Finding the needle in the haystack with ELK

Elasticsearch for Incident Handlers and Forensic Analysts

by Christophe@Vandeplas.com
Whoami

- Working for the Belgian Government
  - my own company
  - Incident Handling
  - Malware analysis
  - Forensics (network + system)
- Open Source minded
- Creator of MISP – Malware Information Sharing Platform
- Creator of pystemon – pastebin monitoring tool
- Core organizer of the FOSDEM conference for many years
- Contact me: christophe@vandeplas.com
Finding the needle in the haystack with ELK

Elasticsearch for Incident Handlers and Forensic Analysts

by Christophe@Vandeplas.com
What tools do you use?

- Text logs
- notepad
- Grep
- awk / sed / cut
- MS Excel / OOo Calc
grep -F log.txt

zgrep -F log.txt

zgrep -f patterns.txt -F log.txt

find "$LOGS_DIR" -iname "*.gz" -print0 | parallel --gnu -0 -n1 -P8 zgrep -f patterns.txt -F > result-all.txt

Fast for single search, however no column lookup!
Optimizing

- MySQL / MS Access
- Splunk
  - free = 500MB/day
- ELSA – Enterprise Log Search and Archive
  - Limitation of the # of columns
- ${COMMERCIAL_TOOL}
Trick for Splunk Addicts

- Limit is 500 MB /day
- 3 license violations allowed per month
- Set the date to 00:01 AM
- Index as much as possible 24h/day for 3 days (while loops are your friend)
- Enjoy searching
Trick for all = ELK

- Elasticsearch  Logstash  Kibana
- Index as much as you want
- **No limit** on volume, speed or position of the moon
- Open Source, Free to use, commercial support

logstash  kibana
Configurations

- [https://github.com/cvandeplas/ELK-forensics](https://github.com/cvandeplas/ELK-forensics)
- Repository with Logstash and Kibana configurations
- Mactime, BlueCoat, Mail IMSS, IWSVA, IIS, SuperTimeline, Plaso, ...

- Our focus today:
  - Forensics and Incident Handling
  - Batch-Import
### CLIENT INFO

#### T_C_IP
<table>
<thead>
<tr>
<th>Term</th>
<th>Count</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.0.1.196</td>
<td>39838</td>
<td>Q, Q, Q</td>
</tr>
<tr>
<td>10.1.337</td>
<td>35860</td>
<td>Q, Q, Q</td>
</tr>
<tr>
<td>10.1.337</td>
<td>22604</td>
<td>Q, Q, Q</td>
</tr>
<tr>
<td>10.8.1.37</td>
<td>10782</td>
<td>Q, Q, Q</td>
</tr>
<tr>
<td>192.168.1.337</td>
<td>457</td>
<td>Q, Q, Q</td>
</tr>
<tr>
<td>192.168.1.337</td>
<td>4445</td>
<td>Q, Q, Q</td>
</tr>
<tr>
<td>192.168.1.337</td>
<td>4166</td>
<td>Q, Q, Q</td>
</tr>
<tr>
<td>10.1.331</td>
<td>3617</td>
<td>Q, Q, Q</td>
</tr>
<tr>
<td>10.1.331</td>
<td>3242</td>
<td>Q, Q, Q</td>
</tr>
<tr>
<td>10.1.331</td>
<td>2599</td>
<td>Q, Q, Q</td>
</tr>
<tr>
<td>Missing field</td>
<td>0</td>
<td>Q, Q, Q</td>
</tr>
<tr>
<td>Other values</td>
<td>6133</td>
<td>Q, Q, Q</td>
</tr>
</tbody>
</table>

#### X-FORWARDED-FOR
<table>
<thead>
<tr>
<th>Term</th>
<th>Count</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.0.1.196</td>
<td>1643</td>
<td>Q, Q, Q</td>
</tr>
<tr>
<td>10.1.337</td>
<td>9148</td>
<td>Q, Q, Q</td>
</tr>
<tr>
<td>10.1.337</td>
<td>7198</td>
<td>Q, Q, Q</td>
</tr>
<tr>
<td>10.8.1.37</td>
<td>4031</td>
<td>Q, Q, Q</td>
</tr>
<tr>
<td>192.168.1.337</td>
<td>3885</td>
<td>Q, Q, Q</td>
</tr>
<tr>
<td>10.1.331</td>
<td>3794</td>
<td>Q, Q, Q</td>
</tr>
<tr>
<td>10.1.331</td>
<td>2780</td>
<td>Q, Q, Q</td>
</tr>
<tr>
<td>10.1.331</td>
<td>2780</td>
<td>Q, Q, Q</td>
</tr>
<tr>
<td>10.1.331</td>
<td>2242</td>
<td>Q, Q, Q</td>
</tr>
<tr>
<td>Missing field</td>
<td>0</td>
<td>Q, Q, Q</td>
</tr>
<tr>
<td>Other values</td>
<td>90971</td>
<td>Q, Q, Q</td>
</tr>
</tbody>
</table>

#### USER AGENT

<table>
<thead>
<tr>
<th>Term</th>
<th>Count</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mozilla/5.0 (compatible; MSIE 9.0; Windows NT 6.1; Trident/5.0)</td>
<td>60126</td>
<td>Q, Q, Q</td>
</tr>
<tr>
<td>FSORSP/1.1.25.2280 (Win32)</td>
<td>27633</td>
<td>Q, Q, Q</td>
</tr>
<tr>
<td>-</td>
<td>15451</td>
<td>Q, Q, Q</td>
</tr>
<tr>
<td>TMULF/5.0</td>
<td>8681</td>
<td>Q, Q, Q</td>
</tr>
<tr>
<td>Mozilla/5.0 (Windows NT 6.1) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/27.0.1453.110 Safari/537.36</td>
<td>4961</td>
<td>Q, Q, Q</td>
</tr>
<tr>
<td>Mozilla/5.0 (Windows NT 6.1; rv:26.0) Gecko/20100101 Firefox/26.0</td>
<td>4577</td>
<td>Q, Q, Q</td>
</tr>
<tr>
<td>Mozilla/5.0 (Windows NT 6.1; Trident/4.0; InfoPath.1; .NET CLR 2.0.50727; .NET4.0C; .NET4.0E)</td>
<td>3307</td>
<td>Q, Q, Q</td>
</tr>
<tr>
<td>Mozilla/5.0 (Windows NT 6.1; WOW64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/31.0.1650.63 Safari/537.36</td>
<td>2809</td>
<td>Q, Q, Q</td>
</tr>
<tr>
<td>Mozilla/4.0 (compatible; MSIE 8.0; Windows NT 5.1; Trident/4.0; InfoPath.1; .NET CLR 2.0.50727; .NET4.0C; .NET4.0E)</td>
<td>1826</td>
<td>Q, Q, Q</td>
</tr>
<tr>
<td>Mozilla/4.0 (compatible; MSIE 5.5; Win32)</td>
<td>1349</td>
<td>Q, Q, Q</td>
</tr>
<tr>
<td>Missing field</td>
<td>5</td>
<td>Q, Q, Q</td>
</tr>
<tr>
<td>Other values</td>
<td>15004</td>
<td>Q, Q, Q</td>
</tr>
</tbody>
</table>

### C_DNS

<table>
<thead>
<tr>
<th>Term</th>
<th>Count</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>Missing field</td>
<td>145719</td>
<td>Q, Q, Q</td>
</tr>
<tr>
<td>Other values</td>
<td>0</td>
<td>Q, Q, Q</td>
</tr>
</tbody>
</table>

### CS_USER

<table>
<thead>
<tr>
<th>Term</th>
<th>Count</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>Missing field</td>
<td>145719</td>
<td>Q, Q, Q</td>
</tr>
<tr>
<td>Other values</td>
<td>0</td>
<td>Q, Q, Q</td>
</tr>
</tbody>
</table>

### BROWSER

<table>
<thead>
<tr>
<th>Term</th>
<th>Count</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chrome/31.0.1650.63 Safari/537.36</td>
<td>2809</td>
<td>Q, Q, Q</td>
</tr>
<tr>
<td>Mozilla/5.0 (Windows NT 6.1; WOW64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/31.0.1650.63 Safari/537.36</td>
<td>2809</td>
<td>Q, Q, Q</td>
</tr>
<tr>
<td>Mozilla/4.0 (compatible; MSIE 8.0; Windows NT 5.1; Trident/4.0; InfoPath.1; .NET CLR 2.0.50727; .NET4.0C; .NET4.0E)</td>
<td>1826</td>
<td>Q, Q, Q</td>
</tr>
<tr>
<td>Mozilla/4.0 (compatible; MSIE 5.5; Win32)</td>
<td>1349</td>
<td>Q, Q, Q</td>
</tr>
<tr>
<td>Missing field</td>
<td>5</td>
<td>Q, Q, Q</td>
</tr>
<tr>
<td>Other values</td>
<td>15004</td>
<td>Q, Q, Q</td>
</tr>
<tr>
<td>Term</td>
<td>Count</td>
<td>Action</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-------</td>
<td>--------</td>
</tr>
<tr>
<td><a href="http://www.verkeerscentrum.be">www.verkeerscentrum.be</a></td>
<td>10085</td>
<td>Q, Ø</td>
</tr>
<tr>
<td>lwava55-en-url.trendmicro.com</td>
<td>8681</td>
<td>Q, Ø</td>
</tr>
<tr>
<td>d94e61c2.des1.onsp.f-secure.com</td>
<td>7146</td>
<td>Q, Ø</td>
</tr>
<tr>
<td>d94e61c2.des1.onsp.f-secure.com</td>
<td>5417</td>
<td>Q, Ø</td>
</tr>
<tr>
<td>d94e61c6.des1.onsp.f-secure.com</td>
<td>5001</td>
<td>Q, Ø</td>
</tr>
<tr>
<td>rldh.be</td>
<td>4575</td>
<td>Q, Ø</td>
</tr>
<tr>
<td>d94e61c8.des2.onsp.f-secure.com</td>
<td>3414</td>
<td>Q, Ø</td>
</tr>
<tr>
<td>d94e61c6.des1.onsp.f-secure.com</td>
<td>2578</td>
<td>Q, Ø</td>
</tr>
<tr>
<td>d94e61c8.des2.onsp.f-secure.com</td>
<td>2347</td>
<td>Q, Ø</td>
</tr>
<tr>
<td><a href="http://www.buienrader.nl">www.buienrader.nl</a></td>
<td>2178</td>
<td>Q, Ø</td>
</tr>
<tr>
<td>Other values</td>
<td>94317</td>
<td>Q, Ø</td>
</tr>
</tbody>
</table>

**COUNTRIES**

<table>
<thead>
<tr>
<th>Term</th>
<th>Count</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>Germany</td>
<td>32264</td>
<td>Q, Ø</td>
</tr>
<tr>
<td>United States</td>
<td>24233</td>
<td>Q, Ø</td>
</tr>
<tr>
<td>Belgium</td>
<td>18720</td>
<td>Q, Ø</td>
</tr>
<tr>
<td>Europe</td>
<td>10896</td>
<td>Q, Ø</td>
</tr>
<tr>
<td>France</td>
<td>9306</td>
<td>Q, Ø</td>
</tr>
<tr>
<td>Ireland</td>
<td>4794</td>
<td>Q, Ø</td>
</tr>
<tr>
<td>Netherlands</td>
<td>4666</td>
<td>Q, Ø</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>2977</td>
<td>Q, Ø</td>
</tr>
<tr>
<td>Sweden</td>
<td>657</td>
<td>Q, Ø</td>
</tr>
<tr>
<td>Denmark</td>
<td>171</td>
<td>Q, Ø</td>
</tr>
<tr>
<td>Other values</td>
<td>739</td>
<td>Q, Ø</td>
</tr>
</tbody>
</table>
How does it work?
Trick for all = ELK

- Elasticsearch
- Logstash
- Kibana

- Index as much as you want

- No limit on volume, speed or position-of-the-moon-licensing

- Open Source, Free to use, commercial support
Inputs

.inputs & codecs

- collectd, drupal_dblog, elasticsearch, eventlog, exec, file, ganglia, gelf, gemfire, generator, graphite, heroku, imap, invalid_input, irc, jmx, log4j, lumberjack, pipe, puppet_facter, rabbitmq, rackspace, redis, relp, s3, snmptrap, sqlite, sqs, stdin, stomp, syslog, tcp, twitter, udp, unix, varnishlog, websocket, wmi, xmpp, zenoss, zeromq

- cloudtrail, collectd, compress_spooler, dots, edn, edn_lines, fluent, graphite, json, json_lines, json_spooler, line, msgpack, multiline, netflow, noop, oldlogstashjson, plain, rubydebug, spool

Outputs

Filters
I usually don’t use “file” as input
- Keeps a reference to the position in the file

TCP socket is the easiest for me

ncat log01.lab.internal 18001 < logfile.log
Outputs

- Inputs & codecs
- Outputs
  - boundary, circonus, cloudwatch, csv, datadog, datadog_metrics, **elasticsearch**, elasticsearch_http, elasticsearch_river, email, exec, **file**, ganglia, gelf, gemfire, google_bigquery, google_cloud_storage, graphite, graphtastic, hipchat, http, irc, jira, juggernaut, librato, loggly, lumberjack, metriccatcher, mongodb, nagios, nagios_nsca, null, opentsdb, pagerduty, pipe, rabbitmq, rackspace, redis, redmine, riak, riemann, s3, sns, solr_http, sqs, statsd, stdout, stomp, syslog, tcp, udp, websocket, xmpp, zabbix, zeromq
- Filters
Output Example

```ruby
output {
  if [type] == "mactime" {
    elasticsearch {
      index => "logstash-mactime"
      host => localhost
    }
  }
}

# Configuration to save the json output to files.
# The filename and directory depends on the type, and date
# Make sure the directory is writable by logstash !!
# Subdirs are automagically created

output {
  file {
    path => "/mnt/data/jsonarchive/%{[type]}/%{[type]}-%{+YYYY.MM.dd}.gz"
    gzip => true
  }
}
```
Filters

- Inputs & codecs
- Outputs
- Filters
  - advisor, alter, anonymize, checksum, cidr, cipher, clone, collate, csv, date, dns, drop, elapsed, elasticsearch, environment, extractnumbers, fingerprint, gelfify, geoip, grep, grok, grokdiscovery, i18n, json, json_encode, kv, metaevent, metrics, multiline, mutate, noop, prune, punct, railsparallelrequest, range, ruby, sleep, split, sumnumbers, syslog_pri, throttle, translate, unique, urlencode, useragent, uuid, wms, wmts, xml, zeromq
filter {
  if [type] == "mactime" {
    csv {
      separator => "",
      columns => ["timestamp", "size", "macb", "perms", "uid", "gid", "inode", "path"]
    }
    date {
      match => ["timestamp", "EEE MMM dd YYYY HH:mm:ss"]
    }
    # extract deleted info (deleted) and (deleted-realloc)
    if ("(deleted" in [path]) {
      mutate {
        add_tag => ["deleted"]
        gsub => ["path", "\(deleted.*\$", ""]
    }
    # extract macb info
    if ("m" in [macb]) { mutate { add_tag => ["modified"] } }
    if ("a" in [macb]) { mutate { add_tag => ["accessed"] } }
    if ("c" in [macb]) { mutate { add_tag => ["changed"] } }
    if ("b" in [macb]) { mutate { add_tag => ["birth"] } }
    # extract file extension
    grok {
      match => ["path", "(?:<filename>[^/]+?\$"]}
  }
}
# extract macb info
if ("m" in [macb]) { mutate { add_tag => ["modified"] } }
if ("a" in [macb]) { mutate { add_tag => ["accessed"] } }
if ("c" in [macb]) { mutate { add_tag => ["changed"] } }
if ("b" in [macb]) { mutate { add_tag => ["birth"] } }

# extract file extension
fork {
  match => ["path", "(?<filename>[^/]+)?\$"
}
fork {
  match => ["filename", "\.(?<ext>[^./]+)\$"
}
mutable {
  lowercase => ["ext"
  remove_field => ["message", "perms", "uid", "gid"
}
Grok

- Named **regular expressions** to match patterns/extract data.
- Logstash ships with lots of patterns!
  [https://github.com/elasticsearch/logstash/tree/master/patterns](https://github.com/elasticsearch/logstash/tree/master/patterns)

```
55.3.244.1 GET /index.html 15824 0.043

grok {
  match => [ "message", "{IP:client} {WORD:method} {URIPATHPARAM:request} {NUMBER:bytes} {NUMBER:duration}" ]
}
```

- Test app: [http://grokdebug.herokuapp.com](http://grokdebug.herokuapp.com)
Testing complex Groks

c:/documents and settings/file with spaces.exe.jpg

(((\.(?<ext>[^./]+))?))?$

- [ ] Add custom patterns
- [ ] Keep Empty Captures
- [x] Named Captures Only
- [ ] Singles

```json
{
  " ext": [
    [
      "jpg"
    ]
  ]
}
```
Data Enrichment with Filters

- Extract fields: csv, grok, kv
- Extract date
- Modify using mutate
- **Enrich with**
  - Geoip
  - User-agent
  - Urldecode
  - Translate
  - …
"geoip": {
    "ip": "217.110.97.200",
    "country_code2": "DE",
    "country_code3": "DEU",
    "country_name": "Germany",
    "continent_code": "EU",
    "latitude": 51,
    "longitude": 9,
    "timezone": "Europe/Berlin",
    "location": [9, 51]
},
<table>
<thead>
<tr>
<th>Term</th>
<th>Count</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>18998197</td>
<td></td>
</tr>
<tr>
<td>Mozilla/5.0 (compatible; MSIE 10.0; Windows NT 6.1; Trident/6.0)</td>
<td>11775197</td>
<td></td>
</tr>
<tr>
<td>Shockwave Flash</td>
<td>9537116</td>
<td></td>
</tr>
<tr>
<td>Mozilla/5.0 (compatible; MSIE 9.0; Windows NT 6.1; WOW64; Trident/5.0)</td>
<td>8540697</td>
<td></td>
</tr>
<tr>
<td>Mozilla/5.0 (Windows NT 6.1; WOW64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/35.0.1916.153 Safari/537.36</td>
<td>6118773</td>
<td></td>
</tr>
<tr>
<td>Mozilla/5.0 (compatible; MSIE 9.0; Windows NT 6.1; Trident/5.0)</td>
<td>5208000</td>
<td></td>
</tr>
<tr>
<td>Mozilla/5.0 (compatible; MSIE 10.0; Windows NT 6.1; WOW64; Trident/6.0)</td>
<td>4557473</td>
<td></td>
</tr>
<tr>
<td>Mozilla/5.0 (Windows NT 6.1; WOW64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/35.0.1916.114 Safari/537.36</td>
<td>2988231</td>
<td></td>
</tr>
<tr>
<td>Apple-iPhone4C1/1104.201</td>
<td>2001403</td>
<td></td>
</tr>
<tr>
<td>Mozilla/5.0 (Windows NT 6.1) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/35.0.1916.153 Safari/537.36</td>
<td>1662367</td>
<td></td>
</tr>
<tr>
<td>Missing field</td>
<td>203804</td>
<td></td>
</tr>
<tr>
<td>Other values</td>
<td>2311378</td>
<td></td>
</tr>
</tbody>
</table>
User-Agent

Browser:
- Chrome: 17%
- IE: 40%
- Other: 34%

Operating Systems:
- Windows: 57%
- Other: 39%

- Windows 7: 53902029
- Windows Vista: 25400990
- Windows XP: 578523
- Windows 2000: 482992
- Android: 404663
- Linux: 83310
- Windows: 81229
- Windows 8: 1931
- Other 4.2.1: 111
- Other values: 125
translate {
    field => "tk_category"
    destination => "category"
    override => true
    dictionary => [
        "1", "Adult/Mature Content",
        "3", "Pornography",
        "4", "Sex Education",
        "5", "Intimate Apparel/Swimsuit",
        "6", "Nudity",
        "8", "Alcohol/Tobacco",
        "9", "Illegal/Questionable",
        "10", "Tasteless",
        "11", "Gambling",
        "14", "Violence/Hate/Racism",
        "15", "Weapons",
        "16", "Abortion",
    ]
}
<table>
<thead>
<tr>
<th>Term</th>
<th>Count</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>Computers/Internet</td>
<td>8405476</td>
<td></td>
</tr>
<tr>
<td>News/Media</td>
<td>4660890</td>
<td></td>
</tr>
<tr>
<td>Search Engines/Portals</td>
<td>2895718</td>
<td></td>
</tr>
<tr>
<td>Web Advertisement</td>
<td>2322018</td>
<td></td>
</tr>
<tr>
<td>Business/Economy</td>
<td>2078554</td>
<td></td>
</tr>
<tr>
<td>Internet Infrastructure</td>
<td>1713403</td>
<td></td>
</tr>
<tr>
<td>Social Networking</td>
<td>1601947</td>
<td></td>
</tr>
<tr>
<td>Untested</td>
<td>1150511</td>
<td></td>
</tr>
</tbody>
</table>
Ruby as last resort

```ruby
# convert the int_IP to a human readable form
ruby {
  code => "
require 'ipaddr'
begins
  client_ip = event['clientip'].to_i
  if client_ip < 0
    client_ip = 4294967296 + client_ip
  end
  event['c_ip'] = IPAddr.new(client_ip, Socket::AF_INET).to_s
rescue
  event['c_ip'] = 'conversionerror'
end
"
}

* There might be a better way to do this, but ruby and I are not really friends yet
Data Enrichment with Filters

- Extract fields: csv, grok, kv
- Extract date
- Modify using mutate
- Enrich with
  - Geoip
  - User-agent
  - Urldecode
  - Translate
  - ...

[Image of a personified tree with a mustache]
Trick for all = ELK

- Elasticsearch  Logstash  Kibana
- Index as much as you want
- **No limit** on volume, speed or season-licensing
- Open Source, Free to use, commercial support
Elasticsearch

- Wikipedia: Elasticsearch is a search server based on Lucene. It provides a distributed, multitenant-capable full-text search engine with a RESTful web interface and schema-free JSON documents. Elasticsearch is developed in Java and is released as open source under the terms of the Apache License.

- Very very fast

- Adding an node = easier than extremely easy
Elasticsearch

- Be cautious
- No security by default
- Auto-discovery, auto-distribution if other node is present
- Elastic HQ plugin
  - cd /usr/share/elasticsearch/bin
  - ./plugin -install royrusso/elasticsearch-HQ
Trick for all = ELK

- Elasticsearch  Logstash  Kibana
- Index as much as you want
- **No limit** on volume, speed or horoscope-licensing
- Open Source, Free to use, commercial support
Kibana

- Fancy GUI
- Extremely easy to build up a dashboard
- Gives good overview over data
- Powerful, but limited in capability
- For more: write a python script or use REST API
DO NOT PRESS THIS BUTTON
Apache Lucene Search syntax

- title:foo    title:"foo bar"
- title:"foo bar" AND body:"quick fox"
- (title:"foo bar" AND body:"quick fox") OR title:fox
- title:foo -title:bar
- title:foo*bar
- time_taken:[10000 TO 999999999]

Load dashboards
### USER AGENT

<table>
<thead>
<tr>
<th>Term</th>
<th>Count</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>FSORSP/1.1.25.2280 (Win32)</td>
<td>5616</td>
<td>Q</td>
</tr>
<tr>
<td>Mozilla/5.0 (Windows NT 6.1) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/30.0.1599.69 Safari/537.36</td>
<td>2032</td>
<td>Q</td>
</tr>
<tr>
<td>Mozilla/5.0 (compatible; MSIE 9.0; Windows NT 6.1; Trident/5.0)</td>
<td>1489</td>
<td>Q</td>
</tr>
<tr>
<td>Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36 (KHTML, like Gecko) Ubuntu Chromium/31.0.1650.63 Chrome/31.0.1650.63 Safari/537.36</td>
<td>714</td>
<td>Q</td>
</tr>
<tr>
<td>Mozilla/4.0 (compatible; MSIE 5.5; Win32)</td>
<td>634</td>
<td>Q</td>
</tr>
</tbody>
</table>

### FILTERING

- **time** must
  - field: `@timestamp`
  - from: now-5m
  - to: now

- **terms** must
  - field: `tk_operation.raw`
  - value: POST
Performance
Performance goals

- Focus Incident Handling and Forensics
- Max speed of indexing
- Max speed of searching
- During indexation search may be slow
- No need for redundancy
- So don’t use this advice for operations-live-production
Performance Logstash

- **Memory setting:** (`/etc/default/elasticsearch`)
  - `LS_HEAP_SIZE="500m"`

- **Command line flag:**
  - `-w` or `--filterworkers AMOUNT_OF_CORES` (default: 1)

- **Each extra filter slows it down**
  - Grok aka regex = slow
    - Prefer csv, kv
    - Use the least possible wildcards (`*` or `+`)
  - Geoip = slow but very practical
  - User-agent = slow, often practical
Performance Elasticsearch

- **Memory setting** (/etc/default/elasticsearch)
  - `ES_HEAP_SIZE=12g` => set to half of RAM (max 32 GB)

- **Disable redundancy** (/etc/elasticsearch/elasticsearch.yml)
  - `index.number_of_replicas: 0`

- **Shards for number of nodes** (/etc/elasticsearch/elasticsearch.yml)
  - `index.number_of_shards: 1`

- **Increase memory buffer for search**
  - `indices.memory.index_buffer_size: 50%`
Perf. Elasticsearch Indexes

- Open Index = memory usage + disk usage
  Closed Index = disk usage, so close index when not needed

- Per case new indexes
  Similar logs in the same index, but use a field "host" to differentiate investigations
  - system timelines: logstash-%{[case]}-%{[type]}
  - mail logs: logstash-%{[case]}-%{[type]}-%{+YYYY.MM}
  - proxy logs: logstash-%{[case]}-%{[type]}-%{+YYYY.MM.dd}

- curl -XPOST 'localhost:9200/logstash-${case}*/_close'
  curl -XPOST 'localhost:9200/logstash-${case}*/_open'
Performance Kibana

- Each block/graph is extra search
- So 10 graphs equals 10 simultaneous searches

1. First select small date/time window
2. Test your search on small data set
3. Add filters
4. Zoom out on date/time
5. Dig deeper
Keep in mind

- Logstash is *(relatively)* SLOW
- Finished? **Close** the index, do NOT **delete** it
- **Or save JSON to files** *(output plugin Logstash)*, **re-index them** later
- **Node++ = Speed++**
Forensic analysis
Plaso

- Plaso = the new log2timeline and more
- log2timeline.py win7-64-nfury-10.3.58.6.dump /path/to/disk/image
- psort.py -o elastic win7-64-nfury-10.3.58.6.dump
ELK-forensics

- [https://github.com/cvandeplas/ELK-forensics](https://github.com/cvandeplas/ELK-forensics)
- Logstash configs
- Kibana dashboards
- Mactime, Log2timeline csv, BlueCoat, Mail IMSS, IWSVA, IIS
- More to come
Other interesting projects using Elasticsearch

- **Moloch** – Open Source large scale IPv4 full PCAP capturing, indexing and database system. [https://github.com/aol/moloch](https://github.com/aol/moloch)

- **Mozdef** – PoC – automate IH process and facilitate real-time activities - [https://github.com/jeffbryner/MozDef](https://github.com/jeffbryner/MozDef)

- **Suricata** – Exports data in EVE format (JSON). Great to visualize malware activity from sandbox
### Alerts

<table>
<thead>
<tr>
<th>Timestamp</th>
<th>src_ip</th>
<th>alertSignature</th>
<th>alertCategory</th>
</tr>
</thead>
<tbody>
<tr>
<td>2014-03-27T14:39:59.333Z</td>
<td>192.168.1.13</td>
<td>MISP e843 Outgoing HTTP Domain: lemondebe.org</td>
<td>A Network Trojan was detected</td>
</tr>
<tr>
<td>2014-03-27T14:39:38.450Z</td>
<td>192.168.1.13</td>
<td>MISP e843 Outgoing HTTP Domain: lemondebe.org</td>
<td>A Network Trojan was detected</td>
</tr>
<tr>
<td>2014-03-27T14:39:11.221Z</td>
<td>192.168.1.13</td>
<td>MISP e843 Outgoing HTTP Domain: lemondebe.org</td>
<td>A Network Trojan was detected</td>
</tr>
<tr>
<td>2014-03-27T14:38:52.363Z</td>
<td>192.168.1.13</td>
<td>MISP e843 Outgoing HTTP Domain: lemondebe.org</td>
<td>A Network Trojan was detected</td>
</tr>
<tr>
<td>2014-03-27T14:38:51.745Z</td>
<td>192.168.1.13</td>
<td>MISP e843 Outgoing HTTP Domain: lemondebe.org</td>
<td>A Network Trojan was detected</td>
</tr>
<tr>
<td>2014-03-27T14:38:42.353Z</td>
<td>192.168.1.13</td>
<td>MISP e843 Outgoing HTTP Domain: lemondebe.org</td>
<td>A Network Trojan was detected</td>
</tr>
<tr>
<td>2014-03-27T14:38:41.327Z</td>
<td>192.168.1.13</td>
<td>MISP e843 Outgoing HTTP Domain: lemondebe.org</td>
<td>A Network Trojan was detected</td>
</tr>
<tr>
<td>2014-03-27T14:23:28.599Z</td>
<td>192.168.1.13</td>
<td>MISP e843 Outgoing HTTP Domain: lemondebe.org</td>
<td>A Network Trojan was detected</td>
</tr>
<tr>
<td>2014-03-27T14:23:21.386Z</td>
<td>192.168.1.13</td>
<td>MISP e843 Domain: lemondebe.org</td>
<td>A Network Trojan was detected</td>
</tr>
<tr>
<td>2014-03-27T14:23:20.372Z</td>
<td>192.168.1.13</td>
<td>MISP e843 Domain: lemondebe.org</td>
<td>A Network Trojan was detected</td>
</tr>
<tr>
<td>2014-03-27T14:23:19.358Z</td>
<td>192.168.1.13</td>
<td>MISP e843 Domain: lemondebe.org</td>
<td>A Network Trojan was detected</td>
</tr>
</tbody>
</table>

### Alert Types

<table>
<thead>
<tr>
<th>Term</th>
<th>Count</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>MISP e843 Outgoing HTTP Domain: lemondebe.org</td>
<td>8</td>
<td>-</td>
</tr>
<tr>
<td>MISP e843 Domain: lemondebe.org</td>
<td>3</td>
<td>-</td>
</tr>
<tr>
<td>Other values</td>
<td>0</td>
<td>-</td>
</tr>
</tbody>
</table>
Places to be?

- https://github.com/cvandeplas/ELK-forensics
- http://www.elasticsearch.org/overview/elkdownloads/
- http://logstash.net/
- https://groups.google.com/forum/#!forum/logstash-users