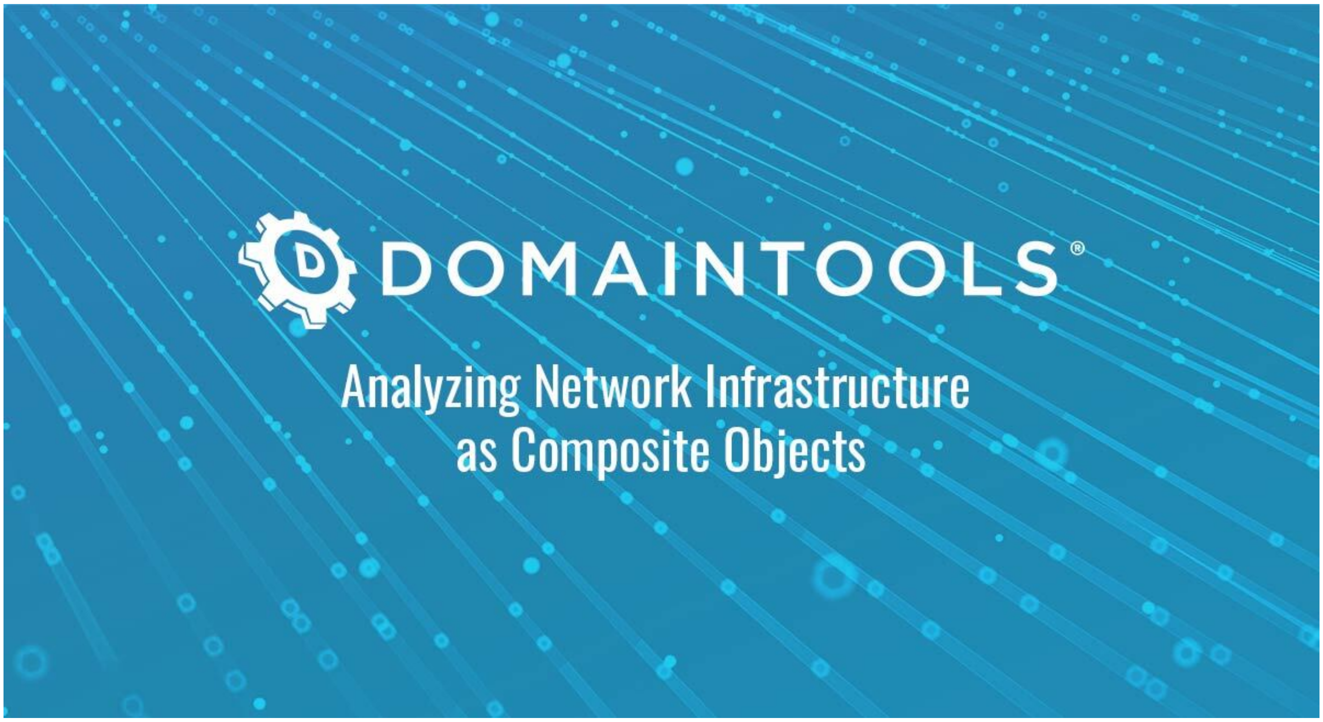
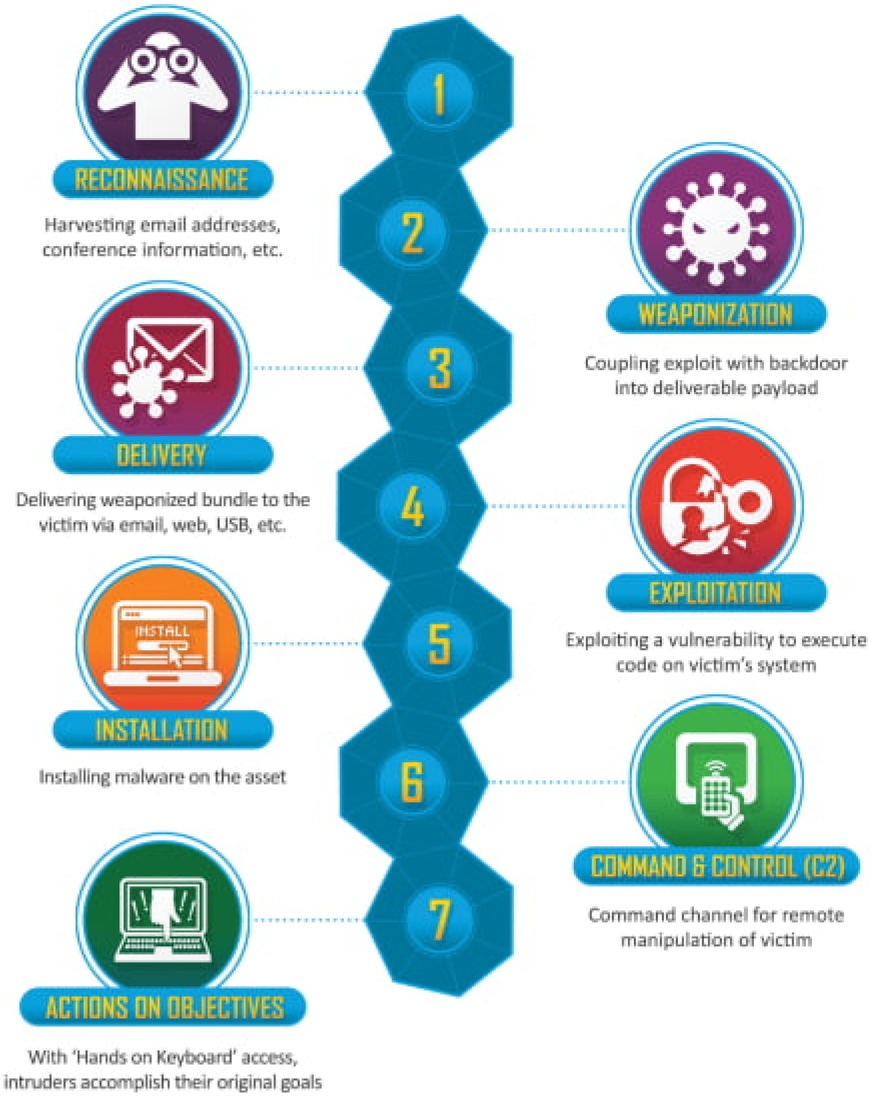
Analyzing Network Infrastructure as Composite Objects

domaintools.com/resources/blog/analyzing-network-infrastructure-as-composite-objects



Introduction

Network infrastructure is one of the primary observables associated with cyber intrusions. From IP addresses serving as scanning or reconnaissance infrastructure through domains functioning as command and control (C2) or exfiltration servers, network infrastructure forms a prerequisite for adversary operations. Viewed through typical models such as the Lockheed-Martin Cyber Killchain, (shown below) network infrastructure observables factor into nearly every stage of the intrusion lifecycle as either a critical dependency or an enabling factor.


1
RECONNAISSANCE
Harvesting email addresses, conference information, etc.
2
WEAPONIZATION
Coupling exploit with backdoor into deliverable payload
3
DELIVERY
Delivering weaponized bundle to the victim via email, web, USB, etc.
4
EXPLOITATION
Exploiting a vulnerability to execute code on victim's system
5
INSTALL
INSTALLATION
Installing malware on the asset
6
COMMAND & CONTROL (C2)
Command channel for remote manipulation of victim
7
ACTIONS ON OBJECTIVES
With 'Hands on Keyboard' access, intruders accomplish their original goals

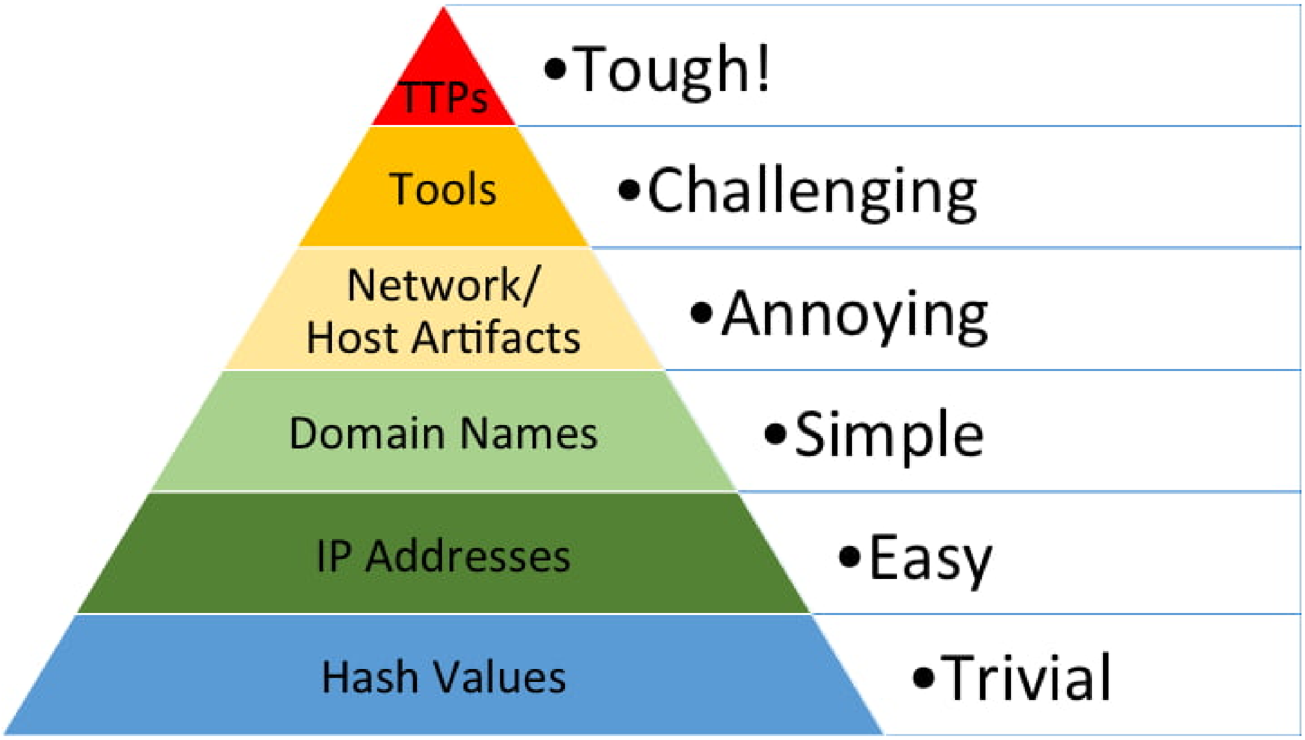
Yet while commonly referred to and almost universally present, network infrastructure observables—specifically domain names and IP addresses—are also frequently derided as atomic, minimally enriched items for defensive purposes. This paper will show that such a view is mistaken and misguided—but only if we expand our understanding of network infrastructure observables and their characteristics.

A thorough examination of network observables shows differentiation between minimally enriched indicators and composite objects which enable more in-depth understanding and analysis. By adopting the latter view, a seemingly atomic object such as a domain name yields a number of linked observations, which enable further analysis and pivoting. By following this methodology, defenders can use a relatively small set of observations to build a robust picture of adversary tendencies and unearth additional campaigns and adversary observations.

Indicators and Atomic Objects

Information security operations live through the ingestion, analysis, and disposition of technical observations. Frequently called “indicators of compromise” (IOCs), these items are in theory composite objects consisting of observations, descriptions, and metadata. Yet in practice, “IOC” refers to a debased, diminished version of the original concept. Rather than containing built-in contextuality, IOCs instead are reduced to mere observables in isolation: an IP address, a domain name, a hash value.

While (debased) IOCs still drive much of everyday network security operations, they are increasingly derided by analysts and industry representatives. Examples include calls to emphasize adversary behaviors over specific technical observations for defense, or theoretical constructs such as the Pyramid of Pain representing the “staying power” of different types of observations.

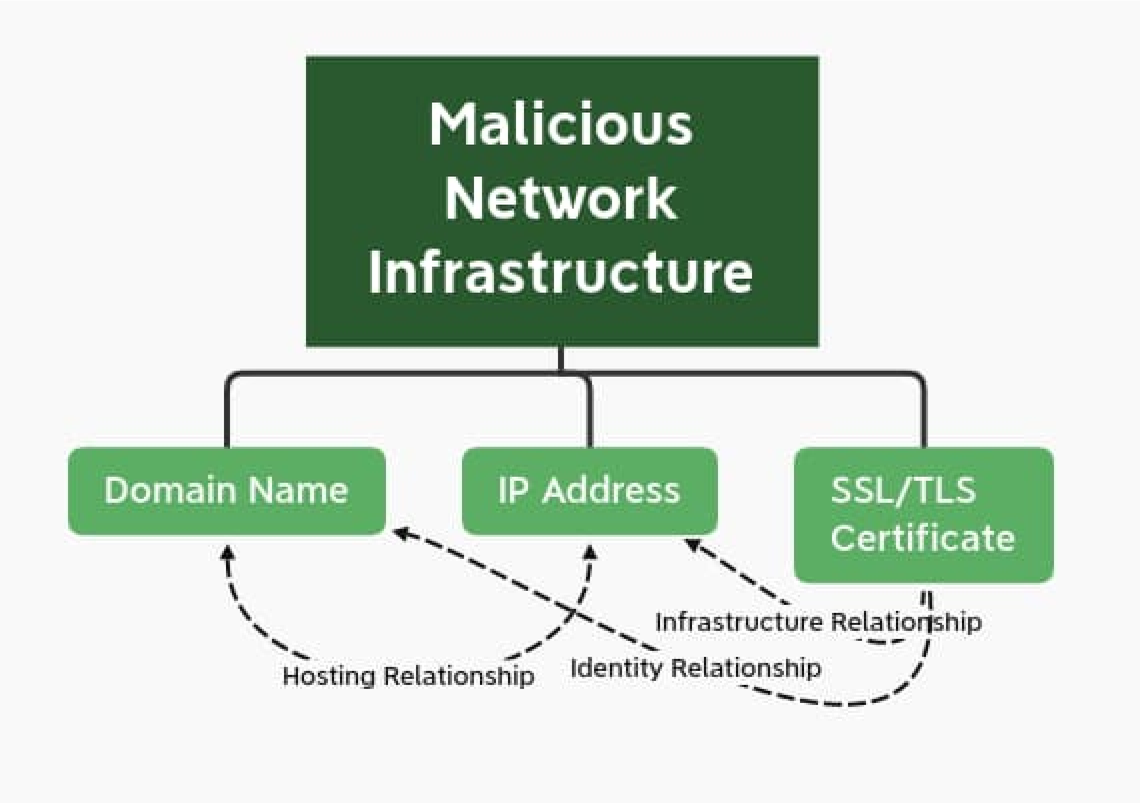


While these arguments are generally correct and insightful for improving the practice of information security, such developments come with an implicit and often ignored cost. In the rush to embrace behavior-based, tactics-techniques-procedures (TTP) focused defense, the value of indicators and their nature may be left behind.

A more thorough understanding of just what exists within an indicator allows us to explore in greater detail the nature and basis for that indicator’s existence. Just as an atom, while representing the fundamental building block of matter, contains subatomic particles that define its characteristics, the same is true for bare indicators. With further analysis and enrichment, we can discover greater details and gain an improved understanding of these observations.

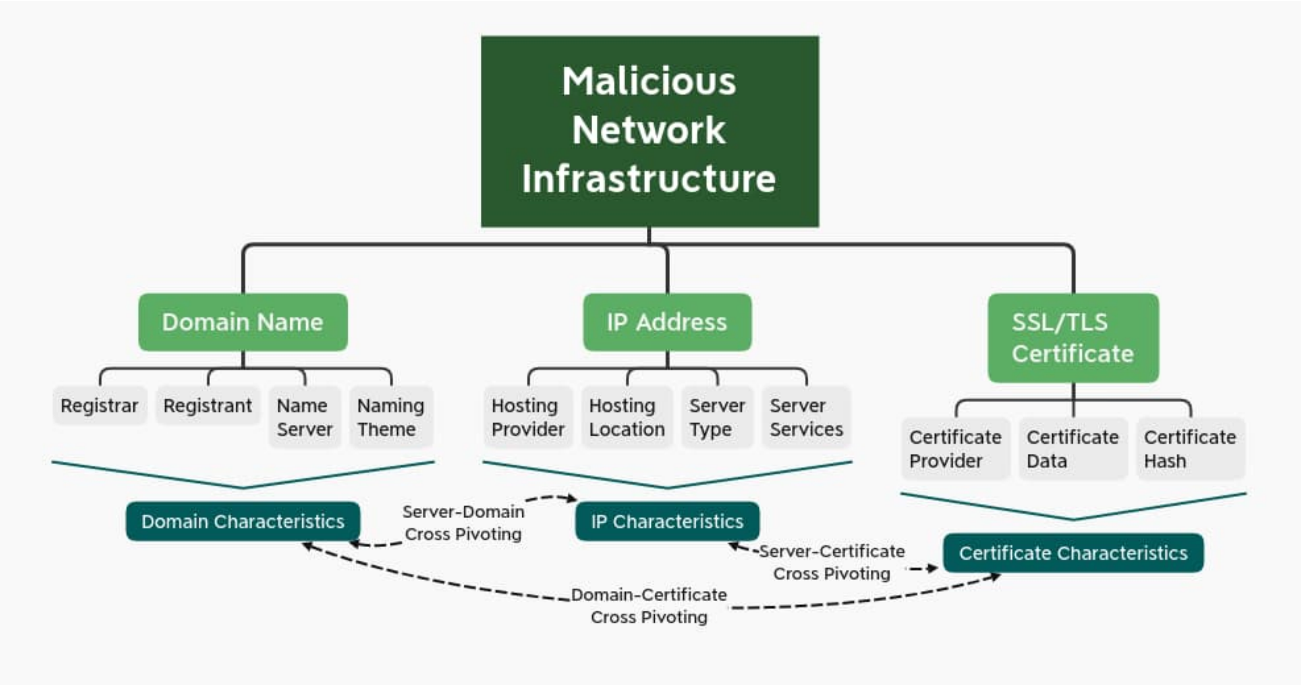
Nature of Network Infrastructure

Network infrastructure observables are those artifacts related to intrusion events or adversary activity linked to delivery, communication, control, and exfiltration among other items. Although not exhaustive, examples of network infrastructure observables include domain names, IP addresses, and SSL/TLS certificates. As shown below, these items are interrelated as they pertain to different aspects of the same overall communication scheme: an IP hosts a domain that uses an SSL/TLS certificate to encrypt traffic.



At first glance, these items appear to be unitary, atomic indicators. As such, they would appear to require enrichment and context outside of themselves to have lasting, meaningful value for understanding adversary behaviors and tendencies. However, further investigation of these items indicates a more complex nature with multiple subcomponents and characteristics that identify these items, under proper analysis, as composite objects.

As shown in the updated image, infrastructure observables contain a wealth of data when properly analyzed and observed. Extending the comparison to atoms made above, each type of network observable effectively “breaks down” into a mixture of subcomponents. Understanding and analyzing these items, their relationships, and patterns of composition yields insights into adversary behaviors which extends and deepens the value of a “bare” network indicator.



Domain Names

A domain name is a natural language, human-readable item designed as a reference to a machine-focused IP address hosting content online. While content or services can be accessed simply through an IP address, domains facilitate the process and allow for a certain degree of "branding" or uniqueness through their name.

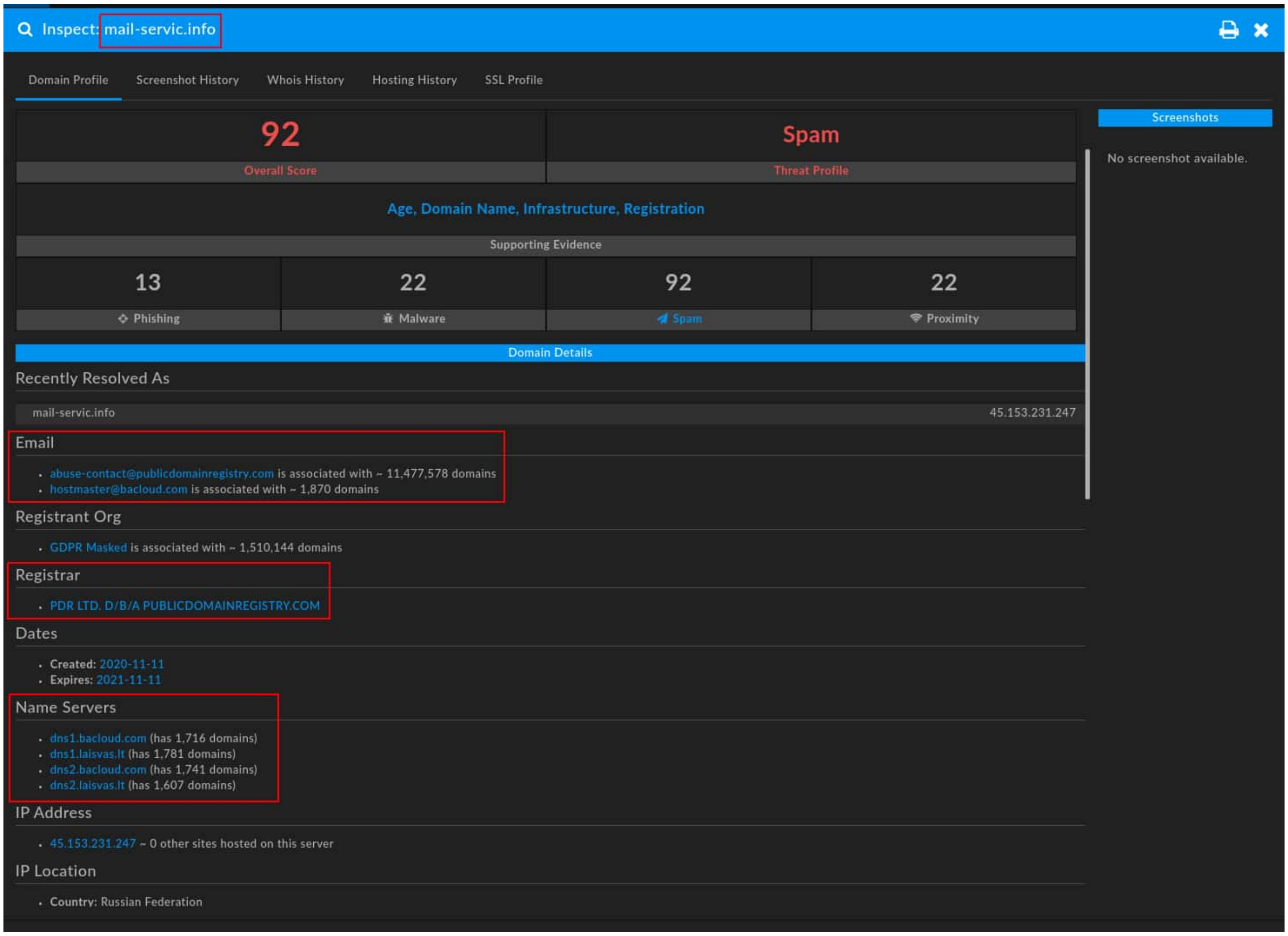
The domain itself is not a unitary object though. Rather, the domain consists of several components or identifying metadata that can be used to "fingerprint" or gain further insight into the domain's nature or creation. In addition to any hosting information related to the domain and its use, covered below, the following items represent characteristics of a domain that analysts and defenders can leverage:

- **Domain Registrar:** In order to create and take ownership of a domain, an individual or entity needs to work through a registrar to secure a domain through one of the registries managing the desired Top Level Domain (TLD - e.g., ".com"). Registrars differ widely in terms of pricing, client scrutiny, and other aspects. As a result of these characteristics and infrastructure preferences, threat actors may prefer or primarily leverage certain registrars over others for infrastructure creation.
- **Domain Registrant:** Domains are created by a given registrant. While this information was historically quite useful, as such information would include contact email addresses and other information that could be used to fingerprint infrastructure creation, the increasing adoption of privacy protection services and the impact of the European Union's General Data Protection Regulation (GDPR) have greatly restricted such information at present. Nonetheless, commonality in privacy protection services across registrations can still be used as a weak link to tie together various domains.

- **Name Server:** Domain resolution to an IP address requires an authoritative name server in order to translate requests. Identifying name servers associated with registration—especially specific authoritative servers—can reveal patterns of infrastructure creation and adversary tendencies.
- **Domain Naming Theme or Convention:** In one of my previous posts, I described how actual domain name selection may be used to infer adversary intent as well as adversary tendencies. Threat actors must pick something for a domain name, whether this is a randomly-generated string, an item matching a theme, or a name matching a target or campaign. Identifying these themes or conventions can be a surprisingly useful mechanism to differentiate domain registrations and identify commonalities for an actor.

The totality of these above items defines a domain. Just as they are components that represent the domain, they are also items that can be used to search for similarly-structured or created infrastructure. For example, an adversary may consistently use the same combination of name server, registrar, and registration privacy protection service that can enable pivoting and identification of additional adversary infrastructure.

As seen in the DomainTools Iris Investigate screenshot below, these items can be readily identified and used for pivoting purposes.



Inspect: mail-servic.info

No screenshot available.

92	Spam
Overall Score	Threat Profile

Age, Domain Name, Infrastructure, Registration

Supporting Evidence

13	22	92	22
Phishing	Malware	Spam	Proximity

Domain Details

Recently Resolved As

mail-servic.info 45.153.231.247

Email

- abuse-contact@publicdomainregistry.com is associated with ~ 11,477,578 domains
- hostmaster@bacloud.com is associated with ~ 1,870 domains

Registrant Org

- GDPR Masked is associated with ~ 1,510,144 domains

Registrar

- PDR LTD. D/B/A PUBLICDOMAINREGISTRY.COM

Dates

- Created: 2020-11-11
- Expires: 2021-11-11

Name Servers

- dns1.bacloud.com (has 1,716 domains)
- dns1.laisvas.lt (has 1,781 domains)
- dns2.bacloud.com (has 1,741 domains)
- dns2.laisvas.lt (has 1,607 domains)

IP Address

- 45.153.231.247 ~ 0 other sites hosted on this server

IP Location

- Country: Russian Federation

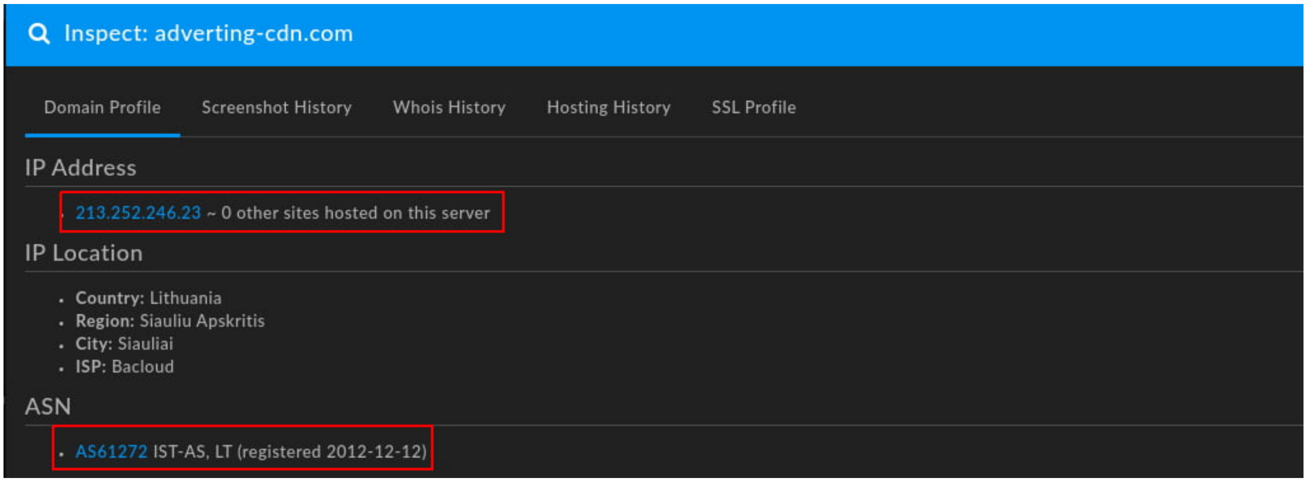
IP Addresses

IP addresses are the identifying schema for Internet Protocol (IP) traffic and are used to designate specific machines or servers to receive traffic. While domain names are not necessary for network communication, IP addresses are required for IP-based communication. An active, communicating domain will always be paired with an IP address, while an IP address need not have a related domain to ensure communication between hosts.

While IP addresses are one of the most commonly seen indicators in CTI reporting, just like domains IP addresses hide multiple subcomponents that identify aspects of adversary tendencies and behaviors:

- **Hosting Provider:** Adversaries need to find reasonably private, non-attributable hosting for network infrastructure. Options include any of the major cloud service providers from Amazon Web Services to DigitalOcean; smaller virtual private server (VPS) providers; or utilizing services such as CloudFlare to mask true hosting from monitoring parties.
- **Hosting Location:** In addition to hosting providers, threat actors also have a degree of choice over hosting location. Cloud, VPS, and other providers typically own infrastructure located in various countries. Adversaries can leverage location specificity for purposes ranging from avoiding potential geographic-based traffic filtering to taking advantage of the legal system of the hosting country to maximize privacy or make defender investigations more difficult.
- **Server Type:** Infrastructure still needs a system on which to run, and the choice of operating system (OS) and version can also be used to fingerprint adversary tendencies. Threat actors can decide between various flavors of Linux to different versions of Windows for the underlying OS. Identifying particular tendencies—especially when related to exposed system services, described below—can reveal patterns of activity that can be used to identify or disposition new infrastructure.
- **Server Services:** To function as a command and control (C2) or other node, a server must listen on some service. The most direct and basic would be HTTP or HTTPS, in which case we as defenders can identify the web server type, version, and, in the case of HTTPS, server SSL/TLS certificates (described further below). Identifying non-standard or atypical services, especially for unique or custom C2 frameworks, can further enable identification and tracking.

To illustrate the above concepts, the suspicious domain “adverting-cdn[.]com” is hosted on a dedicated server at 213.252.246[.]23. Within DomainTools Iris Investigate, we can identify the IP address, hosting provider, and hosting location:



IP Address

- 213.252.246.23 ~ 0 other sites hosted on this server

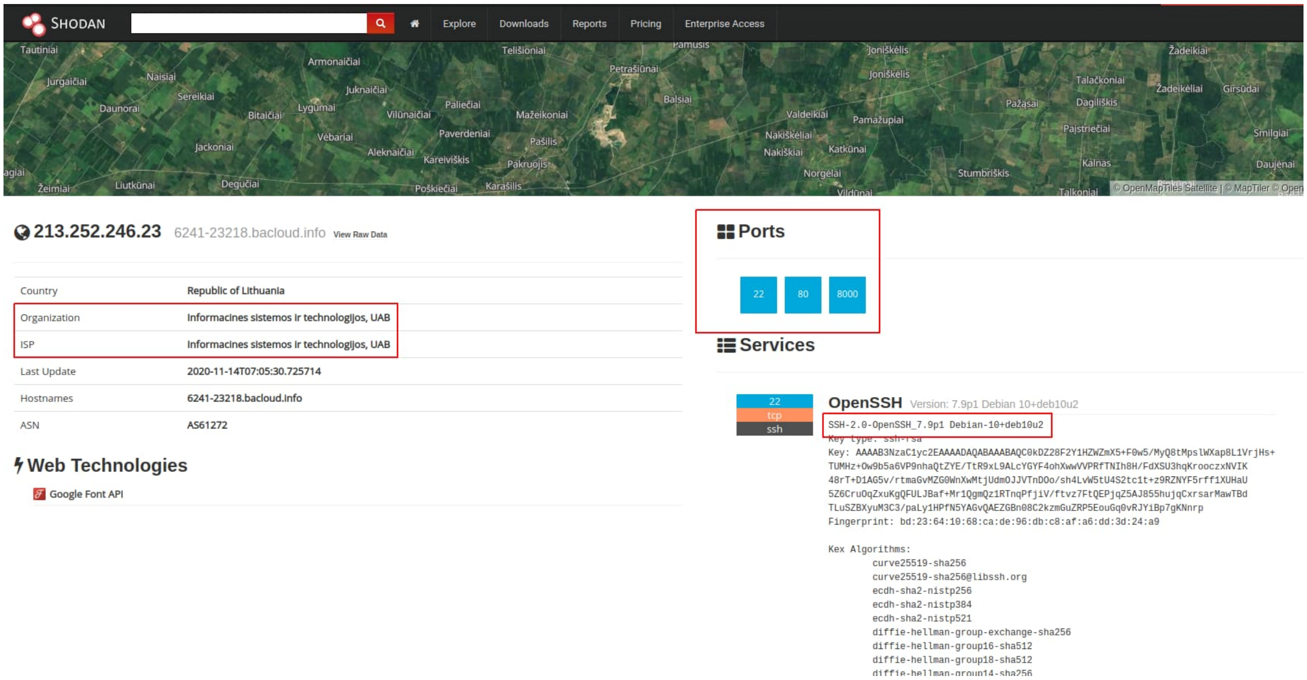
IP Location

- Country: Lithuania
- Region: Siauliu Apskritis
- City: Siauliai
- ISP: Bacloud

ASN

- AS61272 IST-AS, LT (registered 2012-12-12)

Using an internet scanning and enumeration tool such as Shodan, we can further investigate the server to identify services and fingerprint the server's OS:



213.252.246.23 6241-23218.bacloud.info View Raw Data

Country	Republic of Lithuania
Organization	Informacines sistemos ir technologijos, UAB
ISP	Informacines sistemos ir technologijos, UAB
Last Update	2020-11-14T07:05:30.725714
Hostnames	6241-23218.bacloud.info
ASN	AS61272

Web Technologies

Google Font API

Ports

22 80 8000

Services

22 tcp ssh

OpenSSH Version: 7.9p1 Debian 10+deb10u2

```
SSH-2.0-OpenSSH_7.9p1 Debian-10+deb10u2
Key type: ssh-rsa
Key: AAAAB3NzaC1yc2EAAAADAQABAAABAQC0kDZ28F2Y1HZWZmX5+F0w5/MyQ8tMpslWXap8L1VrjHs+
TUMHz+Ow9b5a6VP9nhaQtZYE/TtR9xL9ALcYGYF4ohXwwVVPRfTNIh8H/FdXSU3hqKrooczxNVIK
48rT+D1AG5v/rtmaGvMZG0WnXwMtjUdmOJJVTnDOo/sh4LvW5tU4S2tc1t+z9RZNYF5rff1XUHaU
5Z6CruOqZxuKgQFULJBaf+Mr1QgmQz1RTnqPfjiV/ftvz7FtQEPjqZ5AJ855hujqCxrsarMawTBd
TLuSZBXyuM3C3/paLy1HPfN5YAGvQAEZGBn08C2kzmGuZRP5EouGq0vRJYiBp7gKNnrp
Fingerprint: bd:23:64:10:68:ca:de:96:db:c8:af:a6:dd:3d:24:a9

Kex Algorithms:
        curve25519-sha256
        curve25519-sha256@libssh.org
        ecdh-sha2-nistp256
        ecdh-sha2-nistp384
        ecdh-sha2-nistp521
        diffie-hellman-group-exchange-sha256
        diffie-hellman-group16-sha512
        diffie-hellman-group18-sha512
        diffie-hellman-group14-sha256
```

Based on the above, we can identify a suspicious-looking domain hosted using a specific service in Lithuania, exposing HTTP and SSH as well as HTTP over TCP 8000, and allowing us to fingerprint the server as a Debian Linux machine. Using this information, we could work to identify further infrastructure through both domain and server characteristics.

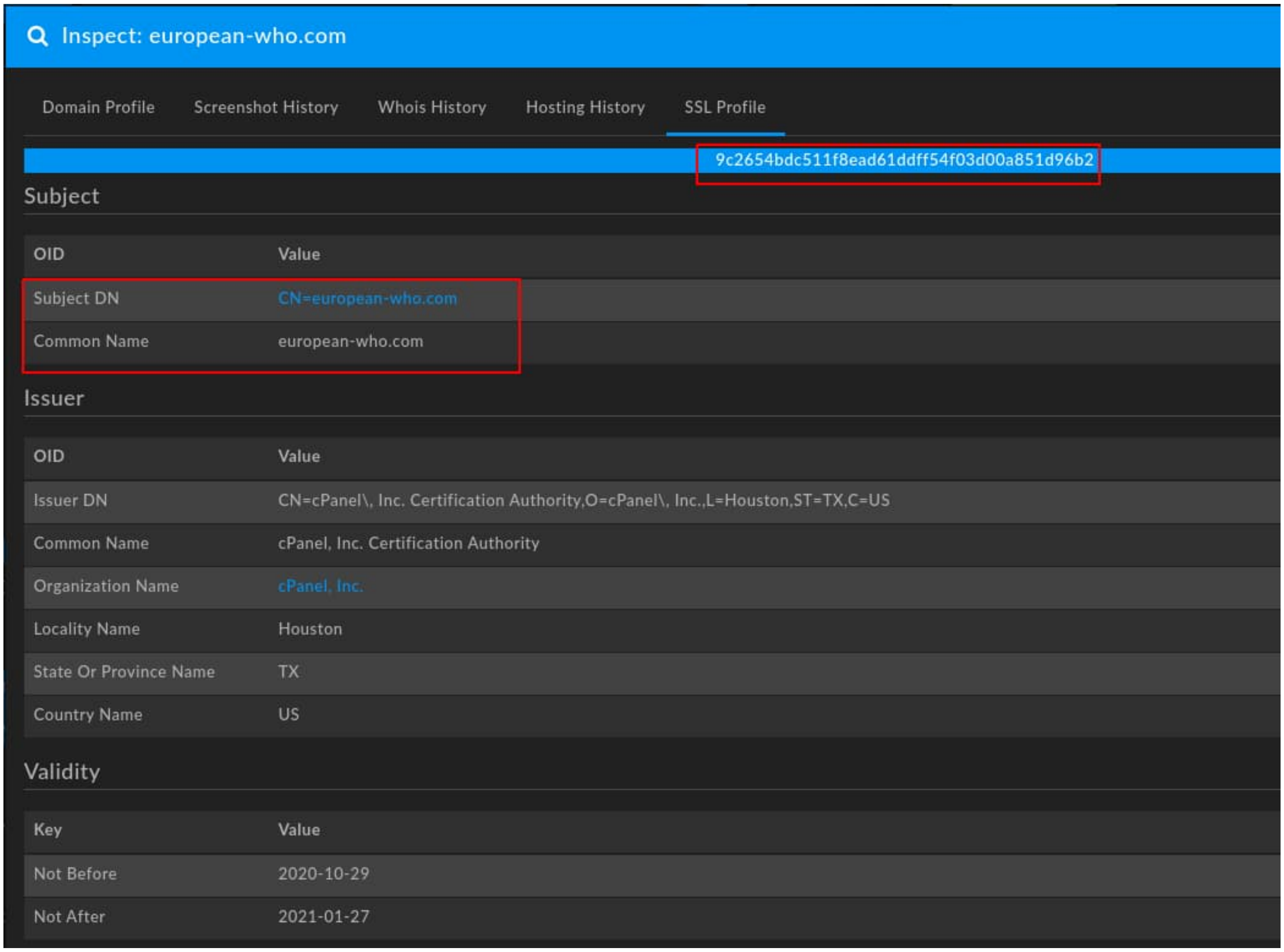
SSL/TLS Certificates

Finally, adversaries (as well as most legitimate web services) frequently employ standard encryption using the Secure Sockets Layer (SSL) or Transport Layer Security (TLS) protocols. While adversaries can certainly use custom encryption or encoding protocols for traffic, the ubiquity of SSL/TLS-wrapped communications and the limited visibility into such communications for most organizations make the publicly-available standard both very effective and significantly cheap for threat actors to deploy.

As a form of public key cryptography, SSL/TLS encryption depends upon certificates for functionality. Certificates can be tracked through various data points, as they typically feature identifying information related to the certificate owner, such as organization or location.

Certificates can take a variety of forms from self-signed, untrusted items to certificates created through free and unvetted services (such as Let’s Encrypt) or sources which perform some degree of vetting when issuing them. While resulting metadata will vary depending on the issuer and the certificate, there is a rich history in using certificate characteristics to track adversary behavior, such as legacy APT28 or Fancy Bear certificate and infrastructure activity.

Looking at the suspicious domain european-who[.]com, we discover a free certificate associated with cPanel services. As such, there is limited data to pivot off of although we can note the limited certificate information and use of a free service to identify this as potentially suspicious, while the certificate hash value provides a way to track the use of this specific item.



9c2654bdc511f8ead61ddff54f03d00a851d96b2

Subject

OID	Value
Subject DN	CN=european-who.com
Common Name	european-who.com

Issuer

OID	Value
Issuer DN	CN=cPanel\, Inc. Certification Authority,O=cPanel\, Inc.,L=Houston,ST=TX,C=US
Common Name	cPanel, Inc. Certification Authority
Organization Name	cPanel, Inc.
Locality Name	Houston
State Or Province Name	TX
Country Name	US

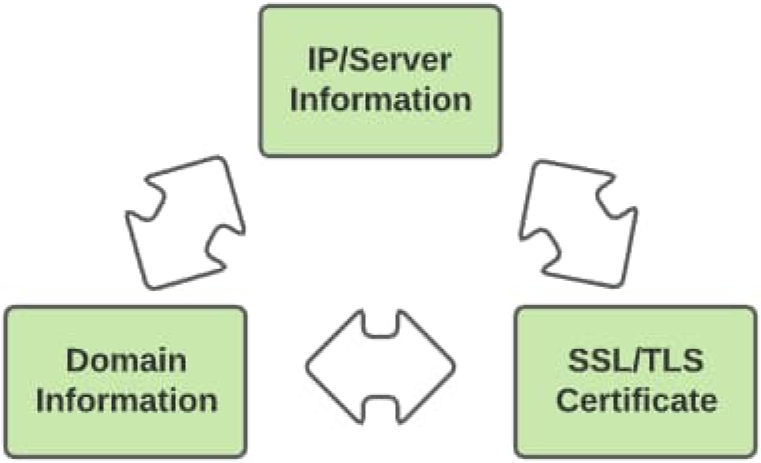
Validity

Key	Value
Not Before	2020-10-29
Not After	2021-01-27

Even though in this case we have limited information for direct pivoting, just identifying the use of self-signed, built-in cPanel certificates can be used to differentiate and identify further infrastructure discovered through domain- or IP-based analysis. Identifying another network object with similar domain or IP characteristics that also deploys similar SSL/TLS certificate creation tendencies can allow us to link infrastructure and begin understanding adversary behaviors.

Composite Details and Pivoting

In addition to pivoting within different types of infrastructure characteristics, as shown in the diagram below, understanding of adversary infrastructure tendencies enables pivoting between items as well.



In this view, insight into domain creation can reveal infrastructure hosting tendencies, which can be leveraged to identify additional domains. Or a persistent pattern in SSL/TLS certificate creation yields additional domains which in turn map to additional infrastructure.

However, while the discovery of new indicators is enticing and potentially useful for defense, this represents an intermediate objective as part of a larger process. Instead of merely attempting to identify more IOCs, cyber threat intelligence (CTI) analysts should leverage this work as a means to identify fundamental adversary tendencies which can be used to continuously identify and disclose infrastructure over time. Indicators will certainly be produced through this work, but they represent an output of a more fundamental process of identifying and observing ground-truth adversary behaviors in creating network infrastructure.

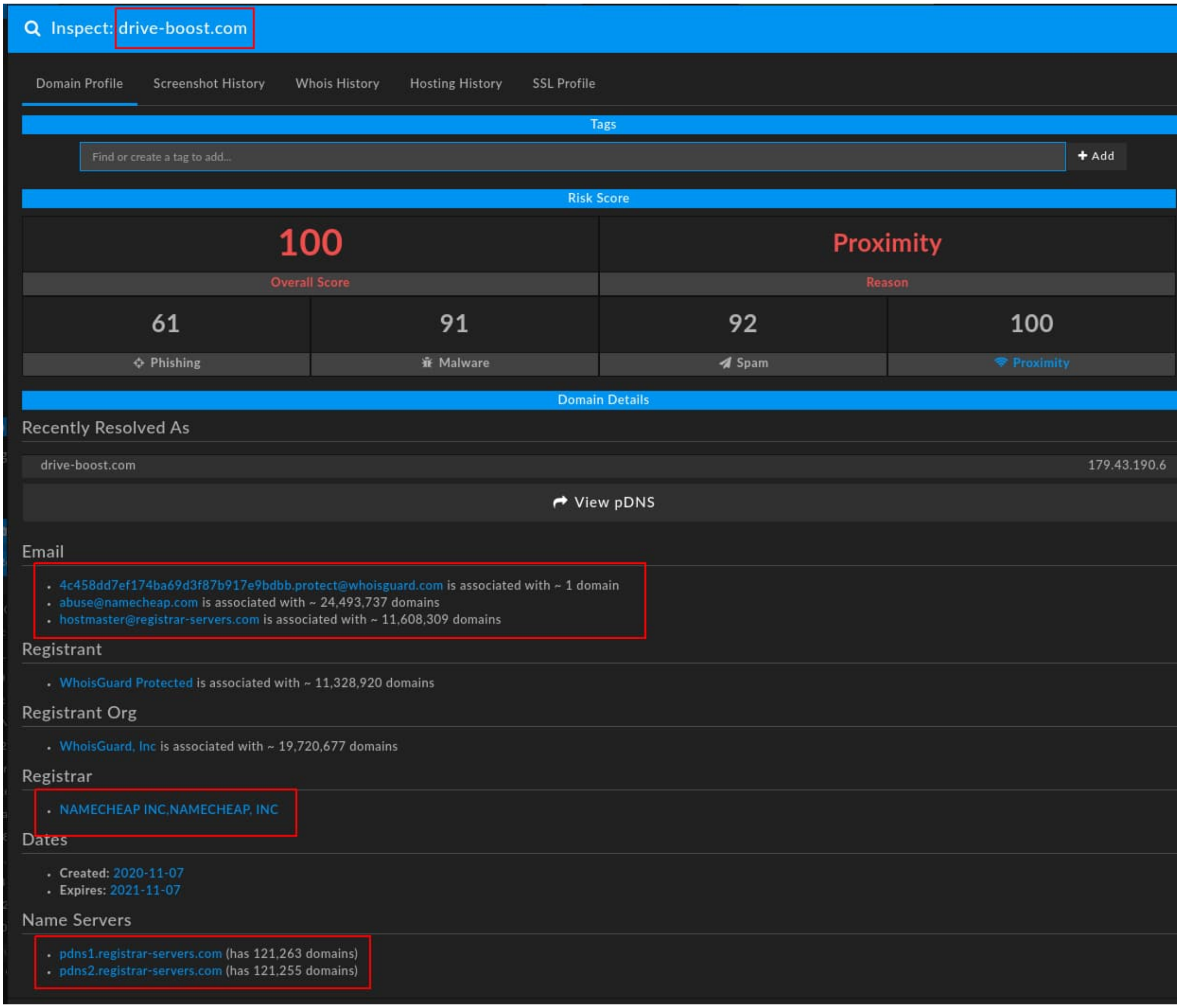
Example: Late 2020 Ryuk Activity

To see an example of the above, we can look at the recent episode of Ryuk ransomware incidents across multiple hospital and health care provider systems in the United States, linked to a group referred to as UNC1878 by information security company FireEye. While utilizing malware and droppers such as BazarLoader and BazarBackdoor, Ryuk deployers still required C2 infrastructure to control and further expand infections in victim environments. Based on information released by FireEye and Kyle Ehmke from ThreatConnect, and confirmed by multiple additional parties, the UNC1878 network infrastructure activity encompassed several hundred domains created from summer 2020 through October 2020. (Please see the linked file for IOCs.)

While seemingly overwhelming, the above domains contained a number of commonalities:

1. Similar registration patterns used over time.
2. Preference for a limited number of hosting providers.
3. Consistency in SSL/TLS certificate characteristics.

For example UNC1878-linked domain “drive-boost[.]com” features the following characteristics:



Inspect: drive-boost.com

100	Proximity
Overall Score	Reason

61	91	92	100
Phishing	Malware	Spam	Proximity

Recently Resolved As

drive-boost.com 179.43.190.6

Email

- 4c458dd7ef174ba69d3f87b917e9bdbb.protect@whoisguard.com is associated with ~ 1 domain
- abuse@namecheap.com is associated with ~ 24,493,737 domains
- hostmaster@registrar-servers.com is associated with ~ 11,608,309 domains

Registrant

- WhoisGuard Protected is associated with ~ 11,328,920 domains

Registrant Org

- WhoisGuard, Inc is associated with ~ 19,720,677 domains

Registrar

- NAMECHEAP INC,NAMECHEAP, INC

Dates

- Created: 2020-11-07
- Expires: 2021-11-07

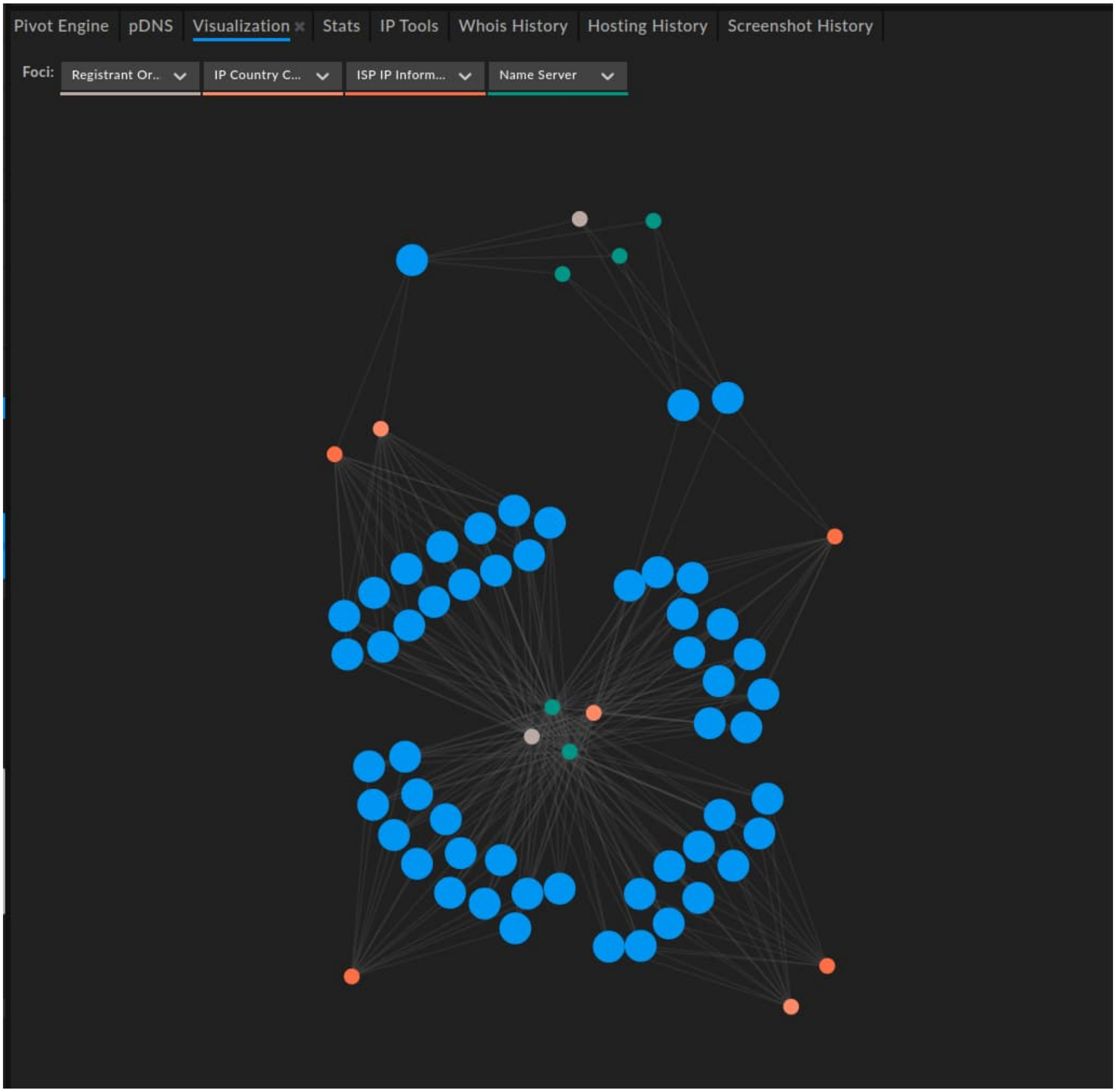
Name Servers

- pdns1.registrar-servers.com (has 121,263 domains)
- pdns2.registrar-servers.com (has 121,255 domains)

Breaking these observables down, we see the following:

- Use of consistent naming “theme” reflecting IT-related concepts or items (e.g., “driver”).

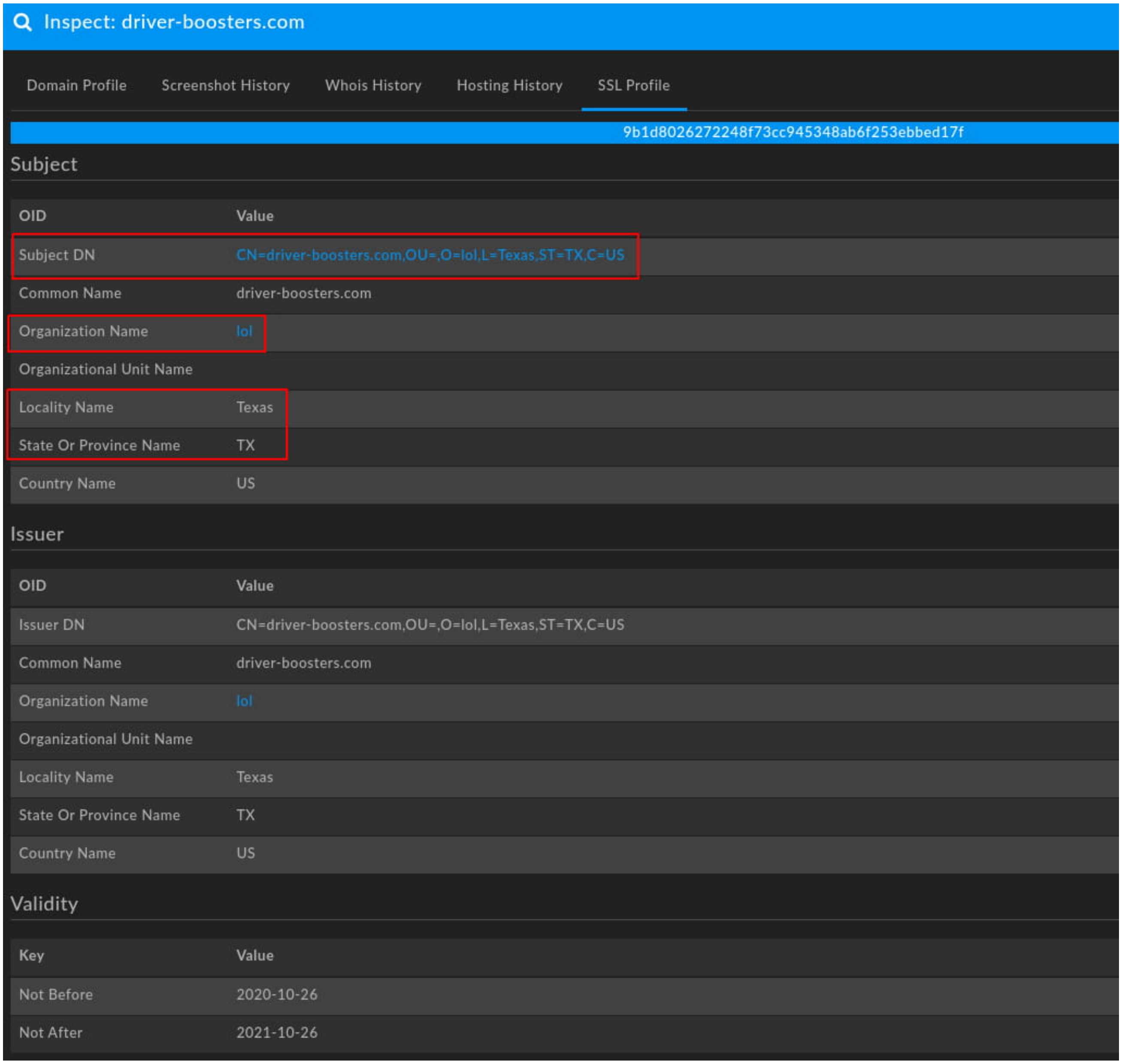
- Use of the Namecheap registrar.
- Consistent use of WhoisGuard privacy protection service.
- Consistent use of registrar-servers[.]com DNS server.



While the above items individually are quite common, taken together they allow us to begin filtering likely related items. Looking at hosting ISP patterns, shown above in a DomainTools Iris Investigate visualization, reveals additional characteristics—namely four distinct "waves" or series of activity centered around the following providers:

- Private Layer Inc., located in Switzerland.
- Psychz Networks, located in the United States.
- Frantech Solutions, located in the United States.
- Combahton GmbH, located in Germany.

Now we have a better way of clustering known activity, and potentially identifying new, similar items. But even further options exist within a subset of domains that feature an existing SSL/TLS certificate. For example, looking at “driver-boosters[.]com” shows the following certificate information:



Inspect: driver-boosters.com

9b1d8026272248f73cc945348ab6f253ebbed17f

Subject

OID	Value
Subject DN	CN=driver-boosters.com,OU=,O=lol,L=Texas,ST=TX,C=US
Common Name	driver-boosters.com
Organization Name	lol
Organizational Unit Name	
Locality Name	Texas
State Or Province Name	TX
Country Name	US

Issuer

OID	Value
Issuer DN	CN=driver-boosters.com,OU=,O=lol,L=Texas,ST=TX,C=US
Common Name	driver-boosters.com
Organization Name	lol
Organizational Unit Name	
Locality Name	Texas
State Or Province Name	TX
Country Name	US

Validity

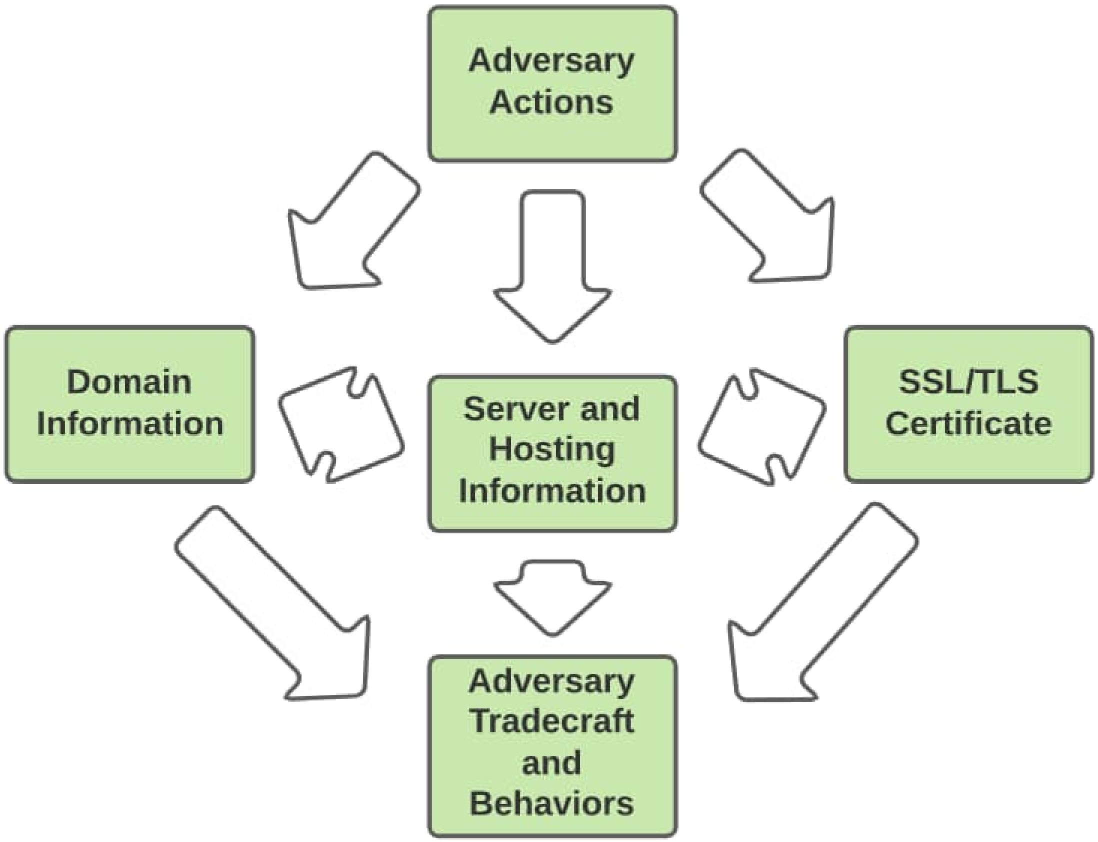
Key	Value
Not Before	2020-10-26
Not After	2021-10-26

In this case, we see a certificate pattern using a Locality Name of “Texas” and an Organization Name of “lol” for a self-signed certificate. Diving into certificate information yields over 100 items which, when looked at in conjunction with other items documented above, allows for high confidence attribution to UNC1878 activity.

Through this process, we have identified a set of characteristics denoting UNC1878 infrastructure. This can be used both for post-incident attribution to determine what entity may be responsible for a breach. Additionally, such methodologies can be used through datasets such as DomainTools and search tools such as Iris Investigate to proactively and preemptively identify infrastructure as it is created. In this fashion, CTI becomes more attuned to adversary tendencies and characteristics while also enabling more direct, proactive support to network defense operations through continuous identification of adversary infrastructure.

Conclusion

While network infrastructure indicators and observables are typically viewed as atomic objects, seeing these items as composites enables powerful analysis able to keep pace with adversary evolution. By understanding the fundamental nature of items such as domain names, IP addresses, and SSL/TLS certificates, analysts can begin understanding fundamental adversary tendencies and tradecraft. When done well, such actions enable defenders to not only accurately disposition new infrastructure as it is discovered, but also to identify new infrastructure as it is created to boost defensive operations.


Adversary Actions
Domain Information
Server and Hosting Information
SSL/TLS Certificate
Adversary Tradecraft and Behaviors

While this process can be quite powerful in identifying and tracking adversary operations, we must also note limitations to this methodology. For example, adversaries may leverage compromised, legitimate infrastructure for communications and similar activity in order to hide their tracks and confuse analysis. This is an increasing trend for many threat actors, which can throw off analytical techniques such as those described above. However, even in these cases, possibilities exist such as identifying underlying server or software commonalities which may be indicative of a vulnerability exploited to gain access to the legitimate server. These commonalities can enable similar types of pivoting to what was discussed previously with respect to adversary-owned-and-operated infrastructure.

By viewing indicators as not just isolated objects but as composites containing multiple components that can be used to better understand the nature, purpose, and composition of the indicator, CTI analysts can unlock a greater understanding of adversary operations. Furthermore, while this article is limited to network observables, the same fundamental concepts are equally applicable to host and file-based indicators as well. By further refining, researching, and enriching indicators, CTI analysts can continuously push the envelope of threat understanding and threat detection, enabling both a better understanding of past events and potential identification of new threat vectors from known adversaries as they emerge.

To learn how to identify and track adversary operations in DomainTools Iris Investigate, visit our product page.

Learn More