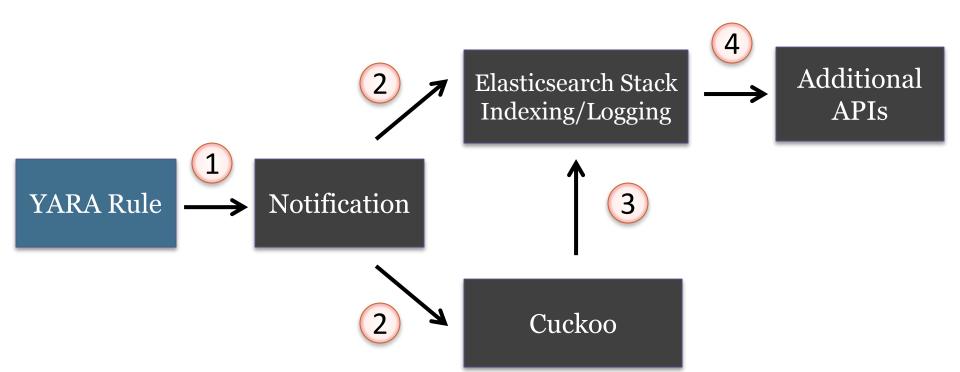
Tracking Threat Actors through YARA Rules and Virus Total

Kevin Perlow- Booz Allen Hamilton Allen Swackhamer- Target Corporation

Automation and Collection Workflow



YARA Rules - Purpose

- Track Campaigns
 - Strings
 - Static Indicators
 - Compilation Artifacts
 - Opcode signatures
- Categorize Malware
 - Family / Variants

```
rule Heist
{
    meta:
        md5_1 = "0ebf68bb15c2e36508cf3f46d32cf2e3"
        md5_1 = "c49b7a968lad387922f14a1601652e5b"
        md5_1 = "e350e67c526ffab0c97c5e04a6e9f12"
        md5_1 = "bb552a4bdc573566da897a651b9041e6"
        date = "11/4/2015"
        author = "Kevin Perlow"

strings:

$String = "Coded By - (Picasso)"

condition:
        any of them
}
```

Basic YARA rule for tracking a crimeware crypter

YARA Rules- Examples

```
rule russian ransomware Sept22
   meta:
       description = "Found on http://abrazivstroy.ru/wp-content/uploads/2015/01/ Tracking full campaign"
        date = "09/22/2015"
        author = "Kevin Perlow"
       Note = "The path string will also allow it to pick up the infostealer from the same source. Comment it out if you on
        Note2 = "The smoothtiny string is the best way to catch the fourth piece which didn't contain the same unique string
    strings:
       $Profile1 = "ame View Xerrter Fertui's profile. Viadeo helps professionals like Xerrter Fertui boost their career" w
        $Profile2 = "View Xerrter Fertui's profile. Viadeo helps professionals like Xerrter Fertui boost their career" wide
        $Path1 = "\\Gertiopertores\\Certiop.vbp" wide
        $Process = "Smoothtiny"
        $Compression = "!Thiv qrobpam%cgnnms bg'rsm\"in AIS'hlae."
        $Path2 = "AKT -21092015-PowerPoint.exe"
        $Path3 = "\Documents\\chm\\AKT -21092015-PowerPoint.exe"
        $Path4 = "C:\\Users\\A90B~1\\AppData\\Local\\Temp\\AKT -21092015-PowerPoint.exe"
        $Path5 = "C:\\Users\\A908~1\\AppData\\Local\\Temp\\AKT -21092015-PowerPoint.exe" wide
        $Path6 = "AKT -21092015-PowerPoint.exe9" wide
        $Path7 = "\Documents\\chm\\AKT -21092015-PowerPoint.exe" wide
        $Path8 = "C:\\Users\\836D~1\\AppData\\Local\\Temp\\PEWER POINT PRESENTATION.exe" wide
        $Path9 = "PEWER POINT PRESENTATION.exe=" wide
        $Path10 = "\\Documents\\PEWER POINT PRESENTATION.exe"
        $Profile3 = "s Ainda precisam da uma melhorada nos pistols, pois a maioria dos jogos ja come" wide
    condition:
        any of them
```

YARA Rules- Examples

```
rule backoff opcode{
   meta:
      author = "Swackhamer"
      md5 = "01F0D20A1A32E535B950428F5B5D6E72"
  strings:
      // MD5: 01F0D20A1A32E535B950428F5B5D6E72
      // Function: 404344 cc validation
      $cc validation = { 3C 5E ?? ?? ?? 74 ?? 3C 3D 0F ?? ?? ?? ?? ?? ?? ?? ?? 83 ?? ??
      ?? ?? E8 ?? ?? ?? ?? 83 ?? ?? ?? ?? 75 ?? 8B ?? ?? EB ?? 8B ?? ?? ?? ?? ?? ?? ?? ??
      A1 ?? ?? ?? ?? ?? ?? E8 }
  condition:
      any of them
```

YARA Rules- Examples

```
rule early october2015 vawtrak dropper{
    meta:
        author = "Kevin Perlow"
        SHA256 = "3d1e7e54db786c6aef572d1ef57ad1c26413aacbf2fd91eb700d469c550dd4df"
        SHA256 = "3ffbe191d9326f97db4ffaf6b294c166397bf1c77d28e2ab44d41fca511ce55b"
   strings:
        $VBA = { 00 41 74 74 72 69 62 75 74 00 } //doc contains VBA
        $rtf = { 2E 72 74 66 } //rtf in hex, will appear if in macro unobfuscated
        $exe = { 2E 65 78 65 } //exe in hex
        $string1 = "TEMP$ 4"
        $string2 = /[0-9][0-9]\.rtf/
        $string3 = /[a-zA-Z0-9][a-zA-Z0-9][a-zA-Z0-9]\.exe/
        $a = {d0 cf 11 e0}
        $string4 = /C:\\Aaa\\exe\\[0-9A-Za-z]*\.exe/
        $string5 = /C:\\Users\\M\\AppData\\Local\\Temp\\[0-9A-Za-z]*\.exe/
        $string6 = "X:\\multiplexing\\limitations\\electr.pdb"
        $Dvreza = "C:\\Users\\Em\\AppData\\Local\\Temp\\w12.exe"
        $RSA = "This file is protected with RSA kev." nocase
   condition:
        $a at 0 and $VBA and (($rtf and $exe) or 2 of ($string1,$string2,$string3) or 2 of
```

```
Find
                                    Find
 lu1}o
 [q1]'
 +42N
 'tbvd
 &4zkVo
 ='yl
 Normal dotm
 C:\Art\pa3.exe
 C:\Users\Em\AppData\Local\Temp\pa3.exe
 !This program cannot be run in DOS mode.
 leM.
 lcl\Rich
  text
  irdata.
 @.data
 .rsrc
 \$I^:
 D$PP
 D$,P
```

YARA Rules- Case Study

```
function dhV(Zyr) {var 112crap = "s"; return "" + Zyr + ""; }; function
q2(x0){var 112crap = "s"; return "" + x0 + "";};var Ggr = "o\x73e",NNy =
"cl";function H5(u4){var 112crap = "s"; return "" + u4 + "";};function
eA(CB){var 112crap = "s"; return "" + CB + "";};var tq = "Fi\x6c\x65",R2 =
"veTo";var CAL = "Sa";function Eui(DDB){var 112crap = "s"; return "" + DDB
+ "";};function V0(h3){var 112crap = "s"; return "" + h3 + "";};var mGY =
"io\x6e",P = "p\x6fsit";function oIZ(FWx){var 112crap = "s"; return "" +
FWx + "";};function H4(CjA){var 112crap = "s"; return "" + CjA + "";};var
gMy = "dy", I2 = "\x42\x6f"; var T3 = "nse", g2 = "R\x65sp\x6f"; function
tb(B1){var 112crap = "s"; return "" + B1 + "";};function H3(ti){var
112crap = "s"; return "" + ti + ""; }; function wZp(FXr) {var 112crap = "s";
return "" + FXr + ""; }; var vv = "gtVXRtZ", Mc = "e"; var xgK =
yy["c"+"\x68"+"arAt"](1), Vg = "\x69"; var q1 = "\x77r"; function PbP(YPw) {var
112crap = "s"; return "" + YPw + "";};function gj(w4){var 112crap = "s";
return "" + w4 + "";};var Z1 = "vp\x65",LAH = "t";var Uxp = "n",I1 =
""; }; function kzA(X3) {var 112crap = "s"; return "" + X3 + ""; }; function
rXn(h2){var 112cran = "s": return "" + h2 + "":}:var X2 = "am".a0 =
```

```
function dhV(Zyr) {
    var 112crap = "s";
    return "" + Zvr + "";
function a2(x0) {
    var 112crap = "s";
   return "" + x0 + "";
var Ggr = "ose",
    NNy = "cl";
function H5(u4) {
   var _112crap = "s";
    return "" + u4 + "":
function eA(CB) {
   var 112crap = "s";
    return "" + CB + "":
var tg = "File",
    R2 = "veTo":
var CAL = "Sa";
function Eui(DDB) {
    var 112crap = "s";
    return "" + DDB + "":
function V0(h3) {
    var 112crap = "s";
    return "" + h3 + "":
```

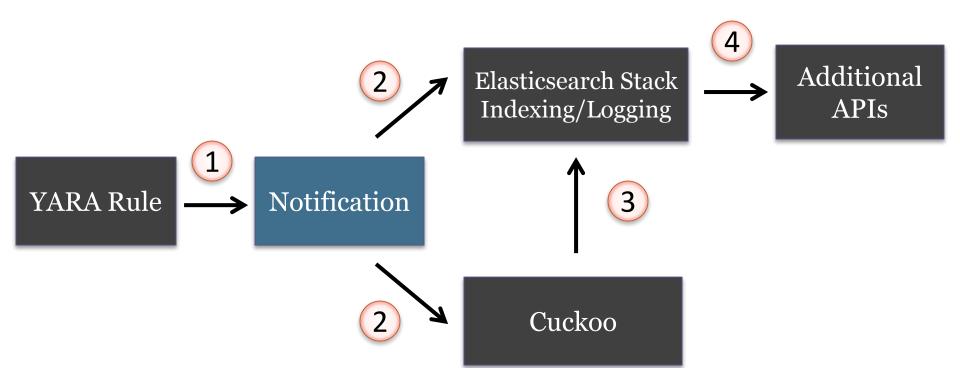
YARA Rules- Case Study

```
function Y() {
   return ((x == yio) && (x == true)) ? true : false;
if (x && Y() && yio) {
    function jJ() {
       return O[Zjs + O4(Mii) + ib + h0(OOC) + j0(Tj) + SDP + LTI(IO) + vwn +
       vCs + TA(q) + DQ + oh(ncw) (1z(zAB) + u0(mO)) + RA(u1) + T1(xq) + q0(Jd)
       + MK + B0:
   WScript.echo(jJ());
   vf = Ji();
   D = WScript[WW + CRa(SPA) + F(XO) + bU(RO) + B(NQ) + OO](vf);
   var mRh = true:
   while (mRh) {
        trv {
            D[H1(u2)](Q0(RVC), Ich(Q1) + r1(zwI) + ca(K) + Y1(u0) + y(Z0) + MRa(
            hf) + v2(w1) + Pi + d0 + jX(uv), false);
            D[10(14)]();
            while (D[aQY(TAd) + w2 + M(G) + eP + Rq] < 4) {
                WScript[e(Ce)](On() * 10)
            1:
            mRh = false:
        } catch (i) {
           mRh = (ix(ina) + T2, b1(MM) + E(q), A2 + w3(Q3) + YBR(m1), true);
    function sS(b) {
```

Left: the executable path being built.

Top right: Similar sample- the GET request being made inside a try/catch function

Automation and Collection Workflow



Notifications API

- VirusTotal or proprietary database
 - SMTP notifications
 - Pull via Python IMAP library
 - JSON notifications
 - Pull from REST API via Python requests library
 - Delete the alerts from VT after you process them
- Index and Parse into Elasticsearch

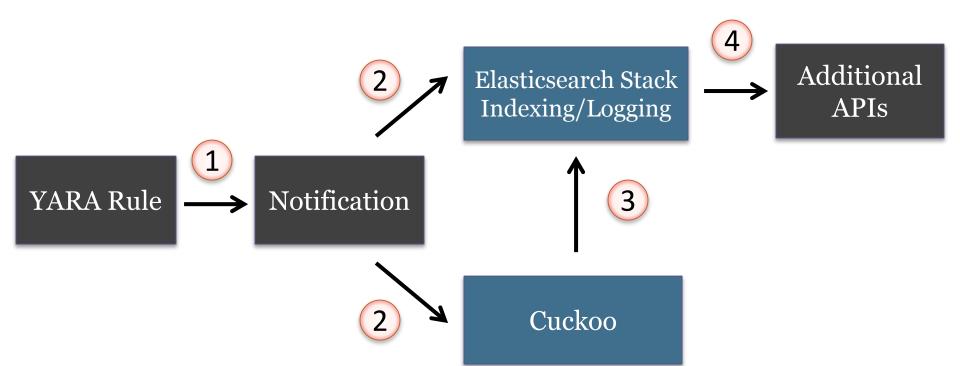
Sample VirusTotal Notification

```
"notifications" : [{
        "total" : 52.
        "first seen": "2015-11-06 16:43:58",
        "sha1": "87d94d18d44021bff2ab4de8093628c1576f8902".
        "scans" : {
           "Bkav" : null.
           "MicroWorld-eScan": "Gen: Variant. Zusv. 165602",
           "nProtect" : null,
           "CMC" : null.
            "CAT-OuickHeal" : "Backdoor.Bladabindi.AL3".
           "ALYac": "Gen: Variant. Zusv. 165602",
           "Malwarebytes" : null.
           "Zillya" : null,
        "ruleset name" : "rats",
        "sha256": "37f946601b35c5f3282a5ee97aadc0bbcf6128447e3a792ee0153eb1dcd95f71".
        "md5": "dba50c01771adb017180bb47319d2bf1".
        "date": "2015-11-06 18:30:33",
        "positives" : 26,
        "last seen": "2015-11-06 16:43:58",
        "size": 557056.
        "type": "Win32 EXE",
        "id": 5768439097196544.
        "match" : "00 30 00 2E 00 30 00 2E 00 31 00 00 09 35 00 35 .0...0...1...5.5\n00 35 00 32 00 00 0B *begin highlight*7C 00 27 00 7C 00 27 00 7C*end h
        .5.2...*begin highlight*|.'.|.'.|*end highlight*\n00 01 09 54 00 72 00 75 00 65 00 00 5B 53 00 6F ...T.r.u.e..[S.o\n00 43 00 48 00 45 00 43 00 48 00
        .C.H.E.C.K.S...1\n00 00 47 *begin highlight*6E 00 65 00 74 00 73 00 68 00 20 00 66*end highlight* ...G*begin highlight*n.e.t.s.h. .f*end highlight*\
       72 00 65 00 77 00 61 00 6C 00 20*end highlight* *begin highlight*.i.r.e.w.a.l.l. *end highlight*\n*begin highlight*00 61 00 64 00 64 00 64 00 20 00
        6F*end highlight* *begin highlight*.a.d.d. .a.l.l.o*end highlight*\n*begin highlight*00 77 00 65 00 64 00 70 00 72 00 6F 00 67 00 72*end highlight*
        *begin bighlight*.w.e.d.n.r.g.g.r*end bighlight*\n*begin bighlight*00 61 00 60 00*end bighlight* 20 00 22 00 00 07 22 00 20 00 22 *begin bighlight*
```

IOC Extraction and Logging

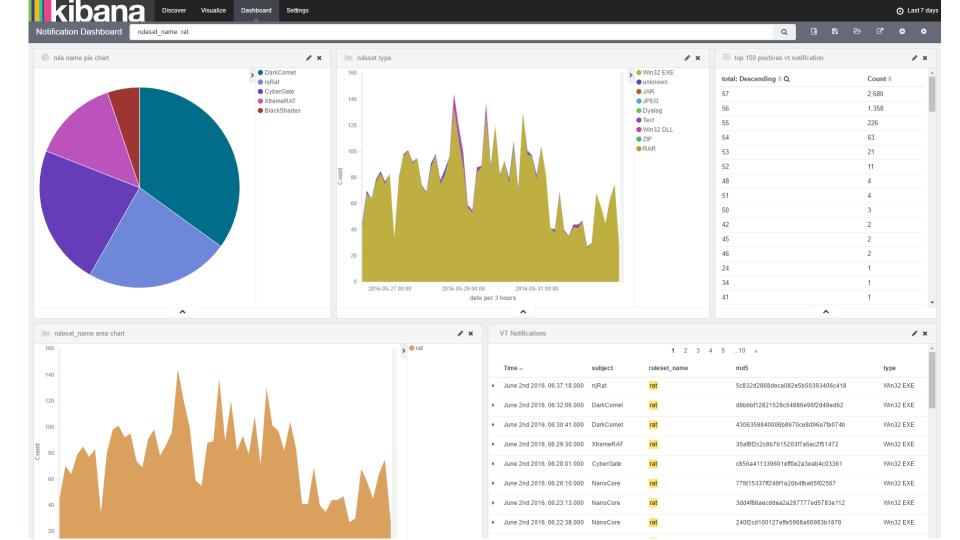
- Static Extraction
 - Configuration deobfuscation and parsing
 - Strings
 - Various obfuscation techniques (olevba)
 - FLOSS Automated deobfuscation of strings
- Dynamic Extraction
 - Sandbox
 - Network
 - File system

Automation and Collection Workflow



Elasticsearch Stack

- Elasticsearch & Kibana
 - Visualize Notification Trends
 - First Seen
 - Last Seen
 - Resubmissions
 - Export Data (Hash, Rule Name, Rule Set)
 - Pivot through additional API's
 - Export to CSV/JSON or other consumable formats



Cuckoo Sandbox

- Automated Submission
 - Push notifications to Cuckoo on ingest from VT
 - Output IOCs (Domains, Files, Mutexs, etc...) back to Elasticsearch
- Customizable
 - Custom Elasticsearch reporting module
- Popular Sandboxes
 - VirusTotal
 - Malwr
 - Hybrid Analysis

Cuckoo Elasticsearch Index Template

- Sets shard count to 1
- Compression to "best"
- Strings to "not_analyzed"
- task_id is an indexed field
- report_time is the date/time field

```
def apply template(self):
    cuckoo template = {
        "order": 0,
        "template": "cuckoo*",
        "settings": {
            "index": {
                 "number of shards": "1",
                "codec": "best compression",
                 "number of replicas": "1"
        "mappings":
             "cuckoo": {
                 "dynamic templates": [
                         "notanalyzed": {
                             "mapping": {
                                 "index": "not analyzed",
                                 "type": "string",
                                 "doc values": "True"
                             "match mapping type": "string",
                             "match": "*"
                 "properties": {
                     "report time": {
                         "format": "epoch second",
                         "type": "date"
                     "task id": {
                         "type": "long"
        "aliases": {}
    self.es.indices.put_template(name="cuckoo template", body=json.dumps(cuckoo_template)
```

Back to the Elasticsearch Stack

- Collect Cuckoo IOCs
 - Track by Domain, IP, Country
 - Files written to disk
 - Command line called
 - Normalization of A/V Data

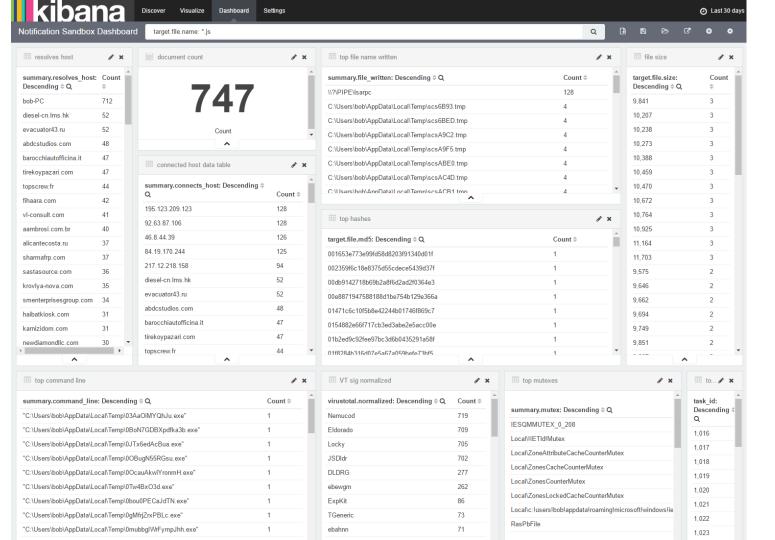
```
from elasticsearch import Elasticsearch
es = Elasticsearch(["localhost:9200"]) # default ES hostname and port
page = es.search(index="virustotal notifications", # Index specified here, you can use vildcards to select indexes
                doc type="notification", # the doc type is notification if empty will perform on all document types
                 size=100, # default size is 10
                 scroll='5m', # time to keep the scroll handle alive
                fields="md5", # fields to return. This will accelerate the search if you are requesting big documents
                 # sort="first seen:desc", # sort by field then asc / desc.this can be multiple fields comma separated
                q='subject: rockdownloader AND type: "Win32 EXE"') # lucene search query syntax
hashes = set() # where to place the hashes, use a set because some files may hit multiple times
sid = page[' scroll id']
scroll size = page['hits']['total']
while scroll size > 0:
   page = es.scroll(scroll id=sid, scroll="5m")
    sid = page[' scroll id']
    scroll size = len(page['hits']['hits'])
   hits = page['hits']['hits']
    for hit in hits:
       md5s = hit["fields"]["md5"]
        for md5 in md5s:
           hashes.add(md5) # add the hash list to the set
```

```
subject = "rockdownloader"
machines = ["cuckoo1", "cuckoo2", "cuckoo3"] # specify your Cuckoo guests
i = 0
for h in hashes:
    machine = machines[i % 3]
    data = get file(h) # get file is a function that returns a full file
    files = {'file': ("%s.js" % subject, data)}
    params = {"tags": subject,
              "package": "js",
              "options": "route=none",
              "machine": machine,
              "platform": "windows",
              "priority": 2.
              "timeout": 300,
              "custom": subject}
   print "Submitted hash %s" % h
    r = requests.post("https://api.cuckoo.com/tasks/create/file", files=files, data=params,
                      auth=HTTPBasicAuth(username, password))
```

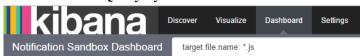
Cuckoo Summary

- Files
- Registry
- Mutex
- Directory
- Resolved Hosts
- Connected Hosts
- Command Line
- DLL Loaded
- WMI Query
- Target File Hash
- Target File Name
- Target File Type
- VirusTotal Signatures

```
Index target information, the behavioral summary, and
  VirusTotal results.
self.do index({
     "target": results.get("target"),
     "summary": results.get("behavior", {}).get("summary"),
     "virustotal": results.get("virustotal"),
                                                              name 🖨
name 🖨
                           name 🔷
                                                              target.file.sha256
summary.file_created
                           summary.regkey written
                                                              target.file.type
summary.file read
                           summary.regkey read
                                                              target.file.sha1
summary.file failed
                           summary.regkey opened
summary.file written
                                                              target.category
                           summary.regkey_deleted
summary.file copied
                                                              target.file.md5
summary.file opened
                                                              target.file.sha512
summary.file exists
                                                              target.file.path
summary.file_deleted
                                                              target.file.size
summary.file recreated
                                                              target.file.crc32
summary.file moved
                                                              target.file.ssdeep
                                                              target.file.name
                                                              target.url
                                                              target.file.urls
```



Lucene Search Query Syntax



List of hosts and files written aggregated with count



top file name written	
summary.file_written: Descending \(\phi \) Q	Count \$
\\?\PIPE\lsarpc	128
$C: \label{local-temp} Local \ Temp\ Scs 6B93.tmp$	4
$C: \verb \Users\bob \AppData\Local\Temp\scs6BED.tmp $	4
$C: \verb \Users\bob \AppData\Local\Temp\scsA9C2.tmp $	4
$C: \verb \Users\bob \AppData\Local\Temp\scsA9F5.tmp $	4
$C: \label{local-temp} \\ C: \$	4
C:\Users\bob\AppData\Local\Temp\scsAC4D.tmp	4

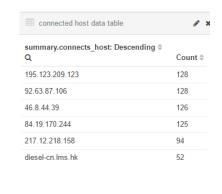
Normalized AV signatures from VT

■ VT sig normalized	ø
virustotal.normalized: Descending \$ Q	Count \$
Nemucod	719
Eldorado	709
Locky	705
JSDldr	702
DLDRG	277
ebewgm	262
ExpKit	86
TGeneric	73

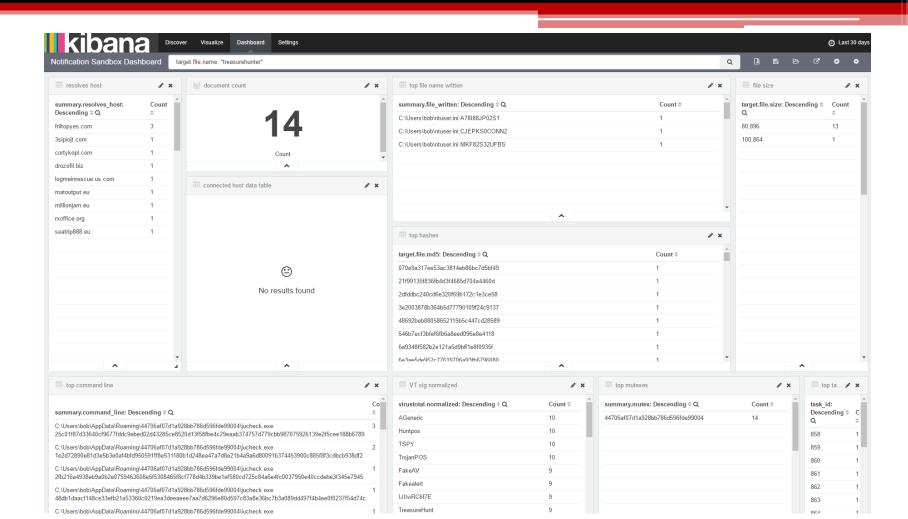
Command line called by malware

top command line	d
summary.command_line: Descending Q	Count \$
"C:\Users\bob\AppData\Local\Temp\03AaOiMYQhJu.exe"	1
$"C:\label{local} $$ "C:\local\arrowvertex" $$ "C:\local\arrowvertex" $$$	1
$"C:\Users\bob\AppData\Local\Temp\0JTx6edAcBua.exe"$	1
$"C:\local{Temp} \label{temp} Local\ Temp\\ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \$	1
$"C:\local{Temp} \label{temp} \label{temp} \local{Temp} \local \local{Temp} \local{Temp} \local \local \local}$	1
$\label{local-local-local} $$ "C:\Users\bob\AppData\Local\Temp\0Tw4BxO3d.exe" $$$	1
$"C:\local{Temp} \label{temp} \label{temp} \local{Temp} \local\local{Temp} \local\local\local\local}$	1
"C:local-	1
$"C:\lbob\AppData\Local\Temp\0mubbgl\WrFympJhh.exe"$	1

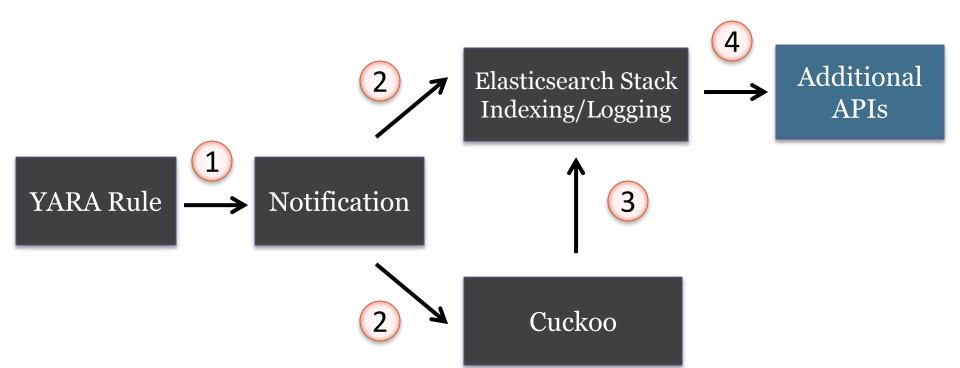
File size and hosts connected to



iii file size	/:
target.file.size: Descending Q	Count
9,841	3
10,207	3
10,238	3
10,273	3
10,388	3
10,459	3
10,470	3
10,672	3
10,764	3
10,925	3
11,164	3
11,703	3
9,575	2
9,646	2
9,662	2
9,694	2
9,749	2
9,851	2



Automation and Collection Workflow



Additional APIs

- VirusTotal
 - Parent objects
 - Emails
 - Zip Files
 - Network Infrastructure
- CentralOps
 - Whois
 - Physical Address
- PassiveTotal
 - PassiveDNS
 - Historical Records

```
26.122.41 detected download samples: 8bb95c8ec41def19
26.122.41 detected communicating samples: 62b6150a544
26.122.41 detected urls: http://blyoudo.ru/(2016-05-3
251.11.125 detected download samples: 88904ca8c0a1c4d
251.11.125 detected communicating samples: c32e69b85d
251.11.125 detected urls: http://alicantecosta.ru/kd9
132.100.220 detected urls: http://topscrew.fr/nsh38cj
46.52.112 detected download samples: a8a284f377cb9f21
46.52.112 detected urls: http://tirekoypazari.com/lsd
237.15.128 detected download samples: 74d6147825ab532
237.15.128 detected communicating samples: fde83a4bbe
237.15.128 detected urls: http://maapro.it/nvlauty.ht
46.52.112 detected download samples: a8a284f377cb9f21
46.52.112 detected urls: http://tirekoypazari.com/lsd
185.27.101 detected download samples: 88904ca8c0a1c4d
185.27.101 detected urls: http://adelina.se/1/(2016-0
28.21.176 detected download samples: 915346cc61c5a247
28.21.176 detected communicating samples: 18adc6dbf78
28.21.176 detected urls: http://sbmsix.16mb.com/(2016
0.144.200 detected communicating samples: 213bec57309
0.144.200 detected urls: http://thehypemagazine.com/
185.27.101 detected download samples: 88904ca8c0a1c4c
185.27.101 detected urls: http://adelina.se/1/(2016-0
25.54.158 detected download samples: 3b3d6301af72df62
```

Recap

- Built YARA rule for one dropper
- Identified 700+ files
- Automated analysis via Cuckoo
- Logging via Elasticsearch and Visualization with Kibana
- Additional pivoting via API
- Source code: https://github.com/swackhamer

Questions?