Amazon Web Service Cloud Computing

Lic. Ismael P. Rodríguez ismael@lidi.info.unlp.edu.ar









- Amazon Web Services (AWS)
 - Amazon EC2
 - Amazon Storage (EBS, S₃, EFS...)
 - Amazon CloudWatch
 - IPs elásticas
- Consultas





- Amazon Web Services (AWS)
 - Amazon EC2
 - Amazon Storage (EBS, S₃, EFS...)
 - Amazon CloudWatch
 - IPs elásticas
- Consultas





AWS comenzó en 2006 a ofrecer servicios de Cloud a empresas.

http://aws.amazon.com/es/about-aws/

- Los servicios se agrupan en diferentes categorías:
 - Procesamiento de datos.
 - Almacenamiento.
 - Bases de datos.
 - Entrega de contenidos.
 - Mensajería.
 - Red.
 - ...



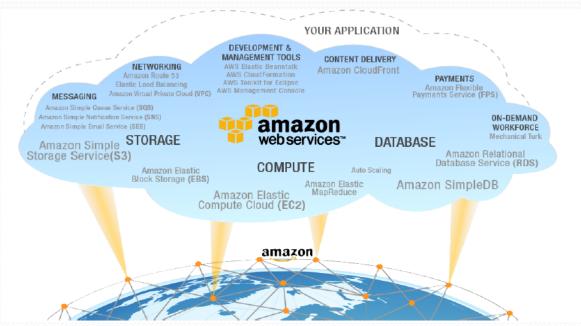


http://aws.amazon.com/es/products/





- Actualmente, ofrece más de 70 servicios que permiten ejecutar una amplia variedad de aplicaciones en la nube:
 - Aplicaciones Big-Data, móviles, alojamiento WEB, arquitecturas tolerantes a fallos, entrega de contenido (Streamming).



Cloud Computing. Cloud Robotics - 2024





Amazon Web Services

Compute & Networking



EC2
Virtual Servers in the Cloud

Route 53
Scalable Domain Name System

VPC
Isolated Cloud Resources

Storage & Content Delivery

CloudFront
Global Content Delivery Network

Glacier
Archive Storage in the Cloud

\$3 Scalable Storage in the Cloud

Storage Gateway
Integrates On-Premises IT Environments with
Cloud Storage

Database

DynamoDB
 Predictable and Scalable NoSQL Data Store

ElastiCache
 In-Memory Cache

RDS

Managed Relational Database Service

Redshift
Managed Petabyte-Scale Data Warehouse
Service

Deployment & Management

CloudFormation Templated AWS Resource Creation

CloudTrail
User Activity and Change Tracking

CloudWatch
Resource and Application Monitoring

Elastic Beanstalk AWS Application Container

P IAM Secure AWS Access Control

OpsWorks
DevOps Application Management Service

Trusted Advisor

AWS Cloud Optimization Expert

Analytics

Data Pipeline Orchestration for Data-Driven Workflows

Elastic MapReduce
Managed Hadoop Framework

KINESIS
Real-time Processing of Streaming Big Data

Mobile Services

Cognito
User Identity and App Data Synchronization

Mobile Analytics
Understand App Usage Data at Scale

SNS
Push Notification Service

App Services

AppStream
Low Latency Application Streaming

CloudSearch
Managed Search Service

Elastic Transcoder
Easy-to-use Scalable Media Transcoding

SES Email Sending Service

SQS

Message Queue Service

Workflow Service for Coordinating Application
Components

Applications

WorkSpaces
Desktops in the Cloud

Zocalo
Secure Enterprise Storage and Sharing
Service







Regiones y número de zonas de disponibilidad:

http://aws.amazon.com/es/about-aws/global-infrastructure/



Región y número de zonas de disponibilidad

China

Pekín (2),

Ningxia (3)

América del Sur

São Paulo (3)

EE.UU. Este

Norte de Virginia (6),

Ohio (3)

Europa EE.UU. Oeste Fráncfort (3).

Norte de

Irlanda (3). California (3), Londres (3). París (3)

Oregón (3)

Asia Pacífico

Mumbai (2),

Seúl (2),

AWS Singapur (3),

GovCloud (EE.UU.-Sídney (3), Oeste) (3)

Tokio (4),

Osaka-Local (1)¹

Canadá

Central (2)



Nueva región (próximamente)

Baréin

Hong Kong, China

Suecia

AWS GovCloud

(EE.UU. Este)









Alta disponibilidad a través de varias zonas de disponibilidad:

Cada región de AWS tiene varias zonas de disponibilidad y varios centros de datos. Esto permite que los clientes interesados en la disponibilidad y el rendimiento de las aplicaciones desean implementarlas en varias zonas de disponibilidad de la misma región a efectos de tolerancia a fallos y baja latencia. Las zonas de disponibilidad están conectadas entre sí con redes de fibra óptica rápidas y privadas, lo que permite diseñar aplicaciones con facilidad que conmute por error entre las zonas de disponibilidad sin interrupciones.







Servicios disponibles por regiones:

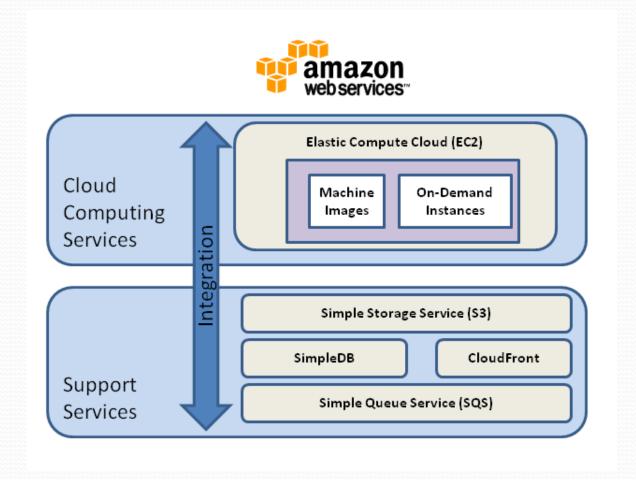
http://aws.amazon.com/es/about-aws/global-infrastructure/regional-product-services/



Servicios ofrecidos:	Norte de Virginia	Oregón	Norte de California	Irlanda	Singapur	Tokio	Sídney	São Paulo	GovCloud	Pekín*
Amazon Elastic Compute Cloud (EC2)	~	•	~	•	~	•	•	•	~	•
Amazon CloudWatch	•	•	•	•	•	•	•	•	•	•
Amazon Virtual Private Cloud (VPC)	~	•	•	~	~	•	~	•	•	~
Amazon Simple Storage Service (S3)	~	~	•	~	~	•	~	•	•	~
Amazon Elastic Block Store (EBS)	~	•	•	•	•	•	•	•	•	•
Auto Scaling	~	~	•	~	~	~	~	~	•	~
Amazon Simple Queue Service (SQS)	~	•	•	•	•	~	•	•	•	•
Amazon Simple Notification Service (SNS)	•	•	•	•	•	•	•	•	~	•
Elastic Load Balancing	~	•	•	~	~	•	•	•	•	~
AWS Support	~	•	•	•	•	~	~	~	•	~
Amazon DynamoDB	~	~	v	•	~	~	~	•	v	~
Amazon Relational Database Service (RDS)	~	•	•	•	~	•	•	•	~	•
Amazon Simple Workflow Service (SWF)	•	•	•	•	•	•	•	•	•	•
Amazon Elastic MapReduce	•	•	~	~	•	•	~	•	~	~
AWS Direct Connect	~	•	•	•	~	•	~	•	•	
AWS CloudFormation	•	•	~	~	~	•	~	•	~	~
VM Import/Export	~	•	v	•	~	~	~	•		~
AWS Elastic Beanstalk	•	•	~	•	•	•	~	•		
AWS Storage Gateway	•	•	*	•	•	•	~	•		•
Amazon SimpleDB								-		











- Capa gratuita de AWS
 - https://aws.amazon.com/es/free/
- 750 horas de uso de instancias **EC2** t2.micro (Linux y/o Windows).
- 5 Gb de almacenamiento S3 estándar.
- 25 Gb de almacenamiento en DinamoDB.
- 750 horas de uso de instancias db-t2-micro Amazon RDS.
- 20 GB de almacenamiento de base de datos Amazon RDS.
- 30 GB de almacenamiento **EBS** (tanto SSD o magnético).
- 10 métricas y 10 alarmas en CloudWatch.
- 250.000 mensajes en AWS IoT.





- Amazon Web Services (AWS)
- Amazon EC2
 - Amazon Storage (EBS, S₃, EFS...)
 - Amazon CloudWatch
 - IPs elásticas
- Consultas





- Permite ejecutar varias instancias ("servidores virtuales") Linux o Windows bajo demanda, permitiendo adquirir tantos equipos como se necesiten para la ejecución de una aplicación.
- Brinda acceso de usuario "root o administrador" al sistema operativo de cada instancia, un Firewall para gestionar el acceso a la red y la libertad para instalar cualquier software.
- Una vez configurada la instancia adecuadamente, se puede guarda la configuración de la VM como una Amazon Machine Image (AMI), permitiendo que se puedan instanciar varias máquinas virtuales bajo demanda con la misma configuración.
- La EC2 API (Query o SOAP) ofrece funcionalidad para iniciar y detener instancias de servidor, configurar permisos de acceso y red o gestionar imágenes de servidor. Cada servidor individual se gestiona usando herramientas Linux o Windows sobre una sesión de shell segura.
- Se factura en función de los recursos consumidos: CPU y transferencia de datos.
- Más información en: http://aws.amazon.com/ec2





Tipos de instancias

Amazon EC2 proporciona una amplia selección de tipos de instancias optimizados para adaptarse a diferentes casos de uso. Los tipos de instancia abarcan varias combinaciones de capacidad de CPU, memoria, almacenamiento y red. Le proporcionan flexibilidad para elegir la combinación de recursos adecuada para sus aplicaciones. Cada tipo de instancia incluye uno o varios tamaños de instancia, lo que le permite escalar sus recursos a los requisitos de su carga de trabajo de destino.

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





Instancia de "Uso/Propósito General" T2 y T3

Las instancias T2 y T3 son <u>instancias de desempeño con ráfagas</u> que proporcionan un nivel base de desempeño de la CPU con la posibilidad de alcanzar ráfagas por encima del nivel básico.

Nombre	vCPU	RAM (GiB)	Créditos de CPU/h
t2.nano	1	0,5	3
t2.micro	1	1,0	6
t2.small	1	2.0	12
t2.medium	2	4,0	24
t2.large	2	8.0	36
t2.xlarge	4	16,0	54
t2.2xlarge	8	32,0	81

Cloud Computing. Cloud Robotics – 2024





Instancia de "Uso/Propósito General" T2 y T3

Name	vCPUs	RAM (GiB)	Baseline Performance/vCPU	CPU Credits earned/hr	Network burst bandwidth (Gbps)	EBS burst bandwidth (Gbps)
t3.nano	2	0.5	5%	6	5	1.50
t3.micro	2	1.0	10%	12	5	1.50
t3.small	2	2.0	20%	24	5	1.50
t3.medium	2	4.0	20%	24	5	1.50
t3.large	2	8.0	30%	36	5	2.05
t3.xlarge	4	16.0	40%	96	5	2.05
t3.2xlarge	8	32.0	40%	192	5	2.05





Créditos máximos por instancia T2 y T3

https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/t2-credits-baselineconcepts.html#t-instance-credit-table

Instance type	CPU credits earned per hour	Maximum earned credits that can be accrued*	vCPUs	Baseline performance per vCPU
t2.nano	3	72	1	5%
t2.micro	6	144	1	10%
t2.small	12	288	1	20%
t2.medium	24	576	2	20%**
t2.large	36	864	2	30%**
t2.xlarge	54	1296	4	22.5%**
t2.2xlarge	81.6	1958.4	8	17%**
t3.nano	6	144	2	5%**
t3.micro	12	288	2	10%**
t3.small	24	576	2	20%**
t3.medium	24	576	2	20%**
t3.large	36	864	2	30%**
t3.xlarge	96	2304	4	40%**
t3.2xlarge	192	4608	8	40%**

Cloud Computing. Cloud Robotics – 2024





Instancia T₂ - Características

- Procesadores Intel Xeon de alta frecuencia con Turbo hasta 3,3 GHz.
- CPU en ráfagas, que se rige por créditos de CPU y rendimiento de base constante.
- Tipo de instancia de uso general a bajo costo e incluida en la capa gratuita (solo t2.micro)
- Equilibrio entre recursos de informática, memoria y red

Las instancias T2 son una buena opción para cargas de trabajo que no usan la CPU por completo, a menudo o de manera constante, pero que de vez en cuando tienen que alcanzar ráfagas (por ejemplo, **servidores web, entornos para desarrolladores y pequeñas bases de datos**).





Instancias de desempeño con ráfagas (T2 y T3)

Las instancias T2 ilimitadas pueden conservar un desempeño de CPU alto durante el tiempo que una carga de trabajo lo necesite. Para la mayoría de las cargas de trabajo generales, las instancias T2 ofrecerán desempeño amplio sin cargo adicional. El precio por hora de la instancia T2 cubre automáticamente todos los picos temporales de uso cuando la utilización de CPU promedio de una instancia T2 es igual o menor al valor de referencia durante un período de 24 horas. Si la instancia requiere un mayor uso de la CPU durante un período prolongado, puede hacerlo con un cargo fijo adicional de 5 centavos de USD por hora de vCPU.





Instancias de desempeño con ráfagas (T2)...

El desempeño base de las instancias T2 y la capacidad de alcanzar ráfagas se rigen por los créditos de la CPU. Toda instancia T2 recibe constantemente créditos de la CPU, cuyo porcentaje depende del tamaño de instancia. Las instancias T2 acumulan créditos de la CPU cuando están inactivas y los utilizan cuando están activas. Un crédito de la CPU proporciona el desempeño de un núcleo de la CPU completa durante un minuto.





Ejemplos de uso de instancias T2 y T3

https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/unlimited-modeexamples.html#t2_unlimited_example

Por ejemplo, una instancia t2.small recibe créditos continuamente con una frecuencia de 12 créditos de CPU por hora. Esta capacidad proporciona un desempeño base equivalente al 20% del núcleo de una CPU (20% x 60 min = 12 min). Si la instancia no utiliza los créditos que recibe, se almacenan en el saldo de créditos de la CPU hasta un máximo de 288 créditos de CPU. Cuando la instancia t2.small necesita alcanzar ráfagas hasta más del 20% de un núcleo, extrae esta capacidad del crédito de la CPU para gestionar este aumento automáticamente.





Con la instancia T2 ilimitada activada, la instancia t2.small puede alcanzar ráfagas por encima del valor de referencia inclusive después de disminuir a cero el saldo de crédito de la CPU. Para la gran mayoría de las cargas de trabajo generales en las que el uso de CPU promedio es igual o menor al desempeño de referencia, el precio básico por hora de la instancia t2.small cubre todas las ráfagas de CPU. Si la instancia se ejecuta con un uso de CPU promedio del 25% (5% por encima del valor de referencia) durante un período de 24 horas después de que el saldo de crédito de CPU haya disminuido a cero, se le cobrará un monto adicional de 5 centavos/hora de vCPU.





Instancias de desempeño con ráfagas (T2 y T3)

- Las instancias T2 se lanzan en modo "standard" por default.
- Las instancias T3 se lanzan en modo "unlimited" por default.
- Un crédito de CPU equivale a un vCPU al 100% de uso en un minuto.
- En una Región se pueden lanzar hasta 20 instancias T2 y 20 instancias T3 simultáneamente.
- Las instancias T₂ y T₃ ilimitadas, pagan 0,05 USD por hora de vCPU para Linux.





Instancia M4 – Para Bases de datos pequeñas y medianas

- Procesadores Intel Xeon® E5-2686 v4 (Broadwell) de 2,3 GHz o Intel Xeon® E5-2676 v3 (Haswell) de 2,4 GHz
- Optimizados para EBS por defecto sin coste adicional
- Soporte para redes mejoradas
- Equilibrio entre recursos de informática, memoria y red

Modelo	vCPU	Memoria (GiB)	Almacenamie nto en SSD (GB)	Rendimiento EBS dedicado (Mbps)
m4.large	2	8	Solo para EBS	450
m4.xlarge	4	16	Solo para EBS	750
m4.2xlarge	8	32	Solo para EBS	1 000
m4.4xlarge	16	64	Solo para EBS	2 000
m4.10xlarge	40	160	Solo para EBS	4 000

Cloud Computing. Cloud Robotics – 2024





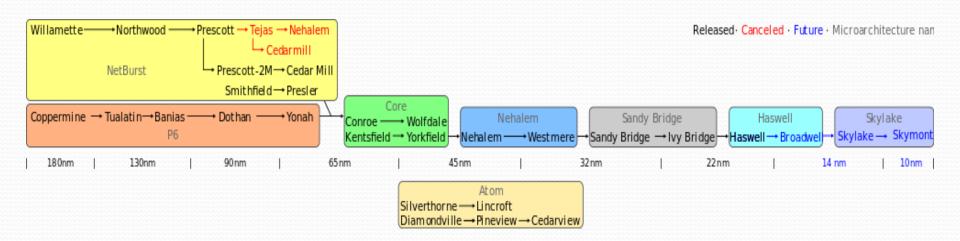
Instancia M5 – De última generación

- Procesadores Intel Xeon® Platinum 8175 de 2,5 GHz.
- Almacenamiento de instancias basado en SSD basado en NVMe para un rápido rendimiento de E/S
- Hasta 25 Gbps de ancho de banda de red con redes mejoradas.





Micro-arquitectura de procesadores



https://en.wikipedia.org/wiki/List_of_Intel_CPU_microarchitectures





Precios de instancias

Más información en: https://aws.amazon.com/es/ec2/pricing/

	CPU virtual	ECU	Memoria (GiB)	Almacenamiento de instancias (GB)	Uso de Linux/UNIX		
Uso general – Generación actual							
t3.nano	2	Variable	0,5 GiB	Solo EBS	0,0052 USD por hora		
t3.micro	2	Variable	1 GiB	Solo EBS	0,0104 USD por hora		
t3.small	2	Variable	2 GiB	Solo EBS	0,0208 USD por hora		
t3.medium	2	Variable	4 GiB	Solo EBS	0,0416 USD por hora		
t3.large	2	Variable	8 GiB	Solo EBS	0,0832 USD por hora		
t3.xlarge	4	Variable	16 GiB	Solo EBS	0,1664 USD por hora		
t3.2xlarge	8	Variable	32 GiB	Solo EBS	0,3328 USD por hora		
t2.nano	1	Variable	0,5 GiB	Solo EBS	0,0058 USD por hora		
t2.micro	1	Variable	1 GiB	Solo EBS	0,0116 USD por hora		
t2.small	1	Variable	2 GiB	Solo EBS	0,023 USD por hora		





AMI (Amazon Machine Instance)

Imágenes preconfiguradas (S.O., Software) usadas para iniciar VMs en EC2.

Existen AMIs gratuitas y pagas. Un usuario/clientes también puede publicar sus propias AMIs (gratuitas o pagas).

AWS marketplace:

https://aws.amazon.com/marketplace/ref=mkt_ste_amis_redirect?b_k=291







Características del servicio

- Elástico: permite aumentar o disminuir la capacidad en cuestión de minutos.
- Control total sobre las instancias: acceso como root a las Vms (detener, reiniciar, terminar instancias).
- **Flexible:** permite ejecución de varios tipos de instancias, S.O., Software, dado que Amazon permite seleccionar la configuración de memoria, CPU, almacenamiento, etc.
- **Fiable:** es un entorno confiable. Amazon tiene un contrato de compromiso a nivel de servicio de disponibilidad a 99,95% en cada región.
- Seguro: Ofrece Firewall y aislación con VPC (Virtual Private Cloud) y acceso por Ipsec VPN.
- **Económico:** ofrece una tarifa accesible.





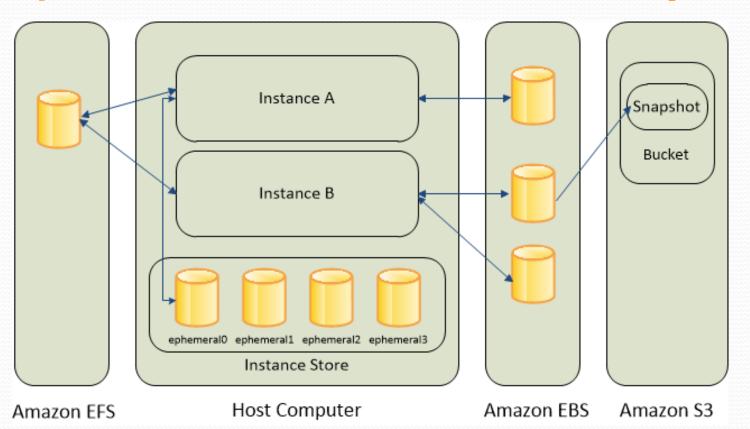
- Amazon Web Services (AWS)
 - Amazon EC2
- Amazon Storage (EBS, S₃, EFS...)
 - Amazon CloudWatch
 - IPs elásticas
- Consultas





Amazon Storage

https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/Storage.html







Amazon Storage - EBS

- Servicio de almacenamiento basado en volúmenes.
 - Es como utilizar un disco duro sin formatear.
 - Está diseñado para ser utilizado con las instancias EC2:
 - Los volúmenes se montan como dispositivos y se accede a ellos a través de la red.
 - Los volúmenes persisten con independencia de las instancias.
 - Una instancia puede montar varios volúmenes (en la misma zona de disponibilidad).

Características

- Tamaños entre 4 GB y 16 TB.
- Replicación dentro de la misma zona de disponibilidad.
- Instantáneas (snapshots) en S3
 - Son incrementales (sólo almacenan los cambios).
 - Se replican entre varias zonas de disponibilidad.
 - Permiten recuperar el volumen en caso de fallo.

https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/AmazonEBS.html





Amazon Storage - EBS

- Volumen "General Purpose SSD (gp2)" permite hasta 3000 IOPS.
- Volumen "Provisioned IOPS SSD (io1), permite hasta 32000 IOPS y un throughput de 500 MB/s.
- Volumen "Magnetic" y HDD (st1 o sc1) permite hasta 250 MB/s de throughput.

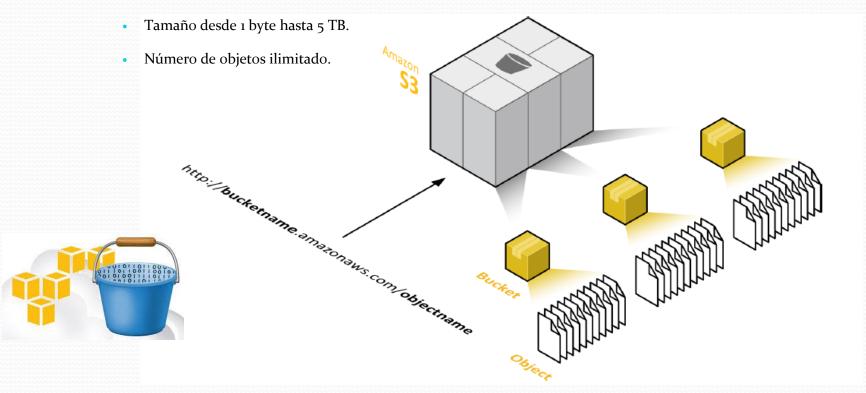
https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/AmazonEBS.html





Amazon Storage - S3

- Almacenamiento de objetos (archivos y carpetas) a escala web.
 - Los objetos están contenidos en depósitos (buckets).
 - Cada objeto puede ser accedido mediante una URL única.







Amazon Storage - S3

Características

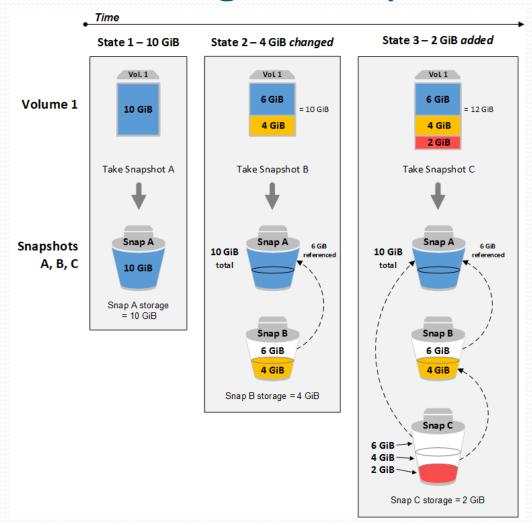
- Redundancia en múltiples dispositivos de diferentes datacenters.
 - Durabilidad: 99,99999999%.
 - Disponibilidad: 99,99% durante un año dado.
- Permite acceso concurrente de lectura/escritura.
 - Los cambios necesitan un tiempo para propagarse a través de las réplicas.
- Los objetos no salen de una región a menos que se transfieran manualmente.
- Usado con EC2:
 - Almacenamiento de AMIs y snapshots.
 - Alojamiento de sitios WEBs estáticos.







Amazon Storage - Snapshots







Amazon Storage – EFS Elastic File System

- Sistema de archivo que debe crearse en la misma VPC (Virtual Private Cloud) que la instancia donde se va a montar el volumen.
- Al momento de lanzar la instancia hay que configurar el script para montar el volumen:

```
#!/bin/bash
yum update -y
yum install -y nfs-utils
FILE_SYSTEM_ID=fs-xxxxxxxx
AVAILABILITY_ZONE=$(curl -s http://169.254.169.254/latest/meta-data/placement/availability-zone )
REGION=${AVAILABILITY_ZONE:0:-1}
MOUNT_POINT=/mnt/efs
mkdir -p ${MOUNT_POINT}
chown ec2-user:ec2-user ${MOUNT_POINT}
echo ${FILE_SYSTEM_ID}.efs.${REGION}.amazonaws.com:/ ${MOUNT_POINT} nfs4
nfsvers=4.1,rsize=1048576,wsize=1048576,hard,timeo=600,retrans=2,_netdev 0 0 >> /etc/fstab
mount -a -t nfs4
```





Amazon Storage – EFS

• Se puede verificar el sistema de archivos montado, desde la consola de la instancia:

\$ df -T					
Filesystem	Type	1K-blocks	Used	Available	Use% Mounted on
/dev/xvda1	ext4	8123812	1949800	6073764	25% /
devtmpfs	devtmpfs	4078468	56	4078412	1% /dev
tmpfs	tmpfs	4089312	0	4089312	0% /dev/shm
efs-dns	nfs4	9007199254740992	0	9007199254740992	0% /mnt/efs



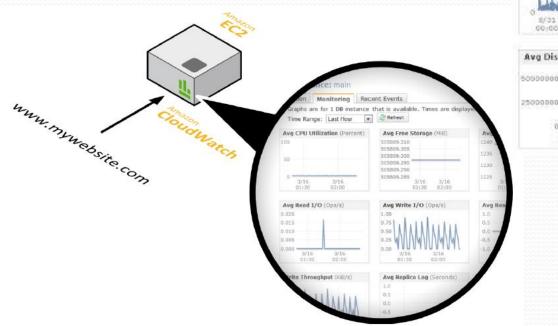


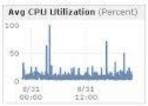
- Amazon Web Services (AWS)
 - Amazon EC2
 - Amazon Storage (EBS, S₃, EFS...)
- Amazon CloudWatch
 - IPs elásticas
- Consultas





- Monitorización de recursos AWS y aplicaciones.
 - Recopila y visualiza gráficamente métricas de diferentes servicios:
 - EC2, EBS, ELB, RDS, aplicaciones de usuario, ..., etc.
 - Permite definir alarmas y realizar acciones automatizadas.





Avg Disk Writes (Bytes)

6/31

25000000



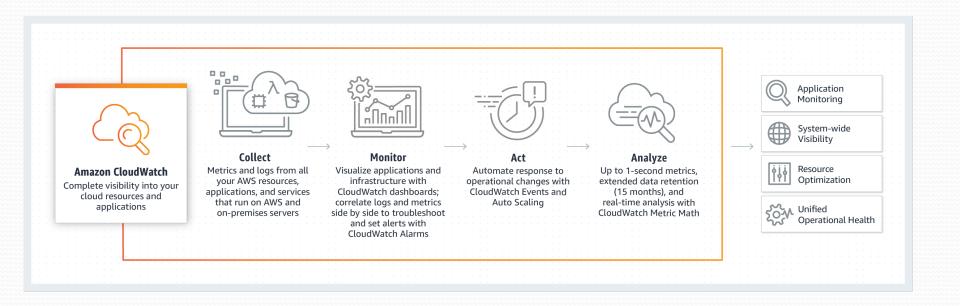
8/31

Avg Disk Reads (Bytes)

5000000







https://aws.amazon.com/es/cloudwatch/features/





- Recopilación: registros y métricas
- Monitoreo: Vista operativa unificada y alarmas.
- Acción: Auto Scaling y automatización de respuestas.
- Análisis: Estadísticias históricas de 15 meses de antigüedad, como análisis en tiempo real.





Capa gratuita de AWS:

Métricas	Métricas de monitoreo básico (frecuencia de 5 minutos) 10 métricas de monitoreo detallado (frecuencia de 1 minuto) 1 millón de solicitudes a API				
Panel	3 paneles para hasta 50 métricas al mes				
Alarmas	10 alarmas (no se aplica a alarmas de alta resolución)				
Registros	5 GB de datos (ingreso y almacenamiento de archivos)				
Eventos	Se incluyen todos los eventos excepto los personalizados				





- Amazon Web Services (AWS)
 - Amazon EC2
 - Amazon Storage (EBS, S₃, EFS...)
 - Amazon CloudWatch
- IPs elásticas
- Consultas





Amazon IPs elásticas

- Direcciones IP asociadas a una cuenta de usuario, no a una instancia en concreto.
- Pueden ser reasignadas dinámicamente.
 - Si la instancia se detiene o se termina, la IP elástica se libera.
- Se resuelven...
 - Desde afuera de AWS a la IP pública de la instancia.
 - Desde adentro de AWS a la IP privada de la instancia.
- Mecanismo de tolerancia a fallos:
 - Si una instancia falla, se asigna su IP elástica a otra y el servicio no se interrumpe.
- Por default están limitadas a 5 direcciones de IP por Región (por limitación del direccionamiento IPv₄).

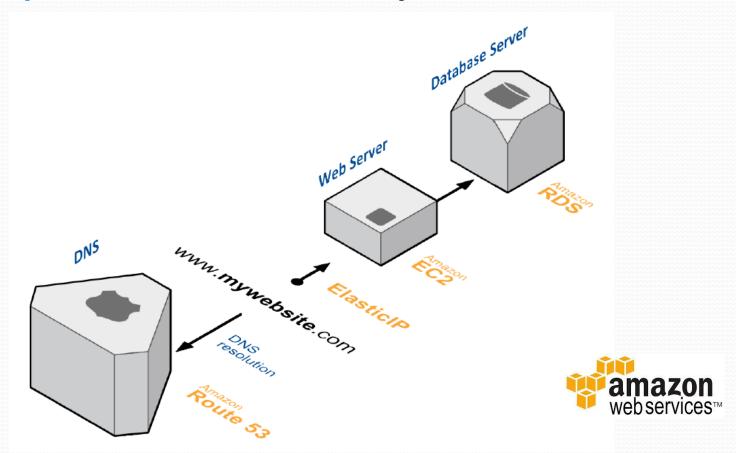






Amazon IPs elásticas

Ejemplo: actualización de un servidor sin interrupción de servicio.

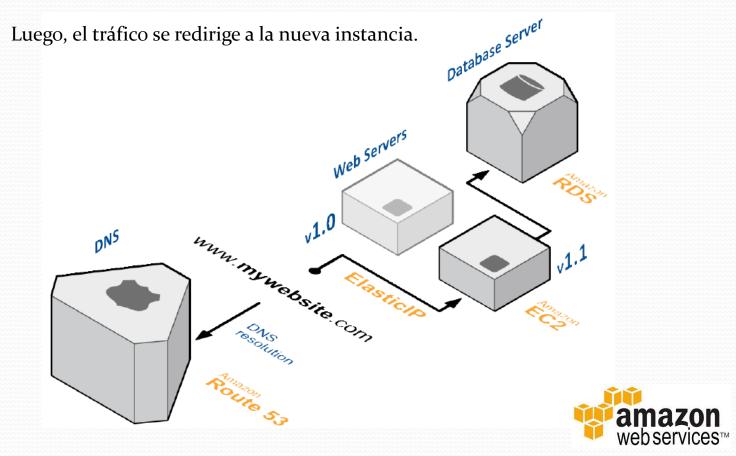






Amazon IPs elásticas

Se reasigna la IP elástica a la instancia con el servidor actualizado.







Amazon Web Services (AWS)

• IMPORTANTE: tener en cuenta los costos:

Entender los cargos de IP elásticas

Precios de las instancias bajo demanda de EC2 – Amazon Web Services





- Amazon Web Services (AWS)
 - Amazon EC2
 - Amazon Storage (EBS, S₃, EFS...)
 - Amazon CloudWatch
 - IPs elásticas
- ¿Consultas?