WHITE PAPER

IS IT A FLASH CROWD OR A DDOS ATTACK?

Protect your users and systems from flash crowds, DDoS attacks and cheaters

TABLE OF CONTENTS

EXECUTIVE SUMMARY

DDoS attacks and flash crowds originate from different motivations, but from a NOC perspective, they look and feel like the same thing. They start with an inrush of traffic accompanied by a swell of new IP addresses that are about to overwhelm the networking and application capabilities of the services. If not handled properly, the result is the same: 503, service not available. This paper will outline the defense strategies that prioritize protecting legitimate users while also ensuring the system isn't overwhelmed during flash-crowd events or a DDoS attack.

503 HTTP Error.
The Service is unavailable.

FLASH CROWDS FROM THE DDOS DEFENDER'S PERSPECTIVE

How to manage flash crowds is an important capability for DDoS defense effectiveness consideration. The fear is that your marketing team hits a home run with a shiny new program that drives troves of customers swarming to the site. Ca-ching! And the unwitting DDoS defense sees the traffic swell as a DDoS attack, blocks the traffic and ends the celebration. Or it recognizes it as legitimate traffic and does nothing until the server dies and the celebrations ends, as well. Oops! Avoiding this career-limiting situation is no different than implementing a user-centric DDoS defense that can distinguish good from bad traffic while protecting the system from total failure. Sounds simple, but not always achievable, with brute-force DDoS defenses.

In the following sections, we will outline how to build DDoS defenses that are resilient to attacks and are flash-crowd ready. Before we jump into the details, we need to first define the objectives of defending networks during a flash-crowd event.

GOALS OF FLASH - CROWD EVENT MANAGEMENT

Flash crowds don't just happen. An event like a new game launch, blowout sale, mega tour announcement, or a significant social event that turns eyes en-masse precedes flash crowds. As a business defender, three other considerations are just as important as making sure the system is ready for the swell of users:

Protect the system – All infrastructures have networking and computational limits. If the crowd grows to the limits, desirable workloads have to be dropped. After all, if the system fails, then it is unavailable to everyone. Putting a cap on the traffic is called destination-traffic shaping.

Stop cheaters – Market-driven flash events draw cheaters that will game the system to get an unfair share. For example, when a concert goes on sale, a scalper will leverage automation to pocket large swaths of the tickets while fans get pushed aside. Do the event organizers care that the fans have to pay outrageous prices from scalpers? I hope so. Cheating can be restrained with source controls, which we will cover later.

Block DDoS and protect users – Unfortunately, when the stakes are the highest, your business is most vulnerable to DDoS attacks. Like cheaters, flash events attract DDoS attacks initiated by spoilers and competitors. It is just a reality of modern life for some industries. This is why you build DDoS defenses, but the tricky part is blocking the attack without inflicting collateral damage against the flash crowd.

SOURCE VS. DESTINATION MITIGATION

Not all DDoS defense or defense strategies are the same. One of the essential principles that distinguish user-centric defenses from ones focused on infrastructure protection is how the system balances source-based and destination-based controls.

Destination controls apply filters and limits to the traffic directed at the protected service. Destination filters are very effective in protecting the system from failing during good or bad traffic swell cycles. Technically, they help IT departments achieve five-nine uptime objectives but also leave a wake of collateral damage against users. Remotely Triggered Black Hole Filtering (RTBH) is one example of destination-rate controls. Zero bad traffic makes it to the service, but so does zero good traffic.

A less draconian method is traffic shaping. Traffic shaping uses layer 3-7 limits of BPS, PPS, SYN, CPS, RPS, or any other system limit that threatens the stability of the service being protected.

Depending on your defense systems capabilities, this could be a single limit for the aggregate ingress traffic to the protected service. For more sophisticated systems, traffic shaping can be refined for each traffic category of the traffic mix (e.g., TCP, UDP, ICMP, GRE, HTTP POST/GET) for granular defense controls that cause less collateral damage.

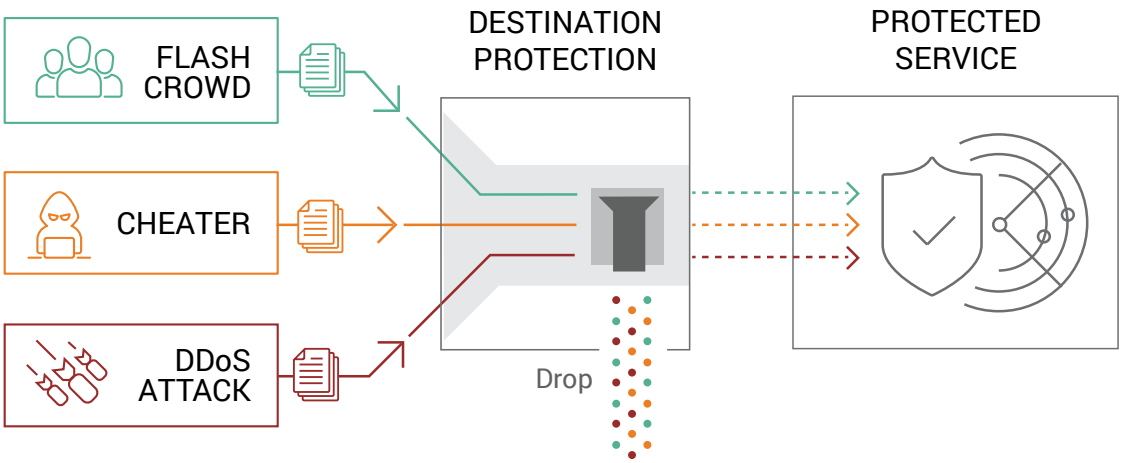

Figure 1: Destination traffic shaping

Alternatively, source controls apply the same principles, but on the requester rather than the destination. The engine looks for anomalous behavior that breaks policies from each of the requesting agents. For example, limiting the number of total sessions allowed or the rate of new sessions started for the scalper would constrain the total number of tickets they can acquire. Today, many internet cheaters are part of a profit-driven criminal organization with cloud computing, powerful hosts, and automation capabilities. Simple access policies that can run at scale will level the field for the real target audience.

"At scale" is the critical point in the context of flash crowds and DDoS management. Slowing down cheater and DDoS agents is vital, but the legitimate flash-crowd participant can't be encumbered.

Source controls are computationally more intensive than destination controls due to the sheer number of individual IP addresses and sessions to track. In addition to L3-7 rate controls, IP reputation, spoof detection, and botnet challenges can be applied to each requester to prevent unwanted requests. Source-based controls, unlike indiscriminate destination controls, are surgical and use cause as a basis for choosing what gets in and what is blocked. This strategy is user-focused, and that makes sense since legitimate users are what drive your business.

SOURCE-BASED LIMITS TO MITIGATE FLASH CROWD, CHEATERS, AND DOS AGENTS

Layer 3-4 Source-base limits

Packet rate (pps)

Bandwidth rate (Kbps)

Connection rate (cps)

Connection limit

Frag packet rate (pps)

SYN rate

Layer 6-7 Source-based limits

SSL request limit

SSL renegotiation limit

POST/GET rate limit

POST/GET rate limit per URI

SIP request type limit

DDOS RESILIENT AND FLASH-CROWD READY DEFENSE

Source and destination control combined improve DDoS defense effectiveness. Destination limits protect the system from being overwhelmed, but they are a tool of last resort, and not the first and only method to be used.

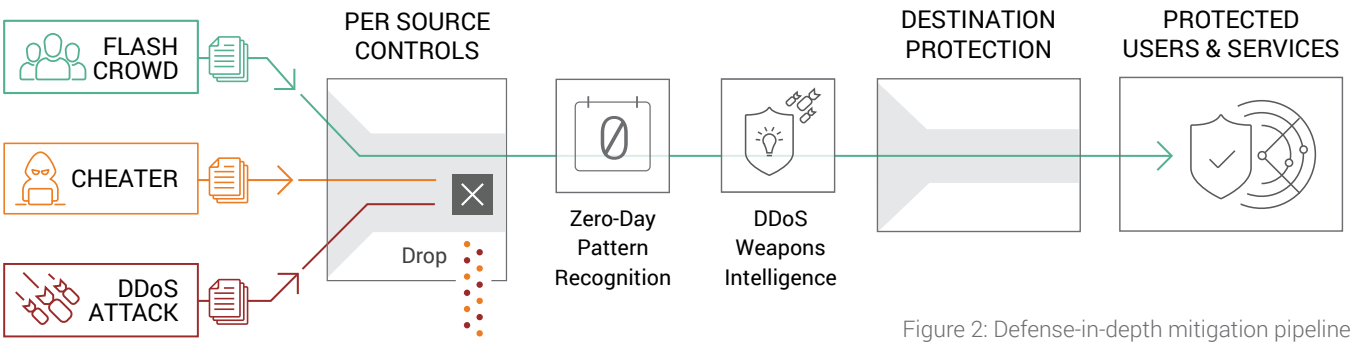

Figure 2: Defense-in-depth mitigation pipeline

Applying a defense-in-depth pipeline of **adaptive source-based controls** with **zero-day attack pattern recognition**, **DDoS weapons IP blocking**, and destination controls provides a holistic defense strategy that will protect legitimate users, restrain cheaters, block attackers and protect the system during a flash-crowd event.

CONCLUSION

Flash crowds don't just happen. For most business, it as frequent as unicorn sightings. For the lucky IT managers who experience them often, follow these recommendations, and you will be ready to delight your customers, block cheaters, thwart DDoS spoilers, and maintain five-nines uptime.

FLASH CROWD DEFENSE GOALS RESULT	RESULT
Ensure availability to users	✓
Protect the system from falling over	✓
Stop cheaters, level the playing field	✓
Block DDoS, protect users	✓
YOUR PROMOTION	👍

ACRONYM GLOSSARY

BPS – Bits per Second

CPS – Connections per Second

DDoS – Distributed Denial of Service

GET – HTTP method

GRE - Generic Routing Encapsulation

ICMP - Internet Control Message Protocol

NOC – Network operation Center

POST – HTTP method

PPS – Packets per Second

RPS – Requests per Second

RTBH – Remote Trigger Blackhole

SYN – Synchronize command in TCP handshake

TCP - Transmission Control Protocol

UDP - User Datagram Protocol

ABOUT A10 NETWORKS

A10 Networks (NYSE: ATEN) provides Reliable SecurityAlways™, with a range of high-performance application networking solutions that help organizations ensure that their data center applications and networks remain highly available, accelerated and secure. Founded in 2004, A10 Networks is based in San Jose, Calif., and serves customers globally with offices worldwide.

For more information, visit: a10networks.com or tweet @a10Networks

LEARN MORE

ABOUT A10 NETWORKS

CONTACT US

a10networks.com/contact

Part Number: A10-WP-21158-EN-01 SEPT 2019