AWS Livre blanc

AWS Meilleures pratiques en matière de DDoS résilience

AWS Meilleures pratiques en matière de DDoS résilience: AWS Livre blanc

Table of Contents

AWS Meilleures pratiques en matière de DDoS résilience

Date de publication : 9 août 2023 (Révisions du document)

Il est important de protéger votre entreprise de l'impact des attaques par déni de service distribué (DDoS), ainsi que d'autres cyberattaques. Garder la confiance des clients dans votre service en maintenant la disponibilité et la réactivité de votre application est une priorité absolue. Vous souhaitez également éviter des coûts directs inutiles lorsque votre infrastructure doit évoluer en réponse à une attaque. Amazon Web Services (AWS) s'engage à vous fournir les outils, les meilleures pratiques et les services nécessaires pour vous défendre contre les acteurs malveillants sur Internet. L'utilisation des services appropriés AWS permet de garantir une disponibilité, une sécurité et une résilience élevées.

Dans ce livre blanc, vous trouverez AWS des DDoS conseils prescriptifs pour améliorer la résilience des applications exécutées sur. AWS Cela inclut une architecture de référence DDoS résiliente qui peut être utilisée comme guide pour protéger la disponibilité des applications. Ce livre blanc décrit également différents types d'attaques, tels que les attaques au niveau de l'infrastructure et les attaques au niveau de la couche application. AWS explique quelles sont les meilleures pratiques les plus efficaces pour gérer chaque type d'attaque. En outre, les services et fonctionnalités qui s'inscrivent dans une stratégie DDoS d'atténuation sont décrits, ainsi que la manière dont chacun peut être utilisé pour protéger vos applications.

Ce paper est destiné aux décideurs informatiques et aux ingénieurs de sécurité familiarisés avec les concepts de base du réseau, de la sécurité et AWS. Chaque section contient des liens vers une AWS documentation qui fournit plus de détails sur les meilleures pratiques ou capacités.

AWS détecte plus d'un million d'DDoSattaques par an et en atténue des milliers chaque jour contre nos clients. Selon notre équipe du Shield Response (SRT), la majorité des clients victimes d'DDoSattaques n'ont pas mis en œuvre les recommandations de ce guide.

Êtes-vous Well-Architected ?

Le AWS Well-Architected Framework vous aide à comprendre les avantages et les inconvénients des décisions que vous prenez lors de la création de systèmes dans le cloud. Les six piliers du Framework vous permettent d'apprendre les meilleures pratiques architecturales pour concevoir et exploiter des systèmes fiables, sécurisés, efficaces, rentables et durables. À l'aide du AWS Well-Architected Tool, disponible gratuitement dans le AWS Management Console(connexion requise),

vous pouvez évaluer votre charge de travail par rapport à ces meilleures pratiques en répondant à une série de questions pour chaque pilier.

Pour obtenir des conseils d'experts supplémentaires et les meilleures pratiques relatives à votre architecture cloud (déploiements d'architecture de référence, diagrammes et livres blancs), consultez le Centre d'architecture.AWS

Présentation des attaques par déni de service

Une attaque ou un événement par déni de service (DoS) est une tentative délibérée de rendre un site Web ou une application indisponible pour les utilisateurs, par exemple en l'inondant de trafic réseau. Les attaquants utilisent diverses techniques qui consomment de grandes quantités de bande passante réseau ou monopolisent d'autres ressources du système, perturbant ainsi l'accès des utilisateurs légitimes. Dans sa forme la plus simple, un attaquant isolé utilise une source unique pour exécuter une attaque DoS contre une cible, comme le montre la figure suivante.

Schéma illustrant une attaque DoS

Lors d'une attaque par déni de service distribué (DDoS), un attaquant utilise plusieurs sources pour orchestrer une attaque contre une cible. Ces sources peuvent inclure des groupes distribués d'ordinateurs, de routeurs, d'appareils IoT et d'autres terminaux infectés par des logiciels malveillants. La figure suivante montre un réseau d'hôtes compromis qui participent à l'attaque, générant un flot de paquets ou de demandes visant à submerger la cible.

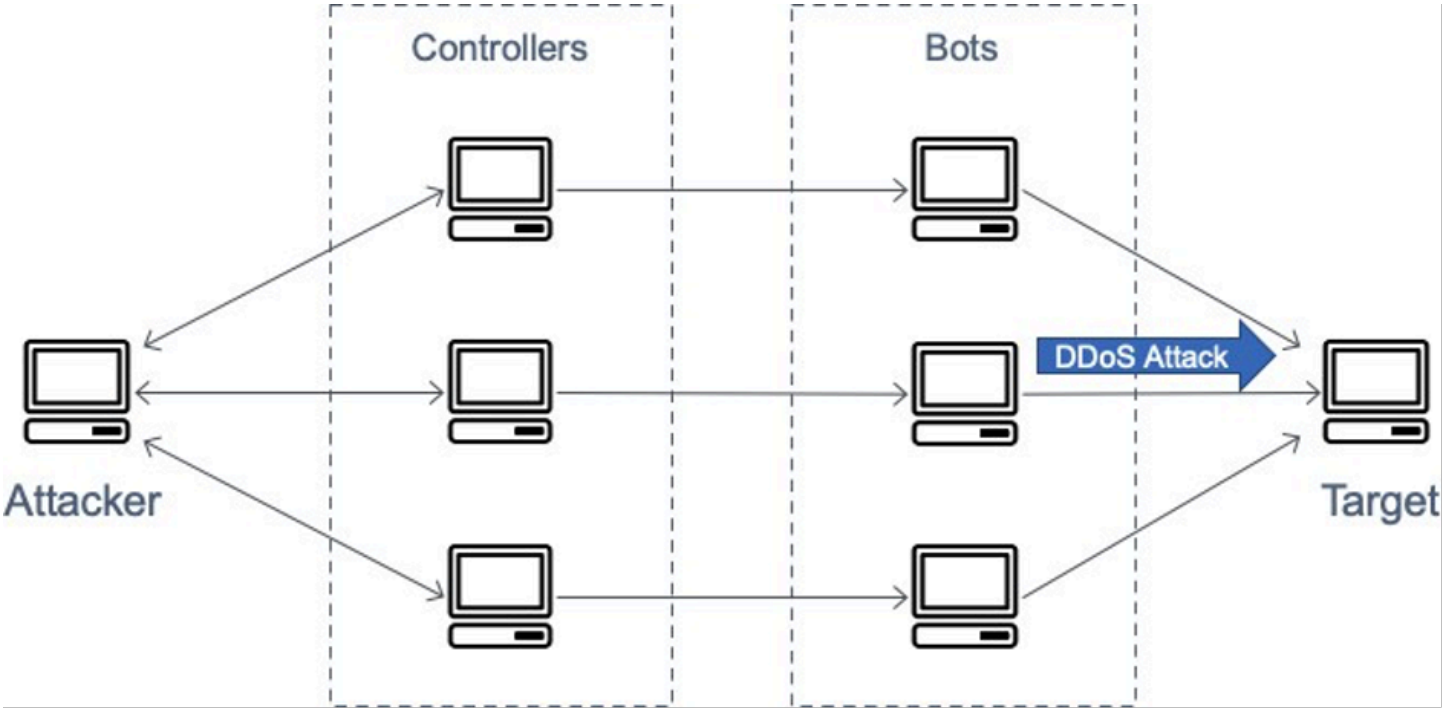


Schéma illustrant une DDoS attaque

Le modèle Open Systems Interconnection (OSI) comporte sept couches, décrites dans le tableau suivant. DDoSles attaques sont plus fréquentes aux couches 3, 4, 6 et 7.

- Les attaques de couches 3 et 4 correspondent aux couches réseau et transport du OSI modèle. Dans ce livre blanc, nous les AWS désignons collectivement sous le terme « attaques de la couche d'infrastructure ».
- Les attaques des couches 6 et 7 correspondent aux couches Présentation et Application du OSI modèle. Ce livre blanc les aborde ensemble en tant qu'attaques de la couche applicative.

Ce paper traite de ces types d'attaques dans les sections qui suivent.

Tableau 1 — OSI modèle

#	Couche	Unité	Description	Exemples vectoriels
7	Application	Données	Processus du réseau jusqu'à la demande	HTTPinondations, inondations DNS de requêtes
6	Présentation	Données	Représentation et chiffrement des données	Abus de sécurité de la couche de transport (TLS)
5	Session	Données	Communication entre hôtes	N/A
4	Transport	Segments	nd-to-end Connexions E et fiabilité	Synchroniser (SYN) les inondations
3	Réseau	Paquets	Détermination du chemin et adressage logique	Attaques par réflexion via le protocole

#	Couche	Unité	Description	Exemples vectoriels
				User Datagram Protocol (UDP)
2	Liaison de données	Frames (Images)	Adressage physique	N/A
1	Physique	Bits	Transmission multimédia, signal et binaire	N/A

Attaques contre la couche d'infrastructure

Les attaques les plus courantes, à savoir DDoS les attaques par réflexion du protocole User Datagram (UDP) et SYN les inondations, concernent la couche d'infrastructure. Un attaquant peut utiliser l'une ou l'autre de ces méthodes pour générer de gros volumes de trafic susceptibles d'inonder la capacité d'un réseau ou d'immobiliser des ressources sur des systèmes tels que des serveurs, des pare-feux, un système de prévention des intrusions (IPS) ou un équilibreur de charge. Bien que ces attaques puissent être faciles à identifier, pour les atténuer efficacement, vous devez disposer d'un réseau ou de systèmes capables d'augmenter la capacité plus rapidement que le flot de trafic entrant. Cette capacité supplémentaire est nécessaire pour filtrer ou absorber le trafic d'attaque, libérant ainsi le système et l'application pour répondre au trafic légitime des clients.

UDPattaques par réflexion

UDPles attaques par réflexion exploitent le fait qu'il UDP s'agit d'un protocole sans état. Les attaquants peuvent créer un paquet de UDP requête valide répertoriant l'adresse IP de la cible de l'attaque comme adresse IP UDP source. L'attaquant a désormais falsifié (usurpé) l'adresse IP source du paquet de requêtes. UDP Le UDP paquet contient l'adresse IP source falsifiée et est envoyé par l'attaquant à un serveur intermédiaire. Le serveur est incité à envoyer ses paquets de UDP réponse à l'adresse IP de la victime ciblée plutôt qu'à l'adresse IP de l'attaquant. Le serveur intermédiaire est utilisé car il génère une réponse plusieurs fois supérieure au paquet de demande, amplifiant ainsi efficacement le volume de trafic d'attaque envoyé à l'adresse IP cible.

Le facteur d'amplification est le rapport entre la taille de la réponse et la taille de la demande, et il varie en fonction du protocole utilisé par l'attaquant : DNS Network Time Protocol (NTP), Simple

Service Directory Protocol (SSDP), Connectionless Lightweight Directory Access Protocol (CLDAP), Memcached, Character Generator Protocol (CharGen) ou Quote of the Day (). QOTD

Par exemple, le facteur d'amplification pour DNS peut être de 28 à 54 fois le nombre d'octets d'origine. Ainsi, si un attaquant envoie une charge utile de requête de 64 octets à un DNS serveur, il peut générer plus de 3 400 octets de trafic indésirable vers une cible d'attaque. UDPles attaques par réflexion sont responsables d'un volume de trafic plus important que les autres attaques. La figure suivante illustre la tactique de réflexion et l'effet d'amplification.

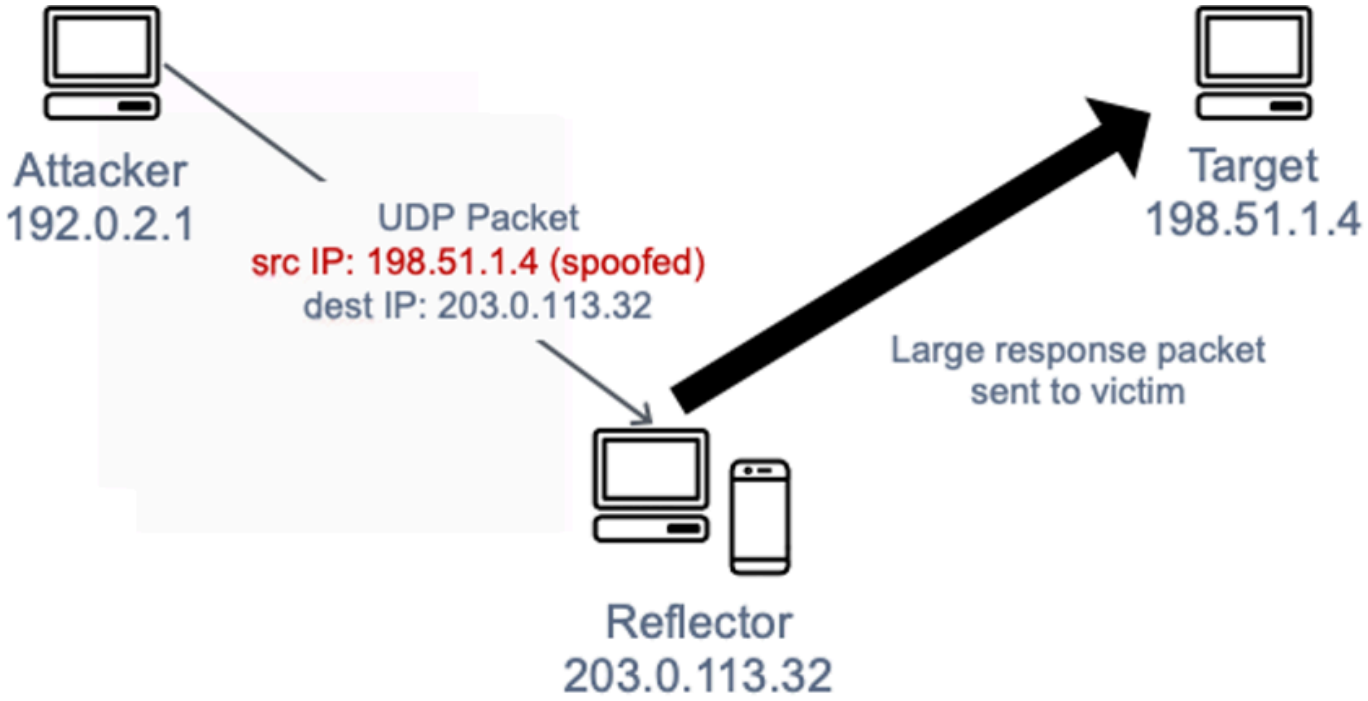


Schéma illustrant une attaque par UDP réflexion

Il convient de noter que les attaques par réflexion, bien qu'elles fournissent aux attaquants une amplification « gratuite », nécessitent une fonction d'usurpation d'adresse IP et, à mesure que de plus en plus de fournisseurs de réseaux adoptent la validation des adresses source partout (SAVE) ou que cette fonctionnalité est supprimée BCP38, obligeant les fournisseurs de DDoS services à mettre fin aux attaques par réflexion ou à se relocaliser vers des centres de données et des fournisseurs de réseaux qui ne mettent pas en œuvre la validation de l'adresse source.

SYNattaques liées aux inondations

Lorsqu'un utilisateur se connecte à un service Transmission Control Protocol (TCP), tel qu'un serveur Web, son client envoie un SYN paquet. Le serveur renvoie un paquet d'accusé de réception de synchronisation (SYN-ACK), et finalement le client répond par un paquet d'accusé de réception

(ACK), qui complète la triple poignée de main attendue. L'image suivante illustre cette poignée de main typique.

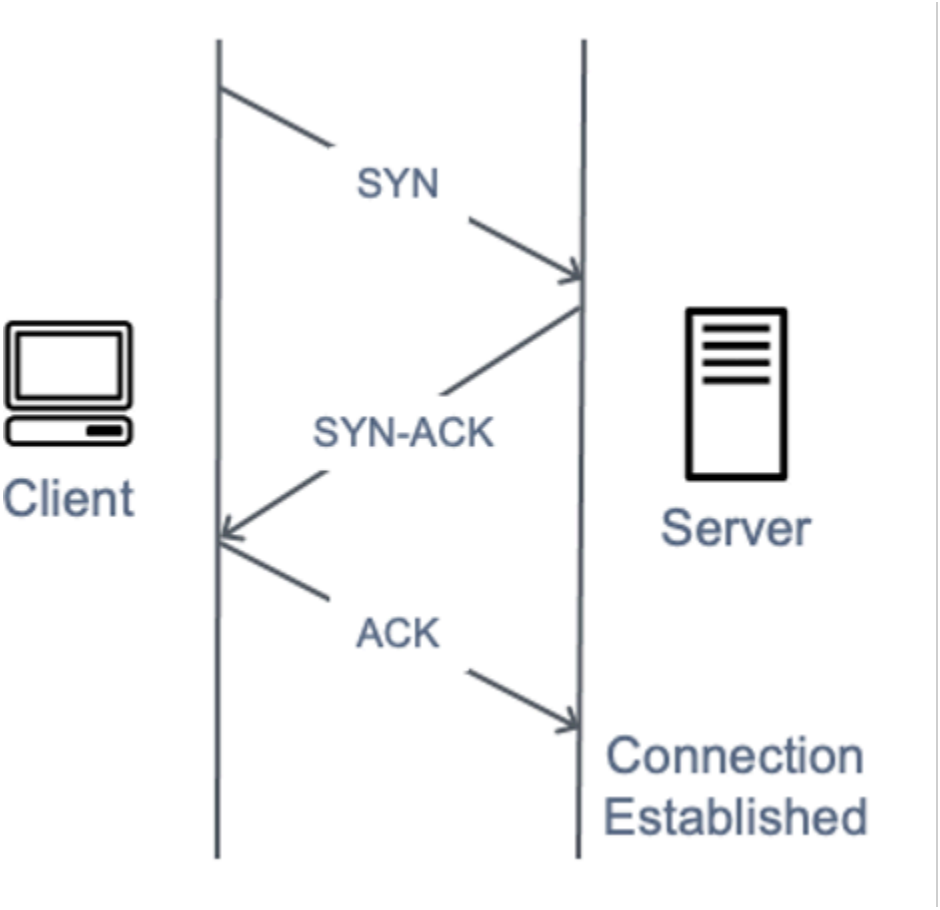


Schéma illustrant une poignée de SYN main à trois

Lors d'une attaque par SYN inondation, un client malveillant envoie un grand nombre de SYN paquets, mais n'envoie jamais les derniers ACK paquets pour terminer les poignées de main. Le serveur est laissé dans l'attente d'une réponse aux TCP connexions semi-ouvertes et l'idée est que la cible finira par manquer de capacité pour accepter de nouvelles TCP connexions, ce qui empêche les nouveaux utilisateurs de se connecter au serveur, mais l'impact réel est plus nuancé. Les systèmes d'exploitation modernes implémentent tous SYN des cookies par défaut comme mécanisme pour empêcher l'épuisement des tables d'états suite à des attaques par SYN inondation. Une fois que la longueur de la SYN file d'attente atteint un seuil prédéterminé, le serveur répond par un SYN - ACK contenant un numéro de séquence initial contrefait, sans créer d'entrée dans sa SYN file d'attente. Si le serveur reçoit ensuite un accusé de réception ACK contenant un numéro d'accusé de réception correctement incrémenté, il est en mesure d'ajouter l'entrée à sa table d'état et de procéder normalement. L'impact réel des SYN inondations sur les équipements cibles est généralement lié à la capacité et à CPU l'épuisement du réseau, mais les périphériques intermédiaires tels que les pare-

feux (ou le suivi des connexions des groupes de EC2 sécurité) peuvent être épuisés dans la table des TCP états et interrompre les nouvelles connexions.

TCPréflexion de la boîte intermédiaire

Ce vecteur d'attaque relativement nouveau a été révélé pour la première fois dans un livre blanc universitaire en août 2021, qui expliquait comment le TCP non-respect des pare-feux des États et du commerce pouvait amener ces derniers à devenir des vecteurs d'amplification. TCP Nous avons été témoins de ces attaques « dans la nature » depuis le début de 2022 et nous continuons d'en être témoins aujourd'hui. Le facteur d'amplification varie en fonction des différentes manières dont les fournisseurs ont implémenté cette « fonctionnalité », mais il peut dépasser l'amplification MemcachedUDP.

Attaques contre la couche applicative

Un attaquant peut cibler l'application elle-même en utilisant une attaque de couche 7 ou de couche d'application. Dans le cadre de ces attaques, similaires aux attaques d'infrastructure par SYN inondation, l'attaquant tente de surcharger des fonctions spécifiques d'une application afin de rendre l'application indisponible ou de ne plus répondre aux utilisateurs légitimes. Cela peut parfois être réalisé avec de très faibles volumes de demandes qui ne génèrent qu'un faible volume de trafic réseau. Cela peut rendre l'attaque difficile à détecter et à atténuer. Parmi les exemples d'attaques contre la couche applicative, on peut citer HTTP les inondations, les attaques par contournement du cache et WordPress XML les inondations. RPC

- Lors d'une attaque d'HTTPinondation, un attaquant envoie des HTTP demandes qui semblent provenir d'un utilisateur valide de l'application Web. Certaines HTTP inondations ciblent une ressource spécifique, tandis que HTTP les inondations plus complexes tentent d'imiter l'interaction humaine avec l'application. Cela peut accroître la difficulté d'utiliser des techniques d'atténuation courantes telles que la limitation du taux de demandes.
- Les attaques par contournement du cache sont un type d'HTTPinondation qui utilise des variations de la chaîne de requête pour contourner la mise en cache du réseau de diffusion de contenu (). CDN Au lieu de pouvoir renvoyer les résultats mis en cache, ils CDN doivent contacter le serveur d'origine pour chaque demande de page, et ces extractions d'origine entraînent une charge supplémentaire sur le serveur Web de l'application.
- Lors d'une WordPress XMLattaque d'RPCinondation, également connue sous le nom de « WordPress pingback flood », un attaquant cible un site Web hébergé sur le logiciel de gestion de WordPress contenu. L'attaquant utilise à mauvais escient la RPC API fonction XML- pour générer

un flot de HTTP requêtes. La fonction de ping back permet à un site Web hébergé sur WordPress (site A) d'avertir un autre WordPress site (site B) par le biais d'un lien créé par le site A vers le site B. Le site B tente ensuite de récupérer le site A pour vérifier l'existence du lien. Lors d'une attaque par pingback, l'attaquant abuse de cette fonctionnalité pour amener le site B à attaquer le site A. Ce type d'attaque possède une signature claire : « `WordPress:` » est généralement présent dans l'agent utilisateur de l'en-tête de la HTTP demande.

Il existe d'autres formes de trafic malveillant qui peuvent avoir un impact sur la disponibilité d'une application. Les robots Scraper automatisent les tentatives d'accès à une application Web pour voler du contenu ou enregistrer des informations concurrentielles, telles que les prix. Les attaques par force brute et par « credential stuffing » sont des actions programmées visant à obtenir un accès non autorisé à des zones sécurisées d'une application. Il ne s'agit pas d'DDoSattaques à proprement parler, mais leur nature automatisée peut ressembler à une DDoS attaque et elles peuvent être atténuées en mettant en œuvre certaines des meilleures pratiques décrites dans ce paper.

Les attaques au niveau de la couche applicative peuvent également cibler les services du système de noms de domaine (DNS). La plus courante de ces attaques est un flot de DNS requêtes dans lequel un attaquant utilise de nombreuses DNS requêtes bien formées pour épuiser les ressources d'un DNS serveur. Ces attaques peuvent également inclure un composant de destruction du cache dans lequel l'attaquant répartit aléatoirement la chaîne de sous-domaine pour contourner le DNS cache local d'un résolveur donné. Par conséquent, le résolveur ne peut pas tirer parti des requêtes de domaine mises en cache et doit contacter à plusieurs reprises le DNS serveur faisant autorité, ce qui amplifie l'attaque.

Si une application Web est fournie via Transport Layer Security (TLS), un attaquant peut également choisir d'attaquer le processus de TLS négociation. TLSétant coûteuse en termes de calcul, un attaquant, en générant une charge de travail supplémentaire sur le serveur pour traiter des données illisibles (ou incompréhensibles (texte chiffré)) dans le cadre d'une poignée de main légitime, peut réduire la disponibilité du serveur. Dans une variante de cette attaque, un attaquant termine la TLS poignée de main mais renégocie perpétuellement la méthode de chiffrement. Un attaquant peut également tenter d'épuiser les ressources du serveur en ouvrant et en fermant de nombreuses TLS sessions.

Techniques d'atténuation

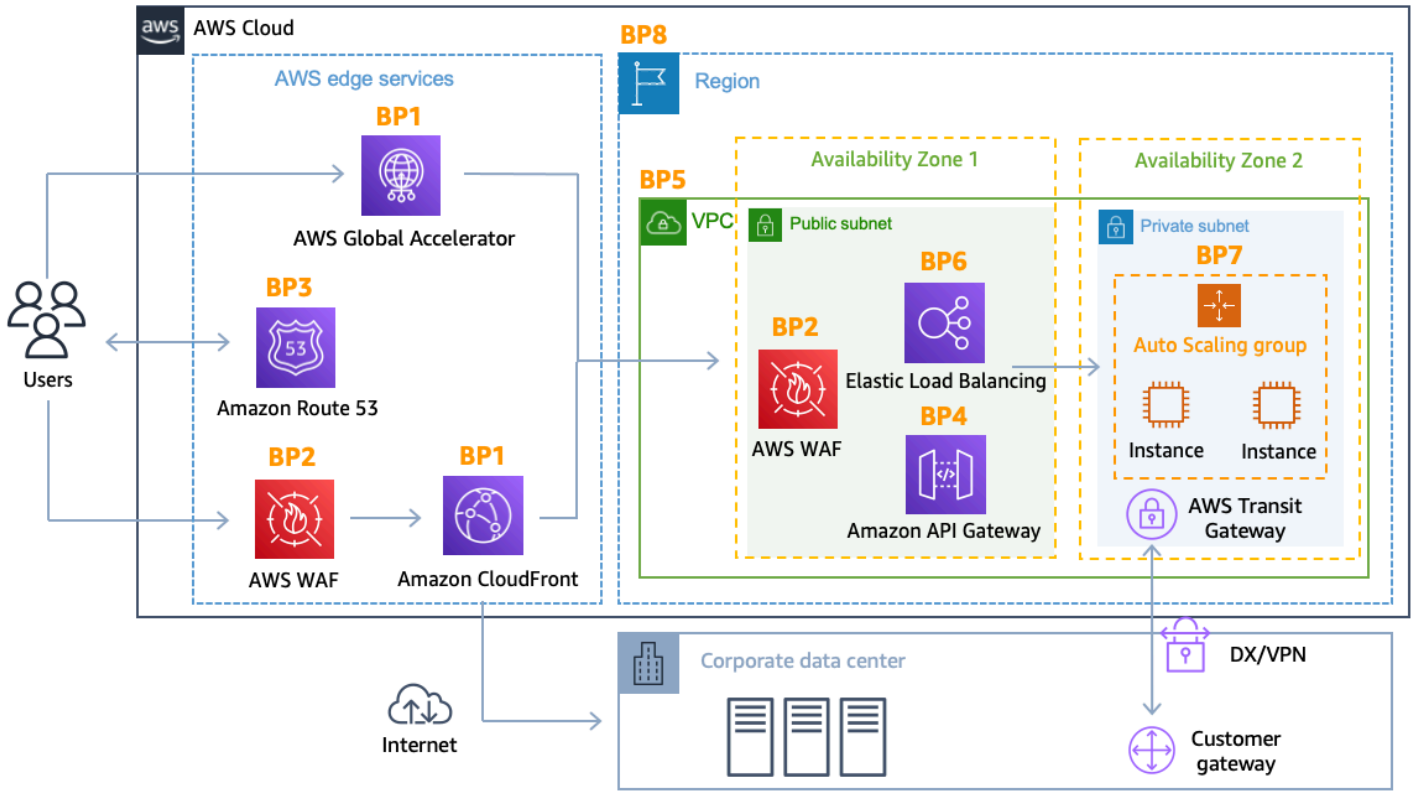
Certaines formes d'DDoSatténuation sont incluses automatiquement dans les AWS services. DDoSla résilience peut être encore améliorée en utilisant une AWS architecture avec des services spécifiques, décrits dans les sections suivantes, et en mettant en œuvre des meilleures pratiques supplémentaires pour chaque partie du flux réseau entre les utilisateurs et votre application.

Vous pouvez utiliser AWS des services qui opèrent depuis des sites périphériques, tels qu'Amazon CloudFront, AWS Global Accelerator et Amazon Route 53, pour créer une protection complète de la disponibilité contre toutes les attaques connues au niveau de l'infrastructure. Ces services font partie du réseau AWS global Edge et peuvent améliorer la DDoS résilience de votre application lorsqu'elle dessert tout type de trafic applicatif à partir d'emplacements périphériques répartis dans le monde entier. Vous pouvez exécuter votre application dans n'importe quel Région AWS environnement et utiliser ces services pour protéger la disponibilité de votre application et optimiser les performances de votre application pour les utilisateurs finaux légitimes.

Les avantages liés à l'utilisation d'Amazon CloudFront, de Global Accelerator et d'Amazon Route 53 incluent :

- Accès à Internet et capacité DDoS d'atténuation sur le réseau AWS Global Edge. Cela est utile pour atténuer les attaques volumétriques de plus grande envergure, qui peuvent atteindre l'échelle du térabit.
- AWS Shield DDoSles systèmes d'atténuation sont intégrés aux services de AWS pointe, ce qui permet time-to-mitigate de passer de quelques minutes à moins d'une seconde.
- Stateless SYN Flood mitigation vérifie les connexions entrantes à l'aide de SYN cookies avant de les transmettre au service protégé. Cela garantit que seules les connexions valides atteignent votre application tout en protégeant vos utilisateurs finaux légitimes contre les pertes de faux positifs.
- Systèmes d'ingénierie automatique du trafic qui dispersent ou isolent l'impact des attaques volumétriques DDoS de grande envergure. Tous ces services isolent les attaques à la source avant qu'elles n'atteignent votre source, ce qui réduit l'impact sur les systèmes protégés par ces services.
- La défense de la couche applicative, CloudFront lorsqu'elle AWS WAFest combinée à cela, ne nécessite pas de modification de l'architecture applicative actuelle (par exemple, dans un centre de données Région AWS ou sur site).

Le transfert de données entrantes est gratuit AWS et vous ne payez pas pour le trafic DDoS d'attaque atténué par. AWS Shield Le schéma d'architecture suivant inclut les services du réseau AWS Global Edge.



DDoS-architecture de référence résiliente

Cette architecture inclut plusieurs AWS services qui peuvent vous aider à améliorer la résilience de votre application Web face aux DDoS attaques. Le tableau suivant fournit un résumé de ces services et des fonctionnalités qu'ils peuvent fournir. AWS a associé à chaque service un indicateur des meilleures pratiques (BP1,BP2) pour faciliter la consultation dans ce document. Par exemple, une prochaine section traite des fonctionnalités fournies par Amazon CloudFront et Global Accelerator, y compris l'indicateur des meilleures pratiquesBP1.

Tableau 2 - Résumé des meilleures pratiques

	AWS Edge			Région AWS		
	Utiliser Amazon CloudFront	Utilisation de Global	Utilisation d'Amazon	Utiliser Elastic Load	Utilisation des groupes de	Utilisation d'Amazon Elastic

	AWS Edge			Région AWS		
	(BP1) avec AWS WAF (BP2)	Accelerator (BP1)	Route 53 (BP3)	Balancing (BP6) avec AWS WAF (BP2)	sécurité et du réseau ACLs dans Amazon VPC (BP5)	Compute Cloud (AmazonEC2) Auto Scaling (BP7)
Atténuation des attaques de couche 3 (par exemple, UDP réflexion)	✓	✓	✓	✓	✓	✓
Atténuation des attaques de niveau 4 (par exemple, SYN inondation)	✓	✓	✓	✓		
Atténuation des attaques de couche 6 (par exempleTLS)	✓	✓	✓	✓		
Réduire la surface d'attaque	✓	✓	✓	✓	✓	

	AWS Edge			Région AWS		
Évoluez pour absorber le trafic de la couche applicative	✓	✓	✓	✓	✓	✓
Atténuation des attaques de couche 7 (couche d'application)	✓	✓(*)	✓	✓	✓(*)	✓(*)
Isolement géographique et dispersion du trafic excédentaire et des DDoS attaques de plus grande envergure	✓	✓	✓			

✓ (*) : Si utilisé AWS WAF avec Application Load Balancer

Vous pouvez également vous préparer à répondre aux DDoS attaques et à les atténuer en vous abonnant à AWS Shield Advanced. Les avantages de l'utilisation AWS Shield Advanced incluent :

- Accès à un support spécialisé 24 h/24 et 7 j/7 de la part de l'équipe d'AWS Shield intervention (AWS SRT) pour vous aider à atténuer les DDoS attaques qui ont un impact sur la disponibilité des applications, y compris une fonctionnalité d'engagement proactif en option
- Seuils de détection sensibles qui acheminent le trafic vers le système DDoS d'atténuation plus tôt et peuvent améliorer time-to-mitigate les attaques contre Amazon EC2 (y compris Elastic Load Balancer) ou Network Load Balancer, lorsqu'ils sont utilisés avec une adresse IP élastique
- Détection personnalisée de couche 7 basée sur les modèles de trafic de référence de votre application lorsqu'elle est utilisée avec AWS WAF
- DDoSAtténuation automatique de la couche applicative grâce à laquelle Shield Advanced répond aux DDoS attaques détectées en créant, en évaluant et en déployant des AWS WAF règles personnalisées
- Accès sans AWS WAF frais supplémentaires pour atténuer les DDoS attaques de la couche application (en cas d'utilisation avec Amazon CloudFront ou Application Load Balancer)
- Gestion centralisée des politiques de AWS Firewall Managersécurité sans frais supplémentaires.
- Protection des coûts qui vous permet de demander un remboursement limité des coûts liés à la mise à l'échelle résultant d'une DDoS attaque.
- Contrat de niveau de service amélioré spécifique aux AWS Shield Advanced clients.
- Des groupes de protection qui vous permettent de regrouper des ressources, offrant ainsi un moyen en libre-service de personnaliser l'étendue de la détection et de l'atténuation pour votre application en traitant plusieurs ressources comme une seule unité. Pour plus d'informations sur les groupes de protection, reportez-vous à la section Groupes de protection Shield Advanced.
- DDoSvisibilité des attaques en utilisant les CloudWatch métriques et les alarmes AWS Management ConsoleAPI,, et Amazon.

Ce service DDoS d'atténuation optionnel permet de protéger les applications hébergées sur n'importe quel site Région AWS. Le service est disponible dans le monde entier pour CloudFront Route 53 et Global Accelerator. Au niveau régional, vous pouvez protéger les adresses IP Application Load Balancer, Classic Load Balancer et Elastic, ce qui vous permet de protéger les instances Network Load Balancer () ou NLBs Amazon. EC2

Pour une liste complète des AWS Shield Advanced fonctionnalités et pour plus d'informations à ce sujet AWS Shield, reportez-vous à la section AWS Shield Fonctionnement.

Bonnes pratiques en matière DDoS d'atténuation

Dans les sections suivantes, chacune des meilleures pratiques recommandées en matière DDoS d'atténuation est décrite plus en détail. Pour un easy-to-implement guide rapide sur la création DDoS d'une couche d'atténuation pour les applications Web statiques ou dynamiques, consultez Comment protéger les applications Web dynamiques contre les DDoS attaques à l'aide d'Amazon CloudFront et d'Amazon Route 53.

Défense de la couche d'infrastructure (BP1BP3,BP6,,BP7)

Dans un environnement de centre de données traditionnel, vous pouvez atténuer les DDoS attaques au niveau de l'infrastructure en utilisant des techniques telles que le surprovisionnement de la capacité, le déploiement de systèmes DDoS d'atténuation ou le nettoyage du trafic à l'aide de services d'atténuation. DDoS Oui AWS, les fonctionnalités DDoS d'atténuation sont automatiquement fournies, mais vous pouvez optimiser la DDoS résilience de votre application en faisant des choix d'architecture qui tirent le meilleur parti de ces fonctionnalités et vous permettent également de vous adapter au trafic excédentaire.

Les principales considérations à prendre en compte pour atténuer les DDoS attaques volumétriques incluent la garantie d'une capacité de transit et d'une diversité suffisantes et la protection AWS des ressources, telles que les EC2 instances Amazon, contre le trafic d'attaque.

Certains types d'EC2instances Amazon prennent en charge des fonctionnalités permettant de gérer plus facilement de gros volumes de trafic, par exemple des interfaces de bande passante réseau allant jusqu'à 100 Gbit/s et une mise en réseau améliorée. Cela permet d'éviter la congestion de l'interface pour le trafic qui a atteint l'EC2instance Amazon. Les instances qui prennent en charge la mise en réseau améliorée offrent des performances d'entrée/sortie (E/S) supérieures, une bande passante plus élevée et une CPU utilisation plus faible par rapport aux implémentations traditionnelles. Cela améliore la capacité de l'instance à gérer de gros volumes de trafic et, en fin de compte, la rend très résiliente face à la charge de paquets par seconde (pps).

Pour permettre ce haut niveau de résilience, il est AWS recommandé d'utiliser des instances Amazon EC2 Dedicated, ou des EC2 instances Amazon avec un débit réseau plus élevé, dotées du suffixe N « » et prenant en charge la mise en réseau améliorée avec jusqu'à 100 Gbit/s de bande passante réseau, par exemple, `c5n.18xlarge` et/ou `c6gn.16xlarge` des instances métalliques (telles que). `c5n.metal`

Pour plus d'informations sur les EC2 instances Amazon qui prennent en charge les interfaces réseau 100 Gigabit et la mise en réseau améliorée, consultez la section Types d'EC2instances Amazon.

Le module requis pour une mise en réseau améliorée et le jeu d'`enaSupport`attributs requis sont inclus dans Amazon Linux 2 et les dernières versions d'Amazon LinuxAMI. Par conséquent, si vous lancez une instance avec une version matérielle de machine virtuelle (HVM) d'Amazon Linux sur un type d'instance pris en charge, la mise en réseau améliorée est déjà activée pour votre instance. Pour plus d'informations, reportez-vous aux sections Tester si la mise en réseau améliorée est activée et Mise en réseau améliorée sous Linux.

Amazon EC2 avec Auto Scaling (BP7)

Un autre moyen d'atténuer les attaques au niveau de l'infrastructure et de la couche applicative consiste à opérer à grande échelle. Si vous possédez des applications Web, vous pouvez utiliser des équilibreurs de charge pour distribuer le trafic vers un certain nombre d'EC2instances Amazon surapprovisionnées ou configurées pour évoluer automatiquement. Ces instances peuvent gérer des pics de trafic soudains qui se produisent pour quelque raison que ce soit, notamment un flash crowd ou une DDoS attaque au niveau de la couche applicative. Vous pouvez configurer les CloudWatch alarmes Amazon pour lancer Auto Scaling afin d'ajuster automatiquement la taille de votre EC2 flotte Amazon en réponse à des événements que vous définissez CPURAM, tels que les E/S réseau et même des métriques personnalisées.

Cette approche protège la disponibilité des applications en cas d'augmentation inattendue du volume de demandes. Lorsque vous utilisez Amazon CloudFront, Application Load Balancer, Classic Load Balancers ou Network Load Balancer avec votre applicationTLS, la négociation est gérée par la distribution (CloudFrontAmazon) ou par l'équilibreur de charge. Ces fonctionnalités aident à protéger vos instances contre l'impact d'attaques TLS basées en les adaptant pour gérer les demandes légitimes et les attaques TLS abusives.

Pour plus d'informations sur l'utilisation d'Amazon CloudWatch pour invoquer Auto Scaling, consultez la section Surveillance CloudWatch des métriques Amazon pour vos groupes et instances Auto Scaling.

Amazon EC2 fournit une capacité de calcul redimensionnable afin que vous puissiez rapidement augmenter ou diminuer en fonction de l'évolution des besoins. Vous pouvez effectuer une mise à l'échelle horizontale en ajoutant automatiquement des instances à votre application en redimensionnant la taille de votre groupe Amazon EC2 Auto Scaling, et vous pouvez effectuer une mise à l'échelle verticale en utilisant des types d'EC2instances plus grands.

En utilisant Amazon RDS Proxy, vous pouvez autoriser vos applications à regrouper et à partager des connexions de base de données afin d'améliorer leur capacité à évoluer et à gérer les pics imprévisibles du trafic de base de données. Vous pouvez également activer l'auto-scaling du

stockage pour une instance de base de RDS données Amazon. Consultez Gestion automatique de la capacité avec Amazon RDS Storage Autoscaling pour plus d'informations.

Elastic Load Balancing (BP6)

DDoSLes attaques de grande envergure peuvent dépasser la capacité d'une seule EC2 instance Amazon. Avec Elastic Load Balancing (ELB), vous pouvez réduire le risque de surcharge de votre application en répartissant le trafic entre de nombreuses instances de backend. Elastic Load Balancing peut évoluer automatiquement, ce qui vous permet de gérer des volumes plus importants lorsque vous êtes confronté à un trafic supplémentaire imprévu, par exemple en raison d'une affluence soudaine ou d'DDoSattaques. Pour les applications créées au sein d'un AmazonVPC, trois types sont ELBs à prendre en compte, en fonction de votre type d'application : Application Load Balancer (ALB), Network Load Balancer () et Classic Load Balancer NLB (). CLB

Pour les applications Web, vous pouvez utiliser l'Application Load Balancer pour acheminer le trafic en fonction du contenu et n'accepter que des requêtes Web bien formulées. Application Load Balancer bloque de nombreuses DDoS attaques courantes, telles que les attaques par SYN inondation ou par UDP réflexion, protégeant ainsi votre application contre ces attaques. Application Load Balancer s'adapte automatiquement pour absorber le trafic supplémentaire lorsque ces types d'attaques sont détectés. La mise à l'échelle des activités liées aux attaques visant la couche d'infrastructure est transparente pour AWS les clients et n'a aucune incidence sur votre facture.

Pour plus d'informations sur la protection des applications Web avec Application Load Balancer, reportez-vous à Getting Started with Application Load Balancers.

Pour les HTTPS applications HTTP autres que/, vous pouvez utiliser Network Load Balancer pour acheminer le trafic vers des cibles (par exemple, des EC2 instances Amazon) avec une latence extrêmement faible. L'une des principales considérations relatives à Network Load Balancer est que tout TCP SYN le UDP trafic qui atteint l'équilibreur de charge sur un écouteur valide sera acheminé vers vos cibles et non absorbé. Toutefois, cela ne s'applique pas aux TLS -listeners qui mettent fin à la connexion. TCP Pour les équilibreurs de charge réseau dotés d'TCPécouteurs, nous recommandons de déployer Global Accelerator pour les protéger contre SYN les inondations.

Vous pouvez utiliser Shield Advanced pour configurer DDoS la protection des adresses IP Elastic. Lorsqu'une adresse IP élastique est attribuée par zone de disponibilité au Network Load Balancer, Shield Advanced applique les DDoS protections appropriées au trafic du Network Load Balancer.

Pour plus d'informations sur la protection TCP et UDP les applications avec Network Load Balancer, reportez-vous à Getting started with Network Load Balancers.

Note

Selon la configuration du groupe de sécurité, la ressource utilisant le groupe de sécurité doit utiliser le suivi des connexions pour suivre les informations sur le trafic. Cela peut affecter la capacité de l'équilibreur de charge à traiter les nouvelles connexions, car le nombre de connexions suivies est limité.

Une configuration de groupe de sécurité qui contient une règle d'entrée acceptant le trafic provenant de n'importe quelle adresse IP (par exemple, `0.0.0.0/0` ou `::/0`), mais qui n'a pas de règle correspondante pour autoriser le trafic de réponse, amène le groupe de sécurité à utiliser les informations de suivi des connexions pour autoriser l'envoi du trafic de réponse. En cas d'DDoSattaque, le nombre maximum de connexions suivies peut être épuisé. Pour améliorer la DDoS résilience de votre Application Load Balancer ou Classic Load Balancer destiné au public, assurez-vous que le groupe de sécurité associé à votre équilibreur de charge est configuré pour ne pas utiliser le suivi des connexions (connexions non suivies), afin que le flux de trafic ne soit pas soumis aux limites de suivi des connexions.

Pour cela, configurez votre groupe de sécurité avec une règle qui autorise la règle entrante à accepter les TCP flux provenant de n'importe quelle adresse IP (`0.0.0.0/0`ou`::/0`), et ajoutez une règle correspondante dans le sens sortant permettant à cette ressource d'envoyer le trafic de réponse (autoriser la plage sortante pour n'importe quelle adresse IP `0.0.0.0/0` ou`::/0`) pour tous les ports (0-65535), afin que le trafic de réponse soit autorisé sur la base de la règle du groupe de sécurité et non sur les informations de suivi. Grâce à cette configuration, Classic et Application Load Balancer ne sont pas soumis aux limites de suivi des connexions épuisées qui peuvent affecter l'établissement de nouvelles connexions à ses nœuds d'équilibreur de charge, et leur permet d'évoluer en fonction de l'augmentation du trafic en cas d'attaque. DDoS Vous trouverez de plus amples informations sur les connexions non suivies à l'adresse suivante : Suivi des connexions des groupes de sécurité : connexions non suivies.

Éviter le suivi des connexions du groupe de sécurité n'est utile que dans les cas où le DDoS trafic provient d'une source autorisée par le groupe de sécurité. Le DDoS trafic provenant de sources non autorisées dans le groupe de sécurité n'affecte pas le suivi des connexions. Dans ces cas, il n'est pas nécessaire de reconfigurer vos groupes de sécurité pour éviter le suivi des connexions, par exemple, si votre liste d'autorisations de groupes de sécurité comprend des plages d'adresses IP en lesquelles vous avez un haut degré de confiance, comme un pare-feu d'entreprise ou une VPN sortie IPs sécurisée ou. CDNs

Utiliser les emplacements AWS Edge pour la mise à l'échelle (BP1,BP3)

L'accès à des connexions Internet diversifiées et à grande échelle peut augmenter considérablement votre capacité à optimiser la latence et le débit pour les utilisateurs, à absorber les DDoS attaques et à isoler les défaillances tout en minimisant l'impact sur la disponibilité de votre application. AWS les emplacements périphériques fournissent une couche supplémentaire d'infrastructure réseau qui offre ces avantages à toute application Web utilisant Amazon CloudFront, Global Accelerator et Amazon Route 53. Grâce à ces services, vous pouvez protéger de manière complète en périphérie vos applications qui s'exécutent Régions AWS.

Livraison d'applications Web à la périphérie (BP1)

Amazon CloudFront est un service qui peut être utilisé pour diffuser l'intégralité de votre site Web, y compris du contenu statique, dynamique, en streaming et interactif. Les connexions persistantes et les paramètres variables time-to-live (TTL) peuvent être utilisés pour décharger le trafic de votre source, même si vous ne diffusez pas de contenu pouvant être mis en cache. L'utilisation de ces CloudFront fonctionnalités réduit le nombre de demandes et de TCP connexions renvoyant à votre origine, ce qui contribue à protéger votre application Web des HTTP inondations.

CloudFront n'accepte que les connexions bien établies, ce qui permet d'empêcher de nombreuses DDoS attaques courantes, telles que SYN les inondations et les attaques par UDP réflexion, d'atteindre votre point d'origine. DDoSles attaques sont également isolées géographiquement à proximité de leur source, ce qui empêche le trafic d'avoir un impact sur d'autres sites. Ces fonctionnalités peuvent considérablement améliorer votre capacité à continuer à fournir du trafic aux utilisateurs lors d'DDoSattaques de grande envergure. Vous pouvez l' CloudFront utiliser pour protéger une origine sur Internet AWS ou ailleurs.

Si vous utilisez Amazon Simple Storage Service (Amazon S3) pour diffuser du contenu statique sur Internet, nous vous AWS recommandons d'utiliser Amazon CloudFront pour protéger votre compartiment en offrant les avantages suivants :

- Restreint l'accès au compartiment Amazon S3 afin qu'il ne soit pas accessible au public.
- Veille à ce que les spectateurs (utilisateurs) puissent accéder au contenu du bucket uniquement par le biais de la CloudFront distribution spécifiée, c'est-à-dire qu'il les empêche d'accéder au contenu directement depuis le bucket ou par le biais d'une distribution involontaire. CloudFront

Pour ce faire, configurez CloudFront pour envoyer des demandes authentifiées à Amazon S3, et configurez Amazon S3 pour autoriser uniquement l'accès aux demandes authentifiées provenant de. CloudFront CloudFront propose deux méthodes pour envoyer des demandes authentifiées à une origine Amazon S3 : le contrôle d'accès à l'origine (OAC) et l'identité d'accès à l'origine (OAI). Nous vous recommandons de l'utiliser OAC car il prend en charge :

- Tous les compartiments Amazon S3 en tout Régions AWS, y compris les régions optionnelles lancées après décembre 2022
- Chiffrement côté serveur Amazon S3 avec AWS KMS (SSE-) KMS
- Demandes dynamiques (`PUT` et `DELETE`) vers Amazon S3

Pour plus d'informations sur OAC etOAI, reportez-vous à Restreindre l'accès à l'origine Amazon S3.

Pour plus d'informations sur la protection et l'optimisation des performances des applications Web avec Amazon CloudFront, consultez Getting Started with Amazon CloudFront.

Protégez le trafic réseau plus loin de votre point d'origine à l'aide de AWS Global Accelerator (BP1)

Global Accelerator est un service réseau qui améliore la disponibilité et les performances du trafic des utilisateurs jusqu'à 60 %. Cela se fait en faisant entrer le trafic à l'emplacement périphérique le plus proche de vos utilisateurs et en le dirigeant vers votre application via l'infrastructure réseau AWS mondiale, qu'elle s'exécute en une ou plusieurs Régions AWS applications.

Global Accelerator achemine TCP et achemine le UDP trafic vers le point de terminaison optimal en fonction des performances Région AWS au plus proche de l'utilisateur. En cas de défaillance d'une application, Global Accelerator assure le basculement vers le meilleur point de terminaison suivant dans les 30 secondes. Global Accelerator utilise la vaste capacité du réseau AWS mondial et les intégrations avec Shield, telles que la fonctionnalité de SYN proxy apatride qui défie les nouvelles tentatives de connexion et ne sert que les utilisateurs finaux légitimes, afin de protéger les applications.

Vous pouvez mettre en œuvre une architecture DDoS résiliente qui offre bon nombre des mêmes avantages que les meilleures pratiques de diffusion d'applications Web à la périphérie, même si votre application utilise des protocoles non pris en charge par CloudFront ou si vous utilisez une application Web qui nécessite des adresses IP statiques globales.

Par exemple, vous pouvez avoir besoin d'adresses IP que vos utilisateurs finaux peuvent ajouter à la liste d'autorisation de leur pare-feu et qui ne soient utilisées par aucun autre AWS client. Dans ces scénarios, vous pouvez utiliser Global Accelerator pour protéger les applications Web exécutées sur Application Load Balancer et, conjointement, pour détecter et atténuer les inondations de demandes AWS WAF au niveau de la couche d'application Web.

Pour plus d'informations sur la protection et l'optimisation des performances du trafic réseau à l'aide de Global Accelerator, consultez Getting started with Global Accelerator.

Résolution des noms de domaine à la périphérie (BP3)

Rubriques

- Utiliser Route 53 pour vérifier DNS la disponibilité
- Configuration de la Route 53 pour la protection des coûts contre NXDOMAIN les attaques

Utiliser Route 53 pour vérifier DNS la disponibilité

Amazon Route 53 est un service de système de noms de domaine (DNS) hautement disponible et évolutif qui peut être utilisé pour diriger le trafic vers votre application Web. Il inclut des fonctionnalités avancées telles que Traffic Flow, Health Checks and Monitoring, le routage basé sur la latence et la géolocalisation. DNS Ces fonctionnalités avancées vous permettent de contrôler la façon dont le service répond aux DNS demandes afin d'améliorer les performances de votre application Web et d'éviter les pannes du site. C'est le seul AWS service dont le plan de données est disponible à 100 %SLA.

Amazon Route 53 utilise des techniques telles que le shuffle sharding et le anycast striping, qui peuvent aider les utilisateurs à accéder à votre application même si le DNS service est la cible d'une attaque. DDoS

Avec le shuffle sharding, chaque serveur de noms de votre ensemble de délégations correspond à un ensemble unique d'emplacements périphériques et de chemins Internet. Cela permet une meilleure tolérance aux pannes et minimise les chevauchements entre les clients. Si l'un des serveurs de noms du jeu de délégation n'est pas disponible, les utilisateurs peuvent réessayer et recevoir une réponse d'un autre serveur de noms situé à un autre emplacement périphérique.

Le striping Anycast permet à chaque DNS demande d'être traitée à l'emplacement le plus optimal, en répartissant la charge du réseau et en réduisant la latence. DNS Cela permet aux utilisateurs de

réagir plus rapidement. En outre, Amazon Route 53 peut détecter des anomalies dans la source et le volume des DNS requêtes, et hiérarchiser les demandes provenant d'utilisateurs réputés fiables.

Pour plus d'informations sur l'utilisation d'Amazon Route 53 pour rediriger les utilisateurs vers votre application, consultez Getting Started with Amazon Route 53.

Configuration de la Route 53 pour la protection des coûts contre **NXDOMAIN** les attaques

`NXDOMAIN`les attaques se produisent lorsque les attaquants envoient un flot de requêtes à une zone hébergée pour des sous-domaines inexistants, souvent via de « bons » résolveurs connus. Le but de ces attaques peut être d'avoir un impact sur le cache du résolveur récursif et/ou la disponibilité du résolveur faisant autorité, ou peuvent être une forme de DNS reconnaissance visant à tenter de découvrir des enregistrements de zones hébergées. L'utilisation de Route 53 comme résolveur fiable atténue le risque d'impact sur la disponibilité et les performances, mais cela peut se traduire par une augmentation significative des coûts mensuels de Route 53. Pour vous protéger contre les augmentations de coûts, profitez de la tarification Route 53, dans laquelle les DNS requêtes sont gratuites lorsque les deux conditions suivantes sont vraies :

- Le nom de domaine ou de sous-domaine (`example.comoustore.example.com`) et le type d'enregistrement (`A`) de la requête correspondent à un enregistrement alias.
- L'alias cible est une AWS ressource autre qu'un autre enregistrement Route 53.

Créez un enregistrement générique, par exemple, `*.example.com` avec un type `A` (Alias) pointant vers une AWS ressource telle qu'une EC2 instance, Elastic Load Balancer CloudFront ou une distribution, de sorte que lorsqu'une requête `qwerty12345.example.com` est effectuée, l'adresse IP de la ressource soit renvoyée et la requête ne vous soit pas facturée.

Défense de la couche applicative (BP1,BP2)

La plupart des techniques abordées jusqu'ici dans ce paper sont efficaces pour atténuer l'impact des DDoS attaques de la couche d'infrastructure sur la disponibilité de votre application. Pour également vous défendre contre les attaques de la couche applicative, vous devez mettre en œuvre une architecture qui vous permet de détecter, de dimensionner pour absorber et de bloquer spécifiquement les demandes malveillantes. Il s'agit d'une considération importante car les systèmes DDoS d'atténuation basés sur le réseau sont généralement inefficaces pour atténuer les attaques complexes au niveau de la couche applicative.

Détectez et filtrez les requêtes Web malveillantes (BP1,BP2)

Lorsque votre application s'exécute AWS, vous pouvez tirer parti d'Amazon CloudFront (et de sa capacité de HTTP mise en cache) et de Shield Advanced Automatic Application Layer Protection pour empêcher que des demandes inutiles n'atteignent votre source lors d'DDoSattaques contre la couche application. AWS WAF

Amazon CloudFront

Amazon CloudFront peut contribuer à réduire la charge du serveur en empêchant le trafic non Web d'atteindre votre point d'origine. Pour envoyer une demande à une CloudFront application, la connexion doit être établie avec une adresse IP valide par le biais d'une TCP poignée de main complète, qui ne peut pas être falsifiée. En outre, il CloudFront peut fermer automatiquement les connexions en cas d'attaques à lecture lente ou à écriture lente (par exemple, Slowloris).

Mise en cache CDN

CloudFront vous permet de diffuser à la fois du contenu dynamique et du contenu statique à partir d'emplacements AWS périphériques. En diffusant du contenu pouvant être mis en cache par proxy à partir du CDN cache, vous empêchez les demandes d'atteindre votre origine à partir d'un nœud de cache périphérique donné pendant toute la durée de la mise en cacheTTL. En conjonction avec le regroupement des demandes pour du contenu expiré mais pouvant être mis en cache, même les demandes très courtes TTL signifient qu'un nombre négligeable de demandes parviendront à votre origine lors d'un afflux de demandes pour ce contenu. En outre, l'activation de fonctionnalités telles qu'CloudFront Origin Shield peut contribuer à réduire encore la charge de travail sur votre ordinateur d'origine. Tout ce que vous pouvez faire pour améliorer le taux de réussite de votre cache peut faire la différence entre une attaque de type « request flood » ayant un impact et une attaque sans impact.

AWS WAF

À l'aide de AWS WAF, vous pouvez configurer des listes de contrôle d'accès Web (WebACLs) sur vos CloudFront distributions mondiales ou vos ressources régionales pour filtrer, surveiller et bloquer les demandes en fonction des signatures de demande. Pour déterminer s'il convient d'autoriser ou de bloquer les demandes, vous pouvez prendre en compte des facteurs tels que l'adresse IP ou le pays d'origine, certaines chaînes ou modèles de la demande, la taille de certaines parties de la demande et la présence de SQL code ou de script malveillants. Vous pouvez également lancer CAPTCHA des puzzles et des défis de session client silencieux en réponse à des demandes.

Les deux AWS WAF vous permettent CloudFront également de définir des restrictions géographiques pour bloquer ou autoriser les demandes provenant de pays sélectionnés. Cela peut aider à bloquer

ou à limiter le taux d'attaques provenant de zones géographiques où vous ne vous attendez pas à desservir les utilisateurs. Grâce à des règles de correspondance géographique précises AWS WAF, vous pouvez contrôler l'accès jusqu'au niveau de la région.

Vous pouvez utiliser les instructions Scope-down pour réduire la portée des demandes évaluées par la règle afin de réduire les coûts et les « étiquettes » sur les requêtes Web pour permettre à une règle correspondant à la demande de communiquer les résultats du match aux règles évaluées ultérieurement sur le même site Web. ACL Choisissez cette option pour réutiliser la même logique dans plusieurs règles.

Vous pouvez également définir une réponse personnalisée complète, avec le code de réponse, les en-têtes et le corps.

Pour identifier les demandes malveillantes, consultez les journaux de votre serveur Web ou utilisez AWS WAF la journalisation et l'échantillonnage des demandes. En activant la AWS WAF journalisation, vous obtenez des informations détaillées sur le trafic analysé par le Web. ACL AWS WAF prend en charge le filtrage des journaux, ce qui vous permet de spécifier quelles requêtes Web sont enregistrées et quelles demandes sont supprimées du journal après l'inspection.

Les informations enregistrées dans les journaux incluent l'heure à laquelle la demande AWS WAF a été reçue de votre AWS ressource, des informations détaillées sur la demande et l'action correspondante pour chaque règle demandée.

Les exemples de demandes fournissent des informations sur les demandes effectuées au cours des trois dernières heures qui correspondaient à l'une de vos AWS WAF règles. Vous pouvez utiliser ces informations pour identifier les signatures de trafic potentiellement malveillantes et créer une nouvelle règle pour refuser ces demandes. Si plusieurs demandes comportent une chaîne de requête aléatoire, veillez à n'autoriser que les paramètres de chaîne de requête pertinents pour le cache de votre application. Cette technique est utile pour atténuer une attaque de piratage du cache dirigée contre votre ordinateur d'origine.

AWS WAF — Règles basées sur les taux

AWS recommande vivement de se protéger contre les inondations de HTTP demandes en utilisant les règles basées sur le débit AWS WAF pour bloquer automatiquement les adresses IP des acteurs malveillants lorsque le nombre de demandes reçues dans une fenêtre glissante de 5 minutes dépasse un seuil que vous définissez. Les adresses IP des clients incriminées recevront une réponse interdite 403 (ou une réponse d'erreur de blocage configurée) et resteront bloquées jusqu'à ce que le taux de demandes tombe en dessous du seuil.

Il est recommandé de superposer des règles basées sur le taux afin de fournir une protection améliorée afin que vous puissiez :

- Une règle générale basée sur les taux pour protéger votre demande contre HTTP les grandes inondations.
- Une ou plusieurs règles basées sur les taux pour protéger des taux spécifiques URIs plus restrictifs que la règle basée sur les taux généraux.

Par exemple, vous pouvez choisir une règle basée sur un taux global (aucune déclaration de portée réduite) avec une limite de 500 demandes sur une période de 5 minutes, puis créer une ou plusieurs des règles basées sur les taux suivantes avec des limites inférieures à 500 (100 demandes sur une période de 5 minutes) à l'aide d'instructions de portée réduite :

- Protégez vos pages Web à l'aide d'une instruction de portée réduite telle que `if NOT uri_path contains '.'` « » afin de mieux protéger les demandes de ressources sans extension de fichier. Cela protège également votre page d'accueil (`/`), qui est un URI chemin fréquemment ciblé.
- Protégez les points de terminaison dynamiques avec une instruction de portée réduite telle que « » `if method exactly matches 'post' (convert lowercase)`
- Protégez les demandes volumineuses qui atteignent votre base de données ou invoquez un mot de passe à usage unique (OTP) avec une portée descendante du type « » `if uri_path starts_with '/login' OR uri_path starts_with '/signup' OR uri_path starts_with '/forgotpassword'`

La base tarifaire en mode « Bloc » constitue la pierre angulaire de votre defense-in-depth WAF configuration pour vous protéger contre les inondations de demandes et constitue une condition préalable à l'approbation des demandes de protection des AWS Shield Advanced coûts. Nous examinerons d'autres defense-in-depth WAF configurations dans les sections suivantes.

AWS WAF — Réputation IP

Pour empêcher les attaques basées sur la réputation des adresses IP, vous pouvez créer des règles à l'aide de la correspondance d'adresses IP ou utiliser des règles gérées pour AWS WAF.

Le groupe de règles de liste de réputation IP d'Amazon inclut des règles basées sur les renseignements internes d'Amazon sur les menaces. Ces règles recherchent les adresses IP qui sont des robots, effectuant des reconnaissances sur AWS des ressources ou participant activement

à DDoS des activités. La `AWSManagedIPDDoSList` règle a été observée bloquant plus de 90 % des flots de demandes malveillantes.

Le groupe de règles de liste d'adresses IP anonymes contient des règles visant à bloquer les demandes provenant de services qui permettent de masquer l'identité du spectateur. Il s'agit notamment des requêtes provenant de proxysVPNs, de nœuds Tor et de plateformes cloud (à l'exclusion AWS).

En outre, vous pouvez utiliser des listes de réputation IP tierces en utilisant le composant d'analyse des listes d'adresses IP de la solution Security Automations for AWS WAF.

AWS WAF - Atténuation intelligente des menaces

Les botnets constituent une grave menace pour la sécurité et sont couramment utilisés pour mener des activités illégales ou nuisibles, telles que l'envoi de spam, le vol de données sensibles, le lancement d'attaques par ransomware, la fraude publicitaire par le biais de clics frauduleux ou le lancement d'attaques distribuées denial-of-service (DDoS). Pour empêcher les attaques de bots, utilisez le groupe de règles géré par AWS WAF Bot Control. Ce groupe de règles fournit un niveau de protection de base « commun » qui ajoute des étiquettes aux robots qui s'identifient eux-mêmes, vérifie les robots généralement souhaitables et détecte les signatures de robots hautement fiables, ainsi qu'un niveau de protection « ciblé » qui permet de détecter les robots avancés qui ne s'identifient pas eux-mêmes.

Les protections ciblées utilisent des techniques de détection avancées telles que l'interrogation du navigateur, la prise d'empreintes digitales et l'heuristique comportementale pour identifier le trafic de bots défectueux, puis appliquent des contrôles d'atténuation tels que la limitation du débit CAPTCHA et les actions relatives aux règles de défi. Targeted propose également des options de limitation de débit pour appliquer des modèles d'accès de type humain et appliquer une limitation de débit dynamique grâce à l'utilisation de jetons de demande. Pour plus de détails, consultez la section Groupe de règles AWS WAF Bot Control. Pour détecter et gérer les tentatives de prise de contrôle malveillantes sur la page de connexion de votre application, vous pouvez utiliser le groupe de règles AWS WAF Fraud Control account takeover prevention (ATP). Pour ce faire, le groupe de règles inspecte les tentatives de connexion que les clients envoient au point de terminaison de connexion de votre application et examine également les réponses de votre application aux tentatives de connexion, afin de suivre le taux de réussite et d'échec.

La fraude liée à la création de compte est une activité illégale en ligne dans le cadre de laquelle un attaquant tente de créer un ou plusieurs faux comptes. Les attaquants utilisent de faux comptes pour des activités frauduleuses, telles que l'utilisation abusive de bonus promotionnels et d'inscription,

l'usurpation d'identité de quelqu'un et des cyberattaques telles que le phishing. La présence de faux comptes peut avoir un impact négatif sur votre entreprise en portant atteinte à votre réputation auprès des clients et en vous exposant à des fraudes financières.

Vous pouvez surveiller et contrôler les tentatives de fraude liées à la création de comptes en mettant en œuvre la fonctionnalité de prévention des AWS WAF fraudes lors de la création de comptes Fraud Control (ACFP). AWS WAF propose cette fonctionnalité dans le groupe de AWS Managed Rules règles `AWS ManagedRulesACFPRuleSet` avec intégration d'applications complémentairesSDKs.

Apprenez-en davantage sur ces protections dans le cadre de l'atténuation AWS WAF intelligente des menaces.

Atténuez automatiquement les DDoS événements liés à la couche applicative (BP1,,BP2) BP6

Si vous êtes abonné AWS Shield Advanced, vous pouvez activer l'DDoSatténuation automatique de la couche d'application de Shield Advanced. Cette fonctionnalité crée, évalue et déploie automatiquement des AWS WAF règles pour atténuer les DDoS événements de couche 7 en votre nom.

AWS Shield Advanced établit une base de trafic pour chaque ressource protégée associée à un WAF site WebACL. Le trafic qui s'écarte de manière significative de la base de référence établie est signalé comme un événement potentielDDoS. Après la détection d'un événement, AWS Shield Advanced tente d'identifier une signature des requêtes Web qui constituent l'événement, et si une signature est identifiée, des AWS WAF règles sont créées pour atténuer le trafic associé à cette signature.

Une fois que les règles sont évaluées par rapport à la base de référence historique et considérées comme sûres, elles sont ajoutées au groupe de règles géré par Shield, et vous pouvez choisir si les règles sont déployées en mode décompte ou en mode bloc. Shield Advanced supprime automatiquement les AWS WAF règles une fois qu'il a déterminé qu'un événement s'est complètement calmé.

Engage SRT (abonnés Shield Advanced uniquement)

En outre, lorsque vous êtes abonné à Shield Advanced, vous pouvez les utiliser AWS SRT pour vous aider à créer des règles visant à atténuer une attaque qui nuit à la disponibilité de votre application. Vous pouvez accorder un accès AWS SRT limité à votre compte AWS Shield Advanced et AWS WAF APIs. AWS SRTn'y accède pour APIs appliquer des mesures d'atténuation sur votre compte

qu'avec votre autorisation explicite. Pour plus d'informations, reportez-vous à la Support section de ce document.

Vous pouvez l'utiliser AWS Firewall Manager pour configurer et gérer de manière centralisée les règles de sécurité, telles que AWS Shield Advanced les protections et AWS WAF les règles, au sein de votre organisation. Votre compte AWS Organizations de gestion peut désigner un compte administrateur, qui est autorisé à créer des politiques Firewall Manager. Ces politiques vous permettent de définir des critères, tels que le type de ressource et les balises, qui déterminent où les règles sont appliquées. Cela est utile lorsque vous avez plusieurs comptes et que vous souhaitez standardiser votre protection.

Pour plus d’informations sur :

- AWS Managed Rules pour AWS WAF, référez-vous à AWS Managed Rules pour AWS WAF.
- En utilisant les restrictions géographiques pour limiter l'accès à votre CloudFront distribution, reportez-vous à la section Restreindre la distribution géographique de votre contenu.
- En utilisant AWS WAF, reportez-vous à :
  - Commencer avec AWS WAF
  - Enregistrement des informations sur ACL le trafic Web
  - Affichage d'un échantillon de requêtes Web
- Pour configurer les règles basées sur les taux, reportez-vous à Protéger les sites Web et les services à l'aide de règles basées sur les taux pour. AWS WAF
- Pour gérer le déploiement de règles sur l'ensemble de vos AWS ressources avec Firewall Manager, consultez :
  - Commencer à utiliser les AWS WAF politiques de Firewall Manager
  - Commencer à utiliser les politiques avancées de Firewall Manager Shield

Réduction de la surface d'attaque

Lors de l'architecture d'une AWS solution, il est également important de limiter les chances qu'un attaquant cible votre application. Ce concept est connu sous le nom de réduction de la surface d'attaque. Les ressources qui ne sont pas exposées à Internet sont plus difficiles à attaquer, ce qui limite les options dont dispose un attaquant pour cibler la disponibilité de votre application.

Par exemple, si vous ne vous attendez pas à ce que les utilisateurs interagissent directement avec certaines ressources, assurez-vous que ces ressources ne sont pas accessibles depuis Internet. De même, n'acceptez pas le trafic provenant d'utilisateurs ou d'applications externes sur des ports ou des protocoles qui ne sont pas nécessaires à la communication.

Dans la section suivante, vous AWS trouverez les meilleures pratiques pour vous aider à réduire votre surface d'attaque et à limiter l'exposition de votre application à Internet.

Masquer les AWS ressources (,,) BP1 BP4 BP5

En règle générale, les utilisateurs peuvent utiliser rapidement et facilement une application sans avoir besoin que les AWS ressources soient entièrement exposées à Internet.

Groupes de sécurité et réseau ACLs (BP5)

Amazon Virtual Private Cloud (AmazonVPC) vous permet de fournir une section isolée de manière logique AWS Cloud où vous pouvez lancer AWS des ressources dans un réseau virtuel que vous définissez.

Les groupes de sécurité et le réseau ACLs sont similaires en ce sens qu'ils vous permettent de contrôler l'accès aux AWS ressources de votreVPC. Mais les groupes de sécurité vous permettent de contrôler le trafic entrant et sortant au niveau de l'instance, tandis que le réseau ACLs offre des fonctionnalités similaires au niveau du VPC sous-réseau. L'utilisation des groupes de sécurité ou du réseau est gratuiteACLs.

Vous pouvez choisir de spécifier des groupes de sécurité lorsque vous lancez une instance ou d'associer l'instance à un groupe de sécurité ultérieurement. Tout le trafic Internet vers un groupe de sécurité est implicitement refusé, sauf si vous créez une règle d'autorisation pour autoriser le trafic.

Par exemple, lorsque des EC2 instances Amazon sont associées à un Elastic Load Balancer, il n'est pas nécessaire que les instances elles-mêmes soient accessibles au public et doivent être

uniquement privéesIPs. Vous pouvez plutôt fournir à Elastic Load Balancer l'accès aux ports d'écoute cibles requis en utilisant une règle de groupe de sécurité qui autorise l'accès à 0.0.0.0/0 (pour éviter les problèmes de suivi des connexions, voir note ci-dessous) en conjonction avec une liste de contrôle d'accès réseau (NACL) sur le sous-réseau du groupe cible afin d'autoriser uniquement les plages d'adresses IP d'Elastic Load Balancing à communiquer avec les instances. Cela garantit que le trafic Internet ne peut pas communiquer directement avec vos EC2 instances Amazon, ce qui rend plus difficile pour un attaquant de découvrir et d'influencer votre application.

Lorsque vous créez un réseauACLs, vous pouvez définir des règles d'autorisation et de refus. Cela est utile si vous souhaitez refuser explicitement certains types de trafic vers votre application. Par exemple, vous pouvez définir des adresses IP (sous forme de CIDR plages), des protocoles et des ports de destination auxquels l'accès à l'ensemble du sous-réseau est refusé. Si votre application est utilisée uniquement pour TCP le trafic, vous pouvez créer une règle pour refuser tout UDP trafic, ou vice versa. Cette option est utile lorsque vous répondez à DDoS des attaques, car elle vous permet de créer vos propres règles pour atténuer l'attaque lorsque vous connaissez la source IPs ou une autre signature.

Si vous êtes abonné AWS Shield Advanced, vous pouvez enregistrer les adresses IP Elastic en tant que ressources protégées. DDoSles attaques contre les adresses IP Elastic enregistrées en tant que ressources protégées sont détectées plus rapidement, ce qui permet de réduire les délais d'atténuation. Lorsqu'une attaque est détectée, les systèmes DDoS d'atténuation lisent le réseau ACL correspondant à l'adresse IP élastique ciblée et l'appliquent à la frontière du AWS réseau plutôt qu'au niveau du sous-réseau. Cela réduit considérablement le risque d'impact d'un certain nombre d'DDoSattaques au niveau de l'infrastructure.

Pour plus d'informations sur la configuration des groupes de sécurité et du réseau ACLs afin d'optimiser DDoS la résilience, consultez Comment vous préparer aux DDoS attaques en réduisant votre surface d'attaque.

Pour plus d'informations sur l'utilisation de Shield Advanced avec des adresses IP élastiques en tant que ressources protégées, reportez-vous aux étapes de souscription AWS Shield Advanced.

Protéger votre origine (BP1,BP5)

Si vous utilisez Amazon CloudFront avec une origine qui se trouve dans le vôtreVPC, vous voudrez peut-être vous assurer que seule votre CloudFront distribution peut transmettre les demandes à votre origine. Avec les en-têtes de demande Edge-to-Origin, vous pouvez ajouter ou remplacer la valeur des en-têtes de demande existants lorsque CloudFront vous transférez des demandes à votre origine. Vous pouvez utiliser les en-têtes personnalisés d'Origin, par exemple l'`X-Shared-`

`Secreten-tête`, pour vérifier que les demandes adressées à votre origine ont été envoyées depuis CloudFront.

Pour plus d'informations sur la protection de votre origine avec des en-têtes personnalisés Origin, reportez-vous aux sections Ajouter des en-têtes personnalisés aux demandes d'origine et Restreindre l'accès aux équilibreurs de charge d'application.

Pour un guide sur la mise en œuvre d'un exemple de solution permettant de faire automatiquement pivoter la valeur des en-têtes personnalisés d'Origin pour la restriction d'accès à l'origine, reportez-vous à la section Comment améliorer la sécurité des CloudFront origines d'Amazon avec AWS WAF et Secrets Manager.

Vous pouvez également utiliser une AWS Lambdafonction pour mettre à jour automatiquement les règles de votre groupe de sécurité afin d'autoriser uniquement CloudFront le trafic. Cela améliore la sécurité de votre origine en empêchant les utilisateurs malveillants de contourner CloudFront et AWS WAF d'accéder à votre application Web.

Pour plus d'informations sur la manière de protéger votre origine en mettant automatiquement à jour vos groupes de sécurité et l'`X-Shared-Secreten-tête`, consultez Comment mettre à jour automatiquement vos groupes de sécurité pour Amazon CloudFront et en AWS WAF utilisant AWS Lambda.

Cependant, la solution implique une configuration supplémentaire et le coût d'exécution des fonctions Lambda. Pour simplifier les choses, nous avons introduit une liste de AWS préfixes gérée afin de limiter le HTTPS trafic entrant CloudFront HTTP vers vos origines à partir des seules adresses IP orientées vers l' CloudFrontorigine. AWS-les listes de préfixes gérées sont créées et maintenues par AWS et peuvent être utilisées sans frais supplémentaires. Vous pouvez faire référence à la liste des préfixes gérés CloudFront dans les règles de votre groupe de sécurité (AmazonVPC), dans les tables de routage des sous-réseaux, dans les règles communes des AWS Firewall Manager groupes de sécurité et dans toute autre AWS ressource pouvant utiliser une liste de préfixes gérés.

Pour plus d'informations sur l'utilisation de la liste de préfixes AWS gérée par -pour Amazon CloudFront, consultez Limiter l'accès à vos origines à l'aide de la liste de AWS préfixes -gérée pour Amazon. CloudFront

> **Note**
>
> Comme indiqué dans d'autres sections de ce document, le fait de s'appuyer sur des groupes de sécurité pour protéger votre origine peut ajouter au suivi des connexions des groupes de sécurité un obstacle potentiel lors d'un afflux de demandes. À moins que vous ne soyez

en mesure de filtrer les demandes malveillantes à CloudFront l'aide d'une politique de mise en cache qui active la mise en cache, il peut être préférable de vous fier aux en-têtes personnalisés d'origine, évoqués précédemment, pour vérifier que les demandes adressées à votre origine ont été envoyées depuis CloudFront, plutôt que d'utiliser des groupes de sécurité. L'utilisation d'un en-tête de demande personnalisé avec une règle d'écoute Application Load Balancer permet d'éviter les ralentissements dus aux limites de suivi susceptibles d'affecter l'établissement de nouvelles connexions à un équilibreur de charge, ce qui permet à Application Load Balancer de s'adapter en fonction de l'augmentation du trafic en cas d'attaque. DDoS

Protection des API terminaux () BP4

Lorsque vous devez exposer un API fichier au public, il existe un risque que le API frontend soit la cible d'une DDoS attaque. Pour réduire les risques, vous pouvez utiliser Amazon API Gateway comme point d'accès aux applications exécutées sur Amazon EC2 ou AWS Lambda ailleurs. En utilisant Amazon API Gateway, vous n'avez pas besoin de vos propres serveurs pour le API frontend et vous pouvez masquer d'autres composants de votre application. En rendant plus difficile la détection des composants de votre application, vous pouvez empêcher que ces AWS ressources ne soient ciblées par une DDoS attaque.

Lorsque vous utilisez Amazon API Gateway, vous pouvez choisir entre deux types de API points de terminaison. La première est l'option par défaut : des API points de terminaison optimisés pour les périphériques accessibles via une distribution Amazon. CloudFront Cependant, la distribution est créée et gérée par API Gateway, vous n'en avez donc aucun contrôle. La deuxième option consiste à utiliser un point de API terminaison régional accessible depuis le même point de terminaison que celui Région AWS dans lequel le vôtre REST API est déployé. AWS vous recommande d'utiliser le second type de point de terminaison et de l'associer à votre propre CloudFront distribution Amazon. Cela vous permet de contrôler la CloudFront distribution Amazon et de pouvoir l'utiliser AWS WAF pour la protection de la couche application. Ce mode vous donne accès à une capacité DDoS d'atténuation étendue sur l'ensemble du réseau périphérique AWS mondial.

Lorsque vous utilisez Amazon CloudFront et AWS WAF Amazon API Gateway, configurez les options suivantes :

- Configurez le comportement du cache pour vos distributions afin de transférer tous les en-têtes au point de terminaison régional API Gateway. Ce faisant, CloudFront vous traiterez le contenu comme dynamique et vous éviterez de le mettre en cache.

- Protégez votre API passerelle contre l'accès direct en configurant la distribution pour inclure l'en-tête personnalisé d'origine x-api-key, en définissant la valeur APIclé dans API Gateway.
- Protégez le backend contre le trafic excessif en configurant des limites de fréquence standard ou de fréquence de rafale pour chaque méthode de votre RESTAPIs.

Pour plus d'informations sur la création APIs avec Amazon API Gateway, consultez Amazon API Gateway Getting Started.

Techniques opérationnelles

Les techniques d'atténuation décrites dans ce paper vous aident à concevoir des applications intrinsèquement résilientes face aux DDoS attaques. Dans de nombreux cas, il est également utile de savoir quand une DDoS attaque cible votre application afin de pouvoir prendre des mesures d'atténuation. Cette section décrit les meilleures pratiques pour améliorer la visibilité sur les comportements anormaux, les alertes et l'automatisation, gérer la protection à grande échelle et AWS solliciter une assistance supplémentaire.

Test de charge

Testez régulièrement la charge de votre application en suivant les directives de notre livre blanc sur les applications de test de charge, avec des niveaux de trafic attendus et supérieurs aux attentes, afin de vérifier l'efficacité de votre architecture, le fonctionnement de vos politiques Auto Scaling et le fonctionnement de votre gestion des erreurs. Testez l'augmentation ou la baisse attendues du trafic, mais également le comportement de type « flash-crowd ». Effectuez un nouveau test périodiquement ou avant toute version majeure. Pour les tests de DDoS simulation de couche 3 ou 4, tels que les SYN inondations, suivez notre politique DDoS de test de simulation.

Métriques et alarmes

Il est recommandé d'utiliser des outils de surveillance de l'infrastructure et des applications pour vérifier la disponibilité de votre application afin de vous assurer qu'elle n'est pas affectée par un DDoS événement. En option, vous pouvez configurer des contrôles de santé de l'application et de l'infrastructure Route 53 pour les ressources afin d'améliorer la détection des DDoS événements. Pour plus d'informations sur les contrôles de santé AWS WAF, consultez le guide du développeur Firewall Manager and Shield Advanced.

Lorsqu'un indicateur opérationnel clé s'écarte considérablement de la valeur attendue, il se peut qu'un attaquant tente de cibler la disponibilité de votre application. La connaissance du comportement normal de votre application vous permet d'agir plus rapidement lorsque vous détectez une anomalie. Amazon CloudWatch peut vous aider en surveillant les applications que vous exécutez AWS. Par exemple, vous pouvez collecter et suivre des métriques, collecter et surveiller des fichiers journaux, définir des alarmes et répondre automatiquement aux modifications de vos AWS ressources.

Si vous suivez l'architecture de référence DDoS résiliente lors de l'architecture de votre application, les attaques courantes au niveau de la couche d'infrastructure seront bloquées avant d'atteindre votre application. Si vous êtes abonné AWS Shield Advanced, vous avez accès à un certain nombre de CloudWatch statistiques qui peuvent indiquer que votre application est ciblée.

Par exemple, vous pouvez configurer des alarmes pour vous avertir lorsqu'une DDoS attaque est en cours, afin de vérifier l'état de santé de votre application et de décider si vous souhaitez y participer AWS SRT. Vous pouvez configurer la `DDoSDetected` métrique pour savoir si une attaque a été détectée. Si vous souhaitez être alerté en fonction du volume d'attaques, vous pouvez également utiliser les `DDoSAttackRequestsPerSecond` métriques `DDoSAttackBitsPerSecond`DDoSAttackPacketsPerSecond, ou. Vous pouvez surveiller ces indicateurs en les intégrant CloudWatch à vos propres outils ou en utilisant des outils fournis par des tiers, tels que Slack ou PagerDuty.

Une attaque au niveau de la couche applicative peut augmenter de nombreux CloudWatch indicateurs Amazon. Si vous en utilisez AWS WAF, vous pouvez l'utiliser CloudWatch pour surveiller et activer des alarmes en cas d'augmentation du nombre de demandes que vous avez définies AWS WAF pour être autorisées, comptées ou bloquées. Cela vous permet de recevoir une notification si le niveau de trafic dépasse ce que votre application peut gérer. Vous pouvez également utiliser les métriques Amazon CloudFront, Amazon Route 53, Application Load Balancer, Network Load Balancer, EC2 Amazon et Auto Scaling qui sont CloudWatch suivies pour détecter les modifications susceptibles DDoS d'indiquer une attaque.

Le tableau suivant décrit les CloudWatch indicateurs couramment utilisés pour détecter les DDoS attaques et y réagir.

Tableau 3 : CloudWatch Indications Amazon recommandées

Rubrique	Métrique	Description
AWS Shield Advanced	`DDoSDetected`	Indique un DDoS événement pour un nom de ressource Amazon spécifique (ARN).
AWS Shield Advanced	`DDoSAttackBitsPerSecond`	Le nombre d'octets observés lors d'un DDoS événement pour un événement spécifiqu eARN. Cette métrique n'est

Rubrique	Métrique	Description
		disponible que pour les DDoS événements de couche 3 ou 4.
AWS Shield Advanced	`DDoSAttackPacketsPerSecond`	Le nombre de paquets observés lors d'un DDoS événement pour un événement spécifiqueARN. Cette métrique n'est disponible que pour les DDoS événements de couche 3 ou 4.
AWS Shield Advanced	`DDoSAttackRequestsPerSecond`	Le nombre de demandes observées lors d'un DDoS événement pour un sujet spécifiqueARN. Cette métrique n'est disponible que pour les DDoS événements de couche 7 et n'est signalée que pour les événements de couche 7 les plus importants.
AWS WAF	`AllowedRequests`	Nombre de requêtes Web autorisées.
AWS WAF	`BlockedRequests`	Nombre de requêtes Web bloquées.
AWS WAF	`CountedRequests`	Nombre de requêtes Web comptabilisées.

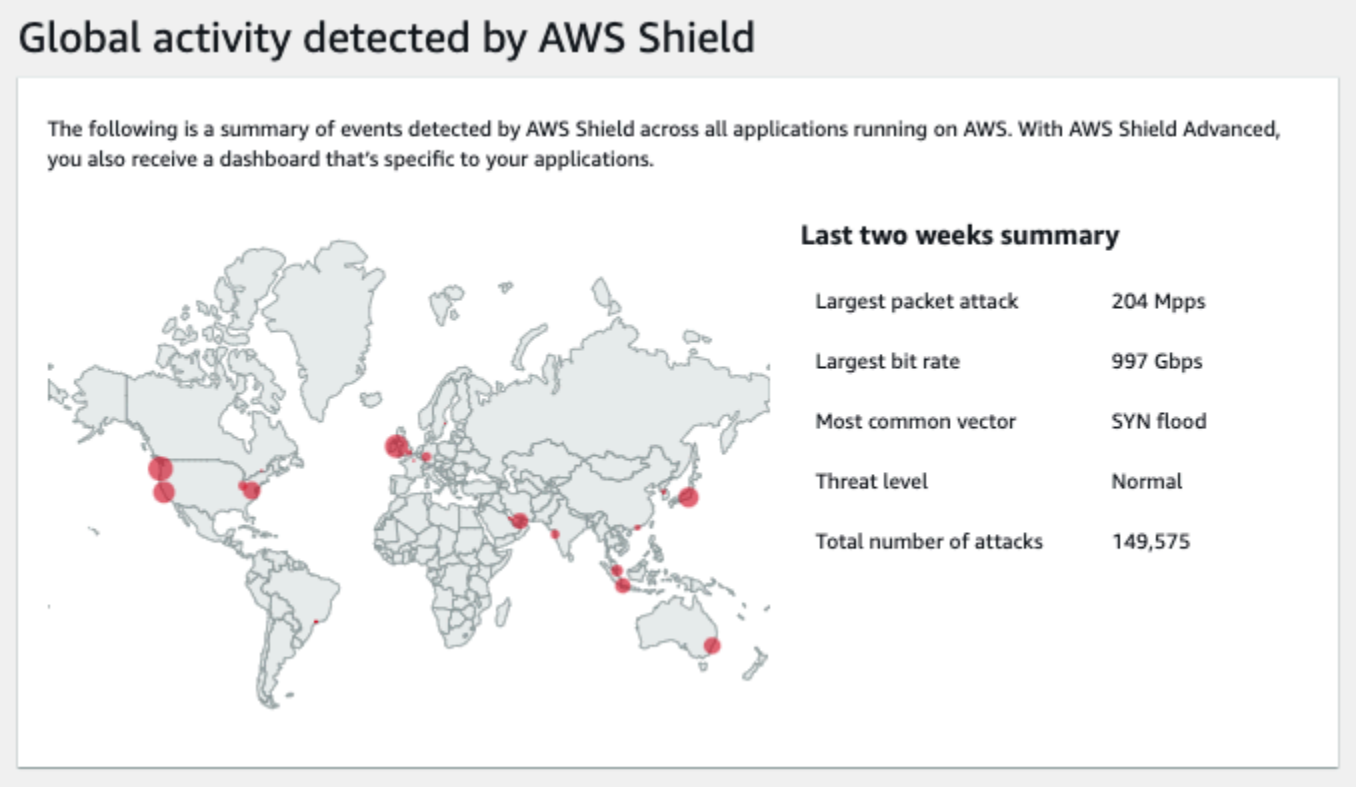
Rubrique	Métrique	Description
AWS WAF	`PassedRequests`	Le nombre de demandes passées. Ceci n'est utilisé que pour les demandes soumises à une évaluation de groupe de règles sans correspondre à aucune des règles du groupe de règles.
Amazon CloudFront	`Requests`	Le nombre de HTTP requêtes /S.
Amazon CloudFront	`TotalErrorRate`	Pourcentage de toutes les demandes pour lesquelles le code de HTTP statut est `4xx` ou`5xx`.
Amazon Route 53	`HealthCheckStatus`	État du point de terminaison du bilan de santé.
Application Load Balancer	`ActiveConnectionCount`	Nombre total de TCP connexions simultanées actives entre les clients et l'équilibreur de charge, et entre l'équilibreur de charge et les cibles.
Application Load Balancer	`ConsumedLCUs`	Le nombre d'unités de capacité de l'équilibreur de charge (LCU) utilisées par votre équilibreur de charge.
Application Load Balancer	`HTTPCode_ELB_4XX_Count` `HTTPCode_ELB_5XX_Count`	Le nombre de codes HTTP `4xx` d'erreur `5xx` client générés par l'équilibreur de charge.

Rubrique	Métrique	Description
Application Load Balancer	`NewConnectionCount`	Nombre total de nouvelles TCP connexions établies entre les clients et l'équilibreur de charge, et entre l'équilibreur de charge et les cibles.
Application Load Balancer	`ProcessedBytes`	Nombre total d'octets traités par l'équilibreur de charge.
Application Load Balancer	`RejectedConnectionCount`	Nombre de connexions rejetées parce que l'équilibreur de charge a atteint le nombre maximal de connexions.
Application Load Balancer	`RequestCount`	Le nombre de demandes traitées.
Application Load Balancer	`TargetConnectionErrorCount`	Nombre de connexions qui n'ont pas été établies avec succès entre l'équilibreur de charge et la cible.
Application Load Balancer	`TargetResponseTime`	Temps écoulé, en secondes, entre le moment où la demande a quitté l'équilibreur de charge et la réception d'une réponse de la cible.
Application Load Balancer	`UnHealthyHostCount`	Nombre de cibles considérées non saines.
Network Load Balancer	`ActiveFlowCount`	Le nombre total de TCP flux (ou connexions) simultanés entre les clients et les cibles.

Rubrique	Métrique	Description
Network Load Balancer	`ConsumedLCUs`	Le nombre d'unités de capacité de l'équilibreur de charge (LCU) utilisées par votre équilibreur de charge.
Network Load Balancer	`NewFlowCount`	Le nombre total de nouveaux TCP flux (ou connexions) établis entre les clients et les cibles au cours de la période.
Network Load Balancer	`ProcessedBytes`	Nombre total d'octets traités par l'équilibreur de charge, y compris les en-têtes TCP /IP.
Global Accelerator	`NewFlowCount`	Le nombre total de nouveaux UDP flux (ou connexions) établis entre les clients TCP et les points de terminaison au cours de la période.
Global Accelerator	`ProcessedBytesIn`	Nombre total d'octets entrants traités par l'accélérateur, y compris les en-têtes TCP /IP.
Auto Scaling	`GroupMaxSize`	Taille maximale du groupe Auto Scaling.
Amazon EC2	`CPUUtilization`	Pourcentage d'unités de EC2 calcul allouées actuellement utilisées.
Amazon EC2	`NetworkIn`	Nombre d’octets reçus par l’instance sur toutes les interfaces réseau.

Pour plus d'informations sur l'utilisation CloudWatch d'Amazon pour détecter DDoS les attaques visant votre application, consultez Getting Started with Amazon CloudWatch.

AWS inclut plusieurs mesures et alarmes supplémentaires pour vous avertir d'une attaque et pour vous aider à surveiller les ressources de votre application. La AWS Shield console ou API fournissez un résumé des événements par compte et des détails sur les attaques détectées.



Activité globale détectée par AWS Shield

En outre, le tableau de bord mondial de l'environnement des menaces fournit des informations récapitulatives sur toutes les DDoS attaques détectées par AWS. Ces informations peuvent être utiles pour mieux comprendre DDoS les menaces pesant sur un plus grand nombre d'applications, en plus de connaître les tendances en matière d'attaques et de les comparer aux attaques que vous avez pu observer.

Si vous êtes abonné AWS Shield Advanced, le tableau de bord du service affiche des mesures de détection et d'atténuation supplémentaires ainsi que des détails sur le trafic réseau pour les événements détectés sur les ressources protégées. AWS Shield évalue le trafic vers votre ressource protégée selon plusieurs dimensions. Lorsqu'une anomalie est détectée, AWS Shield crée un événement et indique la dimension du trafic dans laquelle l'anomalie a été observée. Grâce à une

atténuation placée, cela protège votre ressource contre le trafic excessif et le trafic correspondant à une signature d'DDoSévénement connue.

Les mesures de détection sont basées sur des flux réseau ou des AWS WAF journaux échantillonnés lorsqu'un site Web ACL est associé à la ressource protégée. Les mesures d'atténuation sont basées sur le trafic observé par les systèmes DDoS d'atténuation de Shield. Les mesures d'atténuation sont une mesure plus précise du trafic vers votre ressource.

La métrique des principaux contributeurs du réseau fournit un aperçu de la provenance du trafic lors d'un événement détecté. Vous pouvez afficher les contributeurs les plus importants et les trier par aspects tels que le protocole, le port source et les TCP indicateurs. La métrique des principaux contributeurs inclut des mesures pour l'ensemble du trafic observé sur la ressource selon différentes dimensions. Il fournit des dimensions métriques supplémentaires que vous pouvez utiliser pour comprendre le trafic réseau envoyé à votre ressource lors d'un événement. N'oubliez pas que dans le cas des attaques non réfléchissantes de couche 3 ou 4, les adresses IP sources peuvent avoir été falsifiées et ne sont pas fiables.

Le tableau de bord du service inclut également des détails sur les mesures prises automatiquement pour atténuer les DDoS attaques. Ces informations permettent d'étudier plus facilement les anomalies, d'explorer les dimensions du trafic et de mieux comprendre les mesures prises par Shield Advanced pour protéger votre disponibilité.

Journalisation

Activez une journalisation utile sur tous les services conformément à notre guide de journalisation et de surveillance destiné aux propriétaires d'applications afin d'optimiser la visibilité et de faciliter le dépannage. Cela inclut, mais sans s'y limiter :

- AWS CloudTrail
- AWS WAF Journaux
- CloudFrontjournaux d'accès
- VPCJournaux de flux (voir Enregistrer et afficher les flux de trafic réseau) : incluez un `tcp-flags` champ dans les champs inclus pour optimiser la visibilité
- ELBjournaux d'accès (ALB, CLB, NLB)
- HTTPJournaux d'accès au serveur Web
- Journalisation de sécurité du système d'exploitation

- Journalisation des applications

Gestion de la visibilité et de la protection sur plusieurs comptes

Dans les scénarios où vous opérez sur plusieurs composants Comptes AWS et que vous devez protéger plusieurs composants, l'utilisation de techniques qui vous permettent d'opérer à grande échelle et de réduire les frais opérationnels augmente vos capacités d'atténuation. Lorsque vous gérez des ressources AWS Shield Advanced protégées dans plusieurs comptes, vous pouvez configurer une surveillance centralisée à l'aide de AWS Firewall Manager et AWS Security Hub. Avec Firewall Manager, vous pouvez créer une politique de sécurité qui assure DDoS la conformité de tous vos comptes en matière de protection. Vous pouvez utiliser ces deux services ensemble pour gérer vos ressources protégées sur plusieurs comptes et centraliser la surveillance de ces ressources.

Security Hub s'intègre automatiquement à Firewall Manager, permettant aux clients de Shield Advanced de consulter les résultats de sécurité dans un tableau de bord unique, ainsi que les autres alertes de sécurité prioritaires et les états de conformité.

Par exemple, lorsque Shield Advanced détecte un trafic anormal destiné à une ressource protégée Compte AWS dans l'une des zones concernées, ce résultat sera visible dans la console Security Hub. Si cette configuration est configurée, Firewall Manager peut automatiquement mettre la ressource en conformité en la créant en tant que ressource protégée par le Shield Advanced, puis en mettant à jour Security Hub lorsque l'état de la ressource est conforme.

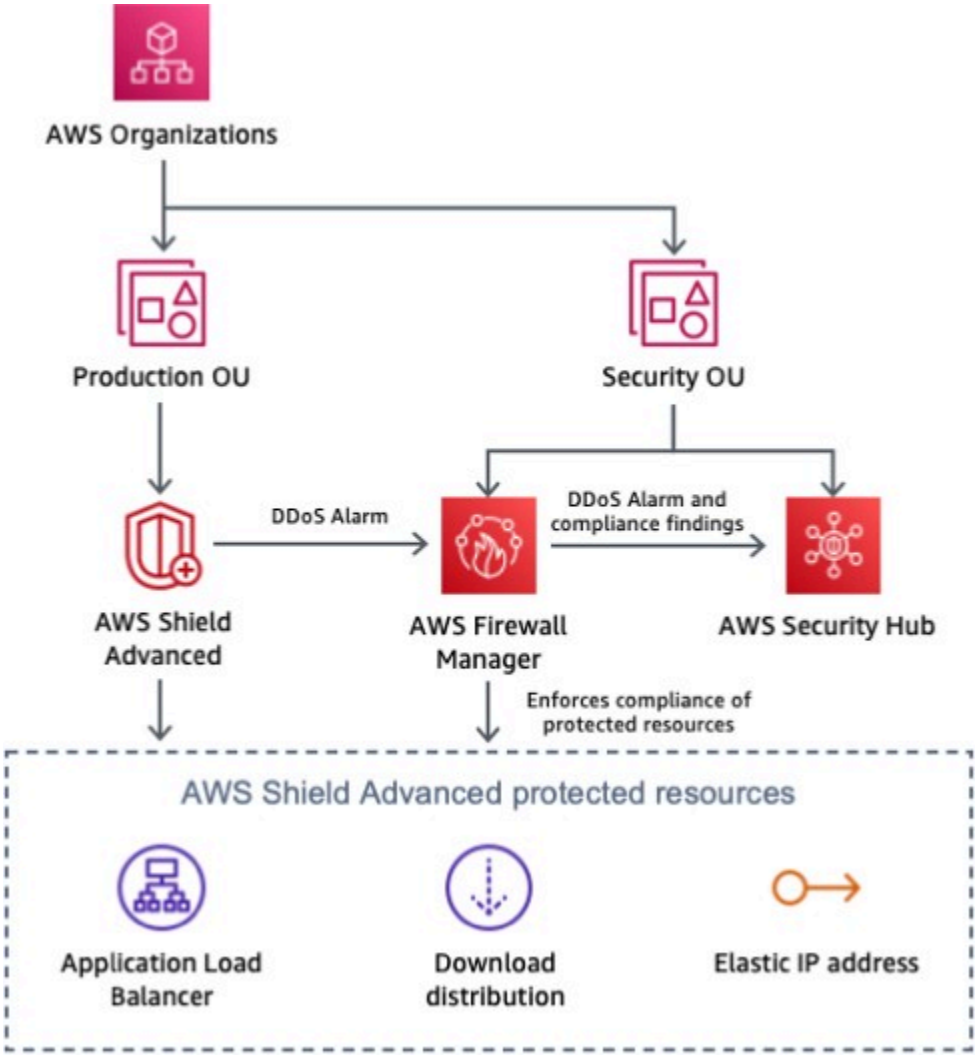


Schéma d'architecture illustrant la surveillance des ressources AWS Shield protégées avec Firewall Manager et Security Hub

Pour plus d'informations sur la surveillance centralisée des ressources protégées par le Shield, reportez-vous à la section Configuration de la surveillance centralisée des DDoS événements et correction automatique des ressources non conformes.

Stratégie de réponse aux incidents et manuels d'exécution

L'élaboration d'une stratégie de réponse aux incidents d'DDoSattaque et la mise en place d'un processus de réponse aux incidents de sécurité autour de cette stratégie sont essentiels pour toutes les organisations. Une approche recommandée consiste à modéliser votre plan de réponse en fonction des étapes suggérées, telles que NIST la collecte de preuves, l'atténuation, le rétablissement et la réalisation d'une analyse post-incident. Par exemple, un manuel de réponse aux attaques

par déni de service ou aux DDoS attaques d'applications Web est fourni à titre d'exemple. Des ressources supplémentaires sont disponibles dans le Guide de réponse aux incidents de AWS sécurité.

Support

Si vous êtes victime d'une attaque, vous pouvez également bénéficier d' AWS une assistance pour évaluer la menace et revoir l'architecture de votre application, ou vous pouvez demander une autre assistance. Il est important de créer un plan de réponse aux DDoS attaques avant qu'un événement ne se produise. Les meilleures pratiques décrites dans ce paper sont destinées à être des mesures proactives que vous mettez en œuvre avant de lancer une application, mais des DDoS attaques contre votre application peuvent tout de même se produire. Passez en revue les options de cette section pour déterminer les ressources de support les mieux adaptées à votre scénario. L'équipe chargée de votre compte peut évaluer votre cas d'utilisation et votre application, et vous aider à résoudre des questions ou à relever des défis spécifiques.

Si vous utilisez des charges de travail de production AWS, pensez à vous abonner au Business Support, qui vous permet d'accéder 24 heures sur 24 et 7 jours sur 7 à des ingénieurs de support cloud qui peuvent vous aider à résoudre les problèmes d'DDoSattaque. Si vous gérez des charges de travail critiques, optez pour le support aux entreprises, qui permet d'ouvrir les dossiers critiques et de recevoir la réponse la plus rapide d'un ingénieur senior du support cloud.

Si vous êtes abonné AWS Shield Advanced et êtes également abonné à Business Support ou Enterprise Support, vous pouvez configurer l'engagement proactif de Shield. Il vous permet de configurer des bilans de santé, de les associer à vos ressources et de fournir les coordonnées des opérations 24 heures sur 24, 7 jours sur 7. Lorsque Shield détecte des signes de dégradation DDoS et que les tests de santé de votre application montrent des signes de dégradation, AWS SRT il vous contacte de manière proactive. Il s'agit du modèle d'engagement que nous recommandons, car il permet d'obtenir les meilleurs délais de AWS SRT réponse et de AWS SRT commencer le dépannage avant même que le contact n'ait été établi avec vous.

Pour plus d'informations, reportez-vous à la section Comparer Support les forfaits.

La fonctionnalité d'engagement proactif vous oblige à configurer un bilan de santé Route 53 qui mesure avec précision l'état de santé de votre application et qui est associé à la ressource protégée par Shield Advanced. Une fois qu'un bilan de santé Route 53 est associé à la console Shield, le système de détection Shield Advanced utilise l'état du bilan de santé comme indicateur de l'état de santé de votre application. La fonction de détection basée sur l'état de santé de Shield Advanced

garantit que vous êtes averti et que des mesures d'atténuation sont prises plus rapidement lorsque votre application ne fonctionne pas correctement. AWS SRTvous contactera pour déterminer si l'application défectueuse est ciblée par une DDoS attaque et pour mettre en place des mesures d'atténuation supplémentaires si nécessaire.

L'achèvement de la configuration de l'engagement proactif inclut l'ajout de coordonnées dans la console Shield. AWS SRTutilisera ces informations pour vous contacter. Vous pouvez configurer jusqu'à dix contacts et fournir des notes supplémentaires si vous avez des exigences ou des préférences spécifiques en matière de contact. Proactif

les contacts d'engagement doivent occuper un rôle 24 heures sur 24, 7 jours sur 7, par exemple en tant que centre des opérations de sécurité ou en tant que personne immédiatement disponible.

Vous pouvez activer un engagement proactif pour toutes les ressources ou pour certaines ressources de production clés pour lesquelles le temps de réponse est essentiel. Cela se fait en attribuant des bilans de santé uniquement à ces ressources.

Vous pouvez également passer à Support AWS SRT en créant un Support dossier à l'aide de la Support console (connexion requise) ou au Support API si un événement DDoS lié à votre application affecte la disponibilité de votre application.

Conclusion

Les meilleures pratiques décrites dans ce paper peuvent vous aider à créer une architecture DDoS résiliente qui protège la disponibilité de votre application en empêchant de nombreuses DDoS attaques courantes liées à l'infrastructure et à la couche applicative. La mesure dans laquelle vous suivez ces bonnes pratiques lors de la conception de votre application influencera le type, le vecteur et le volume des DDoS attaques que vous pourrez atténuer. Vous pouvez intégrer la résilience sans souscrire à un service d'DDoSatténuation. En choisissant de AWS Shield Advanced vous abonner, vous bénéficiez de fonctionnalités supplémentaires de support, de visibilité, d'atténuation et de protection des coûts qui protègent davantage une architecture d'application déjà résiliente.

Collaborateurs

Les personnes qui ont contribué à ce document incluent :

- Rodrigo Ferroni, spécialiste de la sécurité AWS TAM
- Dmitriy Novikov, architecte de solutions AWS
- Achraf Souk, AWS architecte de solutions
- Joanna Knox, Ingénierie Support
- Anuj Butail, AWS architecte de solutions
- Harith Gaddamanugu, spécialiste d'Edge SA AWS

Suggestions de lecture

Pour plus d'informations, reportez-vous à :

- Directives de mise en œuvre AWS WAF (AWS livre blanc)
- NIS301 — Re:inForce 2023 : Comment les informations sur les AWS menaces deviennent des règles de pare-feu gérées (vidéo) YouTube
- NET314- re:INVENT 2022 : Création d'applications DDoS résilientes à l'aide AWS Shield de (vidéo) YouTube
- SEC321- re:Invent 2020 : prenez une longueur d'avance grâce aux escalades des équipes d'DDoSintervention (vidéo) YouTube
- William Hill : DDoS protection haute performance avec AWS - 2020 (YouTubevidéo)
- SEC407 - re:Invent 2019 : une defense-in-depth approche pour créer des applications Web (vidéo) YouTube
- Meilleures pratiques d'DDoSatténuation en AWS 2018 (YouTubevidéo)
- SID324— re:Invent 2017 : Automatiser les DDoS réponses dans le cloud (vidéo) YouTube
- CTD304 — Re:Invent 2017 : Le parcours du Dow Jones et du Wall Street Journal pour gérer les pics de trafic tout en même temps (vidéo) YouTube
- Atténuation DDoS et menaces au niveau de la couche applicative (YouTube vidéo)
- CTD310 — Re:Invent 2017 : Vivre à la limite de la vie est plus sûr que vous ne le pensez ! Renforcer ses capacités avec Amazon (YouTube vidéo)
- CloudFront AWS Shield, et AWS WAF (YouTube vidéo)

Révisions du document

Pour être informé des mises à jour de ce livre blanc, abonnez-vous au fil d'actualité. RSS

Modification	Description	Date
Mise à jour du livre blanc	Ajouté OAC pour une CloudFront protection DNS générique des coûts. Discussion approfondie des techniques opérationnelles, de la mise en cache, des règles basées sur le débit et des groupes de règles gérés. Ajout d'éléments locaux dans le schéma d'architecture, suppression de la duplication et clarification du texte pour éliminer toute ambiguïté.	9 août 2023
Mise à jour du livre blanc	Révisé pour plus de clarté ; mis à jour pour inclure les dernières recommandations et fonctionnalités : suivi des connexions des groupes de sécurité et DDoS atténuati on automatique de la couche d'application Shield Advanced.	13 avril 2022
Mise à jour du livre blanc	Mis à jour pour inclure les dernières recommandations et fonctionnalités. AWS Global Accelerator est ajouté dans le cadre d'une protection complète à la périphérie. AWS Firewall Manager pour une surveillance centralisée	21 septembre 2021

	des DDoS événements et la correction automatique des ressources non conformes.	
Mise à jour du livre blanc	Mise à jour pour clarifier le découpage du cache dans la section Détecter et filtrer les requêtes Web malveilla ntes (BP1,BP2) ELB et son ALB utilisation dans la section Scale to Absorb (BP6). Diagrammes mis à jour et tableau 2, intitulé « Choix de la région ». commeBP8. BP7Section mise à jour avec plus de détails.	18 décembre 2019
Mise à jour du livre blanc	Mise à jour pour inclure la AWS WAF journalisation en tant que meilleure pratique.	1 décembre 2018
Mise à jour du livre blanc	Mis à jour pour inclure AWS Shield les AWS WAF AWS Firewall Manager fonctionn alités et les meilleures pratiques associées.	1er juin 2018
Mise à jour du livre blanc	Ajout de directives d'archite cture prescriptive et mises à jour pour les inclure. AWS WAF	1 juin 2016
Publication initiale	Livre blanc publié.	le 1 juin 2015

Avis

Il incombe aux clients de procéder à une évaluation indépendante des informations contenues dans le présent document. Ce document : (a) est fourni à titre informatif uniquement, (b) représente les offres de AWS produits et les pratiques actuelles, qui sont susceptibles d'être modifiées sans préavis, et (c) ne crée aucun engagement ni aucune assurance de la part de AWS ses filiales, fournisseurs ou concédants de licence. AWS les produits ou services sont fournis « tels quels » sans garanties, déclarations ou conditions d'aucune sorte, qu'elles soient explicites ou implicites. Les responsabilités et obligations AWS de ses clients sont régies par AWS des accords, et ce document ne fait partie d'aucun accord conclu entre AWS et ses clients et ne les modifie pas.

AWS Glossaire

Pour la AWS terminologie la plus récente, consultez le AWS glossaire dans la Glossaire AWS référence.

Les traductions sont fournies par des outils de traduction automatique. En cas de conflit entre le contenu d'une traduction et celui de la version originale en anglais, la version anglaise prévaudra.