Mejores prácticas de AWS para la resistencia al DDoS

Junio de 2018

Avisos

Este documento se proporciona únicamente con fines informativos. Representa las ofertas de productos y las prácticas actuales de AWS en la fecha de emisión de este documento, que están sujetas a cambios sin previo aviso. Los clientes son responsables de realizar su propia evaluación independiente de la información contenida en este documento y de cualquier uso de los productos o servicios de AWS, cada uno de los cuales se proporciona "tal cual" sin garantía de ningún tipo, ya sea expresa o implícita. Este documento no crea ninguna garantía, representación, compromisos contractuales, condiciones o garantías de AWS, sus afiliados, proveedores o licenciantes. Las responsabilidades
y responsabilidades de AWS hacia sus clientes están controlados por los acuerdos de AWS, y este documento no es parte de, ni modifica, ningún acuerdo entre AWS y sus clientes.

Contenido

Introducción1

Ataques de denegación de servicio ***2***

Ataques a la capa de infraestructura **3**

Ataques a la capa de aplicación **5**

Técnicas de mitigació ***7***

Defensa de la capa de infraestructura (BP1, BP3, BP6, BP7) **10**

Defensa de la capa de aplicación (BP1, BP2, BP6) **14**

Reducción de la superficie de ataque ***16***

Ofuscación de los recursos de AWS (BP1, BP4, BP5) **16**

Técnicas operativas ***19***

Visibilidad **19**

Soporte **22**

Conclusión ***24***

Colaboradores ***24***

Revisiones de documentos ***24***

Apéndice A: Recursos adicionales ***26***

Resum

Este documento está dirigido a los clientes que desean mejorar la resistencia de sus aplicaciones que se ejecutan en Amazon Web Services (AWS) contra los ataques de denegación de servicio distribuidos (DDoS). Ofrece una visión general de los ataques DDoS, las capacidades proporcionadas por AWS, las técnicas de mitigación y una arquitectura de referencia resistente a los ataques DDoS que puede utilizarse como guía para ayudar a proteger la disponibilidad de las aplicaciones.

El documento está dirigido a los responsables de la toma de decisiones de TI y a los ingenieros de seguridad que están familiarizados con los conceptos básicos de redes, seguridad y AWS. Cada sección tiene enlaces a la documentación de AWS que ofrece más detalles sobre la práctica recomendada o la capacidad.

Introducció

Usted trabaja para proteger su negocio del impacto de los ataques de denegación de servicio distribuido (DDoS), así como de otros ciberataques. Quiere mantener la confianza de sus clientes en su servicio, manteniendo la disponibilidad y la capacidad de respuesta de su aplicación. Y quiere evitar costes directos innecesarios cuando su infraestructura deba escalar en respuesta a un ataque.

AWS se compromete a proporcionarle herramientas, prácticas recomendadas y servicios que le ayuden a garantizar una alta disponibilidad, seguridad y resistencia para defenderse de los malos actores en Internet.

En este libro blanco, le proporcionamos orientación prescriptiva sobre DDoS. Describimos diferentes tipos de ataques, como los de la capa de infraestructura y los de la capa de aplicación, y explicamos qué prácticas recomendadas son más eficaces para gestionar cada tipo de ataque. También describimos los servicios y características que encajan en una estrategia de mitigación de DDoS, y cómo se puede utilizar cada uno de ellos para ayudar a proteger sus aplicaciones.

Ataques de denegación de servicio

Un ataque de denegación de servicio (DoS) es un intento deliberado de hacer que su sitio web o aplicación no esté disponible para los usuarios, por ejemplo, inundándolo con tráfico de red. Para lograrlo, los atacantes utilizan una variedad de técnicas que consumen grandes cantidades de ancho de banda de la red o atan otros recursos del sistema, interrumpiendo el acceso de los usuarios legítimos. En su forma más simple, un atacante solitario utiliza una única fuente para ejecutar un ataque DoS contra un objetivo, como se muestra en el siguiente diagrama (Figura 1).

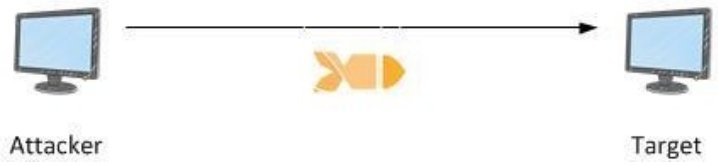

Figura 1: Diagrama de un ataque DOS

Pero en un ataque de denegación de servicio distribuido (DDoS), un atacante utiliza múltiples fuentes -que pueden ser grupos distribuidos de ordenadores infectados con malware, routers, dispositivos IoT y otros puntos finales- para orquestar un ataque contra un objetivo. Como se ilustra en el siguiente diagrama de ataque DDoS (Figura 2), una red de hosts comprometidos participa en el ataque, generando una avalancha de paquetes o solicitudes para abrumar al objetivo.

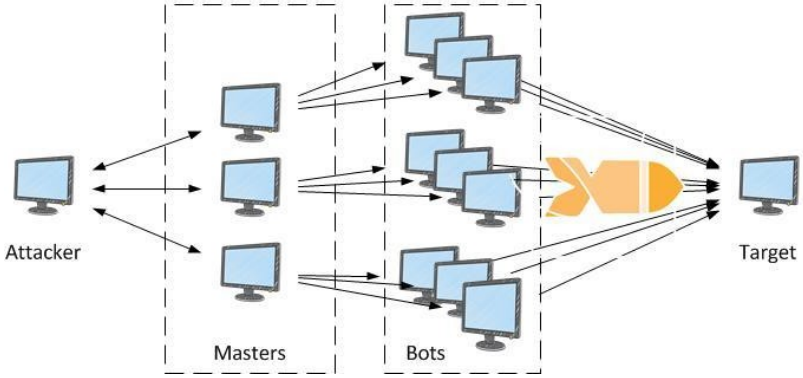

Figura 2: Esquema del ataque DDoS

Los ataques DDoS son más comunes en las capas 3, 4, 6 y 7 del modelo de Interconexión de Sistemas Abiertos (OSI), que se describe en la siguiente tabla (Tabla 1). Los ataques de las capas 3 y 4 corresponden a las capas de Red y Transporte del modelo OSI. Nos referiremos a ellos colectivamente como ataques a *la capa de infraestructura*. Los ataques de las capas 6 y 7 corresponden a las capas de Presentación y capas de aplicación del modelo OSI. Los abordaremos juntos como *ataques a la capa de aplicación*. En las siguientes secciones describiremos ejemplos de estos tipos de ataques.

#	Capa	Unidad	Descripción	Ejemplos de vectores
7	Aplicación	Datos	Proceso de red a la aplicación	Inundaciones HTTP, inundaciones de consultas DNS
6	Presentación	Datos	Representación y cifrado de datos	Abuso del TLS
5	Sesión	Datos	Comunicación entre hosts	N/A
4	Transporte	Segmentos	Conexiones de extremo a extremo y fiabilidad	Inundaciones SYN
3	Red	Paquetes	Determinación de la ruta y direccionamiento lógico	Ataques de reflexión UDP
2	Enlace de datos	Marcos	Direccionamiento físico	N/A
1	Físico	Bits	Medios, señales y transmisión binaria	N/A

Tabla 1: Modelo de interconexión de sistemas abiertos (OSI).

Ataques a la capa de infraestructura

Los ataques DDoS más comunes, los ataques de reflexión UDP y las inundaciones SYN, son ataques de capa de infraestructura. Un atacante puede utilizar cualquiera de estos métodos para generar grandes volúmenes de tráfico que pueden inundar la capacidad de una red o atar los recursos de sistemas como un servidor, un cortafuegos, un IPS o un equilibrador de carga. Aunque estos ataques pueden ser fáciles de identificar, para mitigarlos eficazmente, hay que tener una red o sistemas que aumenten su capacidad más rápidamente que la inundación de tráfico entrante. Esta capacidad adicional sirve para filtrar o absorber el tráfico de ataque, permitiendo a su sistema y aplicación responder al tráfico legítimo de sus clientes.

Ataques de reflexión UDP

Los ataques de reflexión UDP explotan el hecho de que UDP es un protocolo sin estado. Los atacantes pueden crear un paquete de solicitud UDP válido que incluya la IP del objetivo del ataque como dirección IP de origen UDP. El atacante ha falsificado, o "suplantado", la IP de origen del paquete de solicitud UDP. A continuación, el atacante envía el paquete UDP que contiene la IP de origen falsificada a un servidor intermedio. El servidor es engañado para que envíe sus paquetes de respuesta UDP a la IP de la víctima objetivo en lugar de devolverlos al

dirección IP del atacante. El servidor intermedio se utiliza porque genera una respuesta que es varias veces mayor que el paquete de solicitud, amplificando efectivamente la cantidad de tráfico de ataque enviado a la dirección IP objetivo.

El factor de amplificación, que es la relación entre el tamaño de la respuesta y el tamaño de la solicitud, varía en función del protocolo que utilice el atacante: DNS, NTP o SSDP. Por ejemplo, el factor de amplificación para DNS puede ser de 28 a 54 veces el número original de bytes. Así, si un atacante envía una carga útil de solicitud de 64 bytes a un servidor DNS, puede generar más de 3400 bytes de tráfico no deseado hacia un objetivo de ataque. La táctica de reflexión y el efecto de amplificación se ilustran en el siguiente diagrama (Figura 3).

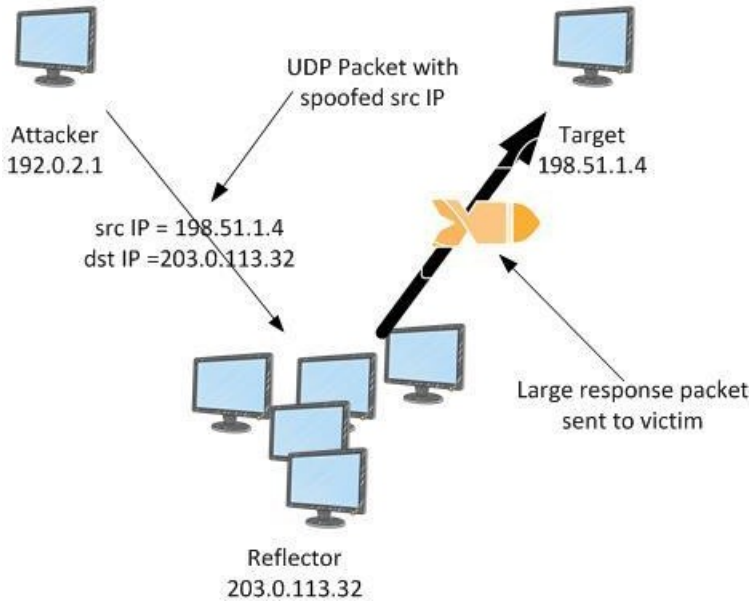

Figura 3: Ataque de reflexión UDP

Ataques de inundación SYN

Cuando un usuario se conecta a un servicio TCP, como un servidor web, su cliente envía un paquete SYN (de sincronización). El servidor devuelve un paquete SYN-ACK como acuse de recibo y, finalmente, el cliente responde con un paquete ACK, que completa el esperado handshake de tres vías. Este típico apretón de manos se ilustra en el siguiente diagrama (Figura 4):

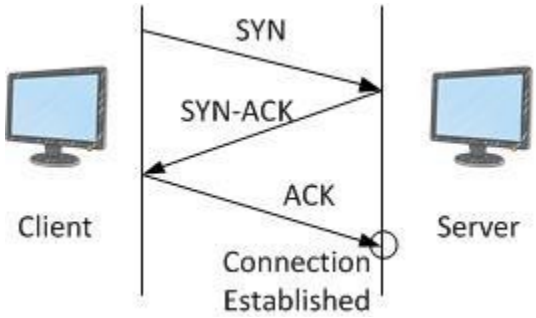

Figura 4: handshake SYN de 3 vías

En un ataque de inundación SYN, un cliente malicioso envía un gran número de paquetes SYN, pero nunca envía los paquetes ACK finales para completar los handshakes. El servidor se queda esperando una respuesta a las conexiones TCP medio abiertas y finalmente se queda sin capacidad para aceptar nuevas conexiones TCP. Esto puede impedir que nuevos usuarios se conecten al servidor. Las inundaciones SYN pueden alcanzar hasta 100s de Gbps, pero el ataque no tiene que ver con el volumen de tráfico SYN, sino con la inmovilización de las conexiones disponibles en el servidor, lo que hace que no haya recursos para las conexiones legítimas.

Ataques a la capa de aplicación

Un atacante puede dirigirse a la propia aplicación utilizando un ataque de capa 7 o de capa de aplicación. En estos ataques, similares a los ataques de infraestructura de inundación SYN, el atacante intenta sobrecargar funciones específicas de una aplicación para que ésta no esté disponible o no responda en absoluto a los usuarios legítimos. A veces esto se puede lograr con volúmenes de solicitudes muy bajos que generan sólo un pequeño volumen de tráfico de red. Esto puede hacer que el ataque sea difícil de detectar y mitigar. Entre los ejemplos de ataques a la capa de aplicación se encuentran las inundaciones HTTP, los ataques de ruptura de caché y las inundaciones XML-RPC de WordPress.

En un **ataque HTTP flood**, un atacante envía peticiones HTTP que parecen ser de un usuario real de la aplicación web. Algunas inundaciones HTTP se dirigen a un recurso específico, mientras que las inundaciones HTTP más complejas intentan emular la interacción humana con la aplicación. Esto puede aumentar la dificultad de utilizar técnicas comunes de mitigación como la limitación de la tasa de solicitudes.

Los ataques de ruptura de caché son un tipo de inundación HTTP que utiliza variaciones en la cadena de consulta para eludir el almacenamiento en caché de la red de distribución de contenidos (CDN). En lugar de poder devolver los resultados almacenados en la caché, la CDN debe contactar con el servidor de origen para cada solicitud de página, y estas recuperaciones de origen causan una tensión adicional en el servidor web de la aplicación.

Con un **ataque de inundación XML-RPC de WordPress**, también conocido como inundación de pingback de WordPress, un atacante utiliza indebidamente la función XML-RPC API de un sitio web alojado en el software de gestión de contenidos WordPress para generar una inundación de peticiones HTTP. La función de pingback permite a un sitio web alojado en WordPress (Sitio A) notificar a otro sitio de WordPress (Sitio B) que el Sitio A ha creado un enlace al Sitio B. El Sitio B intenta entonces buscar el Sitio A para verificar la existencia del enlace. En una inundación de pingback, el atacante utiliza esta capacidad para hacer que el Sitio B ataque al Sitio A. Este tipo de ataque tiene una firma clara: "WordPress" suele estar presente en el "User-Agent" de la cabecera de la petición HTTP.

También hay otras formas de tráfico malicioso que pueden afectar a la disponibilidad de una aplicación. Los bots raspadores automatizan los intentos de acceder a una aplicación web para robar contenidos o registrar información de la competencia, como los precios. Los ataques de fuerza bruta y de relleno de credenciales son esfuerzos programados para obtener acceso no autorizado a las áreas seguras de una aplicación. No son estrictamente ataques DDoS, pero su naturaleza automatizada puede parecerse a un ataque DDoS y pueden mitigarse aplicando algunas de las mismas mejores prácticas que se tratarán más adelante en este documento.

Los ataques a la capa de aplicación también pueden dirigirse a los servicios del sistema de nombres de dominio (DNS). El más común de estos ataques es una inundación de consultas DNS en la que un atacante utiliza muchas consultas DNS bien formadas para agotar los recursos de un servidor DNS. Estos ataques también pueden incluir un componente de ruptura de caché en el que el atacante aleatoriza la cadena de subdominios para eludir la caché DNS local de un determinado resolvedor. Como resultado, el resolvedor no puede aprovechar las consultas de dominio en caché y, en su lugar, debe contactar repetidamente con el servidor DNS autoritativo, lo que amplifica el ataque.

Si una aplicación web se entrega a través de TLS, un atacante también puede optar por atacar el proceso de negociación de TLS. TLS es computacionalmente caro, por lo que un atacante puede reducir la disponibilidad de un servidor enviando datos ininteligibles. En una variación de este ataque, un atacante completa el apretón de manos TLS pero renegocia perpetuamente el método de cifrado. O un atacante puede intentar agotar los recursos del servidor abriendo y cerrando muchas sesiones TLS.

Técnicas de mitigación

Algunas formas de mitigación de DDoS se incluyen automáticamente con los servicios de AWS. Puede mejorar aún más su resistencia a los DDoS utilizando una arquitectura de AWS con servicios específicos e implementando prácticas recomendadas adicionales.

Todos los clientes de AWS se benefician de las protecciones automáticas de AWS Shield Standard, sin cargo adicional. AWS Shield Standard defiende contra la mayoría de los ataques DDoS de la capa de red y de transporte que se producen con frecuencia y que tienen como objetivo su sitio web o sus aplicaciones. Se ofrece en todos los servicios de AWS y en todas las regiones de AWS, sin coste adicional. En las regiones de AWS, los ataques DDoS son detectados por un sistema que realiza automáticamente una línea de base del tráfico, identifica las anomalías y, si es necesario, crea mitigaciones. Este sistema de mitigación proporciona protección contra muchos ataques comunes de la capa de infraestructura. Puede utilizar AWS Shield Standard como parte de una arquitectura resistente a los ataques DDoS para proteger tanto aplicaciones web como no web.

Además, puede aprovechar los servicios de AWS que operan desde ubicaciones de borde, como Amazon CloudFront y Amazon Route 53, para crear una protección de disponibilidad completa contra todos los ataques conocidos de la capa de infraestructura. El uso de estos servicios, que forman parte de la red de borde global de AWS, puede mejorar la resistencia a los DDoS de su aplicación cuando sirve el tráfico de la aplicación web desde ubicaciones de borde distribuidas por todo el mundo.

Entre las ventajas específicas de utilizar Amazon CloudFront y Amazon Route 53 se incluyen las siguientes:

- Los sistemas de mitigación de DDoS de AWS Shield se integran con los servicios de borde de AWS, reduciendo el tiempo de mitigación de minutos a menos de un segundo.
- Técnicas de mitigación de inundaciones SYN sin estado que proxy y verificar las conexiones entrantes antes de pasarlas al servicio protegido.
- Sistemas automáticos de ingeniería de tráfico que pueden dispersar o aislar el impacto de los ataques DDoS de gran volumen.
- Defensa de la capa de aplicación cuando se combina con AWS WAF que no requiere cambiar su arquitectura de aplicación actual (por ejemplo, en una región de AWS o en un centro de datos local).

No se cobra por la transferencia de datos entrantes en AWS y no se paga por el tráfico de ataques DDoS que es mitigado por AWS Shield.

Para una arquitectura de referencia resistente a DDoS que incluya los servicios de AWS Global Edge Network, consulte el siguiente diagrama (Figura 5).

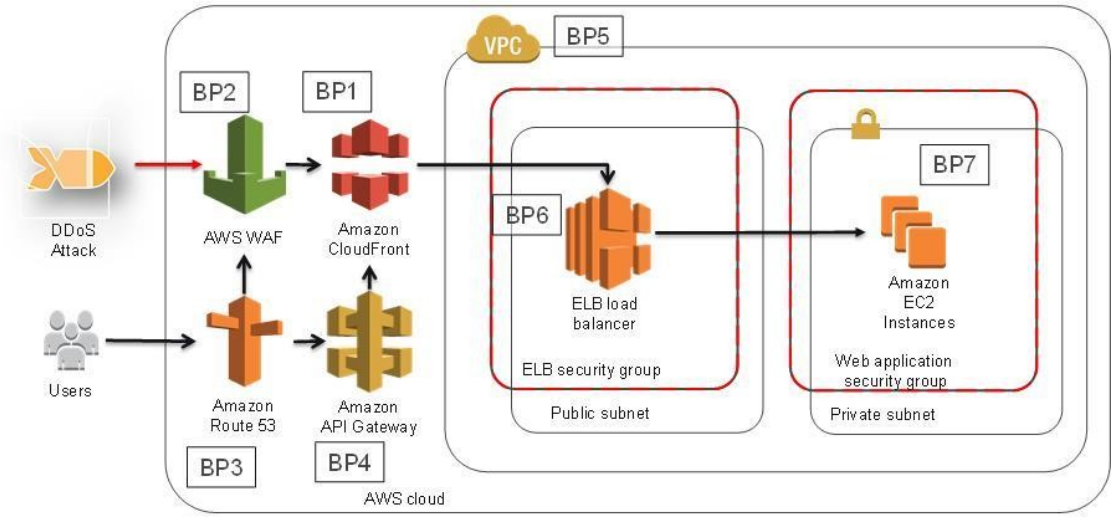

Figura 5: Arquitectura de referencia resistente a DDoS.

Esta arquitectura de referencia incluye varios servicios de AWS que pueden ayudarle a mejorar la resistencia de su aplicación web contra los ataques DDoS. Para ver un resumen de estos servicios y las capacidades que pueden proporcionar, consulte la siguiente tabla (Tabla 2). Hemos etiquetado cada servicio con un "indicador de mejores prácticas" (BP1, BP2, etc.) para que pueda consultar fácilmente cada uno de ellos aquí mientras lee el resto de este documento. Por ejemplo, en una próxima sección se analizan las capacidades proporcionadas por Amazon CloudFront y se incluye el indicador de prácticas recomendadas BP1.

	Ubicaciones de AWS Edge		**Regiones de AWS**			
	Amazon CloudFront (BP1) con AWS WAF (BP2)	Amazon Ruta 53 (BP3)	Elástico Carga Equilibrando (BP6)	Amazon API Puerta de enlace (BP4)	Amazon VPC (BP5)	Amazon EC2 con Auto Escala (BP7)
Mitigación de ataques de capa 3 (por ejemplo, reflexión UDP)	✓	✓	✓	✓	✓	✓

	Ubicaciones de AWS Edge		Regiones de AWS			
Mitigación de ataques de la capa 4 (por ejemplo, inundación SYN)	✓	✓	✓	✓		
Mitigación de ataques de la capa 6 (por ejemplo, TLS)	✓		✓	✓		
Reducir la superficie de ataque	✓	✓	✓	✓	✓	
Escala para absorber el tráfico de la capa de aplicación	✓	✓	✓	✓		✓
Mitigación de ataques de la capa 7 (capa de aplicación)	✓	✓	✓(si se utiliza con AWS WAF)			
Aislamiento geográfico y dispersión del exceso de tráfico, y ataques DDoS de mayor envergadura	✓	✓				

Cuadro 2: Resumen de las mejores prácticas.

Otra forma de mejorar su preparación para responder y mitigar los ataques DDoS es suscribiéndose a AWS Shield Advanced. Este servicio opcional de mitigación de DDoS le ayuda a proteger una aplicación alojada en cualquier región de AWS o alojada fuera de AWS. El servicio está disponible globalmente para Amazon CloudFront y Amazon Route 53. También está disponible en determinadas regiones de AWS para Classic Load Balancer (CLB), Application Load Balancer (ALB) y Elastic IP Addresses (EIP). El uso de AWS Shield Advanced con EIP le permite proteger instancias de Network Load Balancer (NLB) o Amazon EC2.

Con AWS Shield Advanced, obtendrá las siguientes ventajas adicionales:

- Acceso al equipo de respuesta a DDoS de AWS (DRT) para obtener asistencia en la mitigación de ataques DDoS que afectan a la disponibilidad de las aplicaciones.

- Visibilidad de los ataques DDoS mediante la consola de administración de AWS, la API y las métricas y alarmas de Amazon CloudWatch.
- Acceso al panel de control del Entorno Global de Amenazas, que proporciona una visión general de los ataques DDoS observados y mitigados por AWS.
- Acceso a AWS WAF, sin coste adicional, para la mitigación de ataques DDoS en la capa de aplicación (cuando se utiliza con Amazon CloudFront o ALB).
- Línea de base automática de los atributos del tráfico web, cuando se utiliza con AWS WAF.
- Acceso a AWS Firewall Manager, sin coste adicional, para la aplicación automática de políticas. Este servicio permite a los administradores de seguridad controlar y administrar de forma centralizada las reglas de AWS WAF.
- Umbrales de detección sensibles que dirigen el tráfico al sistema de mitigación de DDoS antes y pueden mejorar el tiempo de mitigación de los ataques contra Amazon EC2 o NLB, cuando se utilizan con EIP.
- Protección de costes que le permite solicitar un reembolso limitado de los costes relacionados con el escalado que resulten de un ataque DDoS.
- Acuerdo de nivel de servicio mejorado que es específico para los clientes de AWS Shield Advanced.

Para obtener una lista completa de las características de AWS Shield Advanced y para obtener más información sobre AWS Shield, consulte AWS Shield - Managed DDoS Protection.

En las siguientes secciones, describiremos cada una de las mejores prácticas recomendadas para la mitigación de DDoS en mayor profundidad. Para obtener una guía rápida y fácil de implementar sobre la creación de una capa de mitigación de DDoS para aplicaciones web estáticas o dinámicas, consulte Cómo ayudar a proteger las aplicaciones web dinámicas contra los ataques de DDoS utilizando Amazon CloudFront y Amazon Route 53.

Defensa de la capa de infraestructura (BP1, BP3, BP6, BP7)

En un entorno de centro de datos tradicional, puede mitigar los ataques DDoS de la capa de infraestructura mediante técnicas como el sobreaprovisionamiento de capacidad, la implementación de sistemas de mitigación de DDoS o el lavado de tráfico con la ayuda de servicios de mitigación de DDoS. En AWS, las capacidades de mitigación de DDoS se proporcionan automáticamente, pero puede optimizar la resistencia a los ataques DDoS de su aplicación tomando decisiones de arquitectura que aprovechen mejor esas capacidades y que también le permitan escalar el exceso de tráfico.

Las consideraciones clave para ayudar a mitigar los ataques DDoS volumétricos incluyen asegurar que haya suficiente capacidad de tránsito y diversidad, y proteger sus recursos de AWS, como las instancias de Amazon EC2, contra el tráfico de ataque.

Tamaño de la instancia (BP7)

Amazon EC2 ofrece una capacidad de cómputo redimensionable para que pueda escalar rápidamente hacia arriba o hacia abajo según cambien sus necesidades. Puede escalar horizontalmente añadiendo instancias a su aplicación, y escalar verticalmente utilizando instancias más grandes. Algunos tipos de instancias de Amazon EC2 admiten características que pueden manejar más fácilmente grandes volúmenes de tráfico, por ejemplo, interfaces de red de 25 gigabits y redes mejoradas.

Con 25 interfaces de red Gigabit, cada instancia puede soportar un mayor volumen de tráfico. Esto ayuda a evitar la congestión de la interfaz para el tráfico que ha llegado a la instancia de Amazon EC2.
Las instancias que soportan Enhanced Networking proporcionan un mayor rendimiento de E/S y una menor utilización de la CPU en comparación con las implementaciones tradicionales. Esto mejora la capacidad de la instancia para gestionar el tráfico con mayores volúmenes de paquetes.

La característica de 25 Gigabit está disponible en los tamaños más grandes de instancias, por ejemplo, M4. Para obtener más información sobre las instancias de Amazon EC2 que admiten interfaces de red de 25 gigabits y redes mejoradas, consulte Tipos de instancias de Amazon EC2. Para saber cómo habilitar las redes mejoradas, consulte Cómo habilitar las redes mejoradas en las instancias de Linux en una VPC.

Elección de la región (BP7)

Los servicios de AWS están disponibles en varios lugares del mundo. Estas áreas separadas geográficamente se denominan regiones. Al diseñar su aplicación, puede elegir una o más regiones en función de sus necesidades. Las consideraciones comunes incluyen el rendimiento, el coste y la soberanía de los datos. En cada región, AWS proporciona acceso a un conjunto único de conexiones de Internet y relaciones de peering para proporcionar una latencia y un rendimiento óptimos a los usuarios de esas zonas.

También es importante tener en cuenta la resistencia a los ataques DDoS cuando elija las regiones para su aplicación. Muchas regiones están cerca de los intercambios de Internet, por lo que tienen más conectividad con las principales redes. Estar cerca de los intercambios donde los transportistas internacionales y los grandes pares tienen una fuerte presencia puede ayudar a darle la capacidad de Internet para mitigar los ataques más grandes.

Para saber más sobre cómo elegir una región, consulte Regiones y zonas de disponibilidad. También puede preguntar a su equipo de cuentas sobre las características de cada región que esté considerando, para ayudarle a tomar una decisión informada.

Equilibrio de carga (BP6)

Los grandes ataques DDoS pueden sobrepasar la capacidad de una sola instancia de Amazon EC2, por lo que añadir un equilibrio de carga puede ayudar a su capacidad de recuperación. Existen varias opciones entre las que puede elegir para ayudar a mitigar un ataque equilibrando la carga del exceso de tráfico. Con Elastic Load Balancing (ELB), puede reducir el riesgo de sobrecarga de su aplicación distribuyendo el tráfico entre muchas instancias de backend. El ELB puede escalar automáticamente, lo que le permite administrar volúmenes más grandes cuando tiene un tráfico extra no previsto, por ejemplo, debido a multitudes de flashes o ataques DDoS. Para las aplicaciones construidas dentro de Amazon VPC, hay dos tipos de ELB a considerar, dependiendo de su tipo de aplicación: Application Load Balancer (ALB) o Network Load Balancer (NLB).

En el caso de las aplicaciones web, puede utilizar el ALB para enrutar el tráfico en función de su contenido y aceptar sólo las solicitudes web bien formadas. Esto significa que muchos ataques DDoS comunes, como las inundaciones SYN o los ataques de reflexión UDP, serán bloqueados por ALB, protegiendo su aplicación del ataque. Cuando el ALB detecta este tipo de ataques, se escala automáticamente para absorber el tráfico adicional. Para obtener más información sobre la protección de las aplicaciones web con ALB, consulte Introducción a los balanceadores de carga de aplicaciones.

En el caso de las aplicaciones basadas en TCP, puede utilizar NLB para dirigir el tráfico a las instancias de Amazon EC2 con una latencia ultrabaja. Cuando crea un NLB, se crea una interfaz de red para cada zona de disponibilidad (AZ) que habilite. Tiene la opción de asignar una dirección IP elástica (EIP) por subred habilitada para el balanceador de carga. Una consideración clave con NLB es que cualquier tráfico que llegue al equilibrador de carga en una escucha válida se dirigirá a sus instancias de Amazon EC2, no se absorberá. Para obtener más información sobre la protección de las aplicaciones TCP con NLB, consulte Introducción a los balanceadores de carga de red.

Entregar a escala utilizando las ubicaciones de borde de AWS (BP1, BP3)

El acceso a conexiones de Internet diversas y de gran escala puede aumentar significativamente su capacidad para optimizar la latencia y el desempeño para los usuarios, para absorber los ataques DDoS y para aislar los fallos mientras se minimiza el impacto en la disponibilidad de su aplicación. Las ubicaciones de borde de AWS ofrecen una capa adicional de infraestructura de red que proporciona estos beneficios a cualquier aplicación web que utilice Amazon CloudFront y Amazon Route 53. Cuando utiliza estos servicios, su contenido se sirve y las consultas DNS se resuelven desde ubicaciones que suelen estar más cerca de sus usuarios.

Entrega de aplicaciones web en el borde (BP1)

Amazon CloudFront es un servicio que puede utilizarse para entregar todo su sitio web, incluido el contenido estático, dinámico, de streaming e interactivo. Se pueden utilizar conexiones TCP persistentes y configuraciones variables de tiempo de vida (TTL) para descargar el tráfico de su origen, incluso si no está

servir contenido almacenable en caché. Estas características significan que el uso de Amazon CloudFront reduce el número de solicitudes y conexiones TCP de vuelta a su origen, lo que ayuda a proteger su aplicación web de las inundaciones HTTP. Amazon CloudFront solo acepta conexiones bien formadas, lo que ayuda a evitar que muchos ataques DDoS comunes, como las inundaciones SYN y los ataques de reflexión UDP, lleguen a su origen. Los ataques DDoS también se aíslan geográficamente cerca del origen, lo que impide que el tráfico afecte a otras ubicaciones. Estas capacidades pueden mejorar en gran medida su capacidad para seguir sirviendo tráfico a los usuarios durante grandes ataques DDoS. Puede utilizar Amazon CloudFront para proteger un origen en AWS o en otro lugar de Internet.

Si utiliza Amazon S3 para servir contenido estático en Internet, debe utilizar Amazon CloudFront para proteger su bucket. Puede utilizar Origin Access Identify (OAI) para asegurarse de que los usuarios solo acceden a sus objetos utilizando las URL de CloudFront. Para obtener más información sobre OAI, consulte Uso de una identidad de acceso al origen para restringir el acceso a su contenido de Amazon S3.

Para obtener más información sobre cómo proteger y optimizar el desempeño de las aplicaciones web mediante Amazon CloudFront, consulte Introducción a CloudFront.

Resolución de nombres de dominio en el borde (BP3)

Amazon Route 53 es un servicio de sistema de nombres de dominio (DNS) altamente disponible y escalable que puede utilizarse para dirigir el tráfico a su aplicación web. Incluye características avanzadas como el flujo de tráfico, el enrutamiento basado en la latencia, el DNS geográfico y las comprobaciones de estado y la monitorización que le permiten controlar cómo responde el servicio a las solicitudes de DNS, para mejorar el desempeño de su aplicación web y evitar las interrupciones del sitio.

Amazon Route 53 utiliza técnicas como el shuffle sharding y el anycast striping, que pueden ayudar a los usuarios a acceder a su aplicación incluso si el servicio DNS es el objetivo de un ataque DDoS. Con el shuffle sharding, cada servidor de nombres de su conjunto de delegación corresponde a un conjunto único de ubicaciones de borde y rutas de Internet. Esto proporciona una mayor tolerancia a los fallos y minimiza el solapamiento entre clientes. Si un servidor de nombres del conjunto de delegaciones no está disponible, los usuarios pueden reintentar y recibir una respuesta de otro servidor de nombres en una ubicación de borde diferente. El striping Anycast permite que cada solicitud de DNS sea atendida por la ubicación más óptima, repartiendo la carga de la red y reduciendo la latencia del DNS. A su vez, esto proporciona una respuesta más rápida para los usuarios. Además, Amazon Route 53 puede detectar anomalías en el origen y el volumen de las consultas de DNS y priorizar las solicitudes de los usuarios que se sabe que son fiables.

Para obtener más información sobre el uso de Amazon Route 53 para dirigir a los usuarios a su aplicación, consulte Introducción a Amazon Route 53.

Defensa de la capa de aplicación (BP1, BP2, BP6)

Muchas de las técnicas discutidas en este documento son eficaces para mitigar el impacto que los ataques DDoS de la capa de infraestructura tienen en la disponibilidad de su aplicación. Para defenderse también de los ataques a la capa de aplicación es necesario implementar una arquitectura que permita detectar específicamente, escalar para absorber y bloquear las peticiones maliciosas. Esta es una consideración importante porque los sistemas de mitigación de DDoS basados en la red son generalmente ineficaces para mitigar los ataques complejos a la capa de aplicación.

Detectar y filtrar solicitudes web maliciosas (BP1, BP2)

Cuando su aplicación se ejecuta en AWS, puede aprovechar tanto Amazon CloudFront como AWS WAF para ayudar a defenderse de los ataques DDoS en la capa de aplicación.

Amazon CloudFront le permite almacenar en caché contenido estático y servirlo desde las ubicaciones de borde de AWS, lo que puede ayudar a reducir la carga de su origen. También puede ayudar a reducir la carga del servidor evitando que el tráfico no web llegue a su origen. Además, CloudFront puede cerrar automáticamente las conexiones de los atacantes de lectura o escritura lenta (por ejemplo, Slowloris).

Al utilizar AWS WAF, puede configurar listas de control de acceso web (ACL web) en sus distribuciones de CloudFront o balanceadores de carga de aplicaciones para filtrar y bloquear solicitudes en función de las firmas de las solicitudes. Cada ACL web consta de reglas que puede configurar para que coincidan con cadenas o regex con uno o más atributos de la solicitud, como la URI, la cadena de consulta, el método HTTP o la clave del encabezado. Además, al utilizar las reglas basadas en la tasa de AWS WAF, puede bloquear automáticamente las direcciones IP de los malos actores cuando las solicitudes que coinciden con una regla superan un umbral que usted define.
Las solicitudes procedentes de las direcciones IP de los clientes infractores recibirán respuestas de error "403 Forbidden" y permanecerán bloqueadas hasta que la tasa de solicitudes descienda por debajo del umbral. Esto es útil para mitigar los ataques de inundación HTTP que se disfrazan de tráfico web regular.

Para bloquear los ataques procedentes de direcciones IP conocidas que actúan mal, puede crear reglas mediante condiciones de coincidencia de IP o utilizar las reglas administradas para AWS WAF que ofrecen los vendedores en AWS Marketplace y que bloquearán direcciones IP maliciosas específicas que se incluyen en las listas de reputación de IP. Tanto AWS WAF como Amazon CloudFront también permiten establecer restricciones geográficas para bloquear o poner en la lista blanca las solicitudes procedentes de determinados países. Esto puede ayudar a bloquear los ataques que se originan en ubicaciones geográficas en las que no se espera servir a los usuarios.

Para ayudar a identificar las solicitudes maliciosas, puede revisar los registros de su servidor web o utilizar la función de solicitudes muestreadas de WAF, que proporciona detalles sobre las solicitudes que han coincidido con una de sus reglas de AWS WAF durante un período de tiempo en las últimas 3 horas. Puede utilizar esta información para identificar firmas de tráfico potencialmente maliciosas y crear una nueva regla para denegar esas solicitudes.

Si está suscrito a AWS Shield Advanced, puede contratar al AWS DDoS Response Team (DRT) para que le ayude a crear reglas para mitigar un ataque que esté perjudicando la disponibilidad de su aplicación. El DRT sólo puede obtener un acceso limitado a su cuenta, y sólo con su autorización explícita. Para más información, consulte la sección de Soporte más adelante en este documento.

El uso de AWS Firewall Manager puede ayudar a simplificar la gestión de las reglas de AWS WAF en su organización. Al utilizar AWS Firewall Manager, puede habilitar AWS WAF en muchas cuentas y recursos, incluida la creación de reglas que se aplican automáticamente a las cuentas existentes o nuevas de su organización.

Para obtener más información sobre el uso de la restricción geográfica para limitar el acceso a su distribución de Amazon CloudFront, consulte Restricción de la distribución geográfica de su contenido.

Para obtener más información sobre el uso de AWS WAF, consulte Introducción a AWS WAF y Ver una muestra de las solicitudes web que CloudFront ha reenviado a AWS WAF.

Para obtener más información sobre la configuración de reglas basadas en tasas, consulte Proteger sitios y servicios web mediante reglas basadas en tasas para AWS WAF.

Para saber cómo administrar la implementación de las reglas de AWS WAF en sus recursos de AWS con AWS Firewall Manager, consulte Introducción a AWS Firewall Manager.

Escala para absorber (BP6)

Otra forma de mitigar los ataques a la capa de aplicación es operar a escala. Si tiene aplicaciones web, puede utilizar el ELB para distribuir el tráfico a una serie de instancias de Amazon EC2 sobreaprovisionadas o configuradas para escalar automáticamente. Estas instancias pueden manejar los aumentos repentinos de tráfico que se produzcan por cualquier motivo, incluyendo una multitud de flashes o un ataque DDoS en la capa de aplicación. Puede configurar las alarmas de Amazon CloudWatch para que inicien el escalado automático para escalar automáticamente el tamaño de su flota de Amazon EC2 en respuesta a los eventos que defina. Esto protege la disponibilidad de las aplicaciones cuando se produce un aumento inesperado del volumen de solicitudes. Si utiliza Amazon CloudFront o el Application Load Balancer (ALB) con su aplicación, la negociación de TLS es gestionada por la distribución o el balanceador de carga. Esto ayuda a proteger sus instancias para que no se vean afectadas por los ataques basados en TLS, ya que se escalan para gestionar tanto las solicitudes legítimas como los ataques de abuso de TLS.

Para obtener más información sobre el uso de Amazon CloudWatch para invocar Auto Scaling, consulte Supervisión de sus instancias y grupos de Auto Scaling mediante Amazon CloudWatch.

Reducción de la superficie de ataque

Otra consideración importante cuando se diseña una solución de AWS es limitar las oportunidades que tiene un atacante para atacar su aplicación. Por ejemplo, si no espera que los usuarios interactúen directamente con ciertos recursos, puede asegurarse de que esos recursos no sean accesibles desde Internet. Del mismo modo, si no espera que los usuarios o las aplicaciones externas se comuniquen con su aplicación en determinados puertos o protocolos, puede asegurarse de que ese tráfico no sea aceptado.

Este concepto se conoce como *reducción de la superficie de ataque*. En esta sección, proporcionamos las mejores prácticas para ayudarle a reducir su superficie de ataque y limitar la exposición de su aplicación a Internet. Los recursos que no están expuestos a Internet son más difíciles de atacar, lo que limita las opciones que tiene un atacante para atacar la disponibilidad de su aplicación.

Ofuscación de los recursos de AWS (BP1, BP4, BP5)

Normalmente, los usuarios pueden utilizar rápida y fácilmente una aplicación sin necesidad de que los recursos de AWS estén totalmente expuestos a Internet. Por ejemplo, cuando tiene instancias de Amazon EC2 detrás de un ELB, puede que no sea necesario que las propias instancias sean accesibles al público. En su lugar, podría proporcionar a los usuarios acceso al ELB en determinados puertos TCP y permitir que sólo el ELB se comunique con las instancias. Puede establecer esto configurando grupos de seguridad y listas de control de acceso a la red (NACL) dentro de su Amazon Virtual Private Cloud (VPC). Amazon VPC le permite aprovisionar una sección lógicamente aislada de la nube de AWS en la que puede lanzar recursos de AWS en una red virtual que usted defina.

Los grupos de seguridad y las NACL son similares en el sentido de que le permiten controlar el acceso a los recursos de AWS dentro de su VPC. Pero los grupos de seguridad le permiten controlar el tráfico entrante y saliente a nivel de instancia, mientras que las NACL ofrecen capacidades similares a nivel de subred de la VPC. El uso de los grupos de seguridad o de las NACL no conlleva ningún coste adicional.

Grupos de Seguridad y Listas de Control de Acceso a la Red (NACLs) (BP5)

Puede especificar los grupos de seguridad cuando lanza una instancia, o puede asociar la instancia con un grupo de seguridad en un momento posterior. Todo el tráfico de Internet a un grupo de seguridad se deniega implícitamente a menos que cree una regla de permiso para permitir el tráfico. Por ejemplo, si tiene una aplicación web que utiliza un ELB y varias instancias de Amazon EC2, puede decidir
cree un grupo de seguridad para el ELB ("grupo de seguridad del ELB") y otro para las instancias ("grupo de seguridad del servidor de aplicaciones web"). A continuación, puede crear una regla de *permiso* para permitir el tráfico de Internet.

al grupo de seguridad del ELB, y otra regla para permitir el tráfico desde el grupo de seguridad del ELB al grupo de seguridad del servidor de aplicaciones web. Esto garantiza que el tráfico de Internet no pueda comunicarse directamente con sus instancias de Amazon EC2, lo que hace más difícil que un atacante conozca su aplicación y la afecte.

Cuando creas NACLs, puedes especificar tanto reglas de *permiso* como de *denegación*. Esto es útil si quieres negar explícitamente ciertos tipos de tráfico a tu aplicación. Por ejemplo, puede definir direcciones IP (como rangos CIDR), protocolos y puertos de destino a los que se les niega el acceso a toda la subred. Si su aplicación se utiliza sólo para el tráfico TCP, puede crear una regla para *denegar* todo el tráfico UDP, o viceversa. Esta opción es útil cuando se responde a ataques DDoS porque le permite crear sus propias reglas para mitigar el ataque cuando conoce las IPs de origen u otra firma.

Si está suscrito a AWS Shield Advanced, puede registrar IPs elásticas (EIPs) como recursos protegidos. Los ataques DDoS contra EIPs que han sido registrados como Recursos Protegidos se detectan más rápidamente, lo que puede resultar en un tiempo más rápido de mitigación. Cuando se detecta un ataque, los sistemas de mitigación de DDoS leen la NACL que corresponde al EIP objetivo y la aplican en la frontera de la red de AWS. Esto reduce significativamente el riesgo de impacto de una serie de ataques DDoS en la capa de infraestructura.

Para obtener más información sobre la configuración de grupos de seguridad y NACL para optimizar la resistencia a los ataques DDoS, consulte Cómo prepararse para los ataques DDoS reduciendo la superficie de ataque.

Para obtener más información sobre el uso de AWS Shield Advanced con los EIP como recursos protegidos, consulte Habilitar y configurar AWS Shield Advanced.

Proteger su origen (BP1, BP5)

Si utiliza Amazon CloudFront con un origen que está dentro de su VPC, debe utilizar una función de AWS Lambda para actualizar automáticamente las reglas de su grupo de seguridad para *permitir* únicamente el tráfico de Amazon CloudFront. Esto mejora la seguridad de su origen al ayudar a garantizar que los usuarios maliciosos no puedan eludir Amazon CloudFront y AWS WAF al acceder a su aplicación web.

Para obtener más información sobre cómo proteger su origen mediante la actualización automática de sus grupos de seguridad, consulte Cómo actualizar automáticamente sus grupos de seguridad para Amazon CloudFront y AWS WAF mediante AWS Lambda.

También puede querer asegurarse de que solo su distribución de Amazon CloudFront pueda reenviar solicitudes a su origen. Con los encabezados de solicitud de borde a origen, puede añadir o anular el valor de los encabezados de solicitud existentes cuando Amazon CloudFront reenvía solicitudes a su origen. Puede utilizar el encabezado *X-Shared-Secret* para ayudar a validar que las solicitudes realizadas a su origen se han enviado desde Amazon CloudFront.

Para saber más sobre cómo proteger su origen con una cabecera *X-Shared-Secret*, consulte Reenviar cabeceras personalizadas a su origen.

Protección de los puntos finales de la API (BP4)

Normalmente, cuando debe exponer una API al público, existe el riesgo de que la puerta de entrada de la API sea objeto de un ataque DDoS. Para ayudar a reducir el riesgo, puede utilizar Amazon API Gateway como "puerta de entrada" a las aplicaciones que se ejecutan en Amazon EC2, AWS Lambda o en otro lugar. Al utilizar Amazon API Gateway, no necesita sus propios servidores para el frontend de la API y puede ofuscar otros componentes de su aplicación. Al dificultar la detección de su
de la aplicación, puede ayudar a evitar que esos recursos de AWS sean el objetivo de un ataque DDoS.

Cuando utiliza Amazon API Gateway, puede elegir entre dos tipos de puntos de enlace de API. La primera es la opción predeterminada: puntos de enlace de API optimizados para el borde a los que se accede a través de una distribución de Amazon CloudFront. Sin embargo, la distribución es creada y administrada por API Gateway, por lo que no tiene control sobre ella. La segunda opción es utilizar un punto de enlace de API regional al que se accede desde la misma región de AWS en la que se implementa su API REST. Le recomendamos que utilice el segundo tipo de punto de enlace y que lo asocie con su propia distribución de Amazon CloudFront. De este modo, tendrá control sobre la distribución de Amazon CloudFront y la posibilidad de utilizar AWS WAF para la protección de la capa de aplicación.

Cuando utilice Amazon CloudFront y AWS WAF con Amazon API Gateway, configure las siguientes opciones:

- Configure el comportamiento de la caché para que sus distribuciones reenvíen todos los encabezados al punto final regional de API Gateway. Al hacer esto, CloudFront tratará el contenido como dinámico y omitirá el almacenamiento en caché del contenido.
- Proteja su API Gateway contra el acceso directo configurando la distribución para que incluya la cabecera personalizada de origen *x-api-key*, estableciendo el valor de la clave de API en API Gateway.
- Proteja su backend del exceso de tráfico configurando límites de velocidad estándar o de ráfaga para cada método de sus APIs REST.

  Para obtener más información sobre la creación de API con Amazon API Gateway, consulte Introducción a Amazon API Gateway

Técnicas operativas

Las técnicas de mitigación de este documento le ayudarán a diseñar aplicaciones que sean intrínsecamente resistentes a los ataques DDoS. En muchos casos, también es útil saber cuándo un ataque DDoS se dirige a su aplicación para poder tomar medidas de mitigación. Si experimenta un ataque, también puede beneficiarse de la ayuda para evaluar la amenaza y revisar la arquitectura de su aplicación, o puede solicitar otro tipo de asistencia. En esta sección se analizan las mejores prácticas para obtener visibilidad del comportamiento anormal, las alertas y la automatización, así como para solicitar a AWS asistencia adicional.

Visibilidad

Cuando una métrica operativa clave se desvía sustancialmente del valor esperado, un atacante puede estar intentando atacar la disponibilidad de su aplicación. Si está familiarizado con el comportamiento normal de su aplicación, podrá actuar más rápidamente cuando detecte una anomalía.

Al utilizar Amazon CloudWatch, puede monitorizar las aplicaciones que ejecuta en AWS. Por ejemplo, puede recopilar y realizar un seguimiento de las métricas, recopilar y monitorizar los archivos de registro, establecer alarmas y responder automáticamente a los cambios en sus recursos de AWS.

Si ha diseñado su aplicación siguiendo la arquitectura de referencia resistente a DDoS, los ataques comunes de la capa de infraestructura se bloquearán antes de llegar a su aplicación. Si está suscrito a AWS Shield Advanced, tiene acceso a una serie de métricas de CloudWatch que pueden indicar que su aplicación está siendo atacada. Por ejemplo, puede configurar alarmas para que le notifiquen cuando haya un ataque DDoS en curso, de modo que pueda comprobar su
salud de la aplicación y decidir si se activa la DRT. Puede configurar la métrica DDoSDetected para que le indique si se ha detectado un ataque. Si desea recibir una alerta basada en el volumen de los ataques, también puede utilizar las métricas DDoSAttackBitsPerSecond, DDoSAttackPacketsPerSecond o DDoSAttackRequestsPerSecond. Puede monitorizar estas métricas integrando Amazon CloudWatch con sus propias herramientas o utilizando herramientas proporcionadas por terceros, como Slack o PagerDuty.

Un ataque a la capa de aplicación puede elevar muchas métricas de Amazon CloudWatch. Si utiliza AWS WAF, puede utilizar CloudWatch para monitorizar y alarmar sobre los aumentos de las solicitudes que haya configurado en WAF para ser permitidas, contabilizadas o bloqueadas. Esto le permite recibir una notificación si el nivel de tráfico supera lo que su aplicación puede soportar. También puede utilizar las métricas de Amazon CloudFront, Amazon Route 53, ALB, NLB, Amazon EC2 y Auto Scaling que se rastrean en CloudWatch para detectar cambios que puedan indicar un ataque DDoS.

Para obtener una descripción de las métricas de Amazon CloudWatch que se utilizan habitualmente para detectar y reaccionar ante los ataques DDoS, consulte la siguiente tabla (Tabla 3).

Tema	Métrica	Descripción
Escudo AWS Avanzado	DDoS detectado	Indica un evento DDoS para un nombre de recurso de Amazon (ARN) específico.
Escudo AWS Avanzado	Ataque DDoS Bits por segundo	El número de bytes observados durante un evento DDoS para un nombre de recurso de Amazon (ARN) específico. Esta métrica solo está disponible para eventos DDoS de capa 3/4.
Escudo AWS Avanzado	Paquetes de ataque DDoS por segundo	El número de paquetes observados durante un evento DDoS para un nombre de recurso de Amazon (ARN) específico. Esta métrica solo está disponible para eventos DDoS de capa 3/4.
Escudo AWS Avanzado	Solicitudes de ataques DDoS por segundo	El número de solicitudes observadas durante un evento de DDoS para un nombre de recurso de Amazon (ARN) específico. Esta métrica solo está disponible para los eventos de DDoS de capa 7 y solo se informa de los eventos de capa 7 más significativos.
AWS WAF	Solicitudes permitidas	El número de peticiones web permitidas.
AWS WAF	Solicitudes bloqueadas	El número de solicitudes web bloqueadas.
AWS WAF	Solicitudes contabilizadas	El número de peticiones web contadas.
Amazon CloudFront	Solicita	El número de solicitudes HTTP/S
Amazon CloudFront	Tasa de error total	El porcentaje de todas las solicitudes cuyo código de estado HTTP es 4xx o 5xx.
Ruta 53 del Amazonas	Estado del chequeo de salud	El estado del punto final del chequeo de salud.
ALB	Recuento de conexiones activas	El número total de conexiones TCP concurrentes que están activas desde los clientes al equilibrador de carga, y desde el equilibrador de carga a los objetivos.
ALB	LCUs consumidas	El número de unidades de capacidad del equilibrador de carga (LCU) utilizadas por su equilibrador de carga.

Tema	Métrica	Descripción
ALB	HTTPCode_ELB_4XX_Count HTTPCode_ELB_5XX_Count	El número de códigos de error de cliente HTTP 4xx o 5xx generados por el equilibrador de carga.
ALB	Nuevo recuento de conexiones	El número total de nuevas conexiones TCP establecidas desde los clientes al equilibrador de carga, y desde el equilibrador de carga a los objetivos.
ALB	Bytes procesados	El número total de bytes procesados por el equilibrador de carga.
ALB	Recuento de conexiones rechazadas	El número de conexiones que fueron rechazadas porque el equilibrador de carga había alcanzado su número máximo de conexiones.
ALB	Recuento de solicitudes	El número de solicitudes procesadas.
ALB	Recuento de errores de conexión del objetivo	El número de conexiones que no se han establecido con éxito entre el equilibrador de carga y el objetivo.
ALB	Tiempo de respuesta objetivo	El tiempo transcurrido, en segundos, desde que la solicitud salió del equilibrador de carga hasta que se recibió una respuesta del destino.
ALB	Recuento de huéspedes insalubres	El número de objetivos que se consideran poco saludables.
NLB	Recuento del flujo activo	El número total de flujos TCP concurrentes (o conexiones) desde los clientes a los objetivos.
NLB	LCUs consumidas	El número de unidades de capacidad del equilibrador de carga (LCU) utilizadas por su equilibrador de carga.
NLB	Nuevo recuento de flujo	El número total de nuevos flujos TCP (o conexiones) establecidos desde los clientes a los objetivos en el período de tiempo.
NLB	Bytes procesados	El número total de bytes procesados por el equilibrador de carga, incluyendo las cabeceras TCP/IP.

Escala automática	Tamaño máximo del grupo	El tamaño máximo del grupo de autoescalado
Amazon EC2	Utilización de la CPU	El porcentaje de unidades de computación EC2 asignadas que están actualmente en uso.
Amazon EC2	Red In	El número de bytes recibidos por la instancia en todas las interfaces de red.

Tabla 3: Métricas recomendadas de Amazon CloudWatch.

Para obtener más información sobre el uso de Amazon CloudWatch para detectar ataques DDoS en su aplicación, consulte Introducción a Amazon CloudWatch.

AWS incluye varias métricas y alarmas adicionales para notificarle un ataque y ayudarle a supervisar los recursos de su aplicación. La consola o la API de AWS Shield proporcionan un resumen y detalles sobre los ataques que se han detectado. Además, el panel del entorno de amenazas global proporciona información resumida sobre todos los ataques DDoS que han sido detectados por AWS. Esto puede ser útil para comprender mejor las amenazas DDoS en una población más amplia de aplicaciones, entender las tendencias de los ataques y compararlos con los ataques que pueda haber observado.

Otra herramienta que puede ayudarle a obtener visibilidad sobre el tráfico que se dirige a su aplicación es VPC Flow Logs. En una red tradicional, puede utilizar los registros de flujo de red para solucionar problemas de conectividad y seguridad, y para asegurarse de que las reglas de acceso a la red funcionan como se espera. Al utilizar los registros de flujo de la VPC, puede capturar información sobre el tráfico IP que va hacia y desde las interfaces de red en su VPC.

Cada registro de flujo incluye lo siguiente: direcciones IP de origen y destino, puertos de origen y destino, protocolo y número de paquetes y bytes transferidos durante la ventana de captura. Puede utilizar esta información para ayudar a identificar anomalías en el tráfico de la red y para identificar un vector de ataque específico. Por ejemplo, la mayoría de los ataques de reflexión UDP tienen puertos de origen específicos, como el puerto de origen 53 para la reflexión DNS. Esta es una firma clara que se puede identificar en el registro de flujo. En respuesta, puede optar por bloquear el puerto de origen específico a nivel de instancia o crear una regla NACL para bloquear todo el protocolo si su aplicación
no lo requiere.

Para obtener más información sobre el uso de los registros de flujo de la VPC para identificar anomalías en la red y vectores de ataque DDoS, consulte Registros de flujo de la VPC y Registros de flujo de la VPC: registre y vea los flujos de tráfico de la red.

Soporte

Es importante crear un plan de respuesta para los ataques DDoS antes de un evento real. Las mejores prácticas descritas en este documento están destinadas a ser medidas proactivas que se implementan antes de lanzar una aplicación, pero los ataques DDoS contra su aplicación aún podrían ocurrir. Revise las opciones en esta sección para determinar los recursos de soporte que son mejores para su escenario. Su equipo de cuentas puede evaluar su caso de uso y su aplicación, y ayudarle con preguntas o retos específicos que tenga.

Si ejecuta cargas de trabajo de producción en AWS, considere la posibilidad de suscribirse a Business Support, que le proporciona acceso 24 horas al día, 7 días a la semana, a ingenieros de Cloud Support que pueden ayudarle con los problemas de ataques DDoS. Si ejecuta cargas de trabajo de misión crítica, considere la posibilidad de suscribirse a Enterprise Support, que le ofrece la posibilidad de abrir casos "críticos" y recibir la respuesta más rápida de un
Ingeniero superior de soporte en la nube.

Si está suscrito a AWS Shield Advanced y también está suscrito a Business Support o Enterprise Support, puede escalar al equipo de respuesta a DDoS de AWS (DRT) si tiene un evento relacionado con DDoS que afecta a la disponibilidad de su aplicación. Si la capacidad de respuesta de su aplicación se degrada debido a un ataque DDoS, puede ponerse en contacto en directo con AWS Support. Otra opción es utilizar la función AWS Shield Engagement Lambda para iniciar más rápidamente el contacto con DRT. Por ejemplo, puede utilizar un botón de AWS IoT para activar la función AWS Lambda si tiene una situación de emergencia. Al pulsar el botón, se abre automáticamente un caso con AWS Support y se notifica inmediatamente a DRT. Recibe una respuesta directa para su caso que incluye un puente de conferencia de Amazon Chime al que puede unirse para interactuar con AWS Support y DRT. AWS Shield Engagement Lambda puede utilizarse con cualquier activador admitido por AWS Lambda.

Para obtener más información sobre el compromiso rápido de DRT mediante una función de AWS Lambda, consulte AWS Shield Engagement Lambda.

DRT no suele tener acceso a su cuenta de AWS ni a las solicitudes de muestra de AWS WAF. Puede autorizar a DRT para que acceda a AWS WAF, AWS Shield y a las operaciones relacionadas con la API en su cuenta desde la consola o la API de AWS Shield. Por ejemplo, es posible que quiera permitir a DRT ver sus Sampled Requests o colocar reglas para ayudar a mitigar un ataque DDoS en la capa de aplicación. También puede autorizar a DRT a acceder a los buckets de Amazon S3 que usted especifique. Por ejemplo, puede tener un bucket en el que almacena registros de solicitudes web y le gustaría que DRT tuviera acceso a ellos para analizarlos durante un ataque. DRT sólo accederá a su cuenta o realizará cambios durante un evento escalado y cualquier cambio estará sujeto a su consentimiento. Para obtener más información sobre cómo conceder acceso limitado a la cuenta a DRT, consulte Autorizar al Equipo de Respuesta DDoS.

En algunos casos, DRT puede enterarse de un ataque DDoS e involucrarlo proactivamente. Si hay puntos de contacto específicos que deben participar durante la escalada impulsada por DRT, puede añadirlos en la consola de AWS Shield haciendo clic en "Resumen", seguido de "Editar" en la sección
Sección "Contactos adicionales".

Conclusión:

Las mejores prácticas descritas en este documento pueden ayudarle a construir una arquitectura resistente a los ataques DDoS que puede proteger la disponibilidad de su aplicación mediante la prevención de muchos ataques DDoS comunes de la infraestructura y la capa de aplicación. La medida en que siga estas prácticas recomendadas al diseñar su aplicación influirá en el tipo, el vector y el volumen de los ataques DDoS que puede mitigar. Puede incorporar la resistencia sin suscribirse a un servicio de mitigación de DDoS. O, opcionalmente, puede optar por suscribirse a AWS Shield Advanced para obtener características adicionales de soporte, visibilidad, mitigación y protección de costes que protejan aún más una arquitectura de aplicaciones ya resistente.

Para obtener más información sobre la mitigación de DDoS y las mejores prácticas de resiliencia de DDoS en AWS, consulte el Apéndice A: Recursos adicionales.

Colaboradores

Las siguientes personas y organizaciones han contribuido a este documento:

- Andrew Kiggins, arquitecto de soluciones de AWS
- Jeffrey Lyon, Protección del perímetro de AWS
- Achraf Souk, arquitecto de soluciones de AWS
- Tino Tran, arquitecto de soluciones de AWS
- Yoshihisa Nakatani, arquitecto de soluciones de AWS

Revisión de documentos

Fecha	Descripción
Junio de 2018	Actualizado para incluir AWS Shield, las características de AWS WAF, AWS Firewall Manager y las mejores prácticas relacionadas.
Junio de 2016	Se ha añadido una guía de arquitectura prescriptiva y se ha actualizado para incluir AWS WAF
Junio de 2015	Primera publicación.

Apéndice A: Recursos adicionales

Puede utilizar los siguientes recursos para obtener más información sobre la mitigación de DDoS y las mejores prácticas de resistencia a DDoS con AWS:

- Mejores prácticas para la mitigación de DDoS en AWS
- SID216 - re:Invent 2017: Cloud-Native App Protection: Seguridad de las aplicaciones web en Pearson
- SID324 - re:Invent 2017: Automatización de la respuesta a los DDoS en la nube
- CTD304 - re:Invent 2017: El viaje de Dow Jones y Wall Street Journal para gestionar picos de tráfico mientras se mitigan las amenazas de DDoS y de la capa de aplicación
- CTD310 - re:Invent 2017: Vivir al límite, ¡es más seguro de lo que crees! Construyendo Fuerte con Amazon CloudFront, AWS Shield y AWS WAF