Elastic Common Schema for Cyber Threat Hunting

Mike Paquette
Director of Product, Security Market
April 11, 2019

SECURITY ANALYTICS

Solutions

Automated & Interactive Security Analytics

Threats don't follow templates. Neither should you. The Elastic Stack gives you the edge you need to keep pace with the adversaries of today and tomorrow. Here's how.

Security Events Logging: Learn how Bell Canada turned to Elastic to streamline alerts, deepen log analysis, and address challenges unique to being an ISP. Watch Now

NEW Ingest disparate data types from diverse sources with the open source Elastic Common Schema (ECS) for defining a common set of fields. Learn More

Agenda

elastic

Elastic Common Schema

elastic

elastic

Elastic Common Schema

Background

To define a common set of document fields (and their respective field names) to be used in events (logs and metrics) stored in Elasticsearch as part of any logging or metrics use case of the Elastic Stack, including IT operations analytics, security analytics, and APM.

https://github.com/elastic/ecs

V1.0.0 – GA

Elastic Common Schema (ECS)

Normalize Data to Streamline Analysis

Defines a **common** set of fields and objects to ingest data into Elasticsearch

Enables **cross-source analysis** of diverse data

Designed to be **extensible**

ECS 1.0 is in GA. The Elastic Stack is being transitioned to **default** to ECS

https://github.com/elastic/ecs
Contributions & feedback welcome

Source fields

Source fields describe details about the source of a packet/event.

Source fields are usually populated in conjunction with destination fields.

Field	Description
source.address	Some event source addresses are defined ambiguously. The event will sometimes list an IP, a domain or a unix socket. You should always store the raw address in the `.address` field. Then it should be duplicated to `.ip` or `.domain`, depending on which one it is.
source.ip	IP address of the source. Can be one or multiple IPv4 or IPv6 addresses.
source.port	Port of the source.
source.mac	MAC address of the source.
source.domain	Source domain.
source.bytes	Bytes sent from the source to the destination.
source.packets	Packets sent from the source to the destination.

elastic

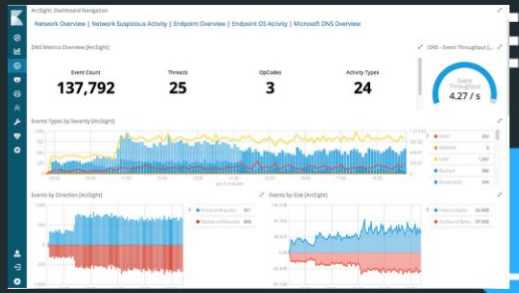
Elastic Common Schema

Without Common Schema

- New queries, dashboards, alerts, ML jobs required for every unique data source

Cisco ASA Firewall Dashboard

Palo Alto Firewall Dashboard



ArcSight SIEM Firewall Dashboard

elastic

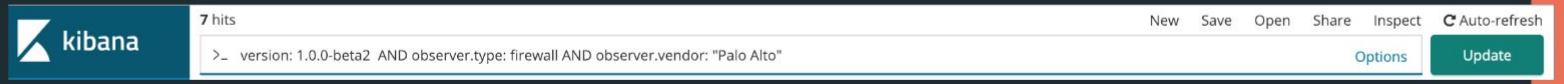
Elastic Common Schema

With Elastic Common Schema

- Queries, dashboards, alerts, ML jobs can be used across many data sources
- Correlation becomes implicit with every search!

Unified Firewall Dashboard

- But you can still filter down to specific device types



Elastic Common Schema (ECS)

Three Levels of Fields

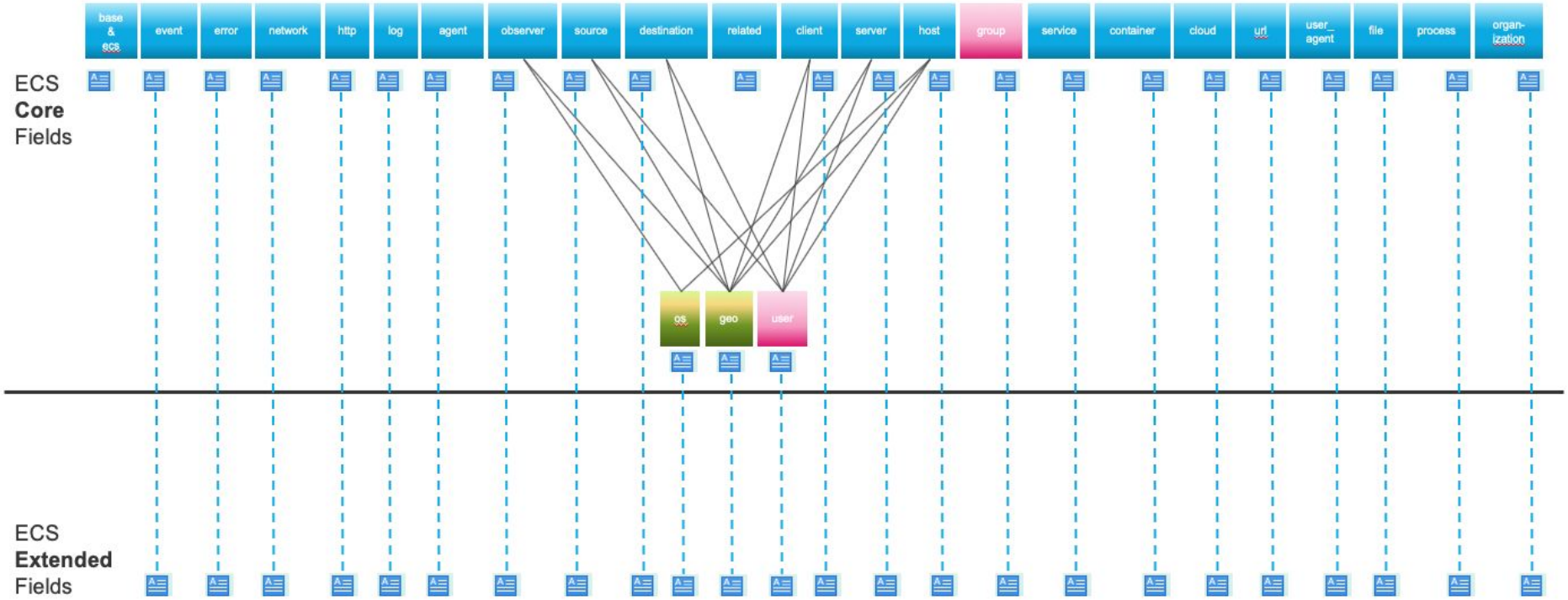
- **ECS-Core:** A fully defined set of field names that exist under a defined set of ECS top-level objects
- **ECS-Extended:** A partially defined set of field names that exist under the same set of ECS top-level objects
- **Custom:** An undefined set of fields that exists under a user-supplied set of Non-ECS top-level objects

Elastic Common Schema

Benefits

Benefits to a user adopting these fields and names in their clusters:

- Ability to **simply correlate** data from different data sources
- Improved ability to **remember** commonly used field names (since there is only a single set, not a set per data source)
- Improved ability to **deduce** unremembered field names (since the field naming follows a small number of rules with few exceptions)
- Ability to **re-use** analysis content (searches, visualizations, dashboards, alerts, reports, and ML jobs) across multiple data sources
- Ability to **use future** Elastic-provided or partner-provided analysis content in their environment without modifications



Current as of ECS v1.0.0

Used only at the top-level of naming hierarchy

Not used at the top-level of naming hierarchy

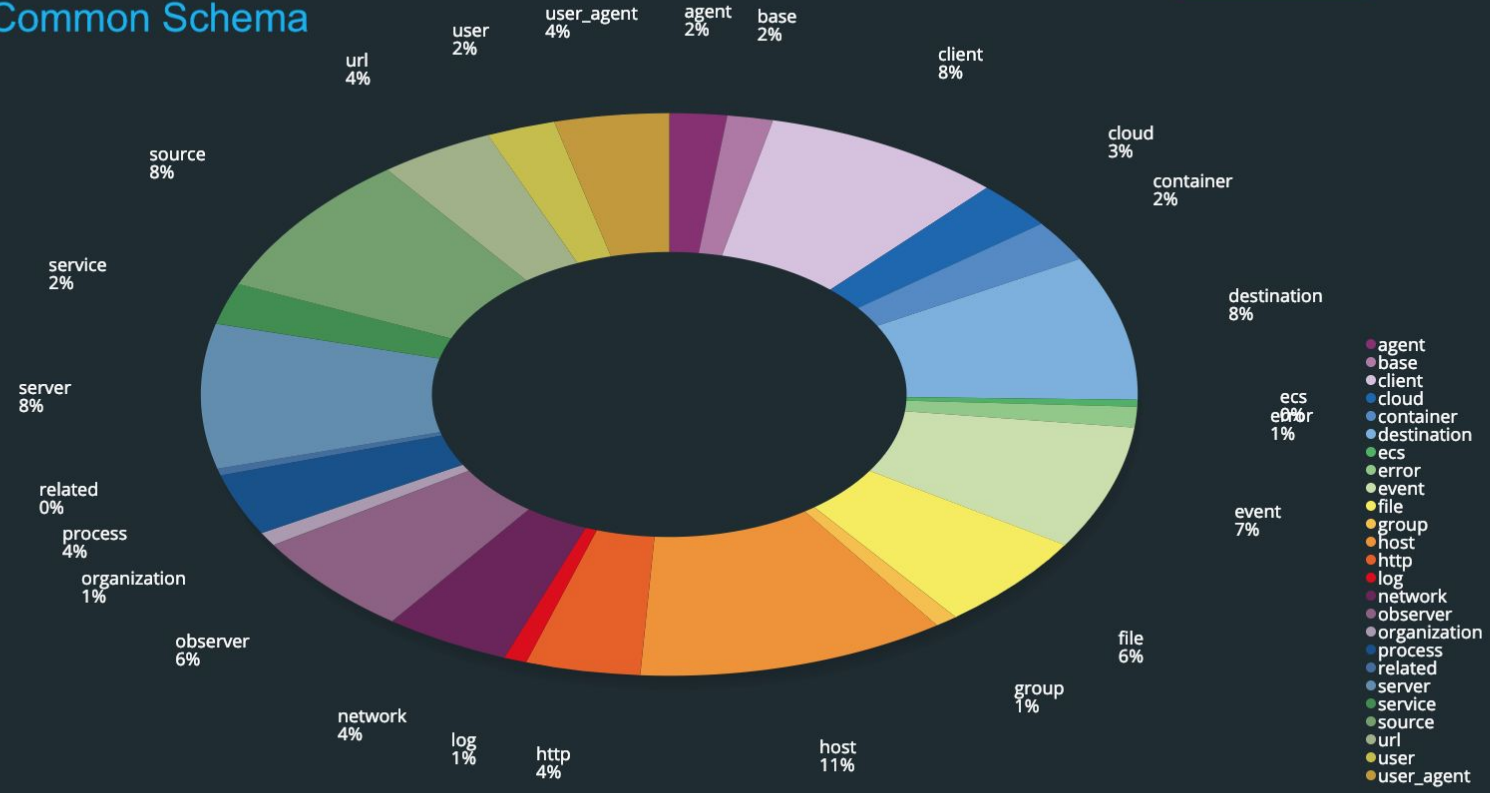
May be used at the top-level of naming hierarchy or nested at lower levels

elastic

elastic

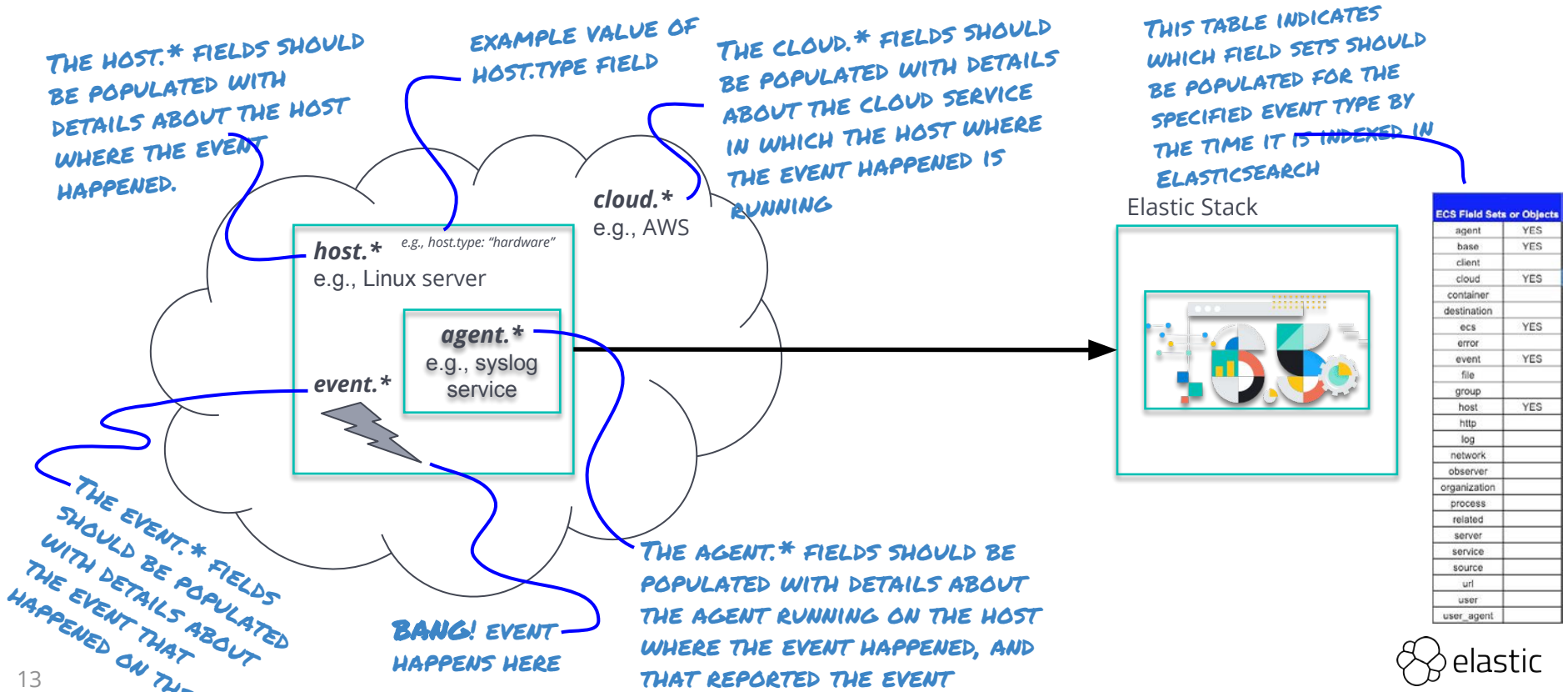
Elastic Common Schema

25
ECS Field Sets



Field Sets, Fields, Names, Datatypes, Oh My!

The ***field.**** labels indicate the entity information, if available, that should be used to populate the fields.

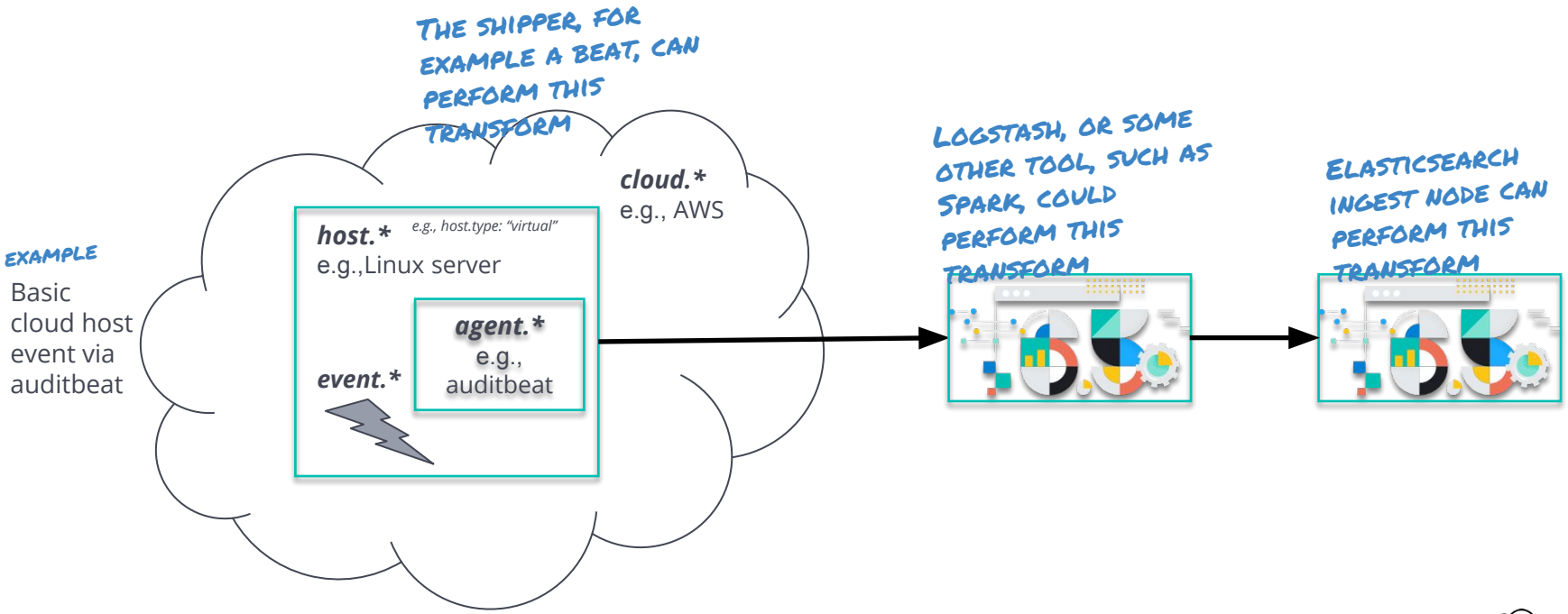


ECS Field Sets or Objects	
agent	YES
base	YES
client	
cloud	YES
container	
destination	
ecs	YES
error	
event	YES
file	
group	
host	YES
http	
log	
network	
observer	
organization	
process	
related	
server	
service	
source	
url	
user	
user_agent	

elastic

Transforming Events into ECS Format

Producing an ECS-compliant event requires one or more transformations which can be performed at the shipper, ETL, or Ingest node.



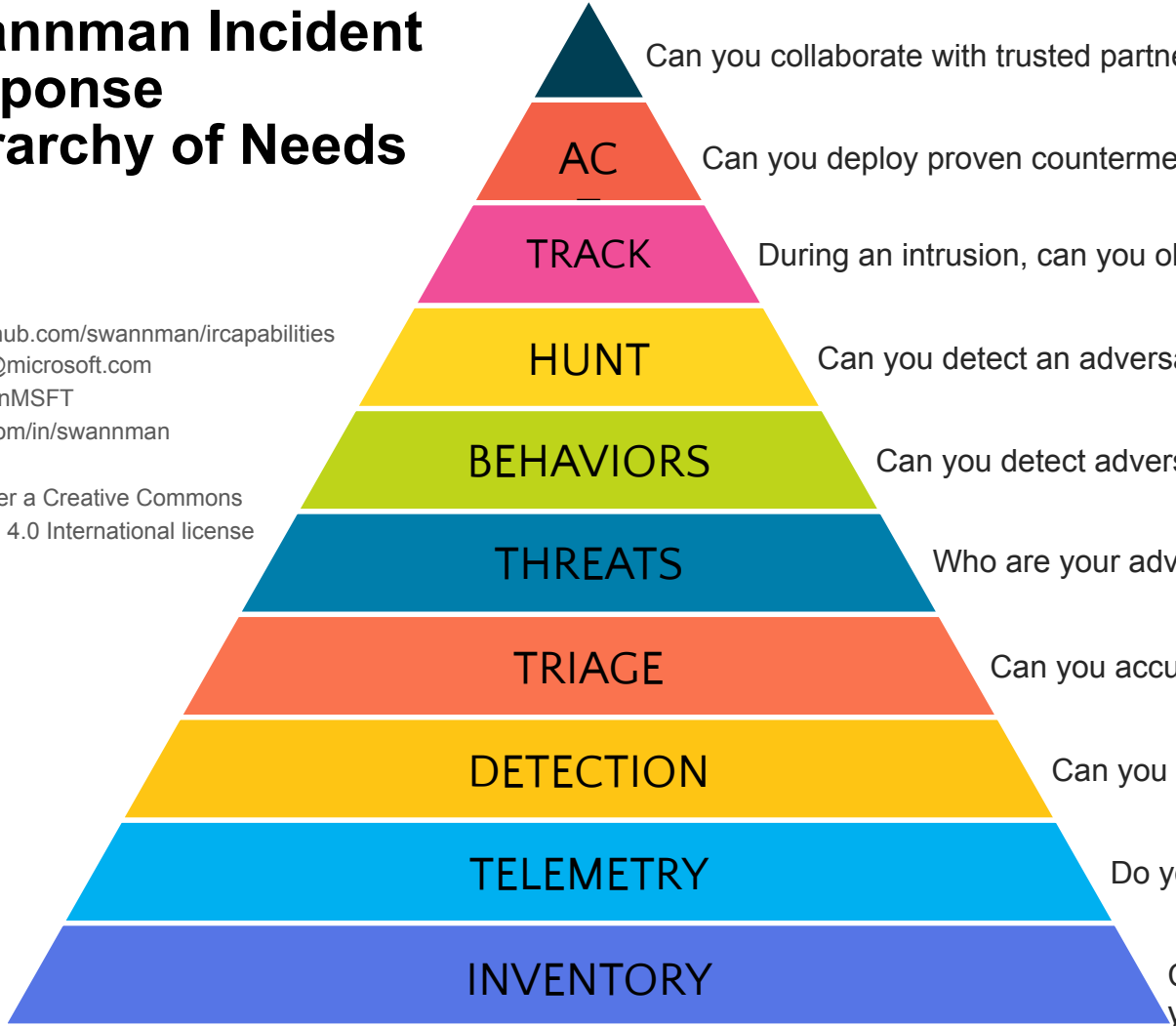
elastic

Threat Hunting with Elastic

elastic

PROACTIVE, RETROSPECTIVE, SEARCHING THROUGH COLLECTED DATA, LOOKING FOR EVIDENCE THAT ADVERSARIES ARE OPERATING IN YOUR ENVIRONMENT

Swannman Incident Response Hierarchy of Needs



Can you collaborate with trusted partners to disrupt adversary campaigns?

Can you deploy proven countermeasures to evict and recover?

During an intrusion, can you observe adversary activity in real time?

Can you detect an adversary that is already embedded?

Can you detect adversary activity within your environment?

Who are your adversaries? What are their capabilities?

Can you accurately classify detection results?

Can you detect unauthorized activity?

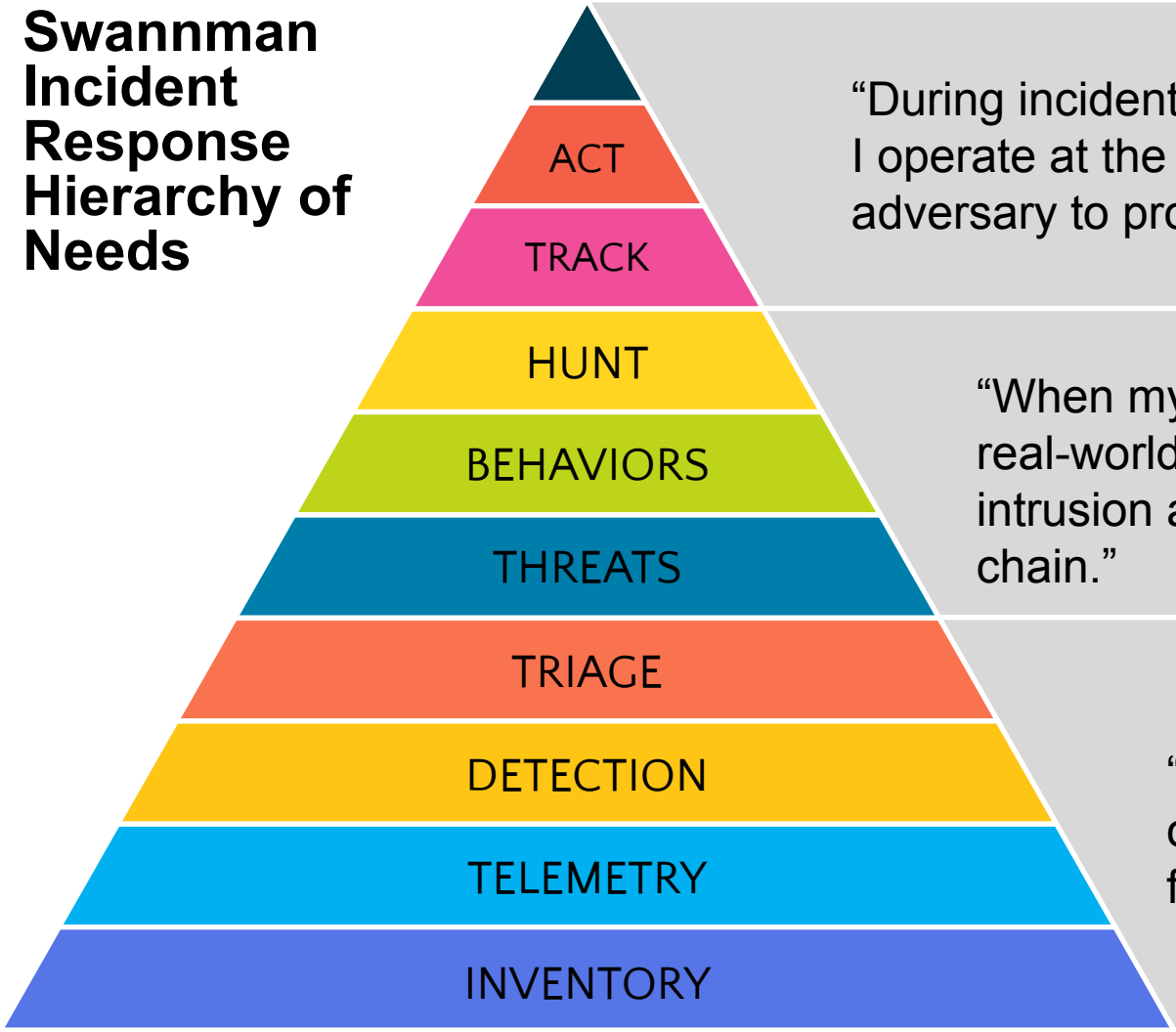
Do you have visibility across your assets?

Can you name the assets you are defending?

https://github.com/swannman/ircapabilities
mswann@microsoft.com
@MSwannMSFT
linkedin.com/in/swannman

elastic

Swannman Incident Response Hierarchy of Needs



"During incident response,
I operate at the same tempo as the adversary to protect my business assets."

"When my red team emulates a real-world adversary, I detect their intrusion at multiple points along the kill chain."

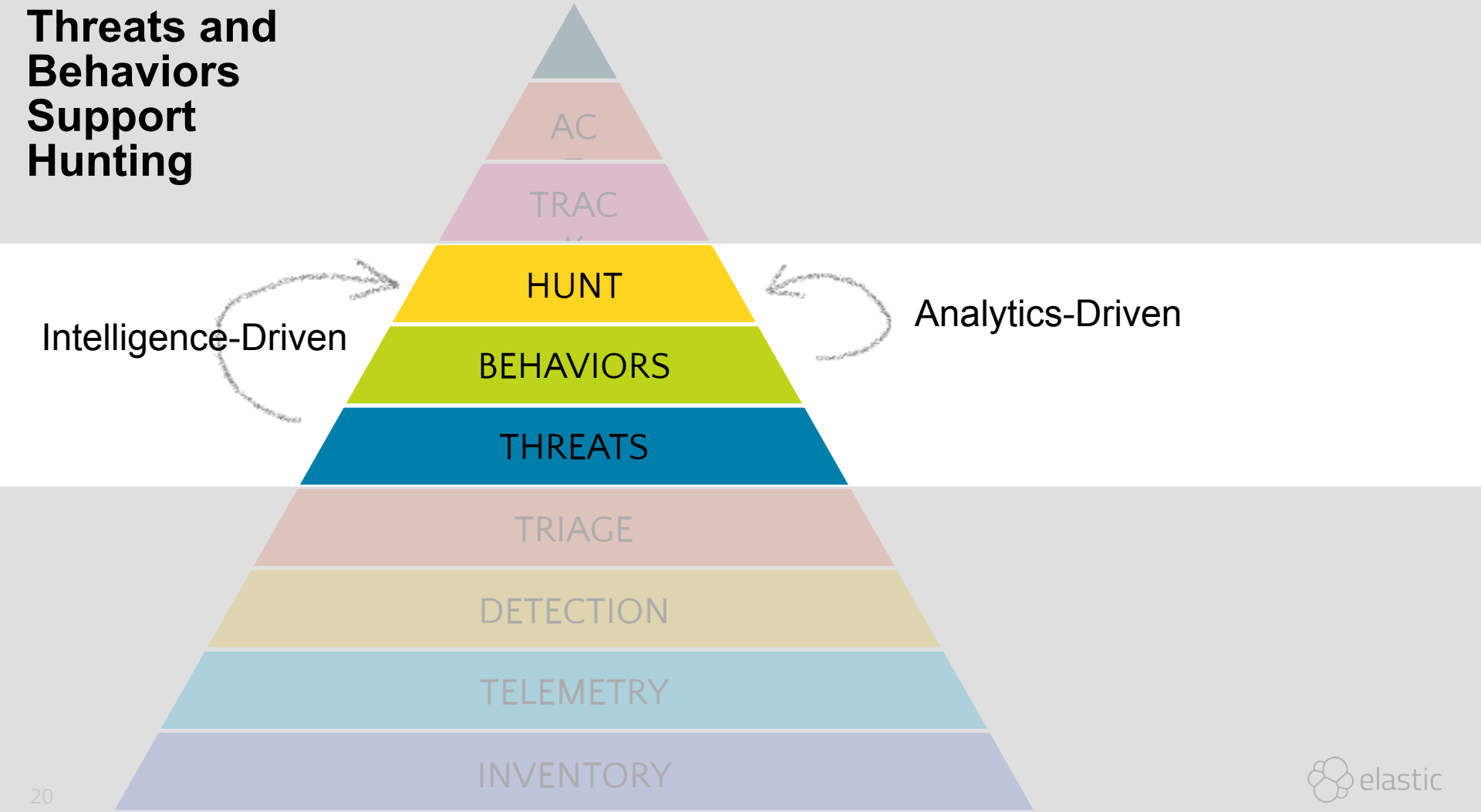
"I detect hygiene issues and operator activity that does not follow best practices."

Threat Hunting Methodologies

https://sqrrl.com/cyber-threat-hunting-1-intro/

- **Analytics-Driven:** "Machine-learning and UEBA, used to develop aggregated risk scores that can also serve as hunting hypotheses"
- **Situational-Awareness Driven:** "Crown Jewel analysis, enterprise risk assessments, company- or employee-level trends"
- **Intelligence-Driven:** "Threat intelligence reports, threat intelligence feeds, malware analysis, vulnerability scans"

Threats and Behaviors Support Hunting



elastic

What are Behaviors?

For Threat Hunting

elastic

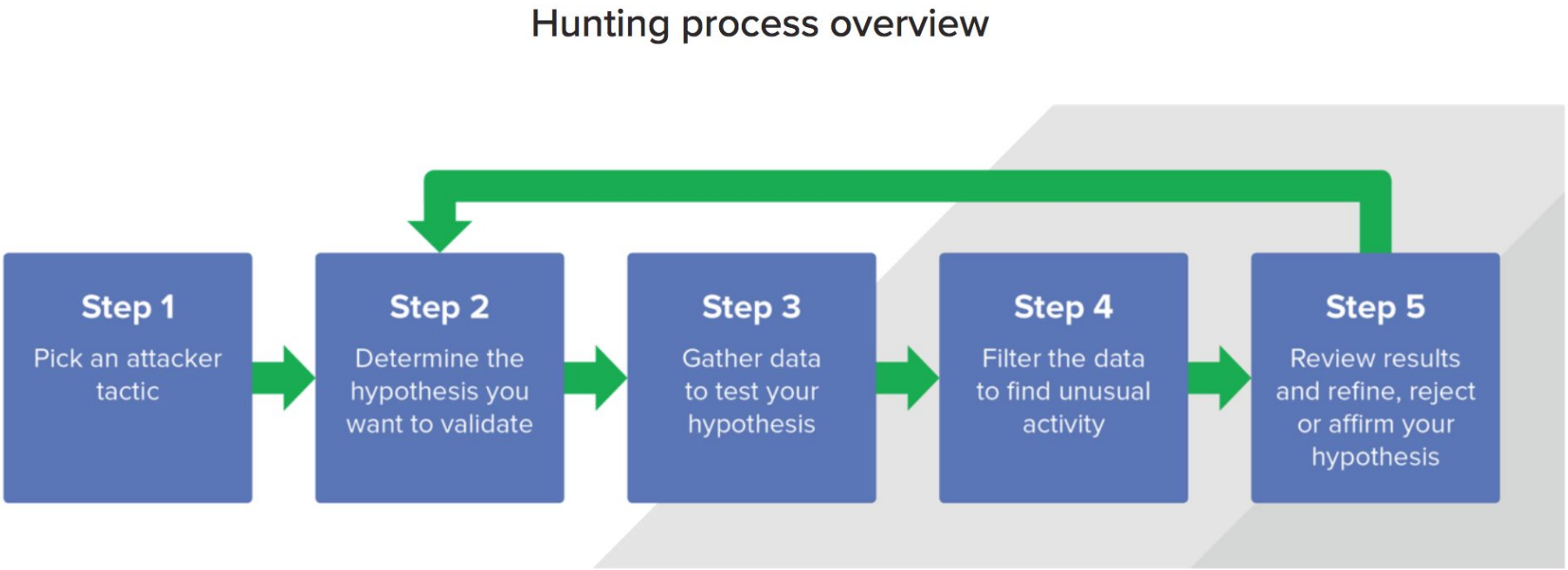
ATT&CK Matrix for Enterprise

The full ATT&CK Matrix below includes techniques spanning Windows, Mac, and Linux platforms and can be used to navigate through the knowledge base.

Initial Access	Execution	Persistence	Privilege Escalation	Defense Evasion	Credential Access	Discovery	Lateral Movement	Collection	Exfiltration	Command and Control
Drive-by Compromise	AppleScript	.bash_profile and .bashrc	Access Token Manipulation	Access Token Manipulation	Account Manipulation	Account Discovery	AppleScript	Audio Capture	Automated Exfiltration	Commonly Used Port
Exploit Public-Facing Application	CMSTP	Accessibility Features	Accessibility Features	BITS Jobs	Bash History	Application Window Discovery	Application Deployment Software	Automated Collection	Data Compressed	Communication Through Removable Media
Hardware Additions	Command-Line Interface	AppCert DLLs	AppCert DLLs	Binary Padding	Brute Force	Browser Bookmark Discovery	Distributed Component Object Model	Clipboard Data	Data Encrypted	Connection Proxy
Replication Through Removable Media	Control Panel Items	AppInit DLLs	AppInit DLLs	Bypass User Account Control	Credential Dumping	File and Directory Discovery	Exploitation of Remote Services	Data Staged	Data Transfer Size Limits	Custom Command and Control Protocol
Spearphishing Attachment	Dynamic Data Exchange	Application Shimming					Logon Scripts	Data from Information Repositories	Exfiltration Over Alternative Protocol	Custom Cryptographic Protocol
Spearphishing Link	Execution through API	Authentication Package					Pass the Hash	Data from Local System	Exfiltration Over Command and Control Channel	Data Encoding
Spearphishing via Service	Execution through Module Load	BITS Jobs					Pass the Ticket	Data from Network Shared Drive	Exfiltration Over Other Network Medium	Data Obfuscation
Supply Chain Compromise	Exploitation for Client Execution	Bootkit					Remote Desktop Protocol	Data from Removable Media	Exfiltration Over Physical Medium	Domain Fronting
Trusted Relationship	Graphical User Interface	Browser Extensions					Remote File Copy	Email Collection	Scheduled Transfer	Fallback Channels
Valid Accounts	InstallUtil	Change Default File Association					Remote Services	Input Capture		Multi-Stage Channels
	LSASS Driver	Component Firmware					Replication Through Removable Media	Man in the Browser		Multi-hop Proxy
	Launchctl	Component Object Model Hijacking					SSH Hijacking	Screen Capture		Multiband Communication
	Local Job Scheduling	Create Account	Image File Execution Options Injection	DLL Side-Loading	Keychain	Security Software Discovery	Shared Webroot	Video Capture		Multilayer Encryption
	Mshta	DLL Search Order Hijacking	Launch Daemon	Deobfuscate/Decode Files or Information	LLMNR/NBT-NS Poisoning	System Information Discovery	Taint Shared Content			Port Knocking
	PowerShell	Dylib Hijacking	New Service	Disabling Security Tools	Network Sniffing	System Network Configuration Discovery	Third-party Software			Remote Access Tools
	Regsvcs/Regasm	External Remote Services	Path Interception	Exploitation for Defense Evasion	Password Filter DLL	System Network Connections Discovery	Windows Admin Shares			Remote File Copy
	Regsvr32	File System Permissions Weakness	Plist Modification	Extra Window Memory Injection	Private Keys	System Owner/User Discovery	Windows Remote Management			Standard Application Layer Protocol
	Rundll32	Hidden Files and Directories	Port Monitors	File Deletion	Securityd Memory	System Service Discovery				Standard Cryptographic Protocol

What is Cyber Threat Hunting?

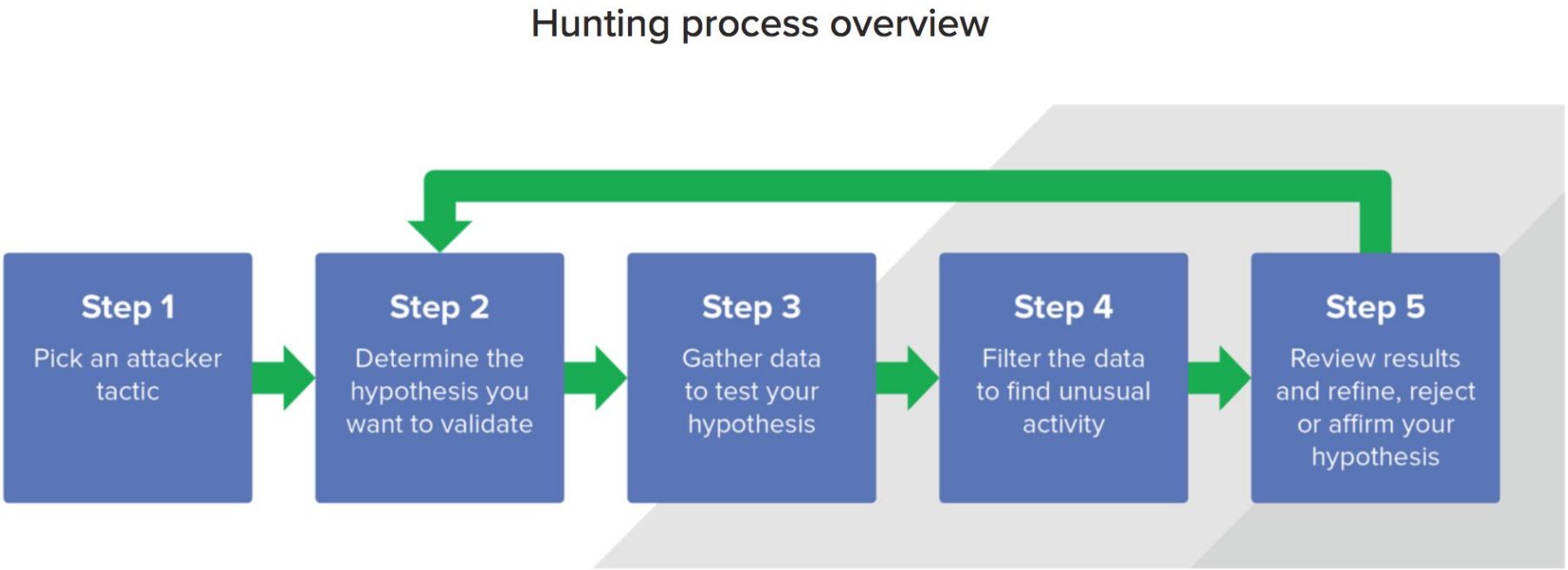
https://expel.io/blog/what-is-cyber-threat-hunting-and-where-do-you-start/



elastic

TIP: Look at Step 3 First – Do We Have the Data?

https://expel.io/blog/what-is-cyber-threat-hunting-and-where-do-you-start/



https://expel.io/blog/what-is-cyber-threat-hunting-and-where-do-you-start/

elastic

Examples

elastic

One Set of Field Names

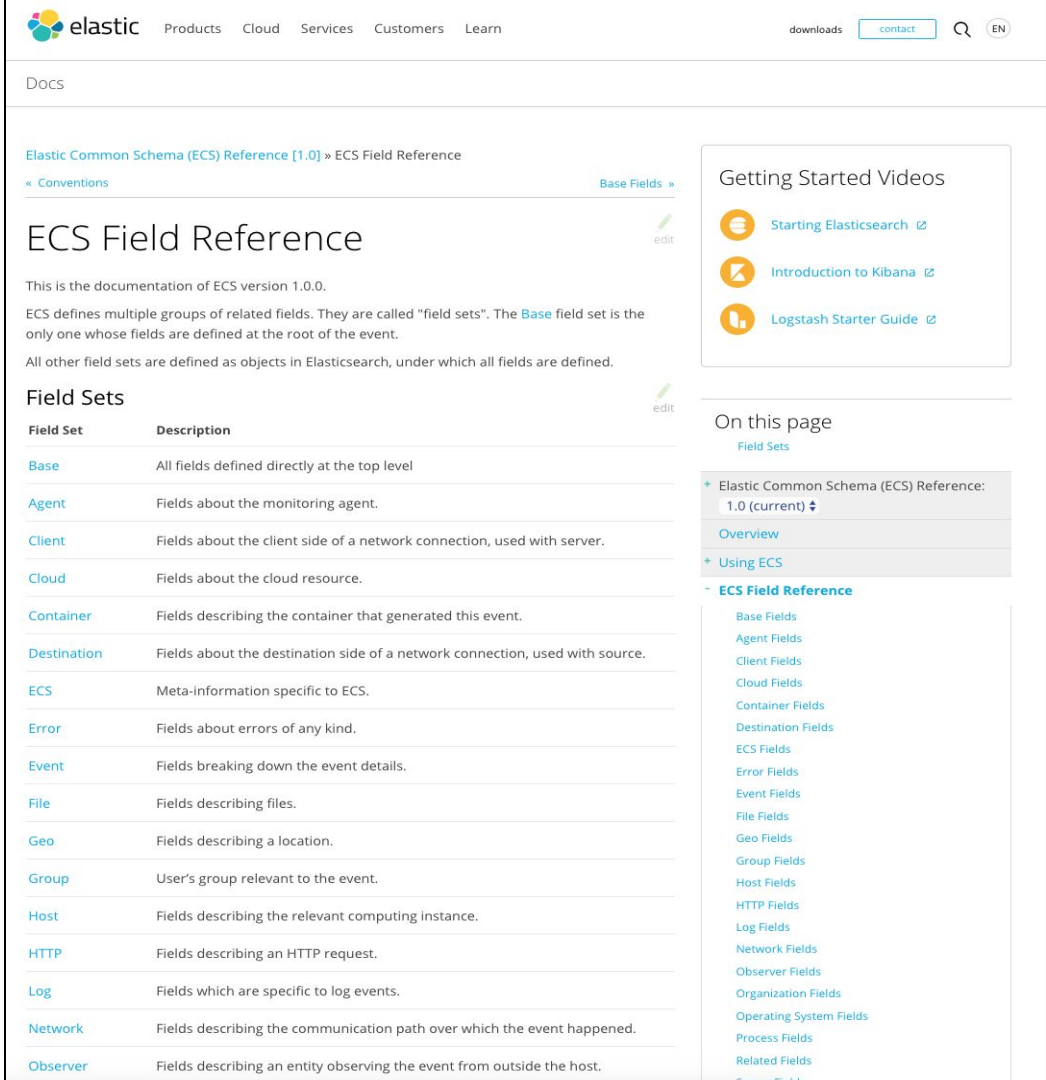
During a hunt, searching for multiple sets of field names can slow you down.

Web Site

Fully documented field reference available online at elastic.co

Offline

Grab generated file from GitHub Repo and use for air-gapped deployments



Leading Wildcard Searches

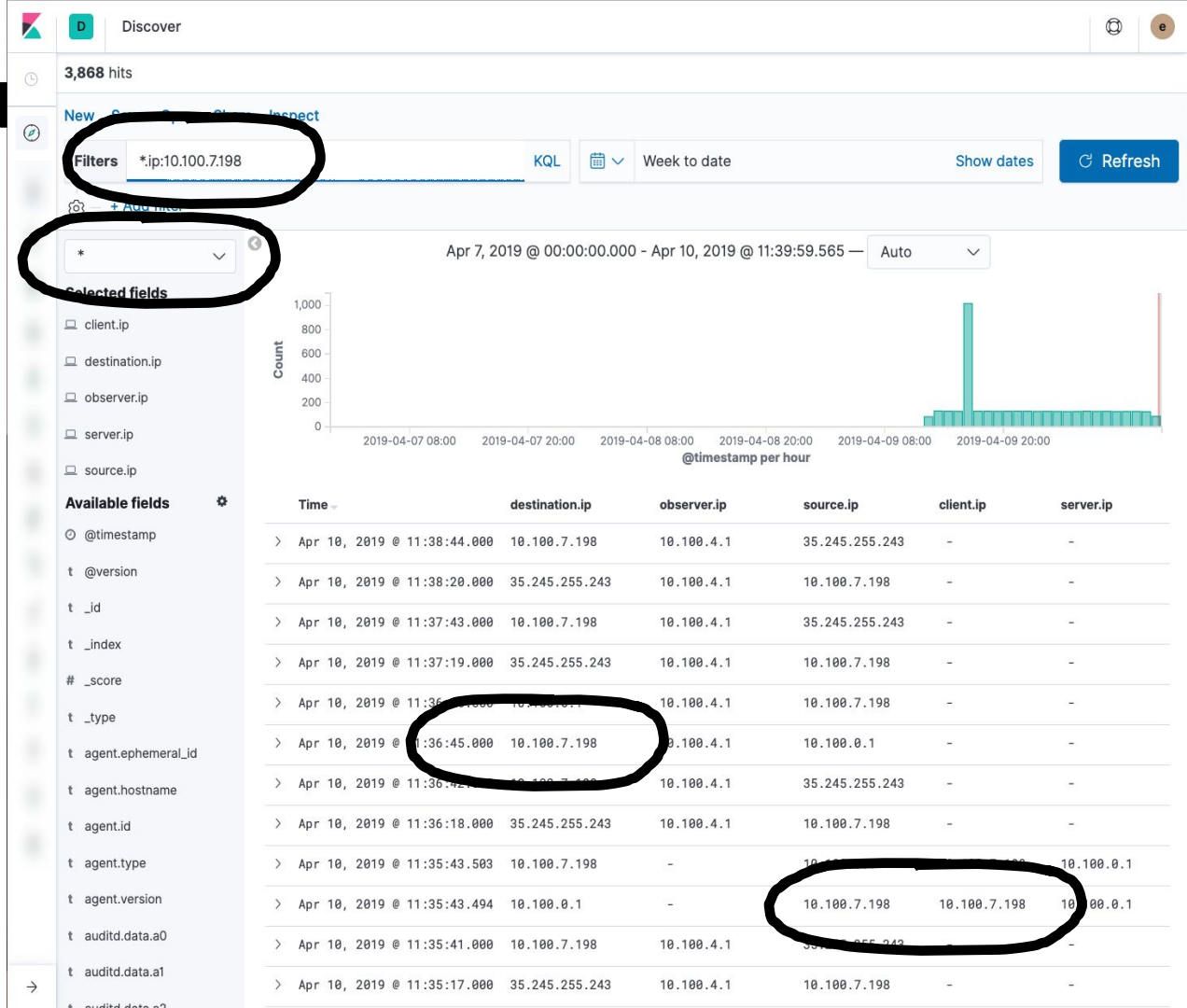
Easily find all log messages associated with a value in an ECS field across indices!

IP Addresses

Whether used as a source.ip, destination.ip, client.ip, server.ip, or observer.ip

Hostnames

Whether used as host.hostname, observer.hostname, or custom field hostname.



Leading Wildcard Searches

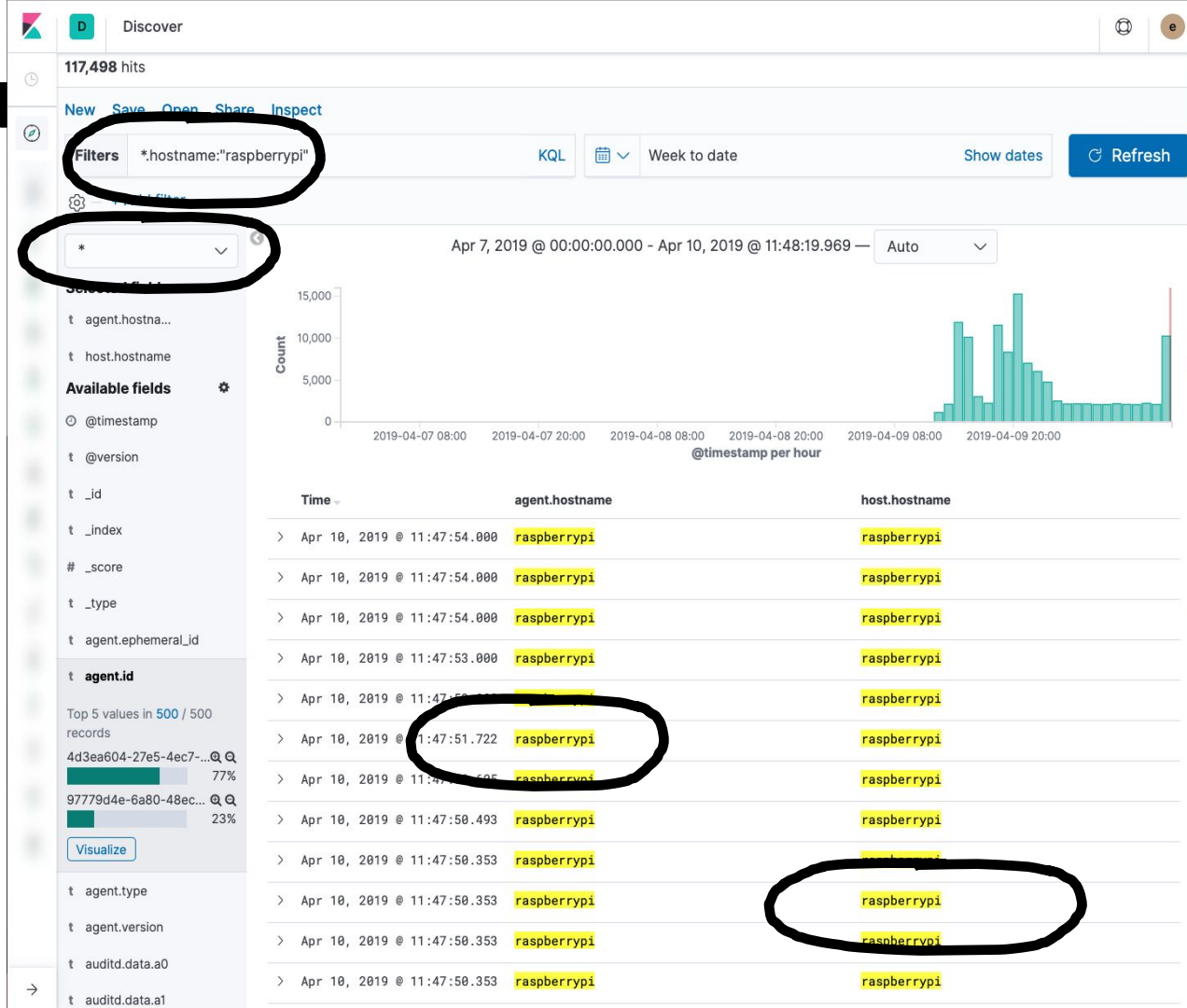
Easily find all log messages associated with a value in an ECS field across indices!

IP Addresses

Whether used as a source.ip, destination.ip, client.ip, server.ip, or observer.ip

Hostnames

Whether used as host.hostname, observer.hostname, or custom field hostname



Event Categorization

Easily find all log messages associated with a certain kind of event.

Event Kinds

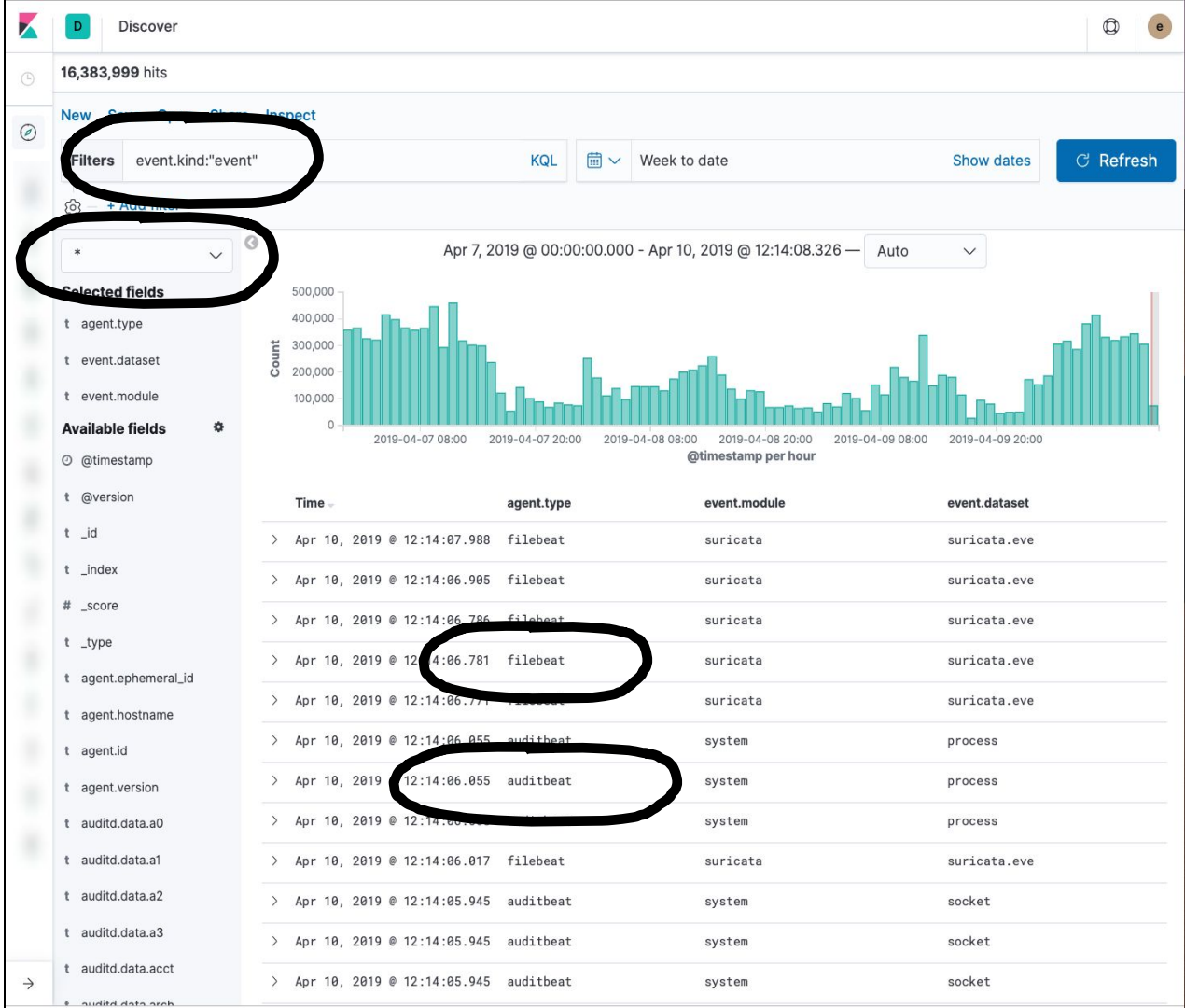
Whether events, alerts, alarms, from multiple sources

Event Actions

Whether user logins, process started, file-opened, from multiple sources

Event Outcomes

Whether user success, failure, or unknown from multiple sources



Thank You

Mike Paquette
mike.paquette@elastic.co

Normalize Data with Elastic Common Schema (ECS)

Searching *without* ECS

```
src:10.42.42.42 OR
client_ip:10.42.42.42 OR
apache2.access.remote_ip:10.
42.42.42 OR
context.user.ip:10.42.42.42
OR src_ip:10.42.42.42
```

Searching *with* ECS

```
source.ip:10.42.42.42
```