Logs, Logs Every Where Nor Any Byte to Grok

Phil Hagen

phil@lewestech.com

@PhilHagen

http://gplus.to/+PhilHagen

THE MOST TRUSTED NAME IN INFORMATION AND SOFTWARE SECURITY TRAINING WORLDWIDE

SANS

Hyderabad

2014

24-29 November

SEC560:
Network Penetration Testing and Ethical Hacking
Instructor: Pieter Danhieux
GIAC Cert: GPEN

SANS

Secure India

2015

23 Feb – 7 Mar | Bangalore

SEC401:
Security Essentials Bootcamp Style
GIAC Cert: GSEC

SEC503:
Intrusion Detection In-Depth
GIAC Cert: GCIA

SEC504:
Hacker Tools, Techniques, Exploits & Incident Handling
GIAC Cert: GCIH

Contact: asiapacific@sans.org

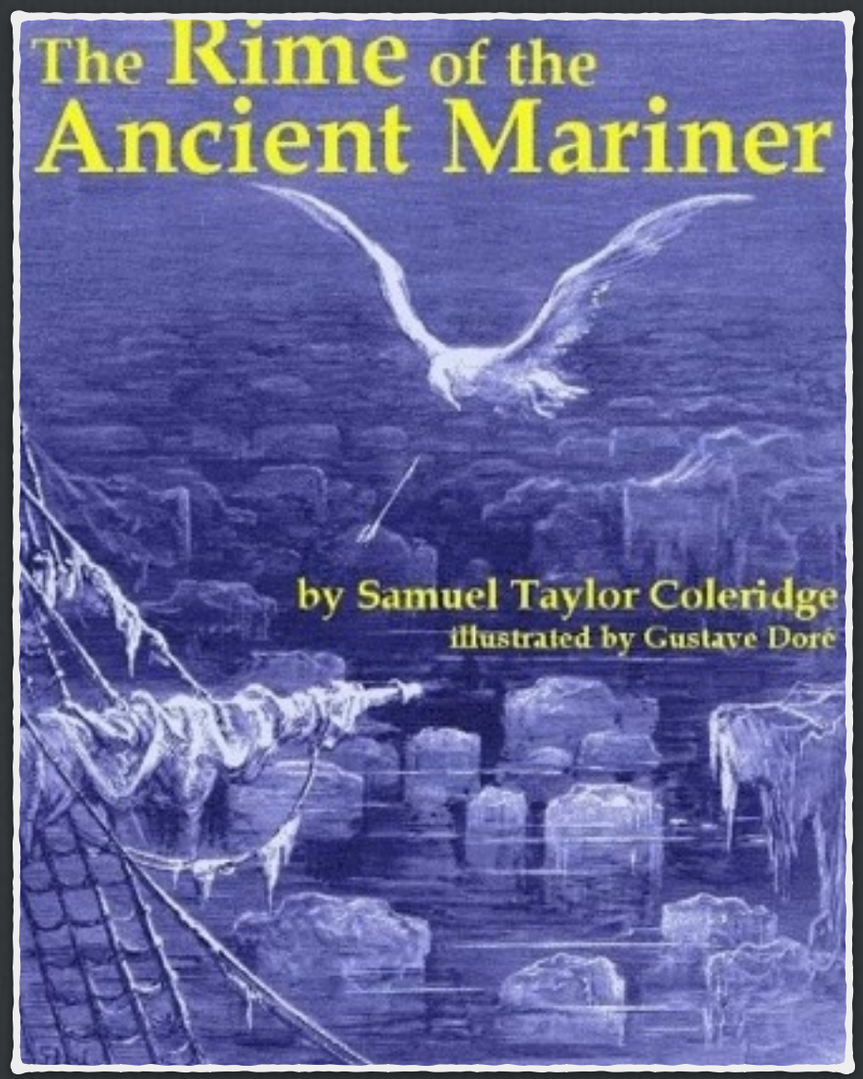
View all SANS training events: sans.org/security-training/by-location/apac

Who is This Dude?!

- ☐ SANS Certified Instructor and Course lead, FOR572: Advanced Network Forensics & Analysis
- ☐ Red Canary Managed Threat Detection Service
- ☐ Forensic/infosec consultant
- ☐ Former DoD/IC/LE contractor, USAF Comm Officer
- ☐ USAFA Computer Science

Phil's Notes™
@ cliffsnotes.com

The Best
New!
Just Got Better



- ☐ Albatross saves ship from Antarctic doom
- ☐ Captain kills albatross… Umm, what?!?
- ☐ Crew blames killing for various misfortunes but they flip-flop a few times
- ☐ Then they all die (one at a time)
- ☐ Captain lives with the burden of an "albatross around his neck"

Don't let log evidence be your forensic albatross!

Relevance to Network Forensics

- Network forensics often lacks long-term evidence archive: pcap, NetFlow, etc may be limited in time or scope (if available!)
- We seek Artifacts of Communication to complete observations
- These artifacts often come in the form of log data
  - Many platforms
  - Multiple observation points
  - Managing lots of sources and formats can be a challenge!

Reasonable rsyslog.conf

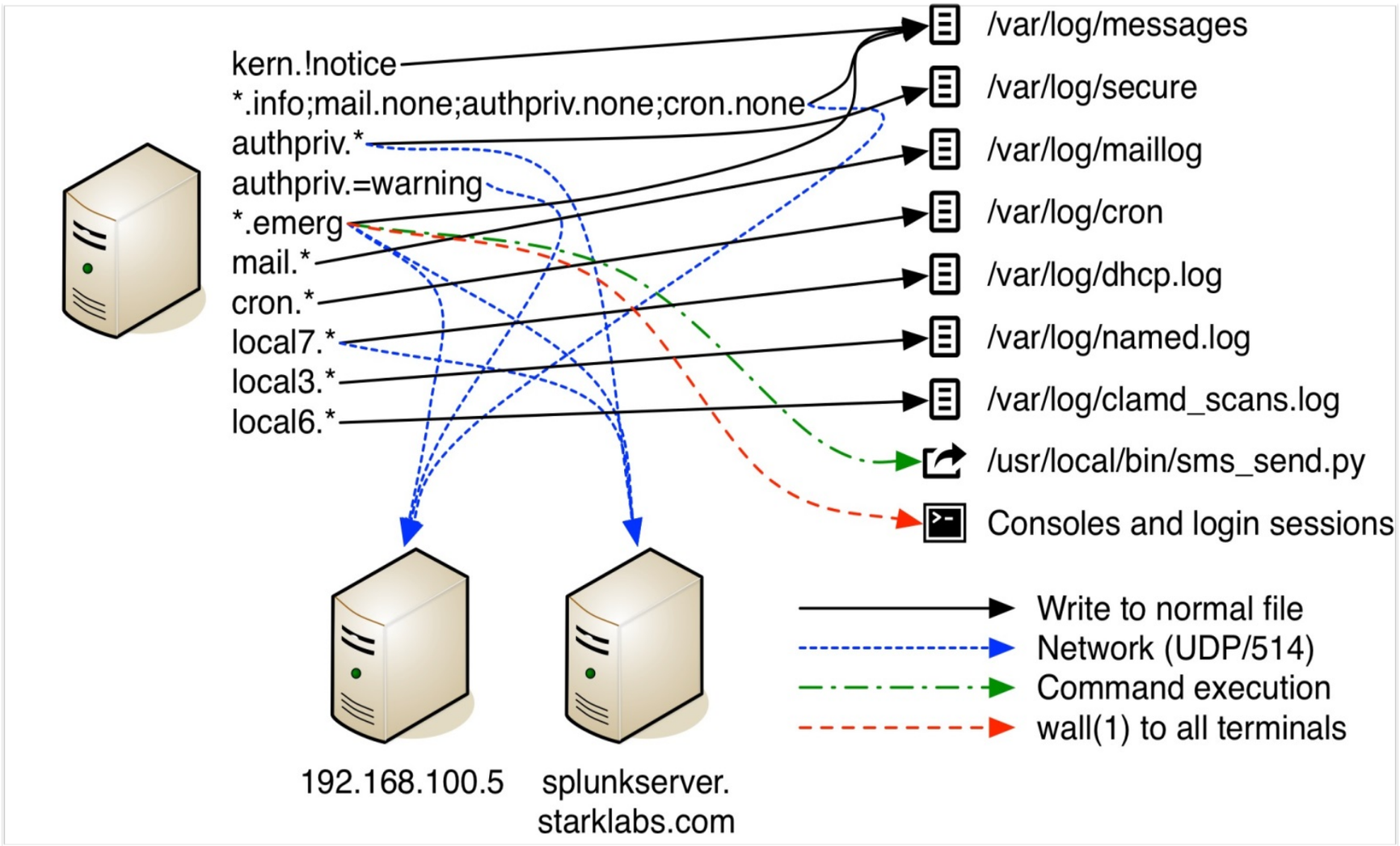
```
$ActionFileDefaultTemplate RSYSLOG_TraditionalFileFormat
$template smsTemplate,"%TIMESTAMP% %HOSTNAME% %syslogtag% %msg%"
$ModLoad imklog
$ModLoad imuxsock

kern.!notice                                    /var/log/messages
*.info;mail.none;authpriv.none;cron.none        /var/log/messages
*.info;mail.none;authpriv.none;cron.none        @192.168.100.5

authpriv.*                                      /var/log/secure
authpriv.=warning                               @192.168.100.5
authpriv.*                                      @splunkserver.starklabs.com

*.emerg                                         *
*.emerg                                         @192.168.100.5
*.emerg                                         @splunkserver.starklabs.com
*.emerg                                         ^/usr/local/bin/sms_send.py;smsTemplate

mail.*                                          /var/log/maillog
cron.*                                          /var/log/cron

local7.*                                        /var/log/dhcp.log
local7.*                                        @splunkserver.starklabs.com
local3.*                                        /var/log/named.log
local6.*                                        /var/log/clamd_scans.log
```

Normal Files

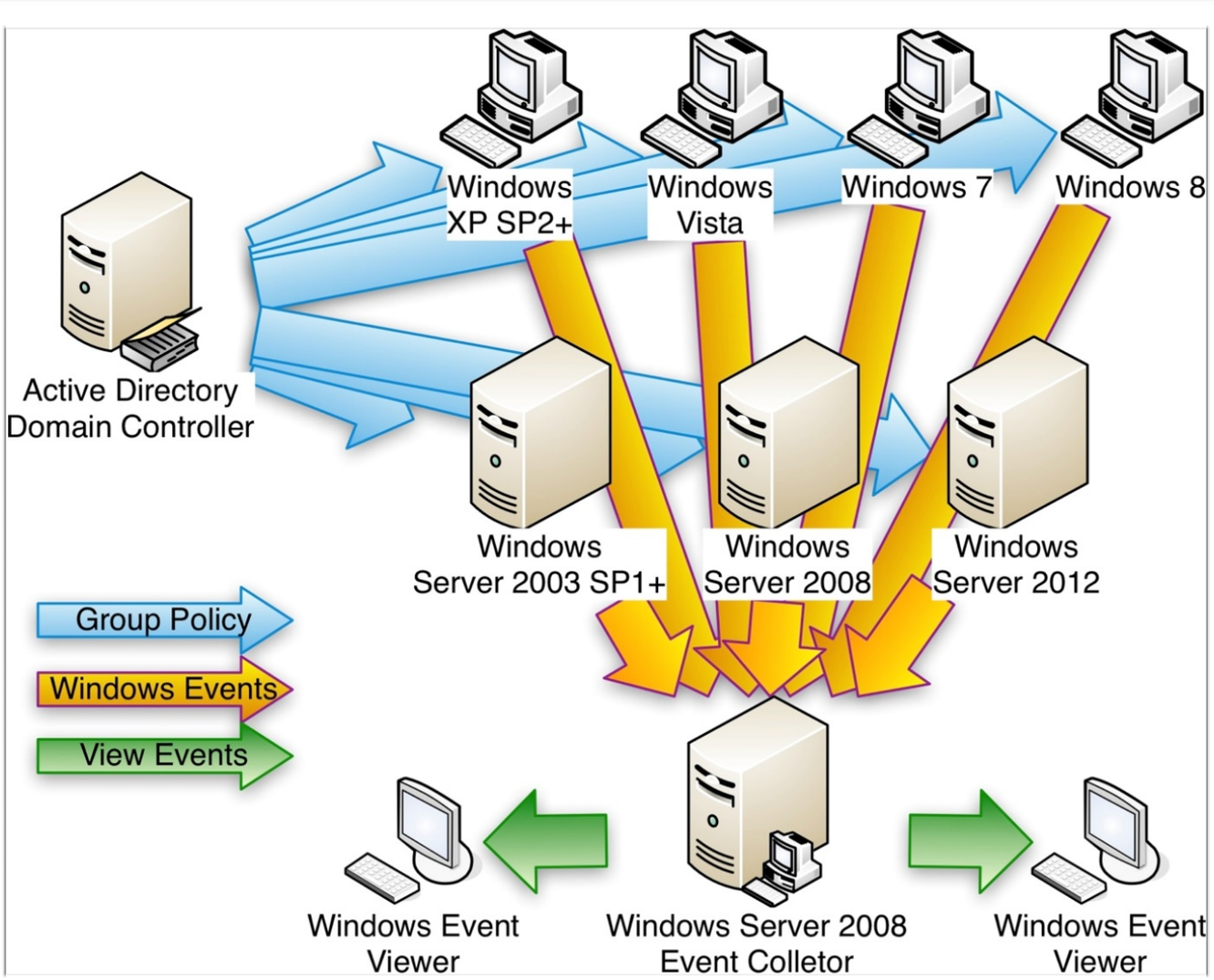
Remote systems

External programs

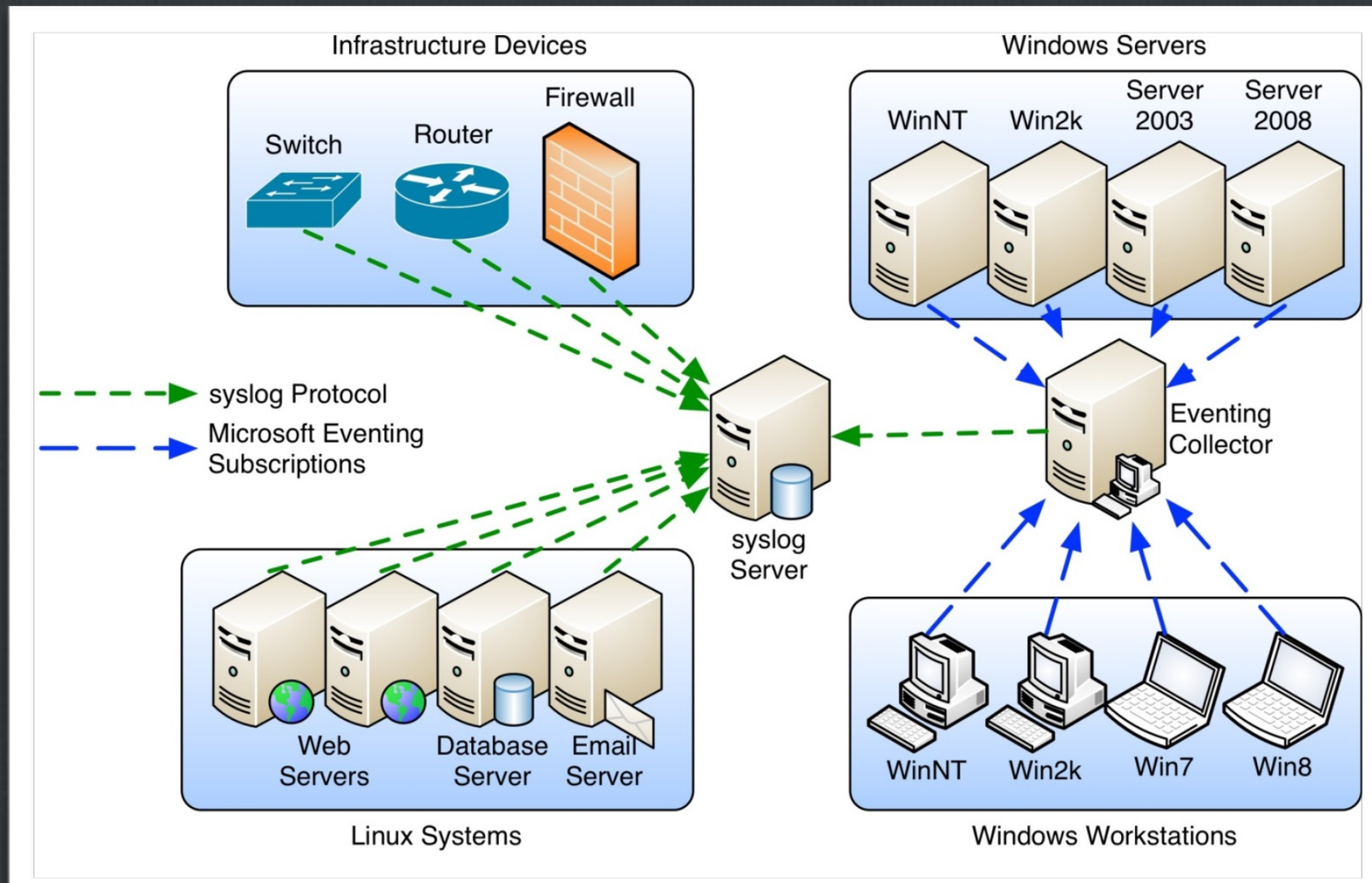
Same Configuration, Visualized



Windows Isn't Much Simpler



Each Camp Might Play Well With Others



Double-Edged Sword…

- The ubiquity of log data can be a HUGE ASSET to the forensicator …

…except when it's the biggest headache you could ever imagine!

Logging Shortfalls

- Network devices tend to have volatile storage
  - Reboot, overflow, corruption? Lost logs
- Distributed log storage = distributed analysis
  - Best case: Collect from far and wide
  - Inefficient use of analysts' time
  - Correlating can be difficult
- Multiple log formats require multiple tools

Live Log Aggregation Wins!

- ☐ Prevent attackers covering tracks
- ☐ Multiple simultaneous analysts/functions
- ☐ Centralized analysis, distributed data sources
  - ☐ Track event across the environment
  - ☐ Correlate event from multiple vantage points
- ☐ Just like paper printer for badge/keypad door

Aggregate to Forensicate (Otherwise, Frustrate)

- ☐ Built-ins: rsyslog, syslog-ng, MS Eventing 6.0
- ☐ SIEMs: Arc Sight, Trustwave, Tenable
- ☐ Pure aggregators: ELSA, Splunk,

Logstash!

Logstash: Super Sweeeeeet

- Free, open-source: http://logstash.net
- Developed at Dreamhost
  - Made to scale huge
- Now part of Elasticsearch
- Kibana web frontend
- Great developer and community support

Logstash at a Glance (1)

- Extensive “grok” pattern-matching syntax
  - Test at http://grokdebug.herokuapp.com
- Can operate on standalone system or in larger hierarchy
- LOTS of inputs: Files, syslog via network socket, NetFlow, Twitter, IMAP, SNMP, named pipes, more
- Filter inputs to match relevant fields
- Many outputs: E-mail, pagers, file, search database, HTTP JSON

Logstash at a Glance (2)

- Not a SIEM, but similar to how forensicators tend to use SIEMs
- Crazy-simple installation
  - Source, apt-get, YUM
  - FOR572 distributes VMware image, incl. config file & relevant parsers
  - README and MD5/SHA256:
    `http://bit.ly/for572-logstash-readme`
- Latest Kibana interface ROCKS
  - Configure/manage/save/share dashboard rows and panels

FOR572 Ex-4.3 Dashboard

Time filter

QUERY

Search query (Apache Lucene)

FILTERING

No filters available

Data filters

TIMELINE

View | count per **1s** | (**0** hits)

Time-based event histogram

SOURCE SYSTEMS

SOURCE SYSTEMS

Term Count Action

SOURCE PROGRAMS

SOURCE PROGRAMS

Term Count Action

Dashboard dials and widget-y things

DOCUMENTS

Fields

All (1) / Current (0)

Type to filter...

Field List

_source (select columns from the list to the left)

0 to **0** of 0 available for paging

Event table

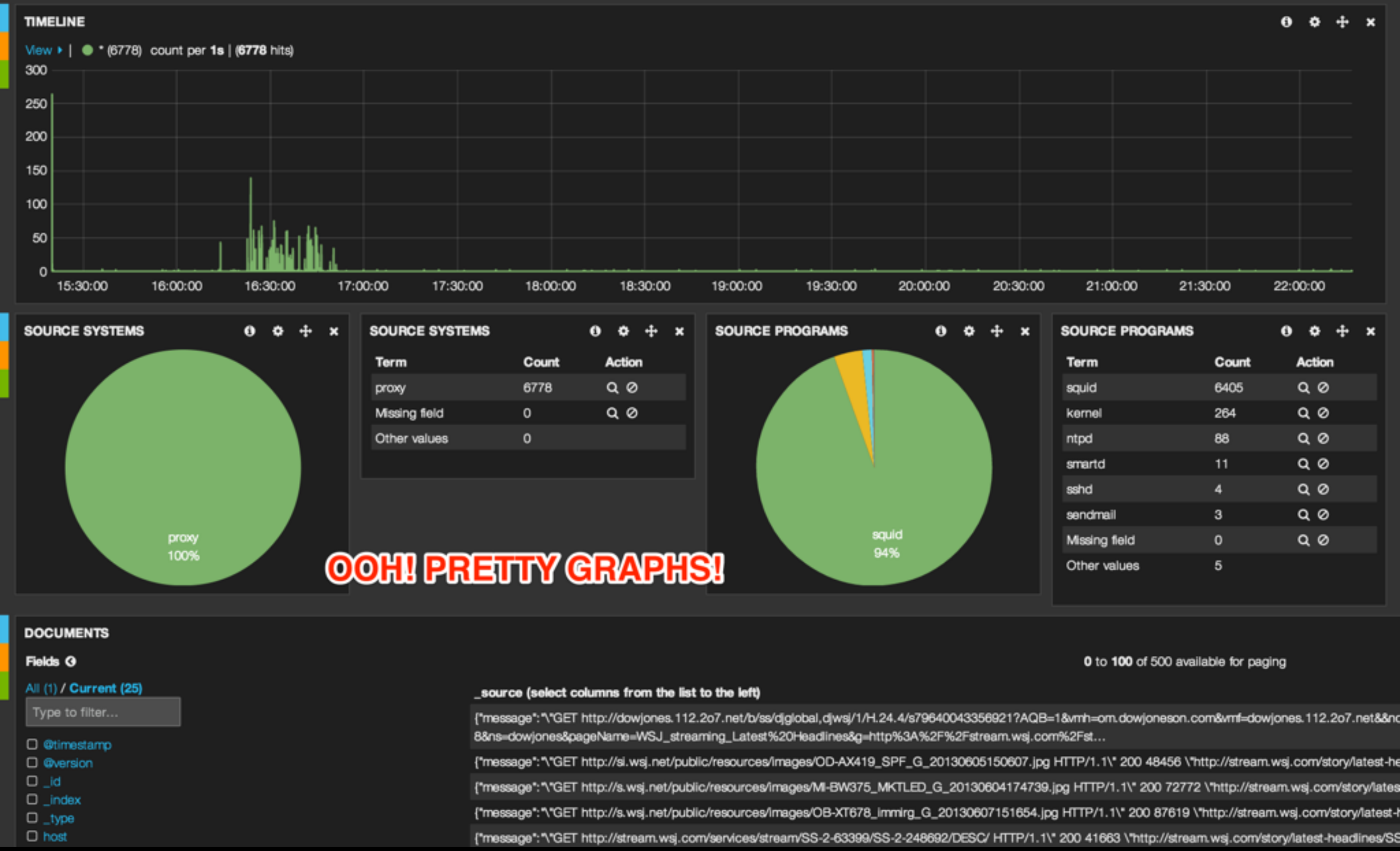
0 to **0** of 0 available for paging

Let's Load Some Data

- Load logs from a squid proxy server (syslog and squid-specific)
  - Caution: syslog doesn't "do" years - must be inferred from metadata! (Ex: FOR572 VM parses the year from directory tree)

```
$ cd /usr/local/logstash-ingest/
$ unzip ~/exercise_source_logs/proxy_logs.zip
Archive:  /home/ls_user/exercise_source_logs/proxy_logs.zip
   creating: 2013/proxy_logs/
   creating: 2013/proxy_logs/proxy/
  inflating: 2013/proxy_logs/proxy/secure
   creating: 2013/proxy_logs/proxy/squid/
  inflating: 2013/proxy_logs/proxy/squid/access_log
  inflating: 2013/proxy_logs/proxy/messages
  inflating: 2013/proxy_logs/proxy/maillog
```

- What happens next?



- This is a FOR572-specific dashboard, but you can create and share row/panel layouts that best suit your needs

Time filter

Enable/disable

FILTERING

time must

field : @timestamp

from : "2013-06-08T16:08:02.452Z"

to : "2013-06-08T17:02:52.401Z"

{"message":"\"GET http://b.scorecardresearch.com/p?c1=8&c2=6035092&c3=0 HTTP/1.1\" 200 418 \"http://www.foxnews.

View: Table / JSON / Raw

Field	Action	Value
@timestamp		2013-06-08T16:31:09.000Z
@version		1
_id		lewxTcUgQa2T0rdYf-BjA
_index		logstash-2013.06.08
_type		syslog
host		proxy
logstash_source		0.0.0.0
message		"GET http://b.scorecardresearch.com/p?c1=8&c2=6035092&c3=0 HTTP/1.1" 200 418 "http Safari/537.36" TCP_MISS:DIRECT
path		/usr/local/logstash-ingest/2013/proxy_logs/proxy/squid/access_log
		2014-04-17 18:27:36 UTC
squid_bytes		418
squid_client		10.3.59.54
squid_fulluri		http://b.scorecardresearch.com/p?c1=8&c2=6035092&c3=0
squid_host		b.scorecardresearch.com
squid_method		GET
squid_referer		http://www.foxnews.com/static/v/all/html/ad-ifr.html?id=frame1-300x250&ns=friendlyComm
squid_request		p?c1=8&c2=6035092&c3=0
squid_returncode		200
squid_useragent		Mozilla/5.0 (Windows NT 6.1; WOW64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/
syslog_hostname		proxy
syslog_pid		2342
syslog_program		squid
syslog_timestamp		Jun 8 12:31:09 2013
tags		got_syslog_timestamp,got_path_year,got_syslog_host,got_syslog_program,got_squid_acce
type		syslog

Original message

squid- and syslog-specific parse fields

squid_bytes		418
squid_client		10.3.59.54
squid_fulluri		http://b.scorecar 8&c2=6035092&c3=0
squid_host		b.scorecardresearch.com
squid_method		GET
squid_referer		http://www.foxnews.com/static/v/all/html/ad-ifr.html?id=frame1-30
squid_request		p?c1=8&c2=6035092&c3=0
squid_returncode		200
squid_useragent		Mozilla/5.0 (Windows NT 6.1; WOW64) AppleWebKit/537.36 (KHTM
syslog_hostname		proxy

Require/exclude as filter

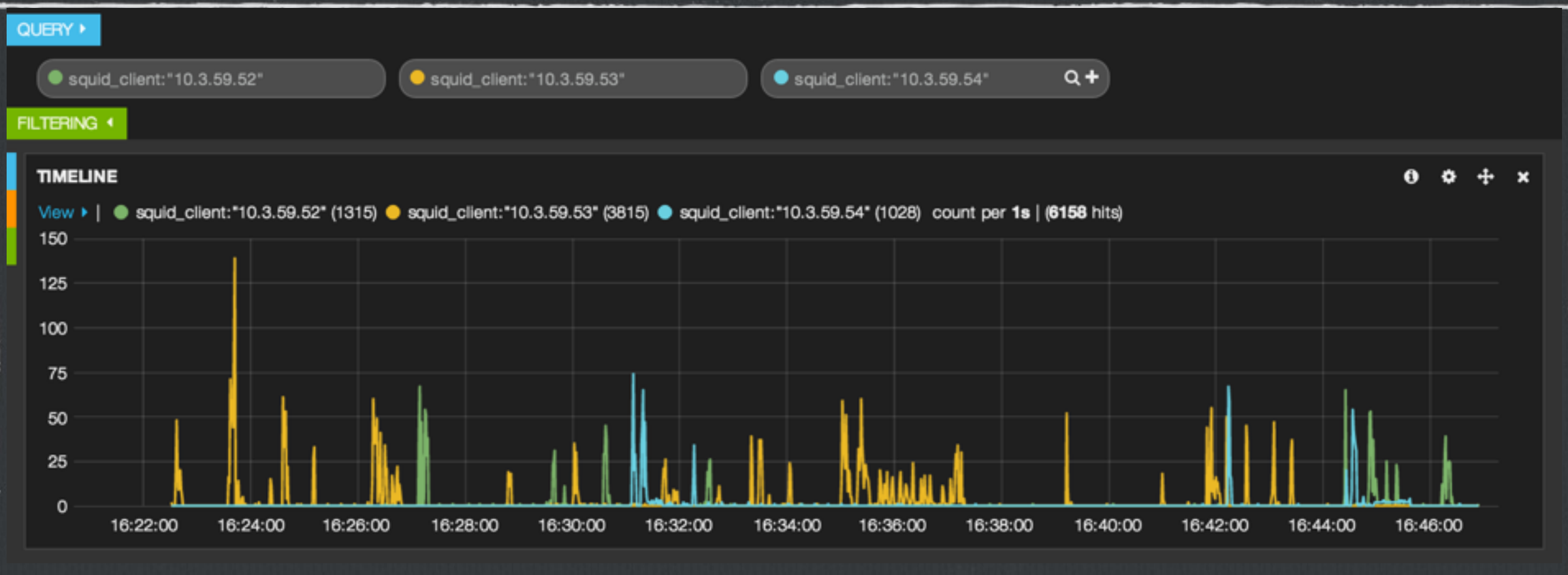
Field filter

FILTERING

field must

field : squid_client

query : "10.3.59.54"



- ☐ Multiple queries are color-coded on histogram = Quick visual pattern recognition
  - ☐ Web proxy activity time frame per client
  - ☐ Earliest/latest occurrence of activity within data set
  - ☐ Trends and peaks/valleys over time

Load Even More Data

- ☐ Windows clients (via SNARE), DHCP, DNS, Firewall, etc.

```
$ cd /usr/local/logstash-ingest/
$ unzip ~/exercise_source_logs/commonuse_windows_logs.zip
...
$ unzip ~/exercise_source_logs/muse_logs.zip
...
$ unzip ~/exercise_source_logs/fw-router_logs.zip
...
```

QUERY

dhcp_hwaddr:"08:00:27:df:33:6d" | dhcp_hwaddr:"08:00:27:ad:38:ca" | dhcp_hwaddr:"08:00:27:c9:40:4f"

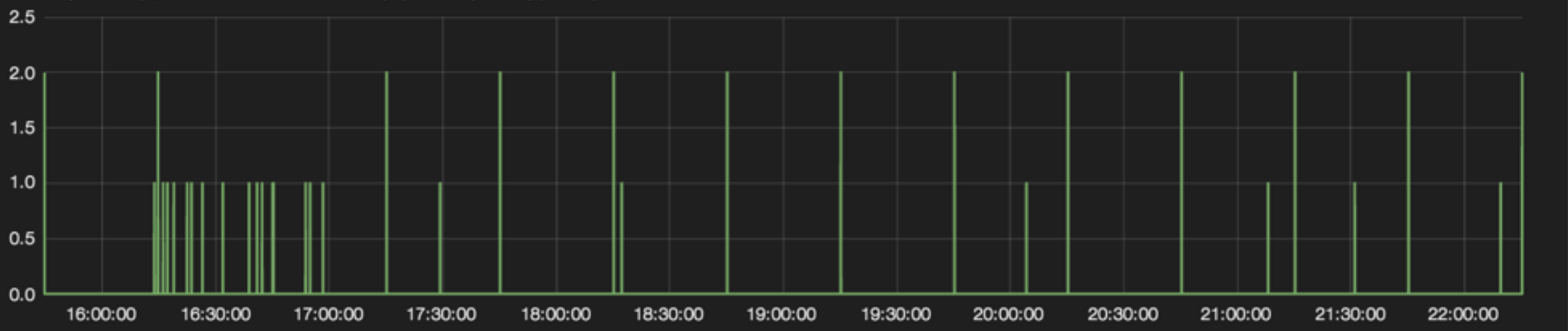
FILTERING

field must
field : dhcp_messagetype
query : "DHCPACK"

Disable queries

TIMELINE

View | dhcp_hwaddr:"08:00:27:df:33:6d" (50) count per **1s** | (**50** hits)



GRAPH

Identify DHCP client assignment by MAC address

DOCUMENTS

Fields

All (1) / Current (19)

Type to filter...

- [x] @timestamp
- [ ] @version
- [ ] _id
- [ ] _index
- [ ] _type
- [x] dhcp_hostname
- [x] dhcp_hwaddr
- [x] dhcp_ip
- [ ] dhcp_messagetype
- [ ] host
- [ ] logstash_source
- [ ] message
- [ ] path

0 to **50** of 50 available for paging

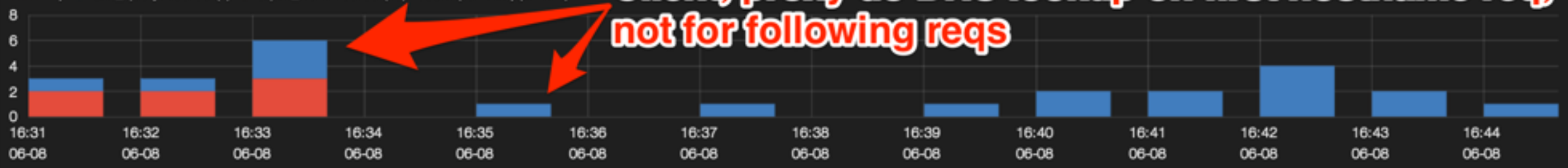
@timestamp	dhcp_hostname	dhcp_hwaddr	dhcp_ip
2013-06-08T19:45:12.000Z	floor2-PC	08:00:27:df:33:6d	10.3.59.53
2013-06-08T18:45:11.000Z	floor2-PC	08:00:27:df:33:6d	10.3.59.53
2013-06-08T18:15:11.000Z	floor2-PC	08:00:27:df:33:6d	10.3.59.53
2013-06-08T16:58:23.000Z		08:00:27:df:33:6d	10.3.59.53
2013-06-08T16:45:15.000Z	floor2-PC	08:00:27:df:33:6d	10.3.59.53
2013-06-08T16:40:59.000Z		08:00:27:df:33:6d	10.3.59.53
2013-06-08T16:14:49.000Z	floor2-PC	08:00:27:df:33:6d	10.3.59.53
2013-06-08T20:45:12.000Z	floor2-PC	08:00:27:df:33:6d	10.3.59.53
2013-06-08T18:17:18.000Z		08:00:27:df:33:6d	10.3.59.53
2013-06-08T18:15:11.000Z	floor2-PC	08:00:27:df:33:6d	10.3.59.53

But how does this apply to a broader data set?

Search for DNS and squid proxy log entries related to the ever-popular “ .co.cc” top-level domain…

TIMELINE

View ▸ | dns_query:*.co.cc (7) squid_host:*.co.cc (19) count per **1m** | (**26** hits)



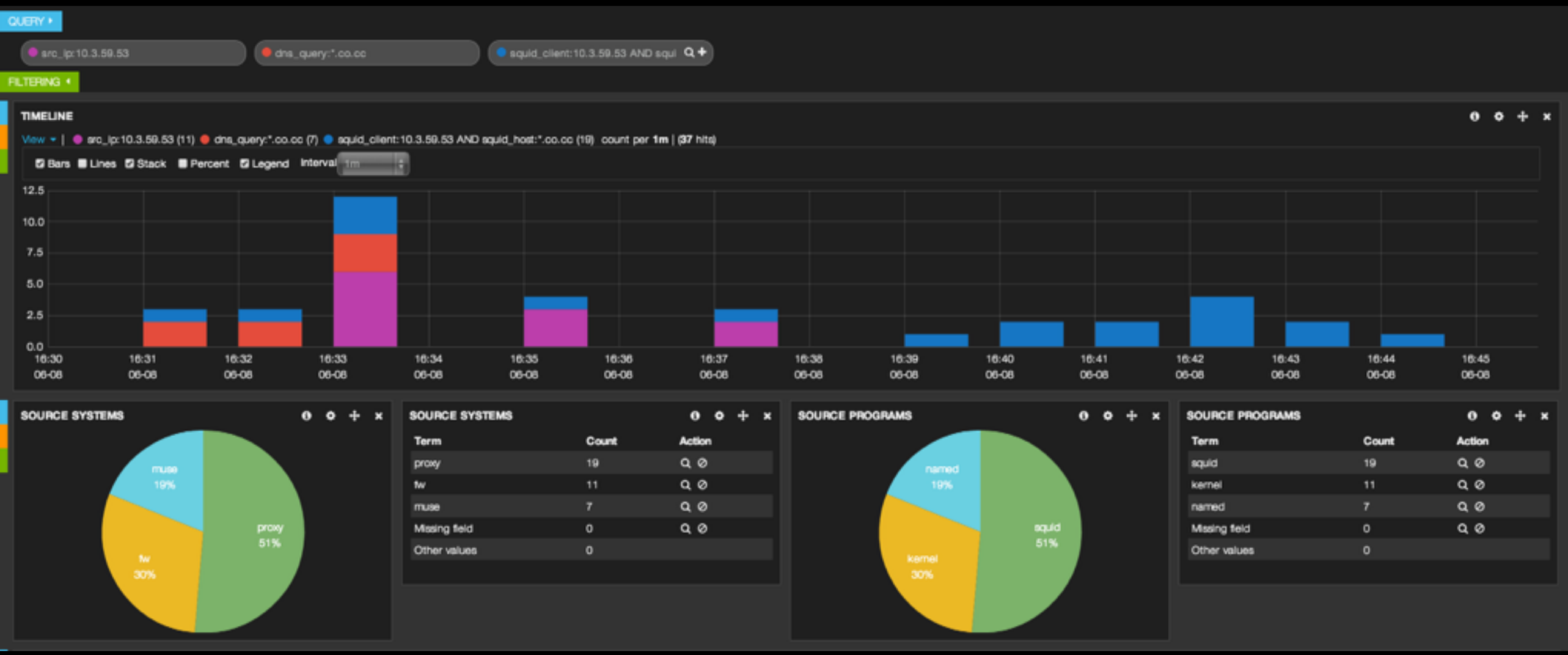
Client, proxy do DNS lookup on first hostname req, not for following reqs

GRAPH

DOCUMENTS

0 to **26** of 26 available for paging

@timestamp	dns_client	dns_query		squid_client	squid_host	squid_request	squid_bytes	squid_useragent
2013-06-08T16:31:44.000Z	10.3.59.53	download.taxforms.usa.gov.ccmktiejgpq.co.c…	a1					
2013-06-08T16:31:50.000Z	10.3.16.11	download.taxforms.usa.gov.ccmktiejgpq.co.c…	a2					
2013-06-08T16:31:54.000Z			HTTP req	10.3.59.53	download.taxforms.usa.gov.ccmktiejgpq.co.c…	taxdocs.exe	4133241	Mozilla/5.0 (Windows NT 6.1; WOW64) A…
2013-06-08T16:32:43.000Z	10.3.59.53	download.taxforms.usa.gov.ccmktiejgpq.co.c…	a3					
2013-06-08T16:32:48.000Z			HTTP req	10.3.59.53	download.taxforms.usa.gov.ccmktiejgpq.co.c…	2014_tax_schedules.pdf	6048145	feevil 18b
2013-06-08T16:32:52.000Z	10.3.59.53	download.taxforms.usa.gov.ccmktiejgpq.co.c…	a4					
2013-06-08T16:33:12.000Z	10.3.59.53	pfumqhbrowahfgk.co.cc	b1					
2013-06-08T16:33:54.000Z	10.3.59.53	qmrigskajgfslqm.co.cc	c1					
2013-06-08T16:33:54.000Z	10.3.16.11	qmrigskajgfslqm.co.cc	d1					
2013-06-08T16:33:55.000Z				10.3.59.53	qmrigskajgfslqm.co.cc	checkin/?nodeID=7a%3A6f%3A6d%3A62%3A69%3A6…	302	feevil 18c
2013-06-08T16:33:56.000Z				10.3.59.53	qmrigskajgfslqm.co.cc	cmdreq/	403	feevil 18c
2013-06-08T16:33:58.000Z				10.3.59.53	qmrigskajgfslqm.co.cc	cmd/?c=orders	309	feevil 18c
2013-06-08T16:35:58.000Z			HTTP reqs	10.3.59.53	qmrigskajgfslqm.co.cc	cmd/?c=orders	309	feevil 18c
2013-06-08T16:37:58.000Z				10.3.59.53	qmrigskajgfslqm.co.cc	cmd/?c=orders	309	feevil 18c
2013-06-08T16:39:58.000Z				10.3.59.53	qmrigskajgfslqm.co.cc	cmd/?c=orders	308	feevil 18c
2013-06-08T16:40:59.000Z				10.3.59.53	qmrigskajgfslqm.co.cc	cmd/?c=orders	367	feevil 18c



- [ ] Add a search for any firewall blocks from the client IP…

BUT WHAT DOES IT ALL MEAN?!

DOCUMENTS

0 to 37 of 37 available for paging

@timestamp	src_ip	dst_ip	proto	dst_port	dns_client	dns_query	squid_client	squid_host	squid_request
2013-06-08T16:31:44.000Z					10.3.59.53	download.taxforms.usa.gov.ccmk...			
2013-06-08T16:31:50.000Z					10.3.16.11	download.taxforms.usa.gov.ccmk...			
2013-06-08T16:31:54.000Z							10.3.59.53	download.taxforms.usa.gov.ccmk...	taxdocs.exe
2013-06-08T16:32:43.000Z					10.3.59.53	download.taxforms.usa.gov.ccmk...			
2013-06-08T16:32:48.000Z							10.3.59.53	download.taxforms.usa.gov.ccmk...	2014_tax_schedules.pdf
2013-06-08T16:32:52.000Z					10.3.59.53	download.taxforms.usa.gov.ccmk...			
2013-06-08T16:33:12.000Z	10.3.59.53	198.46.151.44	TCP	17924					
2013-06-08T16:33:12.000Z					10.3.59.53	ptumqhbrowahfgk.co.cc			
2013-06-08T16:33:15.000Z	10.3.59.53	198.46.151.44	TCP	17924					
2013-06-08T16:33:21.000Z	10.3.59.53	198.46.151.44	TCP	17924					
2013-06-08T16:33:33.000Z	10.3.59.53	198.46.151.41	TCP	17924					
2013-06-08T16:33:36.000Z	10.3.59.53	198.46.151.41	TCP	17924					
2013-06-08T16:33:42.000Z	10.3.59.53	198.46.151.41	TCP	17924					
2013-06-08T16:33:54.000Z					10.3.59.53	qmrigskajgfsiqm.co.cc			
2013-06-08T16:33:54.000Z					10.3.16.11	qmrigskajgfsiqm.co.cc			
2013-06-08T16:33:55.000Z							10.3.59.53	qmrigskajgfsiqm.co.cc	checkin/?nodeID=7a%3A8f%3A6d%3...
2013-06-08T16:33:56.000Z							10.3.59.53	qmrigskajgfsiqm.co.cc	cmdreq/
2013-06-08T16:33:58.000Z							10.3.59.53	qmrigskajgfsiqm.co.cc	cmd/?c=orders
2013-06-08T16:35:06.000Z	10.3.59.53	72.44.47.21	TCP	843					
2013-06-08T16:35:22.000Z	10.3.59.53	72.44.47.21	TCP	843					
2013-06-08T16:35:25.000Z	10.3.59.53	72.44.47.21	TCP	843					
2013-06-08T16:35:58.000Z							10.3.59.53	qmrigskajgfsiqm.co.cc	cmd/?c=orders
2013-06-08T16:37:10.000Z	10.3.59.53	72.44.47.21	TCP	843					
2013-06-08T16:37:13.000Z	10.3.59.53	72.44.47.21	TCP	843					
2013-06-08T16:37:58.000Z							10.3.59.53	qmrigskajgfsiqm.co.cc	cmd/?c=orders
2013-06-08T16:39:58.000Z							10.3.59.53	qmrigskajgfsiqm.co.cc	cmd/?c=orders
2013-06-08T16:40:59.000Z							10.3.59.53	qmrigskajgfsiqm.co.cc	cmd/?c=orders
2013-06-08T16:40:59.000Z							10.3.59.53	qmrigskajgfsiqm.co.cc	cmd/?c=orders
2013-06-08T16:41:29.000Z							10.3.59.53	qmrigskajgfsiqm.co.cc	cmd/?c=orders
2013-06-08T16:41:29.000Z							10.3.59.53	qmrigskajgfsiqm.co.cc	cmd/?c=orders
2013-06-08T16:42:00.000Z							10.3.59.53	qmrigskajgfsiqm.co.cc	cmd/?c=orders
2013-06-08T16:42:03.000Z							10.3.59.53	qmrigskajgfsiqm.co.cc	cmd/?c=orders
2013-06-08T16:42:33.000Z							10.3.59.53	qmrigskajgfsiqm.co.cc	cmd/?c=orders

DNS lookups and 2x proxy downloads

@timestamp	dns_client	dns_query	squid_client	squid_host	squid_request
2013-06-08T16:31:44.000Z	10.3.59.53	download.taxforms.usa.gov.ccmktiejgpq.co.cc			
2013-06-08T16:31:50.000Z	10.3.16.11	download.taxforms.usa.gov.ccmktiejgpq.co.cc			
2013-06-08T16:31:54.000Z			10.3.59.53	download.taxforms.usa.gov.ccmktiejgpq.co.cc	taxdocs.exe
2013-06-08T16:32:43.000Z	10.3.59.53	download.taxforms.usa.gov.ccmktiejgpq.co.cc			
2013-06-08T16:32:48.000Z			10.3.59.53	download.taxforms.usa.gov.ccmktiejgpq.co.cc	2014_tax_schedules.pdf

DNS lookup 2nd host, TCP/17924 to two IPs blocked

@timestamp	dns_client	dns_query	src_ip	dst_ip	proto	dst_port
2013-06-08T16:33:12.000Z			10.3.59.53	198.46.151.44	TCP	17924
2013-06-08T16:33:12.000Z	10.3.59.53	pfumqhbrowahfgk.co.cc				
2013-06-08T16:33:15.000Z			10.3.59.53	198.46.151.44	TCP	17924
2013-06-08T16:33:21.000Z			10.3.59.53	198.46.151.44	TCP	17924
2013-06-08T16:33:33.000Z			10.3.59.53	198.46.151.41	TCP	17924
2013-06-08T16:33:36.000Z			10.3.59.53	198.46.151.41	TCP	17924
2013-06-08T16:33:42.000Z			10.3.59.53	198.46.151.41	TCP	17924

DNS lookup 3rd host, HTTP activity via proxy, TCP/843 to new IP (Amazon EC2) blocked, further HTTP activity via proxy

@timestamp	dns_client	dns_query	squid_client	squid_host	squid_request	src_ip	dst_ip	proto	dst_port
2013-06-08T16:33:54.000Z	10.3.59.53	qmrigskajgtsiqm.co.cc							
2013-06-08T16:33:54.000Z	10.3.16.11	qmrigskajgtsiqm.co.cc							
2013-06-08T16:33:55.000Z			10.3.59.53	qmrigskajgtsiqm.co.cc	checkin/?nodeID=7a%3A6f%3A6d%3A62%3A69%3A65%3A2d%3…				
2013-06-08T16:33:56.000Z			10.3.59.53	qmrigskajgtsiqm.co.cc	cmdreq/				
2013-06-08T16:33:58.000Z			10.3.59.53	qmrigskajgtsiqm.co.cc	cmd/?c=orders				
2013-06-08T16:35:06.000Z						10.3.59.53	72.44.47.21	TCP	843
2013-06-08T16:35:22.000Z						10.3.59.53	72.44.47.21	TCP	843
2013-06-08T16:35:25.000Z						10.3.59.53	72.44.47.21	TCP	843
2013-06-08T16:35:58.000Z			10.3.59.53	qmrigskajgtsiqm.co.cc	cmd/?c=orders				
2013-06-08T16:37:10.000Z						10.3.59.53	72.44.47.21	TCP	843
2013-06-08T16:37:13.000Z						10.3.59.53	72.44.47.21	TCP	843
2013-06-08T16:37:58.000Z			10.3.59.53	qmrigskajgtsiqm.co.cc	cmd/?c=orders				
2013-06-08T16:39:58.000Z			10.3.59.53	qmrigskajgtsiqm.co.cc	cmd/?c=orders				
2013-06-08T16:40:59.000Z			10.3.59.53	qmrigskajgtsiqm.co.cc	cmd/?c=orders				

Future Developments

- Logstash can receive NetFlow Directly
  - http://www.rsreese.com/parsing-netflow-using-kibana-via-logstash-to-elasticsearch
- Or, parse existing NetFlow to text and ingest normally
- Log2timeline/Plaso CSV output into Logstash
  - http://diftdisk.blogspot.com/2014/02/first-look-at-log2timeline-timeline-in.html

Summary

- Lots of value in logs - especially in network forensics
  - Ephemeral evidence and long-term investigation scope… seek Artifacts of Communication
- Incorporate arbitrary data to a single analytic workflow
  - See the whole picture in one tool
- Consider a “to-go” VM that can be used during investigation
  - Very little spin-up needed, consistent processes
  - Current FOR572 README, download link, MD5/SHA256: `http://bit.ly/for572-logstash-readme`

THE MOST TRUSTED NAME IN INFORMATION AND SOFTWARE SECURITY TRAINING WORLDWIDE

SANS

Hyderabad

2014

24-29 November

SEC560:
Network Penetration Testing and Ethical Hacking
Instructor: Pieter Danhieux
GIAC Cert: GPEN

SANS

Secure India

2015

23 Feb – 7 Mar | Bangalore

SEC401:
Security Essentials Bootcamp Style
GIAC Cert: GSEC

SEC503:
Intrusion Detection In-Depth
GIAC Cert: GCIA

SEC504:
Hacker Tools, Techniques, Exploits & Incident Handling
GIAC Cert: GCIH

Contact: asiapacific@sans.org

View all SANS training events: sans.org/security-training/by-location/apac