Search and Discover the Bad Guys in <= 60 Minutes

Purpose

- Apply analytics to your data via Splunk
- Demonstrate simplicity of doing this
- Lower the barrier to entry to explore your data

Not Everyone has effective and well-deployed enterprise security tools capable of telling them where the majority of the bad guys are.

No one

~~Not Everyone~~ has effective and well-deployed enterprise security tools capable of telling them where the majority of the bad guys are.

Doesn't my **new hotness** do that?

- Yes, no, um, well, kinda
- Sec COTS reports on what PM wanted
- Signature-based is helpful but misses a lot
- Anomaly can work, but …
- Amazing products exists, but don't do everything
- Maybe you don't have a new hotness
- “Your network” = “Your responsibility”

Who Am I?

- Sean Wilkerson, Partner/Consultant, Aplura

Who Am I?

- Sean Wilkerson, Partner/Consultant, Aplura
- ~15 Years of Network --> Systems --> **InfoSec**
- A Decade+ of Federal Log-Management
  - Half Spent Deploy/Manage FOSS/SIM/SIEM
  - Half Spent Deploy/Manage FOSS/Splunk
- SANS Log Mgmt Summits
- Splunk Pro Serv Partner Since 2008
- Splunk makes me happy

APLURA

Who Are You?

- You Need Better Visibility into Machine Data
- You Think Splunk is Right for You
- You Know Some Key Splunk Concepts
- You Don't Have or Don't Want to Rely on Infosec COTS Software Alone
- You Are **Analysts**!! <--- *Really important*

Splunk 001

- Complex, Scalable & Fast Search
- Efficiently and Flexibly Mines Machine Data(.)
- Concept of “Fields” (key=value)
- Thousands of Built-in Analytic Combos

  (Learn 5-10 and You Can Do Almost Anything)

Splunk 101

- Boolean Search (eg. NOT 2, y OR z)
- Fields (key=value with CIDR)
- Piped “|” expressions/functions
- Lookups
- The “pivot” ←Where It All Comes Together

Quick Disclaimer

- Data exploration could take until ...
- This pres includes many analytic elements
- Many of which will be immediately useful

Let's explore together

Content Available Now!

aplura.com/splunklive2013

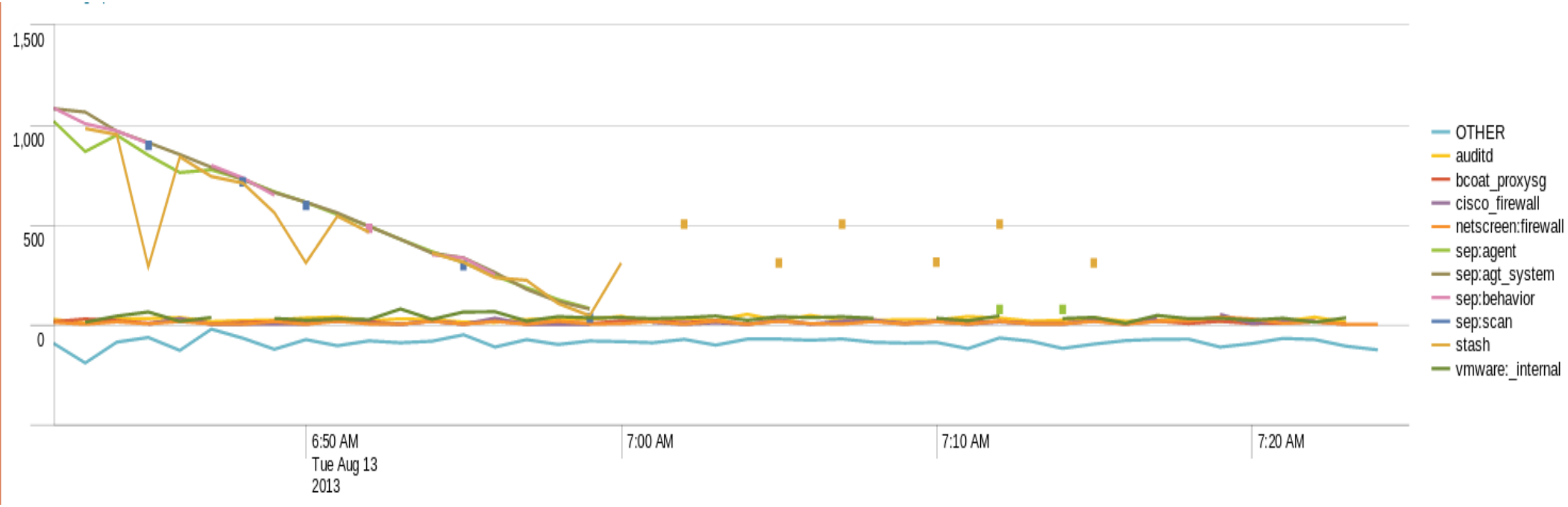
General Analysis

What Time do you Have?

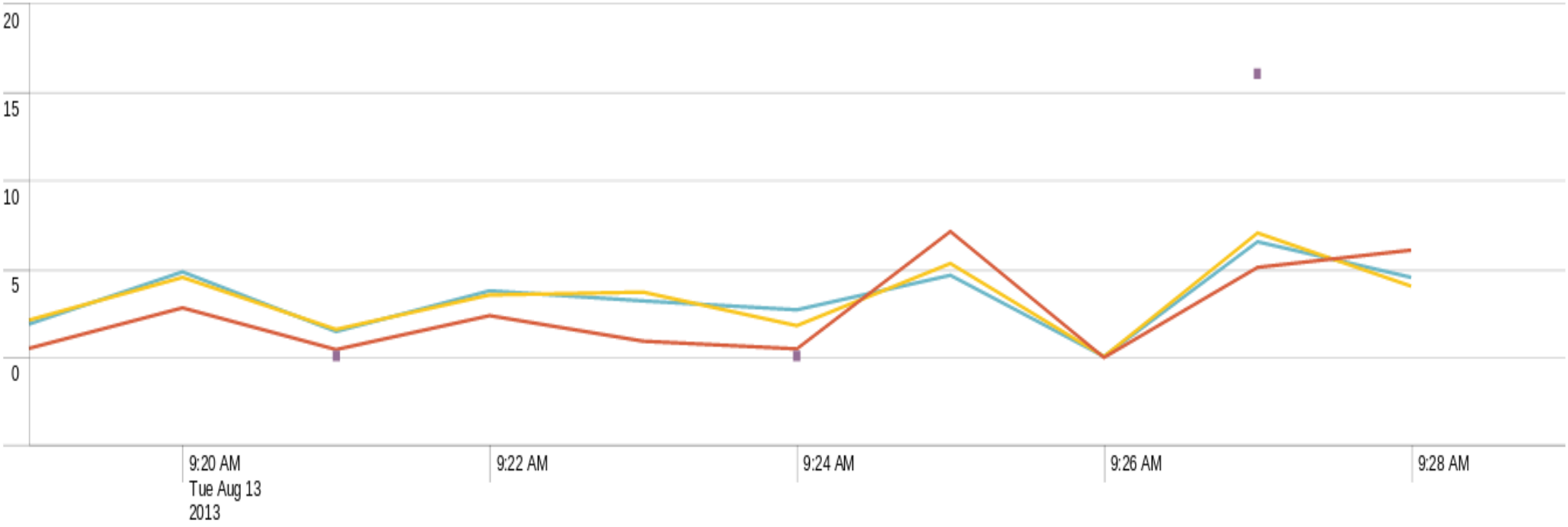
- Concept: Time is bad **everywhere** and this causes havoc during investigations. Do periodic time audits (it takes minutes with Splunk). As an analyst, you can validate your time BEFORE it is too late such as during an investigation.

- *** | eval timeDiff=_indextime-_time | timechart avg(timeDiff) by sourcetype**
- **sourcetype=firewalls | eval timeDiff=_indextime-_time | timechart avg(timeDiff) by host**
- **BONUS: _index_earliest=-h@h _index_latest=@h**

timeDiff by sourcetype



RT:timeDiff by Firewall_host



Off-time Activity

- Concept: Off-time activity can indicate a suspicious system. Splunk events include special built-in fields that are “time” aware

Weekdays after 8PM or before 7AM

- **(NOT (date_wday="Sat" OR date_wday="Sun") AND (date_hour>=20 OR date_hour<7)) OR date_wday="Sat" OR date_wday="Sun"**

Or if you need to “create” those built-in fields

- * | eval date_wday=strftime(_time,"%a") | eval date_hour=strftime(_time, "%H") **| search (NOT (date_wday="Sat" OR date_wday="Sun") AND (date_hour>=20 OR date_hour<7)) OR date_wday="Sat" OR date_wday="Sun"**

Activity by IP Range

- Concept: Groups of systems often have different patterns. So, analyze them by their group.

- **dest_ip=111.109.0.0/16 | top action**
- **sourcetype=checkpoint dest_ip=111.109.0.0/16 action=allowed**

 # eventtypes.conf

[network:all]

search = src_ip="111.109.0.0/16" OR dest_ip="111.109.0.0/16"

- **eventtype=network* | top action by eventtype**

Field Length Analysis

- Concept: Splunk makes analyzing the length of fields really easy. This is valuable to find malicious activity and mis-configurations

- **| eval Length=len(_raw) | where Length>2000**
  - len(http_referer) or len(domain) or len(uri) ...

Field/Data Manipulation

- Concept: During analysis, you often need new fields, or need to manipulate a piece of data to help with analysis.

- **tag=firewall | eval Firewall_Host=orig | top Firewall_Host**
- **tag=firewall | rex "orig\=(?<Firewall_Host>\d+\.\d+\.\d+\.\d+)\|"**
- **tag=firewall | rex "\;policy_name\=(?<policy_name>[^\]]+)\]"**

Use mode=sed to change fields like action or _raw

- **tag=firewall | rex field=action mode=sed "s/reject/blocked/" | top action**

Search

```
index=_internal host=log-prd-* source="/opt/splunk/var/log/splunk/metrics.log" sourceHost=* | timechart span=1m dc(sourceHost) by host
```

All time (real-time)

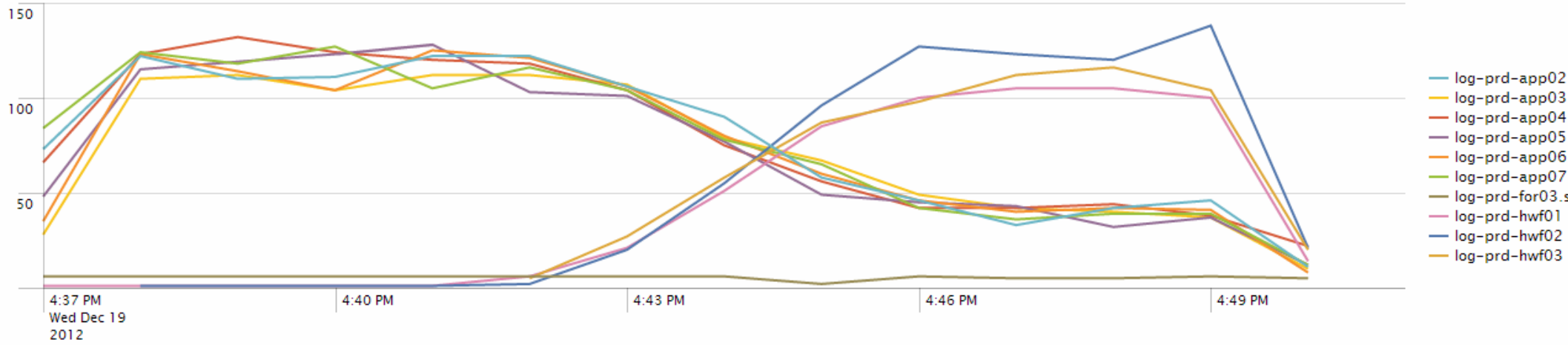
≥ 45,538 matching events | 99,297 scanned events

Save

≥ 14 results in total (real-time)

Options

Formatting options



Search

Smart Mode

```
index=_internal host=log-prd-* source="/opt/splunk/var/log/splunk/metrics.log" NOT host="log-prd-for03.          " sourceHost=* | rex field=host mode=sed "s/log-prd-app\d+([^\ ]+)/Indexer/" | rex field=host mode=sed "s/log-prd-hwf\d+([^\ ]+)/IntermediateHeavyForwarder/" | timechart span=1m dc(sourceHost) by host
```

Custom time

210,500 matching events

Save

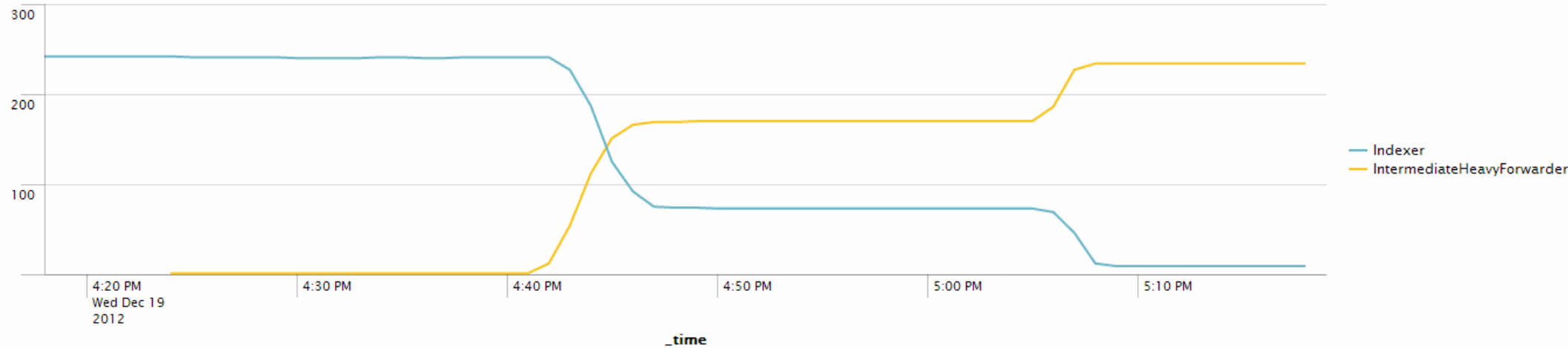
Create

61 results from 4:18:00 PM to 5:18:33 PM on Wednesday, December 19, 2012

Options

20 per page

Formatting options



Date in Search

- Concept: Don't you hate having to take your hands off the keyboard to use your mouse to manipulate the *Timepicker*? Me too.

- **earliest=-3h+22m latest=@h-10m**
- **earliest=-3d@d latest=-2d@d-1s**
- **BONUS: _index_earliest=-h@h _index_latest=@h**

Firewall Analysis

Allows by Previous Drops

- Concept: A FW drop violates policy. Now, let's inspect those offensive “source IPs”

Firewall events with source IPs not from our network that were blocked.

- **NOT eventtype=network:all_src tag=firewall NOT action=allowed | dedup src_ip | table src_ip**

Same as above as a “sub-search” against what WAS allowed.

- **NOT eventtype=network:all_src tag=firewall action=allowed [search NOT eventtype=network:all_src tag=firewall NOT action=allowed | dedup src_ip | table src_ip] | top src_ip by dest_ip**

Bytes Transfer Analysis

- Concept: Whether you are looking for malware payload or data exfiltration, bytes-transferred from your firewall/webpoxy/flow is GOOD!

  - **tag=flow | stats avg(bytes_out) by src_ip,dest_port**

Note: Combine with off-time and IP range for exciting results

Port Scanning Analysis

- Concept: Not all attacks are slow and low. Use Splunk to sniff-out port scanners and add it to your watchlist for later.

Port scanners

- **tag=firewall NOT eventtype=network:all_src | stats dc(dest_port) as Port_Count by src_ip | where Port_Count>50**

Host sweepers

- **tag=firewall NOT eventtype=network:all_src | stats dc(dest_ip) as IP_Count by src_ip | where IP_Count>50**

Webproxy Analysis

Browsed to IP

- Concept: Bare-ip browsing isn't illegal, but it shouldn't give you warm and fuzzies. Not always bad, but combined with other indicators bare ips are suspicious.

- tag=proxy | regex uri_domain="http\:\/\/\d+\.\d+\.\d+\.\d+" | rare uri_domain

Query Watchlist

- Concept: Use getwatchlist to pull any http(s)/ftp accessible delimited file into Splunk
- Tool: Getwatchlist

Pull down a csv of malwaredomains and save it to a lookup

- **| getwatchlist malwaredomains | outputlookup domain_watchlist.csv**

Correlate the lookup to your webproxy data to see if you have hits

- **tag=proxy [| inputlookup watchlist | table domain]**

Web Activity Timing

- Concept: Compare web activity by time and ...
- Tool: Geoip

Pull in geoip location fields to your search data

- **tag=proxy | lookup geoip clientip**

Use geoip data to find clients that browsed to several countries or more in a short period of time

- **tag=proxy | transaction clientip maxspan=60s maxpause=40s | lookup geoip clientip | stats dc(client_country) as Count by clientip | where Count>2**

URI/URL/File/Ext Analysis

- Concept: Evaluate web activity by URI, URL, File, and extension.

How many different file extensions were browsed to by source IP

- **tag=proxy | stats dc(fileextension) as Count by clientip | sort -Count**

How many different web files were downloaded with no referrer or UA

- **tag=proxy http_referer="-" http_user_agent="-" | stats dc(file) as Count by clientip | sort -Count**

Http User Agent Analysis

- Concept: UA strings are incredibly valuable and can be used in a variety of ways.
- Tool: uas_parser

Use uas_parser data to enrich your webproxy events with added fields

- **tag=proxy | lookup uas_lookup http_user_agent | search ua_type="unknown" | stats count by http_user_agent**
- **tag=proxy | lookup uas_lookup http_user_agent | top ua_family**
  - **ua_type, ua_company**

Long URI no Referrer ^M

- Concept: Deep analysis of URI's with no referrer

- tag=proxy http_referer="-" method="GET"
  | strcat uri_path uri_query uri
  | replace *- with * in uri
  | eval uri_length = len(uri)
  | stats count by src_ip dest_ip dest_host http_referrer uri_length uri
  | sort -uri_length

URL Length ^M

- Concept: URL/URI length can be indicative of malicious activity.

- tag=proxy
  | eval "URL Length"=len(uri)
  | eventstats avg(URL Length) AS "Average URL Length" stdev("URL Length") AS "Stdev URL Length"
  | eval Notable=tonumber('Average URL Length')+2 * tonumber('Stdev URL Length')
  | where 'URL Length'> Notable
  | table "domain" Category "URL Length" Notable "Average URL Length" "Stdev URL Length"
  | sort -"URL Length"

DNS Analysis

DNS + Webproxy

- Concept: DNS has significance since it is frequently ignored and a popular C&C vector for bad folks. Correlate DNS and Webproxy?

  - **tag=proxy OR tag=dns [| inputlookup watchlist | table domain]**

More Analysis Ideas

- Compare entropy
- Webproxy: Repeated errors (e.g. 404s and 500), SQL Injection discovery, Unique file/uri
- Auth: Failures to “admin” accounts
- DNS: Deeply nested hosts (lots of dots), TXT queries, Failed lookups, Systems doing a lot of lookups (or failed lookups), End points making strange queries (successive SOA), Odd TLDs, Spikes in lookups, Reverse lookups, Short TTLs, Responses with non-routable IPs

Case Study 1: Provided Trigger

Use known watch-list as a trigger

- **tag=proxy OR tag=dns [| getwatchlist malwaredomains | table domain] outputlookup watchlist**

Inspect URL length of suspicious internal hosts discovered by trigger

- **tag=proxy [| inputlookup watchlist] | eval Length=len(_raw) | where Length>2000**

Hrrm, nothing too bad here, let's check for ex-filtrated data

- **… | stats avg(bytes_out) by src_ip,dest_port**

Woot, we see some of these hosts have a large bytes_out

Time to collect data for forensics and reimage the systems

Case Study 2: BYO Trigger

Look for off-hours activity from critical systems to the Internet

- **tag=proxy eventtype=network:critical (NOT (date_wday="Sat" OR date_wday="Sun") AND (date_hour>=20 OR date_hour<7)) OR date_wday="Sat" OR date_wday="Sun" NOT eventtype=computer_updates | table clientip | outputlookup badclients**

Interesting, we found a few severs that had unexpected off-time activity

Let's look deeper at these hosts

- **tag=proxy [| inputlookup badclients] || transaction clientip maxspan=60s maxpause=40s | lookup geoip clientip | stats dc(client_country) as Count by clientip | where Count>2**

Oh, interesting, we have a few of these hosts that talked to servers in several different countries inside of a very short time-period

It appears as though our servers are being managed from a distributed C&C, now we need to figure out what to do about it.

Wrap Up

- “Lower the cost of exploration” - ^M
- Easily implement/test/evaluate
- Applies common analytic logic
- Easily pivot to validate or adjust strategy

Additional Resources

- docs.splunk.com - Manuals
- splunk-base.splunk.com - User forums
- Cheatsheet - duh!
- #CONF - Annual User Conf – Well-Worth the $

Thanks!

^M = Monzy Merza

Aplura's (Dave Shpritz and Dan Deighton)

Splunk Enterprise Security

Splunk Fed SEs (Mike Wilson, Scott Spencer)

Content Available Now!

Talks and Content:

- Best Practice PDF: aplura.com/splunkbp
- Talk: Security Analysis: aplura.com/splunklive2013
- Talk: Best Practice: aplura.com/splunklive2012
- Talk: SIEM Fails: aplura.com/lookbeforeyousim

FOCUSED INFORMATION SECURITY