RSAConference2016

San Francisco | February 29 – March 4 | Moscone Center

Connect to Protect

SESSION ID: AIR-F03

IOCs are Dead - Long Live IOCs!

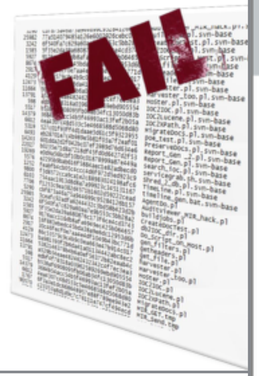
Ryan Kazanciyan

Chief Security Architect
Tanium
@ryankaz42

Yours truly, circa 2010

Before OpenIOC

- Lists of stuff to find evil
  - Easy to create
  - Difficult to maintain
  - Terrible to share
- Lists do not provide context
  - An MD5 of what?
  - Who gave me this?
  - Where is the report?
  - Where is the intelligence??
- **Lists encourage reliance on easily mutable forensic artifacts**



https://buildsecurityin.us-cert.gov/sites/default/files/RyanKazanciyan-APTPanel.pdf

IOC allows this…

```
File Name is sunjre16.exe
File Name is eic16ux.sys
File Name is e216ee.msi
File Name is webserv32.exe
File Name is 60927ux.sys
File Name is b26092.msi
File Name is uddi16.exe
File Name is aic16ux.sys
File Name is b216ee.msi
File MD5 is 5611458A5A03998CB1268190E2818C63
File MD5 is 711F4FE93EA0E8F253FA044BE273FE8B
File MD5 is 4BFDB1ACBB32340E3D4572CD88B9A6FC
File MD5 is CB8990122D2678990C874B4959306793
File MD5 is 8B911B2D6487F24AE6C194D9DA2DDF2C
File MD5 is 402366D97A64CCA71238A0FC771DEE30
File MD5 is 98A9DF9AC08A1765CB3EBE1d4AEA5498
File Name is commdlg64.exe
File Name is a131ux.sys
File Name is b30ee.msi
File Name is smscfg32.exe
File Name is a0c77ux.sys
File Name is b087ee.msi
File MD5 is 1954EB413FDAADE614031B2231E35C7B
File Name contains \Application Data\Microsoft\Media Player\DefaultStore32.exe
File Name contains \Application Data\Microsoft\Media Index\wmplibrary32.db
File Name contains \Favorites\jamny.jpg
Process Handle Name is www.TW0901.2.org
Process Handle Name is www.UG0902.2.org
Process Handle Name is www.UG0905.1.org
Process Handle Name is 1.2.UD0804.1x
Process Handle Name is www.WW0902.1.org
```

…to become this

Name: STISVC.DLL
Author: ryan.kazanciyan@mandiant.com
GUID: 116fc8d2-41b8-4cfc-8590-978d2414

Type	Reference
group	MFR00-0001
report	MA12345
category	Backdoor
grade	Release

Description:

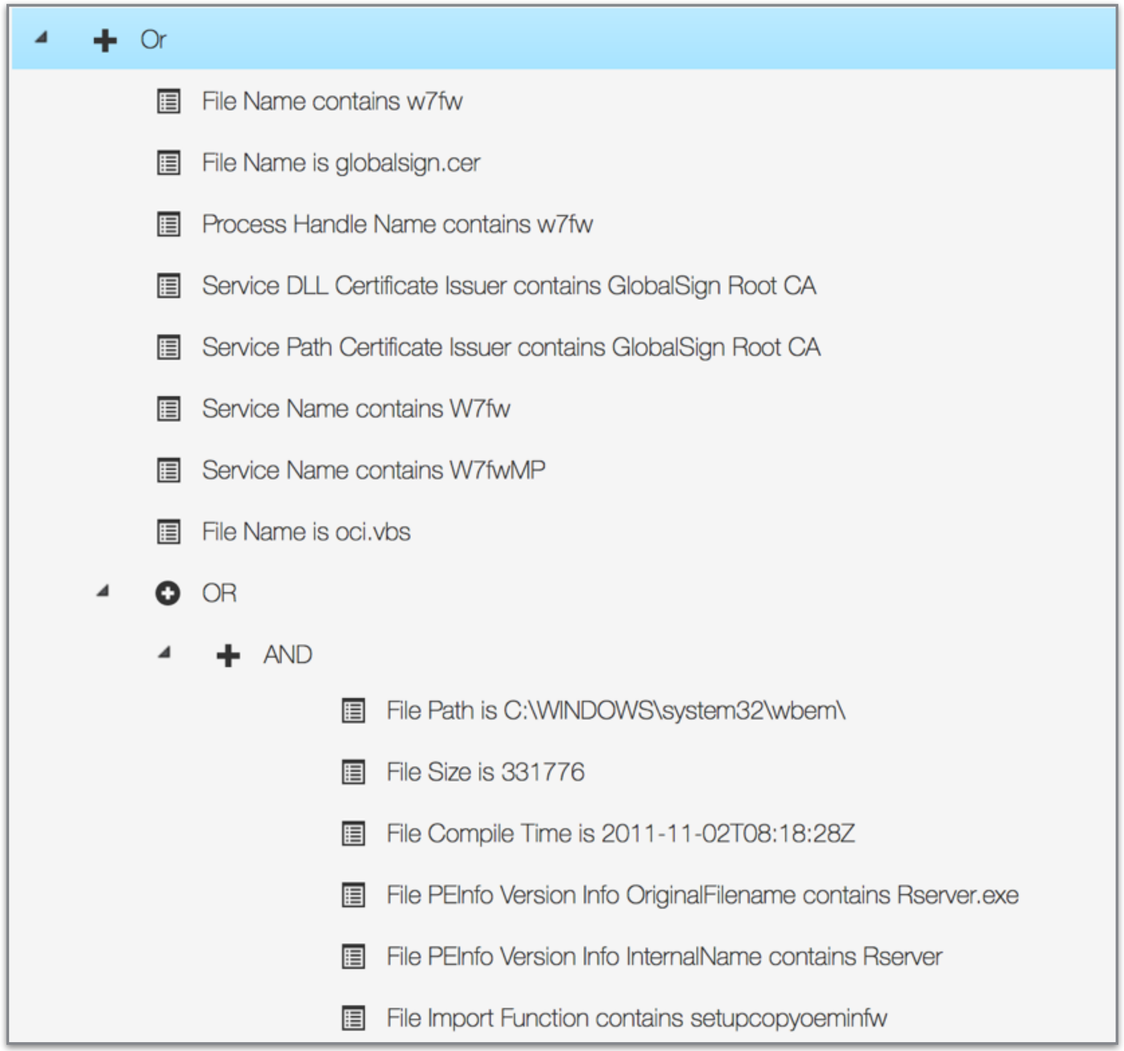
This malware is a backdoor trojan that gets loaded as a Windows Service. It initiates command and control via an SSL encrypted HTTPS session to a hard-coded C2 address. The backdoor features include remote file transfer, command execution, and screen capture

Add: Item AND OR

Definition:

```
OR
    Process Handle Name contains PccGlobalExitEvent
    Network DNS contains evilsite.com
    File Name is backins.exe
    File Name is stisv.dll
    File MD5 is e996c7ff1709e8013765151c1757efe7
    File MD5 is 44e03e0146729b6720ecc9d2f2964865
    AND
        File Import Function is CmdBatNotification
        File Import Function is FlushConsoleInputBuffer
        File Size is [99000] TO [125000]
        File Import Name is wininet.dll
        File Import Name contains urlmon.dll
        File Import Name contains changeserviceconfig2a
    AND
        Service Name contains StiSvc
```

IOCs as advertised



- Human-readable, machine-consumable
- Capture a broad set of forensic artifacts
- Foster information sharing
- Provide context around threats
- Do better than “signatures”

Five years later…

CybOX™

yara

R-CISC RETAIL CYBER INTELLIGENCE SHARING CENTER

NH-ISAC

ONG-ISAC

STIX™

MAEC™

DSIE Defense Security Information Exchange

FINANCIAL SERVICES Information Sharing and Analysis Center

OpenIOC

ICS-ISAC

DIB ISAC
DEFENSE INDUSTRIAL BASE
INFORMATION SHARING AND ANALYSIS CENTER
WWW.DIBISAC.NET

IOC quality and sharing in 2016

Indicators of Compromise (IOCs)

IOC	Type	Notes
91.207.61.208	Destination IP address	Command and control server
109.72.149.42	Destination IP address	Command and control server
130.0.237.22	Destination IP address	Command and control server
5.187.1.198	Destination IP address	Command and control server
ABA833D11679DFEBC95060BD3C557853	File MD5 hash	Malicious driver file
215BDF185C3B35503923FCF8872C75FC	File MD5 hash	Malicious driver file
F9C4E2D38DF8A87F545B6F5BA1F8691B	File MD5 hash	Malicious driver file
F21403B6CF7516B37EFFC17F410CED6F	File MD5 hash	Malicious driver file
6FBD31E7B5A31F5F75BD0D858D3327B5	File MD5 hash	Malicious driver file
68F40544ACD5568BD782434CA0F5AEE5	File MD5 hash	M
540DF6480B393BFA39D2E7CEC608EA12	File MD5 hash	Pa
%SystemRoot%\system32\drivers	Install path	M

usbf4.exe	*e36680a19601f84af6d311e1fb847eef*	Detected as TF STRPADT.A ar *pat2.exe*
vvb.exe	*2a38ff709549b97b4e42b6fae81c6177*	Modifies the au
vvb.sfx.exe	*f747d5f998e48279cad7e9ed46e86a6b*	Drops *VVB.exe*

Indicators of Compromise

Show 10 entries

TYPE	INDICATOR
domain	xxxmobiletubez.com
FileHash-SHA256	7b7eeca21a4aee3768b41b9e194052cbb01835ae3b3503c1d635abbe1
FileHash-SHA256	e6c1621158d37d10899018db253bf7e51113d47d5188fc363c6b5c51a6
FileHash-SHA256	f5bc281ee071f6fb0eb8d25f414770fee67e2ea6e02afe53896a2313f6cfe
IPv4	89.144.14.29
IPv4	185.86.148.188
IPv4	195.3.144.90
IPv4	89.144.14.59
IPv4	5.39.222.162
URL	http://185.86.148.188:2080/forms/

Name	FireEye
Description	This IOC contains indicators detailed in the blog post "Operation Double Tap"

- Or
  - File MD5 is 5a0c4e1925c76a959ab0588f683ab437
  - File MD5 is 492a839a3bf9c61b7065589a18c5aa8d
  - File MD5 is 744a17a3bc6dbd535f568ef1e87d8b9a
  - File MD5 is 5c08957f05377004376e6a622406f9aa
  - File MD5 is 8849538ef1c3471640230605c2623c67
  - DNS Record Name contains join.playboysplus.com
  - DNS Record Name contains inform.bedircati.com
  - DNS Record Name contains pn.lamb-site.com
  - DNS Record Name contains securitywap.com
  - DNS Record Name contains walterclean.com
  - Port Remote IP contains 192.157.198.103

My own point of reference

2009 - 2015: Investigator

- Large-scale, targeted attacks
- Designed, tested, and applied IOCs for proactive and reactive hunting

2015 - Present: Builder

- Designing an EDR platform that includes IOC detection
- Helping orgs build self-sustaining, scalable “hunting” capabilities

The erosion of indicator-based detection

- **Brittle indicators - short shelf-life**
- **Poor quality control in threat data feeds**
- **Hard to build effective homegrown IOCs**
- **Indicator detection tools are inconsistent**
- **IOCs applied to limited scope of data**

“IOCs” vs. “threat data” vs. “intelligence”

- IOCs are structured **threat data**
- **Threat data != threat intelligence**
- Threat intelligence provides context and and analysis
- Threat intelligence is ineffective without quality threat data

IOCs are brittle

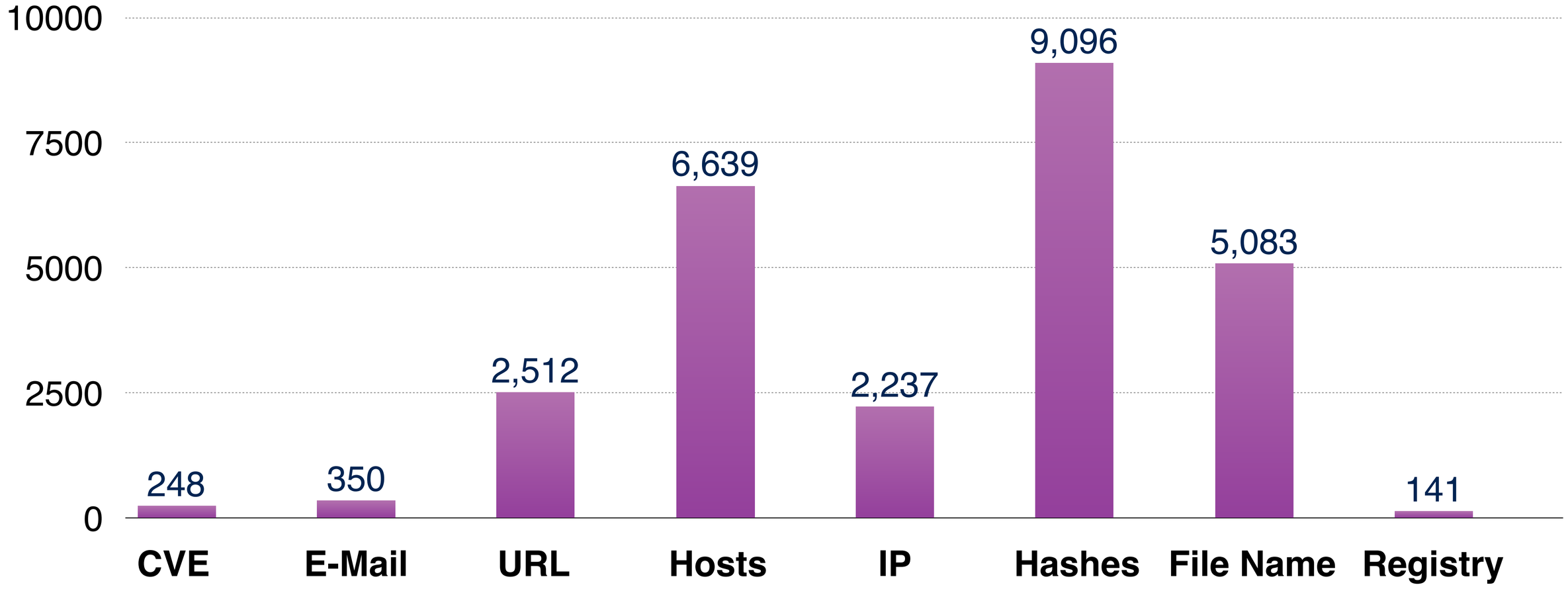
Verizon DBIR 2015: Most shared IOC types

COMMUNITY	IP ADDRESSES	E-MAIL ADDRESSES	FILES	HOSTS	URLS
Common Community	35.9%	1.0%	23.3%	33.0%	6.8%
Event-Based Community #1	77.4%	0.1%	2.5%	19.5%	0.5%
Industry Community #1	16.5%	32.3%	6.3%	43.0%	1.9%
Industry Community #2	47.1%	4.4%	10.3%	29.4%	8.8%
Industry Community #3	8.3%	0.3%	1.2%	87.5%	2.7%
Industry Community #4	25.2%	2.4%	9.0%	58.6%	4.8%
Industry Community #5	50.9%	0.7%	1.3%	22.8%	24.4%
Industry Community #6	66.4%	0.6%	14.0%	13.8%	5.2%
Industry Community #7	59.1%	0.5%	1.4%	23.5%	15.5%
Industry Community #8	39.6%	3.0%	7.7%	36.9%	12.8%
Industry Community #9	51.5%	2.6%	12.6%	23.8%	9.5%
Regional Threat Community #1	49.2%	0.3%	4.5%	42.6%	3.4%
Regional Threat Community #2	50.0%	1.1%	4.5%	30.8%	13.6%
Subscriber Community	45.4%	1.2%	18.4%	24.4%	10.6%
Threat-Based Community #1	50.3%	1.1%	11.0%	24.3%	13.3%

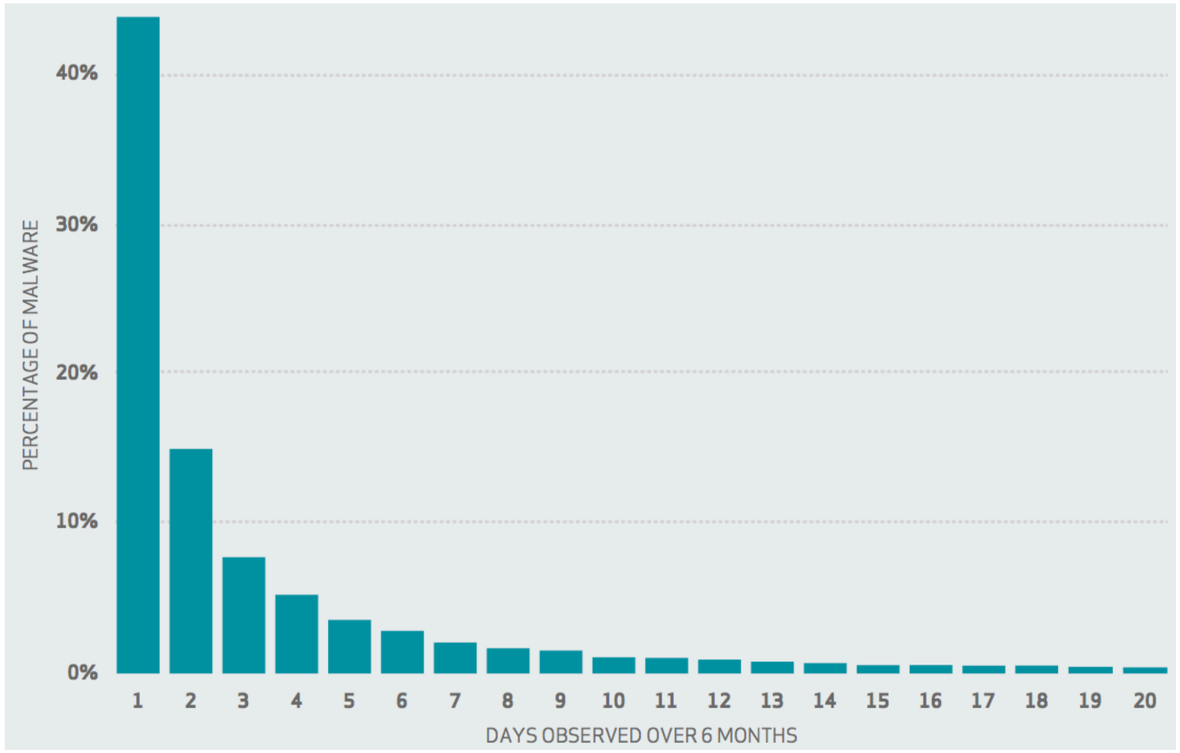
Source: Verizon DBIR 2015

IOCs in the APTnotes data set

Derived from over 340 threat reports (2006 - 2015) archived on https://github.com/kbandla/APTnotes



This will never keep pace…



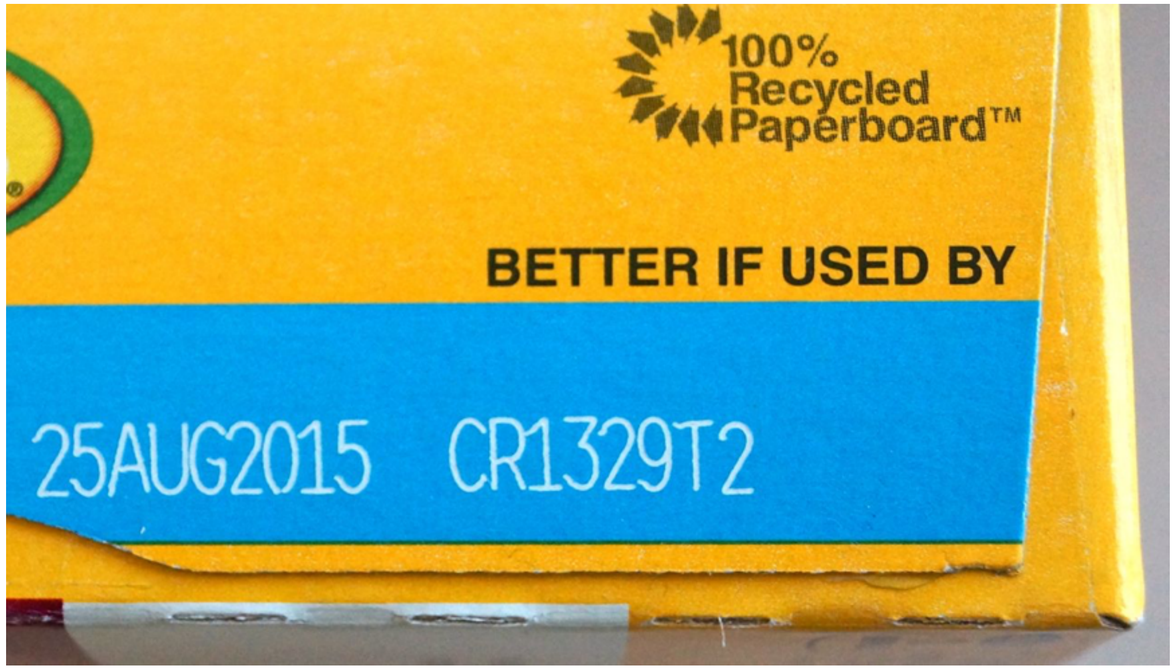
Source: Verizon DBIR 2015

95%

OF MALWARE TYPES SHOWED UP FOR LESS THAN A MONTH, AND FOUR OUT OF FIVE DIDN'T LAST BEYOND A WEEK.

The problem extends beyond file hashes

- Short lifespan of C2 IPs and domains
- Malicious sites co-located on virtual host server IPs
- Low barrier to host malicious content on legitimate providers



Sheer volume does not solve the problem

- 2007: Bit9 FileAdvisor tracked **4 billion unique files**, catalog grew by 50 million entries per day
- 2009: McAfee Global Threat Intelligence tracked reputation data for **140 million IP addresses**, handling **50 million file lookups** per day
- 2011: Symantec Insight tracked **tens of billions** of linkages between users, files, web sites

Seven years of progress?

"…innovating to provide **predictive security**. This approach comprises interconnected security technology at multiple layers in the technology stack, backed by **global threat intelligence**. Predictive security will allow security products to **intelligently block attacks much sooner than is currently possible**…"

"…an **intelligence-led approach** to security will be key in detecting the most sophisticated threats and **responding** to them **quickly** and **effectively**."

Paid IOCs != quality IOCs

Have you assessed your feeds?

- 74% of enterprise cybersecurity professionals say that it is extremely difficult or somewhat difficult to determine the quality and efficacy of each individual threat intelligence feed.

Jon Oltsik / ESG, http://www.networkworld.com/article/2951542/cisco-subnet/measuring-the-quality-of-commercial-threat-intelligence.html

My (incredibly scientific) methodology

- Chose two top-tier paid threat feed services
- Retrieved the most recent ~20 indicators from each
- Spent 15 minutes eyeballing their contents

What are you paying for?

Too specific - malware hash AND'd with a filename

(Real IOC from a commercial feed)

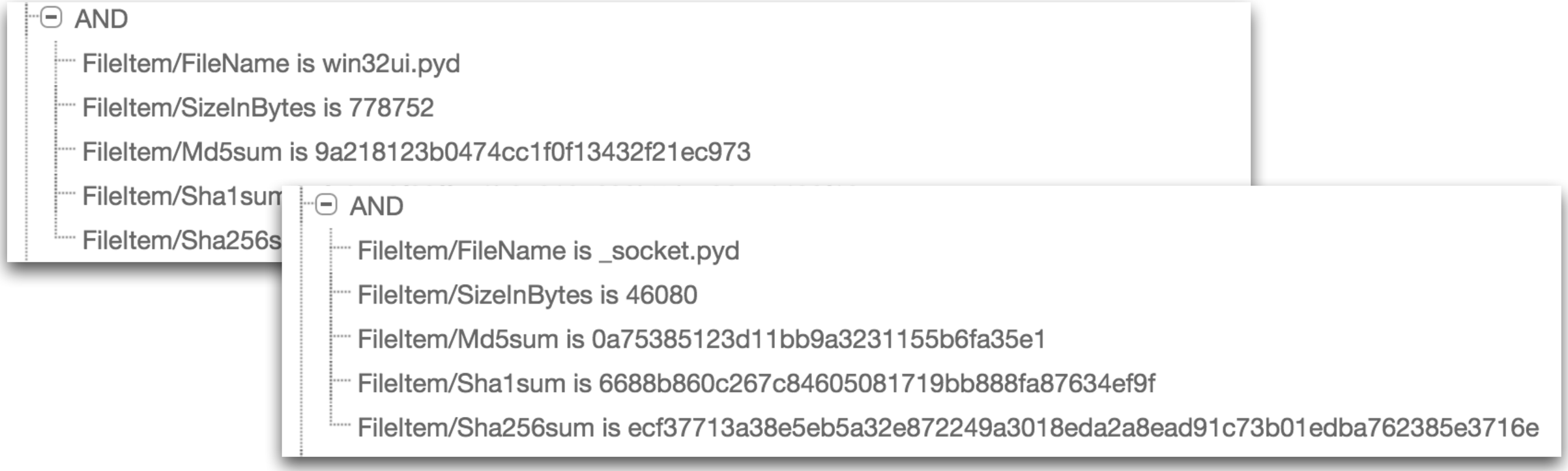
What are you paying for?

Too specific - LNK files are unique per-system

(Real IOC from a commercial feed)

What are you paying for?

Too noisy - matches component of legitimate software

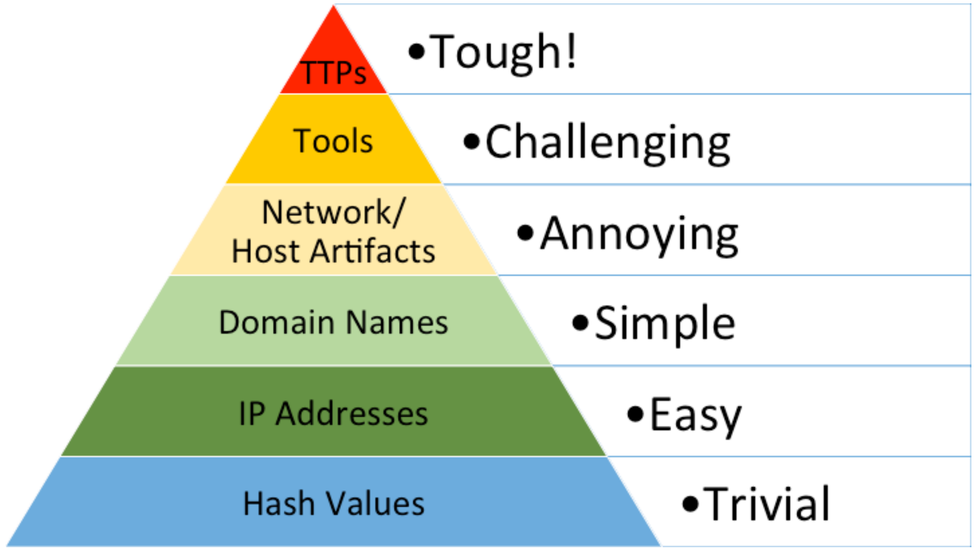


(Real IOC from a commercial feed)

Building good IOCs is hard

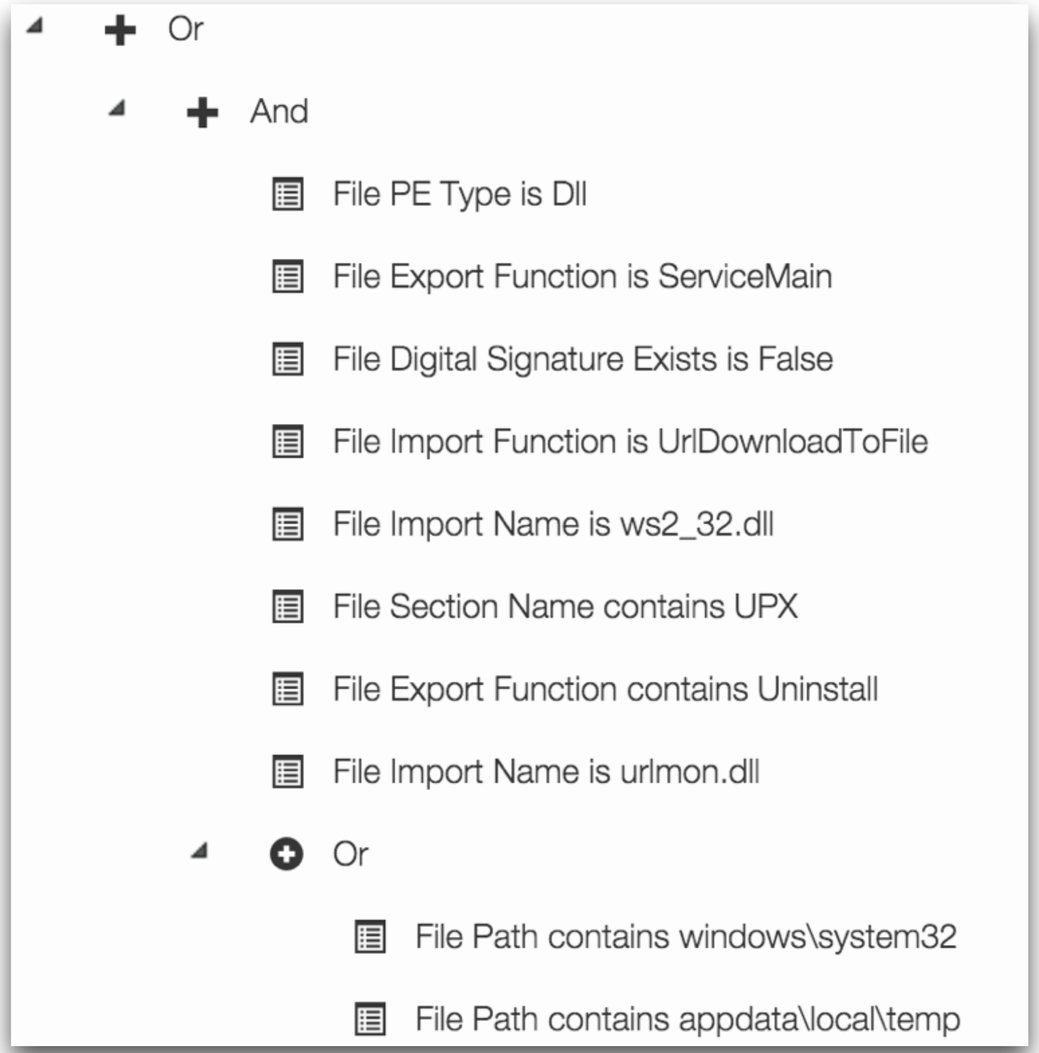
Challenges with IOC development

- Easy to build high-fidelity IOCs
  - (may yield high false-negatives)
- Hard to build robust IOCs
  - (may yield higher false-positives)
- Easy to build IOCs that don't evaluate properly
  - (tools have inconsistent matching logic)



"Pyramid of Pain", David Bianco
http://detect-respond.blogspot.co.uk/2013/03/the-pyramid-of-pain.html

Running aground on a robust IOC

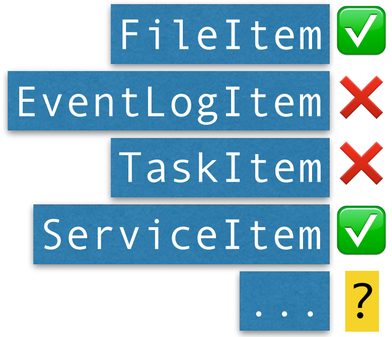


Too broad - may match on uncommon but legitimate binaries

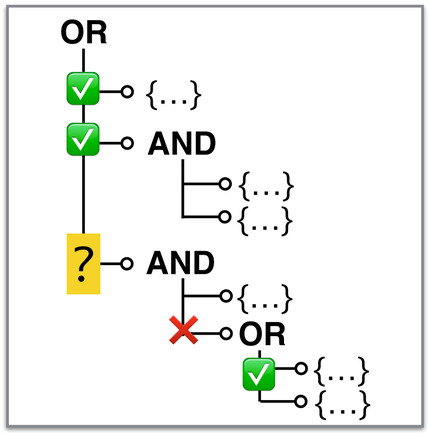
How much time do your analysts have to continuously build, test, and refine IOCs like this?

Inconsistencies in IOC detection tools

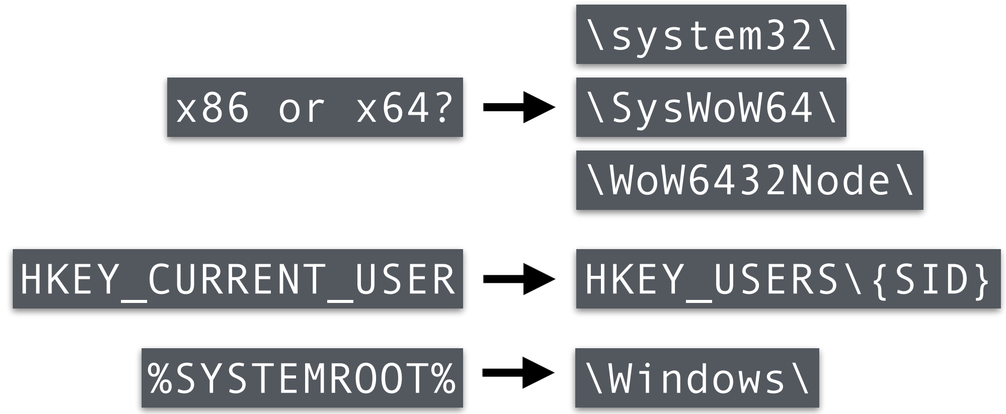
Supported Observables



Logic Handling



Data Normalization



- STIX & CybOX have a few tools to help with this:
  - maec-to-stix
  - python-cybox/normalize.py

Issues specific to OpenIOC

What happens when you try to turn a proprietary tool's unique output schema into a "standard"…

`ProcessItem/PortList/PortItem/process` **"Process Port Process"**

`FileItem/PEInfo/DetectedAnomalies/string` **"File PE Detected Anomali**

`FileItem/PEInfo/DetectedEntryPointSignature/Name` **"File EntryPoint Sig Name"**

Issues specific to OpenIOC

Example: Registry evidence in OpenIOC

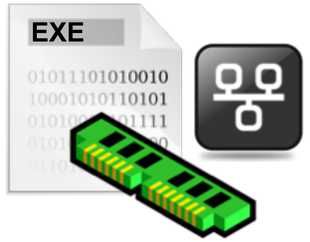
```
Key:   HKLM\SOFTWARE\Microsoft\Windows\CurrentVersion\Run
Value: Backdoor
Data:  C:\path\to\malware.exe
```

```
RegistryItem/Path:      HKLM\SOFTWARE\Microsoft\Windows\CurrentVersion\Run\Backdoor
RegistryItem/KeyPath:   \SOFTWARE\Microsoft\Windows\CurrentVersion\Run
RegistryItem/Value:     C:\path\to\malware.exe
RegistryItem/ValueName: Backdoor
RegistryItem/Text:      C:\path\to\malware.exe
```

Broadening the scope of endpoint indicator usage

Focusing on scope of data, not tools

- What are you matching your endpoint IOCs against?
- What's your cadence of detection?
- Where are your gaps?



Current Activity
(Processes, Network Connections, Memory)

Historical Activity
(Telemetry, logs, alerts, historical data)

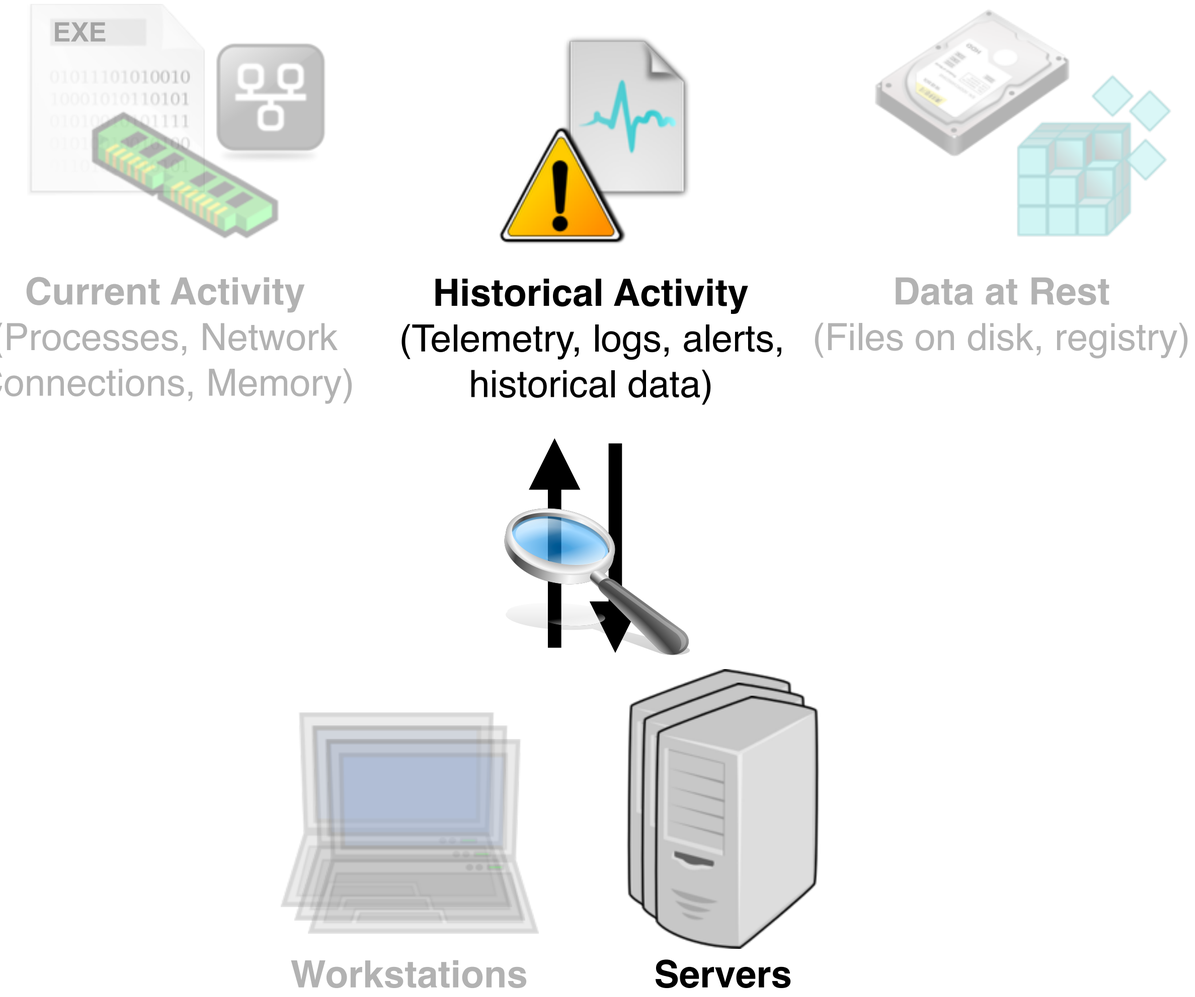
Data at Rest
(Files on disk, registry)

Workstations

Servers

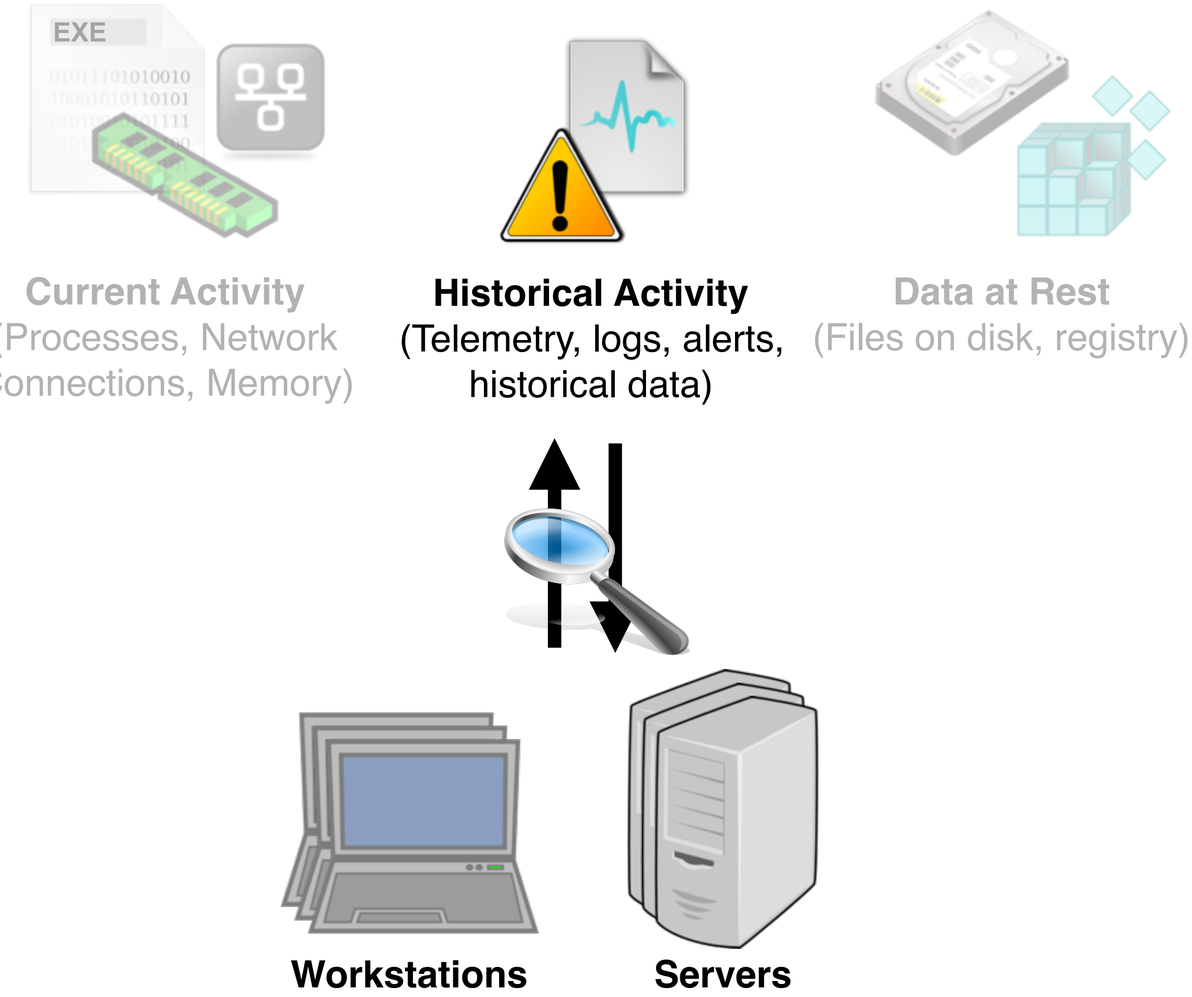
Matching on SIEM / centralized logging

- Most common endpoint data in SIEM:
  - Anti-virus / anti-malware alerts (all systems)
  - Event log data (subset of systems - usually servers)
- **Resource impact** of large-scale event forwarding & storage **limits endpoint coverage & scope of data**



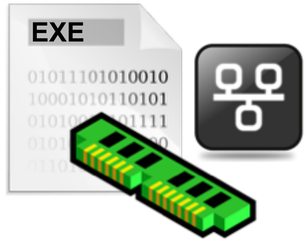
Matching on forensic telemetry

- Process execution, file events, network connections, registry changes
- Preserves historical data, short-lived events
- Expensive to centralize in large environments
- **Limited scope of data** for IOC matching



Matching on live endpoints

- Potentially the broadest set of available data
- Considerations
  - Endpoint impact
  - Availability
  - Time-to-assess
  - Scalability



Current Activity
(Processes, Network Connections, Memory)

Historical Activity
(Telemetry, logs, alerts, historical data)

Data at Rest
(Files on disk, registry)

Workstations

Servers

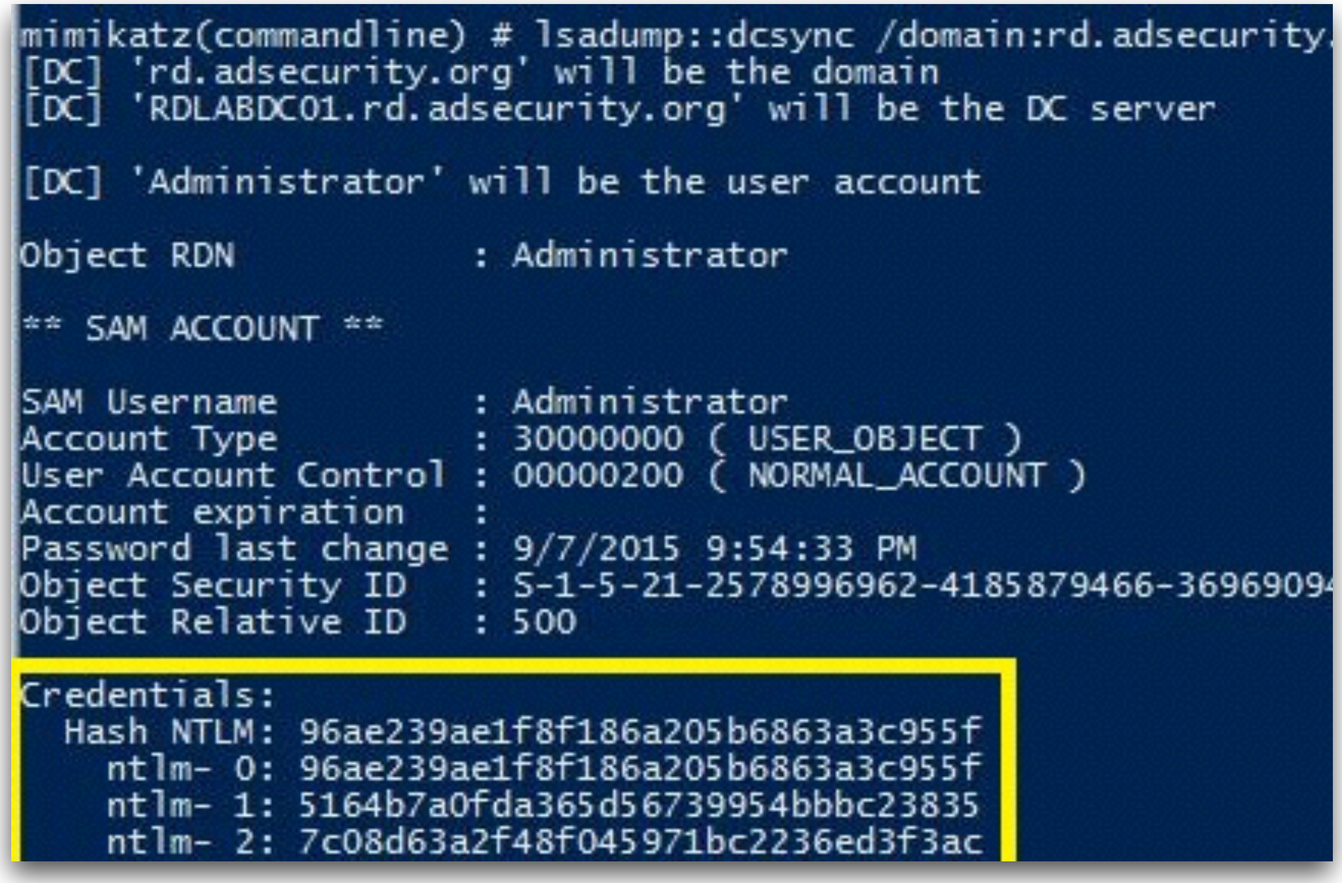
The ideal combination

- Goal: Maximize the value of brittle IOCs
- Telemetry for efficiency, historical data
- On-endpoint to maximize current state & at-rest data
- Increase cadence as tools & resources permit
- Don’t take shortcuts on scope of coverage!

“I only need to check important systems”

An example of why this fails:

- Credentials can be harvested from anywhere on a Windows network
- No need to run malicious code on admin systems or DCs
- By the time they get to “crown jewels”, attackers are already authenticating with legitimate accounts



Source: https://adsecurity.org/?p=1729

Shrinking the detection gap

Doing better with what we've got

"The desire to take a technical feed and simply dump it into our security infrastructure doesn't equate to a threat intelligence win...

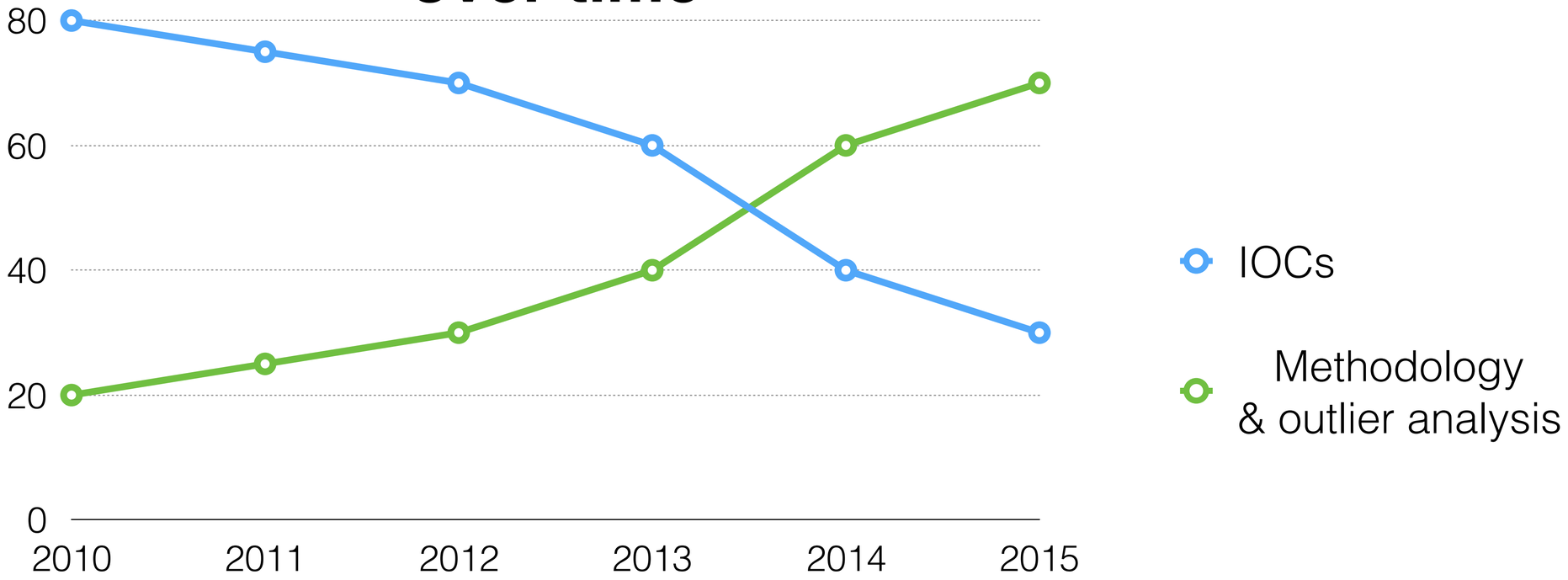
You cannot get more relevant threat intelligence than what you develop from within your own environment. This should then be enriched with external intelligence"

-Rick Holland, Forrester, 2016 CTI Summit

Source:
https://www.digitalshadows.com/blog-and-research/another-sans-cyber-threat-intelligence-summit-is-in-the-books/

My own point of reference

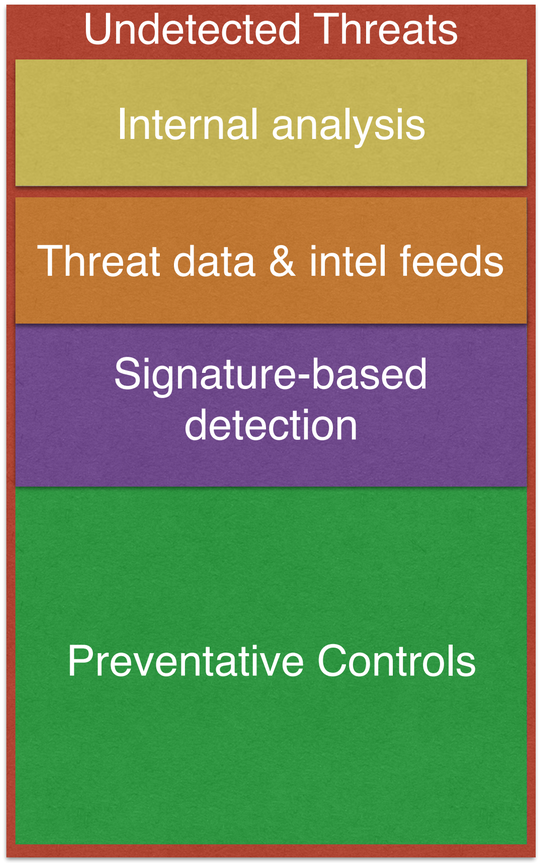
As an investigator:
Relative efficacy of IOCs vs. methodology & outlier analysis over time



(Rough approximation for the sake of having a pretty graph)

Resetting expectations

Expectation



Reality

High-quality threat data and intelligence can help you…

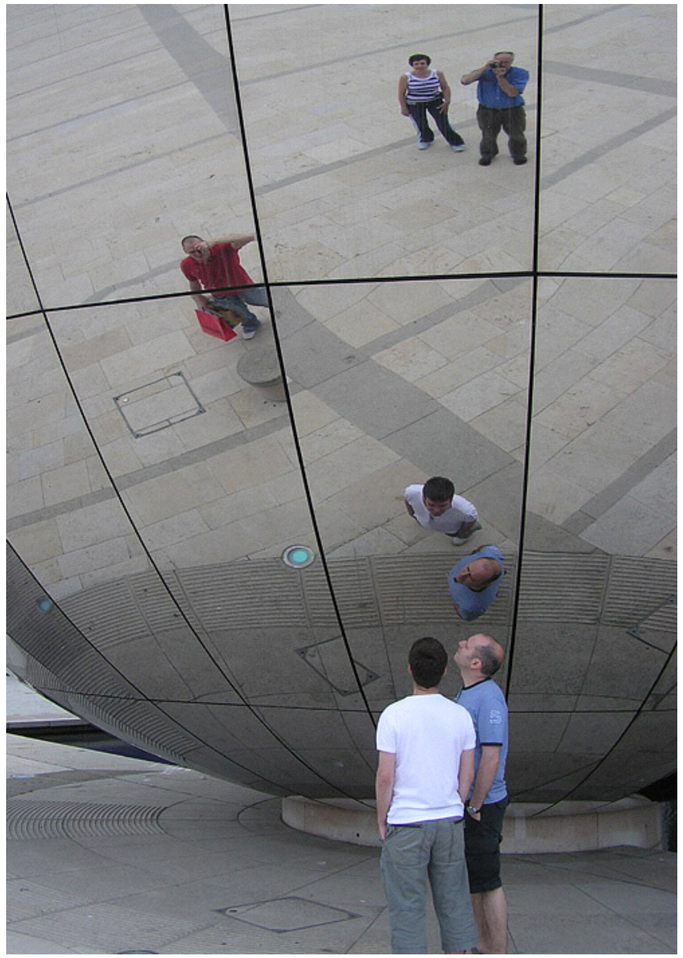
- Categorize and contextualize known threats, streamline response
- Provide additional layer of automated detection

...but it cannot...

- Tell you what's normal in your own environment
- Exceed the benefits of well-implemented preventative controls
- Close the gap of undetected threats

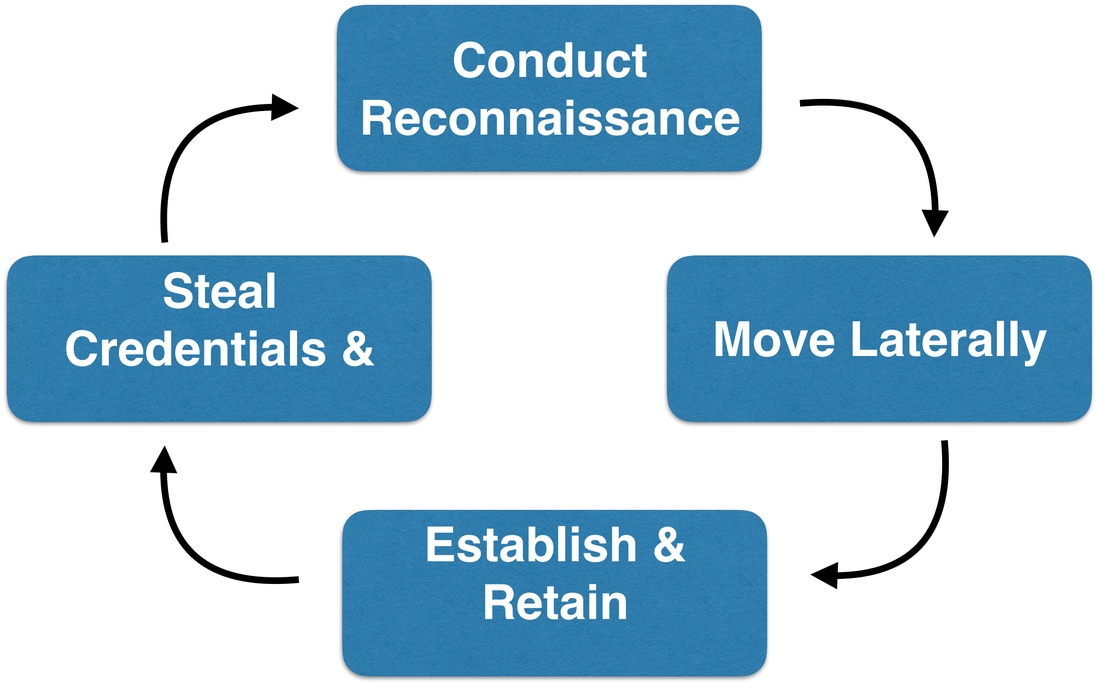
Looking inward to hunt

- Derive intelligence from what's "normal"
- Build repeatable analysis tasks
- Combine with automated use of IOCs and threat data
- More is not always better!
  - Easy to overwhelm yourself
  - Take on discrete, high-value data sets one at a time



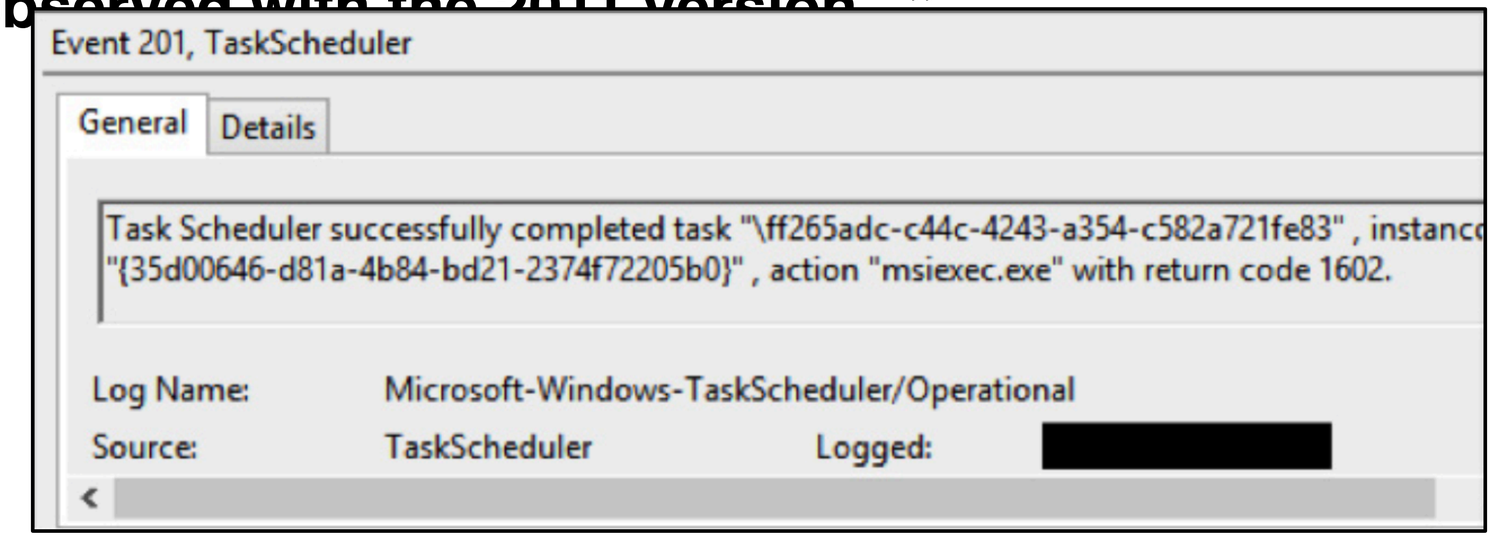
Aligning to the attack lifecycle

- What are the "lowest common denominators" across targeted intrusions?
- What readily-available evidence do they leave behind?
- What easily-observable outlier conditions do they create?



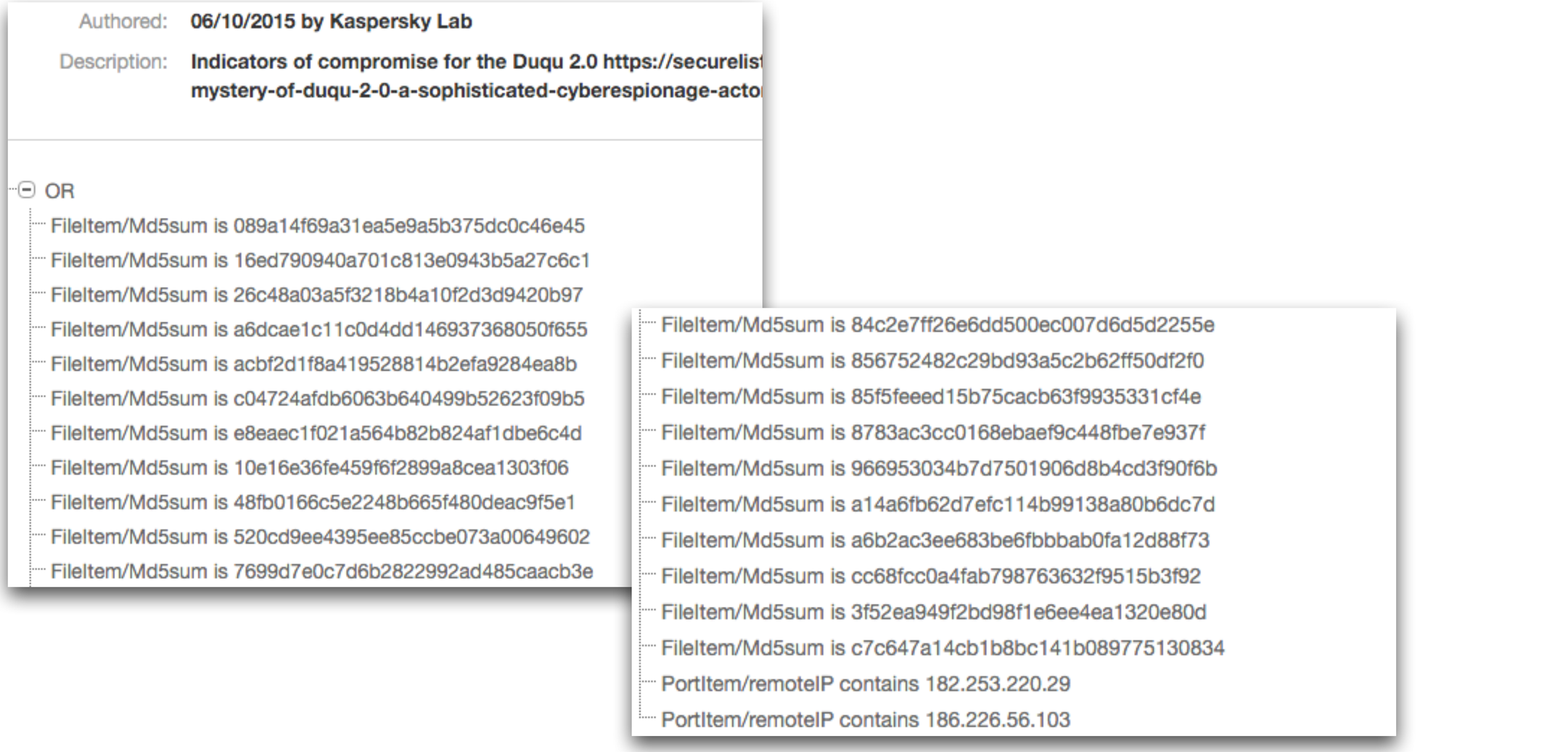
Example: Hunting for Duqu 2.0

“In addition to creating services to infect other computers in the LAN, attackers can also use the Task Scheduler to start ‘msiexec.exe’ remotely. The usage of Task Scheduler during Duqu infections for lateral movement was also observed with the 2011 version…”



Source:
https://securelist.com/files/2015/06/The_Mystery_of_Duqu_2_0_a_sophisticated_cyber espionage_actor_returns.pdf

What was the shared IOC?



How could we do better?

- We could just add a specific **TaskItem** to the IOC...
- …but what about other variants?
- How can we find evidence of other malicious activity that abuses the same (incredibly common) lateral movement technique?

Example: Lateral command execution

Attacker

Scheduled Tasks

WinRM & PowerShell

PsEx

Sources of Evidence

Logon & service events

Process history

Other forensic artifacts

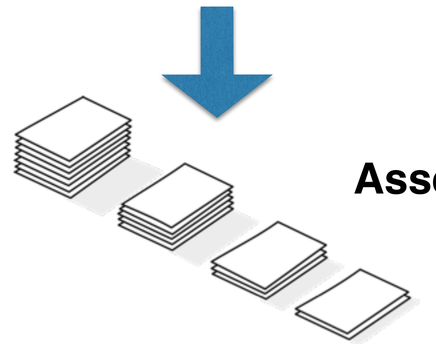
Analysis Criteria

Who? — Accounts used

Where? — Source & target systems

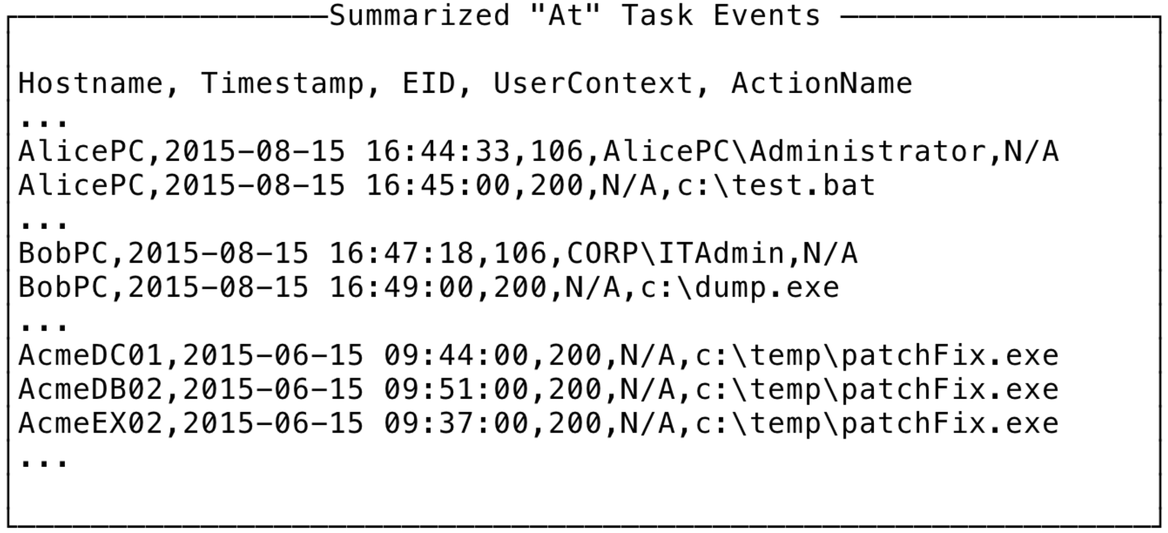
When? — Time & frequency

What? — Executed commands, dropped files, etc.

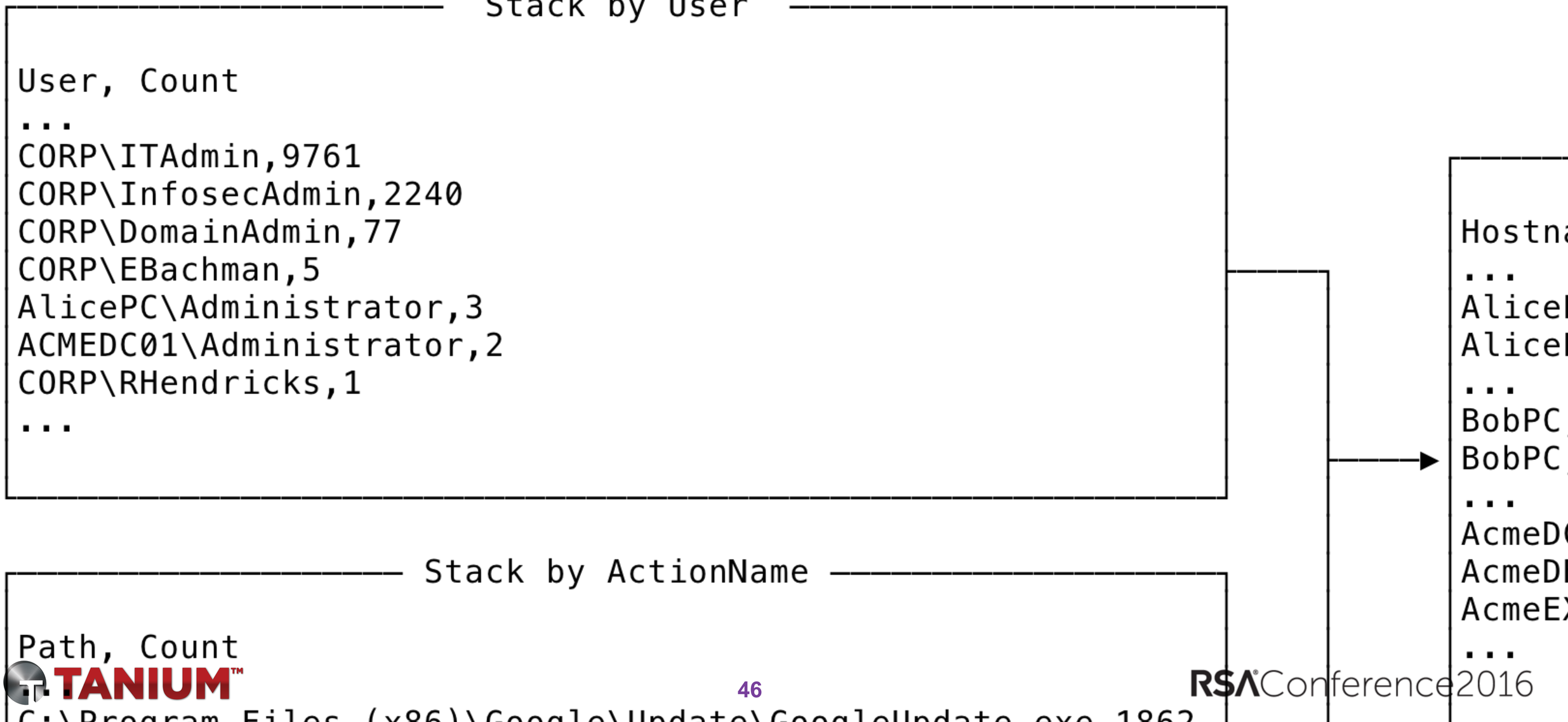


Assess outliers

Resulting stack analysis



Resulting stack analysis



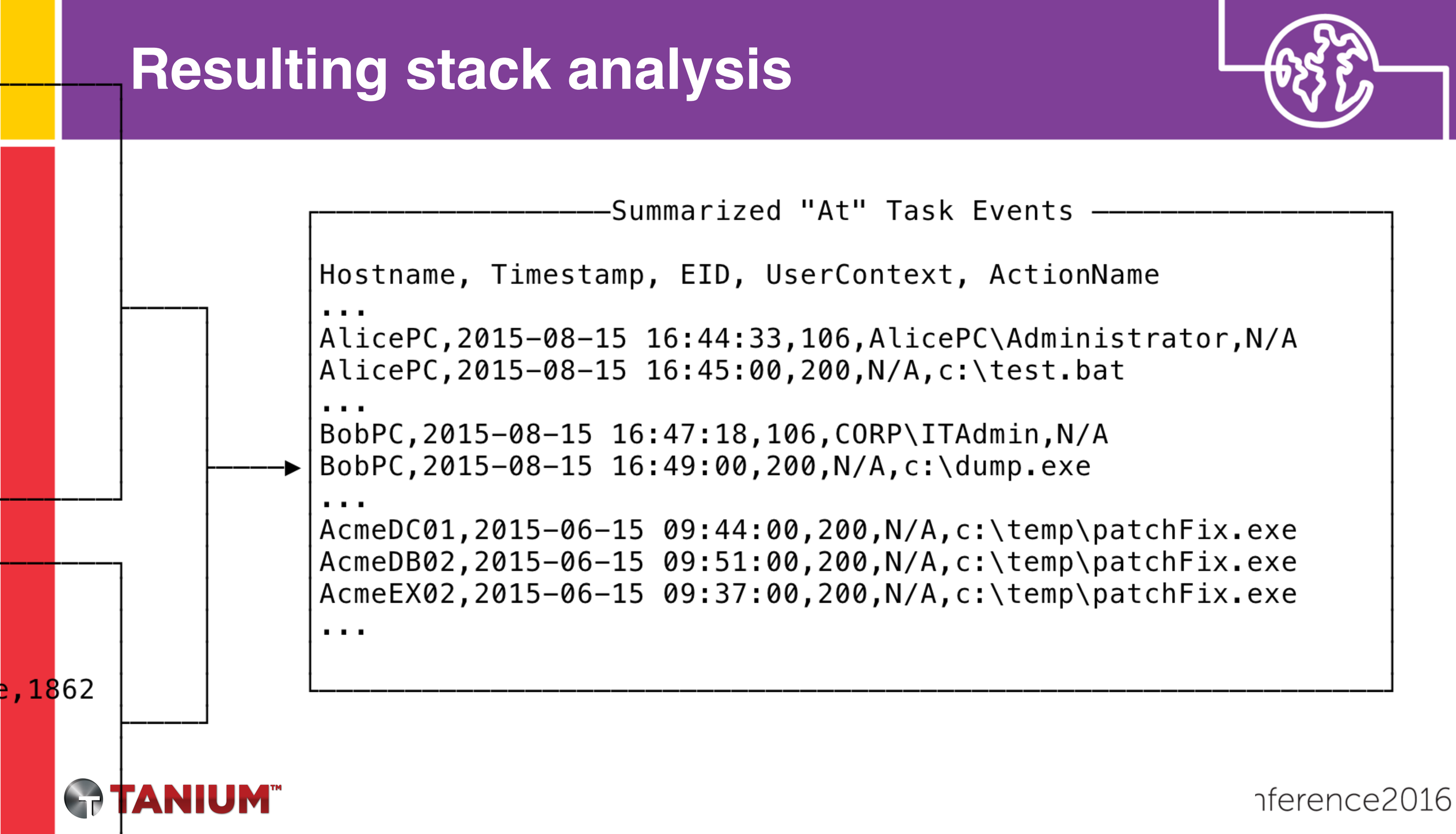
Resulting stack analysis

```
ACMEDC01\Administrator,2
CORP\RHendricks,1
...
```

Stack by ActionName

```
Path, Count
...
C:\Program Files (x86)\Google\Update\GoogleUpdate.exe,1862
c:\temp\patchFix.exe,310
...
c:\test.bat,15
c:\dump.bat,7
c:\temp\1.bat,3
...
```

Resulting stack analysis



For additional examples

- “Hunting in the Dark”
  - https://speakerdeck.com/ryankaz
- Includes coverage of:
  - More task analysis
  - ShimCache and process history
  - Service Events
  - WMI event consumers
  - Alternative authentication mechanisms

Closing thoughts and takeaways

Evolving standards & platforms

- Platforms
  - MISP
    http://www.misp-project.org
- Hubs and exchanges
  - Facebook ThreatExchange
    https://threatexchange.fb.com
- Standards
  - CybOX 3.0 refactoring and simplification

Quantitative assessment of threat feeds

- Few efforts to-date - this is difficult!
- **Threat Intelligence Quotient Test** (tiq-test)
  - Statistical analysis of IPs and domains in threat feeds
  - References:
    https://github.com/mlsecproject

    https://defcon.org/images/defcon-22/dc-22-presentations/Pinto-Maxwell/DEFCON-22-Pinto-and-Maxwell-Measuring-the-IQ-of-your-threat-feeds-TIQtest-Updated.pdf

Ask your threat feed vendor

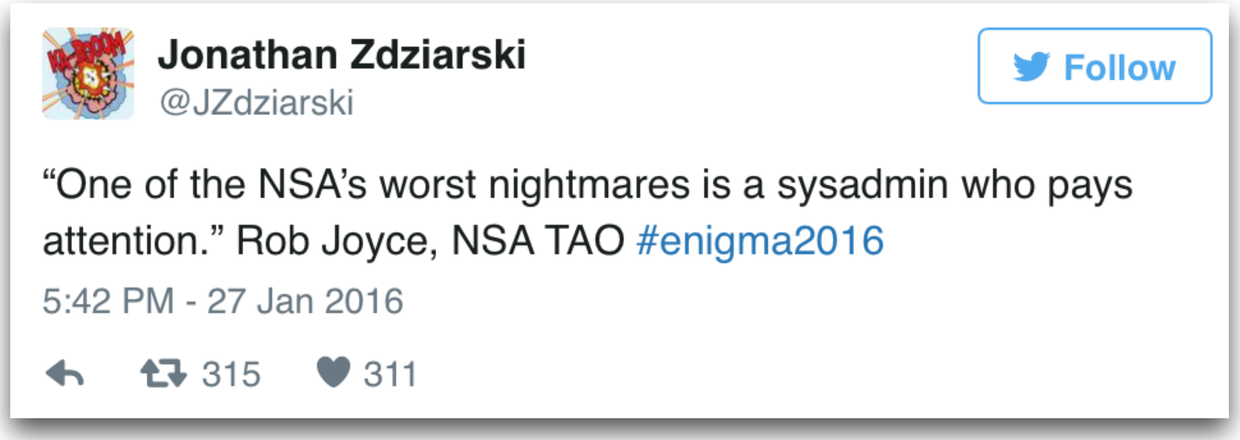
- **Where's the intel coming from?**
  - Professional services
  - Managed security services
  - Partners
  - Honeypots
  - "Open source" data gathering
  - Auto-generated sandbox data
- **What's the breakdown of observable types?**
- **What QC is in place?**
  - Test-cases
  - Documentation
  - Spot-checking

Maximize your IOCs & threat data

- Where are your gaps in endpoint & network **visibility**?
- Can you expand the **scope of data** made available for endpoint IOC matching in your environment?
- Are your tools and threat data sources **fully compatible**?
- How **quickly** are you consuming new threat data? At what **scale**?

Have your investments made you more secure?

- Even the best sources of threat data will never keep pace with emerging attacks
- Know your network above all
- Invest in attack surface reduction and “hygiene”. It really does make a difference.



RSAConference2016

San Francisco | February 29 – March 4 | Moscone Center

Connect to Protect

SESSION ID: AIR-F03

Thank you!

Ryan Kazanciyan

Chief Security Architect
Tanium
@ryankaz42