

RSAConference2016

San Francisco | February 29 – March 4 | Moscone Center

SESSION ID: SPO2-W05

Bots—Fast Growing Bane of the Web: Crawlers, Scrapers and Account Checkers



Connect **to**
Protect

John Summers

Vice President Security Business
Unit, Akamai Technologies



#RSAC

Security always gets the tough nuts.

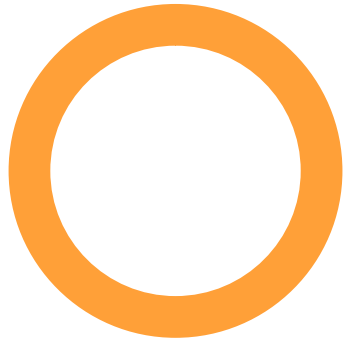
But are bots just a security problem?



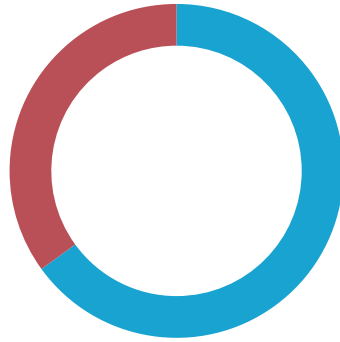


- Understanding the Bot Problem
 - Size
 - Operator
 - Types of Bots
- Business and IT Impact.
- The Right Strategy for Each Use Case
- Lessons Learned

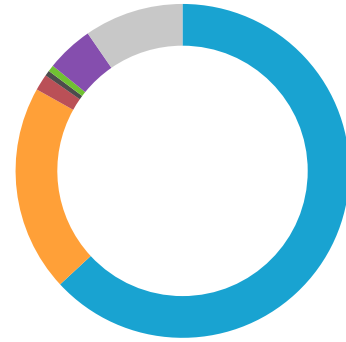
Understanding the “Bot Problem”



Your site traffic

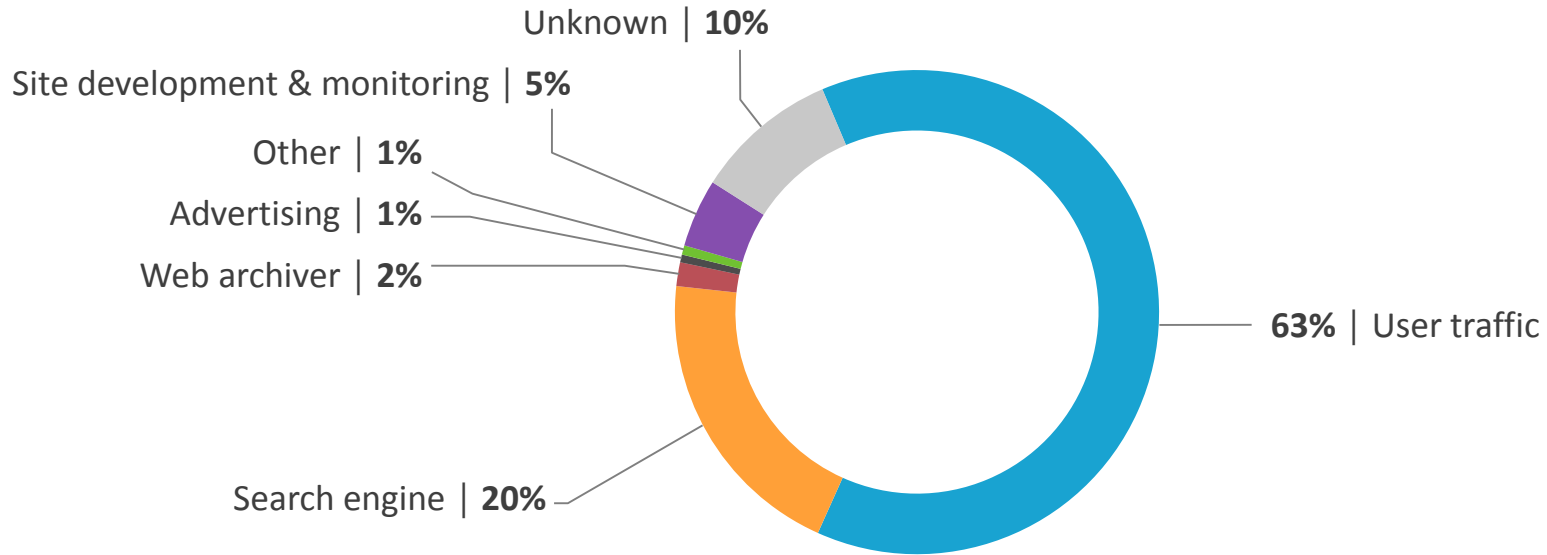


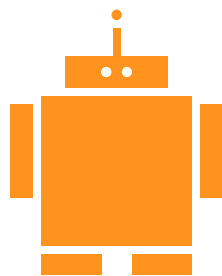
**What you think your
traffic looks like**



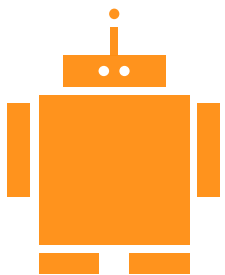
**What your traffic
actually looks like**

Understanding the “Bot Problem”





This is a bot.



A bot is just an automated tool.



More important is the bot operator, which programs the tool to perform a repetitive task on the Internet.

Understanding bots requires understanding the bot operator.

How big is the Bot problem



#RSAC

Clothing company: **15.3%** BW

Car dealer: **32.3%** BW

Hi-Tech distrib: **70.3%** BW

FROM: 21 Jan 2016, 8:14 pm TO: 28 Jan 2016, 8:14 pm

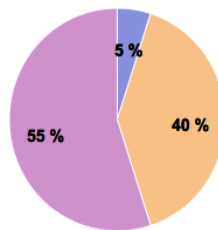
Bot Traffic (% of Total)

48.7% Edge Hits
24.4M Requests 50.2M Total

50.3% Edge Page Views
17.7M Page Views 35.1M Total

70.3% Edge Bandwidth
3TB 4.2TB Total

Types of Bots Detected

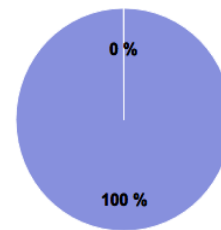


Customer Known Bots
(1,106,878) 5%

Akamai Known Bots
(9,744,135) 40%

Unknown Bots Detected
(13,307,656) 55%

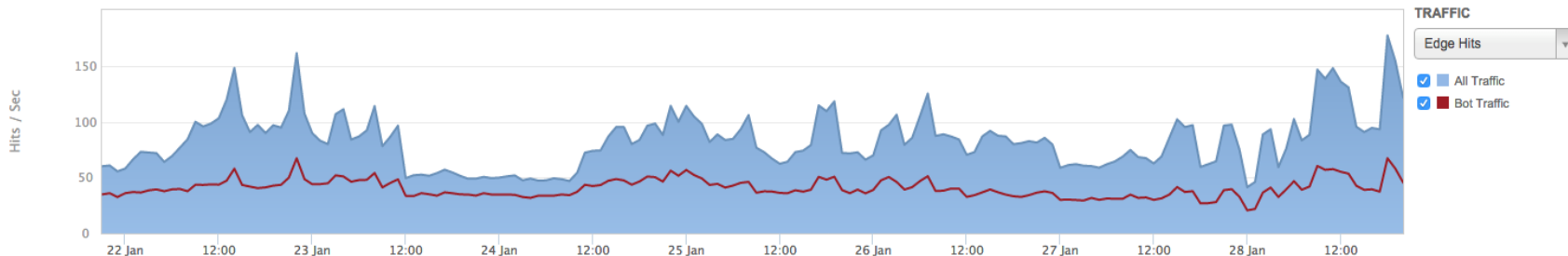
Action Applied to Bots



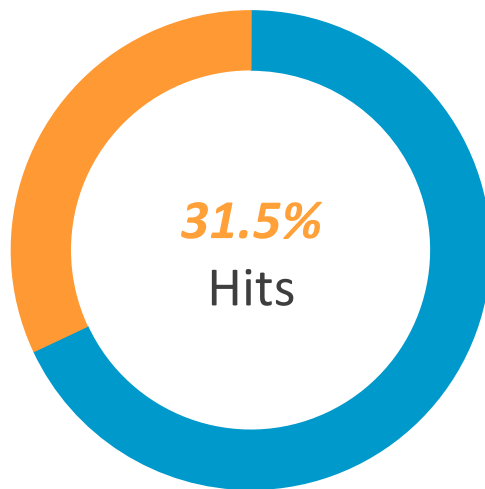
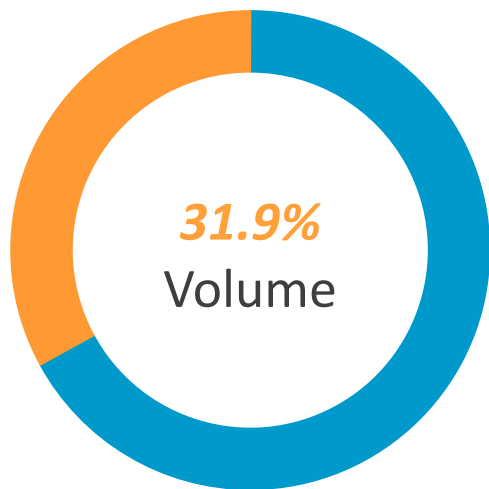
Alert
(24,158,493) 100%

Deny
(176) 0%

Trend of Bot vs Total Traffic



How big is the Bot problem



21st – 28th January 2016
MEDIAN across all companies
using BotMan

Bot Traffic vs. **Total** Traffic

Types of Bots



#RSAC

Security Configuration

Configuration Versions | v157 Quick Links: Custom Rules Management | R

Bot Management

This configuration has been activated and is not editable.

Configure Protections

Manage Components

Default Security Configuration v157

Bot Protection

On

General Settings

Customer-Known Bots

4 Bot Categories

Akamai-Known Bots

15 Bot Categories

Bot Category Expand All Collapse All

Number of Bots

Action When Detected

Academic or Research Bots

37

Monitor

Business Intelligence Bots

es use bots to collect data from various web sites for academic or research purposes. Their automated web bots vary in purpose and can include topic-specific etc.

E-Commerce Search Engine Bots

Enterprise Data Aggregator Bots

Financial Services Bots

Job Search Engine Bots

Media or Entertainment Search Bots

News Aggregator Bots

Online Advertising Bots

RSS Feed Reader Bots

SEO, Analytics or Marketing Bots

Site Monitoring and Web Development Bots

Social Media or Blog Bots

Web Archiver Bots

Web Search Engine Bots

Financial Services Bots

Job Search Engine Bots

32

Monitor

124

Monitor

64

Monitor


11

Monitor

25

Monitor

More than 300 known bots



What's wrong with...



#RSAC



Business and IT Impacts

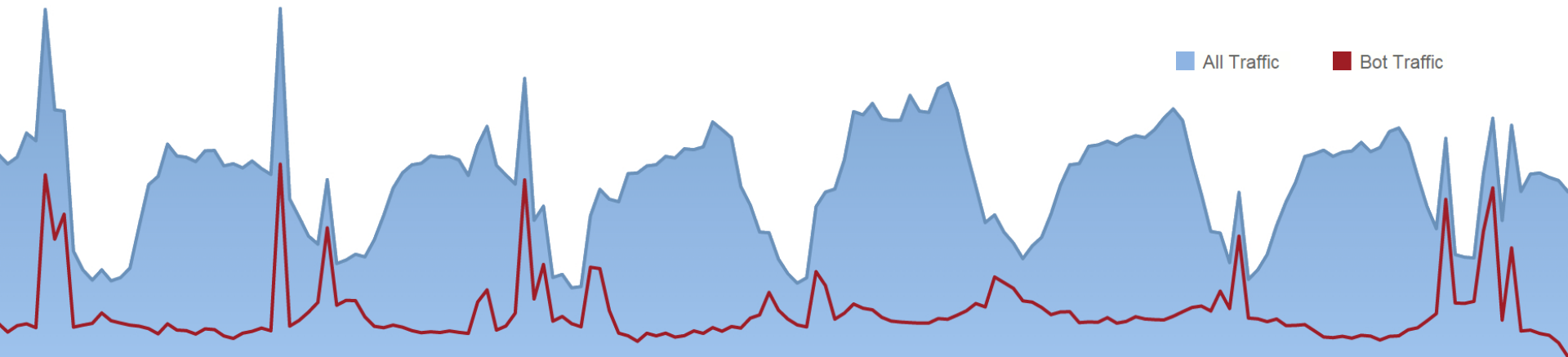


#RSAC



Origin load

Poor site performance and higher infrastructure costs



Business and IT Impacts



Origin load

Poor site performance and higher infrastructure costs



Content aggregation

Loss of opportunity and customer relationships



Inventory grabbing

Loss of sales and customer relationships



Form / comment spam

False records and defaced online communities



Web scraping

IP theft and loss of competitive advantage



Web analytics

Effective monitoring of marketing metrics



Good bot management

Search engines, partners and internal bots

Business and IT Impacts. Stakeholders



⚙️ **Origin load** Infrastructure Manager, IT Manager, CIO

⚙️ **Content aggregation**

Head of Sales, Content Manager

⚙️ **Web scraping**

CISO, Legal, VP of Products

⚙️ **Inventory grabbing**

Head of Sales, Procurement, Operations

⚙️ **Web analytics**

VP of Marketing, VP of Sales

⚙️ **Form / comment spam**

VP of Marketing, Community Manager

⚙️ **Good bot management**

SEO, VP of Channels, Procurement

Manage, Not Mitigate



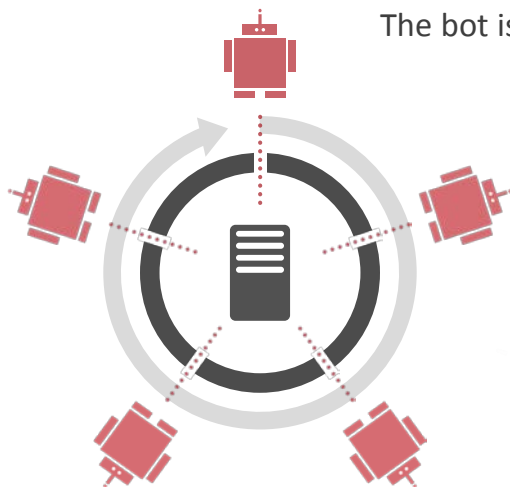
#RSAC

Whack-a-mole

The bot returns but is now better hidden from detection

Evasion

Operator modifies the bot to evade detection / mitigation



Motivation

The bot is here to get something

Blocking

Prevents the bot from getting what it came for

Awareness

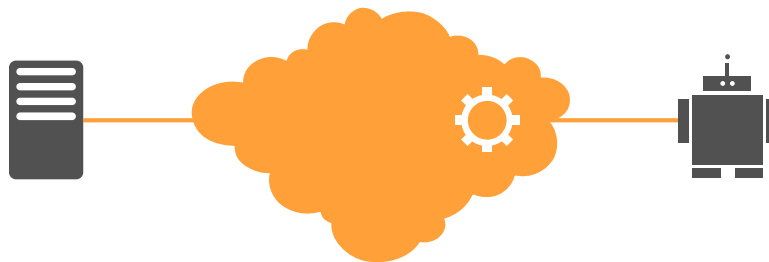
Blocking also alerts bot operator

Let's Talk about Use Cases



#RSAC

What are the impacts on
your website and business?



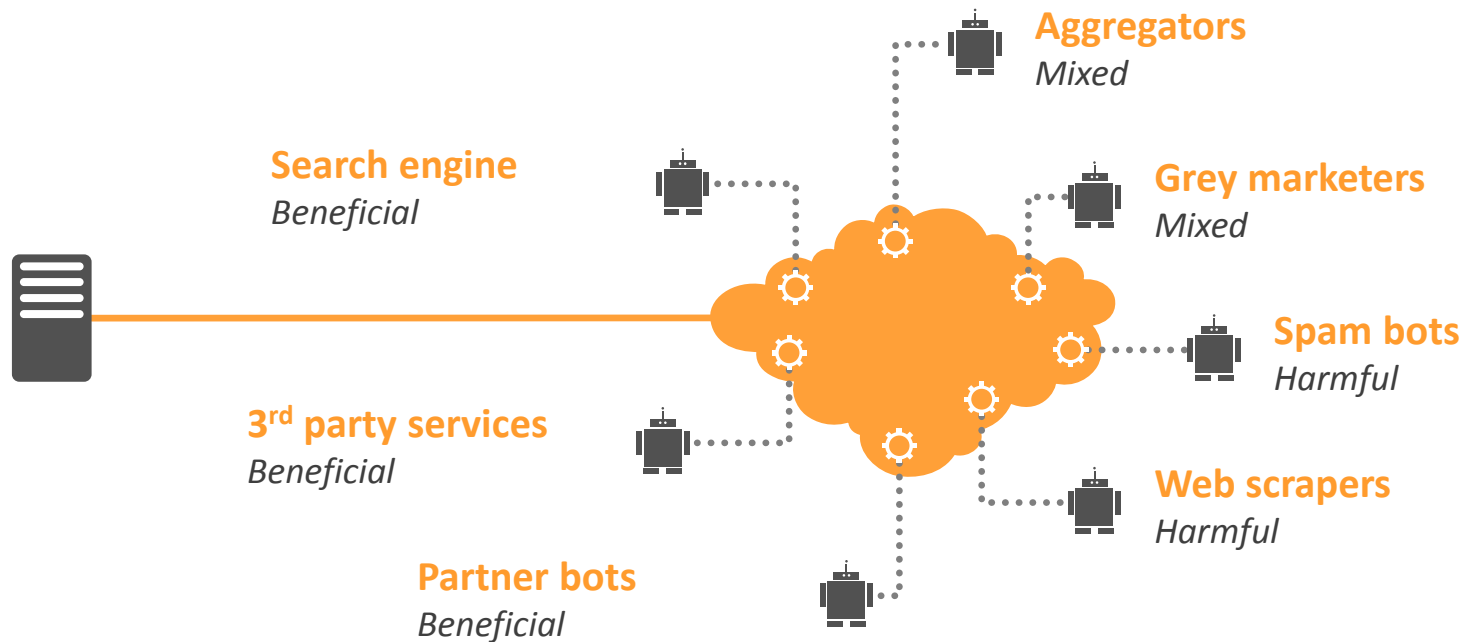
What are the incentives
and goals of the operator?

A bot use case is a combination of the bot
operator and the business and IT impact.

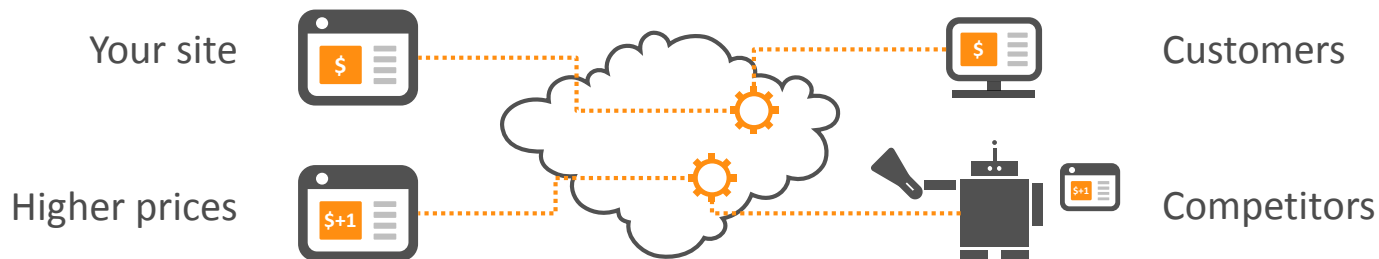
Let's Talk about Use Cases



#RSAC



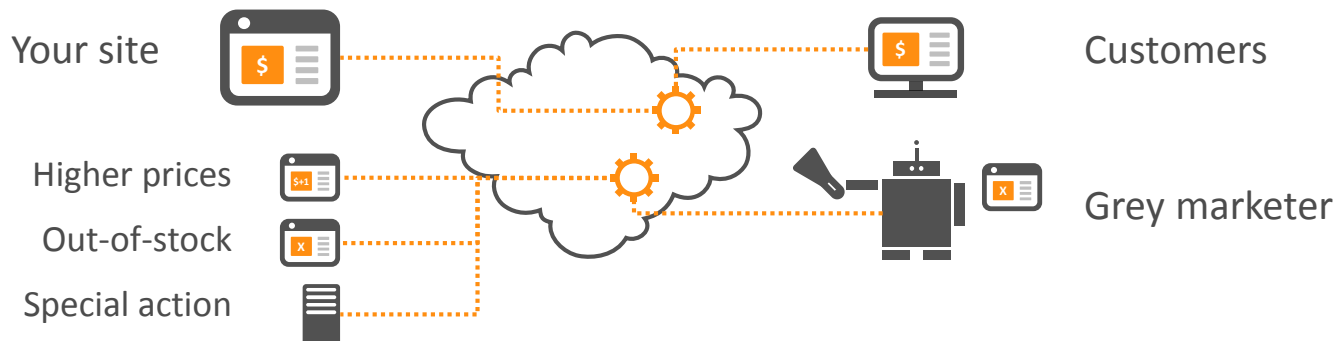
Hoodwink the Price Scrapers?



IMPACT: competitors automatically match your online prices, steal your customers and increase their sales at your expense

VALUE: feed competitors incorrect pricing to maintain competitive advantage, keep your customers and maximize your online sales

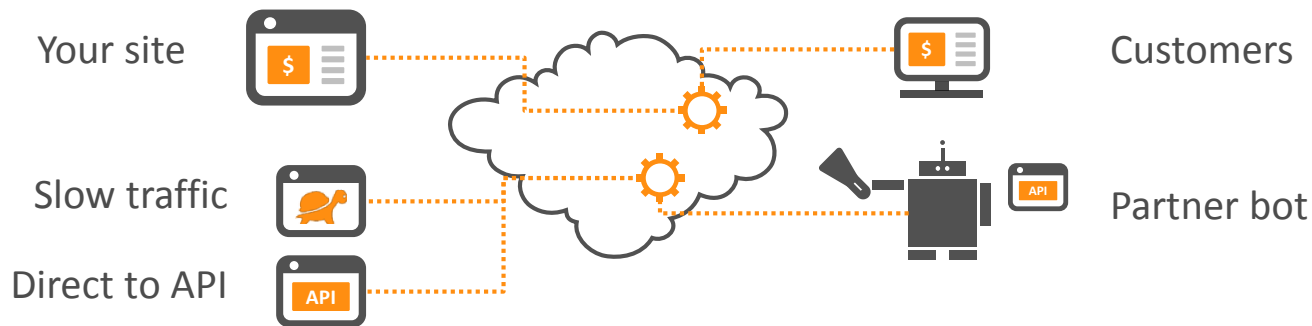
Fool the Grey Marketers?



IMPACT: 3rd-parties purchase or hold scarce inventory, preventing you from selling to real customers and reducing customer satisfaction

VALUE: slow, fool with higher prices or out-of-stock pages, or signal to your origin to take special action on transactions from 3rd-parties

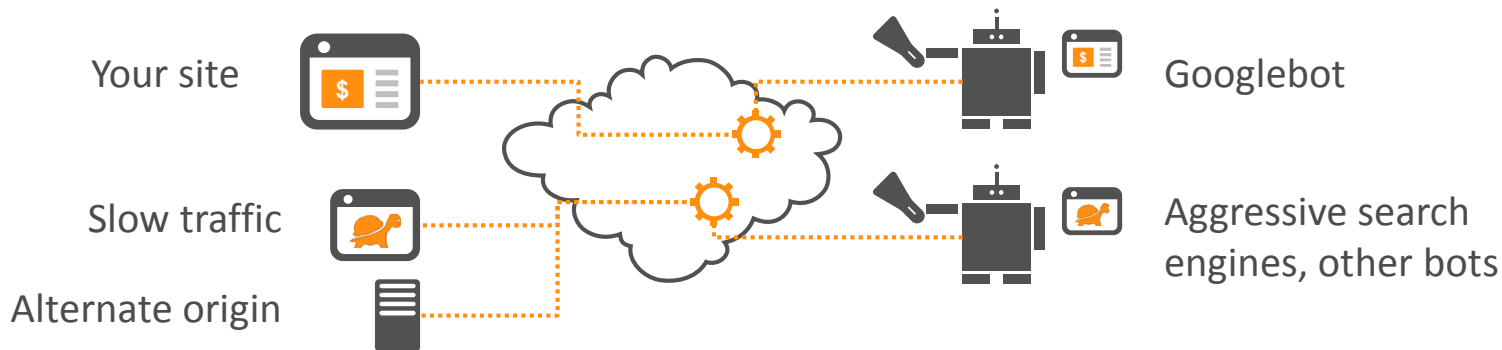
Make Your Partners Behave?



IMPACT: Your own partners are aggressively scraping your site to keep up with the latest pricing and content updates

VALUE: provide partners with the updates they need while reducing associated origin load, or redirect partners to use your API

Show Google the Love?



IMPACT: Heavy bot traffic and aggressive search engines (e.g., Baidu) reduce site performance for Googlebot

VALUE: Slow or direct aggressive search engines and other bots to alternate origin to maximize performance for Googlebot

Bot Management Is a PROCESS



EDUCATION

Concern

Has some understanding of a “bot problem”

Step back

Identify stakeholders with other “bot problems”

ANALYSIS

Visibility

Identify bot traffic and monitor behavior

Analysis

Identify different bots and understand impact

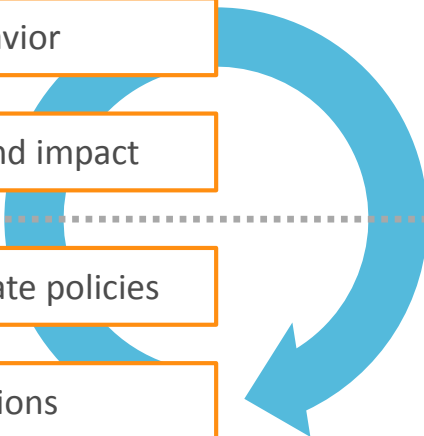
IMPLEMENT

Manage

Determine and implement appropriate policies

Measure

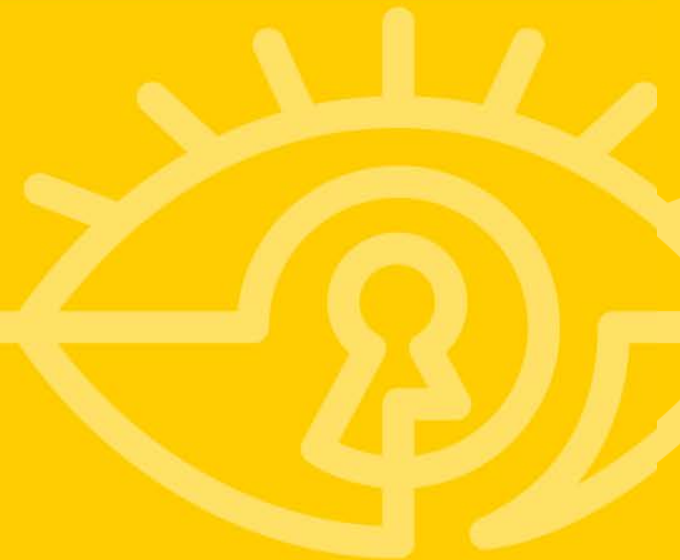
Measure results of management actions



What have we learned?



- Bots are **not only** a security problem. Business and IT are impacted
- There are **good** bots, **bad** bots, and '**so-so**' bots.
- ...and there is **always** a bot operator with a specific **target**
- Just blocking bots is the **wrong** thing to do
- The way you **detect**, **categorize** and **respond** to bots is key to get the benefit of the **good** bots and avoid the problem of the **bad** bots
- Managing bots is an **ongoing process**, not a one time event



Thank you!
Questions