

AMAZON EC2

1. INTRODUCCIÓN	2
2. BENEFICIOS	2
2.1. INFORMÁTICA A ESCALA WEB ELÁSTICA.....	2
2.2. TOTALMENTE CONTROLADO.....	2
2.3. SERVICIOS DE HOSPEDAJE EN LA NUBE FLEXIBLES.....	2
2.4. INTEGRADO.....	2
2.5. DE CONFIANZA.....	3
2.6. SEGURIDAD.....	3
2.7. ASEQUIBILIDAD.....	3
2.8. FÁCIL DE COMENZAR.....	3
3. TIPOS DE INSTANCIAS	3
3.1. USO GENERAL	3
3.1.1. T2	4
3.1.2. M5.....	4
3.1.3. M4.....	5
3.2. OPTIMIZADAS PARA INFORMÁTICA	6
3.2.1. C5	6
3.2.2. C4.....	7
3.3. OPTIMIZADAS PARA MEMORIA.....	8
3.3.1. X1e.....	8
3.3.2. X1.....	8
3.3.3. R4.....	9
3.4. INFORMÁTICA ACELERADA.....	10
3.4.1. P3	10
3.4.2. P2	10
3.4.3. G3	11
3.4.4. F1	12
3.5. OPTIMIZADAS PARA ALMACENAMIENTO	12
3.5.1. H1.....	12
3.5.2. I3.....	13
3.5.3. D2.....	14
4. CASOS DE ESTUDIO.....	14
4.1. EXPEDIA	14
4.2. AIRBNB.....	15
4.3. NETFLIX	16
5. REFERENCIAS	17

1.INTRODUCCIÓN

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 facilitar a los desarrolladores el uso de la informática en la nube a escala de la Web.

La sencilla interfaz de servicios web de Amazon EC2 permite obtener y configurar la capacidad con una fricción mínima. Proporciona un control completo sobre los recursos informáticos y puede ejecutarse en el entorno informático acreditado de Amazon. Amazon EC2 reduce el tiempo necesario para obtener y arrancar nuevas instancias de servidor en cuestión de minutos, lo que permite escalar rápidamente la capacidad, ya sea aumentándola o reduciéndola, según cambien sus necesidades. Amazon EC2 cambia el modelo económico de la informática, ya que solo tendrá que pagar por la capacidad que realmente utilice. Amazon EC2 les brinda a los desarrolladores las herramientas necesarias para crear aplicaciones resistentes a errores y para aislarlas de los casos de error comunes.

2.BENEFICIOS

- INFORMÁTICA A ESCALA WEB ELÁSTICA

Amazon EC2 permite aumentar o disminuir la capacidad en minutos, no en horas ni en días. Puede enviar una, cientos o incluso miles de instancias de servidor simultáneamente. También puede usar Auto Scaling de Amazon EC2 para conservar la disponibilidad de su flota de EC2 y aumentar o disminuir automáticamente la escala de su flota en función de sus necesidades para maximizar el nivel de desempeño y minimizar los costos. Para ajustar la escala de varios servicios, puede usar AWS Auto Scaling.

- TOTALMENTE CONTROLADO

Tiene control total de las instancias, incluido el acceso a raíz y la capacidad para interactuar con estas como lo haría con cualquier máquina. Puede detener cualquier instancia y mantener los datos en su partición de arranque para luego reiniciar la misma instancia a través de las API del servicio web. Las instancias se pueden reiniciar de manera remota con las API del servicio web y también tendrá acceso a los resultados de su consola.

- SERVICIOS DE HOSPEDAJE EN LA NUBE FLEXIBLES

Tendrá la posibilidad de elegir entre varios tipos de instancia, sistemas operativos y paquetes de software. Amazon EC2 permite seleccionar una configuración de memoria, CPU y almacenamiento de la instancia, así como el tamaño de la partición de arranque óptimo para su sistema operativo y su aplicación. Por ejemplo, entre las opciones de sistemas operativos se incluyen varias distribuciones de Linux y Microsoft Windows Server.

- INTEGRADO

Amazon EC2 se integra con la mayoría de los servicios de AWS, como Amazon Simple Storage Service (Amazon S3), Amazon Relational Database Service (Amazon RDS) y Amazon Virtual Private Cloud (Amazon VPC) para suministrar una solución segura y completa para informática, procesamiento de consultas y almacenamiento en la nube en una amplia variedad de aplicaciones.

- DE CONFIANZA

Amazon EC2 ofrece un entorno de elevada confianza en el que las instancias de sustitución se pueden enviar con rapidez y anticipación. El servicio se ejecuta en los centros de datos y la infraestructura de red acreditados de Amazon. El compromiso del Acuerdo de nivel de servicios de Amazon EC2 es ofrecer una disponibilidad del 99.99% en todas y cada una de las regiones de Amazon EC2.

- SEGURIDAD

Para AWS, la seguridad en la nube es la principal prioridad. Como cliente de AWS, se beneficiará de una arquitectura de red y un centro de datos diseñados para satisfacer los requisitos de seguridad de las organizaciones más exigentes. Amazon EC2 funciona junto con Amazon VPC para proporcionar una funcionalidad de red sólida y segura para sus recursos informáticos.

- ASEQUIBILIDAD

Amazon EC2 le permite disfrutar de los beneficios financieros del escalado de Amazon. Pagará una tarifa muy baja por la capacidad de cómputo que realmente utilice. Consulte las opciones de compra de las instancias de Amazon EC2 para obtener más información.

- FÁCIL DE COMENZAR

Existen varias maneras de comenzar a utilizar Amazon EC2. Puede usar la consola de administración, las herramientas de línea de comandos (CLI) o los SDK de AWS. Puede empezar a usar AWS de manera gratuita. Para obtener más información, consulte nuestros tutoriales.

3.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 redes. 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 según los requisitos de la carga de trabajo de destino.

USO GENERAL

- T2

Las instancias T2 son instancias de desempeño con ráfagas que proporcionan un nivel base de desempeño de la CPU con la posibilidad de alcanzar ráfagas por encima del nivel básico.

- Procesadores Intel Xeon de alta frecuencia
- CPU en ráfagas, que se rige por créditos de CPU y desempeño de base constante
- Tipo de instancia de uso general de menor costo e incluida en la capa gratuita
- Equilibrio entre recursos de informática, memoria y red

Casos de uso: Sitios y las aplicaciones web, entornos de desarrollo, servidores de versiones, repositorios de código, microservicios, entornos de pruebas y reproducción, y aplicaciones empresariales.

Modelo	CPU virtual	Créditos por hora de la CPU	Memoria (GiB)	Almacenamiento
t2.nano	1	3	0,5	Solo EBS
t2.micro	1	6	1	Solo EBS
t2.small	1	12	2	Solo EBS
t2.medium	2	24	4	Solo EBS
t2.large	2	36	8	Solo EBS
t2.xlarge	4	54	16	Solo EBS
t2.2xlarge	8	81	32	Solo EBS

- M5

- Procesadores Intel Xeon® Platinum 8175 de 2,5 GHz con conjunto nuevo de instrucciones de extensiones vectoriales avanzadas de Intel (AVX-512)
- Nuevo tamaño de instancia más grande, m5.24xlarge, que ofrece 96 vCPU y 384 GiB de memoria
- Optimizado para EBS de manera predeterminada y desempeño de EBS mayor en tamaños de instancia menores
- Hasta 25 Gbps de ancho de banda de red con redes mejoradas
- Necesita una AMI HVM que incluye controladores para ENA y NVMe
- Con tecnología del nuevo sistema Nitro ligero, una combinación de hardware dedicado e hipervisor ligero

Casos de uso: Las bases de datos pequeñas y medianas, las tareas de procesamiento de datos que requieren memoria adicional, las flotas de almacenamiento en caché y la ejecución de servidores back-end para SAP, Microsoft SharePoint, la informática en clústeres y otras aplicaciones empresariales.

Modelo	CPU virtual	Memoria (GiB)	Almacenamiento en SSD (GB)	Ancho de banda de EBS dedicado (Mbps)
m5.large	2	8	Solo EBS	Hasta 2 120
m5.xlarge	4	16	Solo EBS	Hasta 2 120
m5.2xlarge	8	32	Solo EBS	Hasta 2 120
m5.4xlarge	16	64	Solo EBS	2 120
m5.12xlarge	48	192	Solo EBS	5 000
m5.24xlarge	96	384	Solo EBS	10 000

- M4
 - 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 de manera predeterminada sin costos adicionales
 - Soporte para redes mejoradas
 - Equilibrio entre recursos de informática, memoria y red

Casos de uso: Las bases de datos pequeñas y medianas, las tareas de procesamiento de datos que requieren memoria adicional, las flotas de almacenamiento en caché y la ejecución de servidores back-end para SAP, Microsoft SharePoint, la informática en clústeres y otras aplicaciones empresariales.

Modelo	CPU virtual	Memoria (GiB)	Almacenamiento en SSD (GB)	Ancho de banda de EBS dedicado (Mbps)
m4.large	2	8	Solo EBS	450
m4.xlarge	4	16	Solo EBS	750
m4.2xlarge	8	32	Solo EBS	1 000
m4.4xlarge	16	64	Solo EBS	2 000
m4.10xlarge	40	160	Solo EBS	4 000
m4.16xlarge	64	256	Solo EBS	10 000

OPTIMIZADAS PARA INFORMÁTICA

- C5
 - Procesadores Intel Xeon Platinum de 3.0 GHz con conjuntos de instrucciones de extensiones vectoriales avanzadas de Intel 512 (AVX-512)
 - Ejecuta cada núcleo a hasta 3.5 GHz al usar Intel Turbo Boost Technology
 - Nuevo tamaño de instancia más grande, c5.18xlarge, que ofrece 72 vCPU y 144 GiB de memoria
 - Hasta 25 Gbps de ancho de banda de red al usar Elastic Network Adapter (ENA) basado en la red mejorada
 - EBS optimizado de forma predeterminada
 - Necesita una AMI HVM que incluye controladores para ENA y NVMe

Casos de uso: Servidores web de alto desempeño, modelado científico, procesamiento de lote, análisis distribuido, informática de alto desempeño (HPC), inferencia de aprendizaje profundo/automático, entrega de anuncios, juego para múltiples jugadores altamente escalables y codificación de video.

Modelo	CPU virtual	Memoria (GiB)	Almacenamiento	Ancho de banda de EBS dedicado (Mbps)
c5.large	2	4	Solo EBS	Hasta 2 250
c5.xlarge	4	8	Solo EBS	Hasta 2 250
c5.2xlarge	8	16	Solo EBS	Hasta 2 250
c5.4xlarge	16	32	Solo EBS	2 250
c5.9xlarge	36	72	Solo EBS	4 500
c5.18xlarge	72	144	Solo EBS	9 000

- C4

- Procesadores Intel Xeon E5-2666 v3 (H – ell) de alta frecuencia optimizados específicamente para EC2.
- Optimización para EBS predeterminada para mayor rendimiento de almacenamiento sin costo adicional
- Rendimiento de red superior con redes mejoradas compatibles con Intel 82599 VF
- Requiere Amazon VPC, Amazon EBS y AMI HVM de 64 bits

Casos de uso: Las flotas front-end de alto desempeño, los servidores web, el procesamiento por lotes, los análisis distribuidos, las aplicaciones de ingeniería y científicas de alto desempeño, la entrega de publicidad, los juegos MMO y la codificación de vídeo.

Modelo	CPU virtual	Memoria (GiB)	Almacenamiento	Ancho de banda de EBS dedicado (Mbps)
c4.large	2	3,75	Solo EBS	500
c4.xlarge	4	7.5	Solo EBS	750
c4.2xlarge	8	15	Solo EBS	1 000
c4.4xlarge	16	30	Solo EBS	2 000
c4.8xlarge	36	60	Solo EBS	4 000

OPTIMIZADAS PARA MEMORIA

- X1e
 - Procesadores Intel Xeon E7-8880 v3 (H – ell) de alta frecuencia
 - Uno de los precios más bajos por GiB de RAM
 - Hasta 3 904 GiB de memoria de instancia basada en DRAM
 - Almacenamiento SSD y optimizadas para EBS de manera predeterminada sin costo adicional
 - Capacidad para controlar configuraciones de estados C y P de procesador en las instancias x1e.32xlarge, x1e.16xlarge y x1e.8xlarge

Casos de uso: Bases de datos de alto desempeño, bases de datos en memoria (por ejemplo, SAP HANA) y aplicaciones de uso intensivo de memoria. La instancia x1e.32xlarge está certificada por SAP para la ejecución de Business Suite S/4HANA de próxima generación, Business Suite en HANA (SoH), Business Warehouse en HANA (BW) y soluciones de data mart en HANA en la nube de AWS.

Modelo	CPU virtual	Memoria (GiB)	Almacenamiento en SSD (GB)	Ancho de banda de EBS dedicada (Mbps)
x1e.32xlarge	128	3 904	2 x 1 920	14 000
x1e.16xlarge	64	1 952	1 x 1 920	7 000
x1e.8xlarge	32	976	1 x 960	3 500
x1e.4xlarge	16	488	1 x 480	1 750
x1e.2xlarge	8	244	1 x 240	1 000
x1e.xlarge	4	122	1 x 120	500

- X1
 - Procesadores Intel Xeon E7-8880 v3 (H – ell) de alta frecuencia
 - Uno de los precios más bajos por GiB de RAM
 - Hasta 1 952 GiB de memoria de instancia basada en DRAM
 - Almacenamiento SSD y optimizadas para EBS de manera predeterminada sin costo adicional

- Posibilidad de controlar la configuración del estado C y P del procesador

Casos de uso: Bases de datos en memoria (p. ej.: SAP HANA), motores de procesamiento de big data (p. ej.: Apache Spark o Presto), informática de alto rendimiento (HPC). Certificado por SAP para ejecutar Business Warehouse en HANA (BW), soluciones de centros de datos en HANA, Business Suite en HANA (SoH), Business Suite S/4HANA.

Modelo	CPU virtual	Memoria (GiB)	Almacenamiento en SSD (GB)	Ancho de banda de EBS dedicado (Mbps)
x1.32xlarge	128	1 952	2 x 1 920	14 000
x1.16xlarge	64	976	1 x 1 920	7 000

- R4
 - Procesadores Intel Xeon E5-2686 v4 (Broadwell) de alta frecuencia
 - Memoria DDR4
 - Soporte para redes mejoradas

Casos de uso: Bases de datos de alto desempeño, extracción y el análisis de datos, bases de datos en memoria, almacenamiento en memoria caché a escala de red distribuida, aplicaciones que ejecutan el procesamiento en tiempo real de big data sin estructurar, clústeres Hadoop/Spark y otras aplicaciones empresariales.

Modelo	CPU virtual	Memoria (GiB)	Desempeño de red	Almacenamiento en SSD (GB)
r4.large	2	15,25	Hasta 10 Gigabits	Solo EBS
r4.xlarge	4	30,5	Hasta 10 Gigabits	Solo EBS
r4.2xlarge	8	61	Hasta 10 Gigabits	Solo EBS
r4.4xlarge	16	122	Hasta 10 Gigabits	Solo EBS
r4.8xlarge	32	244	10 gigabits	Solo EBS
r4.16xlarge	64	488	25 gigabits	Solo EBS

INFORMÁTICA ACELERADA

- P3
 - Hasta 8 GPU V100 de NVIDIA Tesla, cada una sincronizada con 5 120 núcleos CUDA y 640 núcleos Tensor
 - Procesadores Intel Xeon E5-2686 v4 (Broadwell) de alta frecuencia
 - Compatible con NVLink para la comunicación de GPU de pares
 - Proporcionan redes mejoradas con el Elastic Network Adapter de hasta 25 Gbps de ancho de banda de red adicional en un grupo de ubicación

Casos de uso: Aprendizaje profundo/automático, informática de alto rendimiento, dinámica de fluidos computacional, finanzas computacionales, análisis sísmico, reconocimiento de voz, vehículos autónomos, descubrimiento de fármacos.

Modelo	GPU	CPU virtual	Memoria (GiB)	Memoria de GPU (Gib)	GPU de pares
p3.2xlarge	1	8	61	16	-
p3.8xlarge	4	32	244	64	NVLink
p3.16xlarge	8	64	488	128	NVLink

- P2
 - Procesadores Intel Xeon E5-2686 v4 (Broadwell) de alta frecuencia
 - GPU NVIDIA K80 de alto desempeño, cada unidad con 2 496 núcleos de procesamiento paralelo y 12 GiB de memoria GPU
 - Compatible con GPUDirect™ para las comunicaciones de GPU de pares
 - Proporciona redes mejoradas con Elastic Network Adapter (ENA) de hasta 25 Gbps de ancho de banda de red adicional en un grupo de ubicación
 - Optimizados para EBS de manera predeterminada sin costos adicionales

Casos de uso: Aprendizaje virtual, bases de datos de alto desempeño, dinámica fluida computacional, finanzas computacionales, análisis sísmico, modelado molecular, genómica, renderizado y otras cargas de trabajo de informática GPU del lado del servidor.

Modelo	GPU	CPU virtual	Memoria (GiB)	Memoria GPU (GiB)
p2.xlarge	1	4	61	12
p2.8xlarge	8	32	488	96
p2.16xlarge	16	64	732	192

- G3
 - Procesadores Intel Xeon E5-2686 v4 (Broadwell) de alta frecuencia
 - GPU NVIDIA Tesla M60, cada unidad con 2 048 núcleos de procesamiento paralelo y 8 GiB de memoria de video
 - Permite utilizar las características de NVIDIA GRID Virtual Workstation, incluida la compatibilidad para 4 monitores con resoluciones de hasta 4 096 x 2 160. Cada GPU incluida en la instancia tiene licencia para un usuario conectado al mismo tiempo
 - Permite utilizar las capacidades de NVIDIA GRID Virtual Application para el software de virtualización de aplicaciones, como Citrix XenApp Essentials y VMware Horizon, con compatibilidad para hasta 25 usuarios al mismo tiempo por GPU
 - Cada GPU incorpora un codificador de video de hardware integrado en la placa, diseñado para admitir hasta 10 transmisiones H.265 (HEVC) de 1 080 p 30 y hasta 18 transmisiones H.264 de 1 080 p 30, lo que permite un registro y una codificación de fotogramas de baja latencia y experiencias de streaming interactivas de gran calidad
 - Redes mejoradas con el adaptador de red elástico (ENA) que ofrece hasta 25 Gbps de ancho de banda total en un grupo de ubicación.

Casos de uso: Visualizaciones 3D, estaciones de trabajo remotas con un uso intensivo de gráficos, renderización 3D, transmisión de aplicaciones, codificación de vídeo y otros trabajos de gráficos en el lado del servidor.

Modelo	GPU	CPU virtual	Memoria (GiB)	Memoria GPU (GiB)
g3.4xlarge	1	16	122	8
g3.8xlarge	2	32	244	16
g3.16xlarge	4	64	488	32

- F1
 - Procesadores Intel Xeon E5-2686 v4 (Broadwell) de alta frecuencia
 - Almacenamiento SSD NVMe
 - Soporte para redes mejoradas
 - Características de la FPGA:
 - Xilinx Virtex UltraScale+ VU9P FPGAs
 - 64 GiB de memoria protegida por ECC en 4x DDR4
 - Interfaz dedicada PCI-Express x16
 - Aproximadamente 2,5 millones de elementos lógicos
 - Aproximadamente 6 800 motores de procesamiento de señales digitales (DSP)
 - AMI para desarrolladores de la FPGA

Casos de uso: Investigaciones genómicas, análisis financiero, procesamiento de vídeo en tiempo real, búsqueda y análisis de big data, y seguridad.

Modelo	FPGA	CPU virtual	Memoria (GiB)	Almacenamiento en SSD (GB)
f1.2xlarge	1	8	122	470
f1.16xlarge	8	64	976	4 x 940

OPTIMIZADAS PARA ALMACENAMIENTO

- H1
 - Con tecnología de procesadores Intel® Xeon® E5 2686 v4 con 2,3 GHz (nombre de código Broadwell)
 - Hasta 16 TB de almacenamiento en HDD
 - Alto rendimiento de disco
 - Redes mejoradas con ENA de hasta 25 Gbps

Casos de uso: Cargas de trabajo basadas en MapReduce, sistemas de archivos distribuidos como HDFS y MapR-FS, sistemas de archivos de red, aplicaciones de procesamiento de datos o logs como Apache Kafka, y clústeres de cargas de trabajo de big data.

Modelo	CPU virtual	Memoria (GiB)	Desempeño de redes	Almacenamiento (GB)
h1.2xlarge	8	32	Hasta 10 Gigabits	1 x 2 000 HD
h1.4xlarge	16	64	Hasta 10 Gigabits	2 x 2 000 HD
h1.8xlarge	32	128	10 Gigabits	4 x 2 000 HD
h1.16xlarge	64	256	25 Gigabits	8 x 2 000 HD

- I3
 - Procesadores de alta frecuencia Intel Xeon E5-2686 v4 (Broadwell) con una frecuencia base de 2,3 GHz
 - Hasta 25 Gbps de ancho de banda de red mediante el uso de redes mejoradas basadas en adaptador de red elástico (ENA)
 - Alto desempeño de E/S aleatoria y alto rendimiento de lectura secuencial

Casos de uso: Bases de datos NoSQL (p. ej.: Cassandra, MongoDB, Redis), bases de datos en memoria (p. ej.: Aerospike, escala de bases de datos transaccionales, almacenamiento de datos, Elasticsearch, cargas de trabajo de análisis.

Modelo	CPU virtual	Memoria (GiB)	Desempeño de redes	Almacenamiento (TB)
i3.large	2	15,25	Hasta 10 Gigabits	1 x 0,475 SSD NVMe
i3.xlarge	4	30,5	Hasta 10 Gigabits	1 x 0,95 SSD NVMe
i3.2xlarge	8	61	Hasta 10 Gigabits	1 x 1,9 SSD NVMe
i3.4xlarge	16	122	Hasta 10 Gigabits	2 x 1,9 SSD NVMe
i3.8xlarge	32	244	10 gigabits	4 x 1,9 SSD NVMe
i3.16xlarge	64	488	25 gigabits	8 x 1,9 SSD NVMe

- D2
 - Procesadores Intel Xeon E5-2676v3 (Hwsell) de alta frecuencia
 - Almacenamiento HDD
 - Alto desempeño consistente en el momento del lanzamiento
 - Alto rendimiento de disco
 - Soporte para redes mejoradas

Casos de uso: Almacenamiento de datos Massively Parallel Processing (MPP), informática distribuida de MapReduce y Hadoop, sistemas de archivos distribuidos, sistemas de archivos de red, aplicaciones de log o procesamiento de datos.

Modelo	CPU virtual	Memoria (GiB)	Almacenamiento (GB)
d2.xlarge	4	30,5	3 x 2 000 HDD
d2.2xlarge	8	61	6 x 2 000 HDD
d2.4xlarge	16	122	12 x 2 000 HDD
d2.8xlarge	36	244	24 x 2 000 HDD

4.CASOS DE ESTUDIOS

Expedia

Expedia es una agencia de viajes en Internet y tiene sus oficinas centrales en Estados Unidos con delegaciones en 31 países. Reserva billetes de avión, hotel, alquiler de vehículos, cruceros, paquetes vacacionales y varios parques de atracciones a través de la web y teléfono.

Expedia está obligada a una continua innovación, tecnología, y mejoras en su plataforma para crear una gran experiencia para sus usuarios. La organización Expedia Worldwide Engineering (EWE) acoge todas las páginas web bajo la marca de Expedia. Expedia empezó a usar Amazon Web Services (AWS) en 2010 para lanzar Expedia Suggest Service (ESS), un servicio de sugerencias typeahead que ayuda al usuario a rellenar el viaje, buscar y localizar información correctamente. De acuerdo a las métricas de la compañía, una página de error es la principal razón de abandono del sitio. Expedia quería que los usuarios globales encontrarán lo que buscaban de una forma rápida y sin errores. Por entonces, Expedia ofrecía todos sus servicios desde centros de datos en Chandler, AZ. El equipo de ingenieros se dio cuenta que tenían que lanzar ESS en localizaciones físicamente cerca a los usuarios para habilitar un servicio rápido y sensible con una latencia mínima de conexión.

Expedia había considerado virtualizaciones on-premises y otros proveedores de servicios en la nube como solución, pero finalmente eligieron AWS porque era la única solución con la infraestructura global capaz de soportar los usuarios de Asia Pacífica. “Desde una perspectiva arquitectónica, infraestructura, automatización, y proximidad a los usuarios fueron los factores claves”, explica Murari Gopalan, Director de Tecnología. “No había otra forma de solucionar el problema que no fuese usando AWS”.

“Usando AWS, fuimos capaces de construir y lanzar el servicio ESS en 3 meses”, dijo Magesh Chandramouli, Arquitecto Principal. ESS usa algoritmos basados en la localización del usuario y datos de compra y reserva de usuarios previos para enseñar sugerencias cuando un usuario empieza a escribir. Por ejemplo, si un usuario de Seattle escribe mar cuando reserva un vuelo, el servicio enseñará Seattle, SeaTac, y otros destinos relevantes.

Expedia lanzó las instancias de ESS inicialmente en la región de Asia Pacífica (Singapur) y justo después replicó el servicio en el Oeste de EEUU (California del Norte) y regiones europeas (Irlanda). Los ingenieros de Expedia inicialmente usaban Apache Lucene y otras herramientas de open source para construir el servicio, pero con el tiempo desarrollaron herramientas potentes internas para guardar índices y consultas.

Con el desarrollo de ESS en AWS, Expedia fue capaz de mejorar su servicio a usuarios de la región de Asia Pacífica como en Europa. “La latencia era nuestro principal problema”, dice Chandramouli. “Usando AWS, disminuimos la latencia media de 700 milisegundos a menos de 50 milisegundos”. La figura siguiente demuestra el servicio de sugerencia typeahead ESS usando AWS.

AIRBNB

Airbnb es un mercado de la comunidad que permite a los propietarios y viajeros conectar entre sí para poder alquilar lugares únicos de vacaciones alrededor del mundo. Las actividades de la comunidad de usuarios de Airbnb se fijan en el uso de la página web de la compañía y de sus aplicaciones móviles. La base de Airbnb en San Francisco empezó sus operaciones en 2008 y, en la actualidad, tiene miles de empleados a lo largo del mundo soportando alquileres en cerca de 25000 ciudades en 192 países.

Un año después de que Airbnb se lanzara, la compañía decidió mudar todas sus funciones de computación en la nube a Amazon Web Services (AWS) debido a desafíos de administración del servicio experimentados con su proveedor original. Nathan Blecharczyk, Co-Fundador & CTO de Airbnb dice, “Inicialmente, el encanto de AWS era la facilidad en el manejo y la customización de la pila. Fue genial ser capaz de aumentar el número de servidores sin tener que contactar con nadie y ningún mínimo compromiso de uso. Conforme nuestra compañía continuaba creciendo, también lo hacía nuestra confianza en la nube de AWS y ahora, hemos adoptado casi todas las características que AWS provee. AWS es la respuesta más fácil a cualquier empresa de Internet que desea llegar al siguiente nivel”.

Airbnb ha crecido significativamente a lo largo de los últimos 3 años. Para soportar la demanda, la compañía usa 200 instancias de Amazon Elastic Compute Cloud (Amazon EC2) para sus aplicaciones, memoria cache, y búsqueda en servidores. Dentro de Amazon EC2, Airbnb está usando Elastic Load Balancing, el cual distribuye automáticamente el tráfico llegando entre las múltiples instancias de Amazon EC2. Para el fácil procesamiento y análisis de 50Gigabytes de datos diarios, Airbnb usa Amazon Elastic MapReduce (Amazon EMR). Airbnb también está usando Amazon Simple Storage Service (Amazon S3) para guardar copias de seguridad y archivos estáticos, incluyendo 10 terabytes de fotos de usuarios. Para monitorizar los recursos de sus servidores, Airbnb usa Amazon CloudWatch, que permite a la compañía supervisar de forma fácil todos sus activos de Amazon EC2 a través de AWS Management Console, comandos de línea de comando, o APIs de servicios web.

Además, Airbnb movió sus principales bases de datos de MYSQL a Amazon Relational Database Service (Amazon RDS). Airbnb eligió Amazon RDS porque simplifica mucho del tiempo consumido en tareas administrativas típicamente asociadas a la base de datos. Amazon RDS permite a los procedimientos difíciles, como la replica o el escalado, ser realizados con una llamada a una API básica o a través de AWS Management Console. Actualmente, Airbnb usa el despliegue de Multi-Availability Zone (Multi-AZ) para automatizar su replica de base de datos y aumentar la durabilidad de los datos.

Airbnb fue capaz de completar su migración total de sus bases de datos a Amazon RDS en tan solo 15 minutos de inactividad. Esta transición tan rápida fue muy importante para el creciente Airbnb porque no querían que su comunidad de usuarios se quedaran fuera de su mercado por un periodo largo de tiempo. Tobi Knaup, un ingeniero de Airbnb dice, “Debido a AWS, siempre ha habido una fácil respuesta (en términos de tiempo necesario y coste) para escalar nuestro sitio”.

NETFLIX

El proveedor de contenido en línea Netflix es capaz de soportar su servicio mundial asociándose con Amazon Web Services (AWS) para los servicios de entrega de contenido. AWS permite a Netflix desplegar rápidamente miles de servidores y terabytes de almacenamiento en cuestión de minutos. Los usuarios pueden ver programas y películas de Netflix desde cualquier parte del mundo, incluso en la web, en tabletas o en dispositivos móviles como el iPhone.

Netflix almacena las series y películas en 50 diferentes formas dependiendo de la calidad del video y del audio. Dicha gran cantidad de videos, por cada miles de series y películas que Netflix nos ofrece, ocupan una gran cantidad de espacio de almacenamiento.

Netflix hace uso de la infraestructura de Amazon EC2 y almacenando las copias de las series y películas en Amazon S3.

5.Referencias

<https://aws.amazon.com/es/ec2/>

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

<https://aws.amazon.com/es/ec2/videos/>

<https://aws.amazon.com/es/solutions/case-studies/netflix/>

<https://aws.amazon.com/es/solutions/case-studies/airbnb/>

<https://aws.amazon.com/es/solutions/case-studies/expedia/>