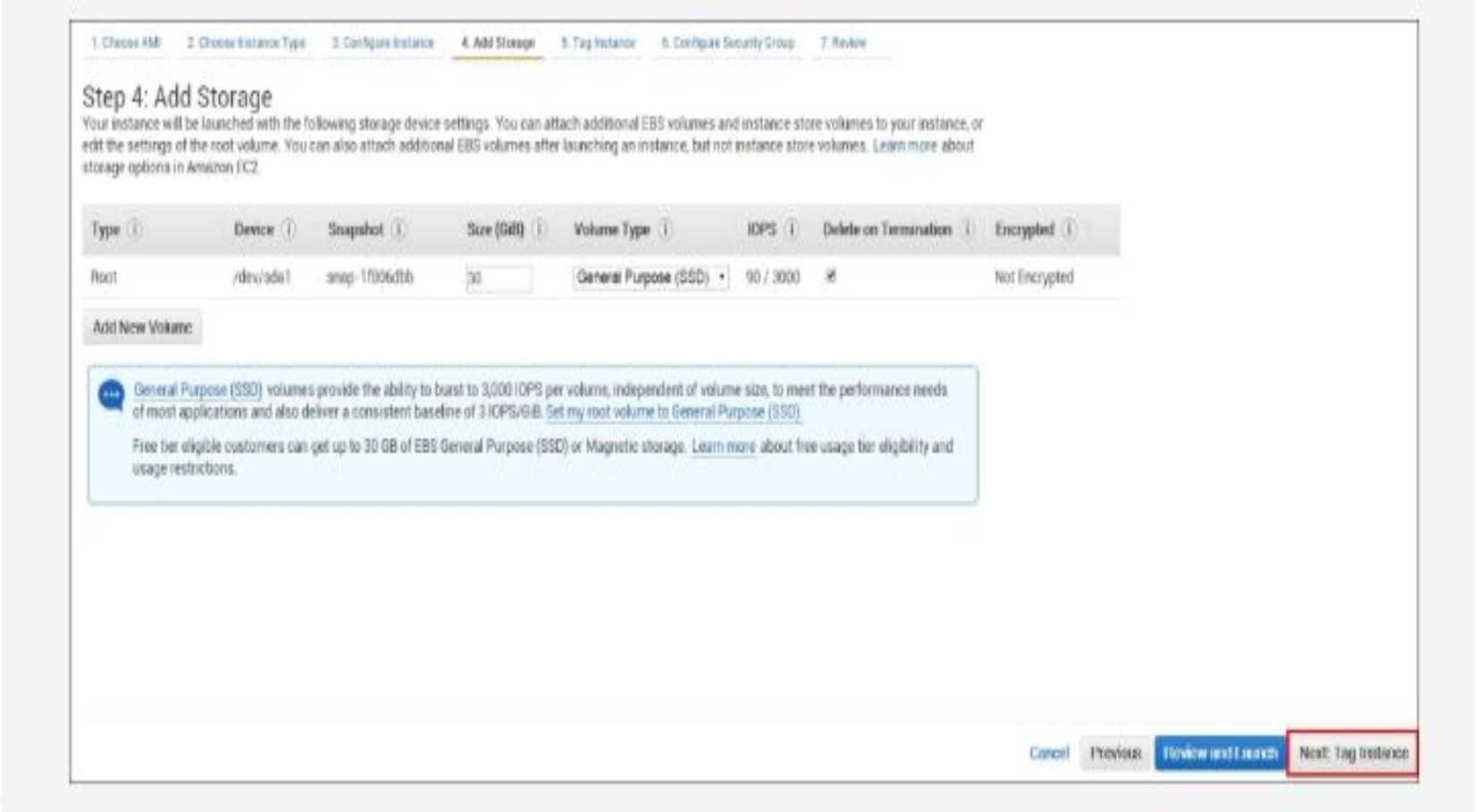
	Family	Type	vCPUs	Memory (GiB)	Instance Storage (GB)	EBS-Optimized Available	Network Performance
☑	General purpose	t2.micro (Free tier eligible)	1	1	EBS only	-	Low to Moderate
☐	General purpose	t2.small	1	2	EBS only	-	Low to Moderate
☐	General purpose	t2.medium	2	4	EBS only	-	Low to Moderate
☐	General purpose	m3.medium	1	3.75	1 x 4 (SSD)	-	Moderate
☐	General purpose	m3.large	2	7.5	1 x 32 (SSD)	-	Moderate
☐	General purpose	m3.xlarge	4	15	2 x 40 (SSD)	Yes	High
☐	General purpose	m3.2xlarge	8	30	2 x 80 (SSD)	Yes	High

Cancel | Previous | Review and Launch | Next: Configure Instance Details

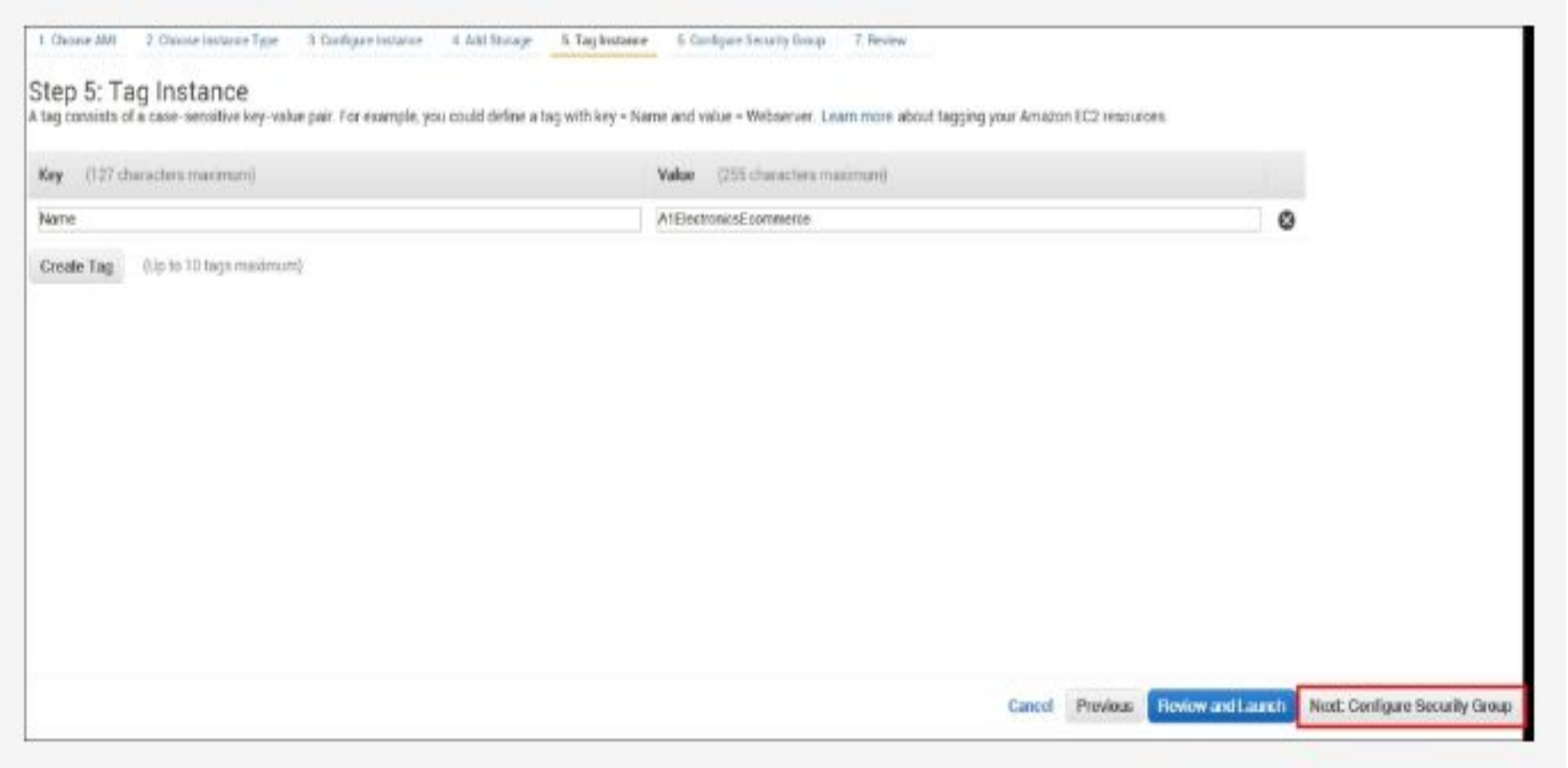
Setting up AWS Infrastructure



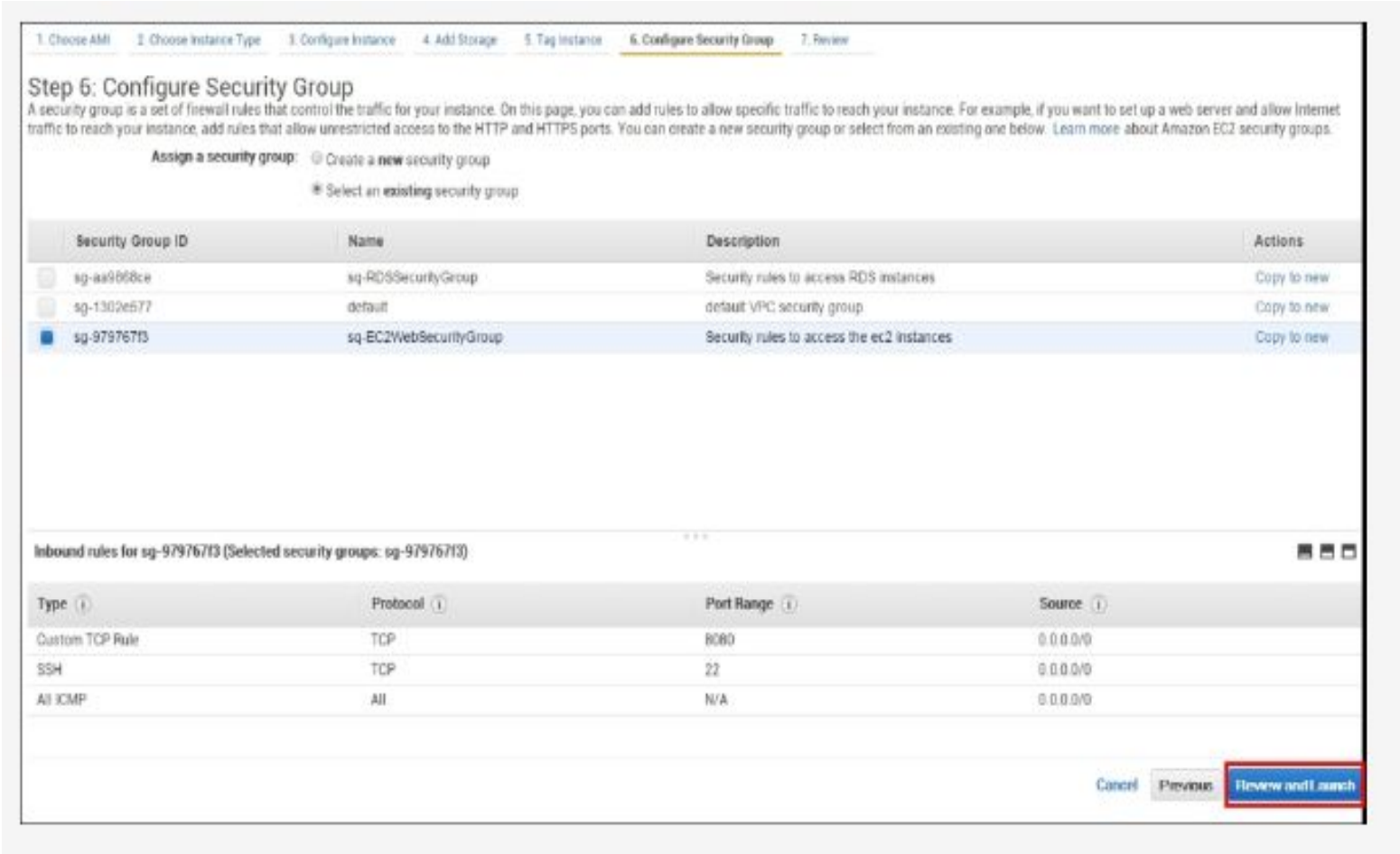
Setting up AWS Infrastructure



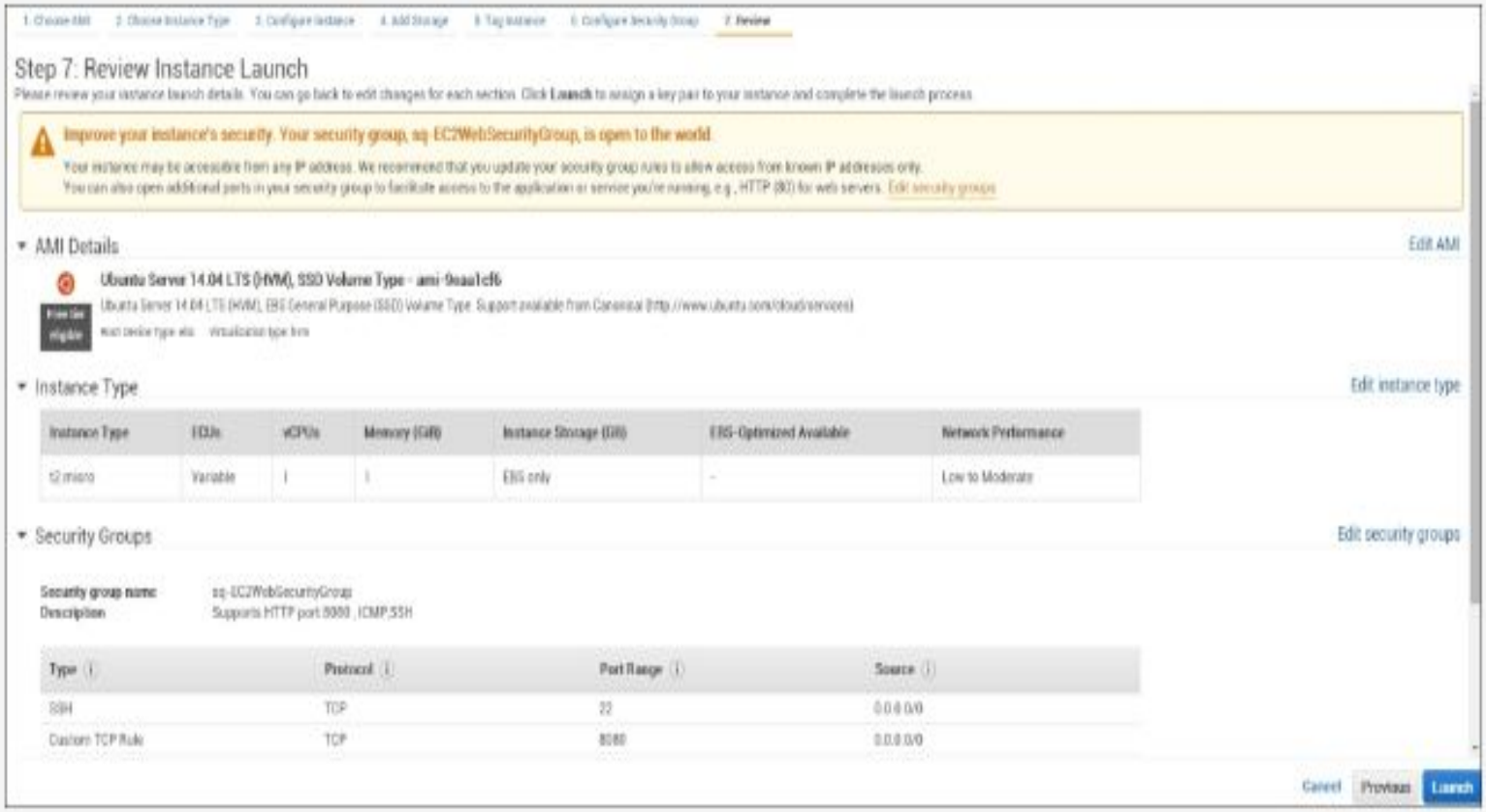
Setting up AWS Infrastructure



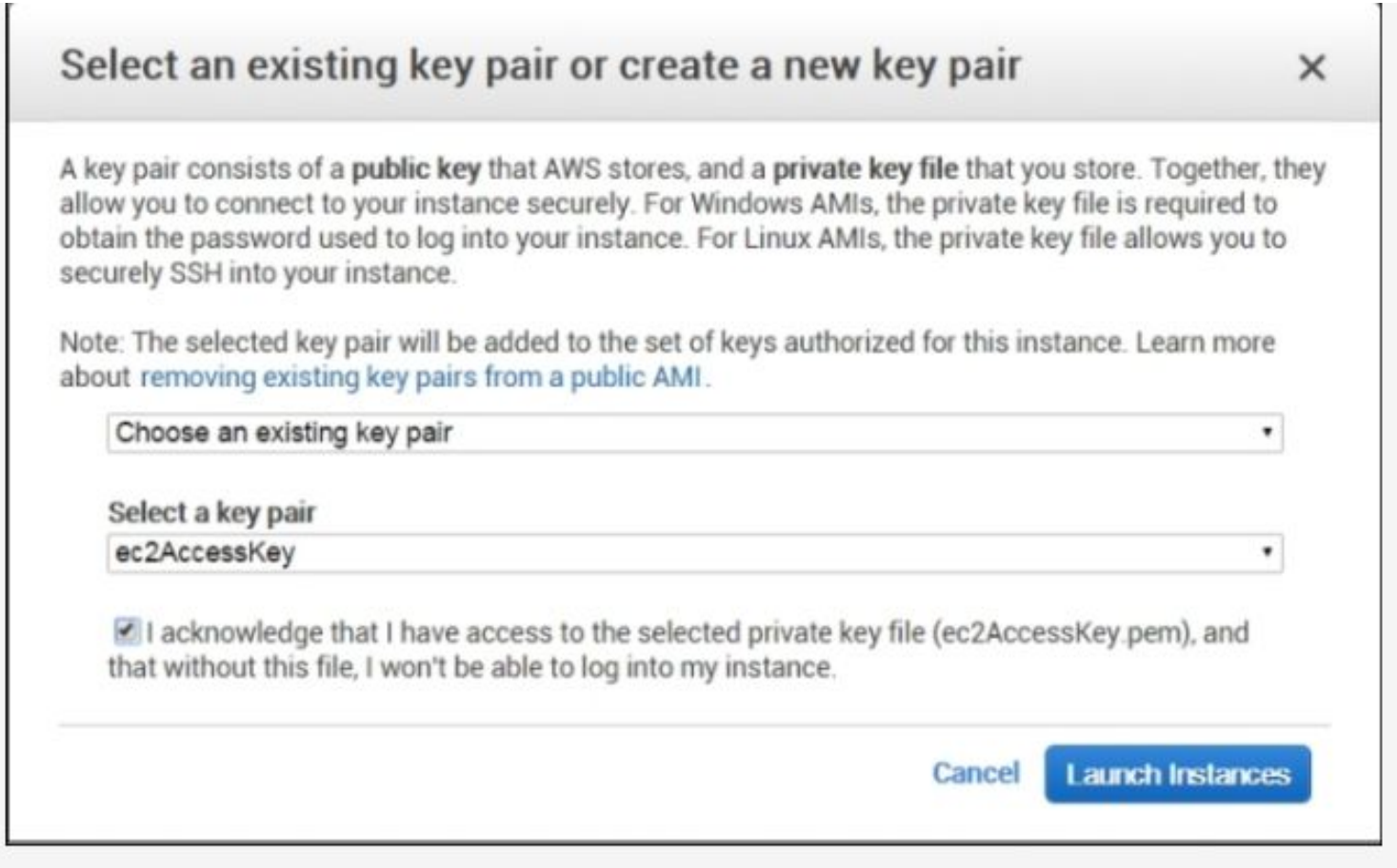
Setting up AWS Infrastructure



Setting up AWS Infrastructure

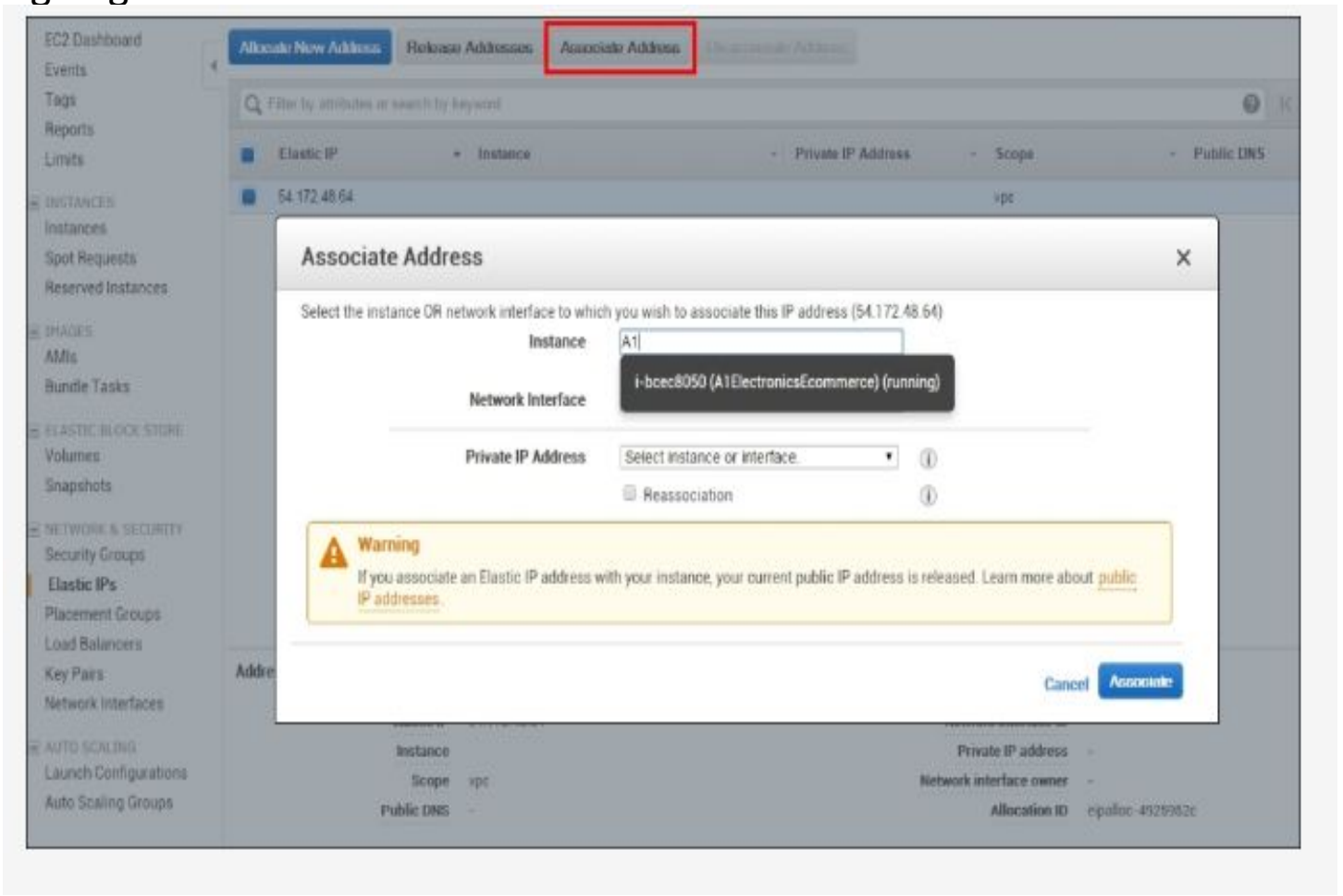


Setting up AWS Infrastructure

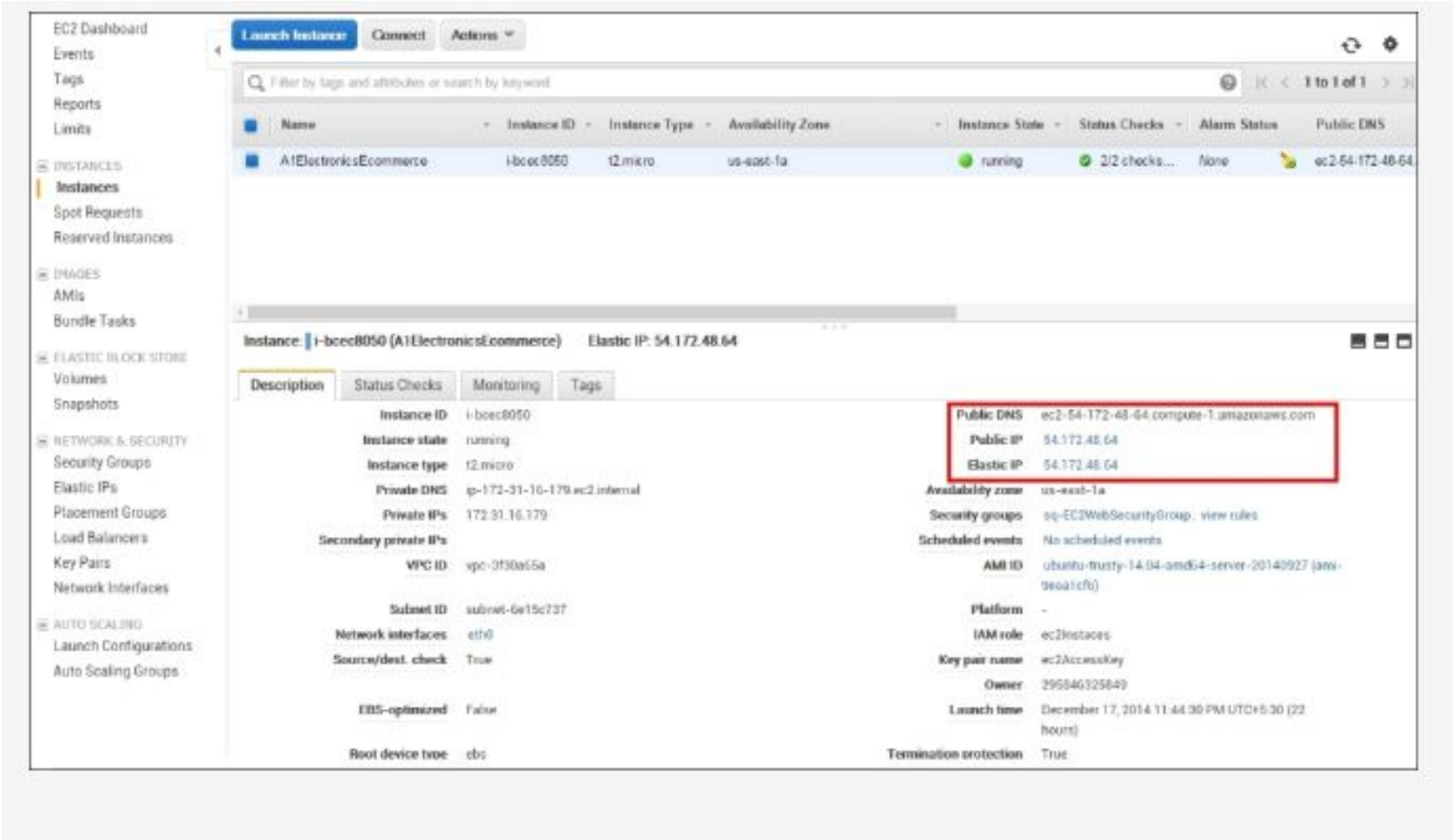


Setting up AWS Infrastructure

Assigning Elastic IP



Setting up AWS Infrastructure



A	: (address) Actual IP address of the Domain
AAAA	: Ipv6 address record which maps the hostname to a 128 bit Ipv6 address
CNAME	: (canonical name) Makes one domain alias for another domain name
MX	: (mail exchange) List of mail exchange servers
PTR	: (pointer record) Maps an Ipv4 address to the CNAME on the hos
NS	: (name server) It highlights which Name Server is authoritative for the domain
SOA	: (State of Authority) It stores important information like timestamp when any change to Domain was made
SRV	: (Service) To define TCP service
TXT	: (Text) This lets the admin to insert text in the DNS record