Introducción a Cloud Computing

Amazon Web Services

Mauricio Améndola / Sebastián Orrego – Profesor Adjunto
Escuela de Tecnología – Facultad de Ingeniería

UNIVERSIDAD ORT Uruguay

AGENDA

1. Infraestructura Global
2. IaaS
3. Servicios IaaS
    a. Compute
    b. Storage
    c. Networking
    d. Seguridad

Infraestructura global

Infraestructura global

- ➔ Regiones
- ➔ Zonas de disponibilidad
- ➔ Edge locations

Regiones

Ubicación física en el mundo donde aws tiene múltiples zonas de disponibilidad.

Zonas de disponibilidad

(Availability Zones) Consisten en dos o más data centers, cada uno con red, energía y conectividad redundantes, y ubicados en espacios físicos independientes.

Edge Locations

Datacenters distribuidos por todo el mundo con el fin de reducir la latencia en la entrega de contenido al usuario.
Este contenido, generalmente es cached.

25 Launched Regions

Each with multiple Availability Zones (AZ's)

81 Availability Zones

8 Local Zones

17 Wavelength Zones

For ultralow latency applications

7 Announced Regions

9 Announced Local Zones

2x More Regions

With multiple AZ's than the next largest cloud provider

245 Countries and Territories Served

108 Direct Connect Locations

230+ Points of Presence

218+ Edge Locations and 12 Regional Edge Caches


Regions
Coming Soon

2021 Magic Quadrant for Cloud Infrastructure & Platform Services

Acceso a Servicios AWS

- ➔ Consola web: acceso desde un browser
- ➔ Interfaz de comandos: Herramienta cli
- ➔ API: Http APIs para realizar todo tipo de tareas

Acceso a Servicios AWS

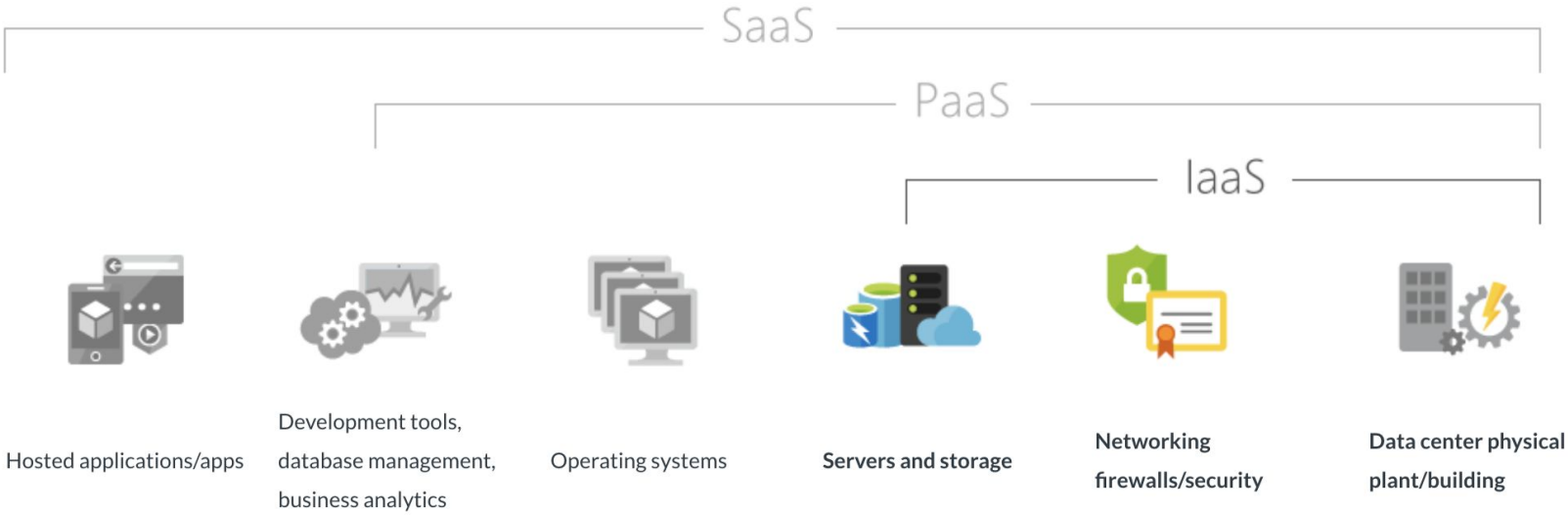
click
click
click

Infrastructure as a Service (IaaS)

IaaS

Se provee al consumidor la capacidad de adquirir bajo demanda recursos de procesamiento, almacenamiento, red, etc. donde se puede desplegar y ejecutar software arbitrario (incluye sistemas operativos y aplicaciones)
En este modelo los desarrolladores crean, configuran y mantienen máquinas virtuales junto con todo el software requerido para ejecutar sus aplicaciones.

IaaS


SaaS
PaaS
IaaS
Hosted applications/apps
Development tools, database management, business analytics
Operating systems
Servers and storage
Networking firewalls/security
Data center physical plant/building

IaaS

- ➔ Elimina la necesidad de inversión inicial y baja el costo
- ➔ Mejora la continuidad de negocio y facilita la recuperación frente a desastres
- ➔ Articula / Facilita la innovación
- ➔ Permite responder rápidamente a cambios de negocio
- ➔ Permite tener foco, liberando al equipo de tareas que no sean específicas del negocio
- ➔ Menor “time to market” (GO Live in seconds)

IaaS en AWS

“Amazon Elastic Compute Cloud (Amazon EC2) es un servicio web que proporciona capacidad informática en la nube segura y de tamaño modificable. Está diseñado para simplificar el uso de la informática en la nube a escala web para los desarrolladores. La sencilla interfaz de servicios web de Amazon EC2 permite obtener y configurar capacidad con una fricción mínima. Proporciona un control completo sobre los recursos informáticos y le permite ejecutarse en el entorno informático probado de Amazon.”

IaaS en AWS

EC2 Foundations

Resources

Instances
Storage
Networking

Availability

Regions and AZs
Placement Groups
Load Balancing
Auto Scaling

Management

Deployment
Monitoring
Administration

Purchase Options

On Demand
Reserved
Spot

Instances

Resources

Instances
Storage
Networking

Availability

Regions and AZs
Placement Groups
Load Balancing
Auto Scaling

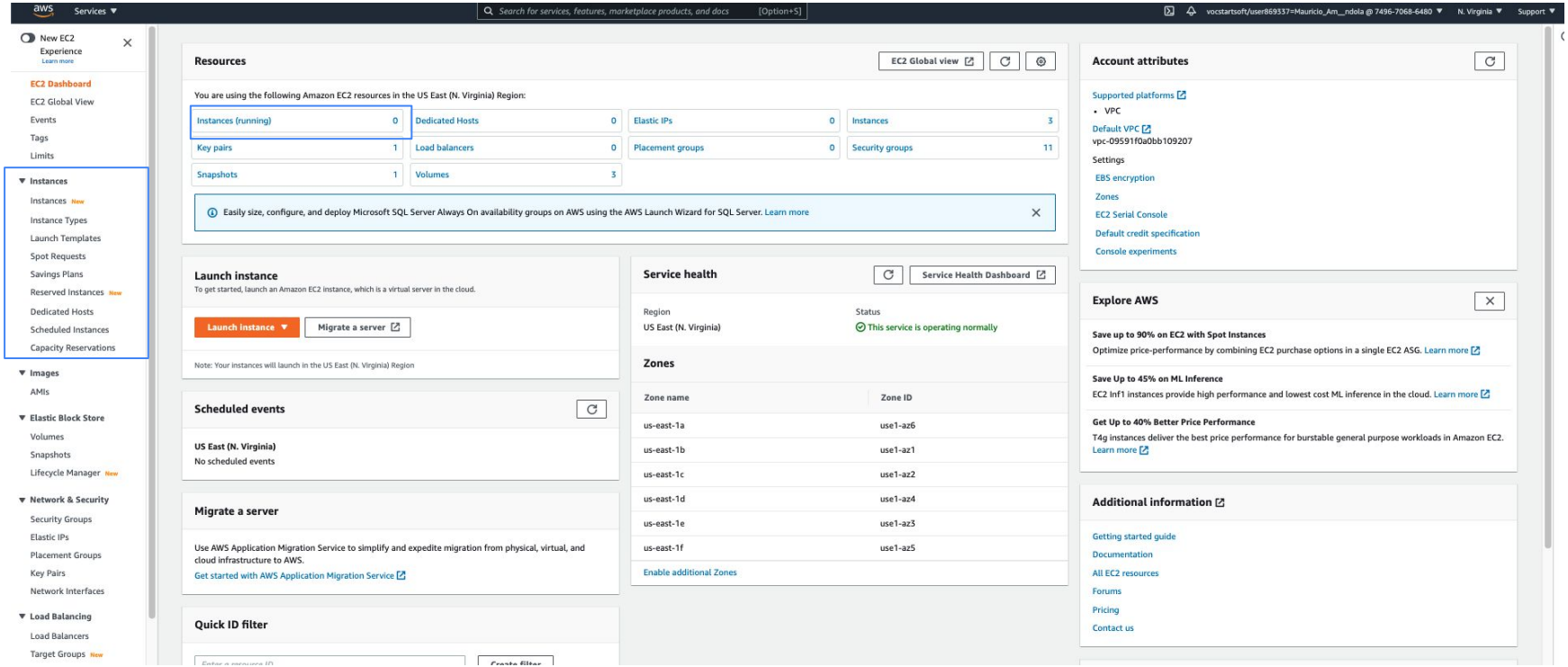
Management

Deployment
Monitoring
Administration

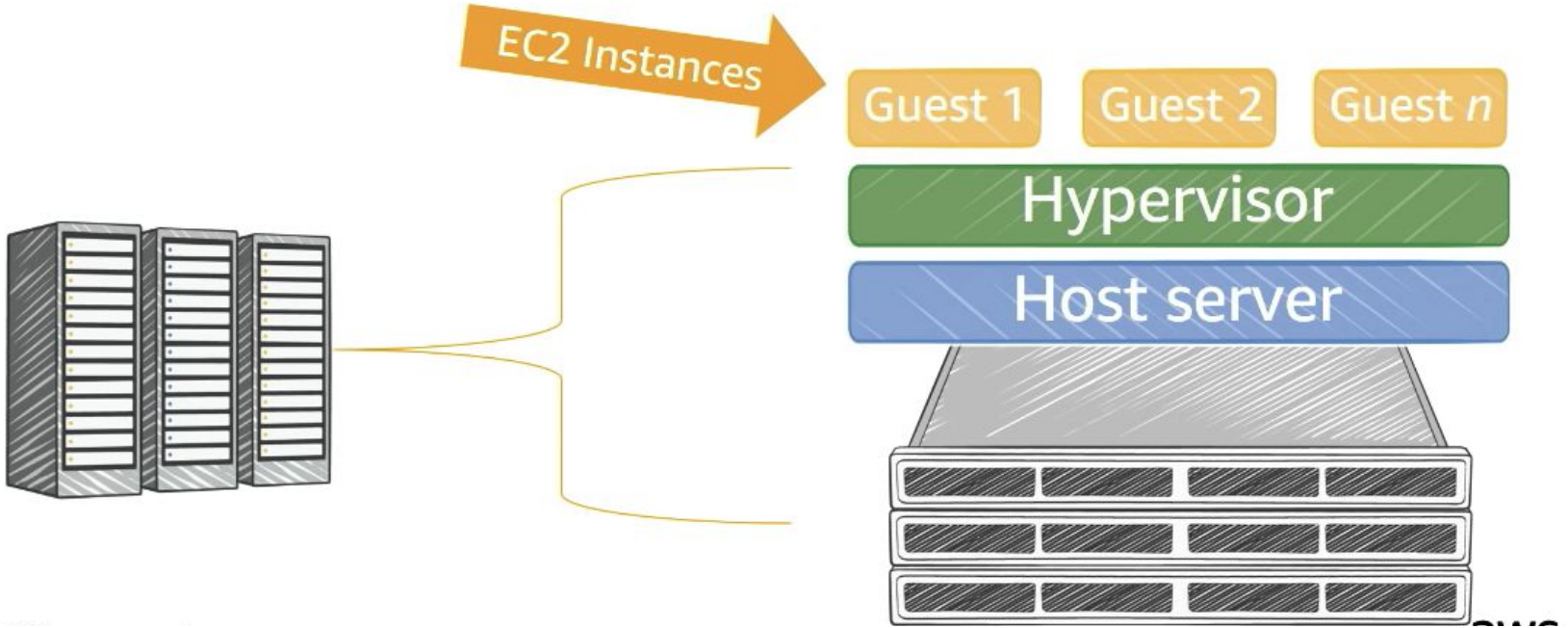
Purchase Options

On Demand
Reserved
Spot

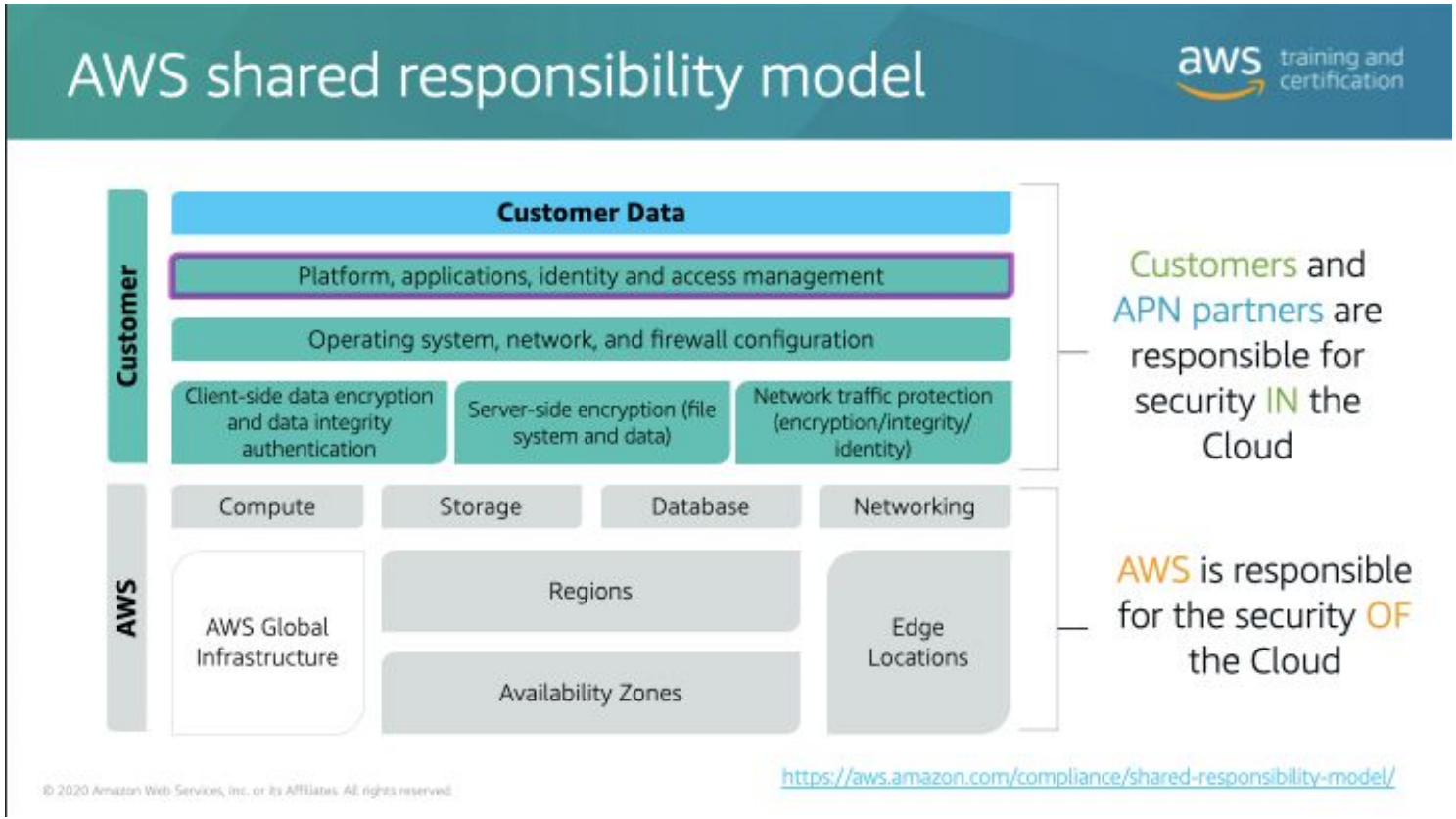
Instances


aws
Services
Search for services, features, marketplace products, and docs [Option+S]
N. Virginia
Support
New EC2 Experience
Learn more
EC2 Dashboard
EC2 Global View
Events
Tags
Limits
Instances
Instances New
Instance Types
Launch Templates
Spot Requests
Savings Plans
Reserved Instances New
Dedicated Hosts
Scheduled Instances
Capacity Reservations
Images
AMIs
Elastic Block Store
Volumes
Snapshots
Lifecycle Manager New
Network & Security
Security Groups
Elastic IPs
Placement Groups
Key Pairs
Network Interfaces
Load Balancing
Load Balancers
Target Groups New
Resources
EC2 Global view
You are using the following Amazon EC2 resources in the US East (N. Virginia) Region:
Instances (running) 0
Dedicated Hosts 0
Elastic IPs 0
Instances 3
Key pairs 1
Load balancers 0
Placement groups 0
Security groups 11
Snapshots 1
Volumes 3
Easily size, configure, and deploy Microsoft SQL Server Always On availability groups on AWS using the AWS Launch Wizard for SQL Server. Learn more
Launch instance
To get started, launch an Amazon EC2 instance, which is a virtual server in the cloud.
Launch instance
Migrate a server
Note: Your instances will launch in the US East (N. Virginia) Region
Scheduled events
US East (N. Virginia)
No scheduled events
Migrate a server
Use AWS Application Migration Service to simplify and expedite migration from physical, virtual, and cloud infrastructure to AWS.
Get started with AWS Application Migration Service
Quick ID filter
Service health
Service Health Dashboard
Region
US East (N. Virginia)
Status
This service is operating normally
Zones
Zone name
Zone ID
us-east-1a
use1-az6
us-east-1b
use1-az1
us-east-1c
use1-az2
us-east-1d
use1-az4
us-east-1e
use1-az3
us-east-1f
use1-az5
Enable additional Zones
Account attributes
Supported platforms
VPC
Default VPC
vpc-09591f0a0bb109207
Settings
EBS encryption
Zones
EC2 Serial Console
Default credit specification
Console experiments
Explore AWS
Save up to 90% on EC2 with Spot Instances
Optimize price-performance by combining EC2 purchase options in a single EC2 ASG. Learn more
Save Up to 45% on ML Inference
EC2 Inf1 instances provide high performance and lowest cost ML inference in the cloud. Learn more
Get Up to 40% Better Price Performance
T4g instances deliver the best price performance for burstable general purpose workloads in Amazon EC2. Learn more
Additional information
Getting started guide
Documentation
All EC2 resources
Forums
Pricing
Contact us

Instances


EC2 Instances
Guest 1
Guest 2
Guest n
Hypervisor
Host server

Instances


AWS shared responsibility model
aws training and certification
Customer
Customer Data
Platform, applications, identity and access management
Operating system, network, and firewall configuration
Client-side data encryption and data integrity authentication
Server-side encryption (file system and data)
Network traffic protection (encryption/integrity/ identity)
Customers and APN partners are responsible for security IN the Cloud
AWS
Compute
Storage
Database
Networking
AWS Global Infrastructure
Regions
Availability Zones
Edge Locations
AWS is responsible for the security OF the Cloud
© 2020 Amazon Web Services, Inc. or its Affiliates. All rights reserved.
https://aws.amazon.com/compliance/shared-responsibility-model/

Instances

- Responsabilidad de AWS en:
  - Controles físicos: ingreso de personal, autorización, etc
  - Controles ambientales: refrigeración DC, electricidad, etc
  - Patch Management (en su infra)
  - Configuration Management (de su infra)
  - Entrenamiento de sus colaboradores

- Responsabilidad del cliente en:
  - Manejo de seguridad a nivel del S.O
  - Cifrado de los datos / servicios (SSL/HTTPS)
  - Entrenamiento de los sysadmin / operadores de los servicios

Instances

Amazon Machine Image (AMI)

Amazon maintained

Set of Linux and Windows images

Kept up-to-date by Amazon in each region

Community maintained

Images published by other AWS users

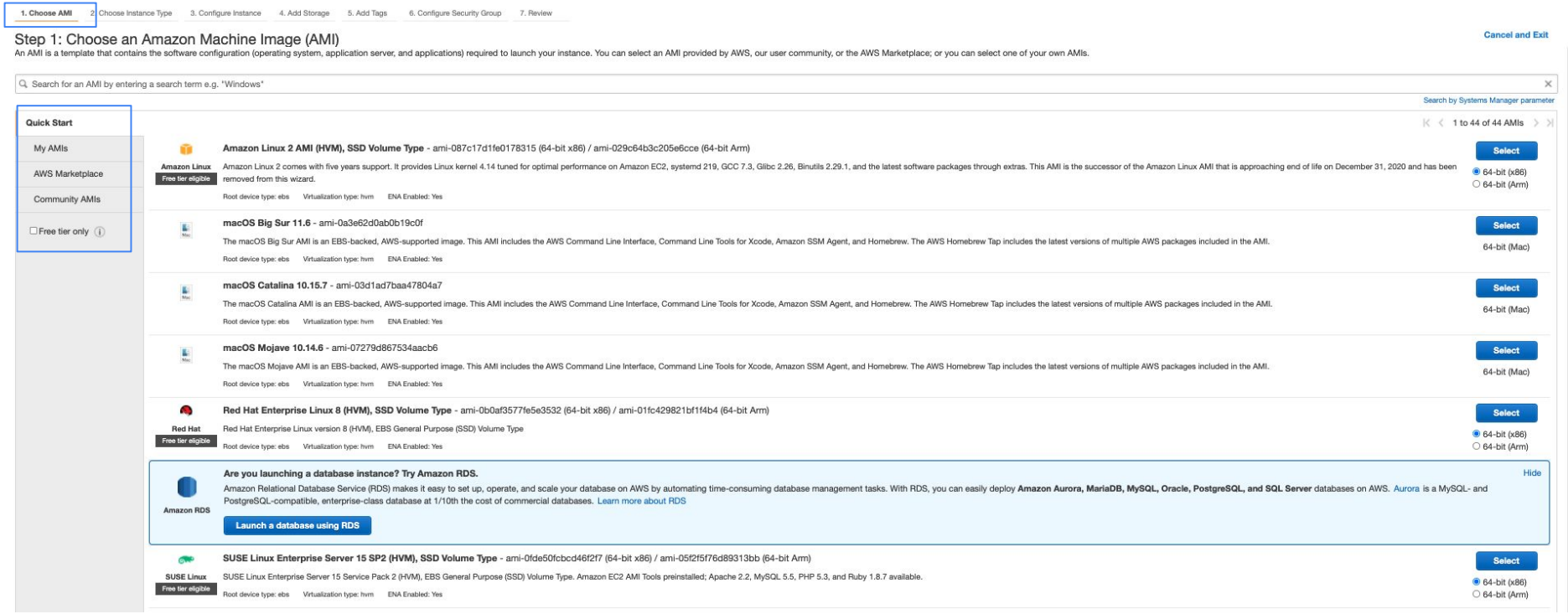
Managed and maintained by Marketplace partners

Your machine images

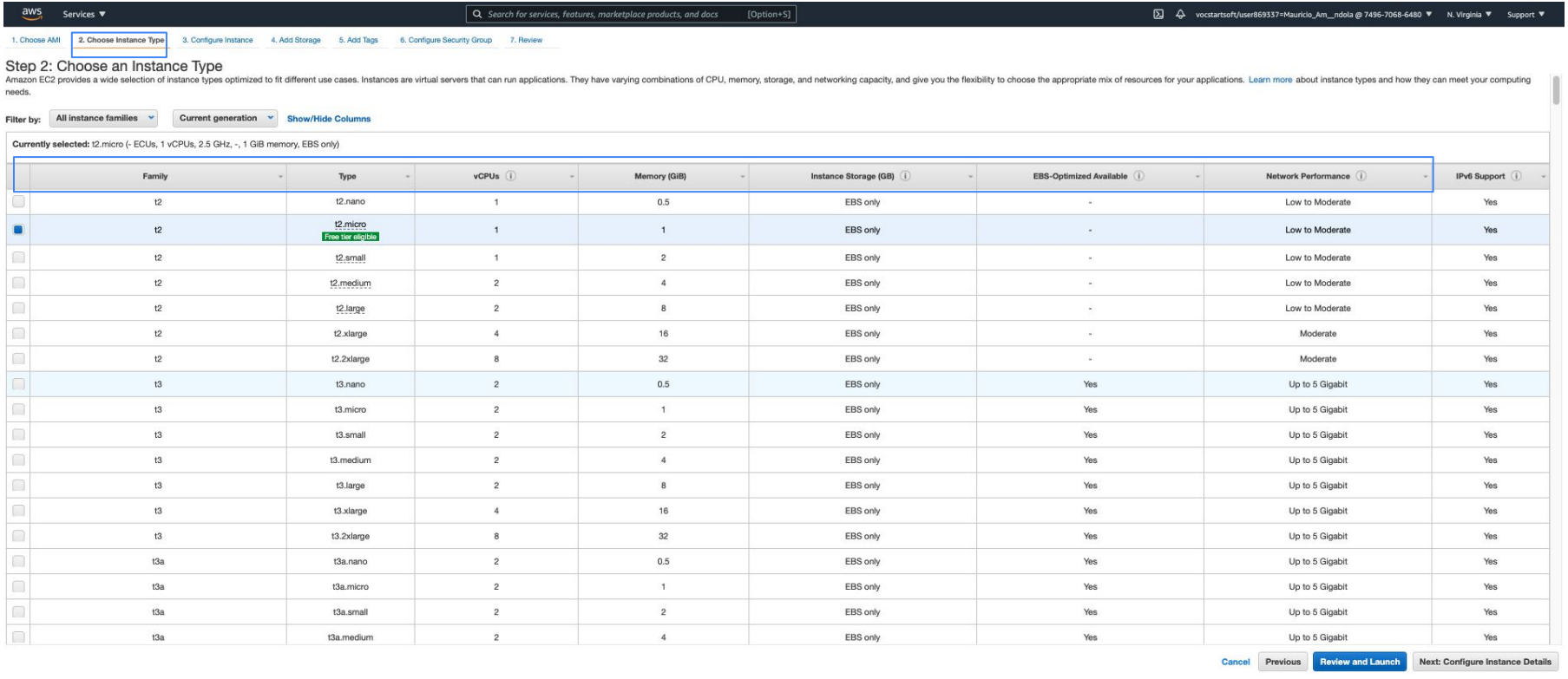
AMIs you have created from EC2 instances

Can be kept private or shared with other accounts

Instances: AMIs



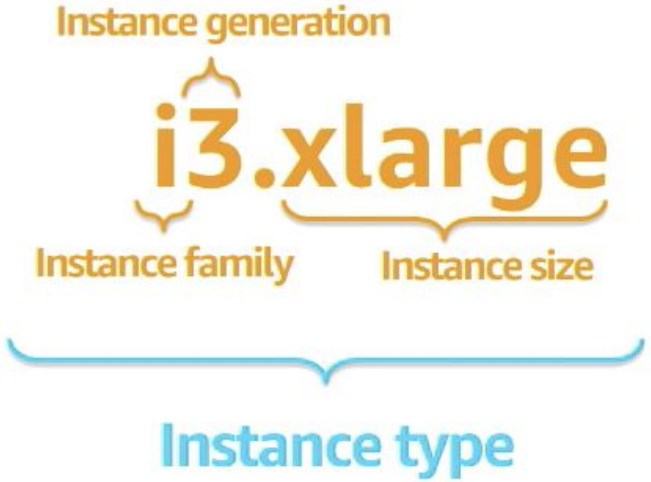
Instances: Tipo



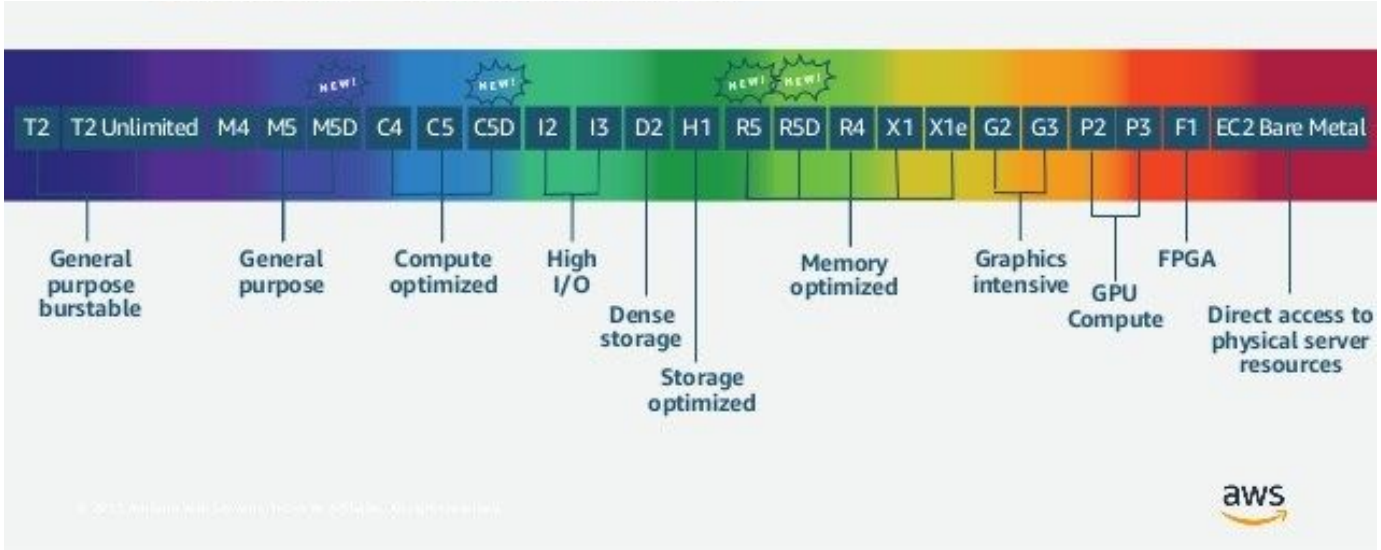
Instances: Tipo

Amplia selección de tipos de instancias optimizados para adaptarse a diferentes casos de uso.
Combinaciones de capacidad de CPU, memoria, almacenamiento y redes.
Cada tipo de instancia incluye uno o varios tamaños de instancia, lo que le permite escalar recursos según los requisitos de la carga de trabajo.

Instances: Tipo

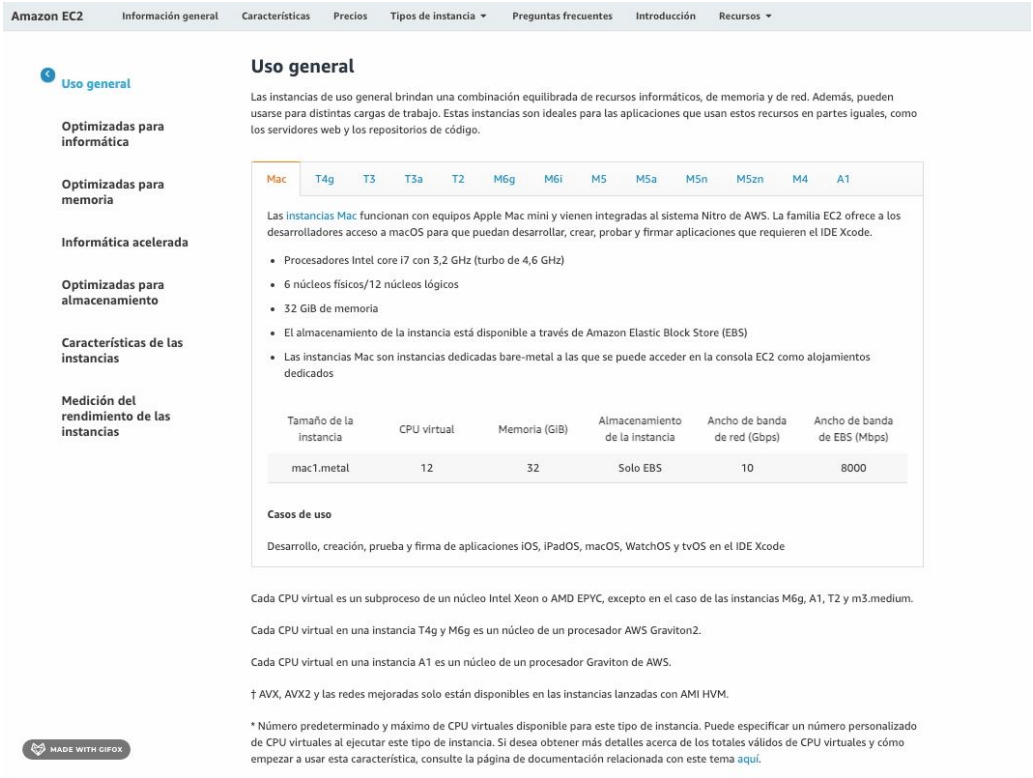


Instances: Tipo



https://aws.amazon.com/es/ec2/instance-types/

Instances: Tipo



[Listado de tipos]

Storage

Resources	Availability	Management	Purchase Options
Instances	Regions and AZs	Deployment	On Demand
Storage	Placement Groups	Monitoring	Reserved
Networking	Load Balancing	Administration	Spot
	Auto Scaling		

Storage

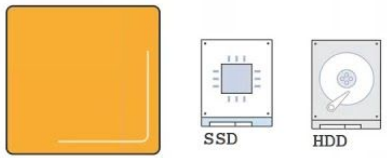
Dos tipos de storage asociados directamente a la instancia:

- ➔ Instance store
- ➔ Elastic Block store

Storage

EC2 Instance Store

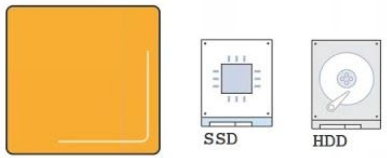
- Local to instance
- Non-persistent data store
- Data not replicated (by default)
- No snapshot support
- SSD or HDD



Storage

EC2 Instance Store

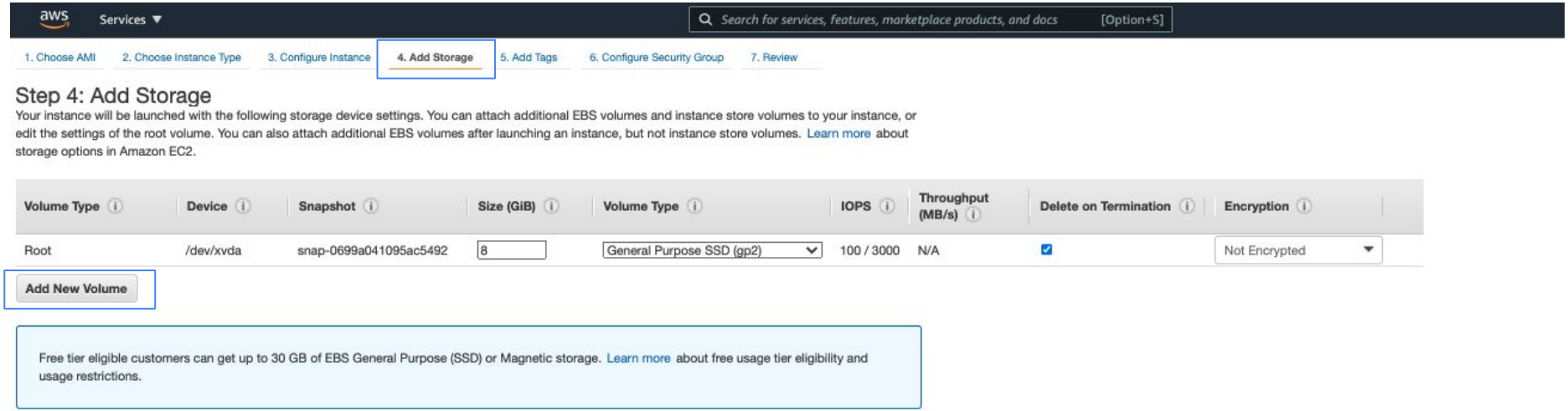
- Local to instance
- Non-persistent data store
- Data not replicated (by default)
- No snapshot support
- SSD or HDD



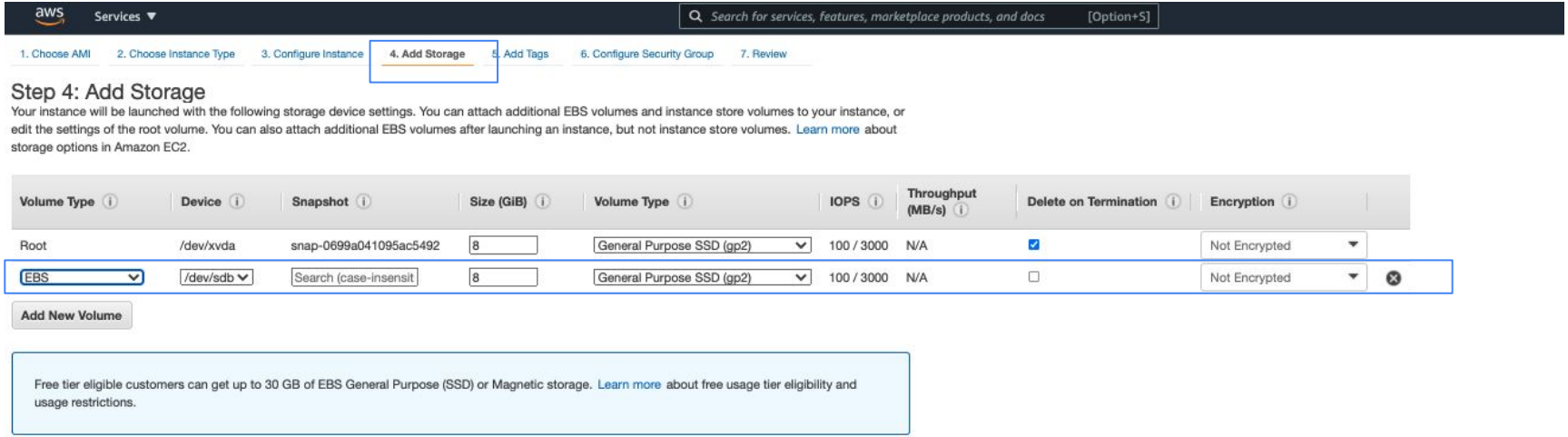
Elastic Block Store

- Persistent block storage volumes
- 99.999% availability
- Automatically replicated within its Availability Zone (AZ)
- Point-in-time snapshot support
- Modify volume type as needs change
- SSD or HDD
- Auto recovery

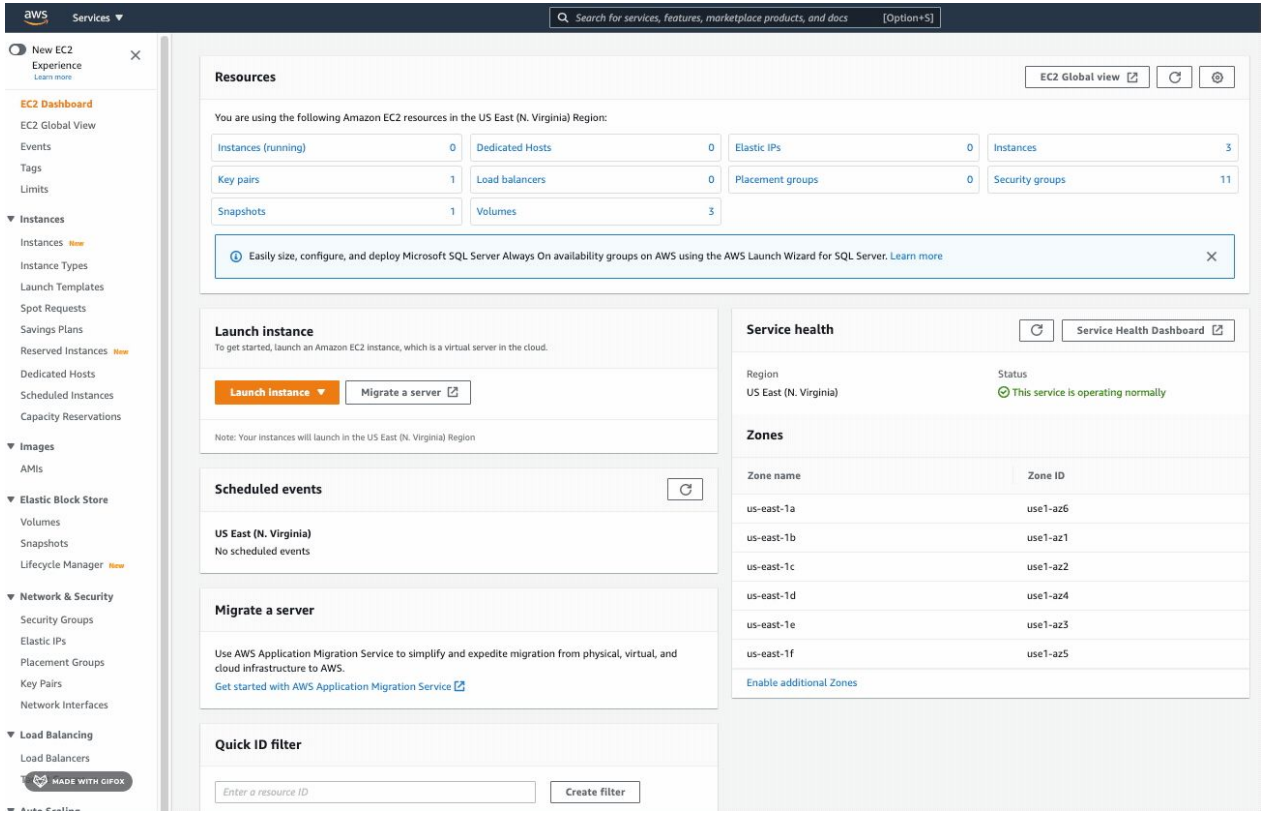
Storage: Instance store



Storage: EBS


aws
Services
Search for services, features, marketplace products, and docs [Option+S]
1. Choose AMI
2. Choose Instance Type
3. Configure Instance
4. Add Storage
5. Add Tags
6. Configure Security Group
7. Review
Step 4: Add Storage
Your instance will be launched with the following storage device settings. You can attach additional EBS volumes and instance store volumes to your instance, or edit the settings of the root volume. You can also attach additional EBS volumes after launching an instance, but not instance store volumes. Learn more about storage options in Amazon EC2.
Volume Type
Device
Snapshot
Size (GiB)
Volume Type
IOPS
Throughput (MB/s)
Delete on Termination
Encryption
Root
/dev/xvda
snap-0699a041095ac5492
8
General Purpose SSD (gp2)
100 / 3000
N/A
Not Encrypted
EBS
/dev/sdb
Search (case-insensit
8
General Purpose SSD (gp2)
100 / 3000
N/A
Not Encrypted
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.

Storage: EBS


aws
Services
Search for services, features, marketplace products, and docs [Option+S]
New EC2 Experience
Learn more
EC2 Dashboard
EC2 Global View
Events
Tags
Limits
Instances
Instances New
Instance Types
Launch Templates
Spot Requests
Savings Plans
Reserved Instances New
Dedicated Hosts
Scheduled Instances
Capacity Reservations
Images
AMIs
Elastic Block Store
Volumes
Snapshots
Lifecycle Manager New
Network & Security
Security Groups
Elastic IPs
Placement Groups
Key Pairs
Network Interfaces
Load Balancing
Load Balancers
MADE WITH GIFOX
Resources
EC2 Global view
You are using the following Amazon EC2 resources in the US East (N. Virginia) Region:
Instances (running) 0
Dedicated Hosts 0
Elastic IPs 0
Instances 3
Key pairs 1
Load balancers 0
Placement groups 0
Security groups 11
Snapshots 1
Volumes 3
Easily size, configure, and deploy Microsoft SQL Server Always On availability groups on AWS using the AWS Launch Wizard for SQL Server. Learn more
Launch instance
To get started, launch an Amazon EC2 instance, which is a virtual server in the cloud.
Launch Instance
Migrate a server
Note: Your instances will launch in the US East (N. Virginia) Region
Scheduled events
US East (N. Virginia)
No scheduled events
Migrate a server
Use AWS Application Migration Service to simplify and expedite migration from physical, virtual, and cloud infrastructure to AWS.
Get started with AWS Application Migration Service
Quick ID filter
Enter a resource ID
Create filter
Service health
Service Health Dashboard
Region
US East (N. Virginia)
Status
This service is operating normally
Zones
Zone name
Zone ID
us-east-1a use1-az6
us-east-1b use1-az1
us-east-1c use1-az2
us-east-1d use1-az4
us-east-1e use1-az3
us-east-1f use1-az5
Enable additional Zones

Networking

Resources

Instances
Storage
Networking

Availability

Regions and AZs
Placement Groups
Load Balancing
Auto Scaling

Management

Deployment
Monitoring
Administration

Purchase Options

On Demand
Reserved
Spot

Networking

Conceptos nuevos:

- ➔ VPC: Virtual Private Cloud
- ➔ Subnets
- ➔ Route Tables
- ➔ Internet Gateway
- ➔ Security Groups
- ➔ Network ACLs

Networking: VPC

Virtual Private Cloud

Provision a logically isolated cloud where you can launch AWS resources into a virtual network

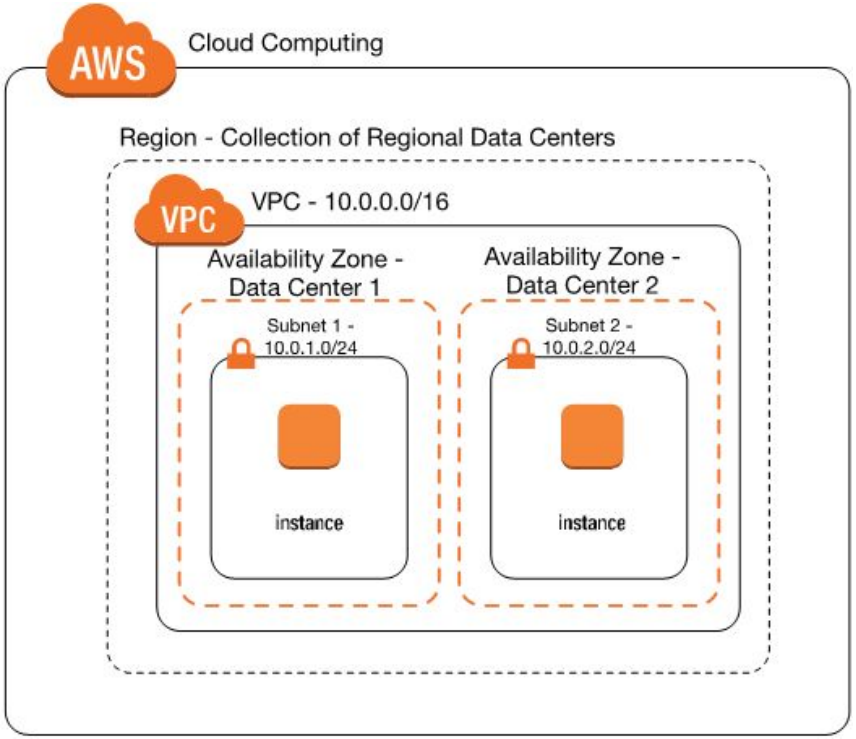
Networking: VPC

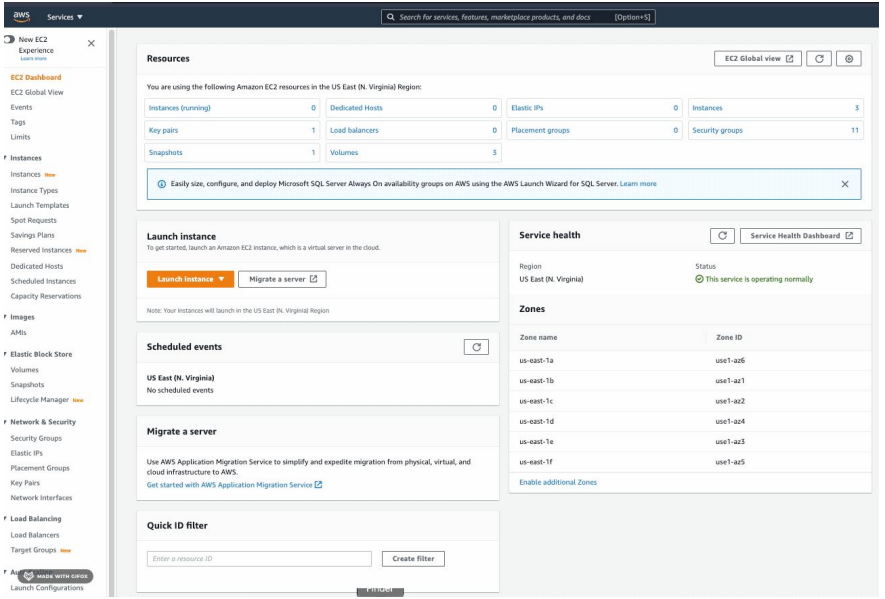
Virtual Private Cloud

Provision a logically isolated cloud where you can launch AWS resources into a virtual network

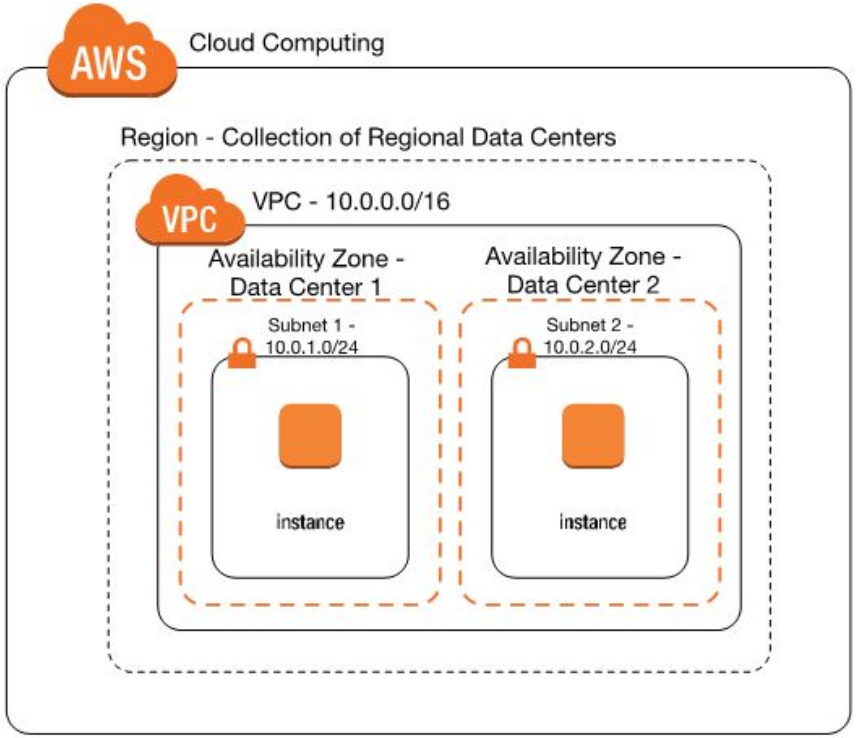
- A private, virtual network in the AWS Cloud
  - Uses same concepts as on premise networking
- Allows complete control of network configuration
  - Ability to isolate and expose resources inside VPC
- Offers several layers of security controls
  - Ability to allow and deny specific internet and internal traffic
- Other AWS services deploy into VPC
  - Services inherent security built into network

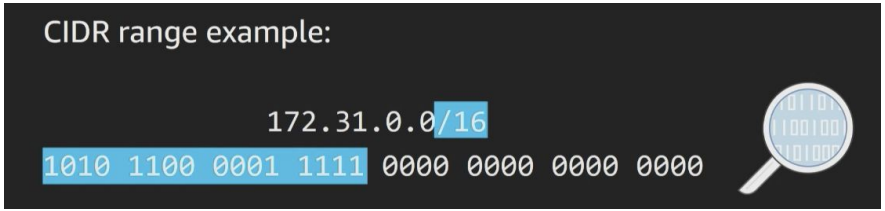
Networking: VPC





Networking: Subnets





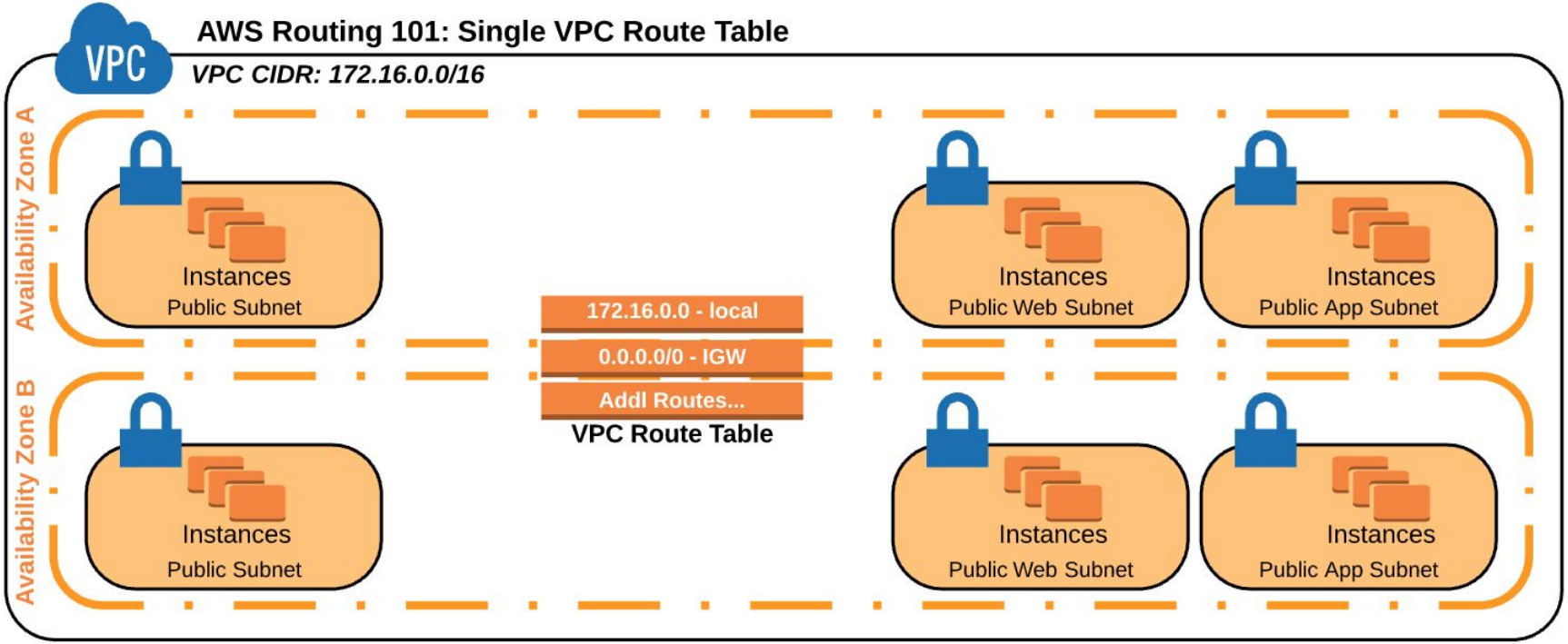
Recomendaciones:

- ➔ /16 VPCs (65,536 direcciones)
- ➔ /24 Subnets (253 direcciones)
- ➔ Usar múltiples zonas de disponibilidad mediante múltiples subnets.

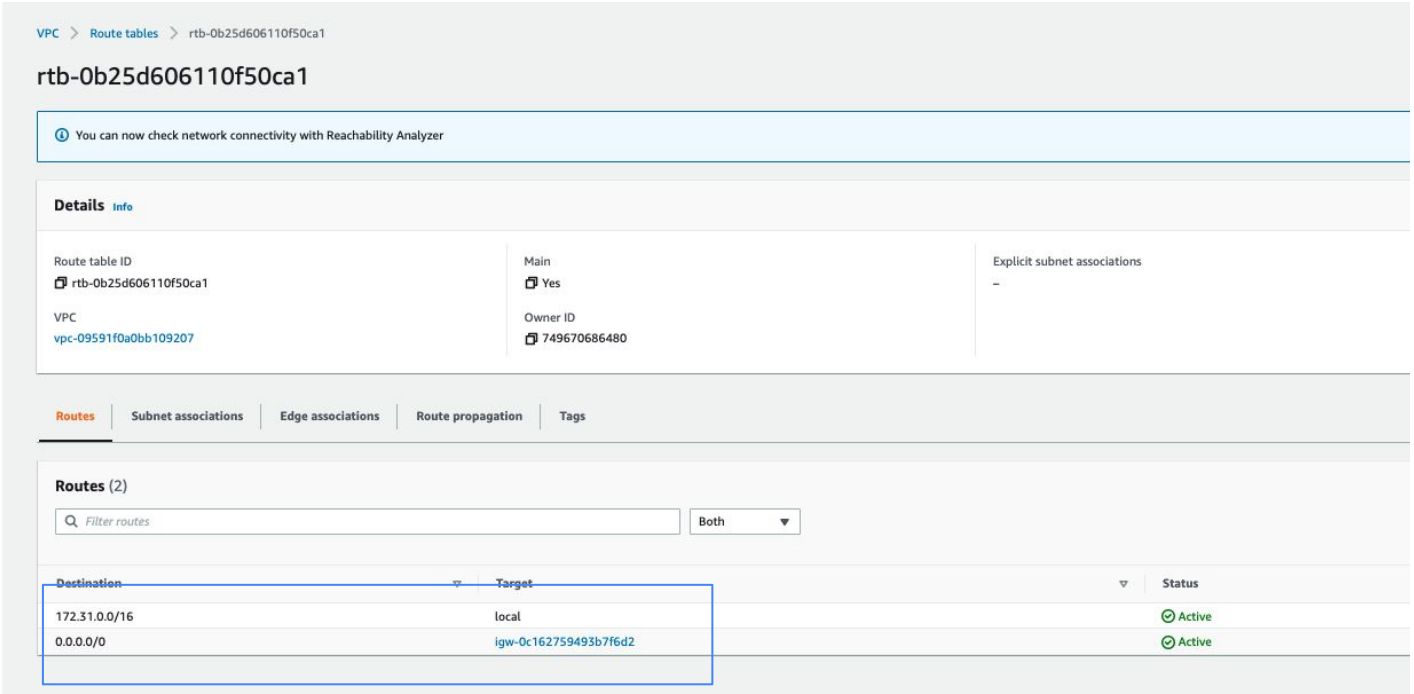
Networking: Route tables

Las tablas de ruteo contienen conjuntos de reglas, denominadas rutas, que se usan para determinar adónde se dirige el tráfico de red desde su subred o gateway.
Toda VPC tiene una regla por defecto
Es posible asignar diferentes tablas de ruteo a diferentes sub-net

Networking: Route tables



Networking: Route tables



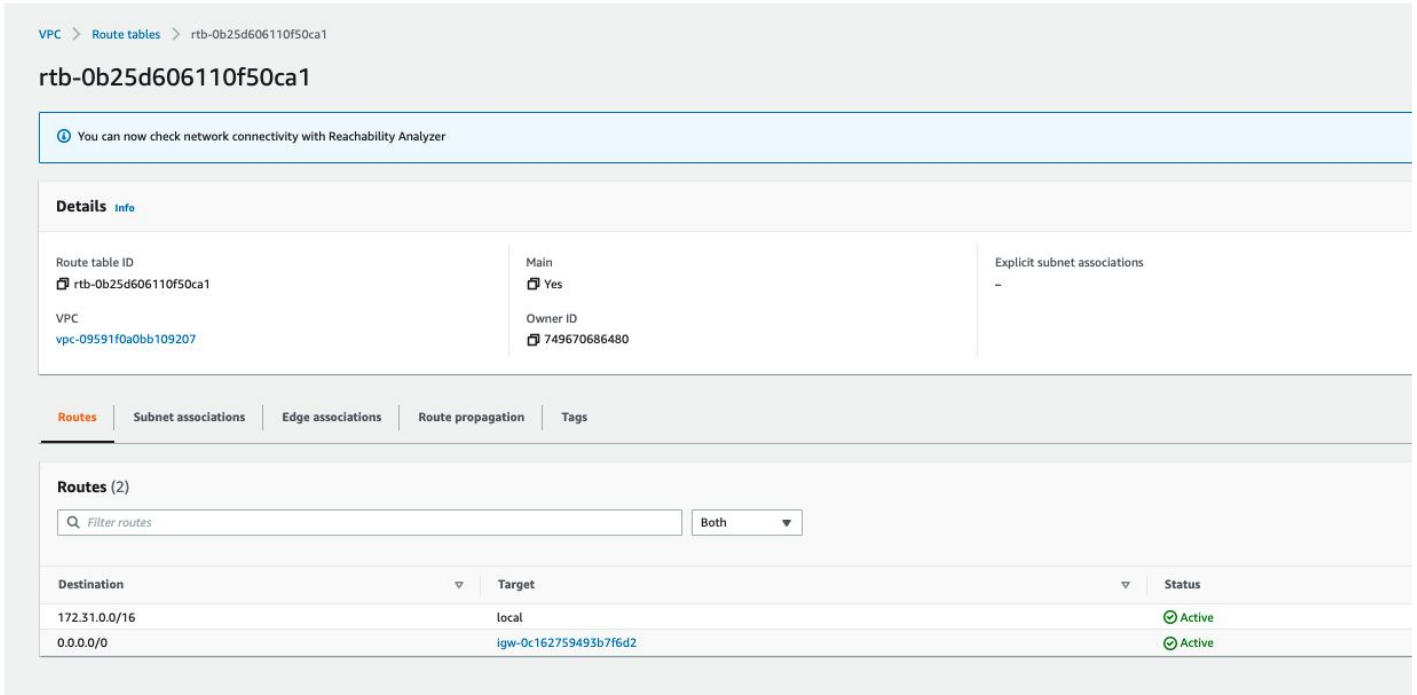
Networking: Internet Gateway

Componente de la VPC de escalado horizontal, redundante y de alta disponibilidad que permite la comunicación entre las instancias de la VPC e Internet.

Sirve para dos fines:

- ➔ Proporcionar un objetivo en tablas de ruteo de VPC para el tráfico direccionable de Internet
- ➔ Realizar la conversión de las direcciones de red (NAT) para las instancias que tengan asignadas direcciones IPv4 públicas.

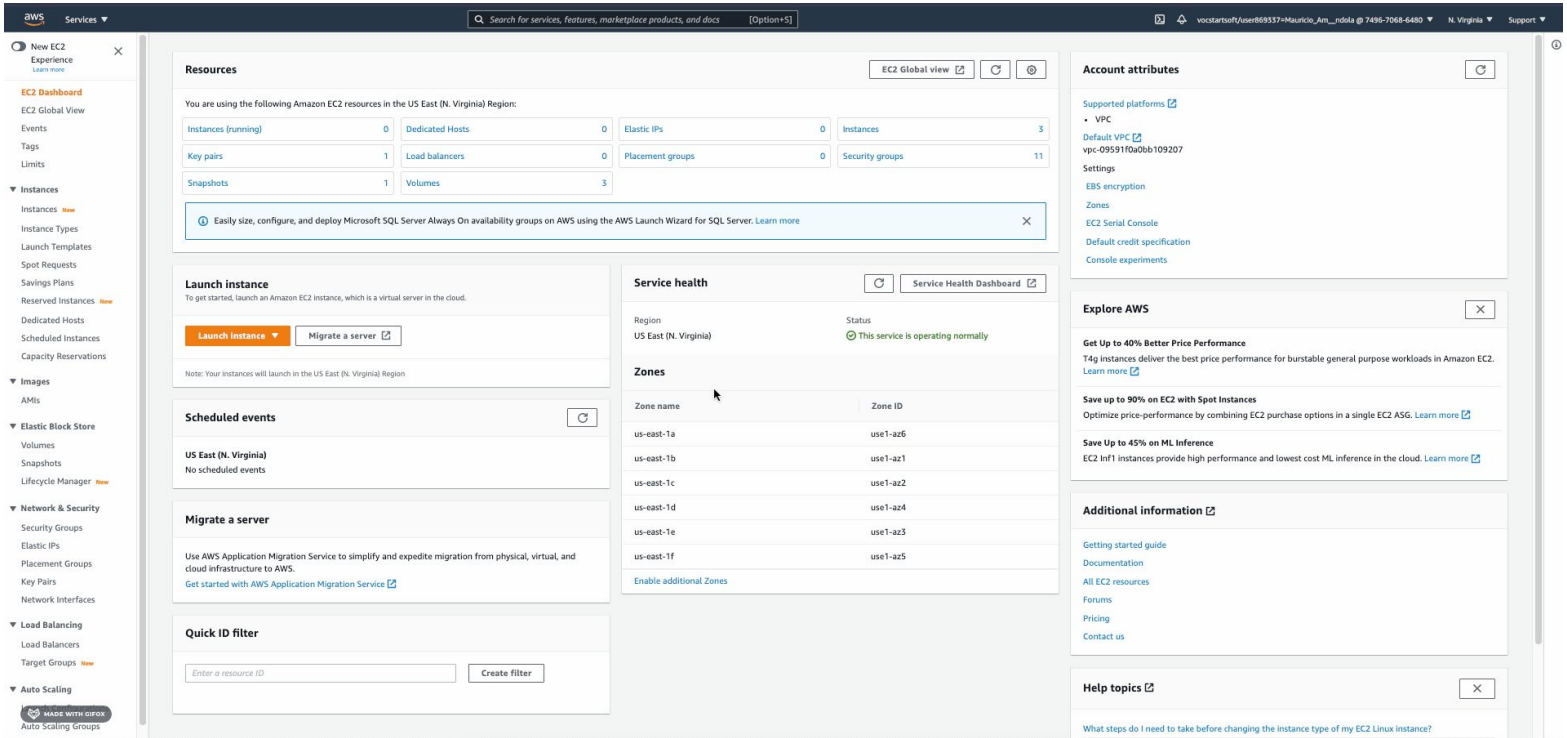
Networking: Internet Gateway



Networking: Security Groups

- ➔ Funcionan como un firewall virtual de la instancia para controlar el tráfico entrante y saliente.
- ➔ Hasta cinco grupos de seguridad por instancia.
- ➔ Actúan a nivel de la instancia, no en el de la subred. Por lo tanto, cada instancia de la subred puede asignarse a distintos conjuntos de grupos de seguridad.
- ➔ Son stateful: cuando se crea una regla de inbound para permitir tráfico, se permite además el tráfico de retorno.

Networking: Security Groups

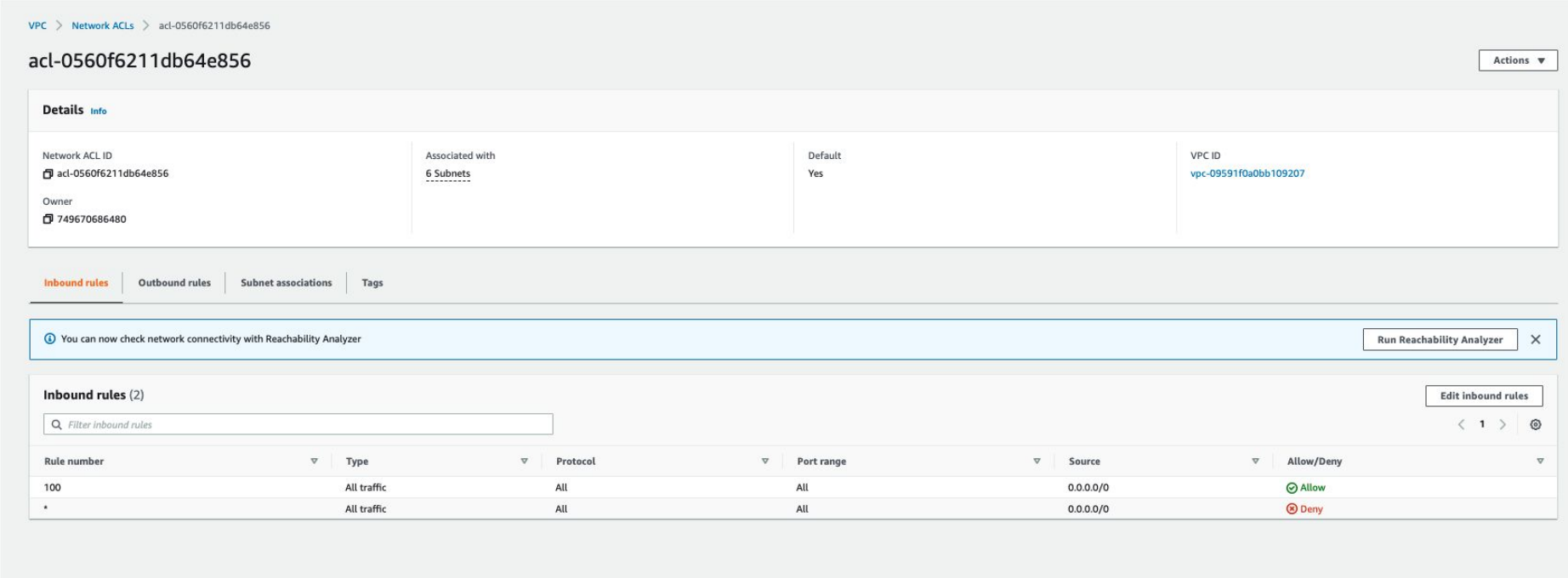


Networking: Network ACLs

Agregan una capa más de seguridad.

- ➔ Las ACLs soportan reglas de denied, por lo que podemos filtrar tráfico proveniente de orígenes no conocidos, botnets, etc.
- ➔ Son stateless, es decir, el tráfico de respuesta de una regla de inbound está sujeto a que exista una regla de outbound
- ➔ Todo VPC viene con una NACL por omisión que habilita todo el tráfico de salida y entrada
- ➔ Se pueden crear NACLs custom, en cuyo caso, la regla por omisión es “denied all”

Networking: Network ACLs



Disponibilidad / Confiabilidad

Resources

Instances
Storage
Networking

Availability

Regions and AZs
Placement Groups
Load Balancing
Auto Scaling

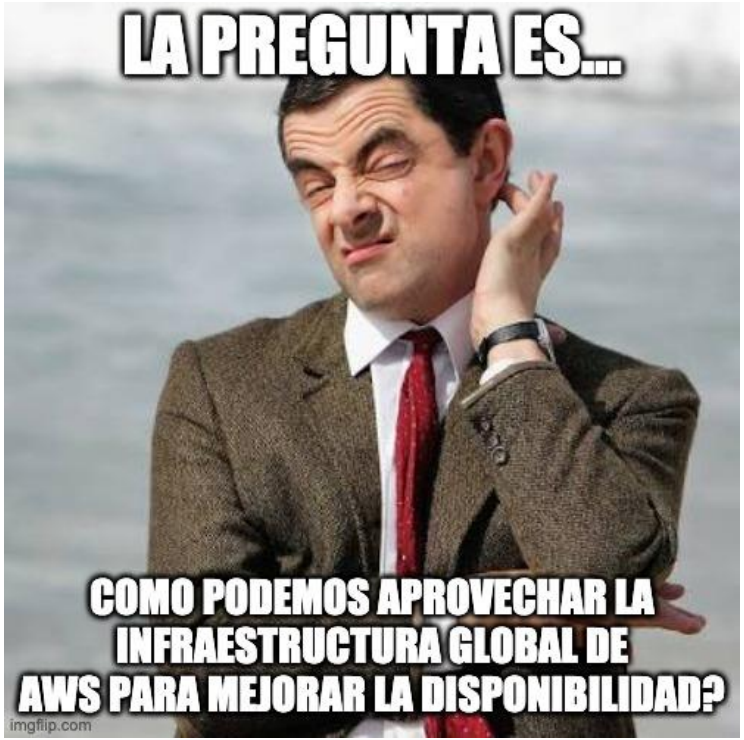
Management

Deployment
Monitoring
Administration

Purchase Options

On Demand
Reserved
Spot

Disponibilidad / Confiabilidad



Disponibilidad / Confiabilidad

AWS provee varios servicios para ayudar en este área:

- ➔ Placement Groups
- ➔ Load Balancers
- ➔ Auto-scaling

Disponibilidad: Placement Groups

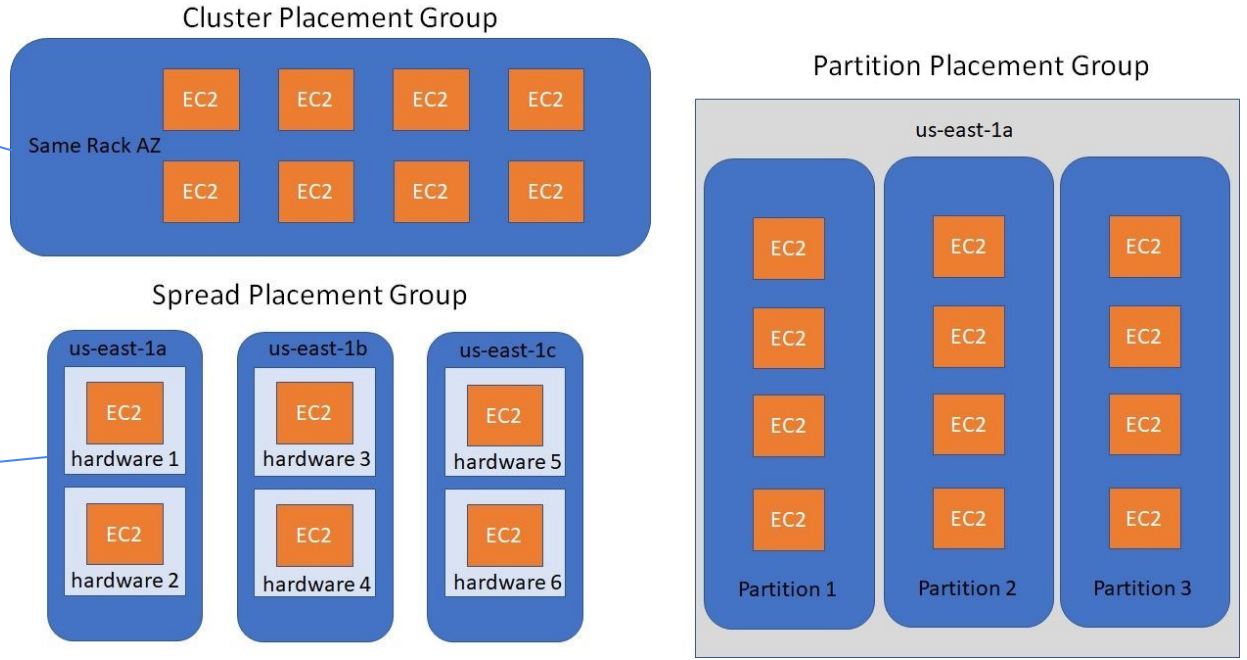
Son el mecanismo para influir, usando distintas estrategias, cómo se distribuyen las instancias en el hardware subyacente.
Tres tipos:

- ➔ Cluster
- ➔ Spread
- ➔ Partition

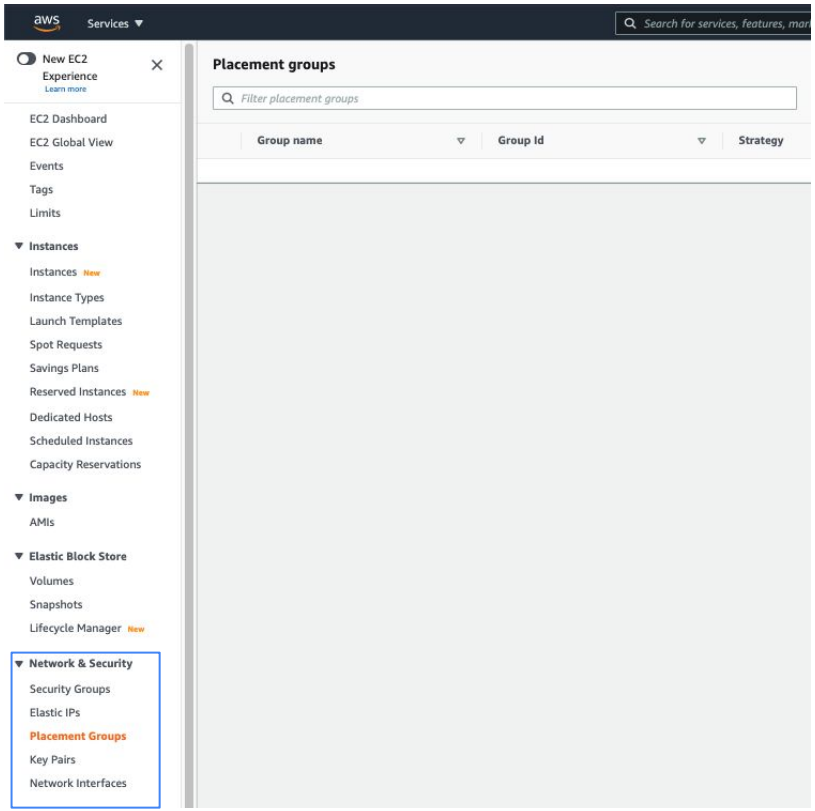
Disponibilidad: Placement Groups

Mejor performance

Mejor disponibilidad



Disponibilidad: Placement Groups

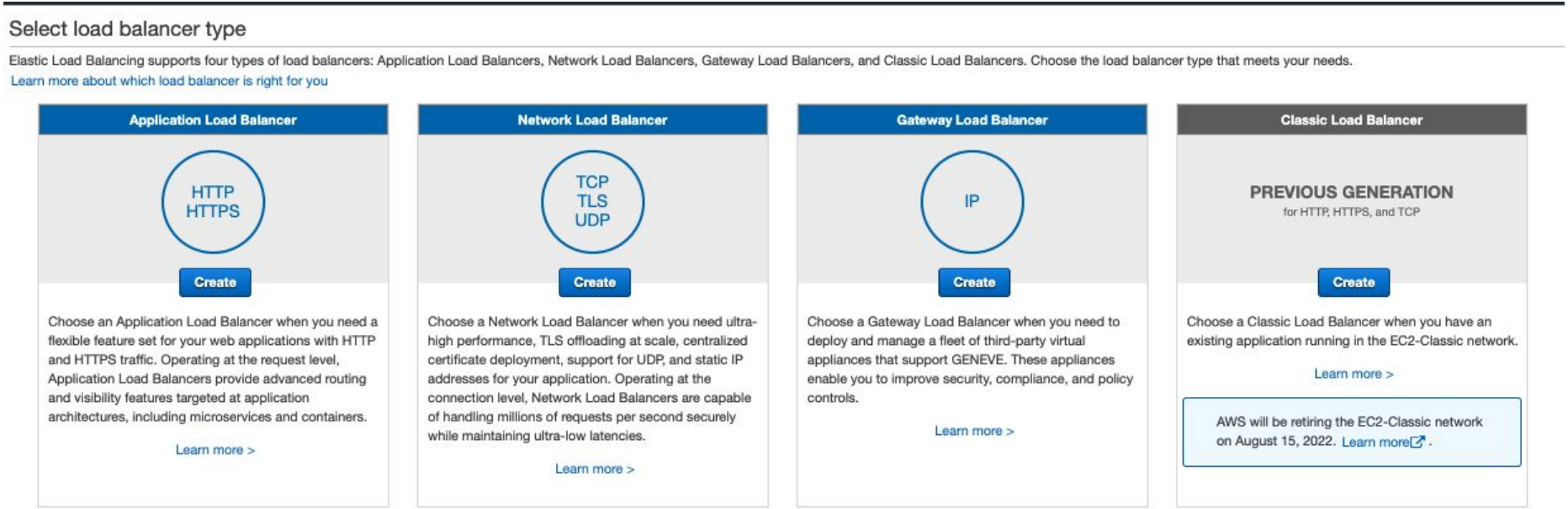


Disponibilidad: Load Balancers

Existen 4 tipos:

- ➔ Clásico
- ➔ De aplicación
- ➔ De red
- ➔ Gateway LB

Disponibilidad: Load Balancers



Disponibilidad: Load Balancers

➔ Clásico

Proporciona equilibrio de carga básico en varias instancias de Amazon EC2 y funciona tanto en el nivel de solicitud como en el nivel de conexión. Load Balancer layer 4.

ATENCIÓN: se retira en Agosto 2022

Disponibilidad: Load Balancers

➔ De Red

Adecuado para equilibrar la carga del tráfico del protocolo de control de transmisión (TCP) y UDP, del protocolo de datagramas de usuario y de Transport Layer Security (TLS)
A nivel de conexión (capa 4), dirige el tráfico hacia destinos dentro de Amazon Virtual Private Cloud (Amazon VPC) y es capaz de controlar millones de solicitudes por segundo

Disponibilidad: Load Balancers

➔ De Aplicaciones

Adecuado para el equilibrio de carga del tráfico HTTP y HTTPS.
A nivel de solicitud individual (capa 7), dirige el tráfico a los destinos dentro de Amazon Virtual Private Cloud (Amazon VPC) en función del contenido de la solicitud.

Es el reemplazo al LB clásico.

Disponibilidad: Load Balancers

➔ Gateway LB

Es un tipo especial de LBs. Se utiliza para balancear appliances de terceros cómo firewalls o detectores de intrusos.

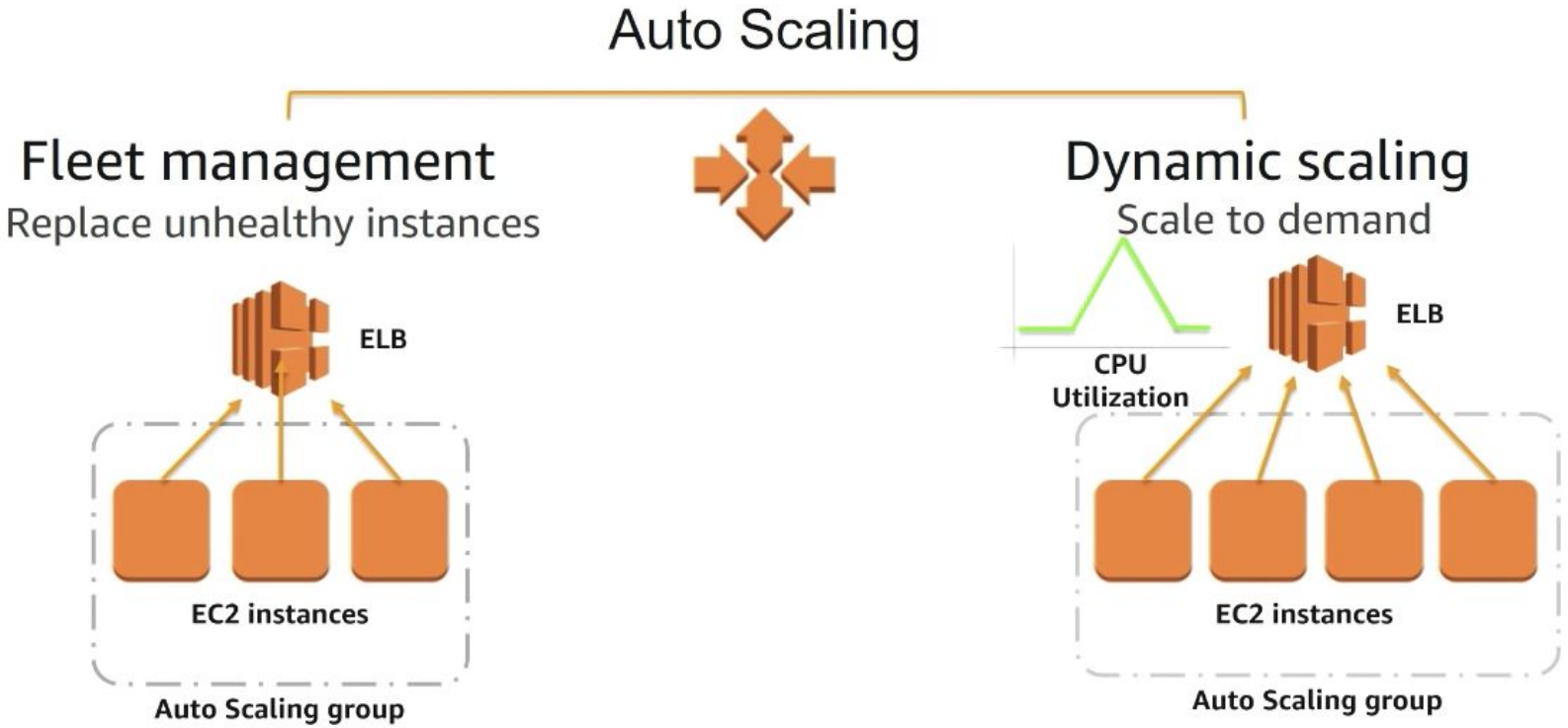
Disponibilidad: Auto-scaling

Es una funcionalidad de EC2 que se utiliza para agregar capacidad de cómputo dinámicamente.

Cumple dos funciones principales:

- ➔ Reemplazar instancias no saludables
- ➔ Escalar instancias de forma horizontal

Disponibilidad: Auto-scaling



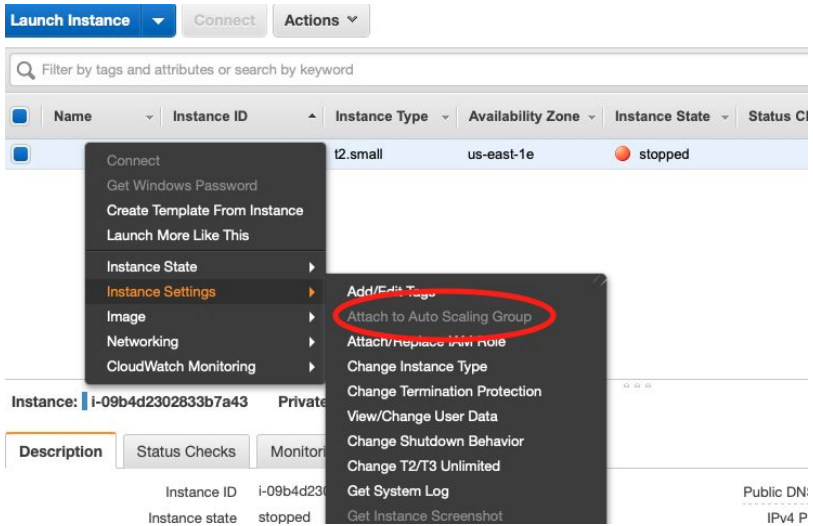
Disponibilidad: Auto-scaling

Los auto scaling groups se pueden crear a partir de una instancia o a partir de un Launch Configuration ó Launch Template

Disponibilidad: Auto-scaling

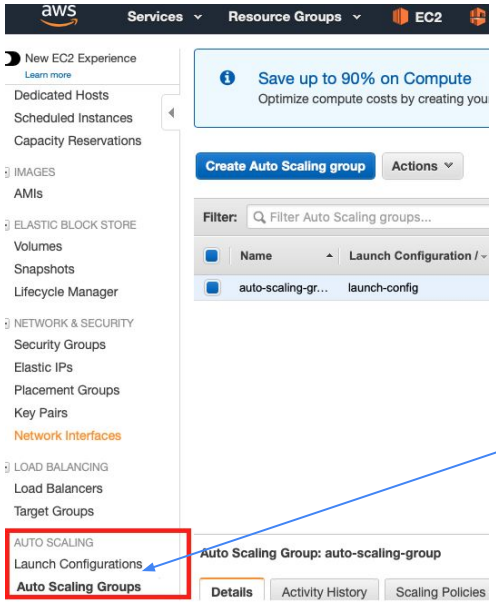
A partir de una instancia

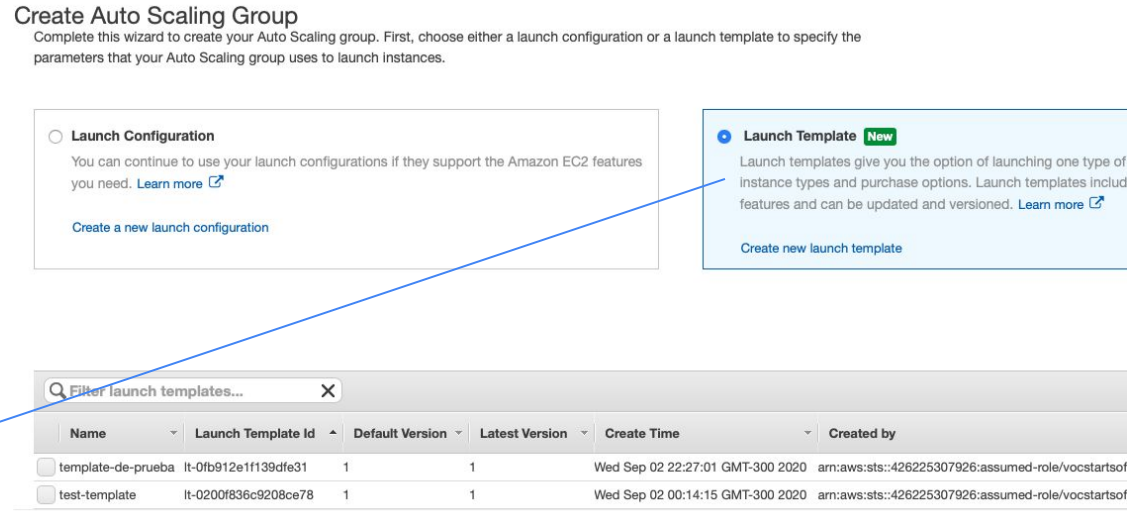
- ➔ La instancia debe estar encendida
- ➔ La AMI debe existir todavía
- ➔ La instancia no puede pertenecer a otro Grupo



Disponibilidad: Auto-scaling

A partir de un launch configuration / Template

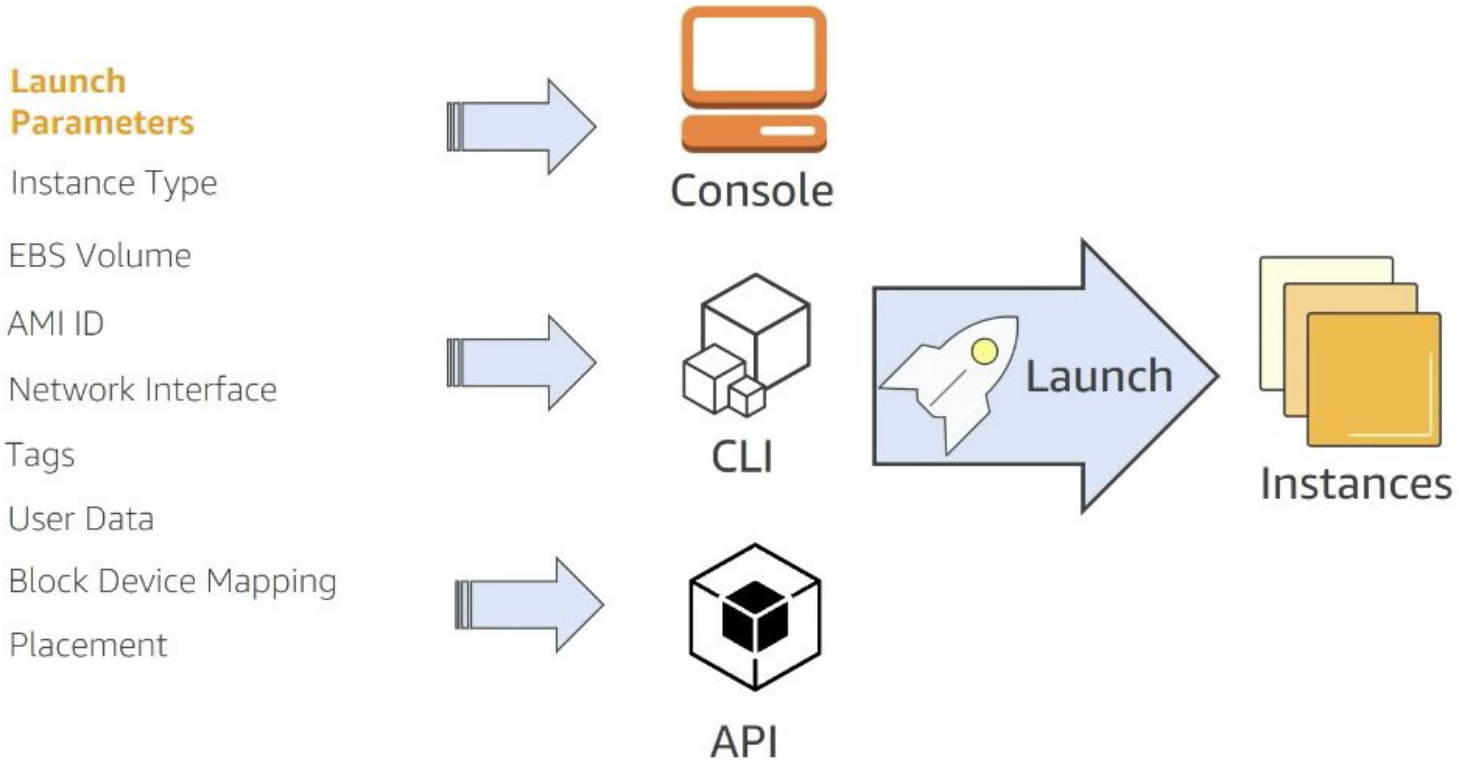




Management

Resources	Availability	Management	Purchase Options
Instances Storage Networking	Regions and AZs Placement Groups Load Balancing Auto Scaling	Deployment Monitoring Administration	On Demand Reserved Spot

Management: Despliegue



Management: Despliegue

Desplegar instancias puede ser tan simple cómo:

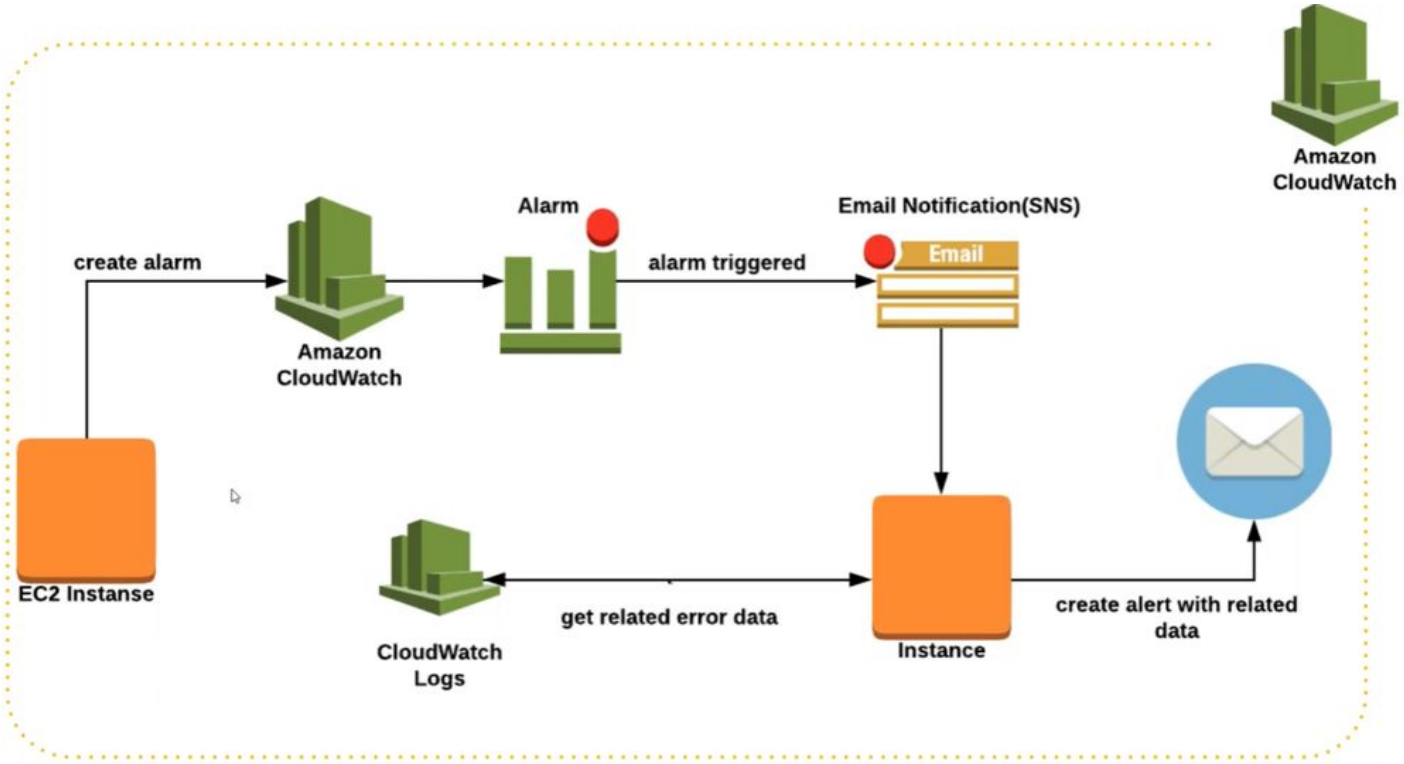
- ➔ Tener un template con todos los datos
- ➔ Usar las APIs y código
- ➔ Embeber en scripts comandos de la Cli

Management: Monitoreo / Trazabilidad

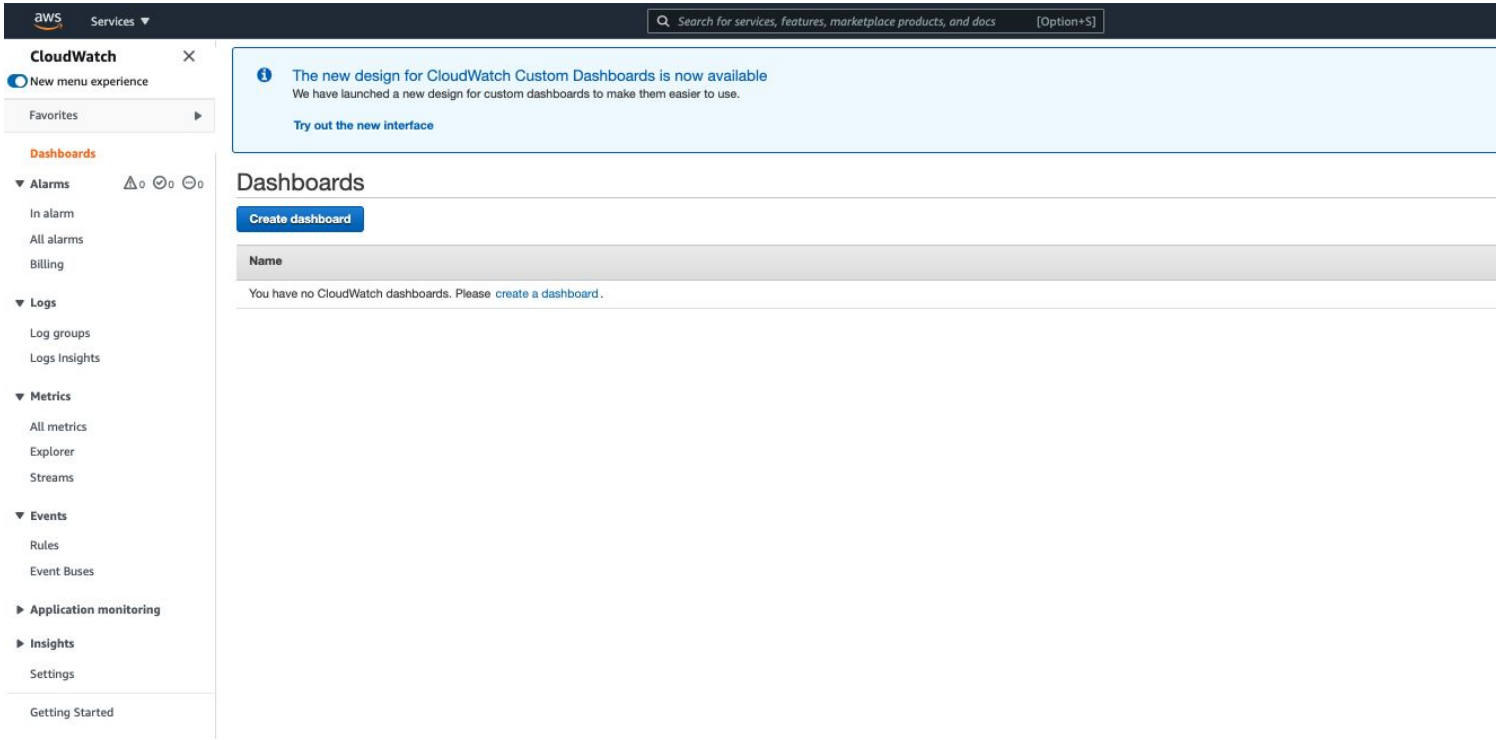
AWS Integra varias herramientas para monitorear nuestra infraestructura y nuestra cuenta:

- ➔ AWS Cloudwatch
- ➔ AWS CloudTrail

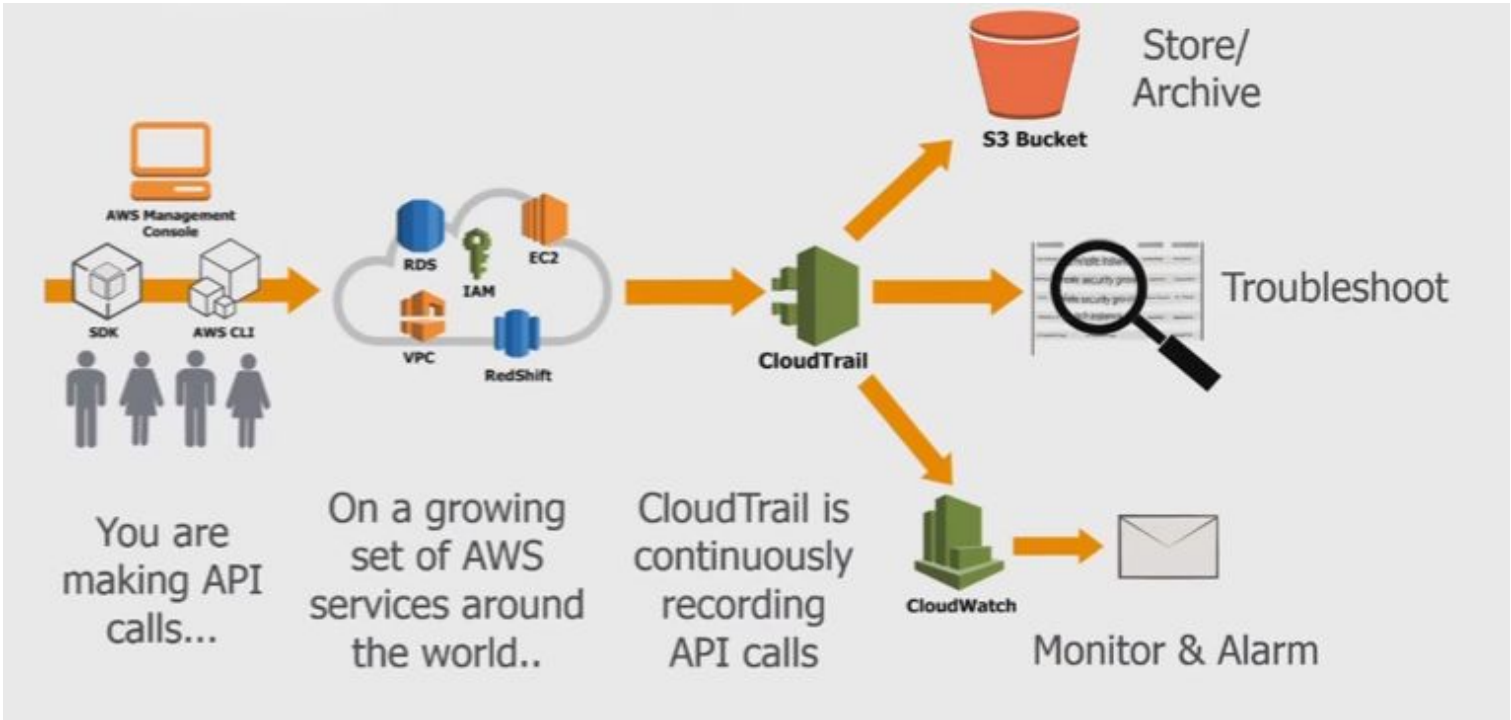
Management: CloudWatch



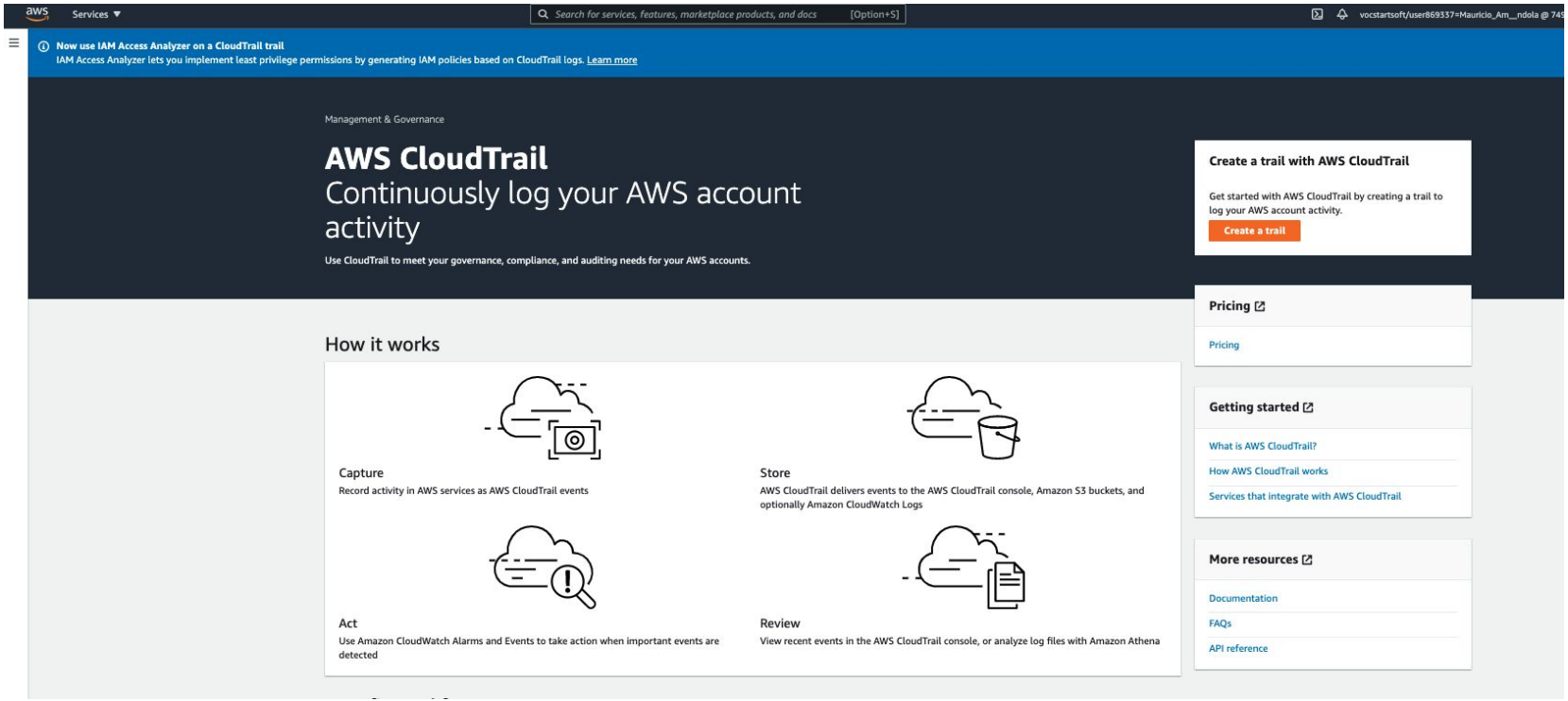
Management: CloudWatch



Management: CloudTrail



Management: CloudTrail



Costos

Resources

Instances
Storage
Networking

Availability

Regions and AZs
Placement Groups
Load Balancing
Auto Scaling

Management

Deployment
Monitoring
Administration

Purchase Options

On Demand
Reserved
Spot

Costos

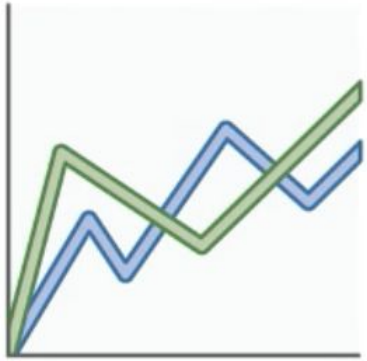
Varias opciones de compra. Dependiendo del estado de madurez de nuestra organización y la adopción de cloud services, es la opción que más nos conviene:

- ➔ On-demand
- ➔ Reserved
- ➔ Spot instances

Costos: On-demand

➔ On-demand

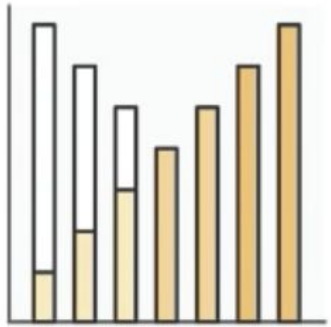
Sin compromisos a largo plazo
Facilidad para escalar
Se paga por segundo!



Costos: Reserved

➔ Reserved

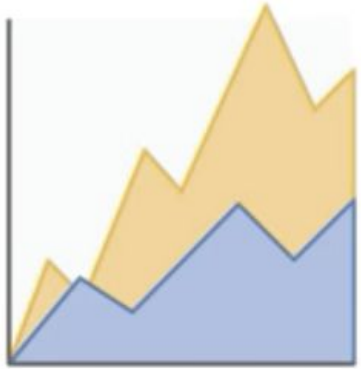
Compromiso de 1 y 3 años
Descuentos significativos en función del tiempo de compromiso
Se paga por segundo!



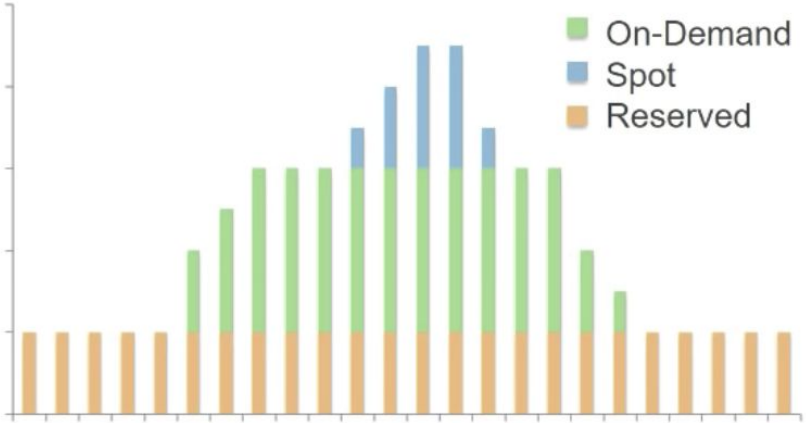
Costos: Spot

➔ Spot

Usar capacidad “de sobra” de ec2
Descuentos de hasta 90% sobre precio a demanda
Trabajos no críticos, tiempo flexible



Costos: Best choice



- ➔ Instancias reservadas para carga conocida o constante
- ➔ On demand para escalar fuera de lo esperado
- ➔ Spot para cosas puntuales, no productivas