Demystifying the Hunt

Threat Hunting Basics and Practical Application

Brian Gittinger

Senior Sales Engineer @ Endgame, Inc.

- InfoSec professional with SOC experience - GE
- Consulting for Fortune 500
- Sales Engineer for Cyber Security Services and Product companies

Agenda for this talk

- Basics of standing up a threat hunting operation
- What skills, data + tools are useful for reaching success
- Can I do this with the staff I have? What about external consulting services?
- Advanced Hunting: Evolving the threat hunting program

Why this Topic?

Threat Hunting 101 – what are we talking about?

- Is this a new concept? What does threat hunting really mean?
- Is this in the ballpark of computer forensics? Is there overlap?
- Is hunting the same as detecting unknown ‘badness’?
- How important is this function within a Security Operations team structure?

Hunt teams still struggle with …

- Aligning hunting campaigns & business priorities
- Providing transparency to senior leadership
- Showing progress over time
- Mapping gaps to data sources and security controls
    - Coverage of adversary techniques is much more than a green check or red “x”
- Assessing the effectiveness of the program and any tools used during engagements
- Developing and implementing parity with ATT&CK
- Just getting started – no joke

Effective Hunting

How are you effective? What does being effective even mean?

____ for effective threat hunting

Efficiency Efficacy

Effectiveness

Efficiency

The way resources are used (or wasted),
How much I make the most of the resources I have

Effectiveness

Accomplishes the goals (to be efficacious) employing the best and most economic methodology (to be efficient).

Efficacy

It doesn't matter how we do it, but only on what we accomplish

Efficiency

- Choosing an adversary model
- Assessing quality of data
  - Do we even have the data?
- Utilizing the right technology
- Applying the right personnel skills
- Prioritizing adversary techniques
- Enhancing data security analytics

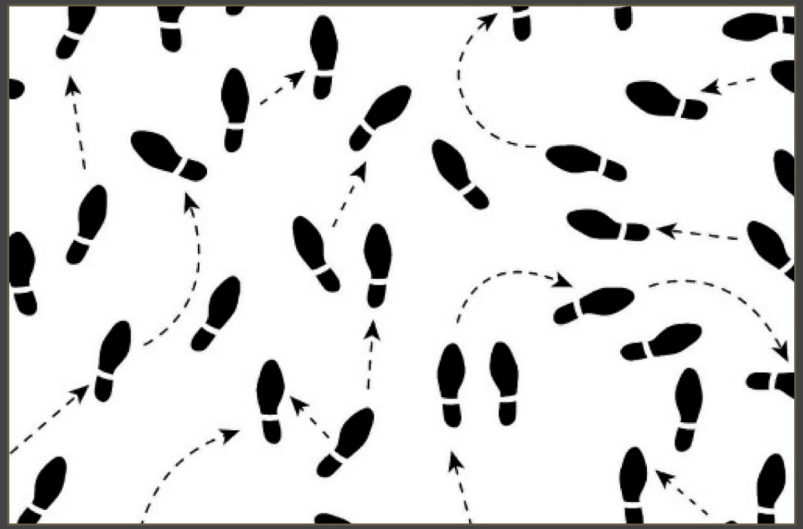
Efficacy

- Let's find evil! Can we detect it? Yes or No?
  - Signatures vs security analytics
  - Are you considering attack variations?
- Uncovering Incidents vs Validating Detection of adversaries

Where do I start?

How are we going to start approaching this?

Two steps back, one step forward: or history repeats itself



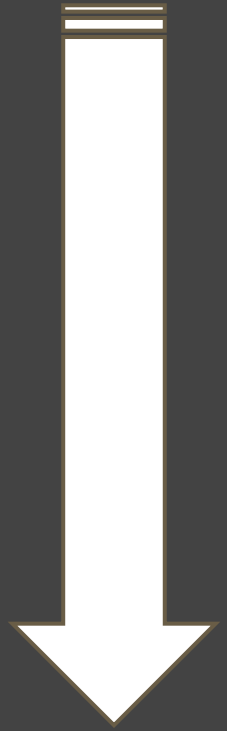
The Evolution of the Hunt HeatMap

How Hot Is Your Hunt Team?

https://cyberwardog.blogspot.com/2017/07/how-hot-is-your-hunt-team.html

Ready to hunt? First, Show me your data!

https://cyberwardog.blogspot.com/2017/12/ready-to-hunt-first-show-me-your-data.html



We’re not ready to measure anything just yet...

What are you potentially measuring already?

Risk Forecasting

- Choose a risk to measure
- Decompose the scenario
- Gather supporting data
- Make forecasts
- Mitigate the potential risk
- Measure again

Scenario: *"An attacker can access destructive AWS IAM permissions in the next 365 days."*

2016 - Q3: 25%

First forecast. We haven't fixed anything yet.

2016 - Q4: 23%

We have limited the destructive capability of keys in production.

2017 - Q1: 16%

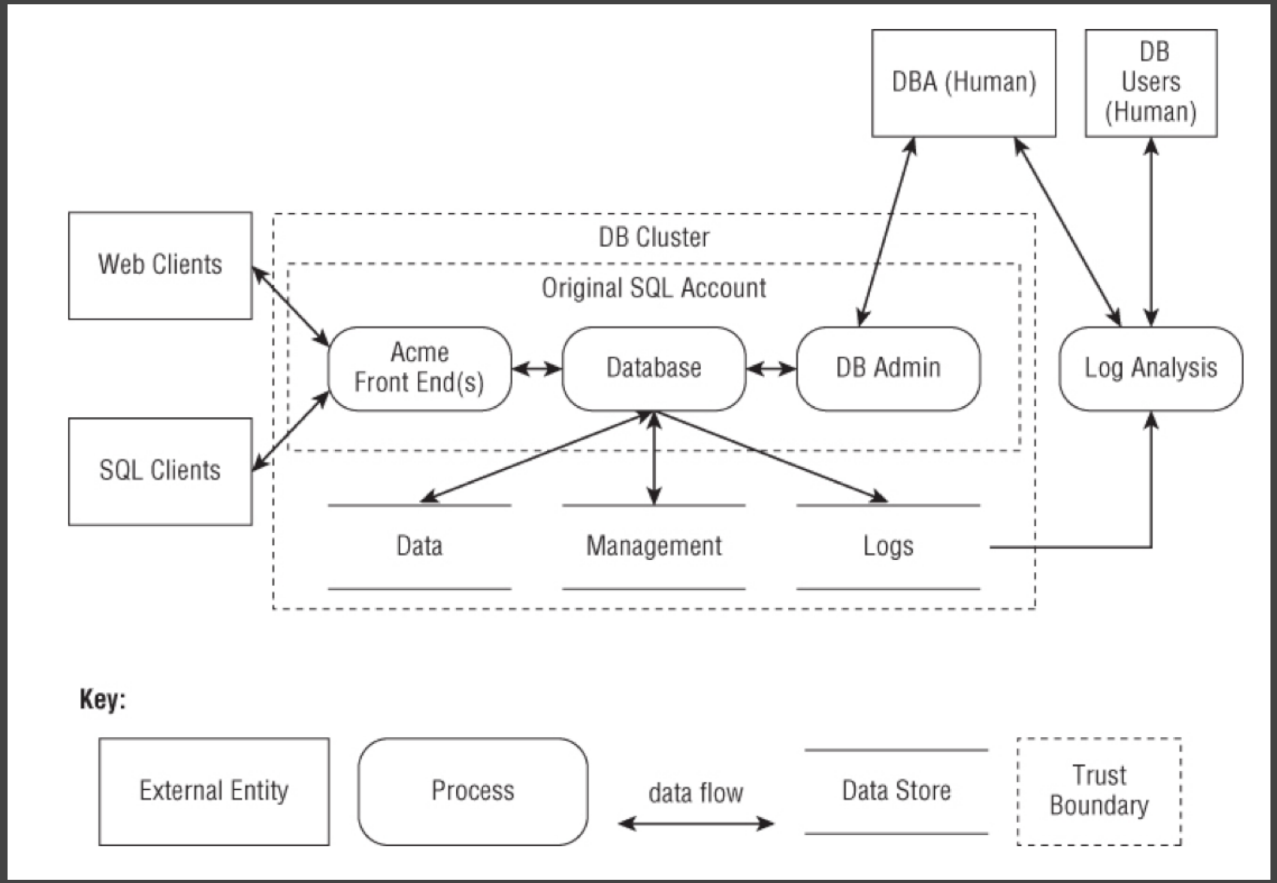
We added multifactor protection to keys used by engineers.

2017 - Q2: 10%

We took keys out of source code and use roles now.

An increase in confidence against this risk of 15%.

https://magoo.github.io/Risk-Forecasting/#173 (Ryan McGeehan @Magoo)



Threat Modeling

- Model the system
- Identify Threats
- Define how threat occurs
- Address threats
- Validate
- Measure again

Shostack, Adam. Threat Modeling: Designing for Security (p. 46). Wiley. Kindle Edition

Where do you fit “hunt”?

Threat Hunting

Identify a technique

Develop a hypothesis

Identify scope and resources

Develop Analytics

Validate & Report

Automate? & Repeat

Risk Forecasting

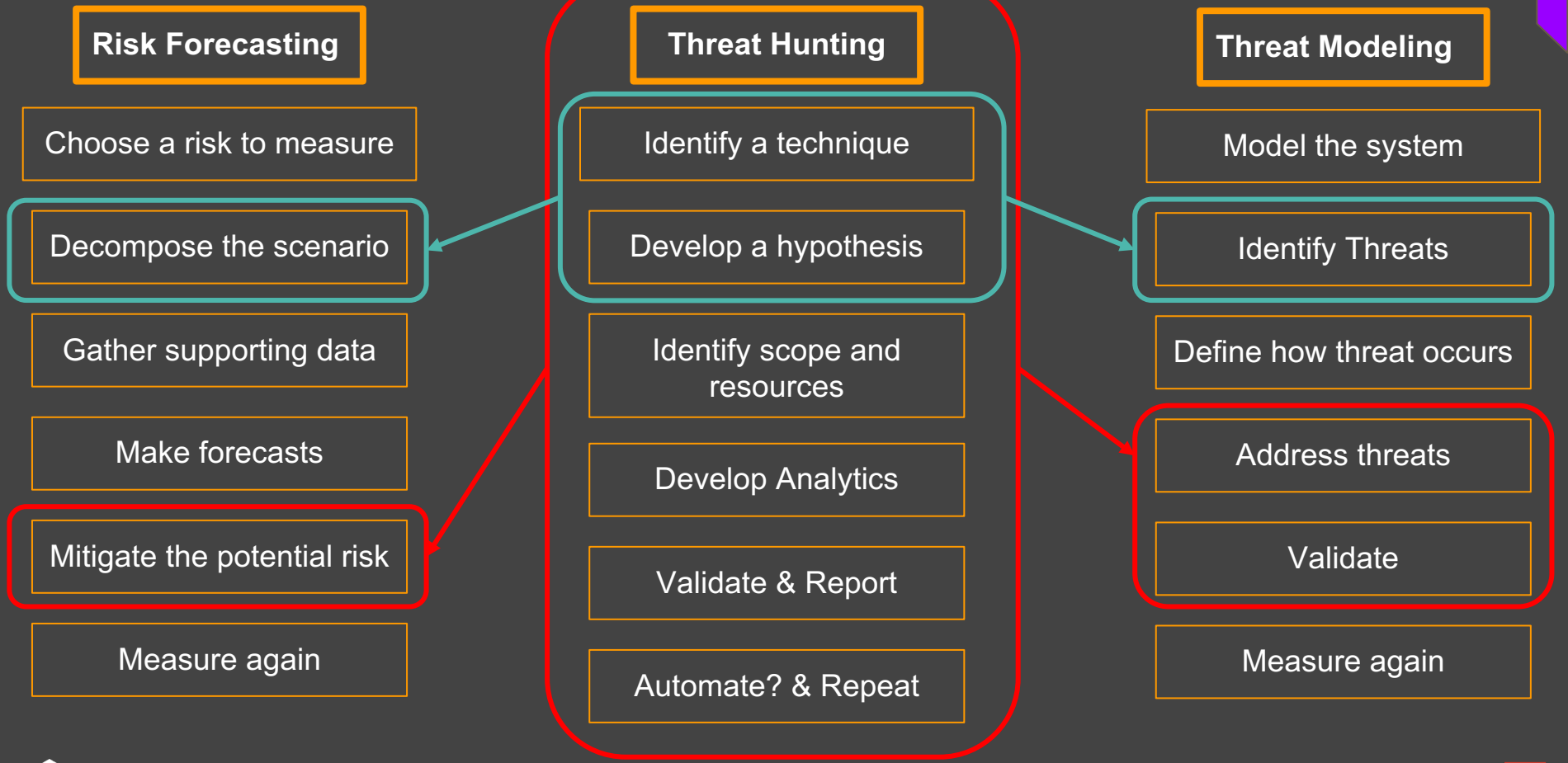
- Choose a risk to measure
- Decompose the scenario
- Gather supporting data
- Make forecasts
- Mitigate the potential risk
- Measure again

Threat Hunting

- Identify a technique
- Develop a hypothesis
- Identify scope and resources
- Develop Analytics
- Validate & Report
- Automate? & Repeat

Threat Modeling

- Model the system
- Identify Threats
- Define how threat occurs
- Address threats
- Validate
- Measure again


Risk Forecasting
Choose a risk to measure
Decompose the scenario
Gather supporting data
Make forecasts
Mitigate the potential risk
Measure again
Threat Hunting
Identify a technique
Develop a hypothesis
Identify scope and resources
Develop Analytics
Validate & Report
Automate? & Repeat
Threat Modeling
Model the system
Identify Threats
Define how threat occurs
Address threats
Validate
Measure again

Risk Forecasting

- Choose a risk to measure
- Decompose the scenario
- Gather supporting data
- Make forecasts
- Mitigate the potential risk
- Measure again

Threat Hunting

- Identify scope and resources
- Develop Analytics
- Validate & Report
- Automate? & Repeat

Threat Modeling

- Model the system
- Identify Threats
- Define how threat occurs
- Address threats
- Validate
- Measure again

What can we measure from a hunt detection?

We need to understand what we are trying to measure from a detection perspective

- Do we have the right resources to validate the detection of identified threats?
  - What percentage of my tools help the most during a hunt?
  - What percentage of data is utilized the most during a hunt?
- How much can we cover with the current resources we have?
  - Percentage of data in relation to detected techniques
  - Percentage of successful analytics for hunt engagements
- Are we reducing the probability of attackers achieving their objective?
  - Percentage reduced each quarter after a hunting engagement. forecasting?

ENTERPRISE ATT&CK

The practitioner's choice of knowledge base

MITRE said it best

“

MITRE’s Adversarial Tactics, Techniques and Common Knowledge (ATT&CK™) is a curated knowledge base and model for cyber adversary behavior, reflecting the various phases of an adversary’s lifecycle and the platforms they are known to target. ATT&CK is useful for understanding security risk against known adversary behavior, for planning security improvements, and verifying defenses work as expected.

”

- MITRE ATT&CK -

We *really* like Enterprise ATT&CK

What's not to like:

- it is threat-agnostic, describing the purpose and effect of many techniques
- contains more than 200 categorized and curated entries
- includes forensic artifacts and references to educate analysts and decrease barrier-to-entry
- techniques are cross-referenced by threat group, *if that's important to your business (it might not be, no judgement)*

ATT&CK STATISTICS (As of April 27, 2018)

- 219 techniques
  - 187 - Windows
  - 130 - MacOS
  - 108 - Linux
- 11 Tactics
- 68 groups
- 187 Tools
- 48 Data Sources
- 39 Contributors
- 21 Bypasses

What parts of ATT&CK are measurable?

Explore ATT&CK

Access Token Manipulation	
Technique	
ID	T1134
Tactic	Defense Evasion, Privilege Escalation
Platform	Windows
Permissions Required	User, Administrator
Effective Permissions	SYSTEM
Data Sources	API monitoring, Access Tokens

PowerShell	
Technique	
ID	T1086
Tactic	Execution
Platform	Windows
Permissions Required	User, Administrator
Data Sources	Windows Registry, File monitoring, Process command-line parameters, Process monitoring
Supports Remote	Yes

https://www.mitre.org/sites/default/files/publications/pr-18-0944-11-mitre-attack-design-and-philosophy.pdf

Explore ATT&CK

The lowest level of permissions the adversary is required

Access Token Manipulation

Technique

ID	T1134
Tactic	Defense Evasion, Privilege Escalation
Platform	Windows
Permissions Required	User, Administrator
Effective Permissions	SYSTEM
Data Sources	API monitoring, Access Tokens

PowerShell

Technique

ID	T1086
Tactic	Execution
Platform	Windows
Permissions Required	User, Administrator
Data Sources	Windows Registry, File monitoring, Process command-line parameters, Process monitoring
Supports Remote	Yes

https://www.mitre.org/sites/default/files/publications/pr-18-0944-11-mitre-attack-design-and-philosophy.pdf

Explore ATT&CK

The lowest level of permissions the adversary is required

Permissions an adversary will attain by performing the technique

Access Token Manipulation

Technique

ID	T1134
Tactic	Defense Evasion, Privilege Escalation
Platform	Windows
Permissions Required	User, Administrator
Effective Permissions	SYSTEM
Data Sources	API monitoring, Access Tokens

PowerShell

Technique

ID	T1086
Tactic	Execution
Platform	Windows
Permissions Required	User, Administrator
Data Sources	Windows Registry, File monitoring, Process command-line parameters, Process monitoring
Supports Remote	Yes

https://www.mitre.org/sites/default/files/publications/pr-18-0944-11-mitre-attack-design-and-philosophy.pdf

Explore ATT&CK

Access Token Manipulation	
	Technique
ID	T1134
Tactic	Defense Evasion, Privilege Escalation
Platform	Windows
Permissions Required	User, Administrator
Effective Permissions	SYSTEM
Data Sources	API monitoring, Access Tokens

The lowest level of permissions the adversary is required

Permissions an adversary will attain by performing the technique

Data recommended to be collected for the detection of an action

PowerShell	
	Technique
ID	T1086
Tactic	Execution
Platform	Windows
Permissions Required	User, Administrator
Data Sources	Windows Registry, File monitoring, Process command-line parameters, Process monitoring
Supports Remote	Yes

https://www.mitre.org/sites/default/files/publications/pr-18-0944-11-mitre-attack-design-and-philosophy.pdf

Explore ATT&CK

Access Token Manipulation

Technique

ID	T1134
Tactic	Defense Evasion, Privilege Escalation
Platform	Windows
Permissions Required	User, Administrator
Effective Permissions	SYSTEM
Data Sources	API monitoring, Access Tokens

The lowest level of permissions the adversary is required

Permissions an adversary will attain by performing the technique

Data recommended to be collected for the detection of an action

If the technique can be used to execute something on a remote system

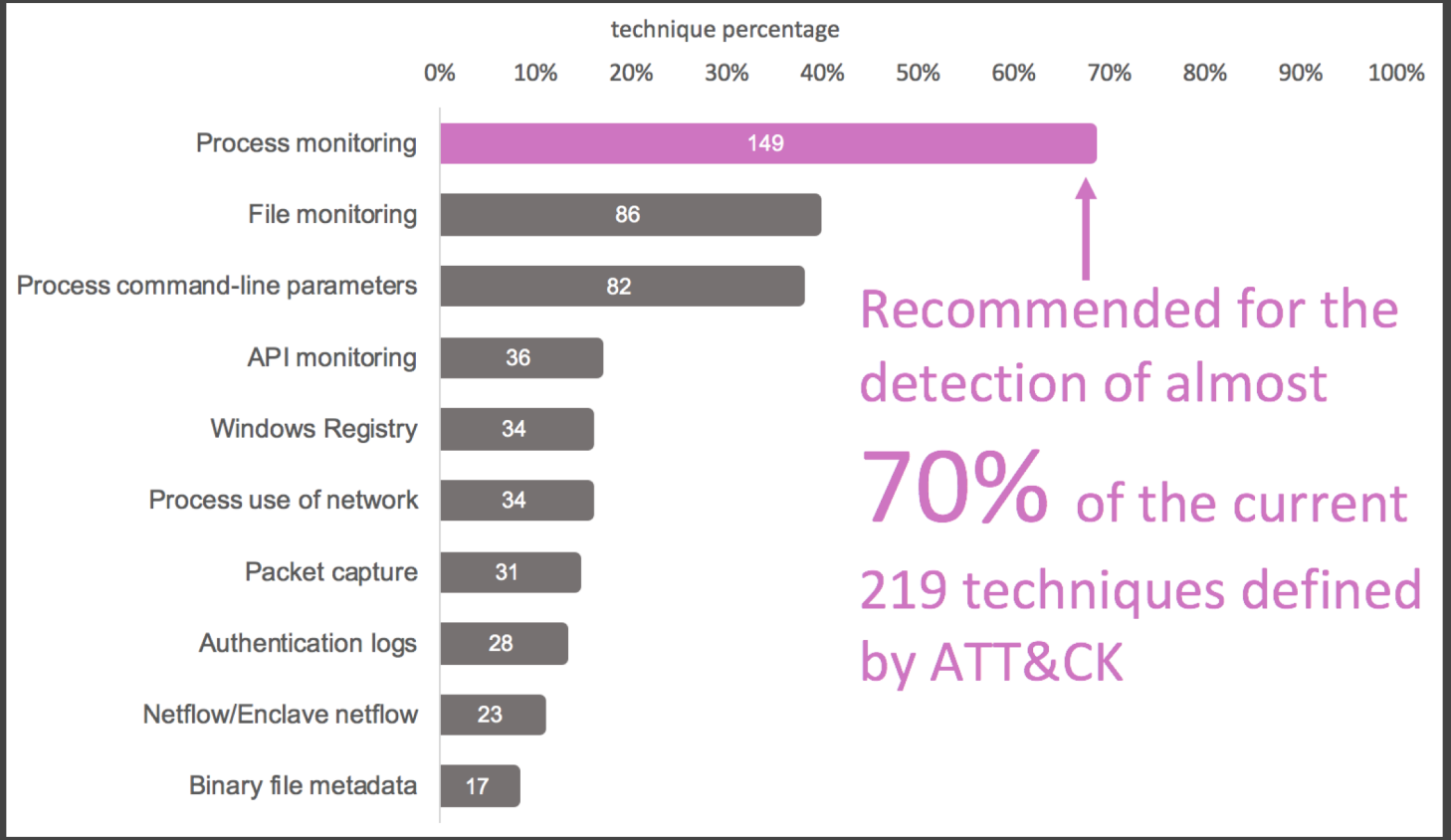
PowerShell

Technique

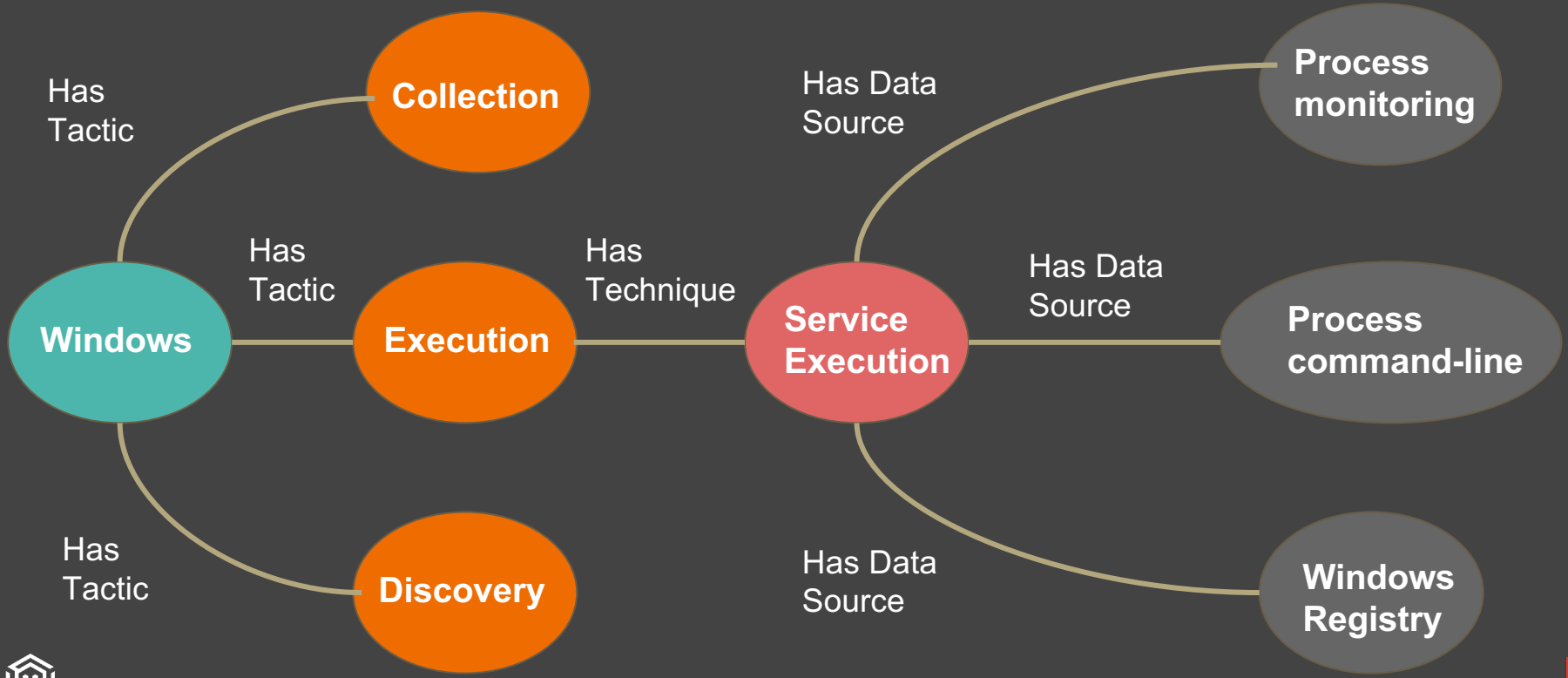
ID	T1086
Tactic	Execution
Platform	Windows
Permissions Required	User, Administrator
Data Sources	Windows Registry, File monitoring, Process command-line parameters, Process monitoring
Supports Remote	Yes

E.

Data Sources -> Adversarial Techniques



Identify Relationships in ATT&CK



Identify Relationships in ATT&CK

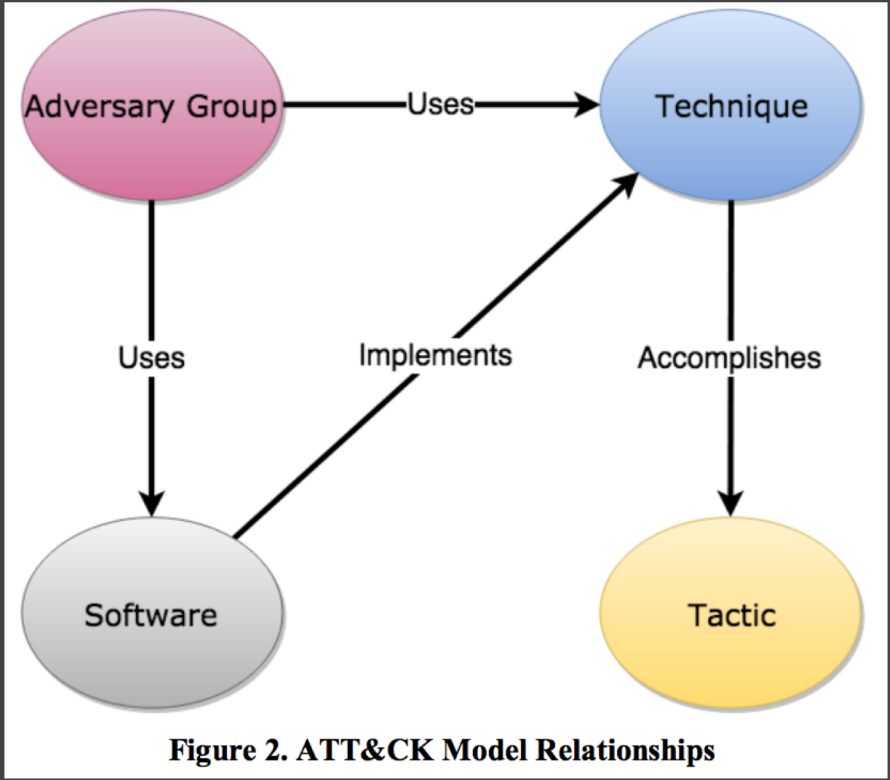
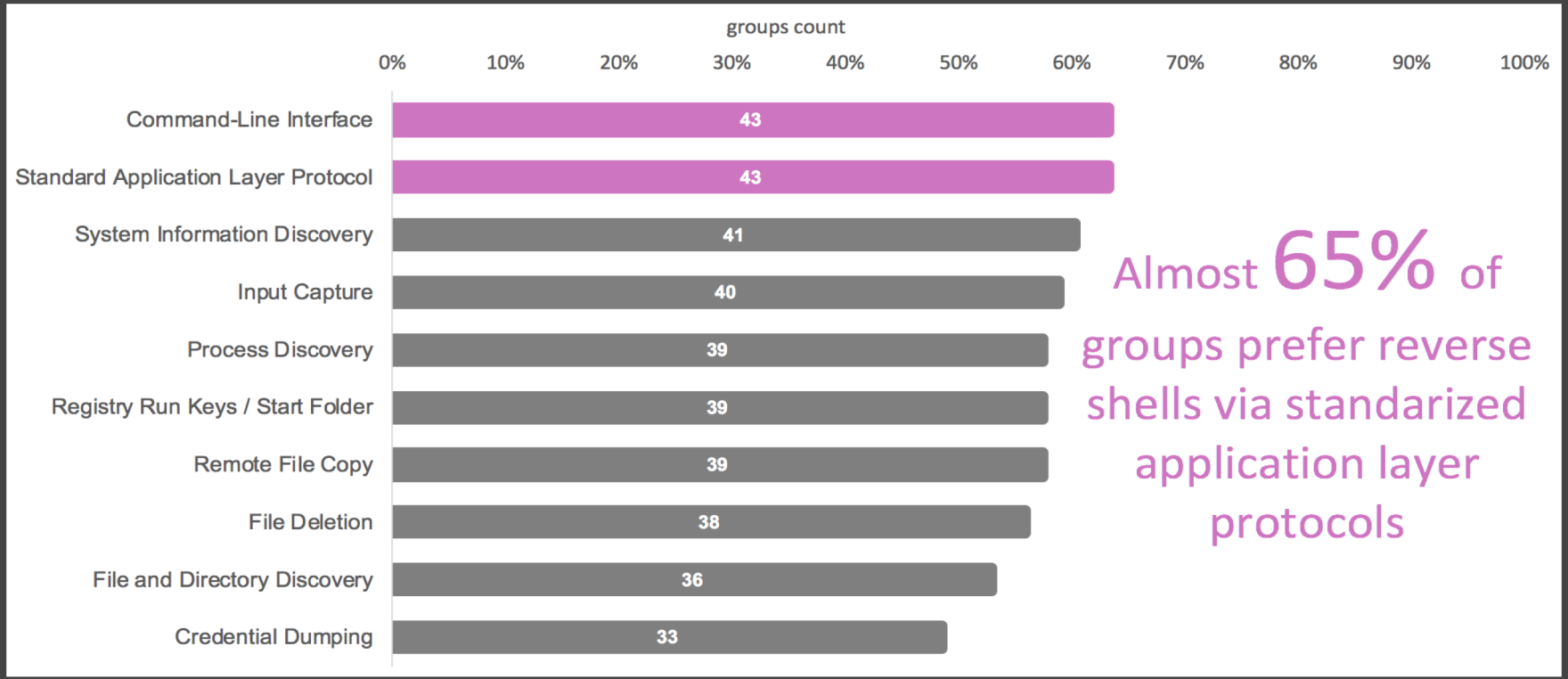


Figure 2. ATT&CK Model Relationships

https://www.mitre.org/sites/default/files/publications/pr-18-0944-11-mitre-attack-design-and-philosophy.pdf

Groups -> Adversarial Techniques



Almost 65% of groups prefer reverse shells via standarized application layer protocols

MITRE has already covered this topic, though

Part 1: Using ATT&CK to Advance Cyber Threat Intelligence

This excellent blogpost by Katie Nickels (@likethecoins) covers:

- An overview of traditional CTI
- Challenges
- How ATT&CK can help provide a way of expressing TTPs, exposing a *common language*
- Using ATT&CK to understand blind spots
- Using TTP counts as a metric to justify your CTI program

Part 2: Using ATT&CK to Advance Cyber Threat Intelligence

The second part in this series focuses on knowledge management and adversary behavior curation, which ATT&CK is perfectly designed to assist with.

Two of the major points to take away:

- Get as close to original information as possible to avoid misinterpreting a tactic or event
- Select *appropriate* information to curate

Hunt team staffing?

Do it with the team you have, outsource or a blend

Considerations when staffing a hunt team

- Frequency of hunt exercises. How often your organization plans on hunting will determine resourcing plans.
- Skillsets and tradecraft experience. It goes without saying that there is not one type of ‘hunter’ profile out there. The team will be made up of varying skillsets (sys admin, db admin, soc analyst, network admin, etc).
- Is this something that you should consider outsourcing to a managed service company? Proactive Hunt services are offered by several leading Security managed services companies and can provide instantaneous results.

Evolving the program - Useful info to consider

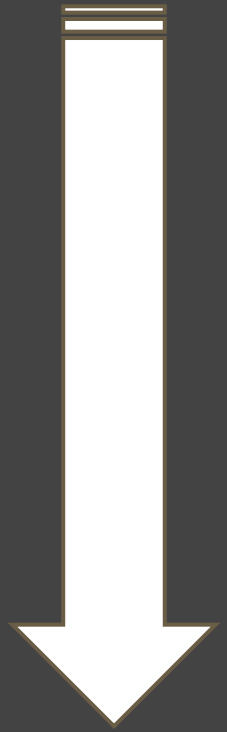
What data sources are recommended?

Access Tokens	Detonation chamber	Loaded DLLs
Anti-virus	Digital Certificate Logs	Mail server
API monitoring	DLL monitoring	Malware reverse engineering
Application Logs	DNS records	MBR
Asset Management	EFI	Named Pipes
Authentication logs	Email gateway	Netflow/Enclave netflow
Binary file metadata	Environment variable	Network device logs
BIOS	File monitoring	Network intrusion detection system
Browser extensions	Host network interface	Network protocol analysis
Data loss prevention	Kernel drivers	Packet capture

What data sources are recommended?

PowerShell logs
Process command-line parameters
Process monitoring
Process use of network
Sensor health and status
Services
SSL/TLS inspection
System calls
Third-party application logs
User interface

VBR
Web application firewall logs
Web logs
Web proxy
Windows Error Reporting
Windows event logs
Windows Registry
WMI Objects



You don’t just need data, you need the right data...

Let's take a look at data sources again:

PowerShell logs
Process command-line parameters
Process monitoring
Process use of network
Sensor health and status
Services
SSL/TLS inspection
System calls
Third-party application logs
User interface

Process object attributes…

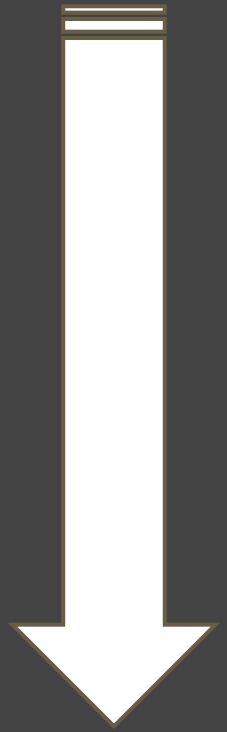
PowerShell logs
Process command-line parameters
Process monitoring
Process use of network
Sensor health and status
Services
SSL/TLS inspection
System calls
Third-party application logs
User interface

Process
process_name
process_command_line
process_path
process_parent_name
user_name
hash_sha256
host_name

So, what can I measure *now*?

Do I know what I have?

Is this data what I need?



Not all data sources are created equal, data quality matters.

If data needed for a hunting engagement does not meet specific requirements defined by the hunt team, then the data is not considered quality data since it is affecting the intended purpose of it.

“Data are of high quality if they are fit for their intended uses in operations, decision making and planning."

”

- Julian’s Quality Handbook -

Threat Hunting vs Detection

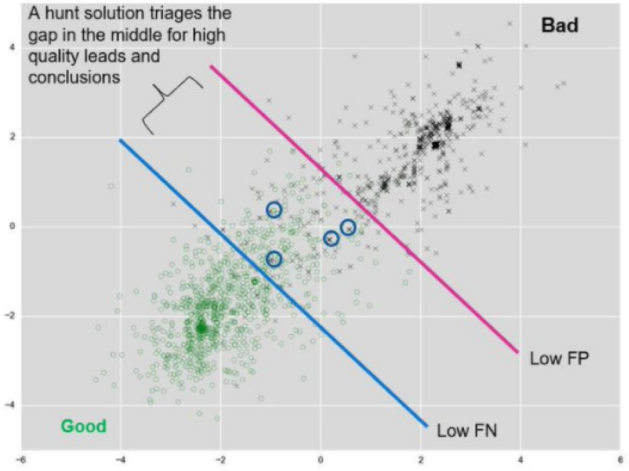
Chris Gerritz @gerritzc · Mar 15

Why most prevention and detection tools make poor #ThreatHunting solutions in one slide. @InfocyteInc

Threat Hunting vs Detection - The Optimization Problem

Why do most defensive tools make poor hunt tools?

- Prevention and real-time detection solutions are **optimized for low False Positive (FP) Alerting**
- Hunt solutions are **optimized for low False Negatives (FN)**
  - For Hunting: Anomalies, Outliers, and Suspicious Activity are **leads**, not FPs to be tuned out
  - A good hunt solution sorts and scores leads, then enables a quick path to verify and investigate to a conclusion



Precision is being tolerant of False Positives

True Positive - a malicious thing you correctly identify as malicious

False Positive - a benign thing you incorrectly identify as malicious

Precision = (True Positives/(True Positives + False Positives))

Example:

- 100 events
- 74 TPs
- 26 FPs
- 0.74 precision

Recall is how well you find malicious activity

True Positive - a malicious thing you correctly identify as malicious

False Negative - a malicious thing you incorrectly identify as benign

Recall = (True Positives/(True Positives + False Negatives))

Example:

- 100 events
- 55 TPs
- 21 FPs
- 24 FNs
- 0.69 recall

Data Quality Dimensions

Data Quality	Characteristics Description	Example Metric
Accuracy	A quality of that which is free of error. A qualitative assessment of freedom from error, with a high assessment corresponding to a small error. (FIPS Pub 11-3)	Percent of values that are correct when compared to the actual value. For example, M=Male when the subject is Male.
Completeness	Completeness is the degree to which values are present in the attributes that require them. (Data Quality Foundation)	Percent of data fields having values entered into them.
Consistency	Consistency is a measure of the degree to which a set of data satisfies a set of constraints. (Data Quality Management and Technology)	Percent of matching values across tables/files/records.
Timeliness	As a synonym for currency, timeliness represents the degree to which specified data values are up to date. (Data Quality Management and Technology)	Percent of data available within a specified threshold time frame (e.g., days, hours, minutes).
Uniqueness	The state of being the only one of its kind. Being without an equal or equivalent.	Percent of records having a unique primary key.
Validity	The quality of data that is founded on an adequate system of classification and is rigorous enough to compel acceptance. (DoD 8320.1-M)	Percent of data having values that fall within their respective domain of allowable values.

http://mitiq.mit.edu/ICIQ/Documents/IQ%20Conference%201996/Papers/DODGuidelinesonDataQualityManagement.pdf

Data Completeness

- How much data that is required/needed is available in my network?
- Are all required/needed data fields and values recorded?

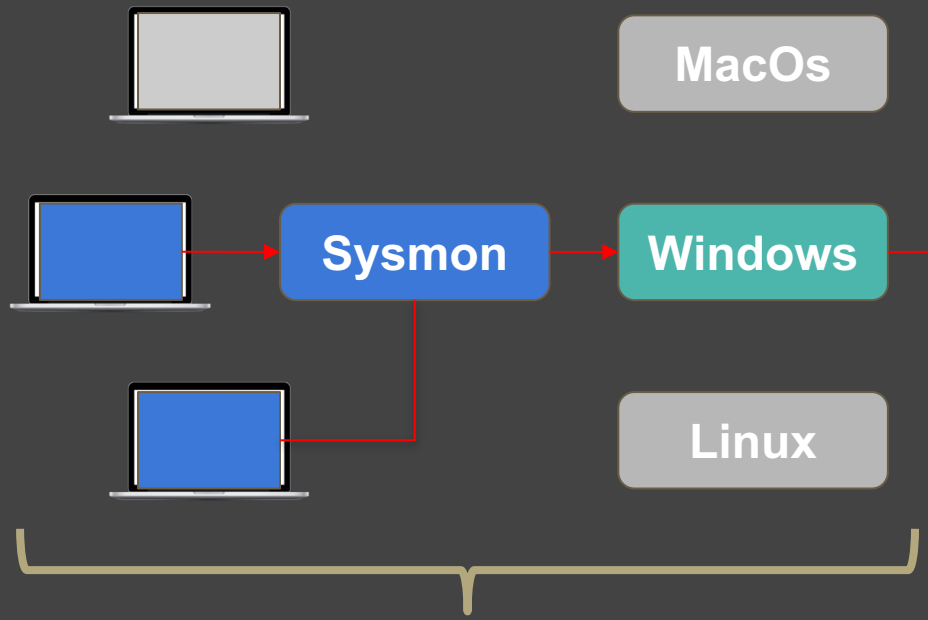
Data Consistency

- Can we match required/needed fields across data sources?

Data Timeliness

- Does my data represent reality?
- How far back in time can I hunt with required/needed data?

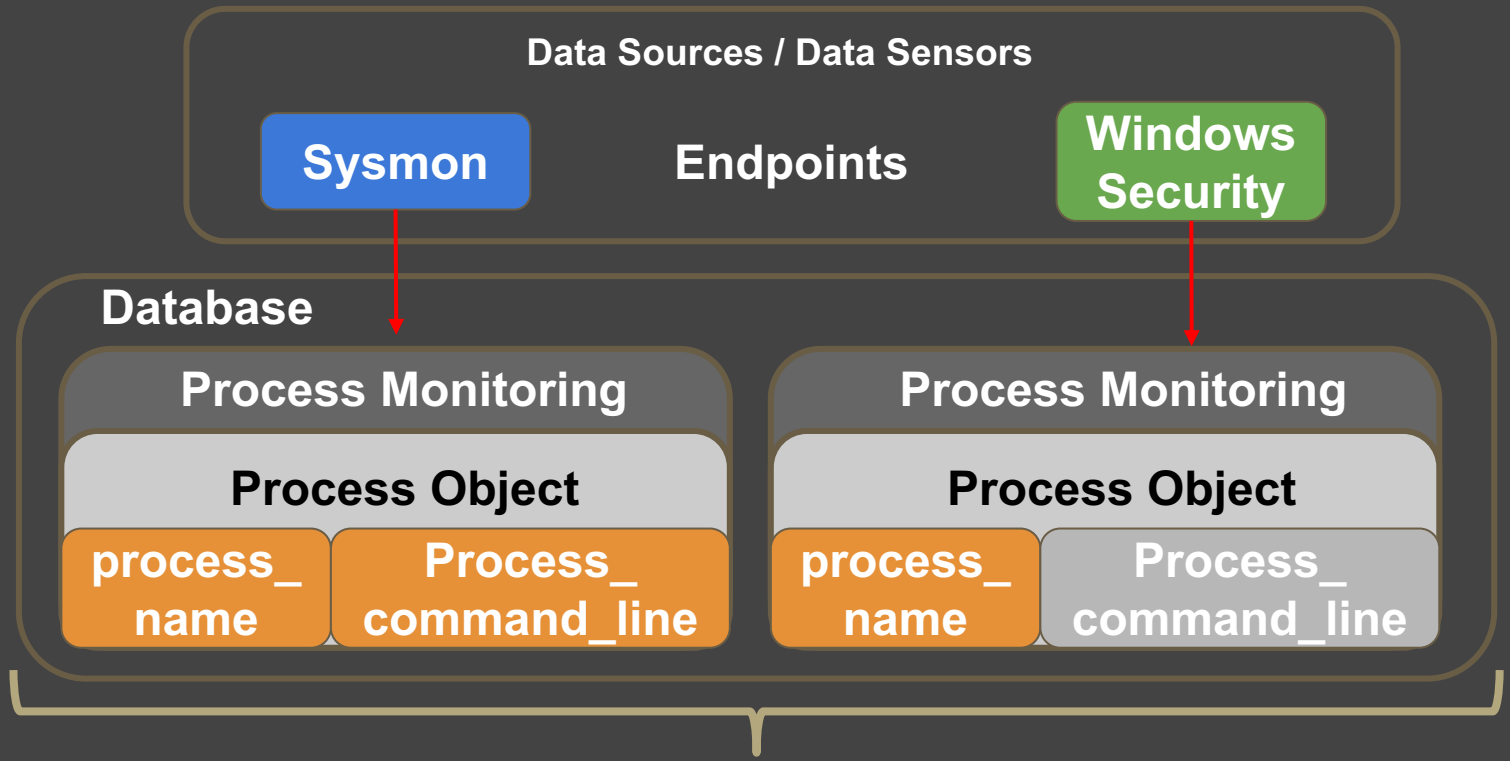
Completeness: Percentage of network covered?



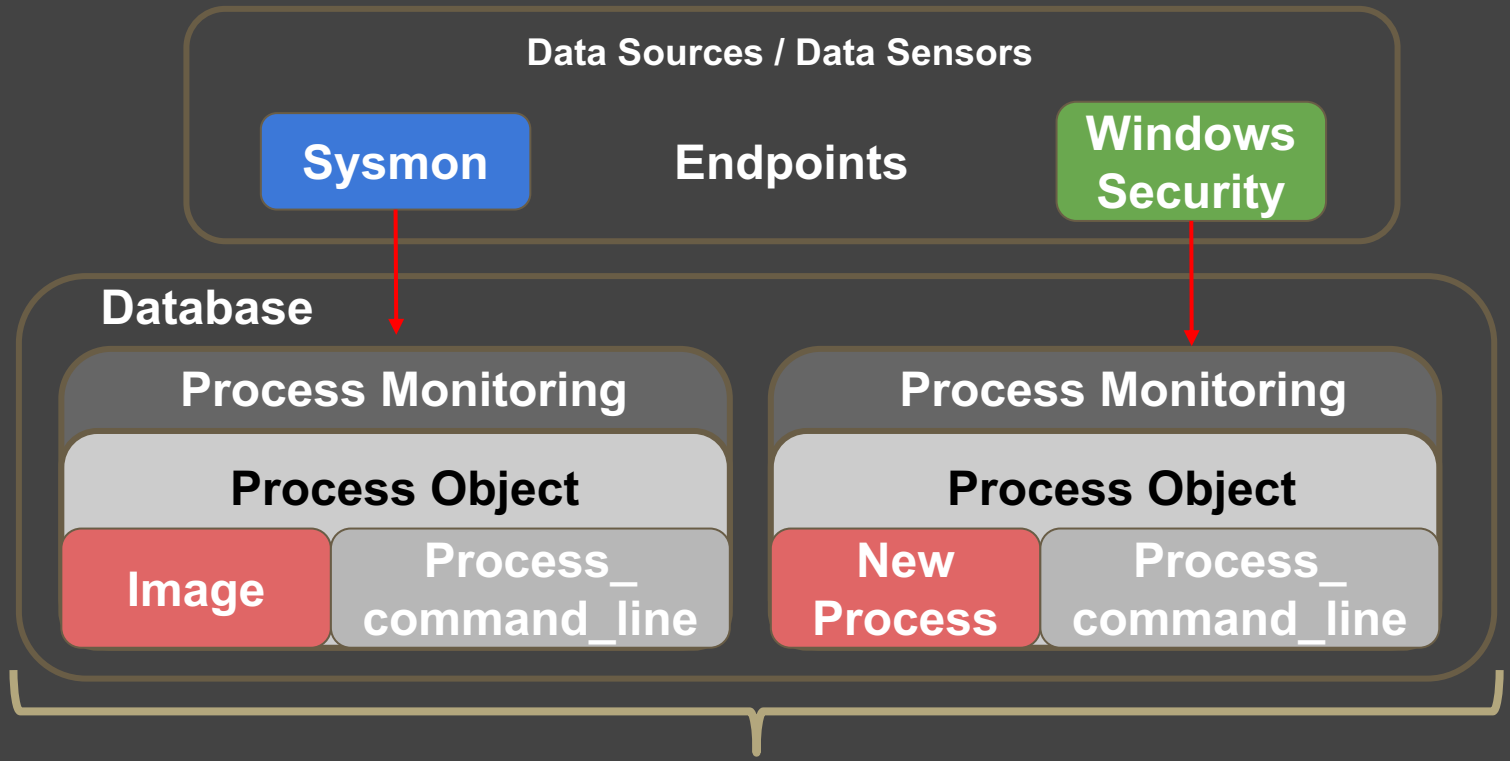
Process Object
process_name
process_command_line
process_path
process_parent_name
user_name
hash_sha256
host_name

Data Completeness

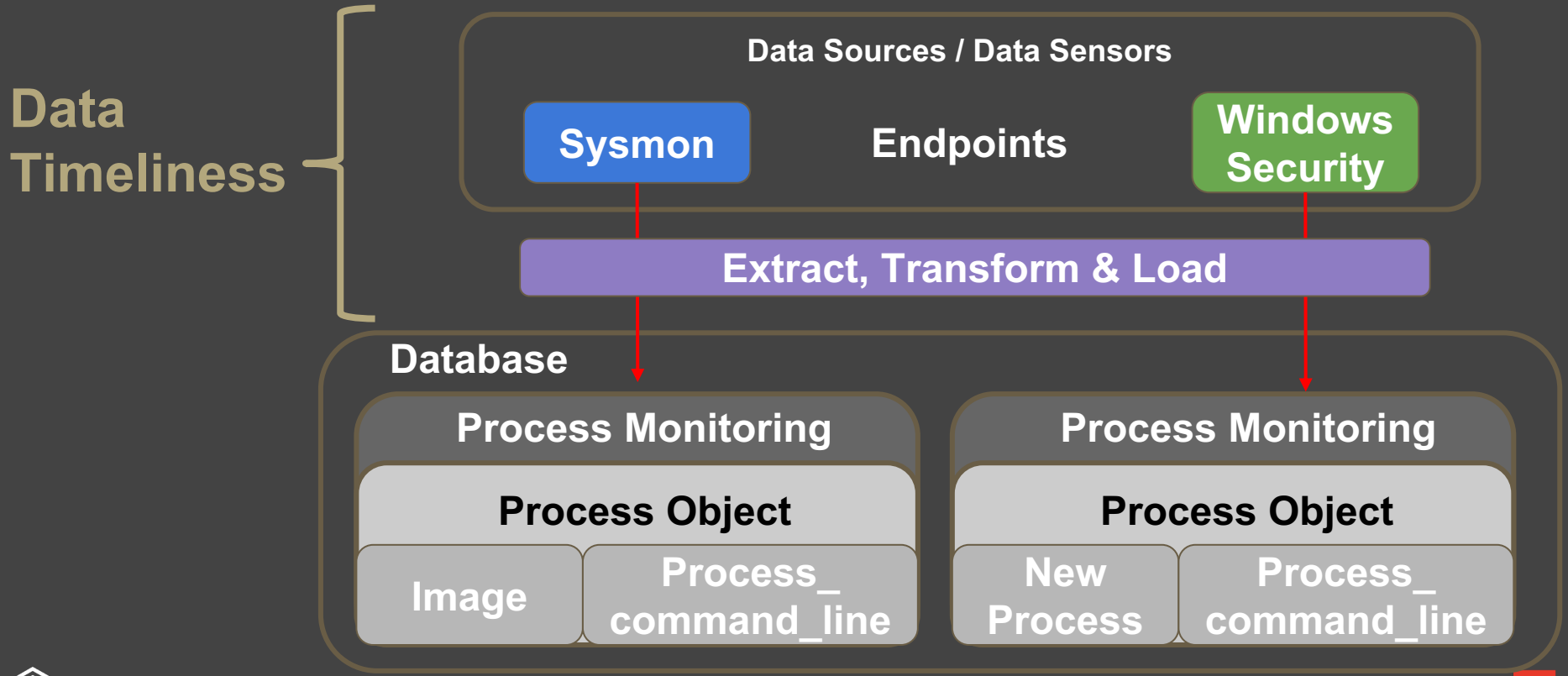
Completeness: Is the expected data complete?



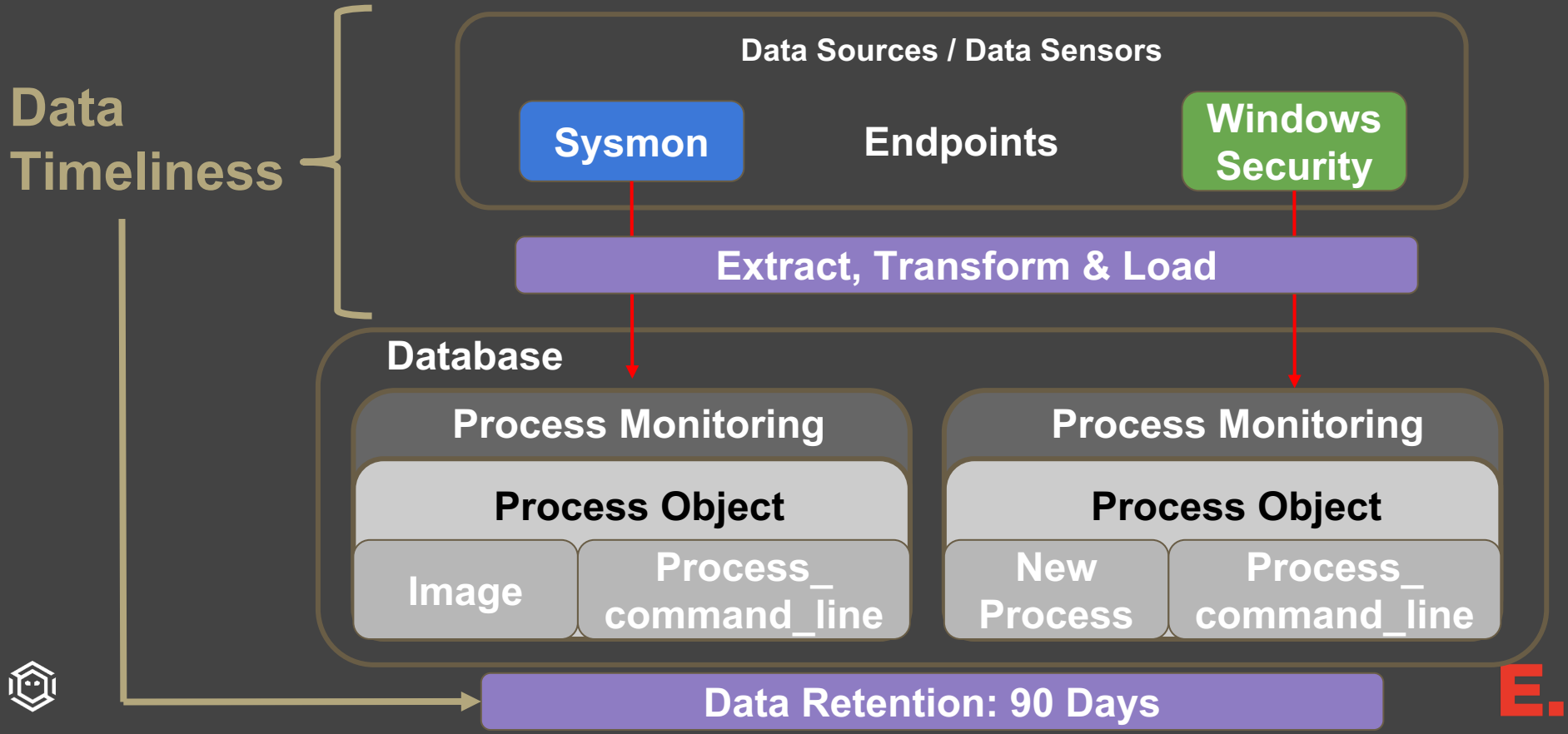
Consistency: Consistency across all data sources?



Timeliness: Does my data represent reality?



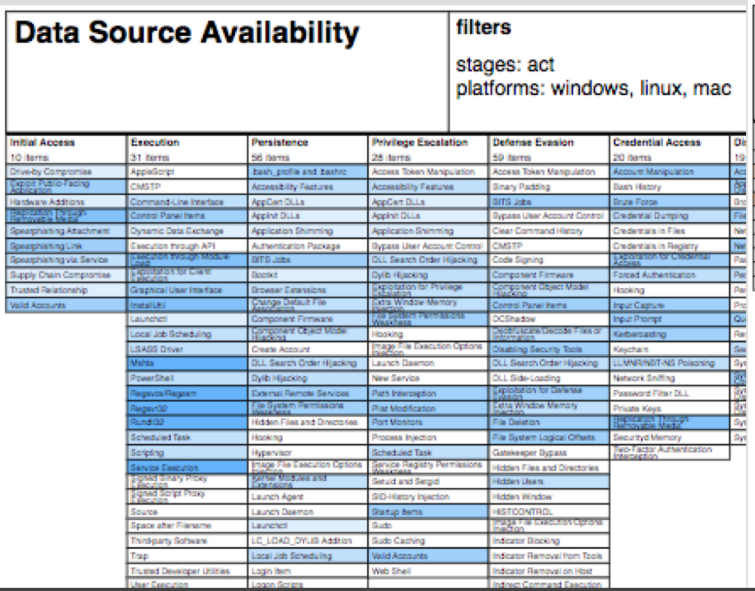
Timeliness: Does my data represent reality?

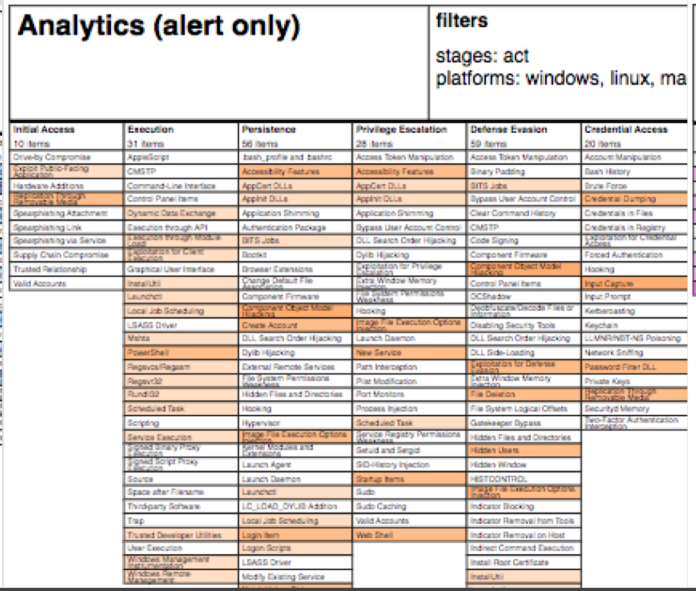


A few hunt metrics you *could* measure:

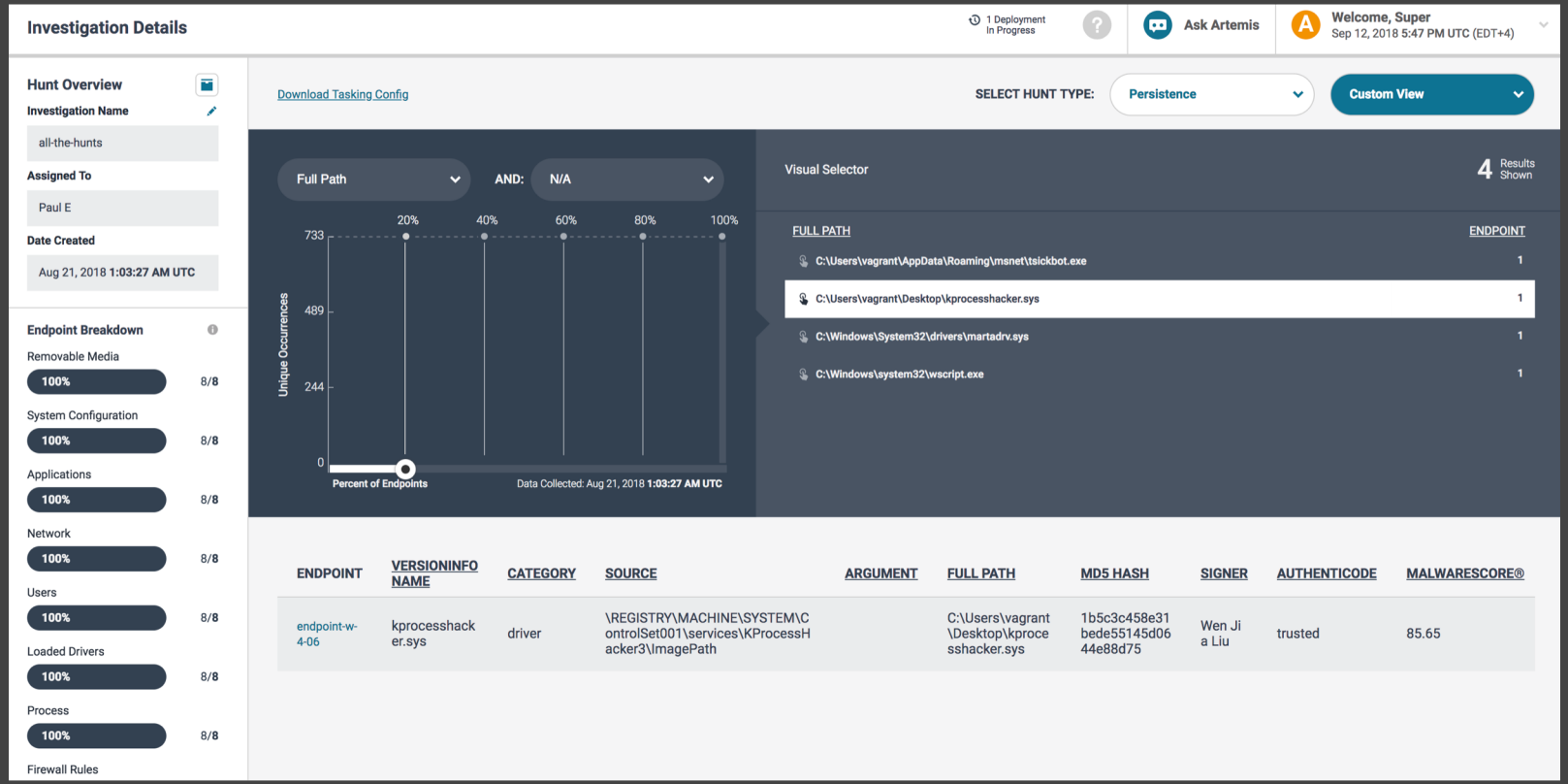
- What percentage of recommended data is available for a hunt?
- What percentage of the expected data is complete for a hunt?
- What percentage of my environment could I cover in an hunt based on the available recommend data?
- How far back in time can I hunt with recommended data?
- What percentage of my data sources are consistent across all the data provided by data sensors?
- Do I have the right technology or skills to hunt?

We visualize each type of thing *separately*





Snapshot of Endgame hunt window



Endgamers' support in the InfoSec community

- Endgame blogs (just a recent few):
  - https://www.endgame.com/blog/technical-blog/introducing-ember-open-source-classifier-and-dataset
  - https://www.endgame.com/blog/technical-blog/opening-machine-learning-black-box-model-interpretability
  - https://www.endgame.com/blog/technical-blog/introducing-endgame-red-team-automation
  - https://www.endgame.com/blog/executive-blog/endgame-presents-hacker-summer-camp-2018

Closing thoughts

If this is something of interest to your organization, wonderful! Come talk to me after.

Thank you

This is an appendix

All this is stuff we wanted you to have

Quick wins

Count	Technique
92	Remote File Copy
92	Standard Application Layer Protocol
91	Command-Line Interface
85	System Information Discovery
75	File and Directory Discovery
70	Credential Dumping
68	Process Discovery
67	Registry Run Keys /Start Folder
62	File Deletion
57	Input Capture

# Techniques	Data Source
149	Process monitoring
86	File monitoring
82	Process command-line parameters
36	API monitoring
34	Windows Registry
34	Process use of network
31	Packet capture
28	Authentication logs
23	Netflow/Enclave netflow
17	Binary File Metadata

Data Sources -> Adversary Techniques

# Techniques	Name
149	Process monitoring
86	File monitoring
82	Process command-line parameters
36	API monitoring
34	Windows Registry
34	Process use of network
31	Packet capture
28	Authentication logs
23	Netflow/Enclave netflow

# Techniques	Name
17	Binary file metadata
16	DLL monitoring
16	Network protocol analysis
14	Windows event logs
12	Loaded DLLs
9	System calls
8	SSL/TLS inspection
8	Malware reverse engineering
6	Anti-virus

Data Sources -> Adversary Techniques

# Techniques	Name
6	Data loss prevention
5	Application Logs
4	Network device logs
4	Windows Error Reporting
4	Network intrusion detection system
4	User interface
4	Web proxy
3	Kernel drivers
3	Services

# Techniques	Name
3	Email gateway
3	Third-party application logs
2	Mail server
2	Detonation chamber
2	MBR
2	Environment variable
2	BIOS
2	Host network interface
1	Web logs

Data Sources -> Adversary Techniques

# Techniques	Name
1	Asset Management
1	Web application firewall logs
1	EFI
1	DNS records
1	Browser extensions
1	Sensor health and status
1	Named Pipes
1	VBR
1	PowerShell logs

# Techniques	Name
1	Access Tokens
1	Digital Certificate Logs
1	WMI Objects

Reference: assessing data visibility

Understanding overall coverage of ATT&CK is related to understanding data availability:

1. Assess data sources (there are 48 in ATT&CK) across the organization
    a. For each data object, document whether you have access to it either centrally or distributed
        i. Take the time to document data object properties and attributes (ask your vendors to help)
    b. Document fleet coverage for each applicable OS
2. Map data sources and attributes to a common data model as defined by MITRE's CAR or OSSEM
    a. Data object -> Data source -> Mapped sensor(s)
        i. If there are operating systems you can't cover, document that
3. Determine the quality of each data object mapped to sources of evidence
    a. Note the longevity of every source of evidence you intend to use
    b. If a source of evidence requires transformation to be usable, note that as well
    c. Use the DoD scale to determine each of 6 metrics

What is a data model?

- A data model basically determines the structure of data and the relationships identified among each other.
- MITRE Data Model:
  - Strongly inspired by CybOX, is an organization of the objects that may be monitored from a host-based or network-based perspective.
  - https://car.mitre.org/wiki/Data_Model
- STIX™ Version 2.0. Part 4: Cyber Observable Objects
  - http://docs.oasis-open.org/cti/stix/v2.0/stix-v2.0-part4-cyber-observable-objects.html

Data Model (Defining Data Objects)

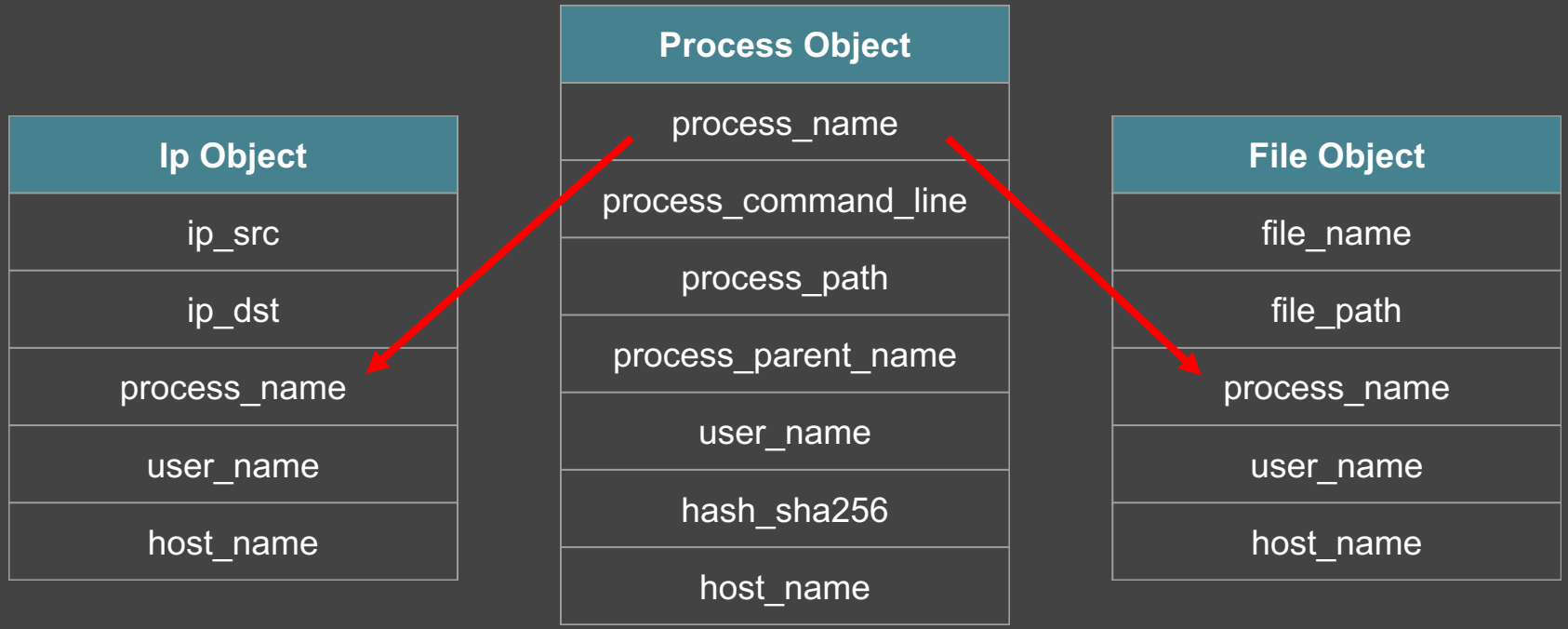
Ip Object
ip_src
ip_dst
process_name
user_name
host_name

Process Object
process_name
process_command_line
process_path
process_parent_name
user_name
hash_sha256
host_name

File Object
file_name
file_path
process_name
user_name
host_name

https://github.com/Cyb3rWard0g/OSSEM/tree/master/data_models

Data Model (Defining Data Objects)



Data Model (Defining Object Relationships)

Applicable Objects (Source)	Relationship	Applicable Objects (Destination)
Process	Created	File, Process, Win Registry Key, Service
File, Process, Win Registry Key, Service	Created_By	Process
Process	Parent_Of	Process
Process	Modified_Properties_Of	File, Win Registry Key, Service
Process	Renamed	File
File	Renamed_By	Process
Process	Connected_To	IP, Hostname

http://stixproject.github.io/data-model/1.2/cyboxVocabs/ObjectRelationshipVocab-1.1/

Example: Process use of network

PowerShell logs
Process command-line parameters
Process monitoring
Process use of network
Sensor health and status
Services
SSL/TLS inspection
System calls
Third-party application logs
User interface

VBR
Web application firewall logs
Web logs
Web proxy
Windows Error Reporting
Windows event logs
Windows Registry
WMI Objects

Process use of network: Process & IP Relationship

Applicable Objects (Source)	Relationship	Applicable Objects (Destination)
Process	Created	File, Process, Win Registry Key, Service
Process	Parent_Of	Process
File, Process, Win Registry Key, Service	Created_By	Process
Process	Modified_Properties_Of	File, Win Registry Key, Service
Process	Renamed	File
File	Renamed_By	Process
Process	Connected_To	IP, Hostname

http://stixproject.github.io/data-model/1.2/cyboxVocabs/ObjectRelationshipVocab-1.1/

Data Source: Process use of network

Process Object
process_name
process_command_line
process_path
process_parent_name
user_name
hash_sha256
host_name

Ip Object
ip_src
ip_dst
process_name
user_name
host_name

Linking it to data sensors (Sysmon)

Process Object
process_name
process_command_line
process_path
process_parent_name
user_name
hash_sha256
host_name

Sysmon (1)
Image
Command_line
User
Hashes
ParentImage

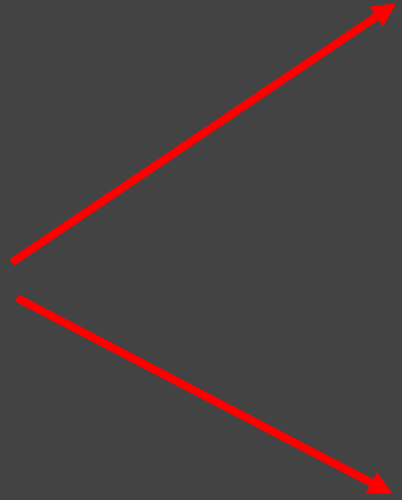
Linking it to data sensors (Windows Security)

Process Object
process_name
process_command_line
process_path
process_parent_name
user_name
hash_sha256
host_name

Windows Security (4688)
NewProcessName
CommandLine
SubjectUserName
ParentProcessName

Do I have what I need?

Process Object
process_name
process_command_line
process_path
process_parent_name
user_name
hash_sha256
host_name



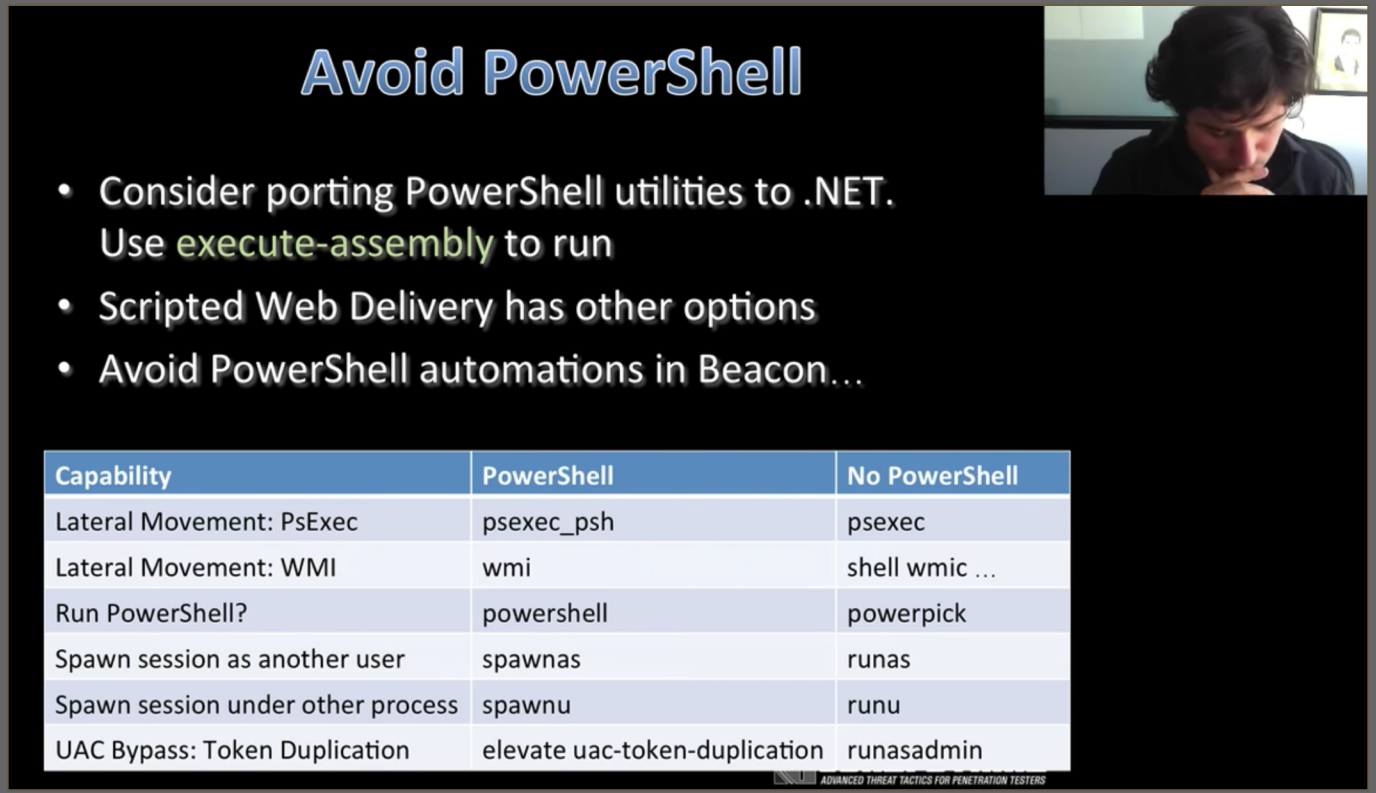
Sysmon
Image
Command_line
User
Hashes
ParentImage

Windows Security (4688)
NewProcessName
CommandLine
SubjectUserName
ParentProcessName

Some guidelines:

1. Keyword searches for process names or network locations aren't hunting and should be automated
2. Know whether a technique is best detected on its own or with other techniques
3. Don't try to score techniques or try to categorize on a sophistication scale - the things that work succeed whether you respect their novelty or not
    a. *Everyone uses PsExec equally*
4. You can't measure things you don't know about
    a. Attack variations
5. Tools, with a small number of exceptions, are not techniques
    a. PowerShell, fwiw, falls into this weird characterization
        i. Detecting PowerShell abuse involves seeing dozens of variations (not just being able to tell "powershell.exe" ran)
            1. Oh, and remember it can be invoked easily through module load and a half-dozen other methods
                a. So *anything* can be PowerShell….***anything***.

Fighting the toolset.. Doing things differently..



Capability	PowerShell	No PowerShell
Lateral Movement: PsExec	psexec_psh	psexec
Lateral Movement: WMI	wmi	shell wmic …
Run PowerShell?	powershell	powerpick
Spawn session as another user	spawnas	runas
Spawn session under other process	spawnu	runu
UAC Bypass: Token Duplication	elevate uac-token-duplication	runasadmin

https://www.youtube.com/watch?time_continue=1&v=RoqVunX_sqA

Defining ATT&CK based analytics

MITRE researchers categorized the ATT&CK-related analytics into four major types:

- **Behavioral** – An analytic to detect a specific adversary behavior
- **Situational Awareness** – what is occurring within a network environment at a given time. Not all analytics need to be geared towards generating alerts
- **Anomaly/Outlier** – Analytics that may detect behavior that is not malicious, but which is unusual and may be suspect
- **Forensic** – Analytics that are most useful when conducting an investigation

CAR Analytics List:
https://car.mitre.org/wiki/Full_Analytic_List
Finding Cyber Threats with ATT&CK based analytics:
https://www.mitre.org/sites/default/files/publications/16-3713-finding-cyber-threats%20with%20att%26ck-based-analytics.pdf

Scoring Table (Basic Example)

		Technology	Talent	Data Quality			Detection
Definition	Score	Hunt Tool	Skills	Completeness	Timeliness	Consistency	capabilities
None	0	I don't know what tools I have or need	A team might not even exist You might be recruting	don't know / not documented	don't know / not documented	don't know / not documented	no ability to detect
Poor	1	Centralizing data across several other tools (Splunk, ELK, EDR, etc). Tools allowing you to run basic queries in order to make sense of the data	Your team focuses only on creating signatures or basic correlation rules to detect IOCs from intel reports (IOC Sweeps)	**Endpoint Coverage** - 0-25 % **Missing Data** - Required data (values or fields) is missing 75%-100%.	**Data Retention -** 0%-25% of the time needed or defined by the organization.	**Standard Field Names** - data standardization only 0-25% across all data sources.	limited ability to detect variations of the technique using basic signatures
Fair	2	Hunters might be running queries and still get a very high amount of events that still need to be analyzed.	Your team identifies the value of ATT&CK beyond data availability.	**Endpoint Coverage** - 25% - 50% **Missing Data** - Required data (values or fields) is missing 50%-75%.	**Data Retention -** 25%-50% of the time needed or defined by the organization.	**Standard Field Names** - data standardization only 25-50% across all data sources.	analytics for some variations of a technique leverage relationships (IOCs, IOAs) over signatures
Good	3	Here is where you star using a few basic Data Science capabilities provided by your tools (i.e ELK Enterprise) Better Automation.	Better understanding of the environment and has documented several parts of the network already.	**Endpoint Coverage** - 50% - 75% **Missing Data** - Required data (values or fields) is missing 25%-50%.	**Data Retention -** 50%-75% of the time needed or defined by the organization.	**Standard Field Names** - data standardization only 50-75% across all data sources.	analytics for most techniques w/ introduction of data-driven capabilities (statistical detections such as outlier analyses and behavioral analytics)
Very Good	4	Here is where the capability of using advanced data science techniques are posible. If you can validate the detection of an adversay technique by just applying basic data anlytics, you might be already in the **"Very Good"** level.	Your team already understands the data sources available to hunt. Better institutional knowledge	**Endpoint Coverage** - 75% - 100% **Missing Data** - Required data (values or fields) is missing 0%-25%.	**Data Retention -** 75%-100% of the time needed or defined by the organization.	**Standard Field Names** - data standardization only 75-100% across all data sources.	analytics for most techniques w/ introduction of data-driven capabilities (statistical detections such as outlier analyses and behavioral analytics)
Excellent	5	Hunt tools helping to automate several procedures expediting the time of analysis and hunting Tool integrating hunting with other security procedures (i.e. Incident Response)	Your team already understand the environment very well and has complete documentation of the network. Helping either the vendor or internal data scientists to improve detections	**Endpoint Coverage** - 75% - 100% **Missing Data** - Required data (values or fields) is missing 0%-25%.	**Data Retention -** 75%-100% of the time needed or defined by the organization.	**Standard Field Names** - data standardization only 75-100% across all data sources.	analytics for all known variations of a technique with data-driven and machine-driven (ML) analytics

An alternative approach to be careful using

Adversary-focused:

- Look at the techniques associated with a group like WINNTI (G0044)
- This group has three techniques listed
  - Process Discovery (T1057); all OS
    - Process monitoring
    - Process command-line parameters
  - Rootkit (T1014); all OS
    - BIOS
    - MBR
    - System calls
  - Code Signing (T1116); MacOS and Windows
    - Binary metadata

This approach makes sense and may impact how quickly you obtain coverage; and relies on incomplete data.

We like this approach less than an adversary-focus

Technique-focused:

- Choose a single technique, like Account Manipulation (T1098)
- Verify the status of each data source, mapping sources and attributes to one or more sensors
  - Authentication logs
  - API monitoring
  - Windows event logs
  - Packet capture
- Assess the quality of each data source, as well as completeness and timeliness
- Begin developing analytics for the technique as previously discussed

This approach involves multiple forms of duplicate effort and does not scale very well beyond a small number of techniques.