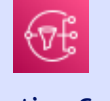
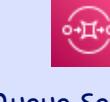



AWS Servicios de Informática / Amazon Elastic Compute Cloud (Amazon EC2)



Servicios de Integración de Aplicaciones

- Simple Notification Service SNS  Envío de Notificaciones o Mensaje entre componentes de SW o Subscriptores, Tópicos, Publicadores
 - Simple Queue Service SQS  Manejo de Colas Mensajes
 - API Gateway  Crear, Publicar, Gestionar, Monitorear API REST y websocket
- Reducir interdependencia
Loose coupling
Ayuda: Microservicios




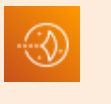
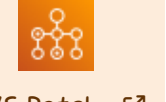
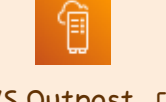
Tipo de Instancias

- Uso General**
 - combinación equilibrada de recursos informáticos, de memoria y de red
 - Para distintas cargas de trabajo por cortos periodos
 - Sitios y aplicaciones web, entornos de desarrollo, servidores de compilación, repositorios de código etc
 - Optimizadas para Informática**
 - procesadores de alto rendimiento
 - procesamiento por lotes, entrega de anuncios, codificación de video, videojuegos, creación de modelos científicos, análisis distribuido e inferencia de machine learning basado en CPU
 - Optimizadas para Memoria**
 - rendimiento rápido para las cargas de trabajo que procesan grandes conjuntos de datos en la memoria
 - Aplicaciones de memoria intensiva como bases de datos de código abierto, cachés en memoria, SAP
 - Informática Acelerada**
 - aceleradores de hardware, o coprocesadores, basadas en GPU
 - Aprendizaje automático, informática de alto rendimiento, dinámica de fluidos computacional, finanzas computacionales, análisis sísmico, reconocimiento de voz, vehículos autónomos ETC
 - Optimizadas para Almacenamiento**
 - acceso de escritura y lectura secuencial alto a conjuntos de datos muy grandes en el almacenamiento local
 - Bases de datos relacionales (p. ej., MySQL, MariaDB, PostgreSQL) y bases de datos que no son de SQL (KeyDB, ScyllaDB, Cassandra), Data Warehousing, Sistemas de archivos Distribuidos
 - Nomenclatura**
 - a – AMD processors
 - g – AWS Graviton processors
 - i – Intel processors
 - d – Instance store volumes
 - n – Network optimization
 - b – Block storage optimization
 - e – Extra storage or memory
 - z – High frequency
- Familia + Generación + Procesador.Tamaño (vCPU, GiB Memoria, Red, Almacenamiento) ejemplo: M4g.medium

Opciones de Compra

- Bajo demanda**
 - Paga, por segundo de encendido de instancias
 - cargas de trabajos flexibles, irregulares
 - sin compromiso a largo plazo
 - Sin costo fijo y pago anticipado
- Spot** 
 - Para cargas de trabajo no críticas y flexibles, tolerantes a fallos, desarrollo/pruebas, tiempo flexible, cargas de trabajo sin estado
 - Nacen de la capacidad excedente (no utilizada)
 - Discount up to 90%
- Reserved Instances** 
 - Contratar instancia por un periodo de 1 o 3 años. Compromiso de pago,
 - Uso comprometido en estado constante
 - Tipos**
 - Estándar**
 - Descuento hasta 72%
 - Se puede cambiar la zona de disponibilidad, el tamaño de la instancia
 - Convertible**
 - Descuento hasta 66%
 - Se puede cambiar familias de instancias, sistemas operativos o tenencias
- Savings Plans**
 - comprometerse a una cantidad constante de uso, en USD por hora, durante un periodo de 1 o 3 años
 - Flexibilidad de cambios: Region, Size, Family, Tenancy, SO,
 - Tipos**
 - EC2**
 - Solo para instancia EC2
 - Compute**
 - Se puede combinar con otros servicios Serverless (Lambda), Container Services
- Dedicated Instances**
 - Pago por hora, por instancias que se ejecutan en hardware de un solo inquilino
- Dedicated Hosts**
 - host físico que esté totalmente dedicado a ejecutar sus instancias y traiga sus licencias de software existentes por socket, por núcleo o por máquina virtual para reducir los costos

Otros Servicios de Informática

- AWS Elastic Beanstalk** 
 - implementar y escalar servicios y aplicaciones web desarrollados con Java, .NET, PHP, Node.js, Python, Ruby, Go y Docker en servidores familiares como Apache, Nginx, Passenger e IIS.
 - PAAS que maneja automáticamente: Infrastructure provisioning and configuration, Deployment, Load balancing, Automatic scaling, Health monitoring, Analysis and debugging, Logging
- AWS Lambda** 
 - servicio informático sin servidor/Serverless controlado por eventos. ejecutar código sin aprovisionar ni administrar servidores
 - Precio por # innovación + duración ms & memoria consumida (28 MiB - 10 GiB de la ejecución del código)
 - Java, Go, PowerShell, Node.js, C#, Python y Ruby,
- Contenedores** 
 - Método de virtualización del sistema operativo que permite ejecutar una aplicación y sus dependencias en procesos aislados de recursos
 - contienen todo lo que el software necesita para ejecutarse, como bibliotecas, herramientas del sistema, código y el tiempo de ejecución.
 - Elastic Container Service ECS
 - Elastic Kubernetes Service EKS
 - Elastic Container Registry ECR
 - alojamiento de alto rendimiento de Imágenes de contenedores
 - AWS Fargate
 - Informática sin servidor para contenedores con ECS o EKS
 - Casos de Usos**
 - Deploy microservices to speed innovation
 - Batch processing
 - Migrate legacy applications without requiring code changes
 - Accelerate machine learning
- Amazon LightSail** 
 - aprovisionamiento de servidores privados virtuales, sin experiencia en programación y administración
 - imágenes con sistemas operativos base, stacks de desarrollo como LAMP, LEMP (Nginx) y SQL Server Express, y aplicaciones como WordPress, Drupal y Magento.
- AWS Batch** 
 - planifica, programa y ejecuta sus cargas de trabajo de informática por lotes en AWS Fargate, Amazon EC2 y las instancias de spot.
- AWS Outpost** 
 - Extender recursos y servicios de AWS a las instalaciones locales (DC onpremis)
 - servidores 1U y 2U, bastidores Rack 42U

Escalabilidad

- Amazon EC2 Auto Scaling** 
 - Agrega o elimina instancias EC2 según las condiciones que defina. Mantener el estado y la disponibilidad de la flota de EC2
 - Supervisar el estado de las instancias en ejecución,
 - Reemplazar instancias deterioradas automáticamente,
 - Equilibrar la capacidad en todas las zonas de disponibilidad
 - instancias en varias zonas de disponibilidad, pero no en regiones de AWS.
 - Tipos**
 - Calendarizadas**
 - El escalado basado en una programación le permite escalar su aplicación antes de los cambios de carga conocidos
 - Dynamic**
 - Se define una alarma a través de una métrica y umbral (ejemplo CPU % utilización) para ajusta horizontalmente la capacidad deseada del grupo de EC2 cuando se activa la alarma
 - Predictive**
 - utiliza el aprendizaje automático para programar el número correcto de instancias EC2 en previsión de los cambios de tráfico que se aproximan.
- Elastic Load Balancer** 
 - Distribuir el tráfico de red para mejorar la escalabilidad de las aplicaciones
 - toma decisiones de enrutamiento en la capa de aplicación (HTTP/HTTPS/gRPC), Modelo OSI de capa 7
 - Application Load Balancer (ELB)**
 - Target: IP, Instance, Lambda
 - toma decisiones de enrutamiento en la capa de transporte (TCP/UDP/SSL). Puede manejar millones de solicitudes por segundo. Modelo OSI de capa 4
 - Network Load Balancer (NLB)**
 - Targte types: IP, Instance, Application Load Balancer
 - Implementar y ejecutar dispositivos virtuales de terceros,
 - Gateway Load Balancer (GLB)**
 - Distribuir tráfico Puerta de enlace de capa 3 + Equilibrio de carga de capa 4
 - Target: IP, Instance

Creado por:
Virginia Gonzalez Carrasquero

