proofpoint

ET Splunk Technical Add-On Quick Start Guide

Table of Contents

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 two primary functions:

1. Automatically Downloads, Installs, and Updates the ET Intelligence reputation list into Splunk.
2. 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 TA's.

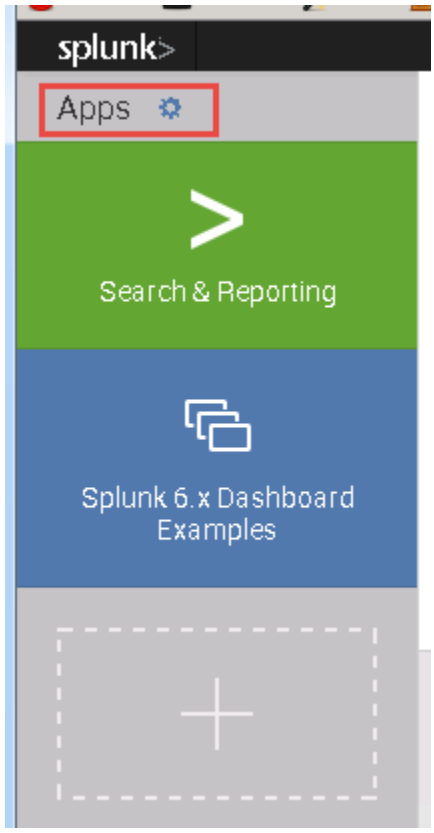
Requirements

The ET-TA is a very lightweight and flexible Technical Add On. It supports Splunk 6.0, and has no requirements for other Apps or TA's to be installed. It can function on any Splunk license, including the Free license. Normal Splunk License limitations apply, e.g. if you only have a 1GB / day license, you can't log more than 1GB/day, but that is completely independent of the ET-TA.

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.

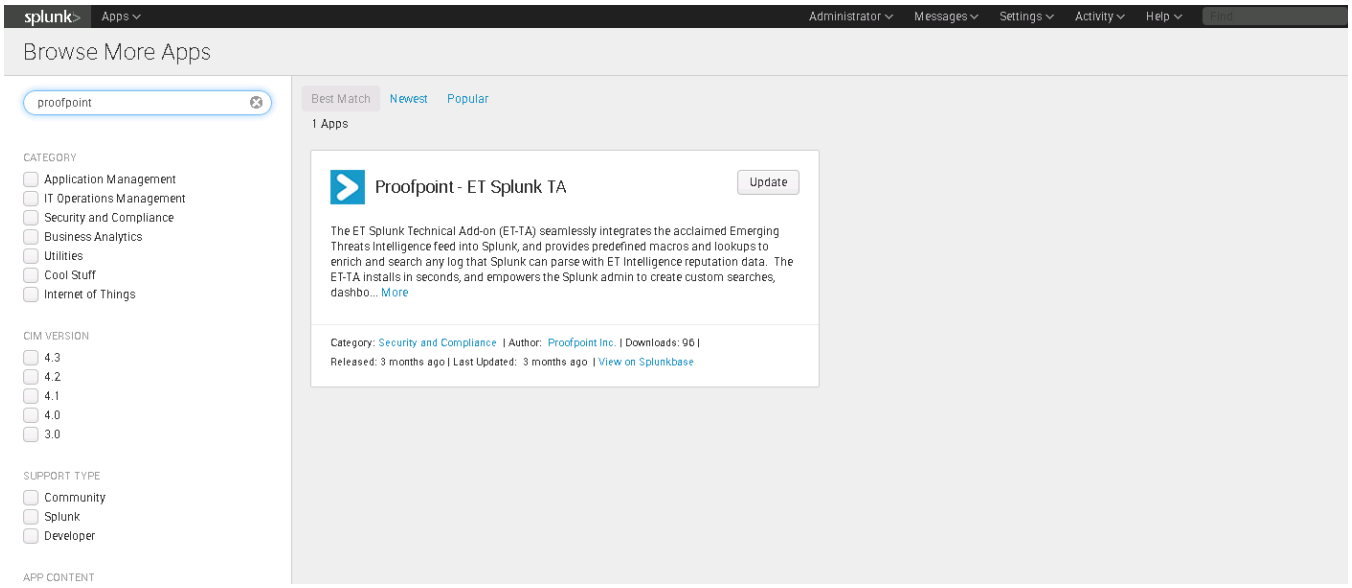


Figure 1: Downloading the Proofpoint Splunk TA

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 k 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.

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:

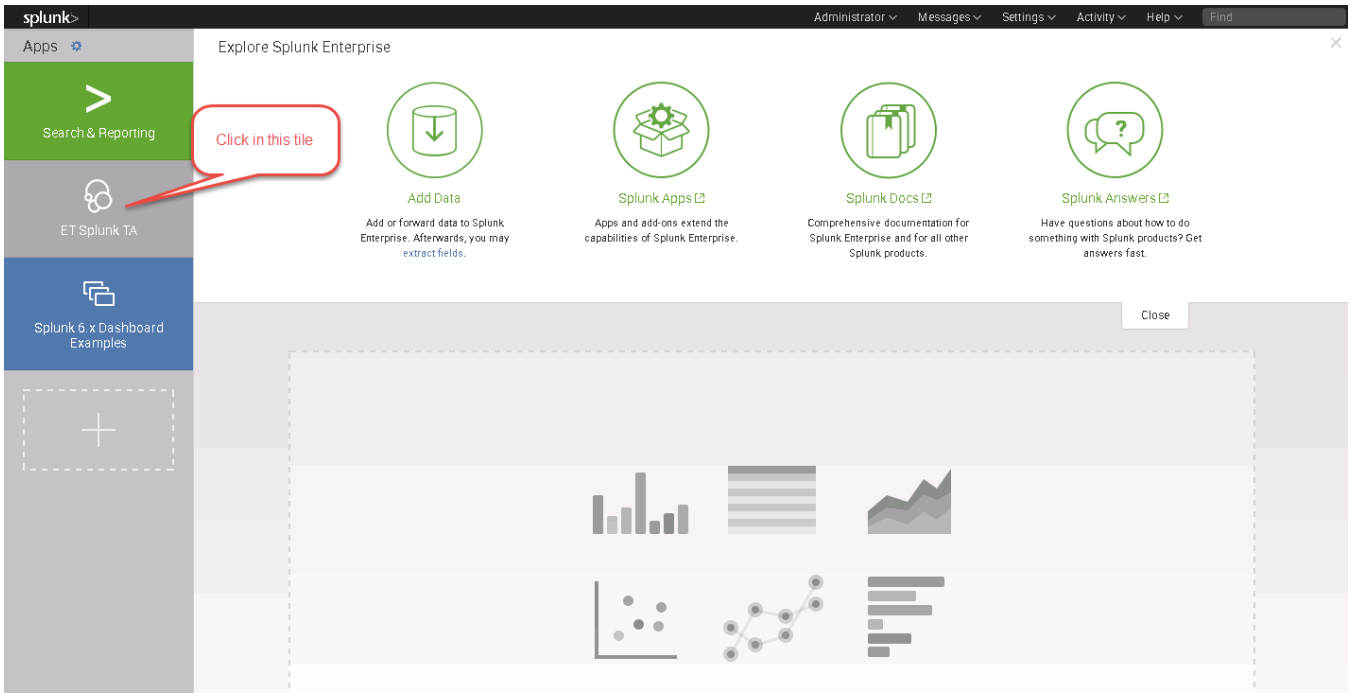


Figure 2. Launch TA from Home Screen

Or

Method 2:

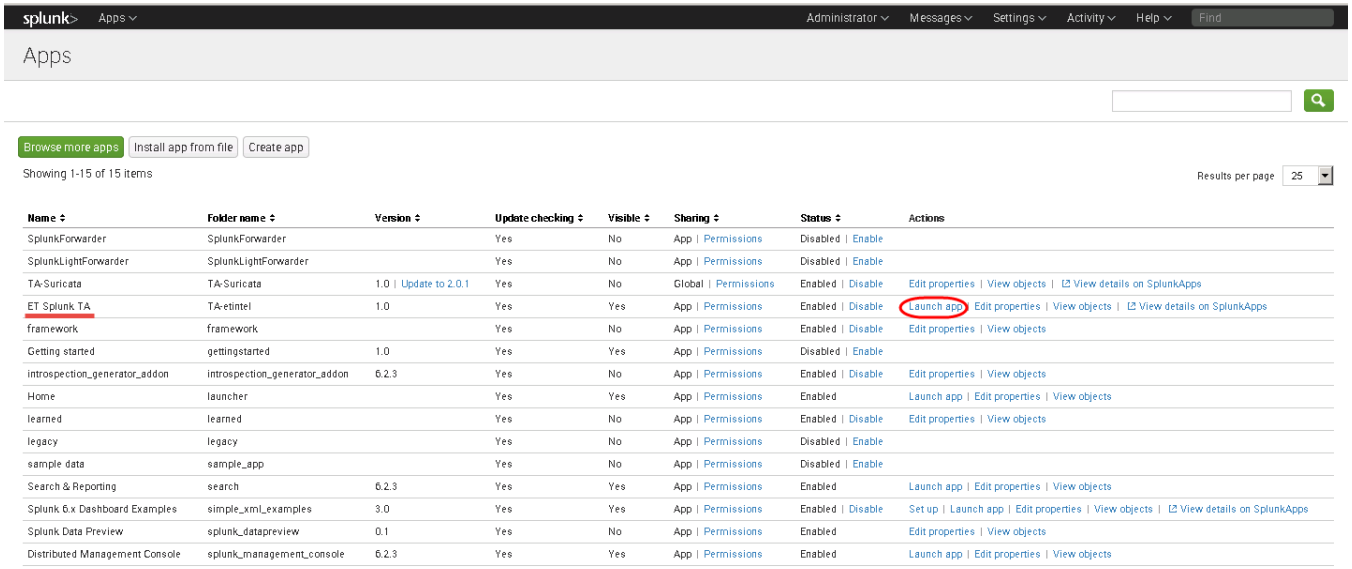


Figure 3: Launch TA from Apps Menu

Once the TA is launched, the first time it will prompt you for your Authorization Code which is provided to you by ET so that you can download the Reputation List. You key is located in https://portal.emergingthreats.net and under the IQRisk List menu.

If you do not have an Authorization Code you can follow the link to the Trial Key site to get a trial key.

***Note: The Authorization Code is NOT the same thing as the ETPro OINK Code. An ETPro OINK code will not allow you to download and install the ET Intelligence reputation data base in Splunk.**

After you enter your Authorization Code, it will take about 1 minute to download and install the ET Intelligence list into your Splunk instance. The TA will automatically go out every hour and check for updates to the reputation list and install them if available.

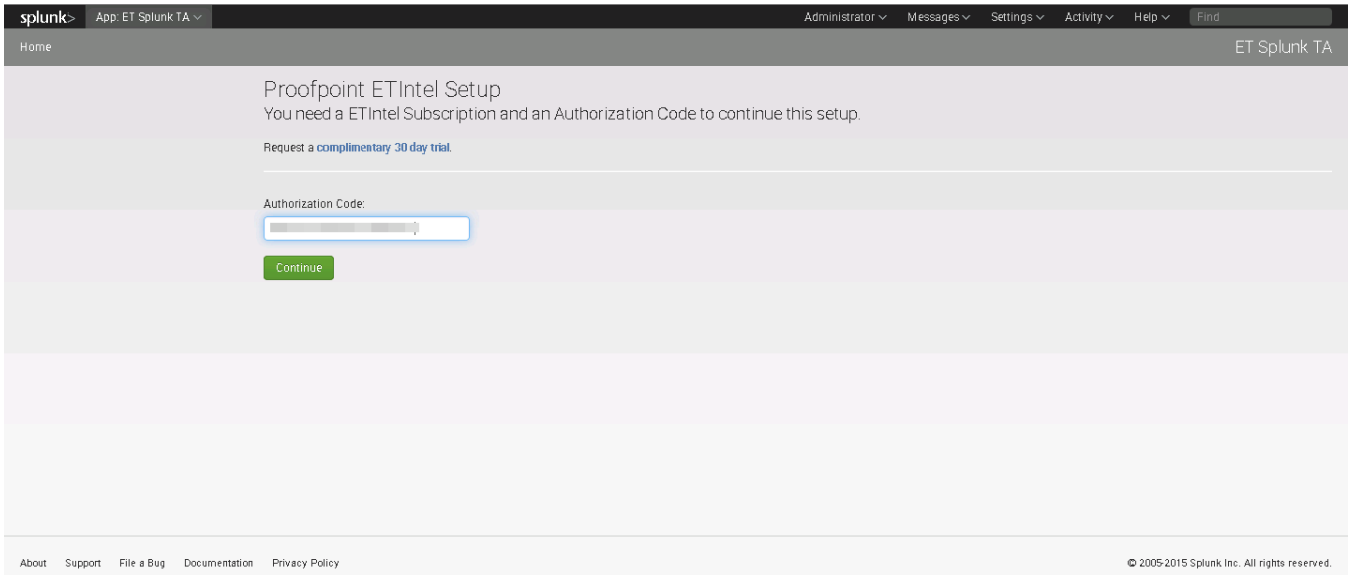


Figure 4: Enter Authorization Code

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.

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:

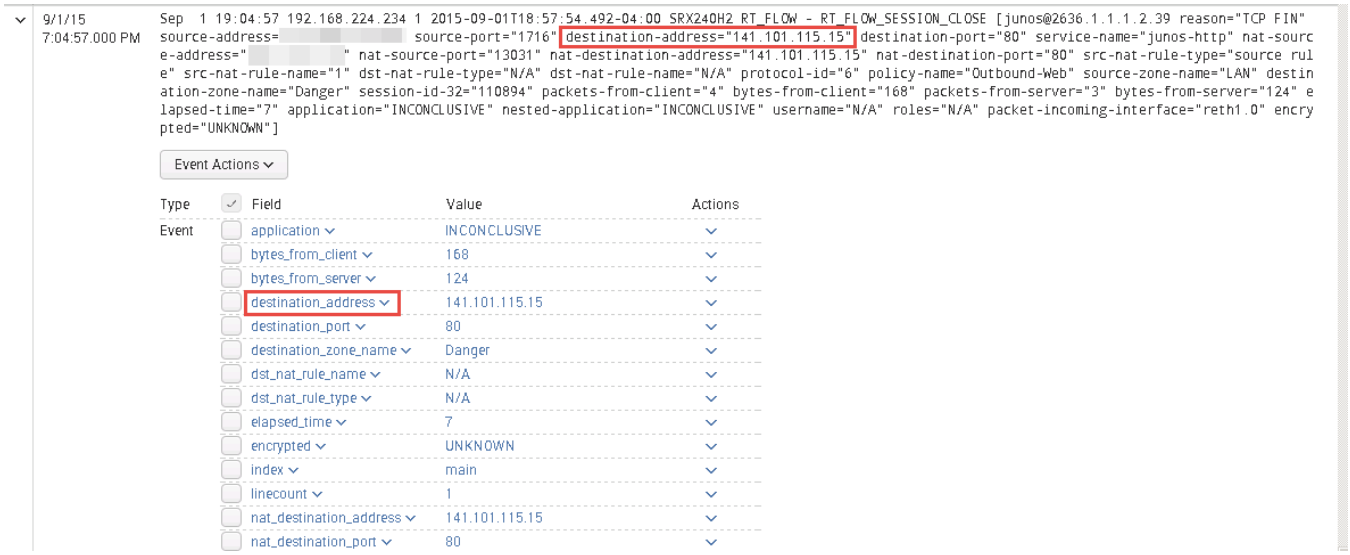


Figure 5: Determining what field to use.

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 "destination-address", while the field name is "destination_address".

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

IP Lookup Macro: et_ip_lookup(IP=<IP field name>)
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=dns.rrname)
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).

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 <select_data> 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.

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.

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 form 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.

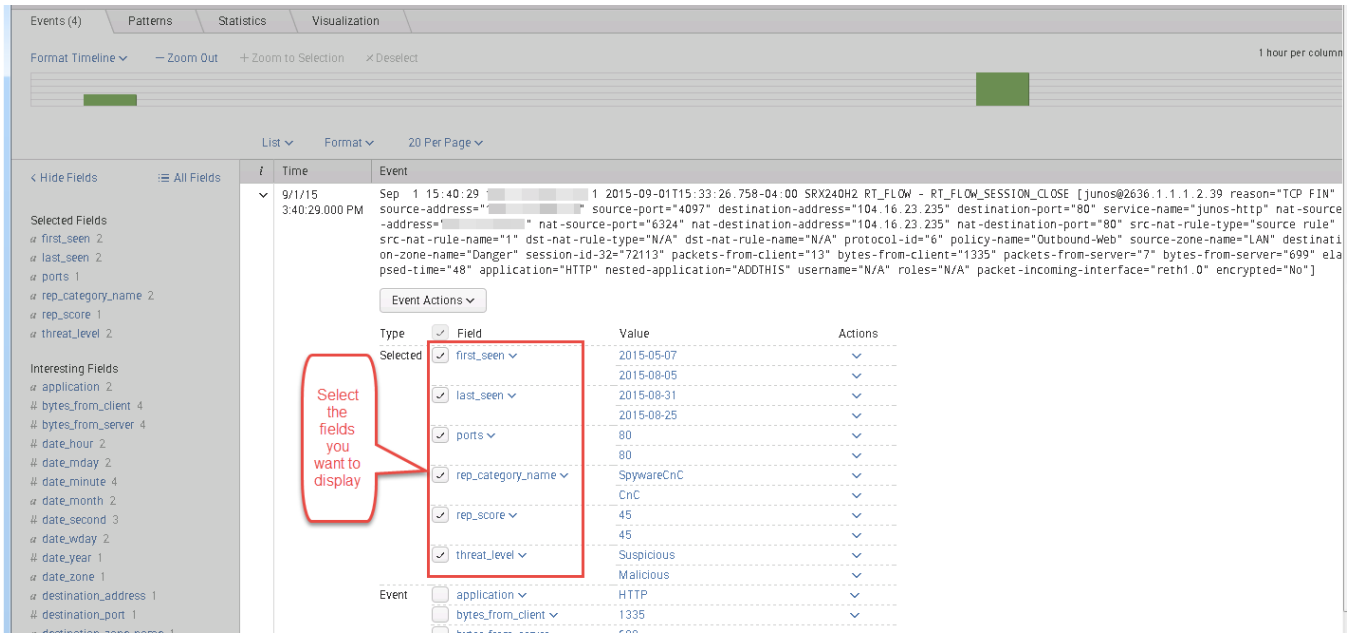


Figure 6: Selecting ET Fields to Display

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.

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 “destination-address”, but remember that Splunk represents it as “destination_address”. Finally we do some additional filtering on the output to match only logs that are Category CNC. Note that in the 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.

a. Simple Query: “host = 192.168.1.1 | `et_ip_lookup(IP=destination_address)` | search rep_category_name = CNC”

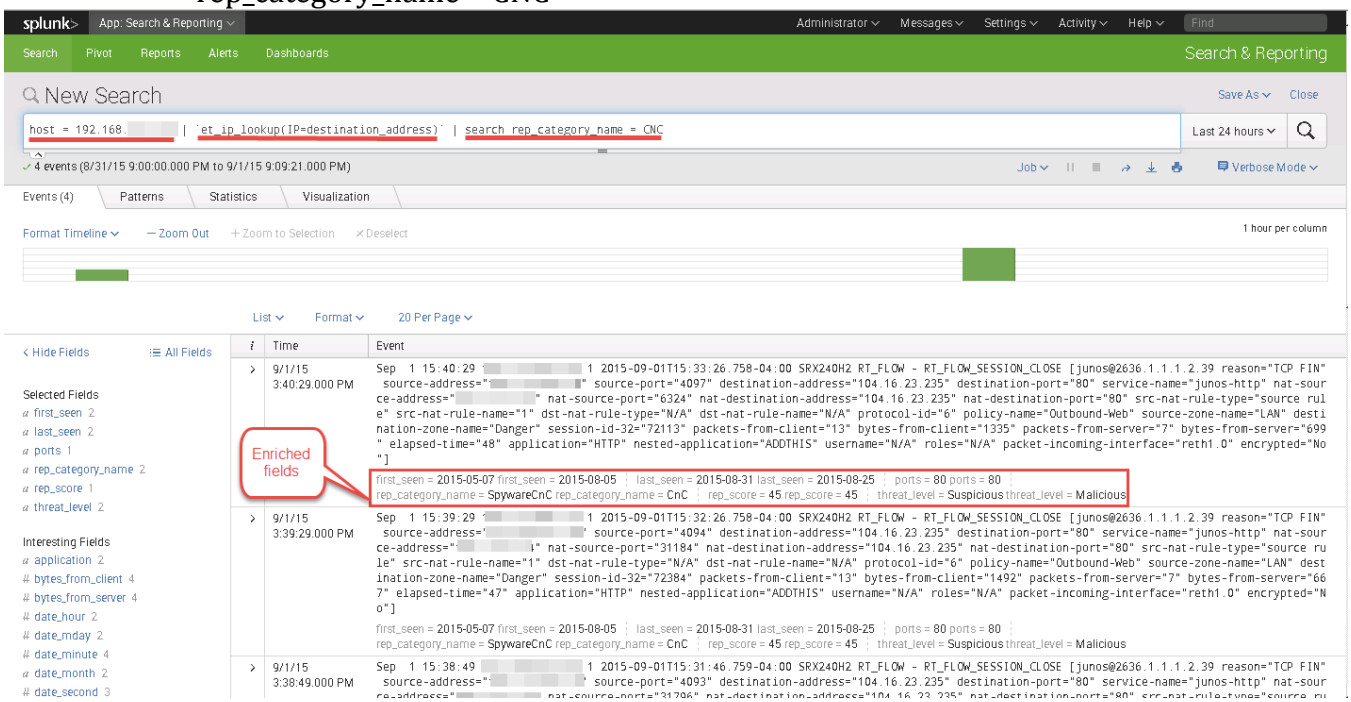


Figure 8: Simple Splunk Search with Macros

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.
    a. Advanced Query: “host = DNS | `et_domain_lookup(IP=DNS_Request)` | search rep_score >= 50 AND rep_category_name = CNC”

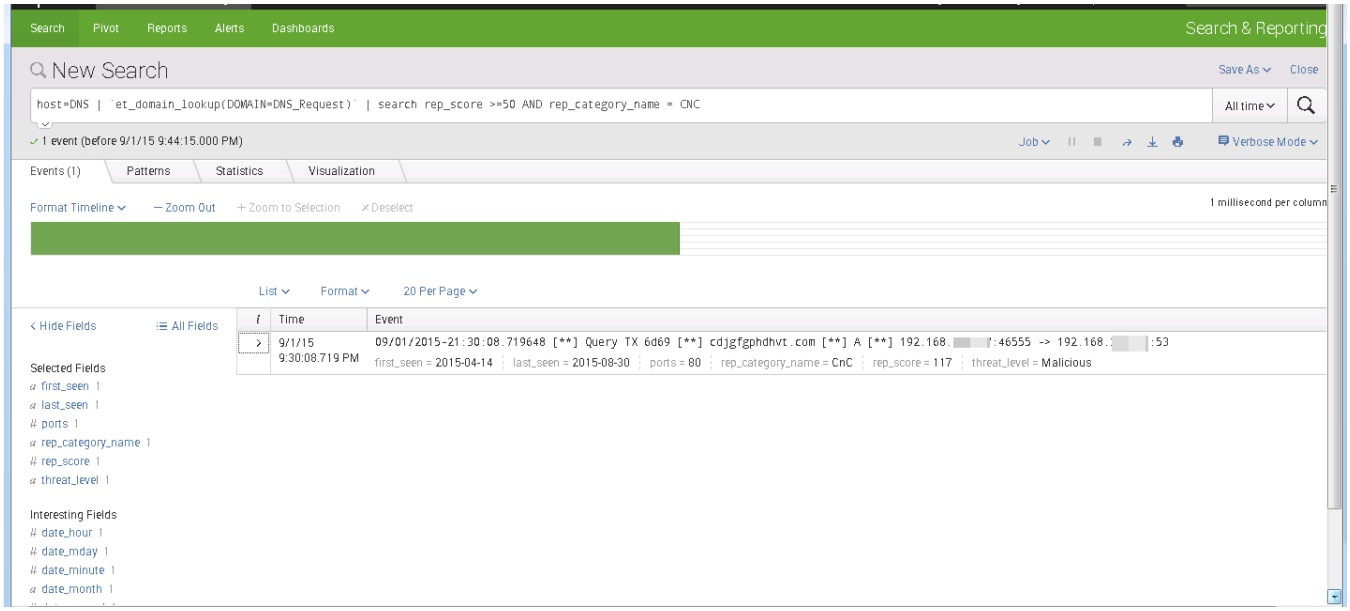


Figure 9: Advanced Splunk Search with Macros

In Figure 9, 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.

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.

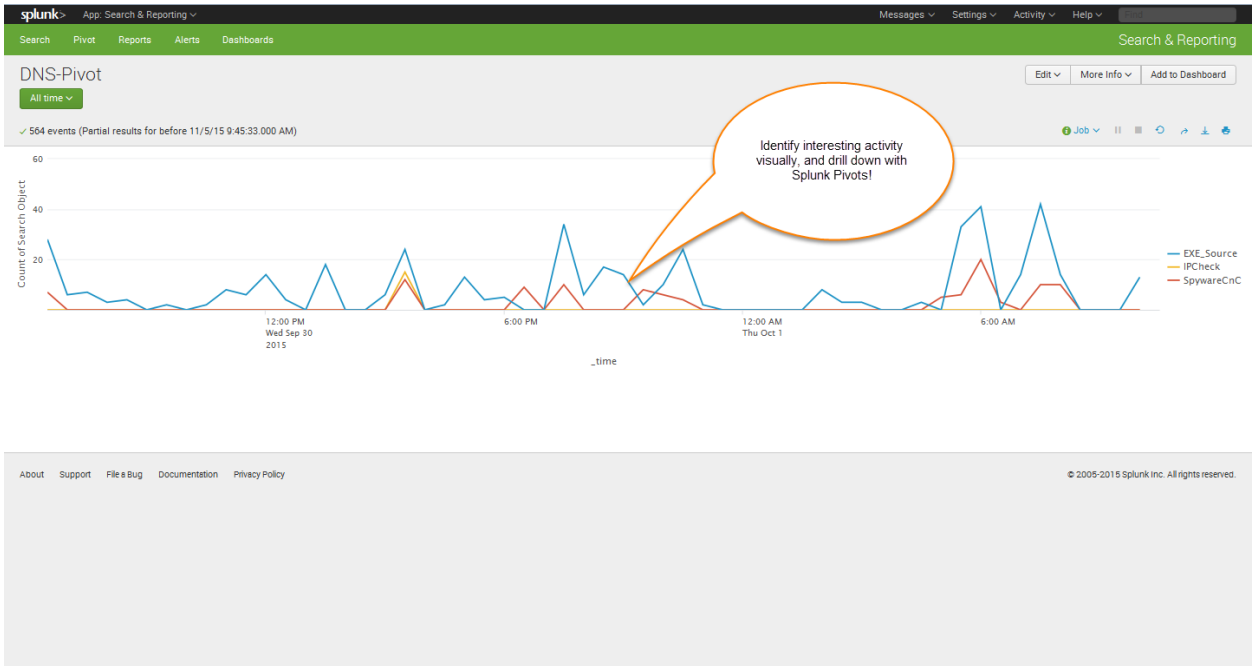


Figure 10: Building Splunk Pivots with ET Intelligence

Appendix

Categories

Category Index Number ,rep_category_name, rep_cat_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
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

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 Number,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
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