



GoDaddy Web Application Firewall (WAF)

Module 5 Assignment Submission

Data Requirements

by
Neel Prajapati

CS6525 – Software Requirements Analysis
University of New Brunswick

DISCLAIMER: This report is not affiliated in any way with GoDaddy, nor does it intend to provide an accurate view of how GoDaddy conducts their business activities. This is a learning exercise at the University of New Brunswick based on unfounded assumptions and “best guesses” made by the author. The author has no inside knowledge of how GoDaddy’s software or business practices work. The objective is to simulate how requirements analysis might have been documented during the development of WAF for GoDaddy’s.

Table of Contents

1. Introduction	3
2. Research Description and Screenshots.....	4
3. Data Area – WAF Operations.....	8
Data Area Description	8
Entity Relationship Model.....	8
Entity & Attribute Descriptions (including sample entity instances)	9
4. Ternary Relationship Analysis.....	21
Selected ER Diagram and Three Entities.....	21
Ternary Relationship ER Diagram	22
Sample List of Triples	22
Six Cardinalities	23
Which is Appropriate: Ternary or Binary?	23

1. Introduction

This document provides the conceptual data requirements for GoDaddy's Web Application Firewall (WAF).

This document attempts to capture the functionality that is to be implemented in the first iteration of WAF for GoDaddy. At present GoDaddy serves hosting with small to medium scale traffic requirements, and as a result its customers did not require detailed control over how security was handled. As it moves in direction of hosting websites with large amounts of traffic coming from various sources, the active control afforded by WAF becomes essential. WAF will allow customers to set custom security rules based on request characteristics and derived metrics to protect their resources. To allow customers to improve their security posture continuously, WAF will also provide detailed analytics based on traffic received and filtered by the WAF.

It is also worth mentioning that this document does not deal with the structure or functionality of ancillary activities such as configuration, support, billing for services, etc.

NOTE: The attributes listed under entities are the once relevant for a standalone WAF functionality. As GoDaddy offers a lot of other services, it might require modifications of some entities listed if WAF were to be made compatible with those exiting services in an actual implementation.

This document presents the data requirements divided into 1 data area, as follows:

WAF Operations concerns data related to the GoDaddy customers such as WAF functionality that they have subscribed to, domains that are being hosted by GoDaddy for them and navigational paths that those websites have. It also deals with data of security rules set by customers, traffic received and filtered by the rules.

2. Research Description and Screenshots

As the Web Application Firewall (WAF) service is not currently offered by GoDaddy, the topic of implementing it for GoDaddy was inspired by the existence of similar functionality in major hosting and Content Delivery Network providers' product offerings. Resources made available by Google Cloud, AWS and Cloudflare have been helpful in getting insights into WAF. The documentation website of Cloudflare, their demo videos and articles amongst other resources have been major sources of knowledge for the assignments including this one. The data model for this assignment has also been largely derived from the WAF service provided by Cloudflare to its enterprise customers.




Following are the screenshots along with descriptions to better describe the research effort.

Note: Some screenshots captured from videos are not clear, as the demo videos were not of suitable resolution. To compensate for this, links to those videos that lead to the time of screenshot capture have been included.

- The entities **Customer**, **Enabled WAF Functionality** and **WAF Functionality** were based on the practice of service providers billing their customers for tailored solution that they request and only for the resources that they use. **Each Customer can be subscribed to many WAF functionalities**, the list of which is represented by a separate entity in the model, with each functionality appearing only once in it and being used by many customers.

Image shows Enterprise customers offered custom solutions with ability to add features.

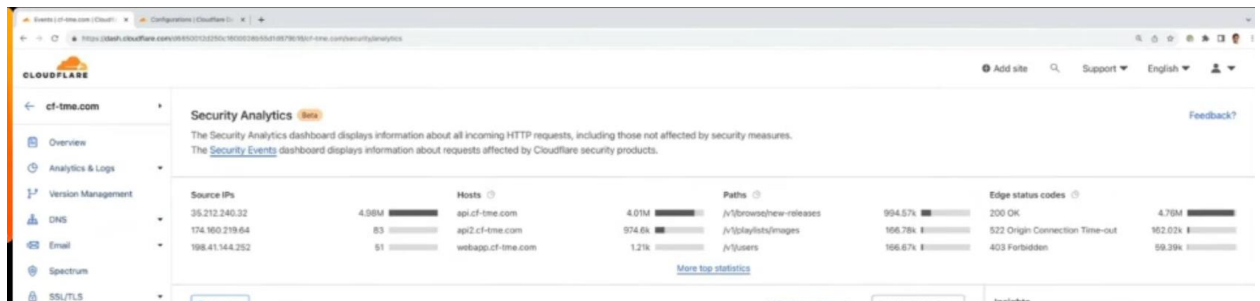
Source: <https://www.cloudflare.com/en-ca/plans/enterprise/#ent-plan-features>

	Business	Enterprise
Includes everything from the Pro plan & much more:	\$200/month <small>When billed annually or \$250/mo billed monthly</small>	Custom <small>Billed annually</small>
	Get Started	Talk to an Expert
Rate Limiting (IP-based)	 	Add-on
Rate Limiting Rules	 5	100 (Add-on)

- As different parts of websites receive different amounts of traffic, noting which paths, i.e., identifier to navigate to parts of website, receive how much traffic is essential. The same is used to control such traffic flow using security rules. For this reason, an **entity recording many paths associated with each domain** has been included.

Image shows paths (center right in image) by the amount of traffic received.

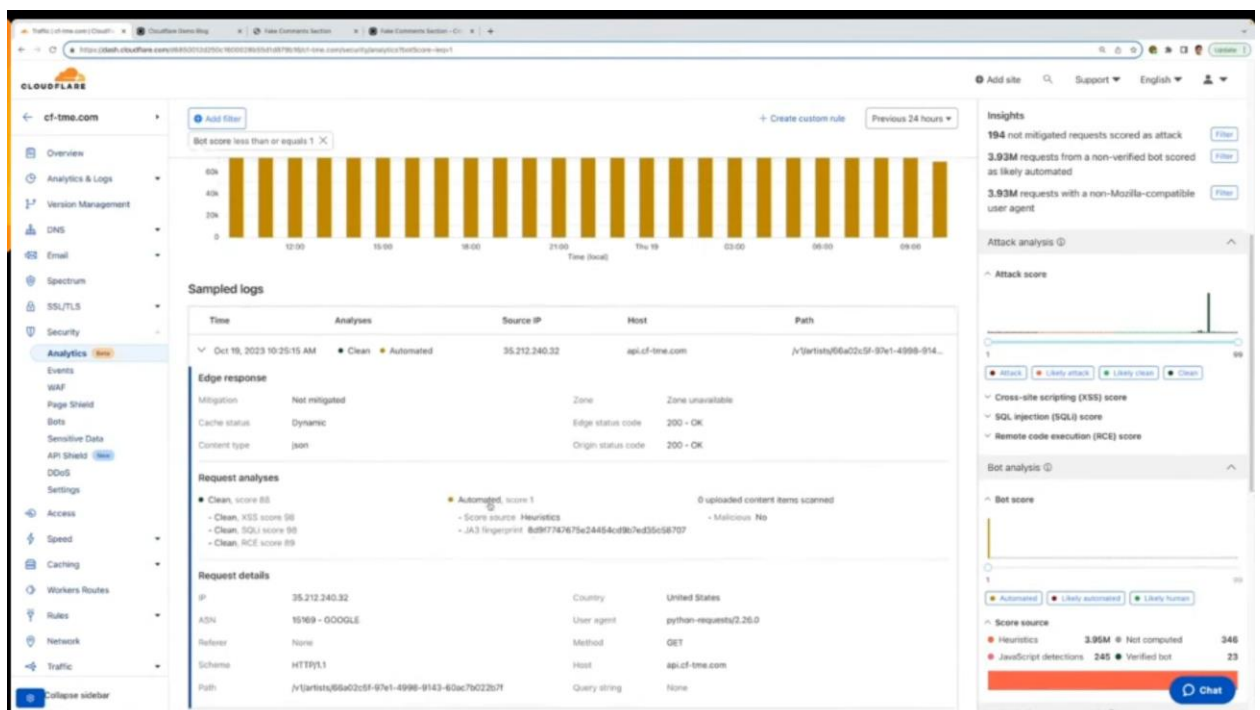
Source: https://youtu.be/0AjlS1JbNzE?si=ZKYqEK_QX56cI5F3&t=1824



- Logging some data of traffic is helpful for analytics and security. Cloudflare also performs security analytics on each request and allocates scores. Along the same lines each request's records are recorded by Request Traffic entity. **Each request is associated with a path and many such requests are generally received by each path.**

Image shows a logged request along with attributes logged, notice Request analysis section in the center left of the image with security scores.

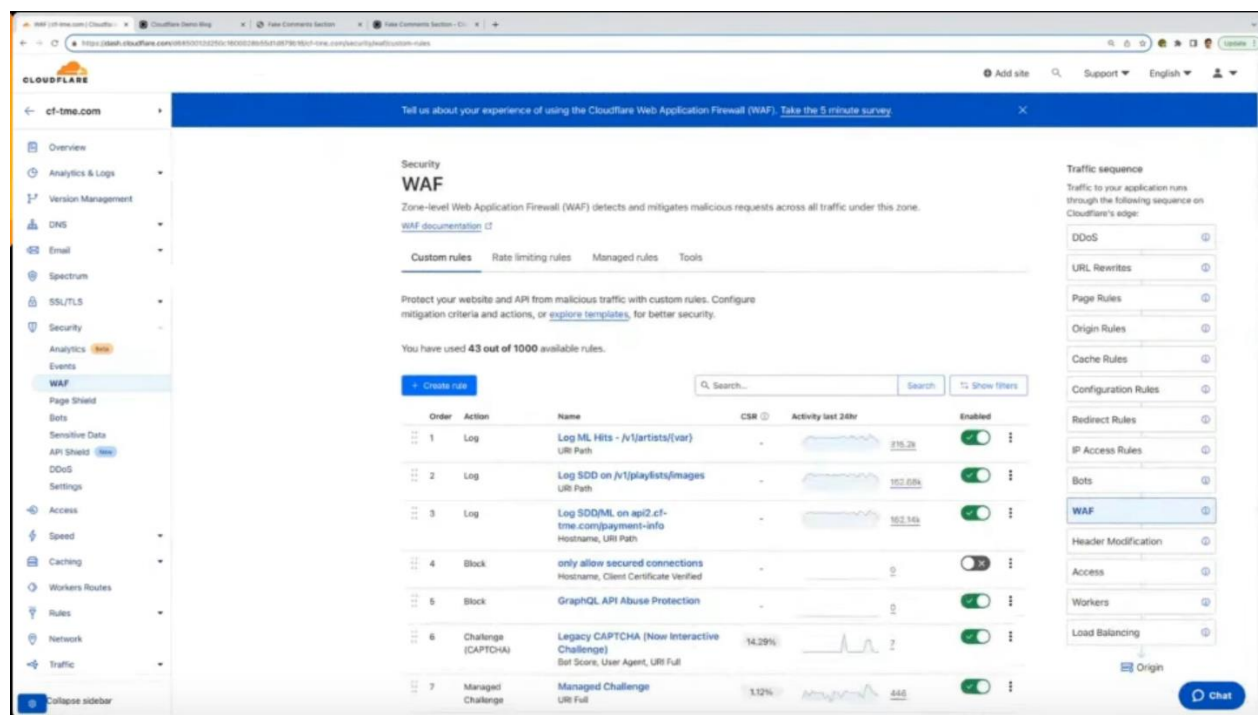
Source: <https://youtu.be/6EnekTohO7I?si=60pGIidHLcvDvYc&t=1708>



- WAF functions by applying rule expressions to requests being received on a customer's account. **Many rules form a ruleset of the customer account**, which is an ordered list, with rules being applied in the order configured by the customer. The rules are expressions written using Rule language with each having an associated action. When a rule expression is matched the action is taken and record is created in the Rule Matched Traffic entity along with payload. The attributes of this entity have not been combined with Request Traffic entity as all traffic is logged but rules are applied for only WAF users.

Image shows the Ruleset of an account.

Sources: <https://youtu.be/6EnekTohO7I?si=g5xzsGCo5xcma3Rk&t=1730>,
<https://developers.cloudflare.com/ruleset-engine/>



- Image shows a request that matched a rule.
Source: <https://youtu.be/6EnekTohO7I?si=vn4aD-pXDxgcYpox&t=2312>

The screenshot displays the Cloudflare dashboard for the domain **cf-tme.com**. The left sidebar shows the navigation menu with 'Events' selected under the 'Security' section. The main content area shows the 'Activity log' for 'Managed Challenge' events, filtered for the 'Previous 24 hours'.

Activity log table:

Date	Action taken	Country	IP address	Service
Oct 19, 2023 10:37:27 AM	Managed Challenge	United States	174.160.219.64	Custom rules

Below the activity log, the 'Matched service' section shows details for the event:

- Service: Custom rules
- Action taken: Managed Challenge
- Rule: ...n74b9e8
- Rule: ...n952f86

The 'Request analyses' section provides a breakdown of the request's security scores:

WAF Attack Score	Bot score	Bot source	Heuristics
90	1		
WAF XSS Attack Score: 98			
WAF SQL Attack Score: 98			
WAF RCE Attack Score: 91			

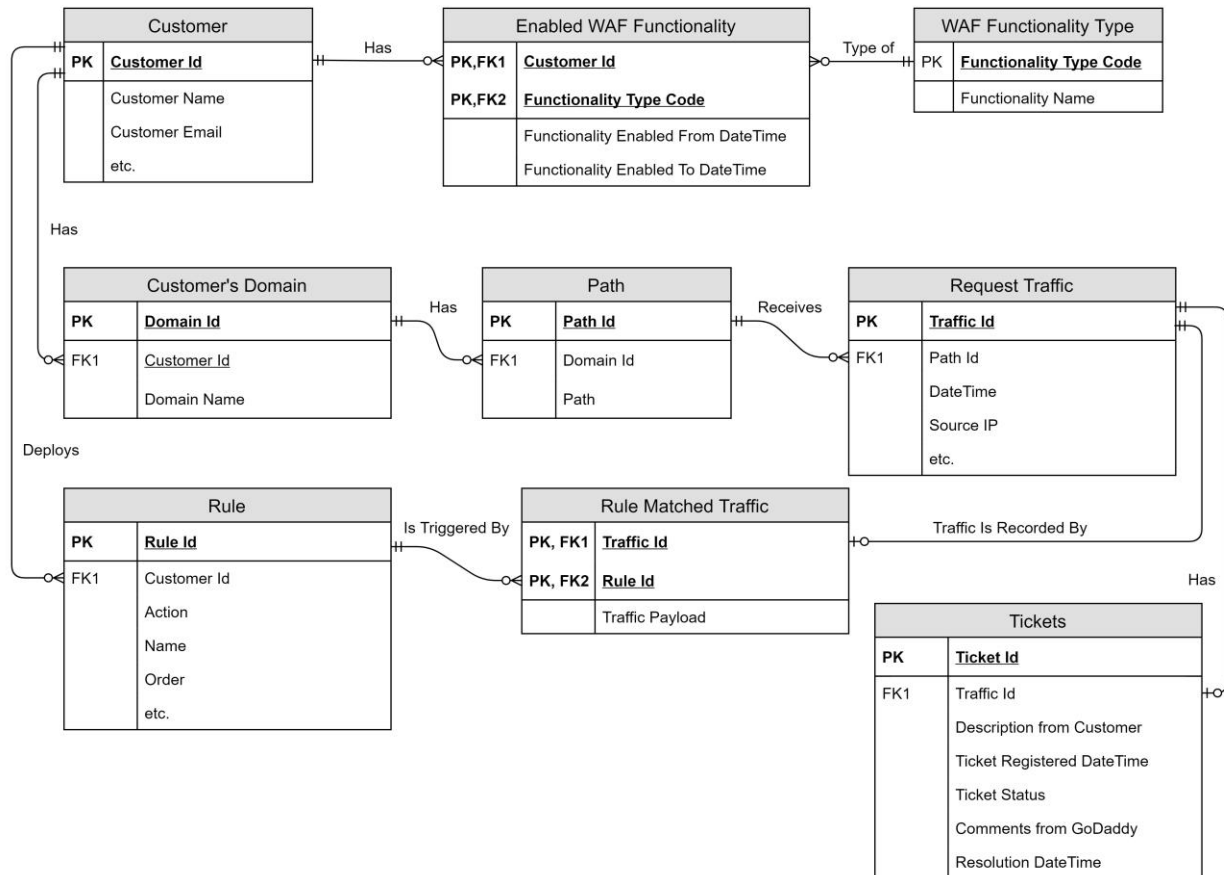
Additional details include the JA3 fingerprint (30b57b9b63283e7c0ad58b0b0d0b051f1) and the JavaScript verification status (Missing).

3. Data Area – WAF Operations

Data Area Description

This data area captures the entire primary data footprint of WAF functionality.

Entity Relationship Model



Entity & Attribute Descriptions (including sample entity instances)

Customer

Once a customer of GoDaddy registers for an account through the Enterprise Products team or through the website, their data is recorded in the Customer entity along with data of other customers. Although the list contains information of all the customers of GoDaddy, the attributes listed under customer are the once relevant for a standalone WAF functionality. Other GoDaddy services might require additional information related to a customer to be recorded in the Customer entity.

Attributes for: **Customer**

Attribute Name	Description (where the meaning is not obvious)
<u>Customer Id</u>	A unique identifier associated with each customer's account.
Customer Name	Name of the entity, person or the organization which operates the account.
Customer Email	Primary email associated with the customer account that is used for communication and login.
Customer Account Recovery Email	Recovery email to allow customers to access their account if the primary email becomes inoperative/inaccessible.
Customer Login Password	Password user for accessing the account.
Customer Phone	Optional: Phone number of the customer with country code.

List of entity instances for: **Customer**

No.	<u>Customer Id</u>	Customer Name	Customer Email
1	6EnekTohO7I	University of New Brunswick	it@unb.ca
2	0Ajs1JbNzE	Airbnb, Inc.	hostingadmin@airbnb.com

No.	Customer Account Recovery Email	Customer Login Password	Customer Phone
1	registrar@unb.ca	3m1!4b1!4m6!3	+1505666454
2	cybersec@airbnb.com	0x4ca42274e7	+1125999854

Enabled WAF Functionality

In this first iteration, the WAF individual functionality and underlying features are enabled for the customers. This entity records specifics of each such access granting transaction. The start and end time record allows billing to be done based on the duration of access.

Attributes for: **Enabled WAF Functionality**

Attribute Name	Description (where the meaning is not obvious)
<u>Customer Id</u>	Identifies the customer. This represents the relationship: <i>Customer Has/Had/Will have access to type Of Functionality</i>
<u>Functionality Type Code</u>	Identifies a single type of functionality with a unique code. This represents the relationship: <i>Functionality Type (is this) Type Of Functionality</i>
Functionality Enabled From DateTime	Records the time when the access to the functionality was enabled or will be enabled.
Functionality Enabled To DateTime	Records the time when the access to the functionality was disabled or will be disabled.

List of entity instances for: **Enabled WAF Functionality**

<u>Customer Id</u>	<u>Functionality Type Code</u>	Functionality Enabled From DateTime	Functionality Enabled To DateTime
6EnekTohO7I	1	2024-06-17 15:56:04.336640	
0Ajls1JbNzE	1	2024-01-12 21:58:24.000422	
0Ajls1JbNzE	2	2024-01-12 21:58:24.000423	2024-03-12 21:58:23.000000
0Ajls1JbNzE	3	2024-01-12 21:58:24.000424	

WAF Functionality Type

A list of WAF functionalities/features that could be enabled for each customer.

WAF functionality can include, for example:

- Bot Management
- Active Attack Mitigation
- Remote Browser Isolation
- Cloud Email Security
- Distributed Denial of Service (DDoS) Attack Protection
- Rate Limiting
- Content Analysis
- Interaction Less Challenge

Attributes for: **WAF Functionality Type**

Attribute Name	Description (where the meaning is not obvious)
<u>Functionality Type Code</u>	A unique identifier associated with each functionality type.
Functionality Name	A descriptive name associated with each WAF functionality type.

List of entity instances for: **WAF Functionality Type**

<u>Functionality Type Code</u>	Functionality Name
1	Bot Management
2	Rate Limiting
3	Content Analysis

Customer's Domain

All customers' accounts allow hosting of multiple websites associated with and identified by separate unique domains. This entity records such associations between customers and the domains that they are hosting with GoDaddy.

Attributes for: **Customer's Domain**

Attribute Name	Description (where the meaning is not obvious)
<u>Domain Id</u>	Unique identifier for each domain.
Customer Id	Identifies the customer. This represents the relationship: Customer <i>Has</i> Domain
Domain Name	Records unique domain name.

List of entity instances for: **Customer's Domain**

<u>Domain Id</u>	Customer Id	Domain Name
CNgEyxLvTRrj	6EnekTohO7I	unb.ca
WUkHgpu25w7P	0Ajls1JbNzE	airbnb.com
3KvnkJdaGxe	0Ajls1JbNzE	airbnb.ca
KSpGHKPHn2N3	0Ajls1JbNzE	airbnb.co.in

Path

Paths are used to navigate to different parts of the website. As the domain name identifies the website itself, paths are used in conjunction with the full domain to reach to appropriate webpages on the Internet. The entity Path records the path strings and associates them with their respective domain names. Each domain can have multiple paths pointing to different parts of the same website.

Attributes for: **Path**

Attribute Name	Description (where the meaning is not obvious)
<u>Path Id</u>	Unique identifier for each Path.
Domain Id	Identifies the domain. This represents the relationship: Domain <i>Has</i> Path
Path	Records path name for a domain name.

List of entity instances for: **Path**

<u>Path Id</u>	Domain Id	Path
b8cpC47bJpKRqFKDK8v6	CNgEyxLvTRrj	/academics/
gC3huyHLy4cRbMFDjwyN	CNgEyxLvTRrj	/admissions/
BdUmhr86hCGKkDHT3xWA	WUkHgpu25w7P	/login/
UXpne9gpzHxR9bADcUay	WUkHgpu25w7P	/canada/stays/

Request Traffic

The security ruleset of the account will be applied to all the request traffic received on the set of websites hosted using that customer's account. The Request Traffic entity records details of every single request. In addition to the basic information such as IP address and time, it also records derived information for each request such as JA3 fingerprint, Bot score, Attack score, Malicious content status, etc.

Attributes for: **Request Traffic**

Attribute Name	Description (where the meaning is not obvious)
<u>Traffic Id</u>	A unique identifier assigned to each request.
Path Id	Identifies the path and the domain which received the request. This represents the relationship: Path <i>Receives</i> Traffic
Source IP	IP address of the client as mentioned in the request.
Date/Time of request	The time when the request was received.
JA3 Fingerprint	JA3 fingerprint is a string of characters that is calculated using the client information received as part of the TLS hand sack such as cipher suite information of client. This helps to identify the requesting client even if the IP of that client changes. NOTE: This is not a full proof identification method.
Host	Host is the server which processed the request. GoDaddy has globally spread Content Delivery Network servers, and traffic requests are routed to them based on the customer configuration and load balancing. Therefore, two requests originating from same client at the same time could be processed by two different servers, known as hosts. This makes it important to record the identifying Id string of that host.
Autonomous system number (ASN)	ASN is number which identifies one of the subcomponent networks that make up the Internet. These networks are known as Autonomous Systems and are generally operated by Internet Service Providers(ISPs) and large organizations.
Country	Records country where the Source IP address is based in.
Method	HTTP POST or GET methods.

User Agent	Identifies the User Agent, which is the client software that is the source of the request traffic.
Malicious Content	When enabled the content analytics will scan the content of the traffic for detecting malware and store the result as Yes (Malware detected) or No (Malware not detected).
Mitigation Status	If the request matched any GoDaddy WAF rules than record Yes else No.
Bot Score	A score of 1(Highly likely) to 99(Not likely) depicting the likeliness of the request coming from an automated client.
Bot Score Source	The mechanism that generated the Bot score. The list of mechanism includes Heuristics, Machine learning, Anomaly detection, JavaScript detections, etc.
Attack Score	A score of 1(Highly likely) to 99(Not likely) depicting the likeliness of the request being a part of an attack.
Cross-site scripting (XSS) Score	A score of 1(Highly likely) to 99(Not likely) depicting the likeliness of the request being a part of a Cross-site scripting attack.
SQL injection (SQLi) Score	A score of 1(Highly likely) to 99(Not likely) depicting the likeliness of the request being a part of an SQL injection attack.
Remote code execution (RCE) score	A score of 1(Highly likely) to 99(Not likely) depicting the likeliness of the request being a part of a Remote code execution attack.

List of entity instances for: **Request Traffic**

No.	Traffic Id	Path Id	Source IP
1	SfnEDRz0RwsroCfFhEx8RV6DDNjrJt	BdUmhr86hCGKkDht3xWA	35.212.240.22
2	sv3C1uLnDqtDpmZL2yz6oKU8CET8Bm	gC3huyHLy4cRbMFDjwyN	174.160.219.64

No.	DateTime of request	JA3 Fingerprint
1	2024-06-17 15:56:04.336640	e4da3b7fbbce2345d7772b0674a318d5
2	2024-06-17 15:56:04.336643	a87ff679a2f3e71d9181a67b7542122c

No.	Host	Autonomous system (ASN)	Country
1	api.cf-tme.com	15169	United States
2	webapp.cf-tme.com	855	Canada

<u>No.</u>	Method	User Agent	Malicious Content	Mitigation Status
1	GET	Python-requests/2.30	NO	YES
2	POST	Mozilla/5.0	NO	NO

<u>No.</u>	Bot Score	Bot Score Source	Attack Score	Cross-site scripting (XSS) Score
1	1	Heuristics	96	99
2	78	Machine Learning	89	73

<u>No.</u>	SQL injection (SQLi) Score	Remote code execution (RCE) score
1	89	99
2	99	99

Rule

The ruleset is structured as an ordered list containing rule expressions, which are applied to every request received for that account in the selected order. Individual rules are a combination of a rule expression and an action. When the ruleset is applied to a request and a rule expression is matched the action is applied. A rule can be a simple expression, or a compound expression formed using logical operators and sub expressions. The actions supported are Interactive Challenge, JS Challenge, Managed Challenge (GoDaddy Selected Challenge), Block, Skip, Log and Execute.

Attributes for: **Rule**

Attribute Name	Description (where the meaning is not obvious)
<u>Rule Id</u>	A unique identifier associated with each Rule in GoDaddy WAF system.
Customer Id	Identifies the customer. This represents the relationship: Customer <i>Deploys</i> Rule
Name	Name of the Rule as set by the customer.
Order	Order(position) of Rule in the Ruleset. The values of order for a customer account starts from 1(First in sequence) and increases.
Rule Expression	The Expression created by system using Rules language of GoDaddy and based on options selected by customer for the rule.
Action	Action that is applied when the rule expression is true.
Rule Start DateTime	Date and time when the Rule was/will be enabled.
Rule End DateTime	Date and time when the Rule was/will be disabled.

List of entity instances for: **Rule**

No.	<u>Rule Id</u>	<u>Customer Id</u>	Name
1	MziPTE29JUC12FWbIpM8	6EnekToh07I	Block Bad Bot
2	f2kHpwEa2M8Zc7kSQPPc	6EnekToh07I	Log Auto Discover path
3	zGn01vXJgXEAJkL1XTUb	6EnekToh07I	Block request to example.com on port 80, 443
4	9aTdlgMZuRnfPzMqKSsT	6EnekToh07I	Skip rules for known countries

No.	Order	Rule Expression	Action
1	01	http.user_agent eq "BadBot/1.0.2 (+http://bad.bot)	Block
2	02	http.request.uri.path matches "/autodiscover\.(xml src)\$"	Log
3	03	host eq www.example.com and not cf.edge.server_port in {80 443}	Block
4	04	(ip.geoip.country in {"CN" "TW" "US" "GB"}) and cf.threat_score gt 0)	Skip

No.	Rule Start DateTime	Rule End DateTime
1	2024-06-17 15:56:04.336640	
2	2024-06-17 15:56:04.336640	2024-06-17 15:56:04.336640
3	2024-06-17 15:56:04.336640	
4	2024-06-17 15:56:04.336640	2024-06-17 15:56:04.336640

Rule Matched Traffic

All requests received for a customer's hosted websites are applied the rules of ruleset. When a Rule is matched it is recorded in this entity Rule Matched Traffic.

Attributes for: **Rule Matched Traffic**

Attribute Name	Description (where the meaning is not obvious)
<u>Traffic Id</u>	Identifies the request amongst all the traffic. This represents the relationship: <i>Matched Request Traffic Is Recorded By Rule Matched Traffic</i>
<u>Rule Id</u>	Identifies the Rule that was matched. This represents the relationship: <i>Rule Is Triggered By Request Traffic</i>
Traffic Payload	Payload stored only if such action is configured by customer.

List of entity instances for: **Rule Matched Traffic**

No.	<u>Traffic Id</u>	<u>Rule Id</u>	Traffic Payload
1	SfnEDRz0RwsroCfFhEx8RV6DDNjrJt	MziPTE29JUC12FWbIpM8	-
2	sv3C1uLnDqtDpmZL2yz6oKU8CET8Bm	f2kHpwEa2M8Zc7kSQPPc	-

Tickets

Customers can raise a Ticket if a False Positive or False Negative is found in traffic. The information of the ticket is to be recorded in this entity.

Attributes for: **Tickets**

Attribute Name	Description (where the meaning is not obvious)
<u>Ticket Id</u>	A unique identifier associated with each ticket.
Traffic Id	Identifies the request for which the ticket was raised. This represents the relationship: Request Traffic <i>Has</i> Ticket
Description from Customer	
Ticket Registered DateTime	
Ticket Status	Status of the ticket as Open/ Resolved/ Escalated
Comments from GoDaddy	
Resolution DateTime	

List of entity instances for: **Tickets**

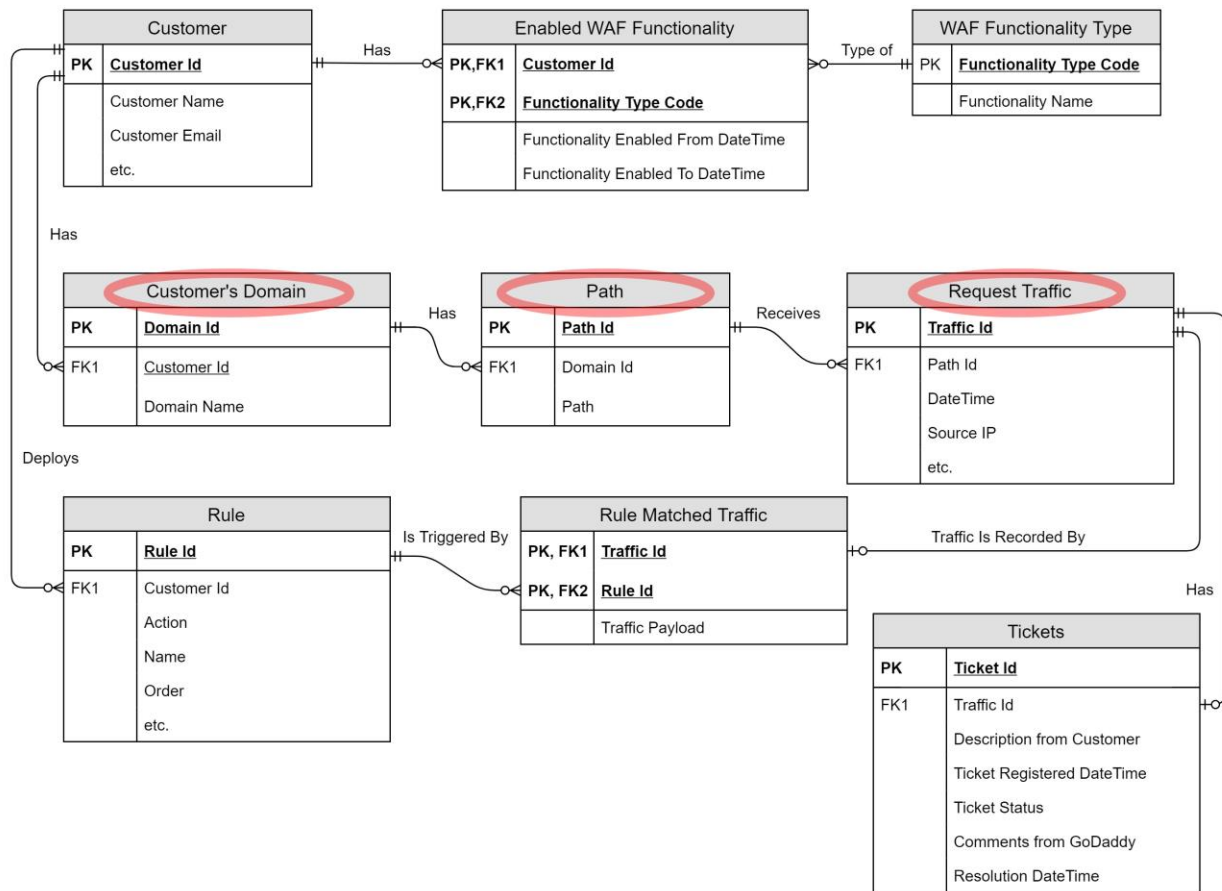
<u>No.</u>	<u>Ticket Id</u>	Traffic Id
1	3605794506	SfnEDRz0RwsroCfFhEx8RV6DDNjrJt
2	2902060333	sv3C1uLnDqtDpmZL2yz6oKU8CET8Bm

<u>No.</u>	Description from Customer	Ticket Registered DateTime	Ticket Status
1	This request was a False Positive.	2024-03-13 14:52:01.000120	Resolved
2	This attack was not mitigated by GoDaddy, please investigate.	2024-06-17 15:56:04.336640	Open

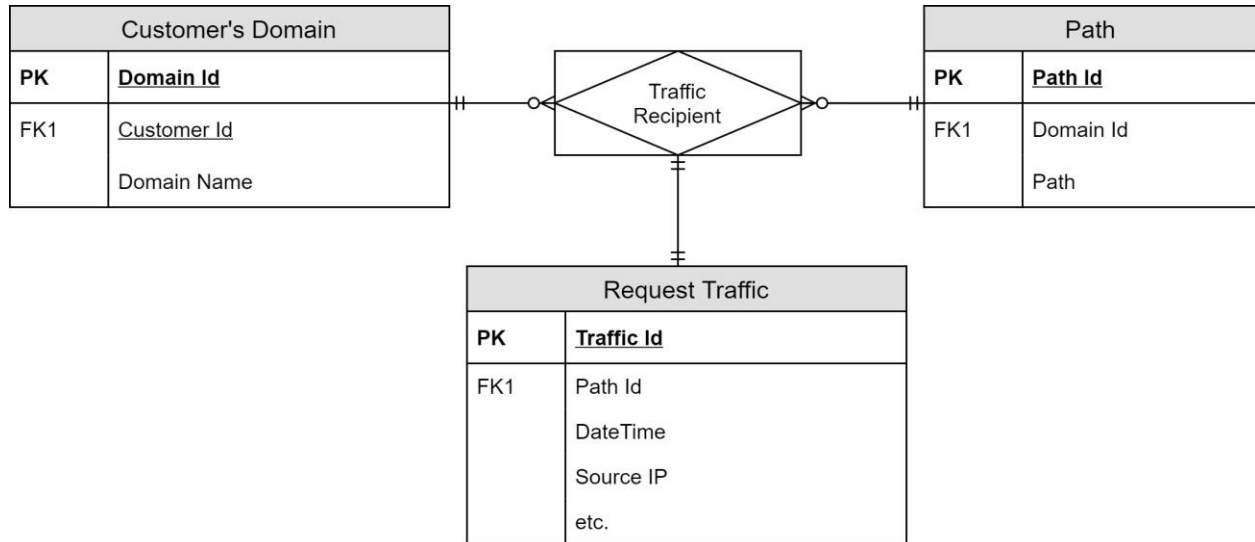
<u>No.</u>	Comments from GoDaddy	Resolution DateTime
1	The new bot signature that was noticed in this request has been inserted into the Bot management.	2024-03-14 23:52:07.044520
2	-	-

4. Ternary Relationship Analysis

Selected ER Diagram and Three Entities



Ternary Relationship ER Diagram



Sample List of Triples

Sample list of **Traffic Recipients** instances

Domain Id	Path Id	Traffic Id
101	2001	30001
101	2001	30002
101	2002	30003
102	2003	30004

Note: The ID values have been simplified for reader's convenience.

Six Cardinalities

For a given instance of **Customer's Domain**, the maximum number of associated **Path** instances is: [Many](#)

For a given instance of **Path**, the maximum number of associated **Customer's Domain** instances is: [1](#)

For a given instance of **Customer's Domain**, the maximum number of associated **Request Traffic** instances is: [Many](#)

For a given instance of **Request Traffic**, the maximum number of associated **Customer's Domain** instances is: [1](#)

For a given instance of **Path**, the maximum number of associated **Request Traffic** instances is: [Many](#)

For a given instance of **Request Traffic**, the maximum number of associated **Path** instances is: [1](#)

Which is Appropriate: Ternary or Binary?

In this instance of determining a relationship between **Customer's Domain**, **Path** and **Request Traffic**; two Binary relationships are appropriate over a single Ternary relationship. Looking at the six cardinalities, 3 questions have an answer 1. As per the instructional video, when any of the answers to the questions is 1, Binary relationship is better suited over Ternary relationship. As simple is better, in having two simple separate relationships between **Customer's Domain & Path** and **Path & Request Traffic**, we do not lose any information. When given a Traffic Id, it is always possible to trace back to Domain Id through Path Id.