Whitepaper AWS

Best practice AWS per la resilienza DDoS

Copyright © Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Best practice AWS per la resilienza DDoS: Whitepaper AWS

Copyright © Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

I marchi e il trade dress di Amazon non possono essere utilizzati in relazione a prodotti o servizi che non siano di Amazon, in qualsiasi modo che possa causare confusione tra i clienti o in qualsiasi modo che denigri o discrediti Amazon. Tutti gli altri marchi non di proprietà di Amazon sono di proprietà dei rispettivi proprietari, che possono o meno essere affiliati, collegati o sponsorizzati da Amazon.

Table of Contents

AWS Best Practices for DDoS Resiliency

Data di pubblicazione: 21 settembre 2021 (Revisioni del documento)

Riassunto

È importante proteggere la tua azienda dall'impatto degli attacchi Distributed Denial of Service (DDoS) e di altri attacchi informatici. Mantenere la fiducia dei clienti nel tuo servizio mantenendo la disponibilità e la reattività della tua applicazione è una priorità assoluta. Vuoi anche evitare costi diretti non necessari quando la tua infrastruttura deve ridursi orizzontalmente in risposta a un attacco. Amazon Web Services (AWS) si impegna a fornire gli strumenti, le best practice e i servizi per difendersi dagli utenti malintenzionati su Internet. L'utilizzo dei servizi AWS giusti contribuisce a garantire alta disponibilità, sicurezza e resilienza.

In questo whitepaper, AWS fornisce una guida DDoS prescrittiva per migliorare la resilienza delle applicazioni in esecuzione su AWS. Ciò include un'architettura di riferimento resiliente agli attacchi DDoS che può essere utilizzata come guida per proteggere la disponibilità delle applicazioni. Questo whitepaper descrive anche diversi tipi di attacco, come gli attacchi a livello di infrastruttura e gli attacchi a livello di applicazione. AWS spiega quali sono le best practice più efficaci per gestire ogni tipo di attacco. Inoltre, vengono descritti i servizi e le caratteristiche che rientrano in una strategia di mitigazione degli attacchi DDoS e viene spiegato come ciascuno di essi può essere utilizzato per proteggere le applicazioni.

Questo documento è destinato ai responsabili delle decisioni IT e ai tecnici della sicurezza che hanno familiarità con i concetti di base di reti, sicurezza e AWS. Ogni sezione contiene collegamenti alla documentazione AWS che fornisce maggiori dettagli sulle best practice o funzionalità.

Introduzione: attacchi Denial of Service

Un attacco Denial of Service (DoS) è un tentativo deliberato di rendere un sito Web o un'applicazione non disponibile per gli utenti, ad esempio inondandolo di traffico di rete. Gli utenti malintenzionati utilizzano diverse tecniche che consumano grandi quantità di larghezza di banda di rete o vincolano altre risorse di sistema, interrompendo l'accesso per gli utenti legittimi. Nella sua forma più semplice, un utente malintenzionato solitario utilizza un'unica fonte per eseguire un attacco DoS contro un bersaglio, come mostrato nell'immagine seguente.

Tabella 1: Diagramma di un attacco DoS

In un attacco DDoS, un utente malintenzionato utilizza più fonti per orchestrare un attacco contro un bersaglio. Queste fonti possono includere gruppi distribuiti di computer, router, dispositivi IoT e altri endpoint infetti da malware. Il diagramma seguente mostra una rete di host compromessi che partecipa all'attacco, generando un flusso di pacchetti o richieste per sopraffare il bersaglio.

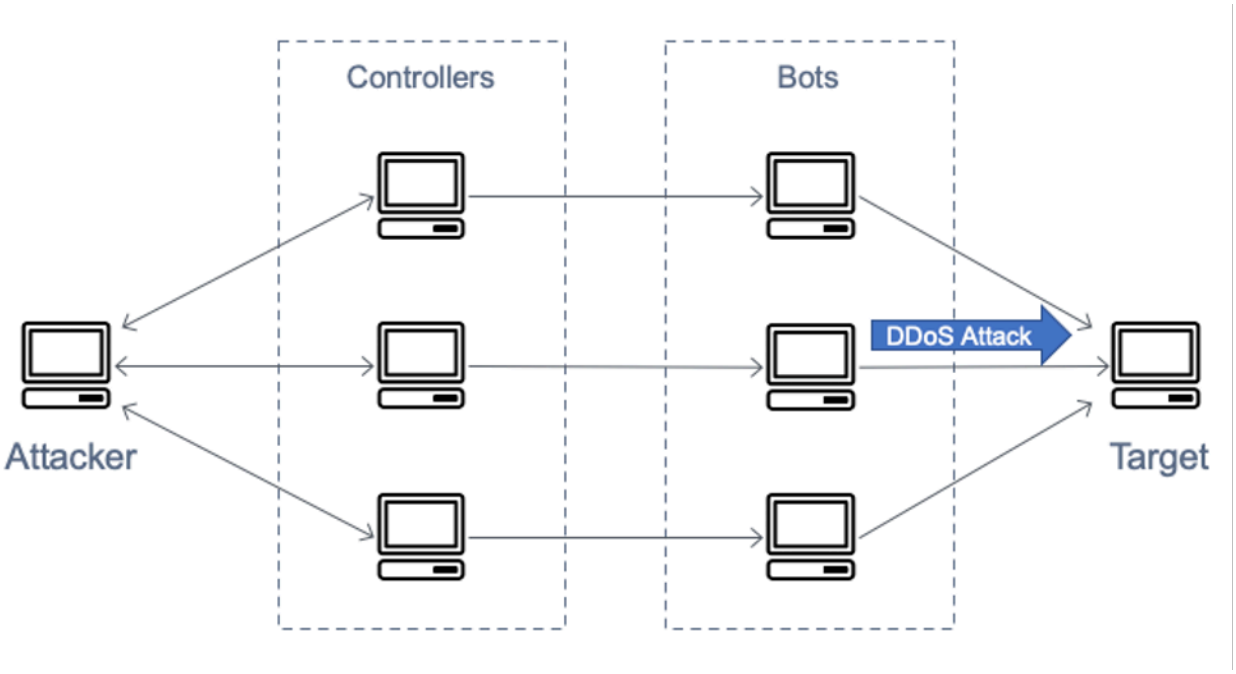

Diagramma di un attacco DDoS

Il modello OSI (Open Systems Interconnection) presenta sette livelli, descritti nella tabella delmodello OSI (Open Systems Interconnection). Gli attacchi DDoS sono più comuni ai livelli tre, quattro, sei e sette. Gli attacchi di livello tre e quattro corrispondono ai livelli di rete e trasporto del modello OSI. In questo documento,AWS li definisce collettivamente attacchi a livello di infrastruttura. Gli attacchi

di livello sei e sette corrispondono ai livelli di presentazione e applicazione del modello OSI. AWS li affronterà insieme come attacchi a livello di applicazione. Gli esempi di questi tipi di attacco sono discussi nelle sezioni seguenti.

Modello Open Systems Interconnection (OSI)

#	Livello	Unità	Descrizione	Esempi vettoriali
7	Applicazione	Dati	Processo di rete e applicazione	Flussi HTTP, flussi di query DNS
6	Presentazione	Dati	Rappresentazione di dati e crittografia	Uso illecito di TLS
5	Sessione	Dati	Comunicazione tra host	N/D
4	Trasporto	Segmenti	Connessioni end-to-end e affidabilità	SYN Flood
3	Rete	Pacchetti	Determinazione del percorso e indirizzamento logico	Attacchi di riflessione UDP
2	Collegamento dati	Fotogramma	Indirizzamento fisico	N/D
1	Fisica	Bit	Trasmissione media, segnale e binaria	N/D

Argomenti

Attacchi a livello di infrastruttura

Gli attacchi DDoS più comuni, gli attacchi di riflessione UDP (User Datagram Protocol) e i SYN Flood (Synchronize), sono attacchi a livello di infrastruttura. Un utente malintenzionato può utilizzare uno di questi metodi per generare grandi volumi di traffico in grado di inondare la capacità di una rete o bloccare risorse su sistemi come server, firewall, sistema di prevenzione delle intrusioni (IPS) o bilanciatore del carico. Sebbene questi attacchi possano essere facili da identificare, per mitigarli in modo efficace è necessario disporre di una rete o di sistemi che aumentino la capacità più rapidamente rispetto al flusso di traffico in entrata. Questa capacità aggiuntiva è necessaria per filtrare o assorbire il traffico di attacco, liberando il sistema e l'applicazione per rispondere al traffico legittimo dei clienti.

Argomenti

Attacchi di riflessione UDP

Gli attacchi di riflessione UDP (User Datagram Protocol) sfruttano il fatto che l'UDP è un protocollo senza stato. Gli aggressori possono creare un pacchetto di richieste UDP valido che elenca l'indirizzo IP del bersaglio dell'attacco come indirizzo IP dell'origine UDP. L'aggressore ha falsificato (sottoposto a spoofing) l'IP sorgente del pacchetto di richieste UDP. Il pacchetto UDP contiene l'IP sorgente falsificato e viene inviato dall'aggressore a un server intermedio. Il server viene indotto a inviare i suoi pacchetti di risposta UDP all'IP della vittima presa di mira piuttosto che all'indirizzo IP dell'aggressore. Il server intermedio viene utilizzato perché genera una risposta che è diverse volte più grande del pacchetto di richiesta, amplificando di fatto la quantità di traffico di attacco inviato all'indirizzo IP di destinazione.

Il fattore di amplificazione è il rapporto tra la dimensione della risposta e la dimensione della richiesta e varia a seconda del protocollo utilizzato dall'aggressore: DNS, NTP, SSDP, CLDAP, Memcached, CharGen o QOTD. Ad esempio, il fattore di amplificazione per il DNS può essere da 28 a 54 volte il numero originale di byte. Quindi, se un utente malintenzionato invia un payload di richiesta di 64 byte a un server DNS, può generare oltre 3400 byte di traffico indesiderato verso un obiettivo di attacco. Gli attacchi di riflessione UDP sono responsabili di un volume di traffico maggiore rispetto ad altri attacchi. La figura Attacco di riflessione UDP illustra la tattica di riflessione e l'effetto di amplificazione.

Attacchi di riflessione UDP

Attacchi SYN Flood

Quando un utente si connette a un servizio TCP (Transmission Control Protocol), ad esempio un server Web, il client invia un pacchetto di sincronizzazione SYN. Il server restituisce un pacchetto SYN-ACK di riconoscimento e infine il client risponde con un pacchetto di riconoscimento (ACK), che completa l'handshake a tre vie previsto. L'immagine seguente illustra questo tipico handshake.

Handshake a 3 vie SYN

In un attacco SYN Flood, un client malintenzionato invia un gran numero di pacchetti SYN, ma non invia mai i pacchetti ACK finali per completare gli handshake. Il server rimane in attesa di una risposta alle connessioni TCP semiaperte e alla fine esaurisce la capacità di accettare nuove connessioni TCP, impedendo ad altri utenti di connettersi al server. L'attacco sta tentando di bloccare le connessioni al server disponibili in modo che le risorse non siano disponibili per connessioni legittime. Anche se i SYN Flood possono raggiungere centinaia di Gbps, lo scopo dell'attacco non è quello di aumentare il volume di traffico SYN.

Attacchi a livello di applicazione

Un utente malintenzionato può colpire l'applicazione stessa utilizzando un attacco a livello 7 o a livello di applicazione. In questi attacchi, simili agli attacchi all'infrastruttura SYN Flood, l'utente malintenzionato tenta di sovraccaricare funzioni specifiche di un'applicazione per renderla non disponibile o non responsiva per gli utenti legittimi. Talvolta questo risultato può essere ottenuto con volumi di richieste molto bassi che generano solo un piccolo volume di traffico di rete, rendendo l'attacco difficile da rilevare e mitigare. Gli esempi di attacchi a livello di applicazione includono i flood HTTP, gli attacchi di cache-busting e i flood XML-RPC di WordPress.

In un attacco flood HTTP, un utente malintenzionato invia richieste HTTP che sembrano provenire da un utente valido dell'applicazione Web. Alcuni flood HTTP hanno come target una risorsa specifica, mentre i flood HTTP più complessi tentano di emulare l'interazione umana con l'applicazione, aumentando in questo modo la difficoltà di utilizzare tecniche di mitigazione comuni come la limitazione del tasso di richiesta.

Gli attacchi di cache-busting sono un tipo di flusso HTTP che utilizza variazioni nella stringa di query per aggirare la memorizzazione nella cache della rete di distribuzione dei contenuti (CDN). Invece

di essere in grado di restituire risultati memorizzati nella cache, la CDN deve contattare il server di origine per ogni richiesta di pagina e questi recuperi dell'origine causano un ulteriore sforzo sul server Web dell'applicazione.

Con un attacco flood XML-RPC di WordPress, noto anche come pingback flood di WordPress, un utente malintenzionato prende di mira un sito Web ospitato sul software di gestione dei contenuti WordPress. L'utente malintenzionato utilizza in modo improprio la funzione API XML-RPC per generare un flusso di richieste HTTP. La funzione pingback consente a un sito Web ospitato su WordPress (sito A) di notificare un diverso sito WordPress (sito B) tramite un collegamento che il sito A ha creato al sito B. Il sito B tenta quindi di recuperare il sito A per verificare l'esistenza del collegamento. In un pingback flood, l'utente malintenzionato abusa di questa capacità per fare in modo che il sito B attacchi il sito A. Questo tipo di attacco ha una firma chiara: WordPress è tipicamente presente nell'Utente-agente dell'intestazione della richiesta HTTP.

Esistono altre forme di traffico dannoso che possono influire sulla disponibilità di un'applicazione. I bot di scraping automatizzano i tentativi di accesso a un'applicazione Web per sottrarre contenuti o registrare informazioni sulla concorrenza, come i prezzi. Gli attacchi di forza bruta e di credential stuffing sono degli sforzi programmati per ottenere l'accesso non autorizzato alle aree protette di un'applicazione. Non si tratta esclusivamente di attacchi DDoS, ma la loro natura automatizzata può assomigliare a un attacco DDoS e possono essere mitigati implementando alcune delle stesse best practice che verranno trattate in questo documento.

Gli attacchi a livello di applicazione possono anche colpire i servizi DNS (Domain Name System). Il più comune di questi attacchi è un flusso di query DNS in cui un utente malintenzionato utilizza molte query DNS ben formate per esaurire le risorse di un server DNS. Questi attacchi possono anche includere un componente cache-busting in cui l'utente malintenzionato randomizza la stringa del sottodominio per bypassare la cache DNS locale di un determinato resolver. Di conseguenza, il resolver non può sfruttare le query di dominio memorizzate nella cache e deve invece contattare ripetutamente il server DNS autorevole, amplificando l'attacco.

Se un'applicazione Web viene distribuita tramite Transport Layer Security (TLS), un utente malintenzionato può anche scegliere di attaccare il processo di negoziazione TLS. TLS è costoso dal punto di vista computazionale, quindi un utente malintenzionato, generando un carico di lavoro aggiuntivo sul server per elaborare dati illeggibili (o incomprensibili (testo cifrato)) come un handshake legittimo, può ridurre la disponibilità del server. In una variante di questo attacco, un utente malintenzionato completa l'handshake TLS ma rinegozia continuamente il metodo di crittografia. In alternativa, un utente malintenzionato può tentare di esaurire le risorse del server aprendo e chiudendo molte sessioni TLS.

Tecniche di mitigazione

Alcune forme di mitigazione degli attacchi DDoS sono incluse automaticamente nei servizi AWS. La resilienza DDoS può essere ulteriormente migliorata utilizzando un'architettura AWS con servizi specifici, trattati nelle sezioni seguenti, e implementando best practice aggiuntive per ogni parte del flusso di rete tra gli utenti e l'applicazione.

Tutti i clienti AWS possono beneficiare della protezione automatica di AWS Shield Standard, senza costi aggiuntivi. AWS Shield Standard protegge dagli attacchi DDoS a livello di rete e di trasporto più comuni e frequenti a siti Web o applicazioni. Questa protezione è sempre attiva, preconfigurata, statica e non fornisce report o analisi dei dati. È offerto su tutti i servizi AWS e in ogni regione AWS. Nelle regioni AWS, gli attacchi DDoS vengono rilevati e il sistema Shield Standard classifica automaticamente il traffico, identifica le anomalie e, se necessario, crea mitigazioni. È possibile utilizzare AWS Shield Standard come parte integrante di un'architettura resiliente agli attacchi DDoS per proteggere le applicazioni Web e non Web.

Puoi anche utilizzare i servizi AWS che operano da posizioni edge, come Amazon CloudFront, Global Accelerator e Route 53 per creare una protezione completa della disponibilità contro tutti gli attacchi noti a livello di infrastruttura. Questi servizi fanno parte di AWS Global Edge Network e possono migliorare la resilienza DDoS della tua applicazione quando serve qualsiasi tipo di traffico applicativo da posizioni edge distribuite in tutto il mondo. Puoi eseguire l'applicazione in qualsiasi regione AWS e utilizzare questi servizi per proteggere la disponibilità dell'applicazione e ottimizzare le prestazioni della tua applicazione per utenti finali legittimi.

I vantaggi dell'utilizzo di Amazon CloudFront, Global Accelerator e Amazon Route 53 comprendono:

- Accesso a Internet e alla capacità di mitigazione degli attacchi DDoS sulla rete edge globale di AWS. Ciò è utile per mitigare gli attacchi volumetrici più grandi, che possono raggiungere una scala di terabit.
- I sistemi di mitigazione degli attacchi DDoS di AWS Shield sono integrati con i servizi edge di AWS, riducendo i tempi di mitigazione da alcuni minuti a meno di secondo.
- Le tecniche di mitigazione del SYN Flood senza stato usano dei proxy e verificano le connessioni in ingresso prima di passarle al servizio protetto. Ciò garantisce che solo le connessioni valide raggiungano l'applicazione proteggendo gli utenti finali legittimi dalle notifiche di falsi positivi.
- Sistemi automatici di ingegneria del traffico che disperdono o isolano l'impatto di attacchi DDoS volumetrici di grandi dimensioni. Tutti questi servizi isolano gli attacchi alla fonte prima che raggiungano la tua origine, con un impatto minore sui sistemi protetti da questi servizi.

- La difesa a livello di applicazione, se combinata con AWS WAF non richiede la modifica dell'architettura dell'applicazione corrente (ad esempio, in una regione AWS o in un datacenter On-Premise).

Non sono previsti costi per il trasferimento dei dati in entrata su AWS e allo stesso modo non si paga il traffico di attacchi DDoS mitigato da AWS Shield. Il seguente diagramma di architettura include i servizi della rete edge globale AWS.

Questa architettura include diversi servizi AWS che possono aiutarti a migliorare la resilienza della tua applicazione Web dagli attacchi DDoS. La tabella Riepilogo delle best practice fornisce un riepilogo di questi servizi e delle funzionalità che possono fornire. AWS ha taggato ogni servizio con un indicatore di best practice (BP1, BP2) per una più facile consultazione all'interno di questo documento. Ad esempio, una delle prossime sezioni illustra le funzionalità fornite da Amazon CloudFront e Global Accelerator, che include l'indicatore di best practice BP1.

Tabella 2 - Riepilogo delle best practice

AWS Edge	Regione AWS					
	Utilizzo di Amazon CloudFront (BP1) con AWS WAF (BP2)	Utilizzo di Global Accelerator (BP1)	Utilizzo di Amazon Route 53 (BP3)	Utilizzo di Elastic Load Balancing (BP6) con AWS WAF (BP2)	Utilizzo di gruppi di sicurezza e liste di controllo accessi di rete in Amazon VPC (BP5)	Utilizzo di Amazon EC2 Auto Scaling (BP7)
Mitigazione degli attacchi di livello 3 (ad esempio,	✓	✓	✓	✓	✓	✓

AWS Edge	Regione AWS					
riflessione UDP)						
Mitigazione degli attacchi di livello 4 (ad esempio SYN Flood)	✓	✓	✓	✓		
Mitigazione degli attacchi di livello 6 (ad esempio TLS)	✓	✓	✓	✓		
Riduzione della superficie d'attacco	✓	✓	✓	✓	✓	
Scalabilità per assorbire il traffico a livello applicativo	✓	✓	✓	✓	✓	✓

AWS Edge	Regione AWS					
Mitigazione degli attacchi di livello 7 (livello dell'applicazione)	✓	✓(*)	✓	✓	✓(*)	✓(*)
Isolamento geografico e dispersione del traffico in eccesso e attacchi DDoS di maggiori dimensioni	✓	✓	✓			
✓ (*): se utilizzato con AWS WAF con Application Load Balancer						

Un altro modo per migliorare la tua preparazione nel rispondere e mitigare gli attacchi DDoS è sottoscrivere AWS Shield Advanced.

I clienti ricevono un rilevamento personalizzato basato su:

- Modelli di traffico specifici della tua applicazione.
- Protezione dagli attacchi DDoS di livello 7 incluso AWS WAF senza costi aggiuntivi.

- Accesso al supporto specializzato 24x7 da AWS SRT.
- Gestione centralizzata delle policy di sicurezza tramite AWS Firewall Manager.
- Protezione dei costi per proteggersi dai costi di dimensionamento derivanti dai picchi di utilizzo correlati agli attacchi DDoS.

Questo servizio opzionale di mitigazione degli attacchi DDoS aiuta a proteggere le applicazioni ospitate in qualsiasi regione AWS. Il servizio è disponibile a livello globale per CloudFront, Route 53 e Global Accelerator. L'uso di Shield Advanced con indirizzi IP elastici consente di proteggere le istanze Network Load Balancer (NLBS) o Amazon EC2.

I vantaggi dell'utilizzo di AWS Shield Advanced includono:

- Accesso a AWS SRT per assistenza nella mitigazione degli attacchi DDoS che influiscono sulla disponibilità delle applicazioni.
- Visibilità degli attacchi DDoS tramite l'utilizzo di AWS Management Console, dell'API, dei parametri e degli allarmi di Amazon CloudWatch.
- Accesso alla cronologia di tutti gli eventi DDoS degli ultimi 13 mesi.
- Accesso al firewall per applicazioni Web di AWS (AWS WAF) senza costi aggiuntivi per la mitigazione degli attacchi DDoS a livello di applicazione (se utilizzato con Amazon CloudFront o Application Load Balancer).
- Linea di base automatica degli attributi del traffico Web, se utilizzato con AWS WAF.
- Accesso a AWS Firewall Manager senza costi aggiuntivi, per l'applicazione automatica delle policy.
- Soglie di rilevamento sensibili che instradano il traffico nel sistema di mitigazione degli attacchi DDoS in anticipo e possono migliorare i tempi di mitigazione degli attacchi contro Amazon EC2 o Network Load Balancer, se utilizzati con un indirizzo IP elastico.
- Protezione dei costi che consente di richiedere un rimborso limitato dei costi relativi alla scalabilità derivanti da un attacco DDoS.
- Accordo sul Livello di Servizio (SLA) avanzato specifico per i clienti di AWS Shield Advanced.
- Coinvolgimento proattivo da parte di AWS SRT quando viene rilevato un evento Shield.
- Gruppi di protezione che consentono di raggruppare le risorse, offrendo una modalità self-service per personalizzare l'ambito del rilevamento e della mitigazione per l'applicazione trattando più risorse come una singola unità. Il raggruppamento delle risorse migliora l'accuratezza del rilevamento, riduce al minimo i falsi positivi, facilita la protezione automatica delle risorse appena create e accelera il tempo necessario per mitigare gli attacchi contro molte risorse che

comprendono una singola applicazione. Per informazioni sui gruppi di protezione, consulta Shield Advanced protection groups.

Per un elenco completo delle caratteristiche di AWS Shield Advanced e per ulteriori informazioni suAWS Shield, consulta Come funziona AWS Shield.

Argomenti

Best practice per la mitigazione DDoS

Nelle sezioni seguenti viene descritta in modo più approfondito ciascuna delle best practice consigliate per la mitigazione degli attacchi DDoS. Per una guida rapida e facile da implementare sulla creazione di un livello di mitigazione DDoS per applicazioni Web statiche o dinamiche, consulta Come migliorare la protezione di applicazioni Web dinamiche da attacchi di tipo DDoS.

Difesa a livello di infrastruttura (BP1, BP3, BP6, BP7)

In un ambiente di datacenter tradizionale, è possibile mitigare gli attacchi DDoS a livello di infrastruttura utilizzando tecniche come l'overprovisioning della capacità, l'implementazione di sistemi di mitigazione DDoS o lo scrubbing del traffico con l'aiuto di servizi di mitigazione DDoS. Su AWS le funzionalità di mitigazione degli attacchi DDoS vengono fornite automaticamente, ma puoi ottimizzare la resilienza DDoS della tua applicazione effettuando scelte di architettura che sfruttino al meglio tali funzionalità e consentano anche di dimensionare il traffico in eccesso.

Le considerazioni più importanti per contribuire a mitigare gli attacchi DDoS volumetrici comprendono la garanzia che siano disponibili sufficienti capacità di transito e diversità, oltre alla protezione delle risorse AWS, come le istanze Amazon EC2, dal traffico degli attacchi.

Alcuni tipi di istanze Amazon EC2 supportano funzionalità in grado di gestire più facilmente grandi volumi di traffico, ad esempio interfacce con larghezza di banda di rete fino a 100 Gbps e reti avanzate. Questo aiuta a prevenire la congestione dell'interfaccia per il traffico che ha raggiunto l'istanza Amazon EC2. Le istanze che supportano reti avanzate forniscono prestazioni I/O più elevate, maggiore larghezza di banda e minore utilizzo della CPU rispetto alle implementazioni

tradizionali. In questo modo migliora la capacità dell'istanza di gestire grandi volumi di traffico e, in ultima analisi, la rende altamente resiliente rispetto al carico di pacchetti al secondo (pps).

Per consentire questo elevato livello di resilienza, AWS consiglia di utilizzare istanze dedicate di Amazon EC2 o istanze Amazon EC2 con velocità effettiva di rete più elevata con suffisso N e supporto per reti avanzate con un massimo di 100 Gbps di larghezza di banda di rete, ad esempio c6gn.16xlarge e c5n.18xlarge o istanze bare metal (come c5n.metal).

Per ulteriori informazioni sulle istanze Amazon EC2 che supportano interfacce di rete da 100 Gigabit e reti avanzate, consulta Tipi di istanza di Amazon EC2.

Il modulo richiesto per le reti avanzate e il set di attributi enaSupport richiesto sono inclusi in Amazon Linux 2 e nelle versioni più recenti dell'AMI Amazon Linux. Pertanto, se si avvia un'istanza con una versione HVM di Amazon Linux su un tipo di istanza supportato, l'istanza dispone già dell'abilitazione delle reti avanzate. Per ulteriori informazioni, consulta Verifica dell'abilitazione delle reti avanzate. Per ulteriori informazioni su come abilitare le reti avanzate, consulta Reti avanzate su Linux.

Amazon EC2 con Auto Scaling (BP7)

Un altro modo per mitigare gli attacchi a livello di infrastruttura e applicazione è operare su vasta scala. Se disponi di applicazioni Web, puoi utilizzare i bilanciatori del carico per distribuire il traffico verso un numero di istanze Amazon EC2 con overprovisioning o configurate per il dimensionamento automatico. Queste istanze sono in grado di gestire picchi di traffico improvvisi che si verificano per qualsiasi motivo, tra cui un flash crowd o un attacco DDoS a livello di applicazione. Puoi impostare gli allarmi di Amazon CloudWatch per avviare Auto Scaling e scalare automaticamente le dimensioni del parco istanze di Amazon EC2 in risposta agli eventi definiti, come CPU, RAM, I/O di rete e anche parametri personalizzati. Questo approccio protegge la disponibilità delle applicazioni in caso di aumento imprevisto del volume delle richieste. Quando si utilizzano Amazon CloudFront, Application Load Balancer, Classic Load Balancer o Network Load Balancer con l'applicazione, la negoziazione TLS viene gestita dalla distribuzione (Amazon CloudFront) o dal bilanciatore del carico. Queste caratteristiche consentono di proteggere le tue istanze dall'impatto degli attacchi basati su TLS ridimensionandole per gestire richieste legittime e attacchi con uso illecito del TLS.

Per ulteriori informazioni sull'utilizzo di Amazon CloudWatch per richiamare Auto Scaling, consulta Monitoraggio dei parametri di Amazon CloudWatch per i gruppi e le istanze Auto Scaling.

Amazon EC2 offre capacità di calcolo scalabile in modo che tu possa aumentare o ridurre le risorse in base al cambiamento dei requisiti. Puoi dimensionare orizzontalmente aggiungendo

automaticamente istanze alla tua applicazione ridimensionando le dimensioni del tuo gruppo Amazon EC2 Auto Scaling e puoi dimensionare verticalmente utilizzando tipi di istanze EC2 più grandi.

Elastic Load Balancing (BP6)

Gli attacchi DDoS di grandi dimensioni possono sovraccaricare la capacità di una singola istanza Amazon EC2. Con Elastic Load Balancing (ELB), puoi ridurre il rischio di sovraccaricare la tua applicazione distribuendo il traffico su molte istanze di back-end. Elastic Load Balancing è in grado di dimensionare automaticamente, consentendoti di gestire volumi più grandi quando hai traffico supplementare imprevisto, come ad esempio a causa di flash crowd o attacchi DDoS. Per le applicazioni create all'interno di un Amazon VPC, ci sono tre tipi di ELB da considerare, a seconda del tipo di applicazione: Application Load Balancer (ALB), Classic Load Balancer (CLB) e Network Load Balancer (NLB).

Per le applicazioni Web è possibile utilizzare Application Load Balancer per instradare il traffico in base al contenuto e accettare solo richieste Web ben formate. Application Load Balancer blocca molti attacchi DDoS comuni, come i SYN Flood o gli attacchi di riflessione UDP, proteggendo la tua applicazione dagli attacchi. Application Load Balancer si dimensiona automaticamente per assorbire il traffico aggiuntivo quando vengono rilevati questi tipi di attacchi. Le attività di dimensionamento dovute agli attacchi a livello di infrastruttura sono trasparenti per i clienti di AWS e non influiscono sulla fattura.

Per ulteriori informazioni sulla protezione delle applicazioni Web con Application Load Balancer, consulta Nozioni di base di Application Load Balancer

Per le applicazioni basate su TCP puoi utilizzare Network Load Balancer per instradare il traffico verso destinazioni a latenza estremamente bassa (come ad esempio le istanze Amazon EC2) . Una considerazione chiave con Network Load Balancer è che qualsiasi traffico che raggiunge il bilanciatore del carico su un listener valido verrà instradato verso le destinazioni, non assorbito. È possibile utilizzare Shield Advanced per configurare la protezione DDoS per gli indirizzi IP elastici. Quando viene assegnato un indirizzo IP elastico per zona di disponibilità al Network Load Balancer, Shield Advanced applicherà le protezioni DDoS pertinenti per il traffico di Network Load Balancer.

Per ulteriori informazioni sulla protezione delle applicazioni TCP con Network Load Balancer, consulta Nozioni di base sui sistemi Network Load Balancer

Sfrutta le posizioni edge di AWS per il dimensionamento (BP1, BP3)

L'accesso a connessioni Internet diversificate e altamente dimensionate può aumentare significativamente la capacità di ottimizzare la latenza e la velocità effettiva per gli utenti, assorbire gli attacchi DDoS e isolare i guasti riducendo al minimo l'impatto sulla disponibilità delle applicazioni. Le posizioni edge di AWS forniscono un ulteriore livello di infrastruttura di rete che offre questi vantaggi a qualsiasi applicazione Web che utilizza Amazon CloudFront, Global Accelerator e Amazon Route 53. Con questi servizi, puoi proteggere in modo completo all'edge le tue applicazioni in esecuzione da regioni AWS.

Distribuzione di applicazioni Web all'edge (BP1)

Amazon CloudFront è un servizio che può essere utilizzato per distribuire l'intero sito Web, inclusi contenuti statici, dinamici, in streaming e interattivi. Le connessioni persistenti e le impostazioni di durata (TTL) variabile possono essere utilizzate per scaricare il traffico dalla tua origine, anche se non stai servendo contenuti memorizzabili nella cache. L'uso di queste caratteristiche di CloudFront riduce il numero di richieste e connessioni TCP all'origine, contribuendo a proteggere la tua applicazione Web dagli attacchi HTTP Flood. CloudFront accetta solo connessioni ben formate, contribuendo a impedire che molti attacchi DDoS comuni, come i SYN Flood e gli attacchi di riflessione UDP, raggiungano la tua origine. Inoltre, gli attacchi DDoS vengono geograficamente isolati in prossimità dell'origine, impedendo al traffico di influenzare altre sedi. Queste funzionalità possono migliorare notevolmente la possibilità di continuare a distribuire il traffico agli utenti durante attacchi DDoS di grandi dimensioni. Puoi usare CloudFront per proteggere un'origine su AWS o altrove su Internet.

Se utilizzi Amazon S3 per pubblicare contenuti statici su Internet, AWS consiglia di utilizzare Amazon CloudFront per proteggere il tuo bucket. Puoi utilizzare un'identità di accesso origine (OAI) per garantire che gli utenti accedano ai tuoi oggetti utilizzando solo gli URL di CloudFront.

Per ulteriori informazioni sull'OAI, consulta Limitazione dell'accesso a contenuti Amazon S3 utilizzando un'identità di accesso origine.

Per ulteriori informazioni sulla protezione e l'ottimizzazione delle prestazioni delle applicazioni Web con Amazon CloudFront, consulta Nozioni di base di CloudFront.

Proteggi ulteriormente il traffico di rete dalla tua origine utilizzando AWS Global Accelerator (BP1)

Global Accelerator è un servizio di rete che migliora la disponibilità e le prestazioni del traffico degli utenti fino al 60%. Questa operazione viene eseguita con l'ingresso del traffico nella posizione edge più vicina agli utenti e il routing attraverso l'infrastruttura di rete globale di AWS verso la tua applicazione, indipendentemente dal fatto che venga eseguita in una o più regioni AWS.

Global Accelerator instrada il traffico TCP e UDP verso l'endpoint ottimale in base alle prestazioni nella regione AWS più vicina all'utente. Se si verifica un fallimento dell'applicazione, Global Accelerator fornisce il failover all'endpoint migliore successivo entro 30 secondi. Global Accelerator utilizza la grande capacità della rete globale di AWS e le integrazioni con Shield, come una funzionalità proxy SYN senza stato che sfida i nuovi tentativi di connessione e serve solo gli utenti finali legittimi per proteggere le applicazioni.

È possibile implementare un'architettura resiliente agli attacchi DDoS che offre molti degli stessi vantaggi delle best practice per la distribuzione di applicazioni Web all'edge, anche se l'applicazione utilizza protocolli non supportati da CloudFront o se si utilizza un'applicazione Web che richiede indirizzi IP statici globali. Ad esempio, potresti richiedere indirizzi IP che i tuoi utenti finali possono aggiungere all'elenco dei permessi nei loro firewall e non vengono utilizzati da altri clienti di AWS. In questi scenari è possibile utilizzare Global Accelerator per proteggere le applicazioni Web in esecuzione su Application Load Balancer e in combinazione con AWS WAF per rilevare e mitigare anche i flussi di richieste a livello di applicazione Web.

Per ulteriori informazioni sulla protezione e l'ottimizzazione delle prestazioni del traffico di rete utilizzando Global Accelerator, consulta Nozioni di base di Global Accelerator.

Risoluzione dei nomi di dominio a livello di edge (BP3)

Amazon Route 53 è un servizio Domain Name System (DNS) altamente disponibile e scalabile che può essere utilizzato per indirizzare il traffico verso la tua applicazione Web. Include caratteristiche avanzate come Traffic Flow, Controllo dell'integrità e monitoraggio, Routing basato sulla latenza e Geo DNS. Queste caratteristiche avanzate ti consentono di controllare come il servizio risponde alle richieste DNS per migliorare le prestazioni della tua applicazione Web ed evitare interruzioni del sito.

Amazon Route 53 utilizza tecniche come lo sharding casuale e l'anycast striping, in grado di aiutare gli utenti ad accedere alla tua applicazione anche se il servizio DNS è preso di mira da un attacco DDoS.

Con lo sharding casuale, ogni server dei nomi nel tuo set di delega corrisponde a un insieme unico di posizioni edge e percorsi internet, offrendo una maggiore tolleranza ai guasti e riducendo al minimo la sovrapposizione tra i clienti. Se un server dei nomi nel set di delega non è disponibile, gli utenti possono riprovare e ricevere una risposta da un altro server in una posizione edge diversa.

L'anycast striping consente a ciascuna richiesta DNS di essere fornita dalla posizione più ottimale, disperdendo il carico di rete e riducendo la latenza DNS. In questo modo gli utenti ricevono una risposta più rapida. Inoltre, Amazon Route 53 è in grado di rilevare anomalie alla fonte e nel volume delle query DNS e assegnare priorità alle richieste di utenti noti per essere affidabili.

Per ulteriori informazioni sull'utilizzo di Amazon Route 53 per indirizzare gli utenti alla tua applicazione, consulta Nozioni di base di Amazon Route 53.

Difesa a livello di applicazione (BP1, BP2)

Molte delle tecniche illustrate finora in questo documento sono efficaci nel mitigare l'impatto degli attacchi DDoS a livello di infrastruttura sulla disponibilità delle applicazioni. Per difendersi anche dagli attacchi a livello di applicazione, è necessario implementare un'architettura che consenta di rilevare in modo specifico e dimensionare per assorbire e bloccare le richieste dannose. Questa è una considerazione importante perché i sistemi di mitigazione degli attacchi DDoS basati sulla rete sono generalmente inefficaci nel mitigare gli attacchi complessi a livello di applicazione.

Rileva e filtra le richieste Web dannose (BP1, BP2)

Quando la tua applicazione esegue AWS, puoi sfruttare sia Amazon CloudFront che AWS WAF per difenderti dagli attacchi DDoS a livello di applicazione.

Amazon CloudFront ti consente di memorizzare nella cache i contenuti statici e di distribuirli dalle posizioni edge di AWS, contribuendo a ridurre il carico sulla tua origine. Può anche aiutare a ridurre il carico del server impedendo al traffico non Web di raggiungere la tua origine. Inoltre, CloudFront può chiudere automaticamente le connessioni da utenti malintenzionati a lettura lenta o a scrittura lenta (ad esempio, Slowloris).

UtilizzandoAWS WAF, è possibile configurare le liste di controllo accessi Web (ACL Web) nelle distribuzioni CloudFront o negli Application Load Balancer per filtrare e bloccare le richieste in base alle firme delle richieste. Ogni ACL Web è costituito da regole che è possibile configurare per la corrispondenza stringa o regex con uno o più attributi di richiesta, ad esempio l'Uniform Resource Identifier (URI), la stringa di query, il metodo HTTP o la chiave di intestazione. Inoltre, utilizzando le

regole basate sulla frequenza di AWS WAF, è possibile bloccare automaticamente gli indirizzi IP degli utenti malintenzionati quando le richieste corrispondenti a una regola superano una soglia definita.

Le richieste provenienti da indirizzi IP client dannosi riceveranno le risposte di errore 403 Accesso negato e rimarranno bloccate fino a quando le frequenze di richiesta scenderanno al di sotto della soglia. Questa soluzione è utile per mitigare gli attacchi HTTP Flood mascherati da traffico Web regolare. Per bloccare gli attacchi basati sulla reputazione degli indirizzi IP, puoi creare delle regole utilizzando le condizioni di corrispondenza IP o utilizzare le regole gestite per AWS WAF offerte dai venditori nell'AWS Marketplace. AWS WAF offre direttamente le regole gestite AWS come servizio gestito in cui è possibile scegliere i gruppi di regole di reputazione IP. Il gruppo di regole dell'elenco di reputazione IP di Amazon contiene regole basate sull'intelligence di minacce interne Amazon. Ciò è utile se desideri bloccare gli indirizzi IP tipicamente associati a bot o ad altre minacce. Il gruppo di regole Elenco IP anonimi contiene regole per bloccare le richieste dai servizi che consentono l'offuscamento dell'identità del visualizzatore. Questi includono richieste da VPN, proxy, nodi Tor e piattaforme cloud (incluso AWS). AWS WAF e CloudFront consentono inoltre di impostare restrizioni geografiche per bloccare o consentire le richieste provenienti da paesi selezionati. Questo può aiutare a bloccare gli attacchi da posizioni geografiche in cui non ci si aspetta di servire gli utenti.

Per identificare le richieste dannose, rivedi i log del tuo server Web o utilizza le funzionalità di registrazione e richieste campionate di AWS WAF. Abilitando la registrazione di AWS WAF, si ottengono informazioni dettagliate sul traffico analizzato dall'ACL Web. AWS WAF supporta il filtro dei log, consentendo di specificare quali richieste Web vengono registrate e quali richieste vengono eliminate dal log dopo l'ispezione.

Le informazioni registrate nei log includono l'ora in cui AWS WAF ha ricevuto la richiesta dalla risorsa AWS, informazioni dettagliate sulla richiesta e l'azione corrispondente per ogni regola richiesta. Le richieste campionate forniscono dettagli sulle richieste delle ultime tre ore che corrispondevano a una delle regole di AWS WAF. È possibile utilizzare queste informazioni per identificare le firme di traffico potenzialmente dannose e creare una nuova regola per negare tali richieste. Se vedi un numero di richieste con una stringa di query casuale, assicurati di permettere solo i parametri della stringa di query che sono rilevanti per la cache della tua applicazione. Questa tecnica è utile per mitigare un attacco di cache busting rivolto alla tua origine.

Se sei iscritto a AWS Shield Advanced, puoi contattare l'AWS Shield Response Team (SRT) di AWS per aiutarti a creare le regole necessarie a mitigare un attacco che sta danneggiando la disponibilità della tua applicazione. Puoi concedere ad AWS SRT un accesso limitato a Shield Advanced e alle API AWS WAF del tuo account. AWS SRT accede a queste API per applicare misure di mitigazione

sul tuo account solo con la tua autorizzazione esplicita. Per ulteriori informazioni, consulta la sezione Supporto di questo documento.

Puoi utilizzare AWS Firewall Manager per configurare e gestire centralmente le regole di sicurezza, come le protezioni Shield Advanced e le regole AWS WAF, in tutta l'organizzazione. Il tuo account di gestione di AWS Organizations può designare un account amministratore autorizzato a creare policy di Firewall Manager. Queste policy consentono di definire dei criteri, ad esempio il tipo di risorsa e i tag, che determinano dove vengono applicate le regole. Ciò è utile quando si dispone di più account e si desidera standardizzare la protezione.

Per ulteriori informazioni su:

- Regole gestite da AWS perAWS WAF, consulta Regole gestite da AWS perAWS WAF.
- Utilizzo della restrizione geografica per limitare l'accesso alla distribuzione CloudFront, consulta Limitazione della distribuzione geografica del contenuto.
- Usando AWS WAF, consulta
  - Nozioni di base su AWS WAF
  - Registrazione informazioni di traffico ACL Web
  - Visualizzazione di un esempio di richieste Web
- Configurazione di regole basate sulla frequenza, consulta Protect Web Sites and Services Using Rate-Based Rules for AWS WAF
- Come gestire la distribuzione delle regole AWS WAF tra le risorse AWS con Firewall Manager, consulta
  - Nozioni di base sulle policy di Firewall Manager AWS WAF.
  - Nozioni di base sulle policy avanzate di Firewall Manager Shield.

Riduzione della superficie d'attacco

Un'altra considerazione importante quando si progetta una soluzione AWS è limitare le opportunità disponibili a un utente malintenzionato per prendere di mira la tua applicazione. Questo concetto è noto come riduzione della superficie di attacco. Le risorse che non sono esposte a Internet sono più difficili da attaccare, e questo fatto limita le opzioni di un utente malintenzionato per colpire la disponibilità della tua applicazione.

Ad esempio, se non ti aspetti che gli utenti interagiscano direttamente con determinate risorse, assicurati che tali risorse non siano accessibili da Internet. Allo stesso modo, non accettare traffico da utenti o applicazioni esterne su porte o protocolli che non sono necessari per la comunicazione.

Nella sezione seguente AWS fornisce le best practice per aiutarti a ridurre la superficie di attacco e limitare l'esposizione a Internet della tua applicazione.

Argomenti

Offuscare le risorse AWS (BP1, BP4, BP5)

In genere, gli utenti possono utilizzare un'applicazione in modo rapido e semplice senza bisogno che le risorse AWS siano completamente esposte a Internet. Ad esempio, se sono presenti delle istanze Amazon EC2 dietro un Elastic Load Balancing, potrebbe non essere necessario che le istanze stesse siano pubblicamente accessibili. È invece possibile fornire agli utenti l'accesso all'Elastic Load Balancing su determinate porte TCP e consentire solo a Elastic Load Balancing di comunicare con le istanze. Puoi impostarlo configurando i gruppi di sicurezza e le liste di controllo accessi di rete (NACL) all'interno del tuo Amazon Virtual Private Cloud (Amazon VPC). Amazon VPC consente di effettuare il provisioning di una sezione logicamente isolata di AWS Cloud in cui è possibile avviare risorse AWS in una rete virtuale definita dall'utente.

I gruppi di sicurezza e le liste di controllo degli accessi di rete sono simili, in quanto permettono di controllare l'accesso alle risorse AWS all'interno del tuo VPC. Tuttavia, i gruppi di sicurezza consentono di controllare il traffico in entrata e in uscita a livello di istanza, mentre le liste di controllo degli accessi di rete offrono funzionalità simili a livello di sottorete VPC. Non sono previsti costi aggiuntivi per l'utilizzo di gruppi di sicurezza o liste di controllo degli accessi di rete.

Gruppi di sicurezza e liste di controllo degli accessi di rete (NACL) (BP5)

È possibile scegliere se specificare i gruppi di sicurezza all'avvio di un'istanza o se associare l'istanza a un gruppo di sicurezza in un secondo momento. Tutto il traffico proveniente da Internet verso un gruppo di sicurezza viene implicitamente negato, a meno che non venga creata un regola di assenso per consentire il traffico. Ad esempio, se è presente un'applicazione Web che utilizza un Elastic Load Balancing e diverse istanze Amazon EC2, potresti decidere di creare un solo gruppo di sicurezza per Elastic Load Balancing (gruppo di sicurezza Elastic Load Balancing) e uno per le istanze (gruppo di sicurezza del server applicazioni Web). Puoi quindi creare una regola di assenso per consentire il traffico Internet al gruppo di sicurezza ELB e un'altra regola per consentire il traffico dal gruppo di sicurezza ELB al gruppo di sicurezza al server di applicazioni Web. Ciò garantisce che il traffico Internet non possa comunicare direttamente con le istanze Amazon EC2, rendendo più difficile per un utente malintenzionato conoscere la tua applicazione e agire su di essa.

Quando si creano le liste di controllo degli accessi di rete è possibile specificare sia le regole di assenso che quelle di diniego. Questa opzione è utile nel caso in cui desideri negare esplicitamente determinati tipi di traffico sulla tua applicazione. Ad esempio, è possibile definire indirizzi IP (come gli intervalli CIDR), protocolli e porte di destinazione ai quali viene negato l'accesso per l'intera sottorete. Se l'applicazione viene utilizzata solo per il traffico TCP, è possibile creare una regola per il diniego di tutto il traffico UDP o viceversa. Questa opzione è utile quando si risponde agli attacchi DDoS perché consente di creare le proprie regole per mitigare l'attacco quando si conoscono gli IP di origine o un'altra firma.

Se sei iscritto a AWS Shield Advanced, puoi registrare gli indirizzi IP elastici come risorse protette. Gli attacchi DDoS contro gli indirizzi IP elastici registrati come risorse protette vengono rilevati più rapidamente, con la possibilità di tempi di mitigazione più rapidi. Quando viene rilevato un attacco, i sistemi di mitigazione DDoS leggono la lista di controllo degli accessi di rete corrispondente all'IP elastico di destinazione e lo impongono al confine della rete AWS. In questo modo viene ridotto in modo significativo il rischio di impatto di una serie di attacchi DDoS a livello di infrastruttura.

Per ulteriori informazioni sulla configurazione dei gruppi di sicurezza e delle liste di controllo degli accessi di rete per ottimizzare la resilienza DDoS, consulta Come preparare la difesa dagli attacchi DDoS riducendo la superficie di attacco.

Per ulteriori informazioni sull'utilizzo di Shield Advanced con indirizzi IP elastici come risorse protette, consulta la procedura di Sottoscrizione a AWS Shield Advanced.

Protezione della tua origine (BP1, BP5)

Se utilizzi Amazon CloudFront con un'origine che si trova all'interno del tuo VPC, assicurati che solo la tua distribuzione CloudFront possa inoltrare le richieste alla tua origine. Con le intestazioni di richiesta dall'edge all'origine puoi aggiungere o sovrascrivere il valore delle intestazioni di richiesta esistenti quando CloudFront inoltra le richieste al server di origine. Puoi utilizzare le Origin Custom Headers (Intestazioni personalizzate origine), ad esempio l'intestazione X-Shared-Secret, per verificare che le richieste presentate alla tua origine siano state inviate da CloudFront.

Per maggiori informazioni sulla protezione della tua origine con Origin Custom Headers (Intestazioni personalizzate origine), consulta Aggiunta di intestazioni personalizzate alle richieste di origine e Limitazione dell'accesso ai servizi di Application Load Balancers.

Per una guida sull'implementazione di una soluzione di esempio per ruotare automaticamente il valore delle Origin Custom Headers (Intestazioni personalizzate origine) per la limitazione dell'accesso all'origine, consulta How to enhance Amazon CloudFront origin security with AWS WAF and Secrets Manager (Come migliorare la sicurezza dell'origine di Amazon CloudFront con AWS WAF e Secrets Manager).

In alternativa, puoi utilizzare una funzione di AWS Lambda per aggiornare automaticamente le regole del gruppo di sicurezza per consentire solo il traffico CloudFront. Ciò migliora la sicurezza della tua origine contribuendo a garantire che gli utenti malintenzionati non possano aggirare CloudFront e AWS WAF quando accedono alla tua applicazione Web.

Per ulteriori informazioni su come proteggere la tua origine aggiornando automaticamente i tuoi gruppi di sicurezza, consulta l'intestazione X-Shared-Secret, Come aggiornare automaticamente i gruppi di sicurezza per Amazon CloudFront e AWS WAF utilizzando AWS Lambda.

Protezione degli endpoint API (BP4)

In genere, quando è necessario esporre un'API al pubblico, c'è il rischio che il front-end dell'API possa essere preso di mira da un attacco DDoS. Per ridurre il rischio, puoi utilizzare Amazon API Gateway come accesso alle applicazioni in esecuzione su Amazon EC2, AWS Lambda o altrove. Utilizzando Amazon API Gateway, non hai bisogno dei tuoi server per il front-end API e puoi offuscare altri componenti della tua applicazione. Rendendo più difficile il rilevamento dei componenti dell'applicazione, è possibile evitare che le risorse AWS vengano prese di mira da un attacco DDoS.

Quando utilizzi Amazon API Gateway, puoi scegliere tra due tipi di endpoint API. Il primo è l'opzione di default: endpoint API ottimizzati per l'edge a cui si accede tramite una distribuzione Amazon

CloudFront. La distribuzione è creata e gestita da API Gateway, quindi non hai il controllo su di essa. La seconda opzione consiste nell'utilizzare un endpoint API regionale a cui si accede dalla stessa regione AWS in cui viene implementata l'API REST. AWS consiglia di utilizzare il secondo tipo di endpoint e di associarlo alla propria distribuzione di Amazon CloudFront. In questo modo avrai il controllo sulla distribuzione di Amazon CloudFront e la possibilità di utilizzare AWS WAF per la protezione a livello di applicazione. Questa modalità fornisce l'accesso a una capacità di mitigazione degli attacchi DDoS dimensionabile in tutta la rete edge globale di AWS.

Quando utilizzi Amazon CloudFront e AWS WAF con Amazon API Gateway, configura le seguenti opzioni:

- Configura il comportamento della cache per le distribuzioni per inoltrare tutte le intestazioni all'endpoint regionale di API Gateway. In questo modo, CloudFront tratterà il contenuto come dinamico e salterà la memorizzazione nella cache del contenuto.
- Proteggi il tuo API Gateway dall'accesso diretto configurando la distribuzione per includere l'intestazione personalizzata origine di `x-api-key`, impostando il valore della chiave API in API Gateway.
- Proteggi il back-end dal traffico in eccesso configurando limiti di velocità standard o di aumento delle prestazioni per ogni metodo nelle API REST.

Per ulteriori informazioni sulla creazione di API con Amazon API Gateway, consulta Nozioni di base su Amazon API Gateway.

Tecniche operative

Le tecniche di mitigazione in questo documento consentono di progettare applicazioni intrinsecamente resilienti contro gli attacchi DDoS. In molti casi, è anche utile sapere quando un attacco DDoS prende di mira la tua applicazione in modo da poter adottare misure di mitigazione. In questa sezione vengono illustrate le best practice per ottenere visibilità su comportamenti anomali, avvisi e automazione, gestire la protezione su vasta scala e impegnare AWS per ulteriore supporto.

Argomenti

Visibilità

Quando un parametro operativo chiave si discosta sostanzialmente dal valore previsto, un utente malintenzionato potrebbe tentare di indirizzare la disponibilità dell'applicazione. La familiarità con il normale comportamento della tua applicazione significa che puoi agire più rapidamente quando rilevi un'anomalia. Amazon CloudWatch può aiutarti monitorando le applicazioni su cui esegui AWS. Ad esempio, puoi raccogliere e monitorare parametri e file di log, impostare allarmi e reagire automaticamente ai cambiamenti nelle risorse AWS.

Se si segue l'architettura di riferimento resiliente agli attacchi DDoS nella fase di progettazione dell'applicazione, gli attacchi comuni a livello di infrastruttura verranno bloccati prima di raggiungere l'applicazione. Se sei iscritto a AWS Shield Advanced, hai accesso a una serie di parametri CloudWatch che possono indicare che la tua applicazione è stata scelta come target. Ad esempio, puoi configurare gli allarmi per notificarti quando è in corso un attacco DDoS, in modo da poter controllare l'integrità della tua applicazione e decidere se attivare AWS SRT. Puoi configurare la il parametro `DDoSDetected` per sapere se è stato rilevato un attacco. Se vuoi essere avvisato in base al volume di attacco, puoi anche utilizzare i parametri di `DDoSAttackBitsPerSecond`, `DDoSAttackPacketsPerSecond` o `DDoSAttackRequestsPerSecond`. Puoi monitorare questi parametri integrando CloudWatch con i tuoi strumenti o utilizzando strumenti forniti da terze parti, come Slack o PagerDuty.

Un attacco a livello di applicazione può far aumentare molti parametri di Amazon CloudWatch. Se utilizziAWS WAF, puoi impiegare CloudWatch per monitorare e attivare allarmi in caso di

aumento delle richieste che hai impostato come consentite, conteggiate o bloccate in AWS WAF. In questo modo hai la possibilità di ricevere una notifica se il livello di traffico supera quello che la tua applicazione è in grado di gestire. Puoi anche utilizzare i parametri di Amazon CloudFront, Amazon Route 53, Application Load Balancer, Network Load Balancer, Amazon EC2 e Auto Scaling tracciati in CloudWatch per rilevare le modifiche che possono indicare un attacco DDoS.

La tabella Parametri di CloudWatch consigliati elenca le descrizioni dei parametri di CloudWatch comunemente utilizzati per rilevare e reagire agli attacchi DDoS.

Tabella 3: parametri consigliati di Amazon CloudWatch

Argomento	Parametro	Descrizione
AWS Shield Advanced	DDoSDetected	Indica un evento DDoS per uno specifico Amazon Resource Name (ARN).
AWS Shield Advanced	DDoSAttackBitsPerSecond	Il numero di byte osservati durante un evento DDoS per un ARN specifico. Questo parametro è disponibile solo per gli eventi DDoS di livello 3/4.
AWS Shield Advanced	DDoSAttackPacketsPerSecond	Il numero di pacchetti osservati durante un evento DDoS per un ARN specifico. Questo parametro è disponibile solo per gli eventi DDoS di livello 3/4.
AWS Shield Advanced	DDoSAttackRequestsPerSecond	Il numero di richieste osservate durante un evento DDoS per un ARN specifico. Questo parametro è disponibile solo per gli eventi DDoS di livello 7 e viene segnalato solo

Argomento	Parametro	Descrizione
		per gli eventi più importanti di livello 7.
AWS WAF	AllowedRequests	Il numero di richieste Web consentite.
AWS WAF	BlockedRequests	Il numero di richieste Web bloccate.
AWS WAF	CountedRequests	Il numero di richieste Web contate.
AWS WAF	PassedRequests	Il numero di richieste superate. Viene utilizzato solo per le richieste che vengono sottoposte a una valutazione del gruppo di regole senza corrispondere a nessuna regola del gruppo.
Amazon CloudFront	Richieste	Il numero di richieste HTTP/S.
Amazon CloudFront	TotalErrorRate	La percentuale di tutte le richieste per le quali il codice di stato HTTP è 4xx o 5xx.
Amazon Route 53	HealthCheckStatus	Lo stato dell'endpoint di controllo dell'integrità.
Application Load Balancer	ActiveConnectionCount	Il numero totale di connessio ni TCP attive dai client al bilanciatore del carico e dal bilanciatore del carico ai target.

Argomento	Parametro	Descrizione
Application Load Balancer	ConsumedLCUs	Il numero di unità di capacità del bilanciatore del carico (LCU) utilizzate dal tuo bilanciatore del carico.
Application Load Balancer	HTTPCode_ELB_4XX_Count HTTPCode_ELB_5XX_Count	Il numero di codici di errore client HTTP 4xx o 5xx generati dal bilanciatore del carico.
Application Load Balancer	NewConnectionCount	Il numero totale di nuove connessioni TCP stabilite dai client al bilanciatore del carico e dal bilanciatore del carico ai target.
Application Load Balancer	ProcessedBytes	Il numero totale di byte elaborati dal bilanciatore del carico.
Application Load Balancer	RejectedConnectionCount	Il numero di connessioni rifiutate perché il bilanciat ore del carico ha raggiunto il numero massimo di connessio ni.
Application Load Balancer	RequestCount	Il numero di richieste che sono state elaborate.
Application Load Balancer	TargetConnectionErrorCount	Il numero di connessioni che non sono state stabilite con successo tra il bilanciatore del carico e il target.

Argomento	Parametro	Descrizione
Application Load Balancer	TargetResponseTime	Il tempo trascorso, in secondi, dal momento in cui la richiesta lascia il bilanciatore del carico fino a quando non si riceve una risposta dal target.
Application Load Balancer	UnHealthyHostCount	Il numero di target considerati non integri.
Network Load Balancer	ActiveFlowCount	Il numero totale di flussi simultanei (o connessioni) TCP da client a target.
Network Load Balancer	ConsumedLCUs	Il numero di unità di capacità del bilanciatore del carico (LCU) utilizzate dal tuo bilanciatore del carico.
Network Load Balancer	NewFlowCount	Il numero totale di nuovi flussi (o connessioni) TCP stabiliti da client a target nel periodo di tempo.
Network Load Balancer	ProcessedBytes	Il numero totale di byte elaborati dal bilanciatore del carico, incluse le intestazioni TCP/IP.
Global Accelerator	NewFlowCount	Il numero totale di nuovi flussi (o connessioni) TCP e UDP stabiliti da client a endpoint nel periodo di tempo.

Argomento	Parametro	Descrizione
Global Accelerator	ProcessedBytesIn	Il numero totale di byte in entrata elaborati dall'acce leratore, incluse le intestazioni TCP/IP.
Auto Scaling	GroupMaxSize	Dimensione massima del gruppo Auto Scaling.
Amazon EC2	CPUUtilization	Percentuale delle unità di elaborazione EC2 attualmente in uso.
Amazon EC2	NetworkIn	Il numero di byte ricevuti dall'istanza su tutte le interfacc e di rete.

Per ulteriori informazioni sull'utilizzo di Amazon CloudWatch per rilevare gli attacchi DDoS nella tua applicazione, consulta Nozioni di base di Amazon CloudWatch.

Per esplorare un esempio di pannello di controllo creato utilizzando alcuni parametri della tabella precedente, consulta A custom baseline monitoring system

AWS include diversi parametri e allarmi aggiuntivi per informarti di un attacco e per aiutarti a monitorare le risorse della tua applicazione. La console o l'API di AWS Shield forniscono un riepilogo degli eventi per account e dettagli sugli attacchi rilevati.

Inoltre, il pannello di controllo globale dell'ambiente delle minacce fornisce informazioni di riepilogo su tutti gli attacchi DDoS rilevati da AWS. Queste informazioni possono essere utili per comprendere meglio le minacce DDoS in una popolazione più ampia di applicazioni oltre alle tendenze degli attacchi e per il confronto con gli attacchi che potresti aver osservato.

Se sei abbonato a AWS Shield Advanced, il pannello di controllo del servizio visualizza parametri di rilevamento e mitigazione aggiuntivi e dettagli sul traffico di rete per gli eventi rilevati nelle risorse protette. AWS Shield valuta il traffico verso la tua risorsa protetta su più dimensioni. Quando viene rilevata un'anomalia, AWS Shield crea un evento e segnala la dimensione del traffico in cui è stata

osservata l'anomalia. Con una mitigazione posizionata, questo protegge la tua risorsa dalla ricezione di traffico in eccesso e di traffico che corrisponde a una firma di evento DDoS nota.

I parametri di rilevamento si basano su flussi di rete campionati o log di AWS WAF quando un ACL Web è associato alla risorsa protetta. I parametri di mitigazione si basano sul traffico osservato dai sistemi di mitigazione DDoS di Shield. I parametri di mitigazione sono una misurazione più precisa del traffico nella tua risorsa.

Il parametro dei principali collaboratori di rete fornisce informazioni dettagliate sulla provenienza del traffico durante un evento rilevato. È possibile visualizzare i collaboratori con il volume più elevato e ordinare per aspetti come protocollo, porta di origine e flag TCP. Il parametro dei collaboratori principali include i parametri di tutto il traffico osservato sulla risorsa in varie dimensioni. Fornisce le dimensioni dei parametri aggiuntivi utilizzabili per comprendere il traffico di rete inviato alla tua risorsa durante un evento.

Il pannello di controllo del servizio include anche dettagli sulle azioni intraprese automaticamente per mitigare gli attacchi DDoS. Queste informazioni facilitano l'indagine sulle anomalie, l'esplorazione delle dimensioni del traffico e offrono una migliore comprensione delle azioni intraprese da Shield Advanced per proteggere la tua disponibilità.

Un altro strumento che può aiutarti a ottenere visibilità sul traffico destinato alla tua applicazione è il flusso di log VPC. Su una rete tradizionale puoi utilizzare i flussi di log per risolvere i problemi di connettività e sicurezza e assicurarti che le regole di accesso alla rete stiano funzionando come dovuto. Utilizzando i flussi di log VPC, puoi acquisire informazioni sul traffico IP da e per le interfacce di rete nel VPC.

Ogni registro del flusso di log include quanto segue: indirizzi IP di origine e destinazione, porte di origine e destinazione, protocollo e numero di pacchetti e byte trasferiti durante la finestra di acquisizione. È possibile utilizzare queste informazioni per identificare le anomalie nel traffico di rete e per identificare un vettore di attacco specifico. Ad esempio, la maggior parte degli attacchi di riflessione UDP presenta porte di origine specifiche, come la porta di origine 53 per la riflessione DNS. Si tratta di una chiara firma di attacco che è possibile identificare nel registro del flusso di log. In risposta, è possibile scegliere di bloccare la porta di origine specifica a livello di istanza o creare una regola ACL di rete per bloccare l'intero protocollo se l'applicazione non lo richiede.

Per ulteriori informazioni sull'utilizzo flussi di log VPC per identificare anomalie di rete e vettori di attacco DDoS, consulta Flussi di log di VPC e VPC Flow Logs – Log and View Network Traffic Flows.

Gestione della visibilità e della protezione su più account

Negli scenari in cui operi su diversi account AWS e hai più componenti da proteggere, l'utilizzo di tecniche che ti consentono di operare su vasta scala e ridurre i costi operativi aumenta le tue capacità di mitigazione. Quando gestisci risorse protette di AWS Shield Advanced in più account, è possibile impostare il monitoraggio centralizzato utilizzando AWS Firewall Manager e AWS Security Hub. Con Firewall Manager, puoi creare una policy di sicurezza che impone la conformità alla protezione DDoS in tutti i tuoi account. È possibile utilizzare questi due servizi insieme per gestire le risorse protette su diversi account e centralizzare il monitoraggio di tali risorse.

Security Hub si integra automaticamente con Firewall Manager, consentendo ai clienti di Shield Advanced di visualizzare i risultati della sicurezza in un unico pannello di controllo, insieme ad altri avvisi di sicurezza ad alta priorità e stati di conformità. Ad esempio, quando Shield Advanced rileva traffico anomalo destinato a una risorsa protetta in un qualsiasi account AWS all'interno dell'ambito, questo risultato sarà visibile nella console di Security Hub. Se configurato, Firewall Manager può portare automaticamente la risorsa alla conformità creandola come risorsa protetta Shield Advanced, e quindi aggiornare Security Hub quando la risorsa è in uno stato conforme.

Per ulteriori informazioni sul monitoraggio centralizzato delle risorse protette da Shield, consulta Set up centralized monitoring for DDoS events and auto-remediate noncompliant resources.

Supporto

Se si verifica un attacco, è anche possibile beneficiare del supporto di AWS nella valutazione della minaccia e nella revisione dell'architettura dell'applicazione, oppure è possibile richiedere altra assistenza. È importante creare un piano di risposta per gli attacchi DDoS prima di un evento reale. Le best practice descritte in questo documento sono intese come misure proattive da implementare prima di avviare un'applicazione, ma potrebbero comunque verificarsi attacchi DDoS che prendono di mira l'applicazione. Esamina le opzioni in questa sezione per determinare le risorse di supporto più adatte al tuo scenario. Il team del tuo account può valutare il tuo caso d'uso e la tua applicazione e aiutarti con domande o problematiche specifiche.

Se esegui carichi di lavoro in produzione su AWS, valuta la possibilità di iscriverti al Supporto Business, che ti offre accesso 24 ore su 24, 7 giorni su 7, ai tecnici del supporto cloud che possono aiutarti con i problemi degli attacchi DDoS. Se esegui carichi di lavoro mission-critical, prendi in considerazione il supporto Enterprise che offre la possibilità di aprire casi critici e ricevere la risposta più rapida da un Tecnico senior di supporto cloud.

Se sei abbonato a AWS Shield Advanced e sei abbonato anche al Supporto Business o Enterprise, puoi configurare il coinvolgimento proattivo di Shield. Ti consente di configurare controlli dell'integrità, associarti alle tue risorse e fornire informazioni di contatto operative 24 ore su 24, 7 giorni su 7. Quando Shield rileva segni di attacchi DDoS e i controlli dell'integrità delle applicazioni mostrano segni di degrado, AWS SRT ti contatterà in modo proattivo. Questo è il nostro modello di coinvolgimento consigliato perché consente a AWS SRT tempi di risposta più rapidi e la possibilità di avviare la risoluzione dei problemi ancor prima che venga stabilito un contatto con te.

La funzione di coinvolgimento proattivo richiede la configurazione di un controllo dell'integrità di Route 53 che misura accuratamente l'integrità dell'applicazione ed è associato alla risorsa protetta da Shield Advanced. Una volta associato un controllo dell'integrità di Route 53 nella console Shield, il sistema di rilevamento Shield Advanced utilizza lo stato del controllo dell'integrità come indicatore dell'integrità dell'applicazione. La funzione di rilevamento basato sull'integrità di Shield Advanced ti garantisce di ricevere una notifica e che le misure di mitigazione vengano applicate più rapidamente quando la tua applicazione non è integra. AWS SRT ti contatterà per la risoluzione dei problemi se l'applicazione non integra è stata presa di mira da un attacco DDoS e per apportare ulteriori misure di mitigazione, se necessario.

Il completamento della configurazione del coinvolgimento proattivo include l'aggiunta di dettagli di contatto nella console Shield. AWS SRT utilizzerà queste informazioni per contattarti. È possibile configurare fino a 10 contatti e fornire note aggiuntive se si hanno requisiti o preferenze di contatto specifici. I contatti di coinvolgimento proattivo devono avere un ruolo 24 ore su 24, 7 giorni su 7, come un centro operativo di sicurezza o una persona immediatamente disponibile.

È possibile abilitare un coinvolgimento proattivo per tutte le risorse o per determinate risorse di produzione chiave in cui il tempo di risposta è fondamentale. Ciò si ottiene assegnando i controlli dell'integrità solo a queste risorse.

Puoi anche effettuare una segnalazione a AWS SRT creando un caso AWS Support utilizzando la console di AWS Support o l'API del supporto se hai un evento correlato a un attacco DDoS che influisce sulla disponibilità della tua applicazione.

Conclusione

Le best practice delineate in questo documento possono risultare utili per creare un'architettura resiliente agli attacchi DDoS in grado di proteggere la disponibilità della tua applicazione prevenendo molti attacchi comuni a livello di infrastruttura e applicazione. Il grado in cui è possibile seguire queste best practice in fase di progetto dell'applicazione influenza il tipo, il vettore e il volume di attacchi DDoS possibili da mitigare. È possibile incorporare la resilienza senza sottoscrivere un servizio di mitigazione degli attacchi DDoS. Scegliendo la sottoscrizione a AWS Shield Advanced ottieni ulteriori caratteristiche di supporto, visibilità, mitigazione e protezione dei costi che proteggono ulteriormente un'architettura dell'applicazione già resiliente.

Collaboratori

I collaboratori di questo documento includono:

- Jeffrey Lyon, Protezione perimetrale AWS
- Rodrigo Ferroni, Esperto di sicurezza AWS TAM
- Dmitriy Novikov, AWS Solutions Architect
- Achraf Souk, AWS Solutions Architect
- Yoshihisa Nakatani, AWS Solutions Architect

Risorse

Approfondimenti:

- Best Practices for DDoS Mitigation on AWS
- Linee guida per l'implementazione di AWS WAF
- SID324 – re:Invent 2017: Automating DDoS Response in the Cloud
- CTD304 – re:Invent 2017: Dow Jones & Wall Street Journal’s Journey to Manage Traffic Spikes While Mitigating DDoS & Application Layer Threats
- CTD310 – re:Invent 2017: Living on the Edge, It’s Safer Than You Think! Building Strong with Amazon CloudFront, AWS Shield, and AWS WAF
- SEC407 - re:Invent 2019: A defense-in-depth approach to building web applications
- SEC321 - re:Invent 2020: Get ahead of the curve with DDoS Response Team escalations
- William Hill: High-performance DDOS Protection with AWS

Revisioni del documento

Per ricevere una notifica sugli aggiornamenti di questo whitepaper, iscriviti al feed RSS.

update-history-change	update-history-description	update-history-date
Aggiornamento del whitepaper	Aggiornato per includere le raccomandazioni e le caratteri stiche più recenti. AWS Global Accelerator viene aggiunto come parte integrante di una protezione completa a livello di edge. AWS Firewall Manager per il monitoraggio centraliz zato degli eventi DDoS e la correzione automatica delle risorse non conformi.	21 settembre 2021
Aggiornamento del whitepaper	Aggiornato per chiarire il busting della cache nella sezione Rilevare e filtrare le richieste Web dannose (BP1, BP2) e l'utilizzo di ELB e ALB nella sezione Dimensionare per assorbire (BP6). Diagrammi aggiornat i e Tabella 2, contrassegnati come "Scelta della regione", come BP8. Sezione BP7 aggiornata con maggiori dettagli.	18 dicembre 2019
Aggiornamento del whitepaper	Aggiornato per includere la registrazione di AWS WAF come best practice.	1 dicembre 2018

Aggiornamento del whitepaper	Aggiornato per includere AWS Shield, le caratteristiche di AWS WAF, AWS Firewall Manager e le best practice correlate.	1 giugno 2018
Aggiornamento del whitepaper	Aggiunta di linee guida sull'architettura prescrittive e aggiornate per includere AWS WAF.	1 giugno 2016
Pubblicazione iniziale	Whitepaper pubblicato.	1 giugno 2015

Avvisi

I clienti sono responsabili della propria valutazione autonoma delle informazioni contenute in questo documento. Questo documento: (a) è solo a scopo informativo, (b) mostra le offerte e le pratiche attuali dei prodotti AWS soggette a modifiche senza preavviso, e (c) non crea alcun impegno o garanzia da parte di AWS e dei suoi affiliati, fornitori o licenziatari. I prodotti o servizi AWS sono forniti "così come sono" senza garanzie, dichiarazioni o condizioni di alcun tipo, sia esplicite che implicite. Le responsabilità e gli obblighi di AWS verso i propri clienti sono disciplinati dagli accordi AWS e il presente documento non fa parte né modifica alcun accordo tra AWS e i suoi clienti.

© 2021, Amazon Web Services, Inc., o sue affiliate. Tutti i diritti riservati.