Proofpoint ET Splunk TA

Proofpoint Inc

Contents

1	Introduction	2		
2	Requirements	2		
3	Installing the ET Splunk TA			
4	Initial Launch of the ET Splunk TA			
5	ET Splunk TA Macros 5.1 Determining Interesting Fields			
6	Enriching data 6.1 Selecting Predefined Fields	10		
7	Sample Queries 1			
8	Adaptive Response Support 8.1 AR support via Correlation Search 8.2 AR support via Notable Events 8.3 Reviewing the AR Response	14		
9	Reports, Dashboards, Pivots, and Alerts	19		
10	Appendix 10.1 Categories			

1 Introduction

The ET Splunk Technical Add-On (ET-TA) allows ET customers with Splunk implementations to greatly enhance their ability to enrich and search any log with ET Intelligence data. The ET-TA provides three primary functions:

- Automatically Downloads, Installs, and Updates the ET Intelligence reputation list into Splunk.
- Provides several Splunk Macros which allow organizations to build their own complex queries using not just ET, but virtually any data, including with other Splunk features and TAs.
- Enables integration with the Splunk Adaptive Response framework and dynamically uses the ET Intelligence API to enrich Notable Events with rich context from Proofpoint ET Intelligence.

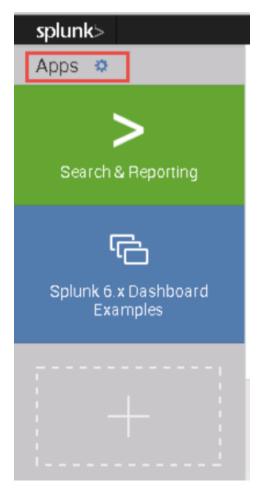
2 Requirements

The ET-TA is a very lightweight and flexible Technical Add On. It can function on any Splunk license, including the Free license. Normal Splunk License limitations apply, e.g. if you only have a 1 GB / day license, you can't log more than 1 GB/day, but that is completely independent of the ET-TA.

3 Installing the ET Splunk TA

The Splunk TA can be installed in under a minute from the Splunk UI. You can easily install the application from the SplunkBase. Please follow the procedure below:

- 1. Log into your Splunk instance at https://<SplunkIP>:8000
- 2. Click the (*) Apps button



- 3. Click "Browse for Apps"
- 4. Enter "Proofpoint" into the browser bar and you should see Proofpoint ET Splunk TA.
- 5. If this is the first time you have installed the ET-TA, then you will be given the "Install" button. If you have already installed the ET-TA then you will be given the "Update" button. Click the Install/Update button to install the current TA version.
- 6. After a moment, Splunk will ask you to restart Splunk, select Restart Now.
- 7. After Splunk restarts, you will be forced to log back into Splunk
- 8. Once again, click the (*) Managed Apps button
- 9. You should now see the ET Splunk TA in the table.
- 10. Click "Launch App" in in the row for ET Splunk TA

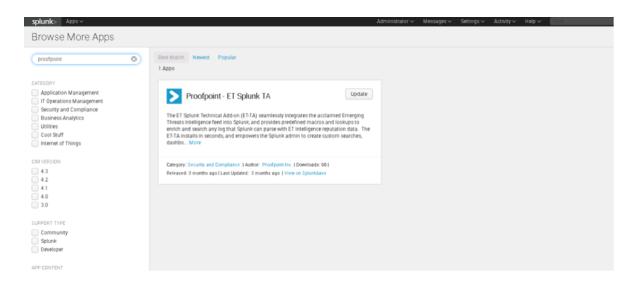


Figure 1: Downloading the ET TA

4 Initial Launch of the ET Splunk TA

Once the Splunk TA is installed and Splunk is started, the next step will be to launch the ET-TA. You can do this simply by clicking on the "ET Splunk TA" tile from the Apps list, or you can go under your "Managed Apps" menu and select "Launch App"

Method 1:

or Method 2:

Once the TA is launched, the first time it will prompt you for your Authorization Code and API key, which are codes used for the ET Intelligence Reputation List and ET Intelligence API access respectively. The Authorization code is a 16-digit number and the API key is 64 alphanumeric characters long. Both the code and the key are available to you from the admin portal.

The new version of ET TA expects both Authorization code and API key.

If you have only one of the Authorization Code or API key, please enter it in the box, however both are required for the Adaptive Response capabilities to work. After you enter your Authorization Code and API key, it will take about a minute to download and install the ET Intelligence list into your Splunk instance. The TA willautomatically check for updates to the reputation list every hour and install a new version if available.

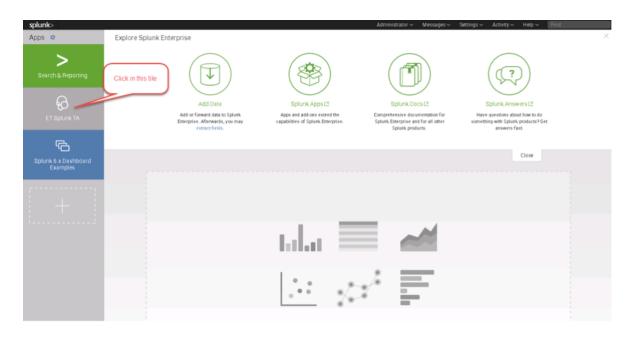


Figure 2: Launch TA from Home Screen

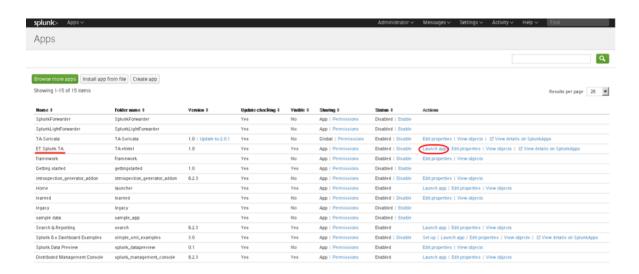


Figure 3: Launch TA from Apps Menu

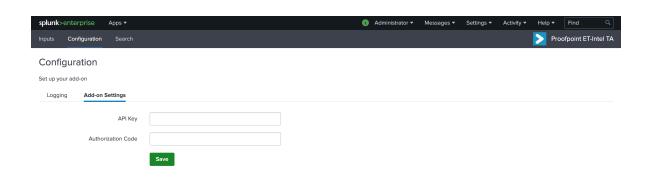


Figure 4: Enter Authorization Code

5 ET Splunk TA Macros

Once the ET-TA is installed, you can immediately begin to leverage the power of the ET-TA. The macros provided will allow you to enrich your logs with ET data at search time, which improves performance and is not reliant on when the logs are received. Additionally, the macros allow you to specify which fields to search for matches, so effectively any field in any log that Splunk can parse can be used to create queries.

5.1 Determining Interesting Fields

Before delving into the TA Macros it is useful to understand your data and how the TA uses it. In each Macro you will provide it with the interesting field to search for the intersection of the ET data with that match of your field (be it IP or DNS entry), and then the TA will enrich your data with the ET Intelligence Reputation information. You can determine any field once you load your logs into Splunk and expand any log:

A few things to note. In the above example there are both Source and Destination address fields, and the ET-TA can match on any of them so long as you define which one to select. One interesting thing to note is that there are some characters which Splunk will replace in the Field name, but display the original. In Figure 5, the dash "-" in any field is replaced by an underscore "_". Look at the log for dest-ip, while the field name is dest ip.

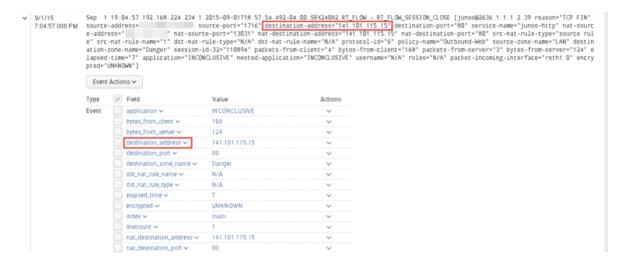


Figure 5: Determining what field to use

There are two types of Macros provided by the ET-TA:

IP Lookup Macro: et_ip_lookup(IP=<IPfieldname>)

This macro takes a single argument which is the IP field name and uses it to search against the ET Intelligence reputation list. If a match is found, that log will be enriched with the ET data for that entry. Typically this will be a field from a Firewall, IPS, Proxy or other log that contains an IPv4 Address. For instance, if your firewall has a field called srcip=192.168.1.1 for Source Address, the macro would be et_ip_lookup(IP=srcip). Again this is only for the field name.

DNS Lookup Macro: et_domain_lookup(DOMAIN=<DNSFieldName>)

The DNS macro takes a single argument which is a field in a log containing a DNS FQDN and searches against the ET Intelligence reputation list to see if there is a match. If there is a match found, the log will be enriched with the ET data for that entry. For instance, if you have a log that has a DNS request field dns-request=time.nist.gov then the macro would be et_domain_lookup(DOMAIN=dns_request).

5.2 Extended Versions

There are two additional types of Macros provided by the ET-TA:

Extended DNS Lookup Macro: et extended domain lookup (DOMAIN=<DNSFieldName>)

The DNS macro takes a single argument which is a field in a log containing a DNS FQDN and searches against the ET Intelligence extended reputation list to see if there is a match. If there is a match found, the log will be enriched with the ET data for that entry.

For instance, if you have a log that has a DNS request field "dns-request=time.nist.gov" then the macro would be et_extended_domain_lookup(DOMAIN=dns-request).

Extended IP Lookup Macro: et_extended_ip_lookup (IP=<IPfieldname>)

This macro takes a single argument which is the IP field name and uses it to search against the extended ET Intelligence reputation list. If a match is found, that log will be enriched with the ET data for that entry. Typically this will be a field from a Firewall, IPS, Proxy or other log that contains an IPv4 Address.

For instance, if your firewall has a field called "srcip=192.168.1.1" for Source Address, the macro would be et_extended_ip_lookup(IP=srcip). Again this is only for the field name.

6 Enriching data

With the TA installed and an understanding of the Macro syntax, it's time for us to start using it live. Typically you would follow the following format for running the macros:

```
select_data | `et_macro()` | additional_filtering | optional_queries_or_macros
```

Where <code><select_data></code> is an optional Splunk query string, but is used to define what data you would like to pass to the ET macro, since you typically want to narrow down your selection in some way (such as by log source or matching some logs ahead of passing it to the macro.) Next we pipe the logs to the Macro.

The Macro is simply finding matches of the IP or Domain field which you pass to it vs. the ET data set, and if there is a match on that log we will enrich it with the additional information we know about that object.

After the macro runs, you may define additional match critiera. Most often, this would be some sort of filter based upon the enriched data. The information that is outputted from that point is then passed to any additional queries or macros that might run and ultimately to the Splunk Search window.

Note: While the <select data> field is optional it is highly recommended for two reasons. First, it allows you to ensure that only logs of a certain datatype are sent to the ET-TA macro. This is important because the TA cannot enrich logs which don't have a matching field. No error will occur, but they won't be enriched. Second, the search time in Splunk is proportional to the number of logs that are passed to it, so by filtering out unnecessary logs, we can improve the search time performance.

6.1 Selecting Predefined Fields

The ET-TA enriches each entry with the several fields. By default these fields will be enriched in the logs, but will not display, so you will need to select which fields you want to display in the UI if you want them to appear. Also please see the Appendix for the list of categories.

IP Address Objects:

- Category: This is the category that the ET Research Team has determined the IP has exhibited.
- Score: This is a score from 0-127 (worst rep) which is the same as what is used in Suricata. The score is a magnitude, but also decays back to 0 if additional events do not occur.
- First Seen: This is the date that the object was first seen as creating interesting activity in the global ET sensornet for that given category.
- Last Seen: This is the date that the object was last seen to be exhibiting interesting activity for that given category.
- Ports: This field is the list of any TCP/UDP ports that we saw the activity on.
- Threat Level: This is defined per category. See appendix.

DNS Objects

- Category: This is the category that the ET Research Team has determined the domain has exhibited.
- Score: This is a score from 0-127 (worst rep) which is the same as what is used in Suricata. The score is a magnitude, but also decays back to 0 if additional events do not occur.
- First Seen: This is the date that the object was first seen as creating interesting activity in the global ET sensornet for that given category.
- Last Seen: This is the date that the object was last seen to be exhibiting interesting activity for that given category.
- Ports: This field is the list of any TCP/UDP ports that we saw the activity on.
- Threat Level: This is defined per category. See appendix.

Extended IP Address Objects

In addition to the IP address objects listed above the extended IP reputation list also contains the following fields:

- ASN: Autonomous System Number globally unique identifier
 - Owner
 - Authorizer
 - Country of IP registration
 - Registration date the date that current
 - Reverse lookup
 - CIDR (Classless Inter Domain Routing)
- Location
 - IP
 - Country Code
 - Country
 - City
 - Latitude
 - Longitude
- Signature data: Which signatures have fired to provide reputation
- Mitre Framework: Knowledge-based framework of adversary tactics and techniques
 - Mitre tactic id: Common identifiers for tactics

- Mitre Tactic name: What the attackers are trying to accomplish like "Defense Evasion"
- Mitre Technique ID: Common identifiers for techniques
- Mitre Technique Name: How they accomplish these steps or goals like "Process Injection"
- Malware family: If there is malware associated with this IP or domain

Extended Domain Objects

In addition to the DOMAIN name objects listed above the extended DOMAIN reputation list also contains the following fields:

- Signature data: Which signatures have fired to provide reputation
- Mitre Framework: Knowledge-based framework of adversary tactics and techniques
 - Mitre tactic id: Common identifiers for tactics
 - Mitre Tactic name: What the attackers are trying to accomplish like "Defense Evasion"
 - Mitre Technique ID: Common identifiers for techniques
 - Mitre Technique Name: How they accomplish these steps or goals like "Process Injection"
- Malware family: If there is malware associated with this IP or domain
- Whois: Information about the registrant of a domain name, the domain registrar, and the domain's status
- Nameservers: Refers to the servers that are responsible for resolving a domain name

6.2 Selecting Interesting Fields to Display

As mentioned, by default the TA won't display the additional enriched fields, so you will want to select them from the drill down if you would like them to be displayed. This has no impact on the actual search, it is purely for visually identifying logs.

6.3 Output Types

Because the TA is empowering you to build your own queries, it can be used to power any integrated Splunk feature such as Reports, Dashboards, Panels, and Alerts. You can also use them to power your own apps.

7 Sample Queries

In this section we will explore a few examples of using the Splunk app. In our example we will be matching logs from a Juniper SRX with our ET data-set.

1. **Simple Query:** Finding all logs where the destination address is known to be a CnC server by ET. In this example we selected what data we wanted to pass to the macro. In this case it was just traffic from a log source defined as host=<log source address>. We then call the et_ip_lookup macro. In this example the log that we were interested on contained an IP address field "dest-ip", but remember that Splunk represents it as "dest_ip". Finally we do some additional filtering on the output to match only logs that are Category CNC. Note that in the

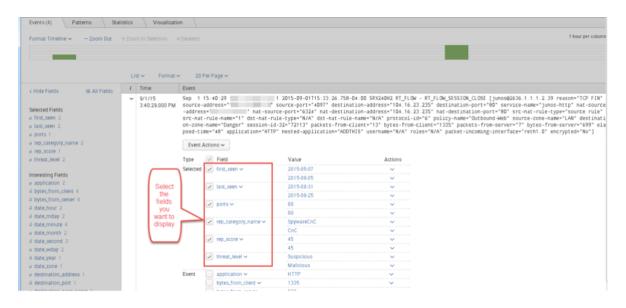


Figure 6: Selecting ET Fields to Display

output you may have multiple matches for categories on a single object. For instance, if an IP address is both SpywareCNC and CNC, as shown below for the object 104.16.23.235 then you would see multiple output fields if you have chosen to display those fields.

host = 192.168.1.1 | `et_ip_lookup(IP=dest_ip)` | search rep_category_name = CNC

2. Advanced Query: In this query we will use the DNS lookup to match DNS requests to malicious domains which may be an indicator of compromise. We will look for objects that not only match the category CNC, but that have a rep_score >= 50. Our log source this time will be Suricata DNS logs. These logs were sent via syslog and were not structured, so we used Splunk to extract our own field which we call DNS_Request which matches the FQDN in any DNS Query.

host = 192.168.1.1 | `et_domain_lookup(domain=DNS_Request)` | search rep_score >= 50 AND rep_category

we start by selecting the log source as DNS, and then use the domain lookup macro (which is no more advanced than the IP lookup macro), however after we find and enrich the data, we then match on multiple criteria that is found in the logs, in this case rep_score >=50 and the rep_category_name = CNC. This brings up a malicious domain cdjgfgphdhvt.com which was found in our DNS logs.

8 Adaptive Response Support

Proofpoint ET Intelligence support for Splunk Adaptive Response further enhances the capabilities of the ET Technical Add-On by not only identifying interesting activity, but also allowing to take action

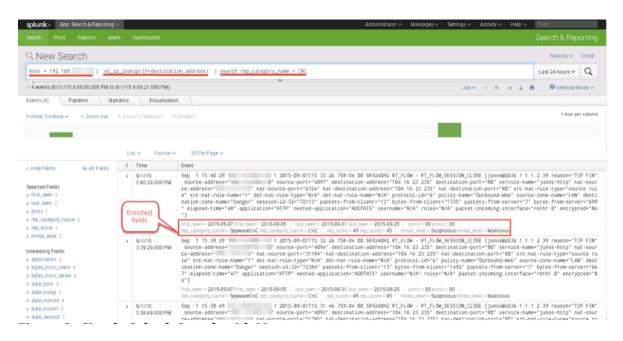


Figure 7: Simple Splunk Search with Macros

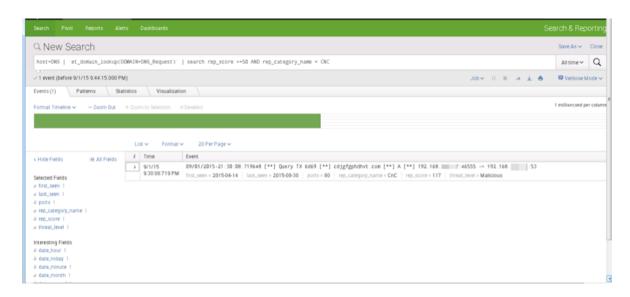


Figure 8: Advanced Splunk Search with Macros

on it. Splunk Adaptive Response actions typically allow the user to gather information or take other action in response to the results of a correlation search or the details of a notable event. Proofpoint ET Intelligence Adaptive Response support falls in the information gathering category where the Technical Add-On acts to get rich Threat Intelligence data on actions specified by the user.

8.1 AR support via Correlation Search

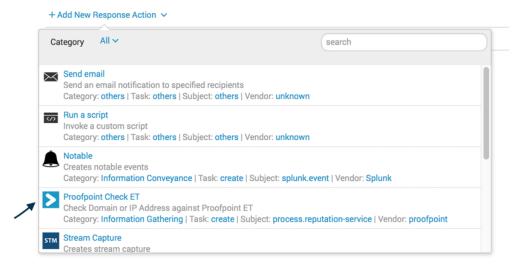
This section describes how to leverage Proofpoint ET Adaptive Response actions via Splunk Correlation Search. Below are the typical steps to create a Correlation Search:

Part 1: Plan the use case for the correlation search. Part 2: Create a new correlation search. Part 3: Create the correlation search in guided mode. Part 4: Schedule the correlation search. Part 5: Choose available adaptive response actions for the correlation search.

More details on creation of Correlation Searches are in the Splunk documentation

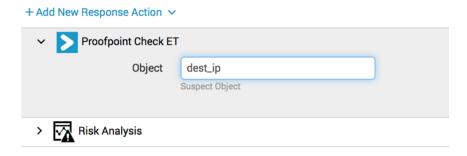
The screenshot below shows an example of how to add the Proofpoint Check ET Adaptive Response action to a correlation search.

Adaptive Response Actions



Choosing 'Proofpoint Check ET' on the above screen requests for an object associated with the action.

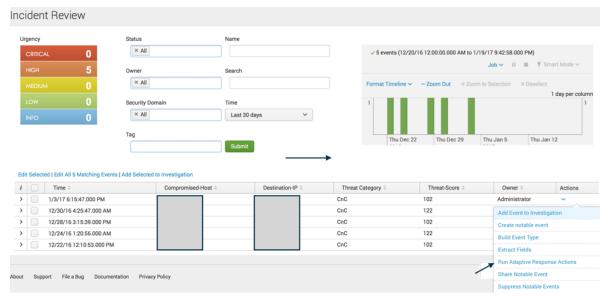
Adaptive Response Actions



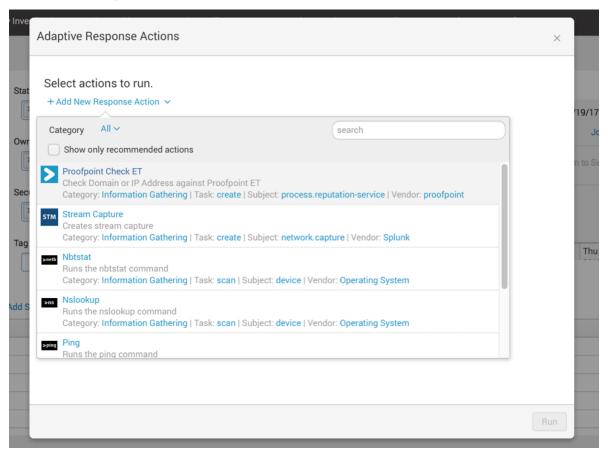
Because Proofpoint ET's AR action is to collect Threat Intelligence tied to the object, the recommended objects are source or destination IP Addresses. In this e.g. the field name is dest_ip.

8.2 AR support via Notable Events

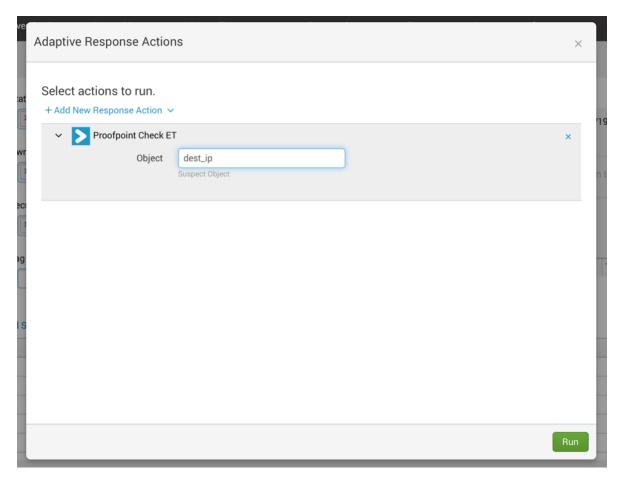
Adaptive Response actions can also be taken manually in the Incident Review tab on Splunk ES. The Incident Review tab displays the Notable Events that are generated in response to Correlation Searches. The screenshot below shows how to add an Adaptive Response action manually for a chosen notable event.



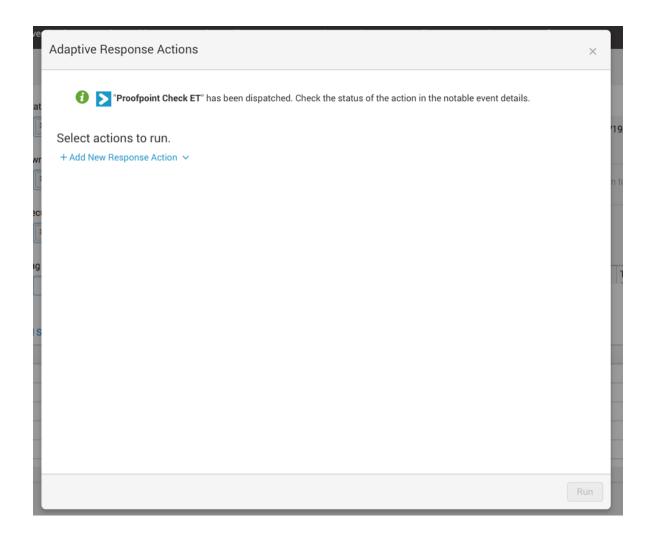
The 'Run Adaptive Response Actions' on the above screen opens a new dialog box that enables the selection of 'Proofpoint Check ET' AR action as shown below.



The Proofpoint Check ET AR further requires an object as described earlier. In the example below, we again associate the AR action with the destination IP Address. As mentioned earlier, the recommended objects for associating Proofpoint Check ET actions are source or destination IP Addresses.

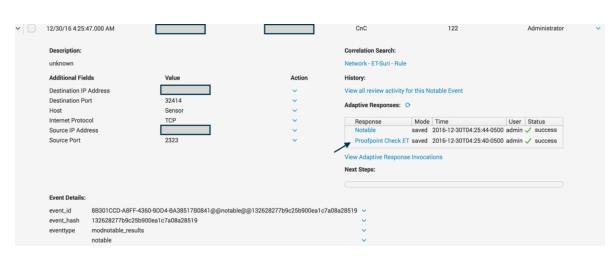


Once this is done, Splunk will notify that the action was successfully added.

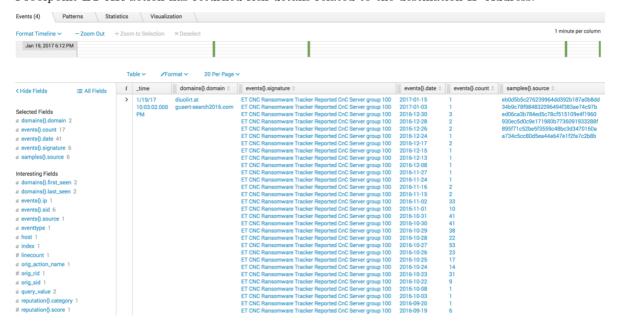


8.3 Reviewing the AR Response

The result of the AR action (triggered either via a correlation search, or manually) is displayed under the details of each notable event.



Clicking on 'Proofpoint Check ET' would display more details tied to this action. In this example, Proofpoint ET AR action has returned rich details related to the destination IP Address.



Note that the output shows two additional domains, six malware sample hashes and the associated IDS events as observed by Proofpoint ET's sensor network.

9 Reports, Dashboards, Pivots, and Alerts

The ET-TA can be leveraged to power any Splunk output mechanism such as Reports, Dashboards, Pivots, and Alerts. These output mechanisms are built into Splunk, and not explicitly part of the ET-TA. However, the ET-TA is essentially a search and enrichment tool it can plug into any query, which can be turned into a Splunk output. Here are a few of the possibilities that you can leverage as part of the ET-TA and Splunk. Start by entering an ET-TA query into the Splunk Search bar:

* | `et_ip_lookup(IP=src_ip)`| search threat_level=Malicious

After the query completed, to create a report select "Save As"

- Report: To save this query as a report which can easily be recalled through the reports menu.
- Dashboard Panel: This is essentially a report or pivot which can be displayed on your home Splunk dashboard screen.
- Alert: Events matching this query can be used to power alerts to generate Splunk alerts when events and thresholds occur.

To generate Pivots, after the query completes select "Visualization" tab, and click the "Pivots" button to generate a new Pivot. You will then be taken to the Splunk Pivot wizard, which allows you to select what format, data filters, and display criteria will be used to generate the Pivot. The pivot will be similar to a graphical report, but goes a step further in allowing the user to drill down or "pivot" into the interesting data by selecting the visual element that the user would like to drill down into.

10 Appendix

10.1 Categories

Cat	Rep Category Name	Rep Category Description
1	CnC	Malware Command and Control Server
2	Bot	Known Infected Bot
3	Spam	Known Spam Source
4	Drop	Drop site for logs or stolen credentials
5	SpywareCnC	Spyware Reporting Server
6	OnlineGaming	Questionable Gaming Site
7	DriveBySrc	Driveby Source
9	ChatServer	POLICY Chat Server
10	TorNode	POLICY Tor Node
13	Compromised	Known compromised or Hostile
15	P2P	P2P Node
16	Proxy	Proxy Host
17	IPCheck	IP Check Services
19	Utility	Known Good Public Utility

Cat	Rep Category Name	Rep Category Description
20	DDoSTarget	Target of a DDoS
21	Scanner	Host Performing Scanning
23	Brute_Forcer	SSH or other brute forcer
24	FakeAV	Fake AV and AS Products
25	DynDNS	Domain or IP Related to a Dynamic DNS Entry or Request
26	Undesirable	Undesirable but not illegal
27	AbusedTLD	Abused or free TLD Related
28	SelfSignedSSL	Self Signed SSL or other suspicious encryption
29	Blackhole	Blackhole or Sinkhole systems
30	RemoteAccessService	GoToMyPC and similar remote access services
31	P2PCnC	Distributed CnC Nodes
33	Parking	Domain or SEO Parked
34	VPN	VPN Server
35	EXE_Source	Observed serving executables
37	$Mobile_CnC$	Known CnC for Mobile specific Family
38	Mobile_Spyware_CnC	Spyware CnC specific to mobile devices
39	Skype_SuperNode	Observed Skype Bootstrap or Supernode
40	Bitcoin_Related	Bitcoin Mining and related
41	DDoSAttacker	DDoS Source

10.2 Category to Threat Level Mapping

Each category defined in the Categories appendix has an associated Threat Level Mapping. The threat levels are provided by Emerging Threats and are understood by Suricata and Snort. You can map the index of the category to the associated threat level below.

Category Index	Threat Level
0	Unknown
1	Malicious
2	Malicious
3	Malicious
4	Malicious
5	Suspicious
6	Suspicious
7	Malicious
8	Other
9	Suspicious
10	Suspicious
11	Other
12	Other

Category Index	Threat Level
13	Malicious
14	Other
15	Suspicious
16	Suspicious
17	Suspicious
18	Other
19	Good
20	Suspicious
21	Malicious
22	Malicious
23	Malicious
24	Malicious
25	Other
26	Suspicious
27	Suspicious
28	Suspicious
29	Malicious
30	Suspicious
31	Malicious
32	Other
33	Suspicious
34	Suspicious
35	Suspicious
36	Other
37	Malicious
38	Suspicious
39	Suspicious
40	Suspicious
41	Malicious

proof point.

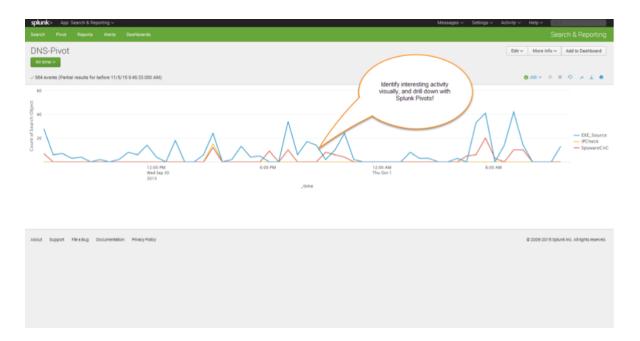


Figure 9: Building Splunk Pivots with ET Intelligence