Documento técnico de AWS

Alojamiento de aplicaciones web en la nube de AWS

Alojamiento de aplicaciones web en la nube de AWS: Documento técnico de AWS

Table of Contents

Alojamiento de aplicaciones web en la nube de AWS

Fecha de publicación: 20 de agosto de 2021 (Revisiones del documento)

Resumen

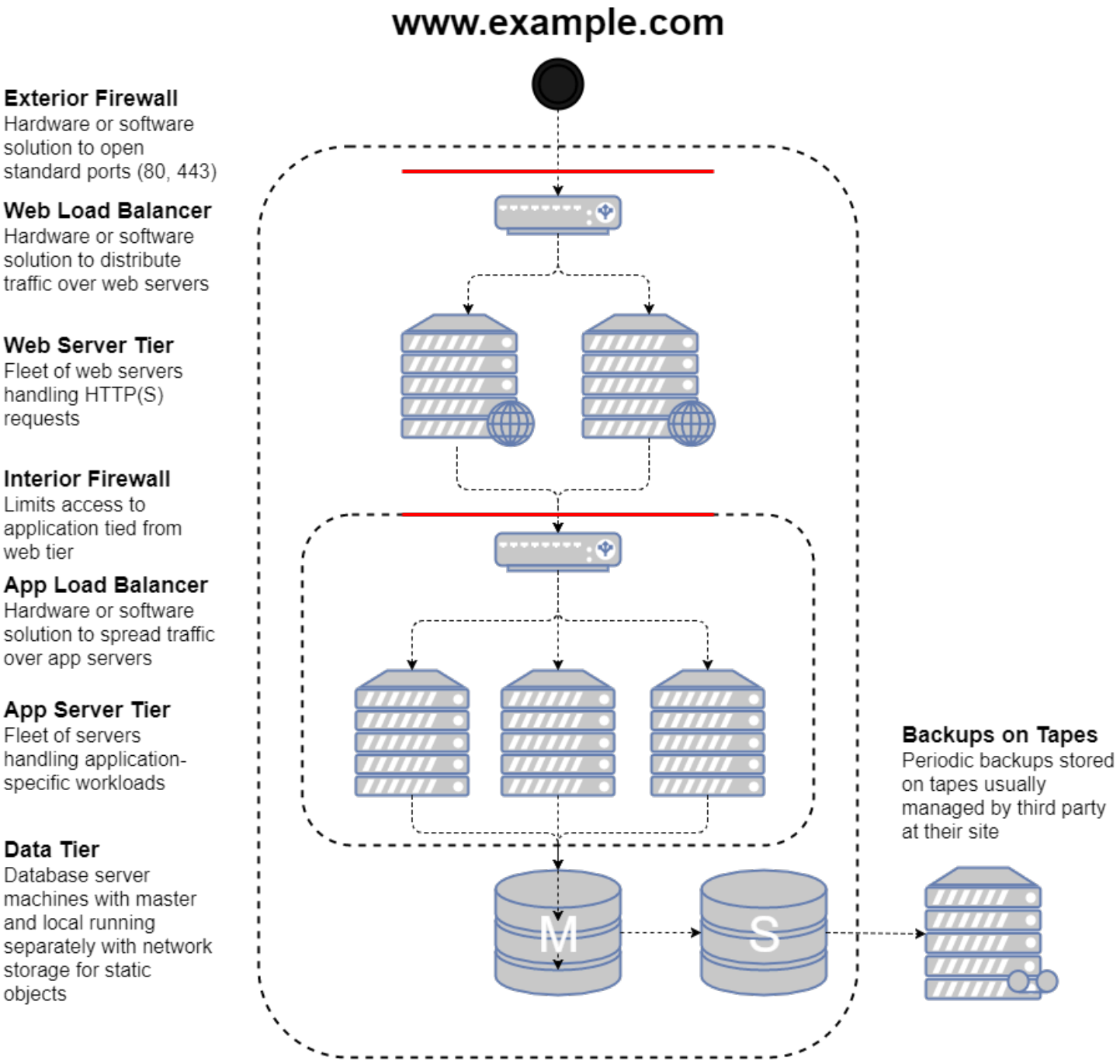
Las arquitecturas web locales tradicionales requieren soluciones complejas y una previsión precisa de la capacidad reservada para garantizar la fiabilidad. Los períodos de picos de tráfico denso y los cambios bruscos en los patrones de tráfico dan como resultado unos índices de uso muy bajos de un hardware caro. Esto genera altos costes operativos al tener que mantener el hardware inactivo y un uso ineficiente del capital del hardware infrautilizado.

Amazon Web Services (AWS) ofrece una infraestructura fiable, escalable, segura y de alto rendimiento para las aplicaciones web más exigentes. Esta infraestructura permite equiparar los costes de TI con los patrones de tráfico de los clientes casi en tiempo real.

Este documento técnico está dirigido a administradores de TI y arquitectos de sistemas que desean aprender a ejecutar arquitecturas web tradicionales en la nube para lograr elasticidad, escalabilidad y fiabilidad.

Información general sobre el alojamiento web tradicional

El alojamiento web escalable es un tema problemático bien conocido. La siguiente imagen muestra una arquitectura de alojamiento web tradicional que implementa un modelo de aplicación web común de tres niveles. En este modelo, la arquitectura se separa en capas de presentación, aplicación y persistencia. La escalabilidad se obtiene añadiendo hosts en estas capas. La arquitectura también tiene integradas características de rendimiento, conmutación por error y disponibilidad. Es fácil realizar la portabilidad de la arquitectura de alojamiento web tradicional a la nube de AWS con solo unas pocas modificaciones.



Una arquitectura de alojamiento web tradicional

En las siguientes secciones, se explica por qué y cómo se debe implementar una arquitectura de este tipo en la nube de AWS.

Alojamiento de aplicaciones web en la nube con AWS

La primera pregunta que debe hacerse se refiere al valor que tiene trasladar una solución de alojamiento de aplicaciones web clásica a la nube de AWS. Si decide que la nube es la adecuada para usted, necesitará una arquitectura apropiada. Esta sección le ayuda a evaluar una solución en la nube de AWS. En ella, se compara la implementación de su aplicación web en la nube con una implementación local, se describe una arquitectura en la nube de AWS para alojar su aplicación y se analizan los componentes clave de la solución de la arquitectura de la nube de AWS.

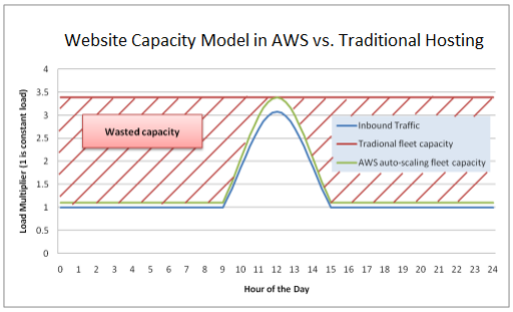
Cómo AWS puede resolver problemas comunes de alojamiento de aplicaciones web

Si tiene la responsabilidad de ejecutar una aplicación web, podría enfrentarse a una gran variedad de problemas de la infraestructura y la arquitectura para los que AWS puede ofrecer soluciones eficaces y rentables. Estos son algunos de los beneficios de usar AWS en lugar de un modelo de alojamiento tradicional.

Una alternativa rentable a las flotas sobredimensionadas que se necesitan para gestionar los picos de tráfico

En el modelo de alojamiento tradicional, debe aprovisionar servidores para gestionar la capacidad en los picos de tráfico. Los ciclos no utilizados se desperdician cuando no hay picos. Las aplicaciones web alojadas en AWS pueden utilizar el aprovisionamiento bajo demanda de servidores adicionales para que pueda ajustar constantemente la capacidad y los costes a los patrones de tráfico reales.

Por ejemplo, el siguiente gráfico muestra una aplicación web que tiene un pico de uso entre las 9:00 y las 15:00 y menos uso durante el resto del día. Con un enfoque de escalado automático basado en las tendencias de tráfico reales, que aprovisione recursos solo cuando sea necesario, se desperdiciaría menos capacidad y se reduciría más de un 50 % del coste.



Ejemplo de capacidad desperdiciada en un modelo de alojamiento clásico

Una solución escalable para gestionar picos de tráfico inesperados

Una consecuencia más grave del aprovisionamiento lento asociado con el modelo de alojamiento tradicional es la incapacidad de responder a tiempo a los picos de tráfico inesperados. Hay muchas historias sobre aplicaciones web que, después de que el sitio web se mencionara en los medios de comunicación más populares, dejaron de estar disponibles debido a un pico inesperado del tráfico. En la nube de AWS, la capacidad bajo demanda que ayuda a las aplicaciones web a escalarse para adaptarse a los picos de tráfico habituales también puede gestionar una carga inesperada. Se pueden lanzar nuevos hosts para que estén disponibles en cuestión de minutos, y se pueden desconectar con la misma rapidez cuando el tráfico vuelve a la normalidad.

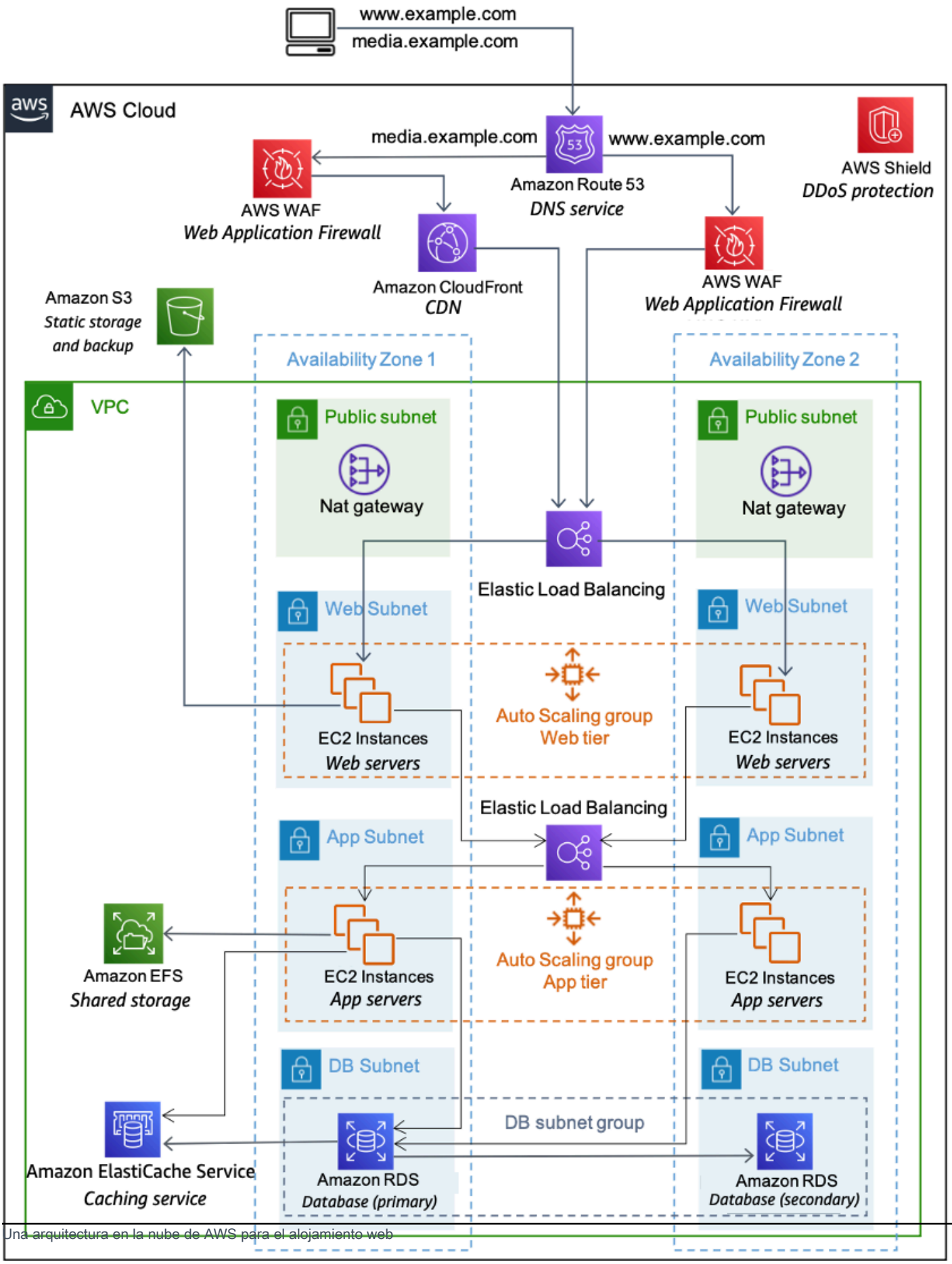
Una solución bajo demanda para entornos de prueba, carga, beta y reproducción

Los costes del hardware para crear y mantener un entorno de alojamiento tradicional para una aplicación web de producción no terminan con la flota de producción. Muchas veces, es necesario crear flotas de preproducción, beta y pruebas para garantizar la calidad de la aplicación web en cada etapa del ciclo de vida del desarrollo. Aunque se pueden realizar varias optimizaciones para garantizar el mayor uso posible de este hardware de prueba, estas flotas paralelas no siempre se usan de la manera óptima y una gran cantidad de hardware muy caro se queda sin usar durante largos períodos de tiempo.

En la nube de AWS, puede aprovisionar flotas de pruebas cuando las necesite. Esto no solo elimina la necesidad de aprovisionar recursos días o meses antes de su uso real, sino que también le ofrece la flexibilidad de desmontar los componentes de la infraestructura cuando no los necesita. Además, puede simular el tráfico de usuarios en la nube de AWS durante las pruebas de carga. También puede utilizar estas flotas paralelas como entorno transitorio para una nueva versión de producción. Esto permite realizar un cambio rápido desde el entorno de producción actual a una nueva versión de la aplicación con pocas interrupciones del servicio o incluso ninguna.

Una arquitectura en la nube de AWS para el alojamiento web

La siguiente figura es otra representación de esa arquitectura clásica de aplicaciones web y cómo puede aprovechar la infraestructura de computación en la nube de AWS.


www.example.com
media.example.com
AWS Cloud
media.example.com
www.example.com
AWS Shield
DDoS protection
AWS WAF
Web Application Firewall
Amazon Route 53
DNS service
Amazon CloudFront
CDN
AWS WAF
Web Application Firewall
Amazon S3
Static storage and backup
Availability Zone 1
Availability Zone 2
VPC
Public subnet
Nat gateway
Public subnet
Nat gateway
Elastic Load Balancing
Web Subnet
Web Subnet
EC2 Instances
Web servers
Auto Scaling group
Web tier
EC2 Instances
Web servers
Elastic Load Balancing
App Subnet
App Subnet
Amazon EFS
Shared storage
EC2 Instances
App servers
Auto Scaling group
App tier
EC2 Instances
App servers
DB Subnet
DB Subnet
DB subnet group
Amazon ElastiCache Service
Caching service
Amazon RDS
Database (primary)
Amazon RDS
Database (secondary)

Un ejemplo de arquitectura de alojamiento web en AWS

1. Servicios de DNS con Amazon Route 53: proporciona servicios de DNS para simplificar la administración de dominios.
2. Almacenamiento en caché de borde con Amazon CloudFront: el borde almacena en caché contenido de gran volumen para reducir la latencia para los clientes.
3. Seguridad de borde para Amazon CloudFront con AWS WAF: filtra el tráfico malicioso, incluido el scripting entre sitios (XSS) y la inyección de SQL a través de reglas definidas por el cliente.
4. Equilibrio de carga con Elastic Load Balancing (ELB): le permite distribuir la carga entre varias zonas de disponibilidad y grupos de AWS Auto Scaling para la redundancia y el desacoplamiento de los servicios.
5. Protección DDoS con AWS Shield: protege su infraestructura contra los ataques de DDoS más comunes en la capa de red y transporte de forma automática.
6. Firewalls con grupos de seguridad: traslada la seguridad a la instancia para proporcionar un firewall de nivel de host con estado para los servidores web y de aplicaciones.
7. Almacenamiento en caché con Amazon ElastiCache: proporciona servicios de almacenamiento en caché con Redis o Memcached para eliminar la carga de la aplicación y la base de datos, y reducir la latencia para solicitudes frecuentes.
8. Base de datos administrada con Amazon Relational Database Service (Amazon RDS): crea una arquitectura de base de datos Multi-AZ de alta disponibilidad con seis motores de base de datos posibles.
9. Almacenamiento estático y copias de seguridad con Amazon Simple Storage Service (Amazon S3): permite el almacenamiento sencillo de objetos basado en HTTP para copias de seguridad y recursos estáticos, como imágenes y vídeo.

Componentes clave de una arquitectura de alojamiento web de AWS

En las siguientes secciones, se describen algunos de los componentes clave de una arquitectura de alojamiento web implementada en la nube de AWS y se explica en qué se diferencian de una arquitectura de alojamiento web tradicional.

Administración de red

En la nube de AWS, la capacidad de separar su red de la de otros clientes permite que la arquitectura sea más segura y escalable. Aunque los grupos de seguridad proporcionan seguridad a nivel de host (consulte la sección Seguridad del host), Amazon Virtual Private Cloud (Amazon VPC) le permite lanzar recursos en una red virtual y aislada lógicamente que haya definido.

Amazon VPC es un servicio que le proporciona un control total sobre los detalles de su configuración de red en AWS. Entre los ejemplos de este control, se incluyen la creación de subredes públicas para servidores web y subredes privadas sin acceso a Internet para sus bases de datos. Además, Amazon VPC le permite crear arquitecturas híbridas mediante el uso de redes privadas virtuales (VPN) de hardware y utilizar la nube de AWS como una extensión de su propio centro de datos.

Amazon VPC también incluye compatibilidad con IPv6, además de la compatibilidad con IPv4 tradicional para su red.

Entrega de contenido

Cuando el tráfico web está disperso geográficamente, no siempre es factible, y desde luego no es rentable, replicar toda la infraestructura a nivel global. Una red de entrega de contenido (CDN) le ofrece la posibilidad de utilizar su red global de ubicaciones de borde para entregar a sus clientes una copia en caché de contenidos web como vídeos, páginas web, imágenes, etc. Para reducir el tiempo de respuesta, la CDN utiliza la ubicación de borde más cercana al cliente o a la ubicación de la solicitud de origen. El rendimiento se incrementa drásticamente dado que los recursos web se entregan desde la caché. Para los datos dinámicos, muchas CDN pueden configurarse para recuperar los datos de los servidores de origen.

Puede usar CloudFront para entregar su sitio web, lo que incluye contenido dinámico, estático y de streaming, mediante una red global de ubicaciones de borde. CloudFront dirige automáticamente las solicitudes de su contenido a la ubicación de borde más cercana para que el contenido se entregue con el mejor rendimiento posible. CloudFront está optimizado para funcionar con otros servicios de AWS, como Amazon S3 y Amazon Elastic Compute Cloud (Amazon EC2). CloudFront también funciona a la perfección con cualquier otro servidor de origen que no sea de AWS, que almacene las versiones originales y definitivas de sus archivos.

Al igual que en otros servicios de AWS, no hay contratos ni compromisos mensuales sobre el uso de CloudFront; solo paga por la cantidad de contenido que distribuya a través del servicio.

Además, cualquier solución existente para el almacenamiento en caché de borde en la infraestructura de aplicaciones web debería funcionar bien en la nube de AWS.

Administración del DNS público

Para trasladar una aplicación web a la nube de AWS, es necesario realizar algunos cambios en el sistema de nombres de dominio (DNS). Para ayudarlo a administrar el enrutamiento de DNS, AWS incluye Amazon Route 53, un servicio web de DNS en la nube escalable y de alta disponibilidad. Route 53 está diseñado para ofrecer a los desarrolladores y las empresas un método enormemente fiable y rentable para dirigir a los usuarios finales a las aplicaciones de Internet convirtiendo nombres como “www.ejemplo.com” en direcciones IP numéricas, como 192.0.2.1, que los equipos utilizan para conectarse entre sí. Route 53 también es completamente compatible con IPv6.

Seguridad del host

Además del filtrado del tráfico de red entrante en el borde, AWS también recomienda que las aplicaciones web filtren el tráfico de red a nivel de host. Amazon EC2 incluye una característica denominada grupos de seguridad. Un grupo de seguridad es algo parecido a un firewall de red entrante, que le permite especificar los protocolos, los puertos y los intervalos de IP de origen que pueden acceder a las instancias de EC2.

Puede asignar uno o más grupos de seguridad a cada instancia de EC2. Cada grupo de seguridad permite el tráfico adecuado en cada instancia. Los grupos de seguridad se pueden configurar de modo que solo las subredes, las direcciones IP y los recursos específicos tengan acceso a una instancia de EC2. Como alternativa, pueden hacer referencia a otros grupos de seguridad para limitar el acceso a las instancias de EC2 que se encuentran en grupos específicos.

En la arquitectura de alojamiento web de AWS de la figura 3, el grupo de seguridad para el clúster de servidores web podría permitir el acceso solo desde el equilibrador de carga de la capa web y solo a través de TCP en los puertos 80 y 443 (HTTP y HTTPS). Por otro lado, el grupo de seguridad del servidor de aplicaciones podría permitir el acceso solo desde el equilibrador de carga de la capa de aplicación. En este modelo, los ingenieros de soporte también necesitarían acceder a las instancias de EC2, lo que se puede realizar con AWS Systems Manager Session Manager. Para obtener un análisis más profundo sobre la seguridad, consulte Seguridad en la nube de AWS, que contiene boletines de seguridad, información sobre certificaciones y documentos técnicos de seguridad en los que se explican las capacidades de seguridad de AWS.

Equilibrio de carga en clústeres

Los equilibradores de carga de hardware son un dispositivo de red que suele utilizarse en las arquitecturas de aplicaciones web tradicionales. AWS proporciona esta capacidad a través del servicio Elastic Load Balancing (ELB). ELB distribuye automáticamente el tráfico de aplicaciones

entrante entre varios destinos, como instancias de EC2, contenedores, direcciones IP, funciones AWS Lambda y dispositivos virtuales. Puede controlar la carga variable del tráfico de su aplicación en una única zona o en varias zonas de disponibilidad. Elastic Load Balancing ofrece cuatro tipos de equilibradores de carga que cuentan con el nivel necesario de alta disponibilidad, escalado automático y seguridad para que sus aplicaciones sean tolerantes a errores.

Búsqueda de otros hosts y servicios

En la arquitectura de alojamiento web tradicional, la mayoría de los hosts tienen direcciones IP estáticas. En la nube de AWS, la mayoría de los hosts tienen direcciones IP dinámicas. Aunque todas las instancias de EC2 pueden tener entradas de DNS tanto públicas como privadas y se pueden direccionar a través de Internet, las entradas de DNS y las direcciones IP se asignan de forma dinámica cuando se lanza la instancia. No se pueden asignar manualmente. Las direcciones IP estáticas (direcciones IP elásticas en la terminología de AWS) se pueden asignar a las instancias en ejecución después de su lanzamiento. Debe usar direcciones IP elásticas para instancias y servicios que requieran puntos de conexión constantes, como bases de datos principales, servidores de archivos centrales y equilibradores de carga alojados en EC2.

Almacenamiento en caché dentro de la aplicación web

Las cachés de aplicaciones en memoria pueden reducir la carga de los servicios y mejorar el rendimiento y la escalabilidad en el nivel de la base de datos al almacenar en caché la información de uso frecuente Amazon ElastiCache es un servicio web que facilita la implementación, el funcionamiento y el escalado de una caché en memoria en la nube. Puede configurar la caché en memoria que cree para que se escale automáticamente con la carga y que reemplace automáticamente los nodos que fallen. ElastiCache es compatible con los protocolos Memcached y Redis, lo que simplifica la migración desde su solución local actual.

Configuración, copia de seguridad y conmutación por error de bases de datos

Muchas aplicaciones web contienen algún tipo de persistencia, generalmente en forma de base de datos relacional o no relacional. AWS ofrece servicios de bases de datos relacionales y no relacionales. Como alternativa, puede implementar su propio software de base de datos en una instancia de EC2. En la siguiente tabla, se resumen estas opciones, que se analizan con más detalle en esta sección.

Tabla 1: Soluciones de bases de datos relacionales y no relacionales

	Soluciones de bases de datos relacionales	Soluciones NoSQL
Servicio de base de datos administrada	Amazon RDS for MySQL, Oracle, SQL Server, MariaDB, PostgreSQL, Amazon Aurora	Amazon DynamoDB , Amazon Keyspaces, Amazon Neptune, Amazon QLDB, Amazon Timestream
Autoadministrada	Alojamiento de un sistema de administración de bases de datos relacionales (DBMS) en una instancia de Amazon EC2	Alojamiento de una solución de base de datos no relacional en una instancia de EC2

Amazon RDS

Amazon Relational Database Service (Amazon RDS) le da acceso a las capacidades de un motor de base de datos tan conocido como MySQL, PostgreSQL, Oracle y Microsoft SQL Server. El código, las aplicaciones y las herramientas que ya utiliza se pueden usar con Amazon RDS. Amazon RDS aplica automáticamente las revisiones en el software de base de datos y realiza una copia de seguridad que se almacena durante un período de retención que define el usuario. También admite la recuperación a un momento dado. De este modo, disfruta de la flexibilidad que supone poder escalar los recursos de computación o la capacidad de almacenamiento asociada con su instancia de base de datos relacional mediante una única llamada a la API.

Las implementaciones Multi-AZ de Amazon RDS aumentan la disponibilidad de la base de datos y la protegen contra interrupciones imprevistas. Las réplicas de lectura de Amazon RDS proporcionan réplicas de solo lectura de su base de datos, por lo que puede escalarlas horizontalmente más allá de la capacidad de implementación de una base de datos única para cargas de trabajo de bases de datos con operaciones de lectura intensiva. Al igual que en todos los servicios de AWS, no es necesario realizar inversiones iniciales y solo se pagan los recursos que se utilizan.

Alojamiento de un sistema de administración de bases de datos relacionales (RDBMS) en una instancia de Amazon EC2

Además de la oferta administrada de Amazon RDS, puede instalar la RDBMS que elija (por ejemplo, MySQL, Oracle, SQL Server o DB2) en una instancia de EC2 y administrarla usted mismo. Los clientes de AWS que alojan una base de datos en Amazon EC2 utilizan de forma eficaz una gran

variedad de modelos de replicación y primarios/en espera, incluida la duplicación para copias de solo lectura y el envío de registros para esclavos pasivos que siempre están preparados.

Al administrar su propio software de base de datos directamente en Amazon EC2, también debe tener en cuenta la disponibilidad de almacenamiento persistente y tolerante a errores. Para ello, recomendamos que las bases de datos que se ejecutan en Amazon EC2 utilicen volúmenes de Amazon Elastic Block Store (Amazon EBS), que son similares al almacenamiento conectado a la red.

Para las instancias de EC2 que ejecutan una base de datos, debe colocar todos los datos y registros de la base de datos en volúmenes de EBS. Seguirán estando disponibles aunque el host de la base de datos falle. Esta configuración sirve en un escenario de conmutación por error sencillo, en el que se puede lanzar una nueva instancia de EC2 si se produce un error en un host y los volúmenes de EBS existentes se pueden conectar a la nueva instancia. La base de datos puede continuar donde se quedó.

Los volúmenes de EBS proporcionan redundancia de forma automática dentro de la zona de disponibilidad. Si el rendimiento de un solo volumen de EBS no es suficiente para las necesidades de sus bases de datos, los volúmenes se pueden fragmentar para aumentar el rendimiento de las operaciones de entrada/salida por segundo (IOPS) de su base de datos.

Para cargas de trabajo exigentes, también puede usar IOPS aprovisionadas de EBS, donde especifica las IOPS que necesita. Si usa Amazon RDS, el servicio administra su propio almacenamiento para que pueda concentrarse en administrar sus datos.

Bases de datos no relacionales

Además de admitir bases de datos relacionales, AWS también ofrece varias bases de datos no relacionales administradas:

- Amazon DynamoDB es un servicio de base de datos NoSQL completamente administrado que ofrece un rendimiento rápido y predecible, así como una escalabilidad perfecta. Con la AWS Management Console o la API de DynamoDB, puede aumentar o reducir la capacidad sin ningún tiempo de inactividad ni una degradación del rendimiento. Dado que DynamoDB se encarga de las cargas administrativas que suponen operar y escalar bases de datos distribuidas en AWS, no tiene que preocuparse por el aprovisionamiento, la instalación y la configuración del hardware, la replicación, la aplicación de revisiones de software o el escalado de clústeres.
- Amazon DocumentDB (compatible con MongoDB) es un servicio de base de datos creado específicamente para la administración de datos JSON a escala, completamente administrado, que se ejecuta con AWS y que está preparado para empresas con alta durabilidad.

- Amazon Keyspaces (para Apache Cassandra) es un servicio de base de datos administrado, de alta disponibilidad y escalable compatible con Apache Cassandra. Con Amazon Keyspaces, puede ejecutar las cargas de trabajo de Cassandra en AWS con las mismas herramientas para desarrolladores y el mismo código de aplicación de Cassandra que utiliza en la actualidad.
- Amazon Neptune es un servicio de base de datos de grafos rápido, fiable y completamente administrado que le permite crear y ejecutar fácilmente aplicaciones que funcionen con conjuntos de datos fuertemente conectados. El núcleo de Amazon Neptune es un motor de bases de datos de grafos de alto rendimiento diseñado expresamente y optimizado para almacenar miles de millones de relaciones y consultar gráficos con una latencia de milisegundos.
- Amazon Quantum Ledger Database (Amazon QLDB) es una base de datos de libro mayor completamente administrada que ofrece un registro de transacciones transparente e inmutable, que se puede verificar de manera criptográfica y que es propiedad de una autoridad central de confianza. QLDB puede utilizarse para registrar cada uno de los cambios que se producen en los datos de las aplicaciones y mantener un historial completo y verificable.
- Amazon Timestream es un servicio de bases de datos de serie temporal rápido, escalable y sin servidor para aplicaciones operativas y de IoT que facilita el almacenamiento y el análisis de billones de eventos al día hasta 1000 veces más rápido y por tan solo una décima parte del coste que las bases de datos relacionales.

Además, puede utilizar Amazon EC2 para alojar otras tecnologías de bases de datos no relacionales con las que esté trabajando.

Almacenamiento y copia de seguridad de datos y recursos

Existen numerosas opciones dentro de la nube de AWS para almacenar, obtener acceso y realizar copias de seguridad de los datos y recursos de su aplicación web. Amazon S3 dispone de un almacén de objetos redundante y de alta disponibilidad. Amazon S3 es una excelente solución de almacenamiento para objetos algo estáticos o que cambian lentamente, como imágenes, vídeos y otros medios estáticos. Amazon S3 también admite almacenamiento en caché de borde y streaming de estos recursos mediante la interacción con CloudFront.

Para un almacenamiento similar a un sistema de archivos conectado, las instancias de EC2 pueden tener volúmenes de EBS conectados. Actúan como discos que se pueden montar para ejecutar instancias de EC2. Amazon EBS es ideal para los datos a los que se debe acceder como un almacenamiento en bloque y que requieren persistencia más allá de la vida útil de la instancia en ejecución, como particiones de bases de datos y registros de aplicaciones.

Además de tener una vida útil independiente de la instancia de EC2, puede realizar instantáneas de los volúmenes de EBS y almacenarlas en Amazon S3. Puesto que las instantáneas de EBS solo realizan copias de seguridad de los cambios que se han producido desde la instantánea anterior, las instantáneas más frecuentes pueden reducir los tiempos de instantánea. También puede utilizar una instantánea de EBS como referencia para replicar datos en varios volúmenes de EBS y conectar esos volúmenes a otras instancias en ejecución.

Los volúmenes de EBS pueden tener hasta 16 TB y se pueden fragmentar en varios volúmenes de EBS para conseguir volúmenes aún más grandes o para aumentar el rendimiento de entrada/salida (E/S). Para maximizar el rendimiento de las aplicaciones que realizan un uso intensivo de operaciones de E/S, puede usar volúmenes de IOPS aprovisionadas. Los volúmenes de IOPS aprovisionadas están diseñados para satisfacer las necesidades de las cargas de trabajo con uso intensivo de operaciones de E/S, y en especial de las cargas de trabajo de bases de datos, que son sensibles al rendimiento del almacenamiento y a la coherencia del rendimiento de las operaciones de E/S de acceso aleatorio.

Especifique una velocidad de IOPS cuando cree el volumen y los aprovisionamientos de Amazon EBS adecuados para la vida útil del volumen. Amazon EBS admite actualmente IOPS por volumen, desde un máximo de 16 000 (para todos los tipos de instancias) hasta 64 000 (para instancias creadas en Nitro System). Puede fragmentar múltiples volúmenes a la vez para ofrecer miles de IOPS por instancia a su aplicación. Aparte de esto, para cargas de trabajo esenciales y de mayor rendimiento que requieran una latencia inferior a un milisegundo, puede utilizar el tipo de volumen io2 block express, que admite hasta 256 000 IOPS con una capacidad de almacenamiento máxima de 64 TB.

Escalado automático de la flota

Una de las diferencias clave entre la arquitectura en la nube de AWS y el modelo de alojamiento tradicional es que AWS puede escalar automáticamente la flota de aplicaciones web bajo demanda para gestionar los cambios en el tráfico. En el modelo de alojamiento tradicional, suelen utilizarse modelos de previsión de tráfico para aprovisionar los hosts de acuerdo con el tráfico previsto. En AWS, las instancias se pueden aprovisionar sobre la marcha de acuerdo con un conjunto de desencadenadores para aumentar y reducir la flota.

El servicio Auto Scaling puede crear grupos de capacidad de servidores que pueden aumentar o reducirse según la demanda. Auto Scaling también funciona directamente con CloudWatch para ofrecer datos de métricas y con Elastic Load Balancing para añadir y eliminar hosts para la distribución de la carga. Por ejemplo, si los servidores web informan que la utilización de CPU es

superior al 80 por ciento durante un período de tiempo, se podría implementar rápidamente un servidor web adicional y, a continuación, añadirlo automáticamente al equilibrador de carga para incluirlo inmediatamente en la rotación del equilibrio de carga.

Como se muestra en el modelo de arquitectura de alojamiento web de AWS, puede crear varios grupos de Auto Scaling para diferentes capas de la arquitectura, de modo que cada capa se pueda escalar de forma independiente. Por ejemplo, el grupo de Auto Scaling del servidor web puede activar el escalado horizontal y vertical en respuesta a los cambios en las operaciones de E/S de la red, mientras que el grupo de Auto Scaling del servidor de aplicaciones podría escalarse horizontal y verticalmente en función del uso de la CPU. Puede establecer mínimos y máximos para ayudar a garantizar la disponibilidad las 24 horas del día y los 7 días de la semana y limitar el uso dentro de un grupo.

Los desencadenadores de Auto Scaling se pueden configurar tanto para aumentar como para reducir la flota total en una capa determinada para adaptar la utilización de los recursos a la demanda real. Además del servicio Auto Scaling, puede escalar flotas de Amazon EC2 directamente a través de la API de Amazon EC2, lo que permite lanzar, terminar e inspeccionar instancias.

Características de seguridad adicionales

El número y la sofisticación de los ataques de denegación de servicio distribuido (DDoS) están aumentando. Estos ataques siempre han sido difíciles de rechazar. Muchas veces, terminan siendo costosos tanto por el tiempo dedicado a mitigarlos y la energía gastada, como por el coste de oportunidad de las visitas al sitio web que se han perdido durante el ataque. Existen varios factores y servicios de AWS que pueden ayudarle a defenderse de estos ataques. Uno de ellos es la escala de la red de AWS. La infraestructura de AWS es bastante grande y le permite aprovechar nuestra escala para optimizar su defensa. Varios servicios, incluidos Elastic Load Balancing, Amazon CloudFront y Amazon Route 53, son eficaces para escalar la aplicación web en respuesta a un gran aumento en el tráfico.

En particular, los servicios de protección de la infraestructura le ayudan en su estrategia de defensa:

- AWS Shield es un servicio de protección de DDoS administrado que ayuda a proteger contra diversos tipos de vectores de ataque de DDoS. La oferta estándar de AWS Shield es gratuita y se activa automáticamente en toda su cuenta. Esta oferta estándar ayuda a defenderse de los ataques más comunes a la capa de red y transporte. Además de este nivel, la oferta avanzada ofrece niveles de protección más altos contra su aplicación web al proporcionarle visibilidad casi en tiempo real de cualquier ataque en curso, así como la integración en niveles más altos con los servicios mencionados anteriormente. Además, le permite acceder al equipo de respuesta de

DDoS (DRT) de AWS para ayudar a mitigar los ataques sofisticados y a gran escala contra sus recursos.

- AWS WAF (Web Application Firewall) está diseñado para proteger sus aplicaciones web de ataques que pueden comprometer la disponibilidad o la seguridad, o consumir recursos excesivos. AWS WAF funciona en línea con CloudFront o Application Load Balancer, junto con sus reglas personalizadas, para defenderse de ataques como scripting entre sitios, inyección SQL y DDoS. Como ocurre con la mayoría de los servicios de AWS, AWS WAF incluye una API con todas las características que puede ayudar a automatizar la creación y edición de reglas para su instancia de AWS WAF a medida que sus necesidades de seguridad cambian.
- AWS Firewall Manager es un servicio de administración de seguridad que permite la configuración y administración centralizadas de reglas de firewall en todas sus cuentas y aplicaciones en AWS Organizations. A medida que se crean nuevas aplicaciones, AWS Firewall Manager facilita el cumplimiento de los nuevos recursos y aplicaciones aplicando un conjunto común de reglas de seguridad.

Conmutación por error con AWS

Otra ventaja clave de AWS con respecto al alojamiento web tradicional son las zonas de disponibilidad que facilitan el acceso a las ubicaciones de implementación redundantes. Las zonas de disponibilidad son ubicaciones físicamente diferentes que están diseñadas para aislarse de fallos en otras zonas de disponibilidad. Proporcionan conectividad de red económica y de baja latencia con las demás zonas de disponibilidad dentro de la misma región de AWS. Como muestra el diagrama de la arquitectura de alojamiento web de AWS, AWS recomienda implementar hosts de EC2 en varias zonas de disponibilidad para que la aplicación web sea más tolerante a errores.

Es importante asegurarse de que existan disposiciones para migrar puntos de acceso únicos en las zonas de disponibilidad en caso de que se produzca un error. Por ejemplo, debe configurar una base de datos en espera en una segunda zona de disponibilidad para que la persistencia de los datos siga siendo coherente y de alta disponibilidad, incluso durante un escenario de error poco probable. Para hacer esto en Amazon EC2 o Amazon RDS, solo tiene que hacer clic en un botón.

Aunque muchas veces es necesario realizar algunos cambios en la arquitectura al trasladar una aplicación web existente a la nube de AWS, hay mejoras importantes en la escalabilidad, la fiabilidad y la rentabilidad que hacen que valga la pena usar la nube de AWS. En la siguiente sección se analizan esas mejoras.

Consideraciones clave al utilizar AWS para el alojamiento web

Existen algunas diferencias clave entre la nube de AWS y un modelo de alojamiento de aplicaciones web tradicional. En la sección anterior, se han destacado muchas de las áreas clave que debe tener en cuenta al implementar una aplicación web en la nube. En esta sección, se señalan algunos de los cambios arquitectónicos clave que debe tener en cuenta al llevar cualquier aplicación a la nube.

Se acabaron los dispositivos de red físicos

No es posible implementar dispositivos de red físicos en AWS. Por ejemplo, los firewalls, enrutadores y equilibradores de carga para las aplicaciones de AWS ya no pueden residir en dispositivos físicos y deben reemplazarse por soluciones de software. Existe una amplia variedad de soluciones de software de calidad empresarial, ya sea para equilibrar la carga o para establecer una conexión VPN. Esto no limita lo que se puede ejecutar en la nube de AWS, sino que se trata de un cambio en la arquitectura de su aplicación si actualmente utiliza estos dispositivos.

Firewalls en todas partes

Donde antes tenía una zona desmilitarizada (DMZ) sencilla y abría las comunicaciones entre sus hosts en un modelo de alojamiento tradicional, AWS utiliza un modelo más seguro, en el que todos los hosts están bloqueados. Uno de los pasos para planificar una implementación de AWS es analizar el tráfico entre los hosts. Este análisis permite tomar decisiones sobre qué puertos deben abrirse exactamente. Puede crear grupos de seguridad para cada tipo de host de su arquitectura. También puede crear una gran variedad de modelos de seguridad simples y en niveles para permitir el acceso mínimo entre los hosts de su arquitectura. El uso de listas de control de acceso a la red en Amazon VPC puede ayudar a bloquear su red en el nivel de subred.

Piense en la posibilidad de tener disponibles varios centros de datos

Considere las zonas de disponibilidad de una región de AWS como varios centros de datos. Las instancias de EC2 en diferentes zonas de disponibilidad están separadas de manera lógica y física, e incluyen un modelo muy sencillo para implementar su aplicación en los centros de datos para

conseguir alta disponibilidad y fiabilidad. Amazon VPC como servicio regional le permite utilizar las zonas de disponibilidad y, al mismo tiempo, mantener todos sus recursos en la misma red lógica.

Trate a los hosts como si fueran efímeros y dinámicos

Probablemente, el cambio más importante en la forma de diseñar su aplicación de AWS sea que los hosts de Amazon EC2 deben tratarse como si fueran efímeros y dinámicos. Ninguna aplicación creada para la nube de AWS debe asumir que un host va a estar disponible siempre, por lo que debe diseñarse con la idea de que los datos de los almacenes instantáneos de EC2 se perderán si falla una instancia de EC2.

Cuando se abre un nuevo host, no debe hacer suposiciones sobre su dirección IP o ubicación dentro de una zona de disponibilidad del host. Su modelo de configuración debe ser flexible y su enfoque sobre el arranque de un host debe tener en cuenta la naturaleza dinámica de la nube. Estas técnicas son fundamentales para crear y ejecutar una aplicación enormemente escalable y tolerante a fallos.

Tenga en cuenta las tecnologías de contenedores y sin servidor

Este documento técnico se centra principalmente en una arquitectura web más tradicional. Sin embargo, considere la posibilidad de modernizar sus aplicaciones web pasándose a las tecnologías de contenedores y sin servidor, para aprovechar servicios como AWS Fargate y AWS Lambda para abstraer el uso de máquinas virtuales para realizar tareas de computación. Con la informática sin servidor, AWS se encarga de las tareas de administración de la infraestructura, como el aprovisionamiento de la capacidad y la aplicación de revisiones, para que pueda crear aplicaciones más ágiles que le permitan innovar y responder a los cambios con mayor rapidez.

Considere la implementación automatizada

- Amazon Lightsail es un servidor virtual privado (VPS) fácil de usar que ofrece todo lo necesario para crear una aplicación o sitio web, además de un plan mensual rentable. Lightsail es perfecto para hacer más sencillas las cargas de trabajo, para implementaciones rápidas y para iniciarse en AWS. Está diseñado para ayudarle a empezar con poco y aumentar a medida que vaya creciendo.
- AWS Elastic Beanstalk es un servicio fácil de utilizar para 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. Solo tiene que cargar el código, y Elastic Beanstalk se encargará de manera automática de la implementación, el aprovisionamiento de la capacidad, el equilibrio de carga, el escalado automático y la supervisión del estado de la

aplicación. Al mismo tiempo, tendrá el control absoluto de los recursos de AWS que hacen posible el funcionamiento de su aplicación y podrá obtener acceso a los recursos subyacentes cuando quiera.

- AWS App Runner es un servicio completamente administrado que facilita a los desarrolladores la implementación rápida de aplicaciones web y API en contenedores, a escala y sin necesidad de experiencia previa en infraestructura. Comience con su código fuente o una imagen de contenedor. App Runner crea e implementa automáticamente la aplicación web y equilibra la carga del tráfico con cifrado. App Runner también se escala vertical u horizontalmente de forma automática para satisfacer las necesidades de tráfico.
- AWS Amplify es un conjunto de herramientas y servicios que se pueden utilizar juntos o de forma individual para ayudar a los desarrolladores de frontend web y móvil a crear aplicaciones de pila completa escalables, con la tecnología de AWS. Con Amplify, puede configurar backends de aplicaciones y conectar la aplicación en cuestión de minutos, implementar aplicaciones web estáticas con tan solo unos clics y administrar el contenido de las aplicaciones fácilmente fuera de la AWS Management Console.

Conclusión y colaboradores

Conclusión

Existen numerosos factores arquitectónicos y conceptuales que hay que tener en cuenta a la hora de migrar una aplicación web a la nube de AWS. Los beneficios de tener una infraestructura rentable, de alta escalabilidad y tolerante a fallos que crezca con su empresa superan con creces el esfuerzo que supone la migración a la nube de AWS.

Colaboradores

Las siguientes personas y organizaciones contribuyeron a redactar este documento:

- Amir Khairalomoum, arquitecto de soluciones sénior, AWS
- Dinesh Subramani, arquitecto de soluciones sénior, AWS
- Jack Hemion, arquitecto de soluciones sénior, AWS
- Jatin Joshi, ingeniero de soporte en la nube, AWS
- Jorge Fonseca, arquitecto de soluciones sénior, AWS
- Shinduri K S, arquitecto de soluciones, AWS

Documentación adicional

- Implementación en Amazon LightSail de una aplicación basada en Django
- Implementación de un sitio web de Drupal de alta disponibilidad en Elastic Beanstalk
- Implementación de una aplicación PHP de alta disponibilidad en Elastic Beanstalk
- Implementación de una aplicación Node.js con DynamoDB en Elastic Beanstalk
- Introducción a las aplicaciones web de Linux en la nube de AWS
- Alojar un sitio web estático
- Alojamiento de un sitio web estático mediante Amazon S3
- Tutorial: Implementación de una aplicación de ASP.NET Core con Elastic Beanstalk
- Tutorial: Cómo implementar una aplicación .NET de ejemplo mediante Elastic Beanstalk

Revisiones del documento

Para recibir notificaciones sobre las actualizaciones de este documento técnico, suscríbase a la fuente RSS.

update-history-change	update-history-description	update-history-date
Documento técnico actualizado	Se han actualizado varias secciones y diagramas con nuevos servicios, caracterí sticas y límites de servicio.	20 de agosto de 2021
Documento técnico actualizado	Etiqueta de icono actualizada para “Almacenamiento en caché con ElastiCache” en la figura 3.	29 de septiembre de 2019
Documento técnico actualizado	Se han añadido y actualizado varias secciones con nuevos servicios. Se han actualizado diagramas para mejorar la claridad y añadir servicios. Adición de VPC como método de red estándar en AWS en “Administración de red”. Se ha añadido una sección sobre protección y mitigación de DDoS en “Características de seguridad adicionales”. Se ha añadido una pequeña sección sobre arquitecturas sin servidor para alojamiento web.	1 de julio de 2017
Documento técnico actualizado	Se han actualizado varias secciones para mejorar la claridad. Se han actualizado	1 de septiembre de 2012

	los diagramas con iconos de AWS. Adición de la sección “Administración de DNS público” para incluir más información sobre Amazon Route 53. Se ha actualizado la sección “Búsqueda de otros hosts y servicios” para mayor claridad. Se ha actualizado la sección “Configuración de bases de datos, copias de seguridad y conmutaci ón por error” para mayor claridad e incluir DynamoDB. Se ha ampliado la sección “Almacenamiento y copia de seguridad de datos y recursos” para cubrir los volúmenes de IOPS aprovisionadas de EBS.	
Publicación inicial	Documento técnico publicado.	1 de mayo de 2010

Avisos

Este documento se ofrece solo con fines informativos. Representa la oferta actual de productos y prácticas de AWS a partir de la fecha de publicación de este documento. Dichas prácticas y productos pueden modificarse sin previo aviso. Los clientes son responsables de realizar sus propias evaluaciones independientes de la información contenida en este documento y de cualquier uso de los productos o servicios de AWS, cada uno de los cuales se ofrece “tal cual”, sin garantía de ningún tipo, ya sea explícita o implícita. Este documento no genera ninguna garantía, declaración, compromiso contractual, condición ni certeza por parte de AWS, sus filiales, proveedores o licenciantes. Las responsabilidades y obligaciones de AWS con respecto a sus clientes se controlan mediante los acuerdos de AWS y este documento no forma parte ni modifica ningún acuerdo entre AWS y sus clientes.