Welcome To HELK!

Elastic Tour 2018

@Cyb3rWard0g

- Adversary Detection Analyst **@SpecterOps**
- Author:
  - ThreatHunter-Playbook
  - Hunting ELK (HELK)
  - ATTACK-Python-Client
  - OSSEM (Open Source Security Event Metadata)
- Former:
  Capital One - USA, Senior Threat Hunter
  DuPont - USA, Security Specialist

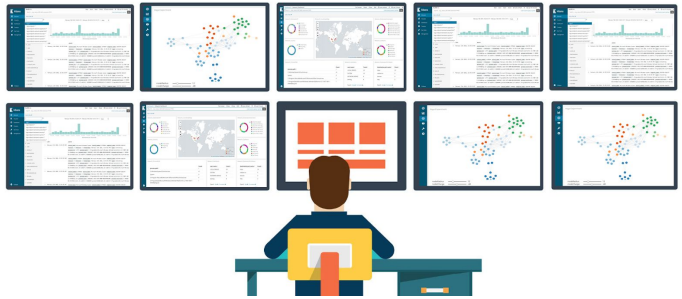
Agenda

- Threat hunting & data
- Threat hunting & hunters
- Threat hunting & SIEMs
- Pre-Hunt activities and data
- HELK
- What's next for HELK?

Effective Threat Hunting

Are you being effective?

Threat Hunting (Expectation)



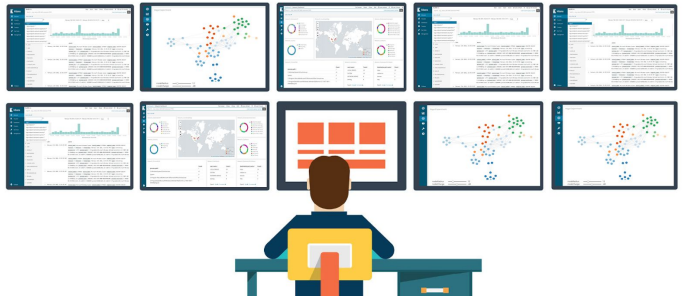
Pre-Hunt

Actual Hunting

Threat Hunting

TIME

Threat Hunting (Reality)



Pre-Hunt Activities

- Identification of adversarial techniques
- Identification of data sources required to validate the detection of adversarial techniques
- Data Documentation (Data Dictionaries)
- Data Modeling
- Data quality Assessment
  - Completeness & Standardization
- Development of data analytics
- Development of playbooks

Pre-Hunt Activities

- Identification of adversarial techniques
- **Identification of data sources** required to validate the detection of adversarial techniques
- **Data Documentation** (Data Dictionaries)
- **Data Modeling**
- **Data quality Assessment**
  - Completeness & Standardization
- **Development of data analytics**
- Development of playbooks

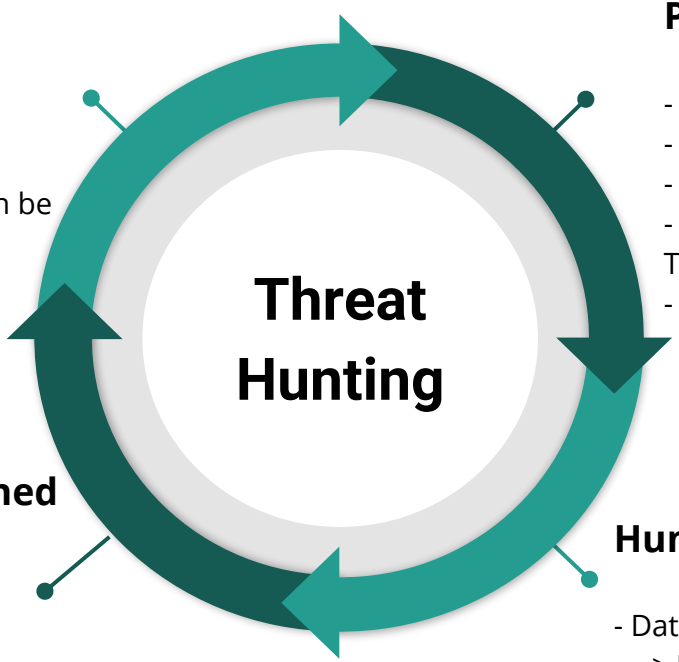
Threat Hunting & Data

LOG IT ALL -> HUNT -> FIND EVIL - REPEAT … Right?, Maybe?

Threat Hunting

What can be automated?

- Not everything can be automated
- Enhance SOC operations



Pre-Hunt

- Define Hunt Model
- Set Scope
- Define Team Roles
- Identify Adversarial Technique
- Develop Hypothesis

Hunt

- Data Analytics
  - > Behavioral
  - > Anomalies/Outliers
- Validate Detection

Lessons Learned

- Metrics
- Report Findings
- Transition to IR?
- What didn't work?

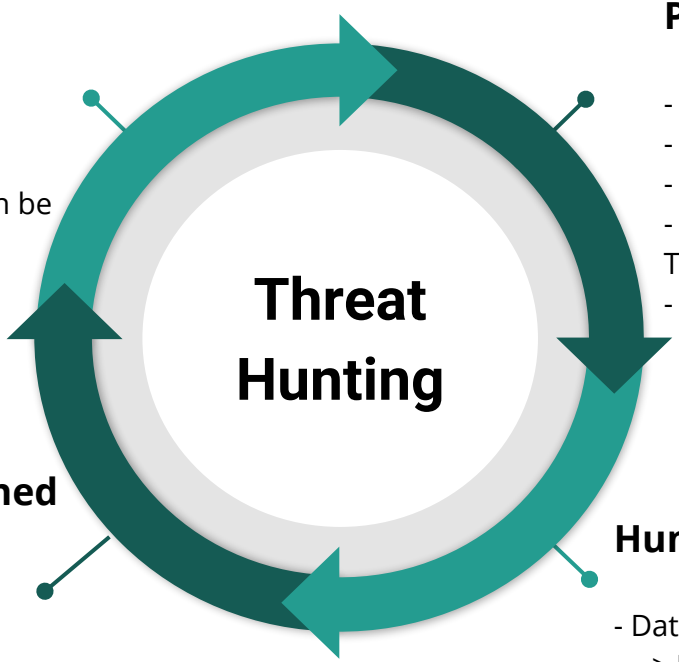
Threat Hunting

What can be automated?

- Not everything can be automated
- Enhance SOC operations

Lessons Learned

- Metrics
- Report Findings
- Transition to IR?
- What didn't work?



Pre-Hunt

- Define Hunt Model
- Set Scope
- Define Team Roles
- Identify Adversarial Technique
- Develop Hypothesis

Hunt

- Data Analytics
  - > Behavioral
  - > Anomalies/Outliers
- Validate Detection

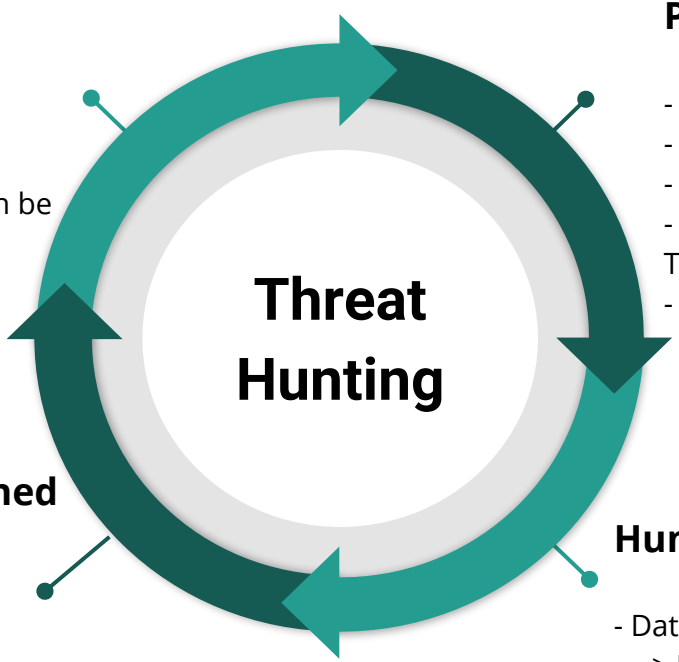
Threat Hunting

What can be automated?

- Not everything can be automated
- Enhance SOC operations

Lessons Learned

- Metrics
- Report Findings
- Transition to IR?
- What didn't work?



Pre-Hunt

- Define Hunt Model
- Set Scope
- Define Team Roles
- Identify Adversarial Technique
- Develop Hypothesis

Hunt

- Data Analytics
  - > Behavioral
  - > Anomalies/Outliers
- Validate Detection

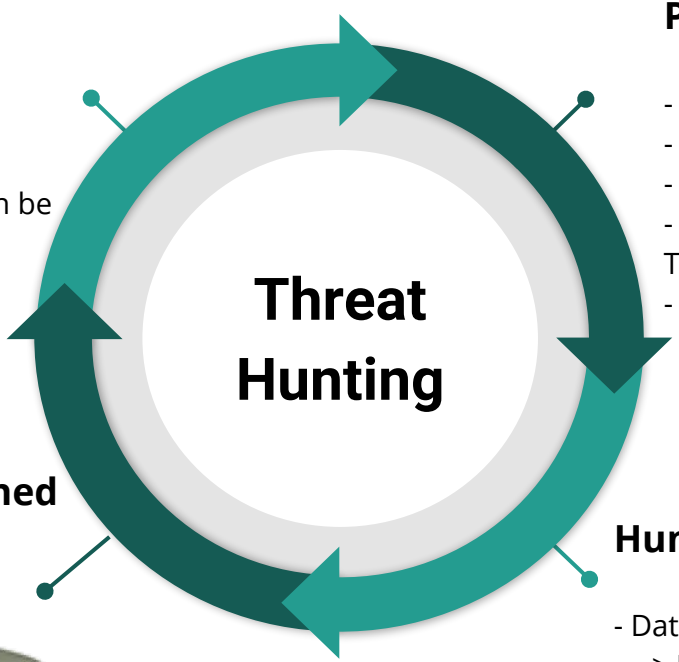
Threat Hunting

What can be automated?

- Not everything can be automated
- Enhance SOC operations

Lessons Learned

- Metrics
- Report Findings
- Transition to IR?
- What didn't work?



Pre-Hunt

- Define Hunt Model
- Set Scope
- Define Team Roles
- Identify Adversarial Technique
- Develop Hypothesis

Hunt

- Data Analytics
  - > Behavioral
  - > Anomalies/Outliers
- Validate Detection

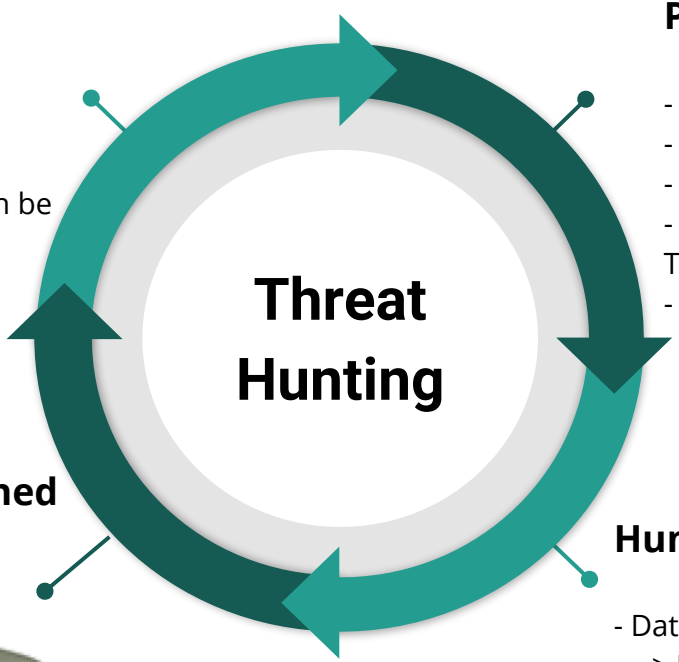
Threat Hunting

What can be automated?

- Not everything can be automated
- Enhance SOC operations

Lessons Learned

- Metrics
- Report Findings
- Transition to IR?
- What didn't work?



Pre-Hunt

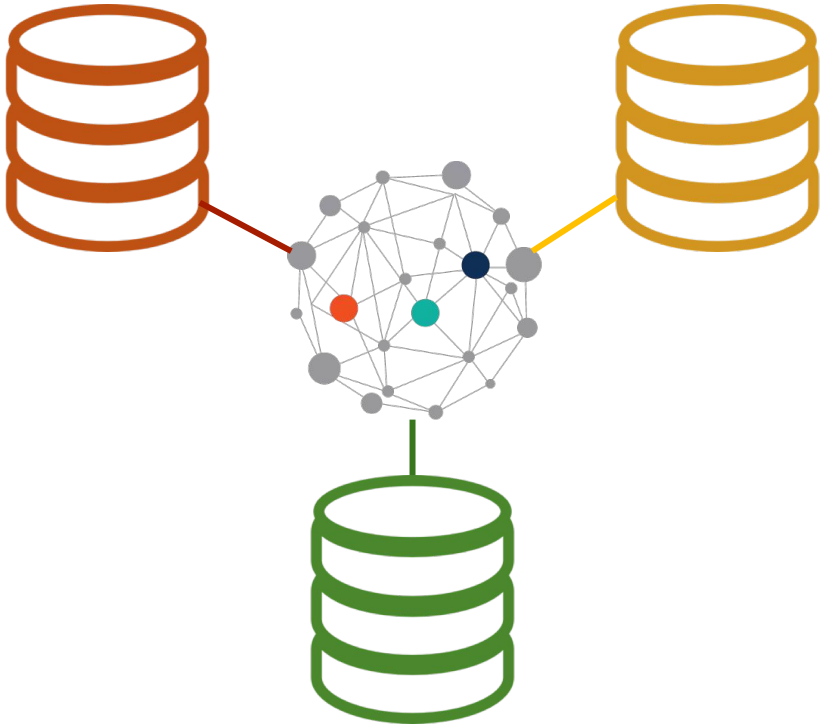
- Define Hunt Model
- Set Scope
- Define Team Roles
- Identify Adversarial Technique
- Develop Hypothesis

Hunt

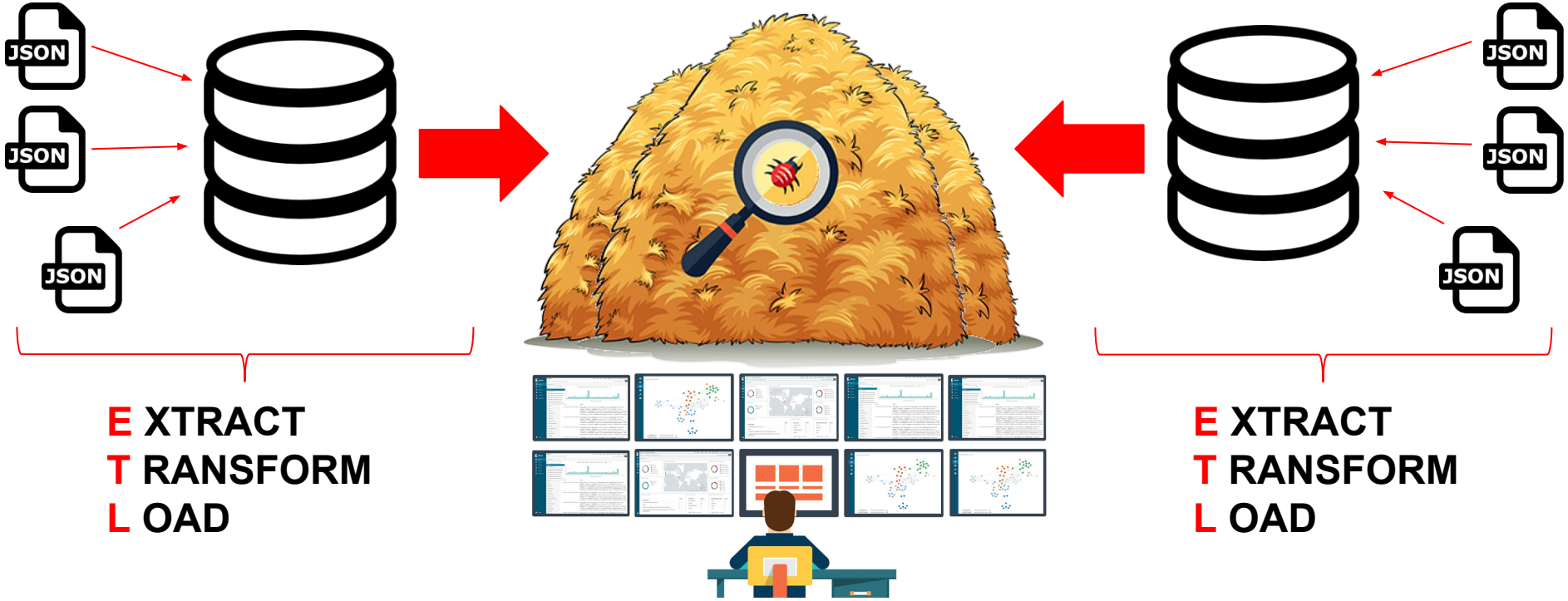
- Data Analytics
  - > Behavioral
  - > Anomalies/Outliers
- Validate Detection

Diverse Attacks Call for Diverse Data Sets

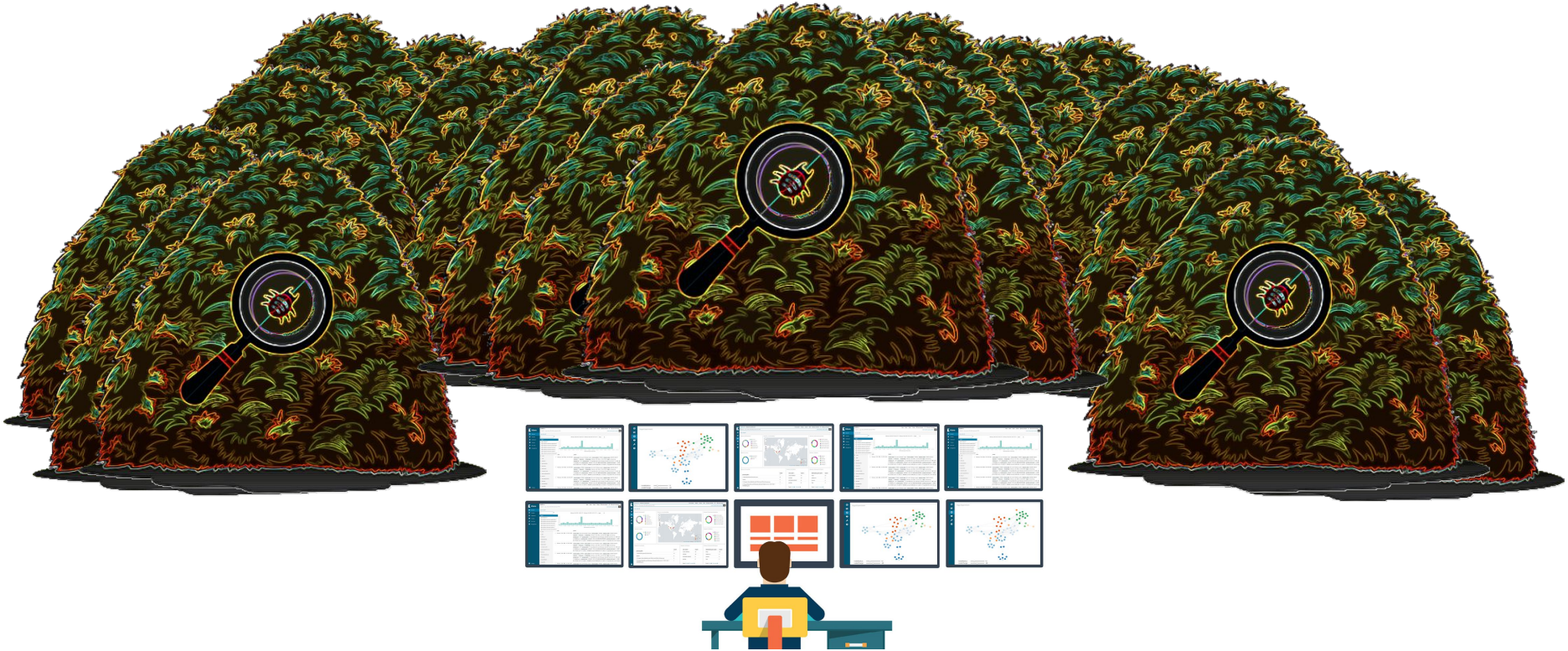
- **Endpoint Data Sets**
  - Built-In Windows Logs
  - Sysmon Logs
  - Antivirus Logs
  - Data Loss Prevention (DLP) Logs
- **Network Data Sets**
  - Proxy Logs
  - Bro **(Now Zeek)** Logs
  - Full Packet Capture



Data Lakes & Threat Hunting (In-Theory)



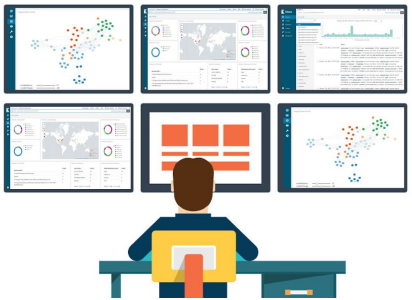
Data Swamps & Threat Hunting (Reality)



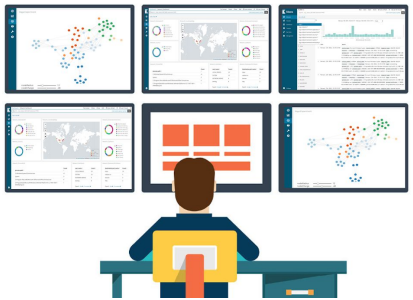
Threat Hunting & Hunters

Can we do better?

Data Engineers & Data Analysts Are Siloed!!

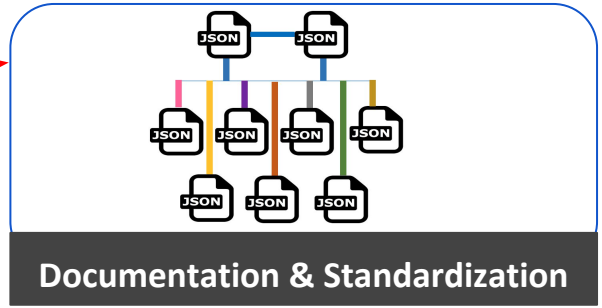


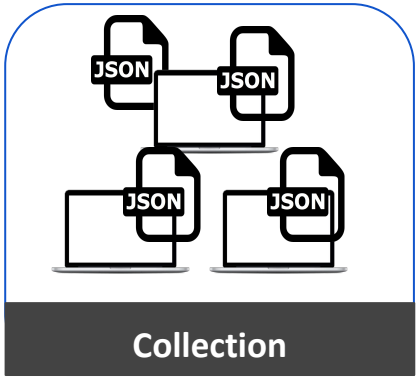
Data Engineer

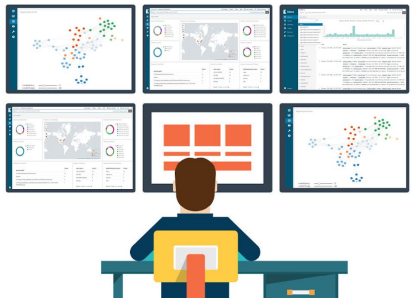


Data Analyst
(Threat Hunter)

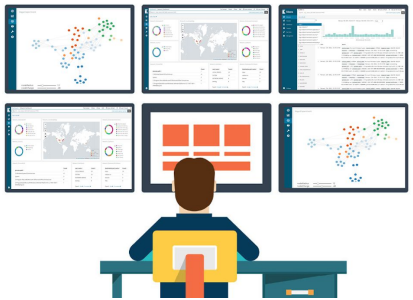
Data Engineers & Data Analysts Are Siloed!!





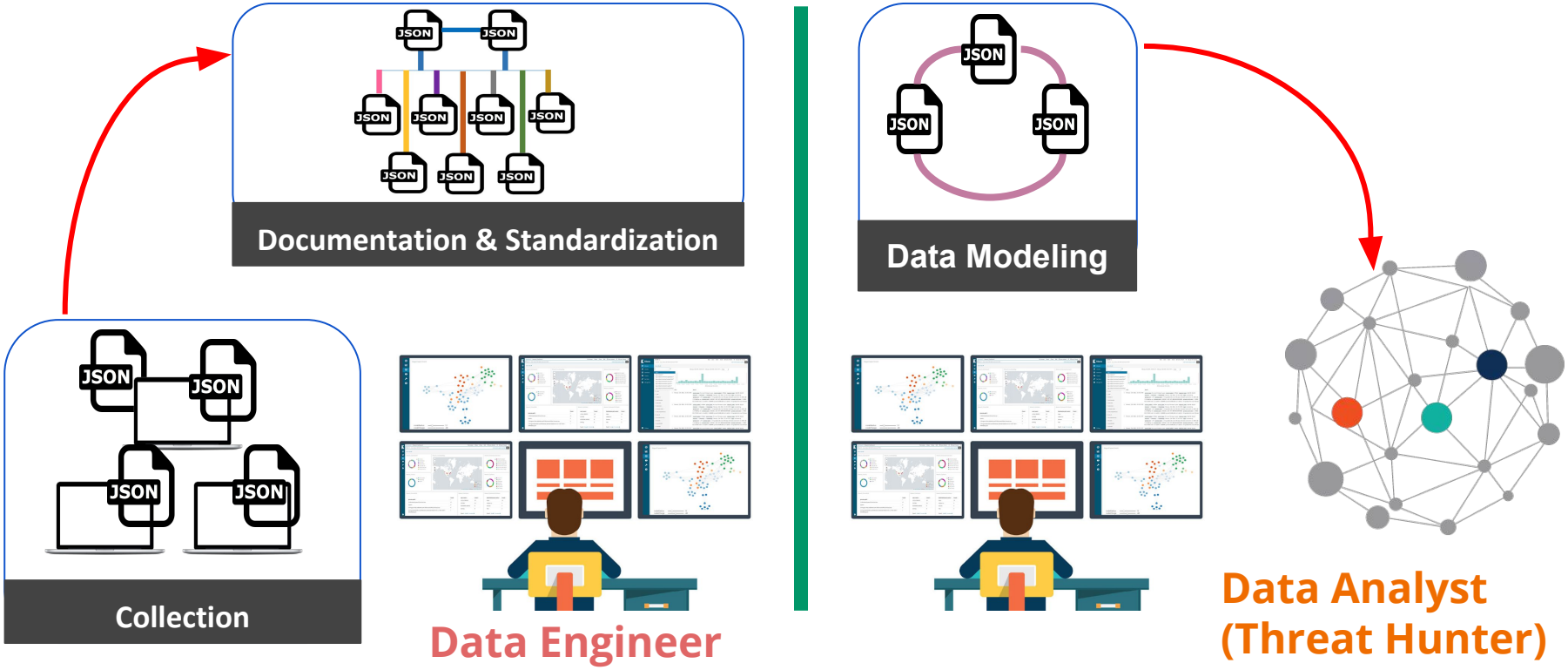


Data Engineer



Data Analyst
(Threat Hunter)

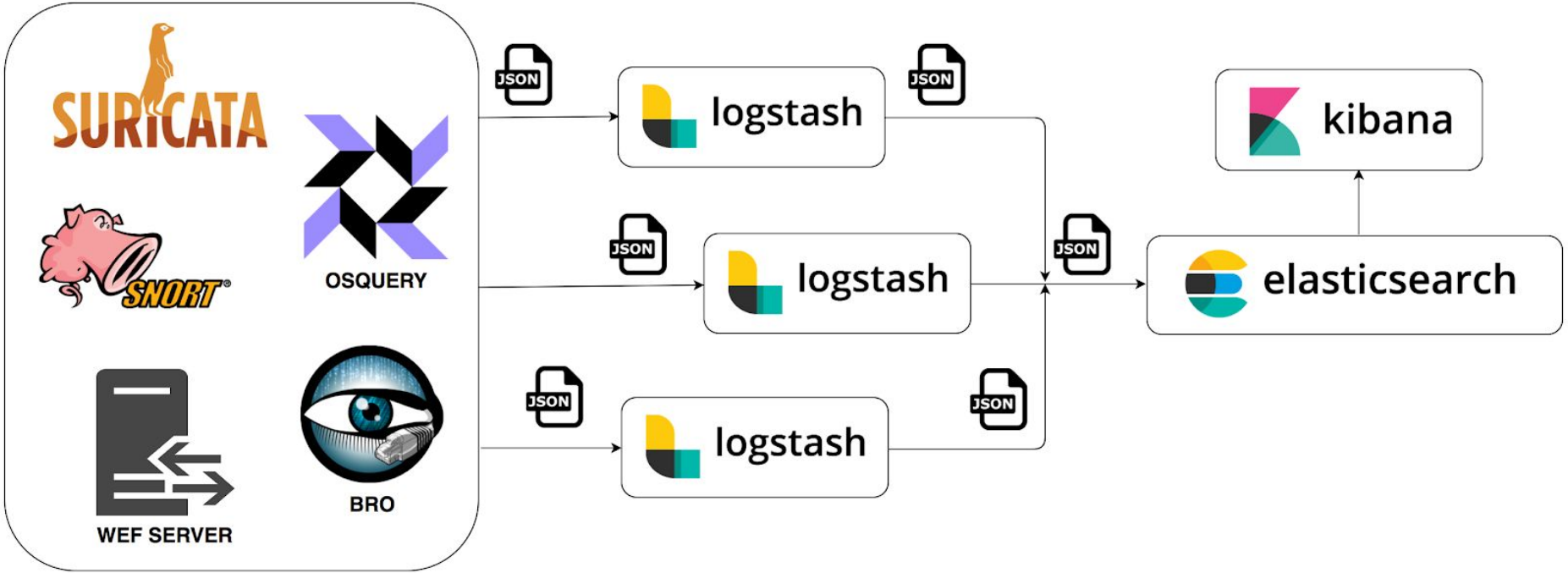
Data Engineers & Data Analysts Are Siloed!!



Threat Hunting & SIEMs

What are we trying to accomplish?

What Is The Main Goal?



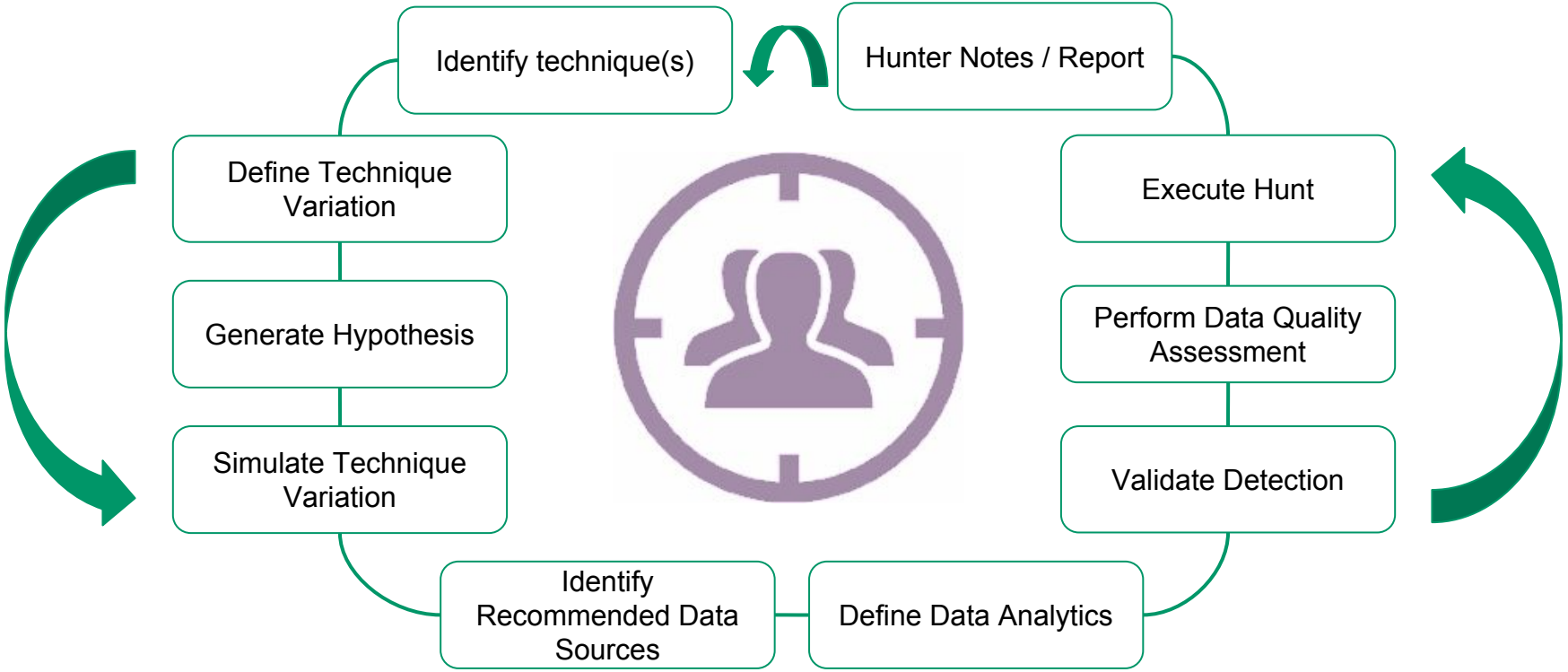
Enable Threat Hunters

- Perform flexible data analysis at scale
- Go beyond Indicators of compromise (IOC)
  - Basic queries to find an IP address from an intel report
- Identify specific structured patterns
  - Streaming, Graphing
- Describe the data in a more intuitive and efficient way
- Integrate with other threat hunting procedures
  - Playbooks
  - Training

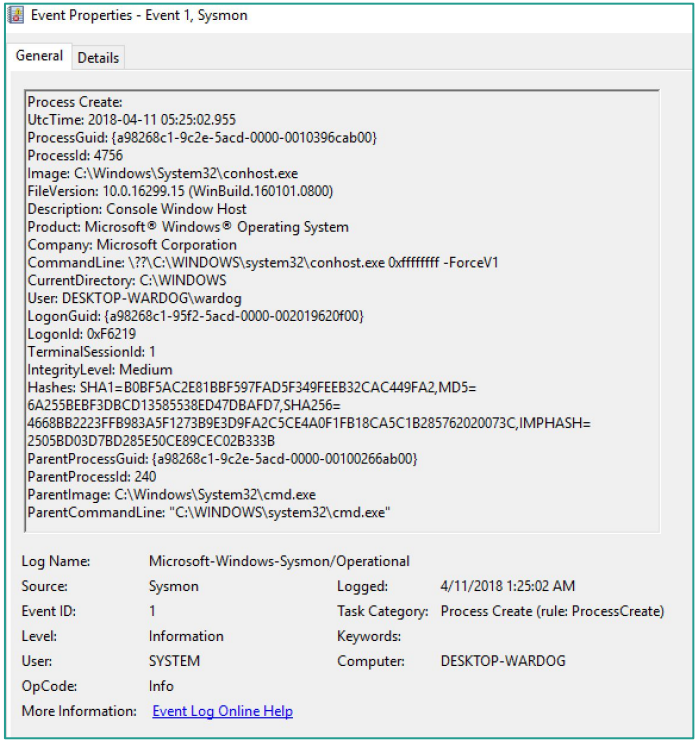
Pre-Hunt activities and data

Bringing Data Engineers and Data Analysts Together

Threat Hunting Approach



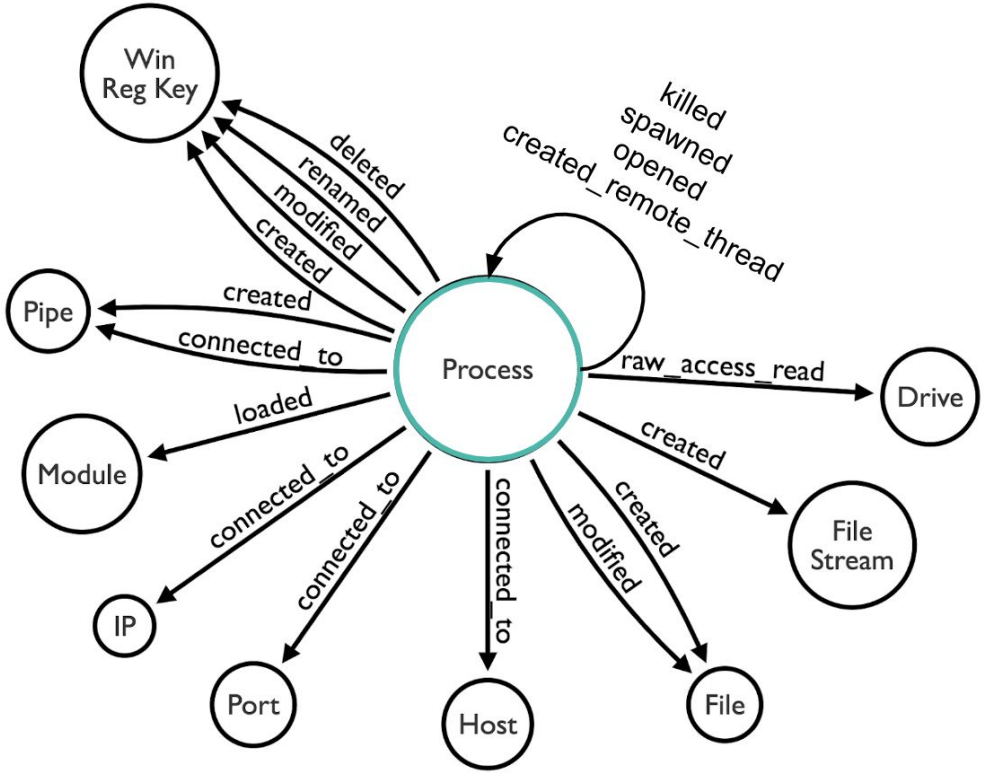
Data Dictionaries (Sysmon Event ID 1)



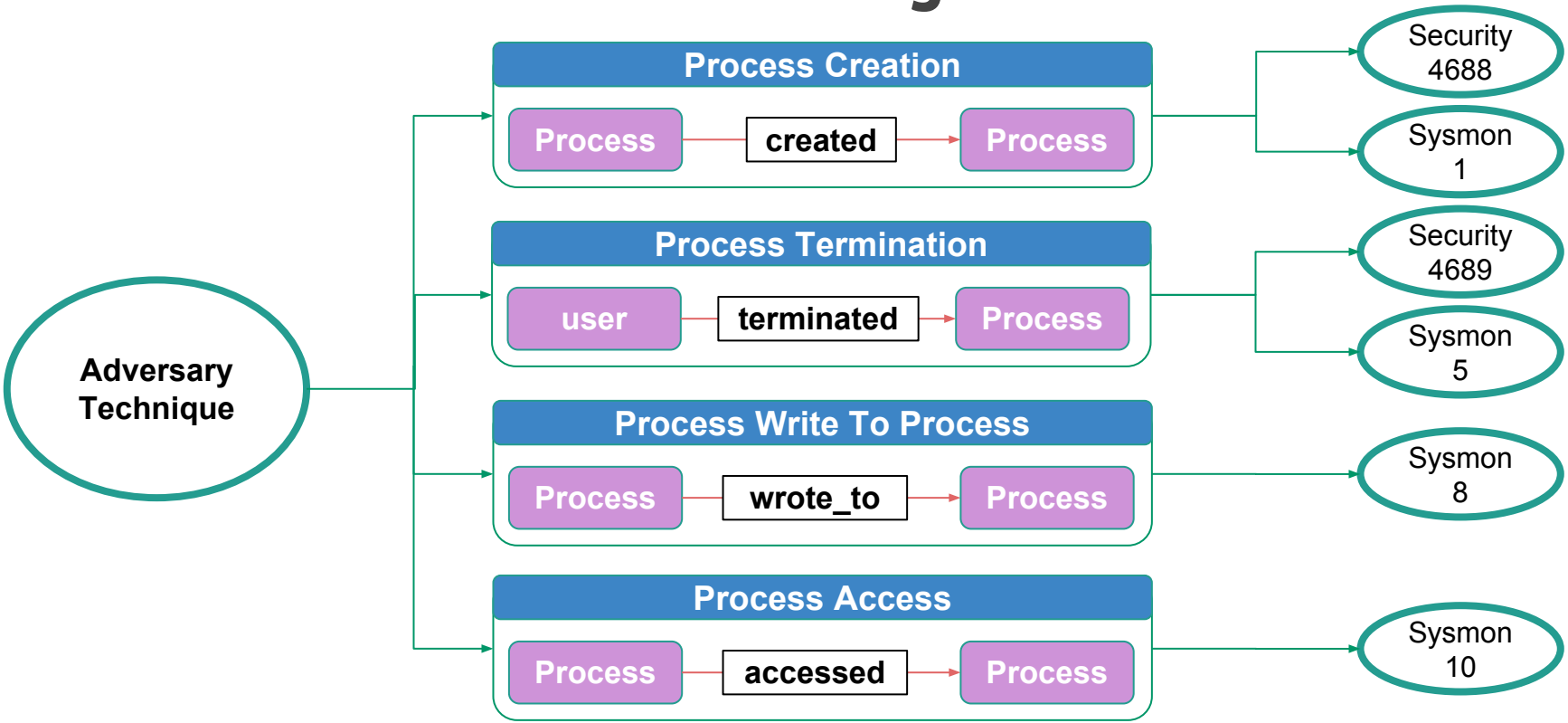
Field	Description
Image	Image path of the process executable
CommandLine	Command line of the process
CurrentDirectory	Current directory of the process
Description	PE version info resource “Description” field
FileVersion	PE version info resource “FileVersion” field
Product	PE version info resource “Product” field

https://github.com/mattifestation/BHUSA2018_Sysmon/blob/master/Whitepaper_Subverting_Sysmon.pdf

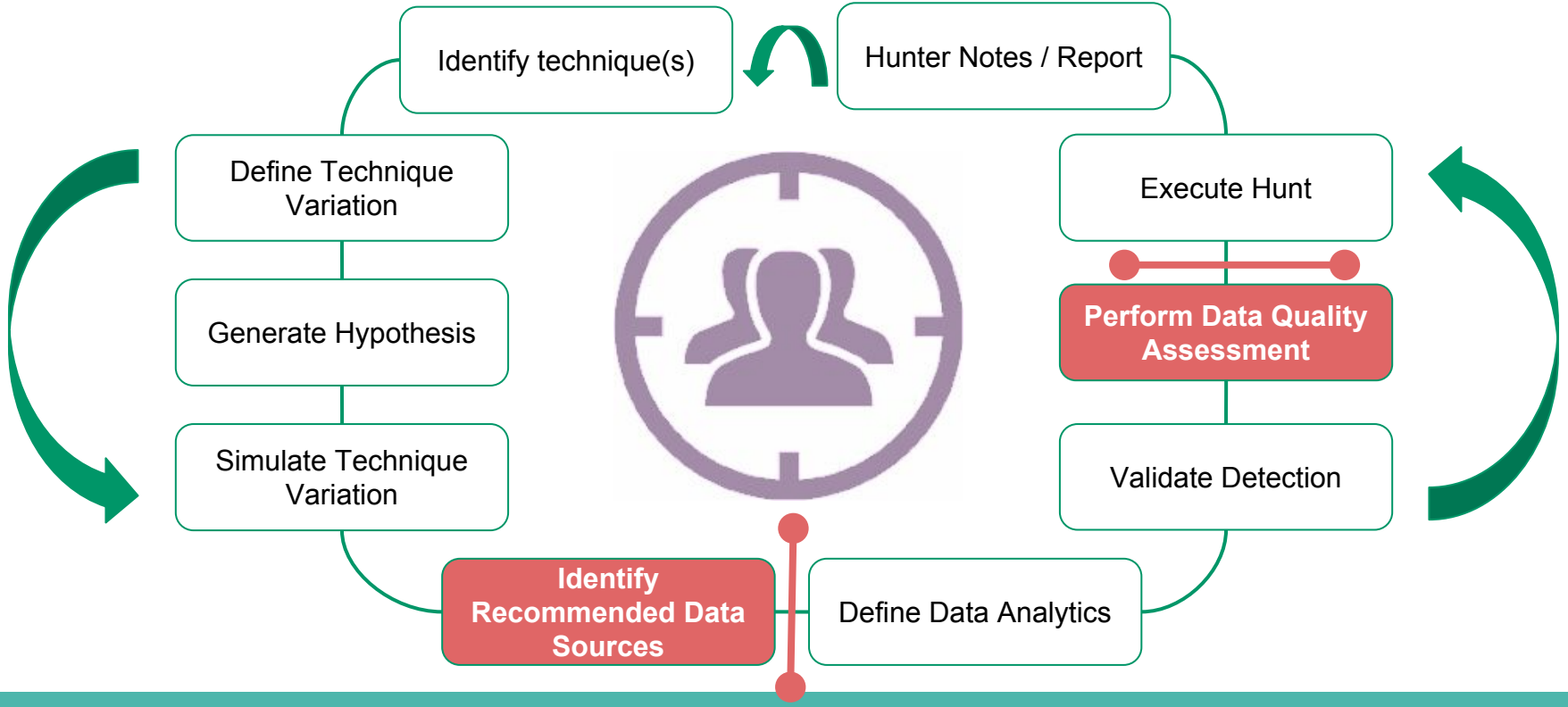
Sysmon Data Model After Documentation



Data Sources & Data Modeling



Threat Hunting & Data



HELK

[Alpha]

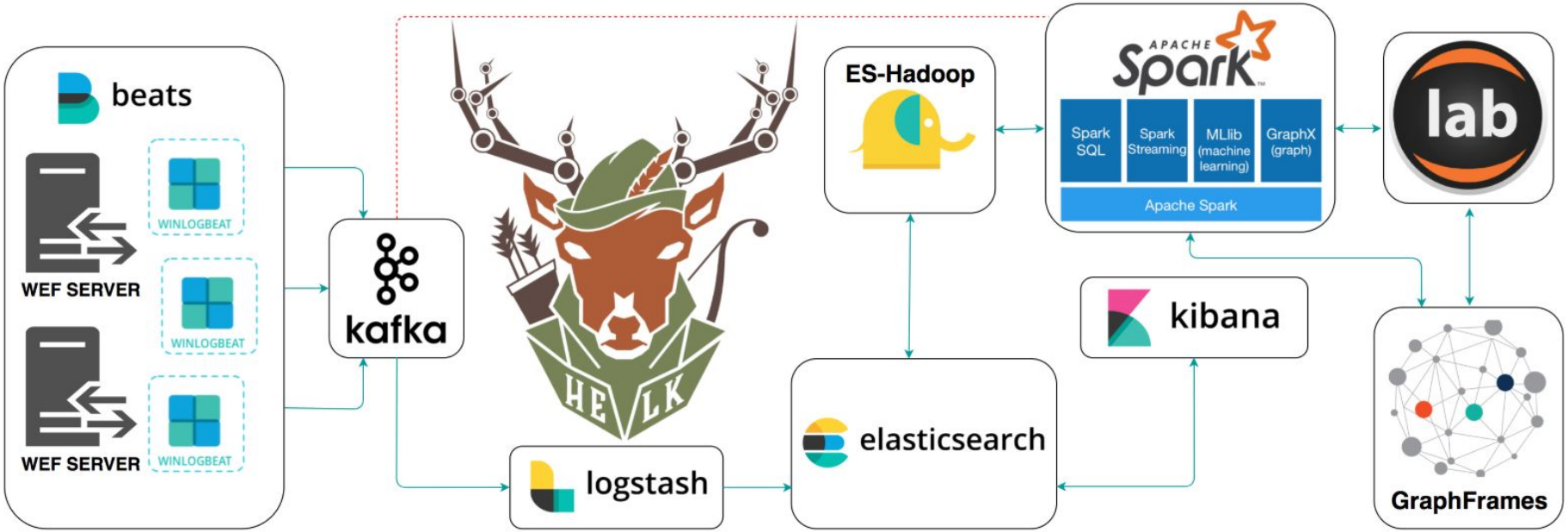
An open source ELK with Advanced Analytics Capabilities

https://github.com/Cyb3rWard0g/HELK

What the HELK?

- An ecosystem composed of several open source frameworks
- Main goal of empowering threat hunters and extending the functionalities of an Elastic ELK stack by enabling advanced analytics capabilities
- First Public Documented & Standardized Pipeline
  - OSSEM Project: https://github.com/Cyb3rWard0g/OSSEM
  - Documenting new events as they show up in new tests
- Awesome additions to the pipeline by **Nate Guagenti @neu5ron**

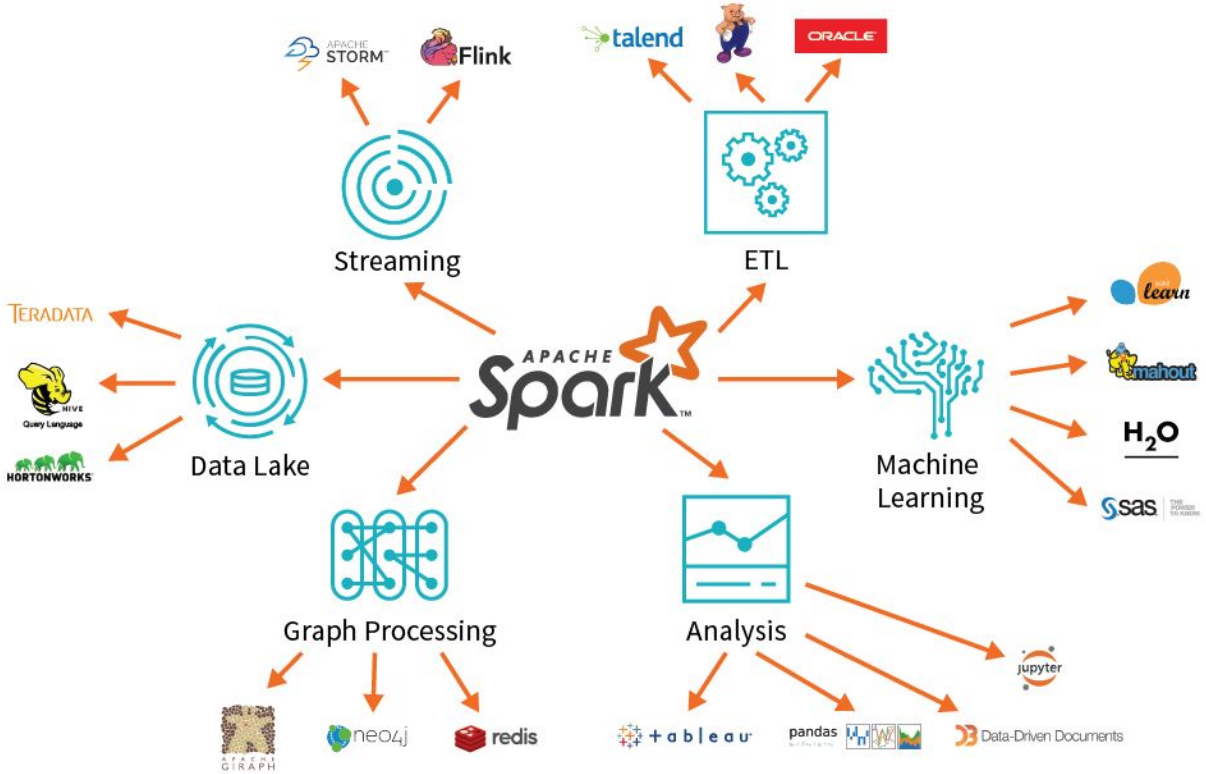
What the HELK?



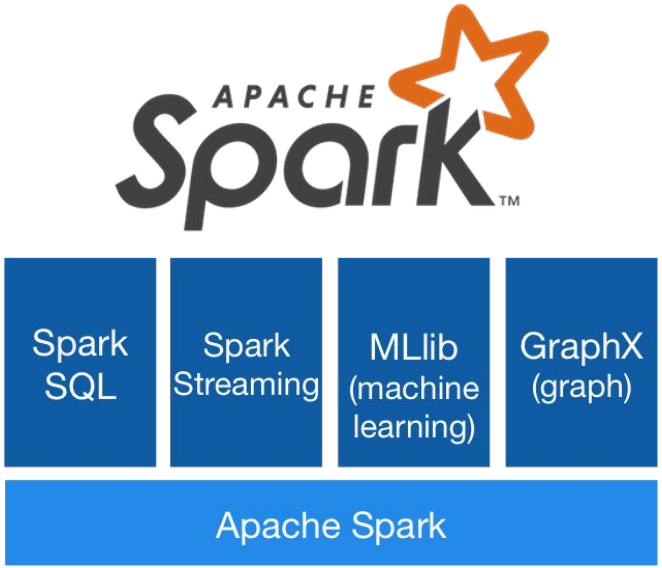
Why Spark?

*"Apache Spark is a **unified computing engine** and **a set of libraries** for parallel data processing on computer clusters"*

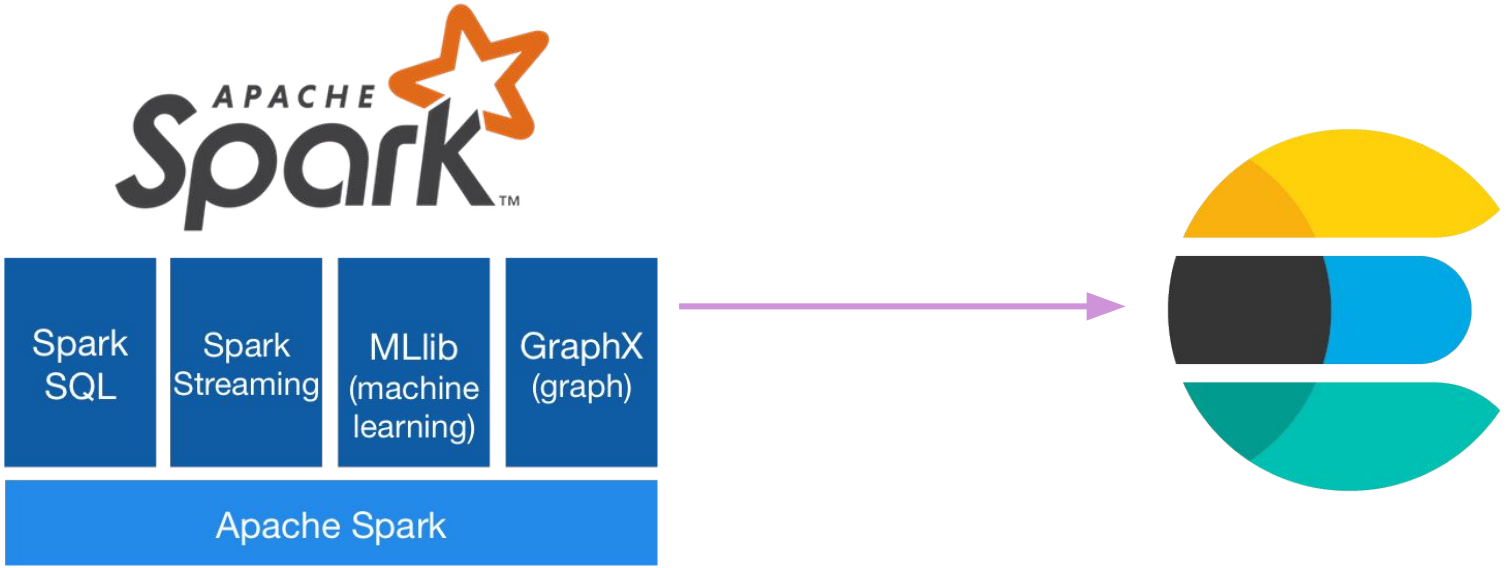
Why Spark?


APACHE STORM
Flink
talend
ORACLE
Streaming
ETL
TERADATA
HIVE
Query Language
HORTONWORKS
Data Lake
APACHE
Spark
learn
mahout
H2O
sas
Machine Learning
Graph Processing
Analysis
jupyter
GIRAPH
neo4j
redis
+ableau
pandas
Data-Driven Documents

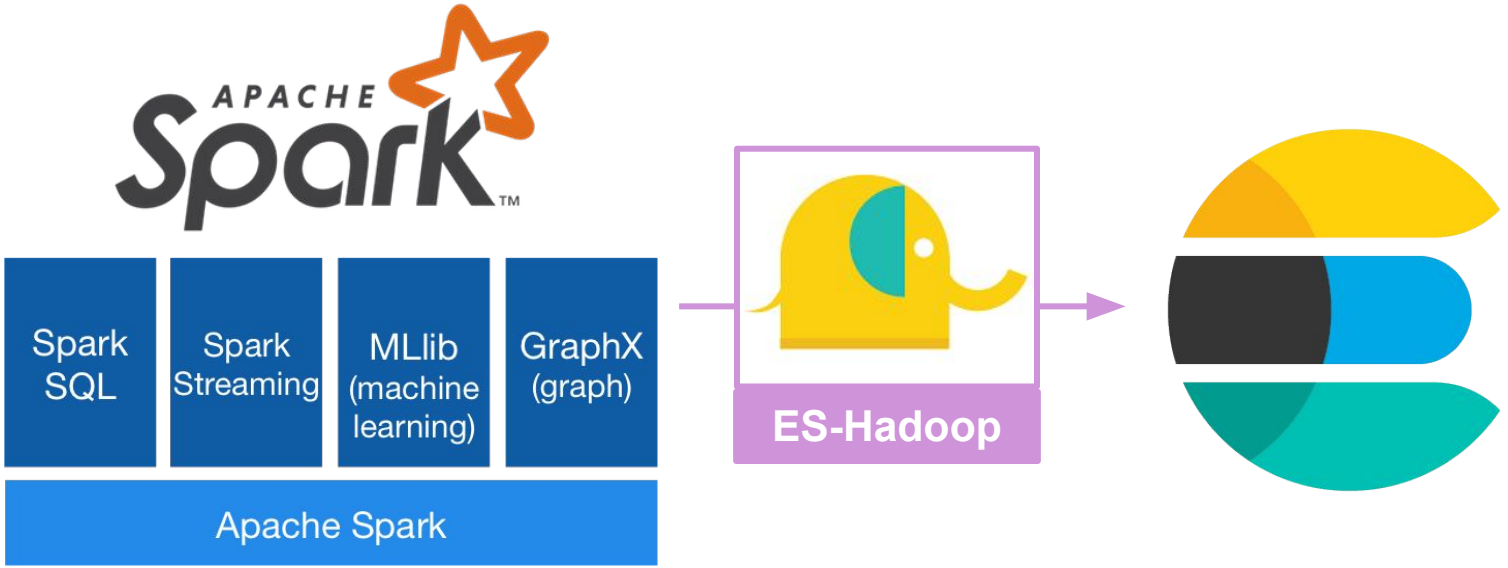
Why Spark?


APACHE
Spark
Spark SQL
Spark Streaming
MLlib (machine learning)
GraphX (graph)
Apache Spark

Why Spark?


APACHE
Spark
Spark SQL
Spark Streaming
MLlib (machine learning)
GraphX (graph)
Apache Spark

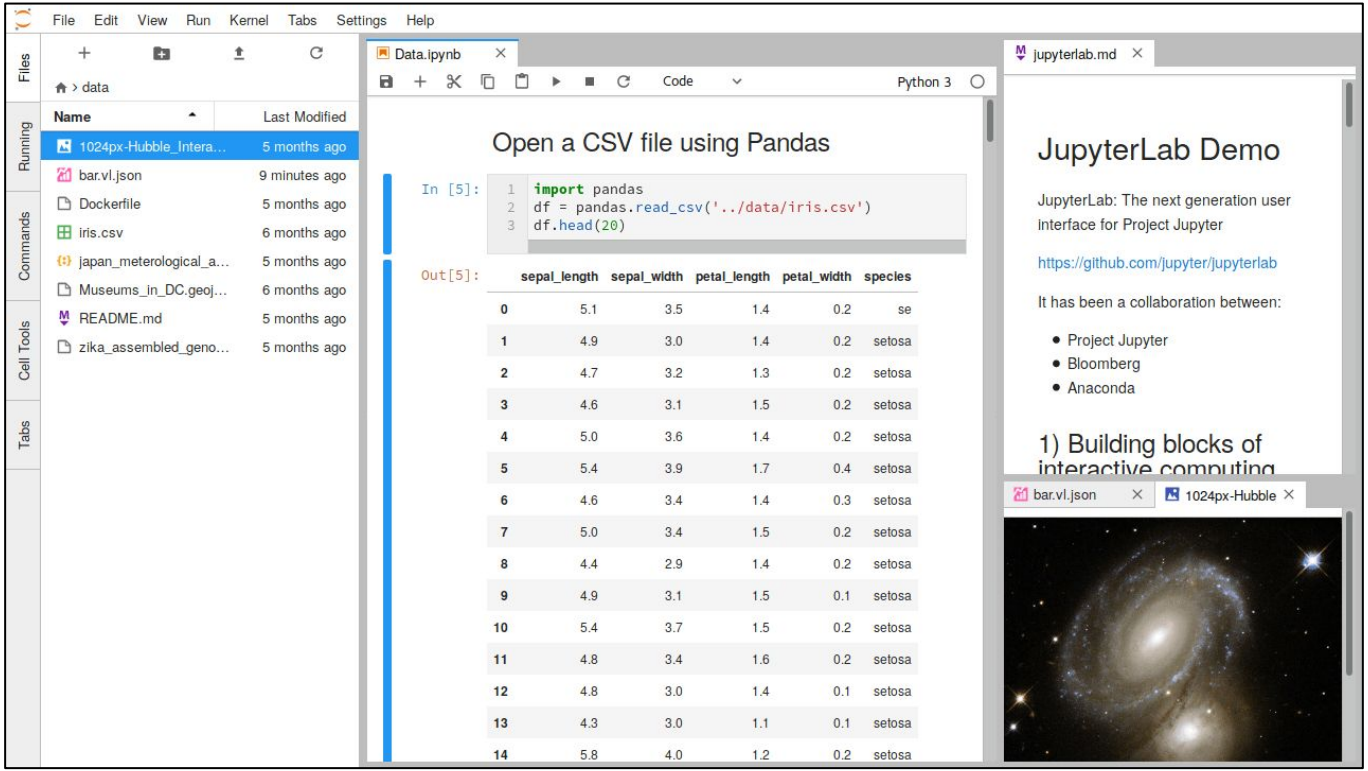
Why Spark?


APACHE
Spark
Spark SQL
Spark Streaming
MLlib (machine learning)
GraphX (graph)
Apache Spark
ES-Hadoop

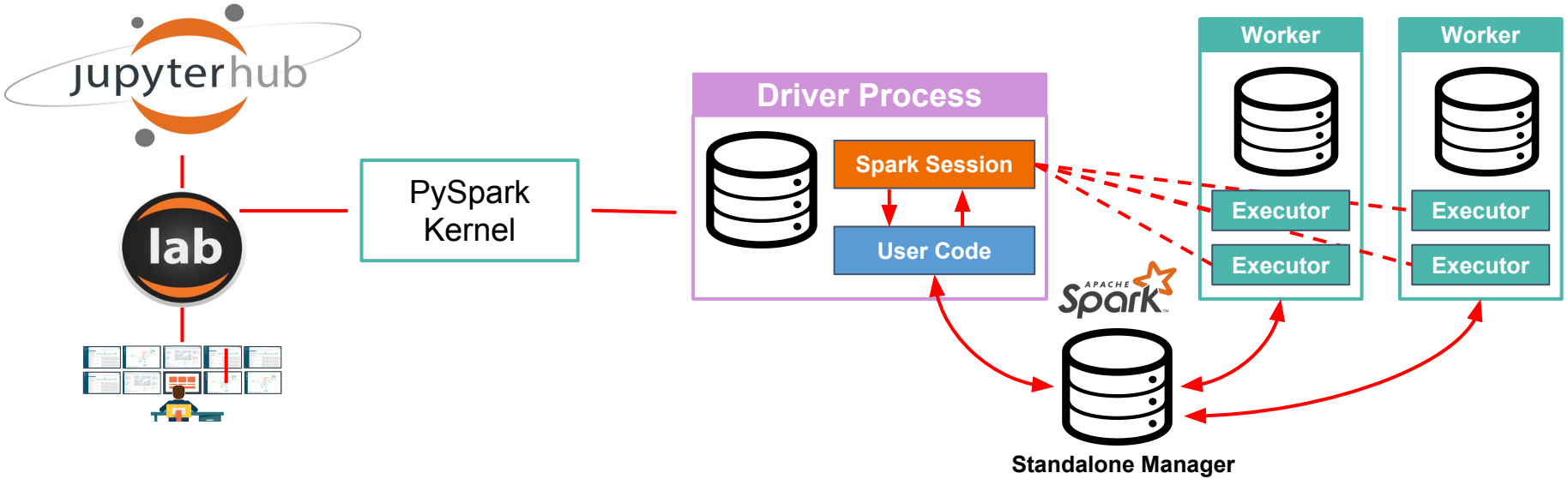
Why Jupyter?

- The Jupyter Notebook is an open-source web application that allows you to create and share documents that contain live code, equations, visualizations and narrative text.
- Uses include:
  - data cleaning and transformation
  - numerical simulation
  - statistical modeling
  - data visualization
  - machine learning, and much more.

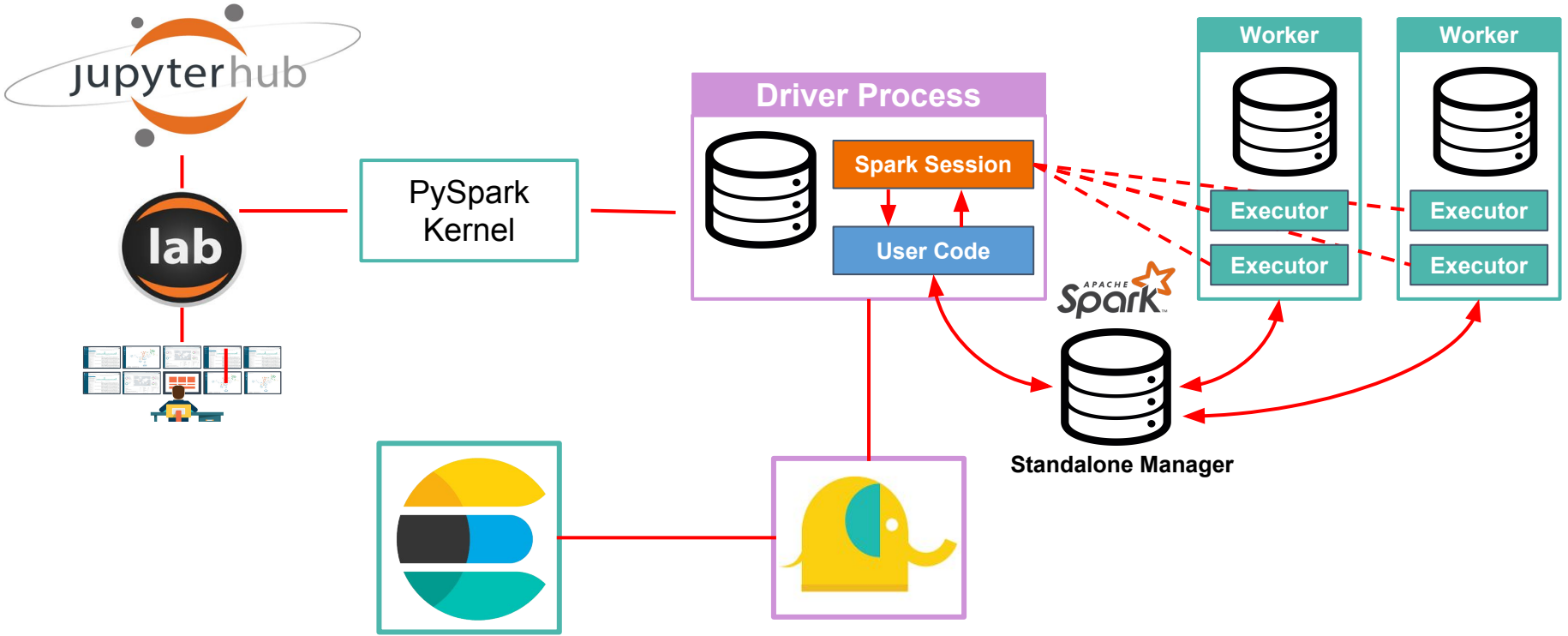
Why Jupyter?


File Edit View Run Kernel Tabs Settings Help
Files
Running
Commands
Cell Tools
Tabs
data
Name
Last Modified
1024px-Hubble_Intera... 5 months ago
bar.vl.json 9 minutes ago
Dockerfile 5 months ago
iris.csv 6 months ago
japan_meterological_a... 5 months ago
Museums_in_DC.geoj... 6 months ago
README.md 5 months ago
zika_assembled_geno... 5 months ago
Data.ipynb
Code
Python 3
Open a CSV file using Pandas
In [5]:
import pandas
df = pandas.read_csv('../data/iris.csv')
df.head(20)
Out[5]:
sepal_length sepal_width petal_length petal_width species
0 5.1 3.5 1.4 0.2 se
1 4.9 3.0 1.4 0.2 setosa
2 4.7 3.2 1.3 0.2 setosa
3 4.6 3.1 1.5 0.2 setosa
4 5.0 3.6 1.4 0.2 setosa
5 5.4 3.9 1.7 0.4 setosa
6 4.6 3.4 1.4 0.3 setosa
7 5.0 3.4 1.5 0.2 setosa
8 4.4 2.9 1.4 0.2 setosa
9 4.9 3.1 1.5 0.1 setosa
10 5.4 3.7 1.5 0.2 setosa
11 4.8 3.4 1.6 0.2 setosa
12 4.8 3.0 1.4 0.1 setosa
13 4.3 3.0 1.1 0.1 setosa
14 5.8 4.0 1.2 0.2 setosa
jupyterlab.md
JupyterLab Demo
JupyterLab: The next generation user interface for Project Jupyter
https://github.com/jupyter/jupyterlab
It has been a collaboration between:
Project Jupyter
Bloomberg
Anaconda
1) Building blocks of
bar.vl.json
1024px-Hubble

Spark + Jupyter



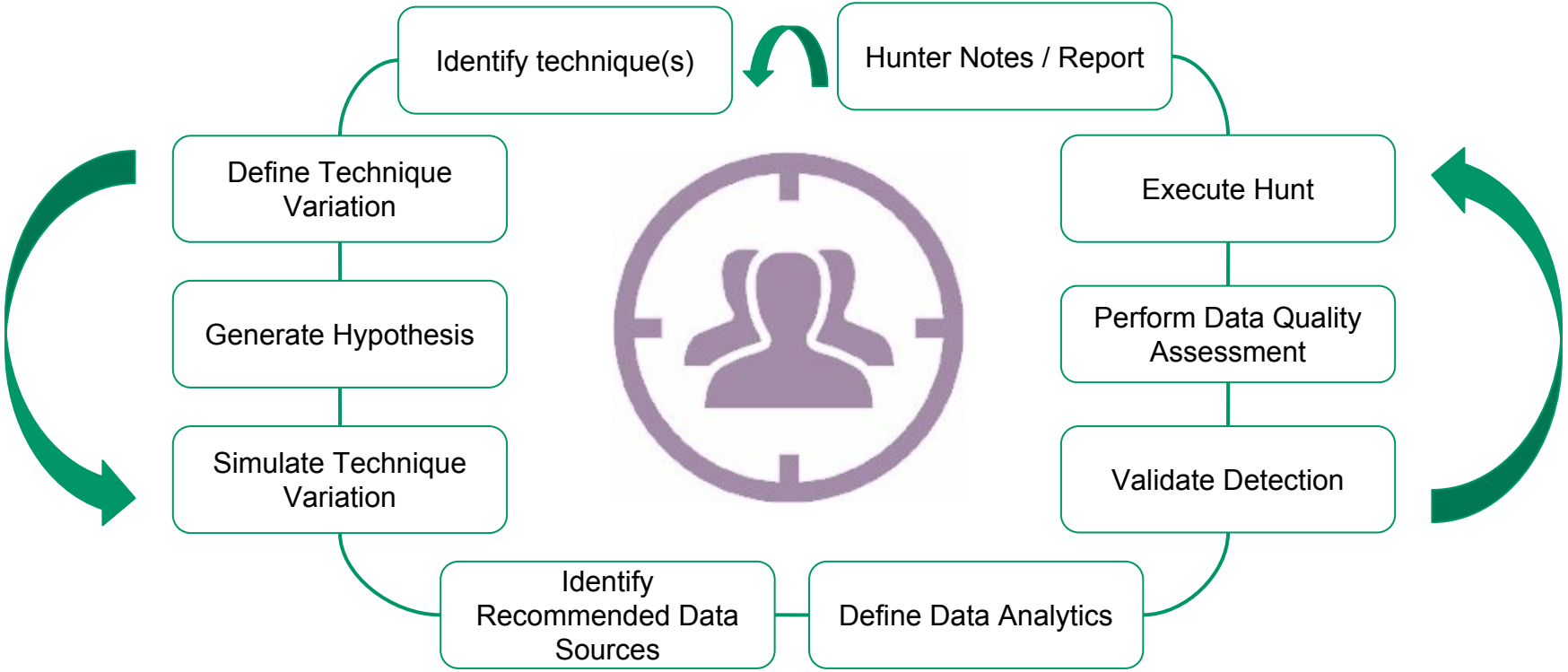
Elasticsearch + Spark + Jupyter



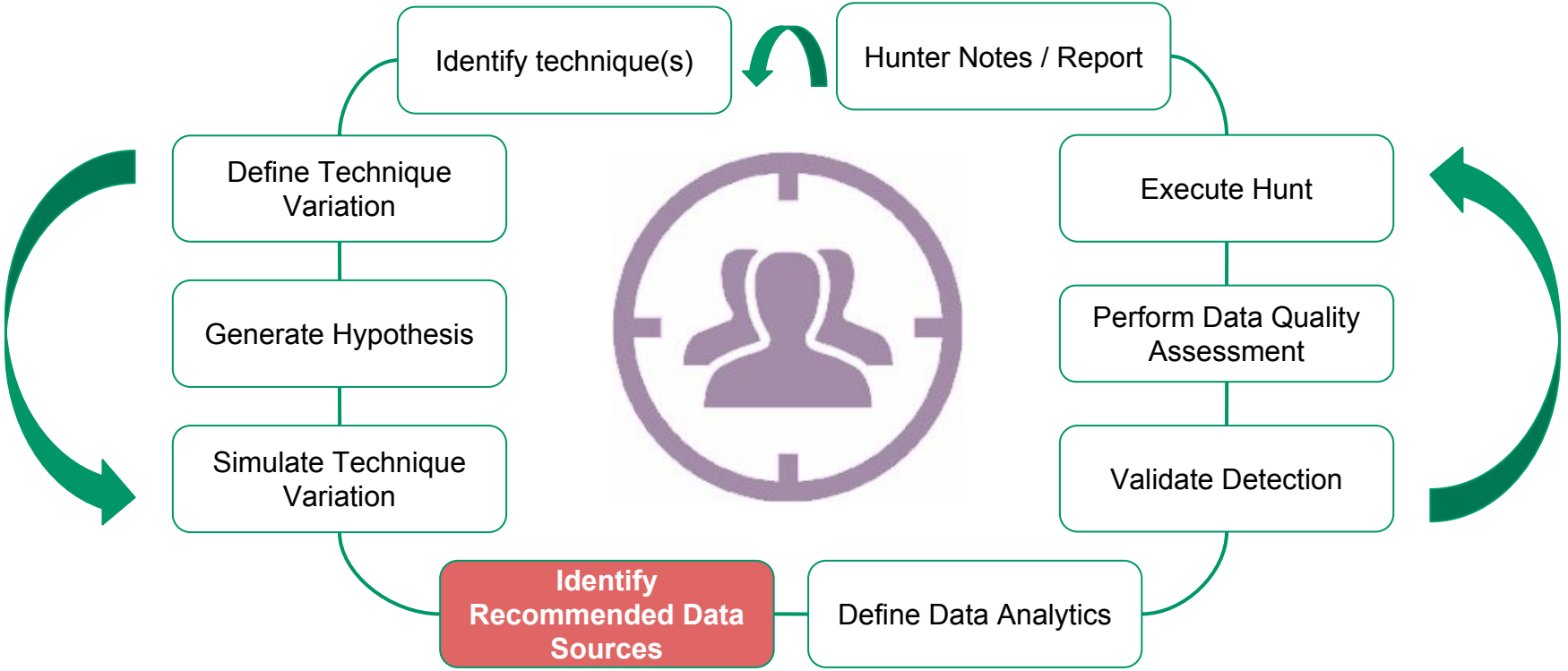
What The HELK?

HELKing the community...

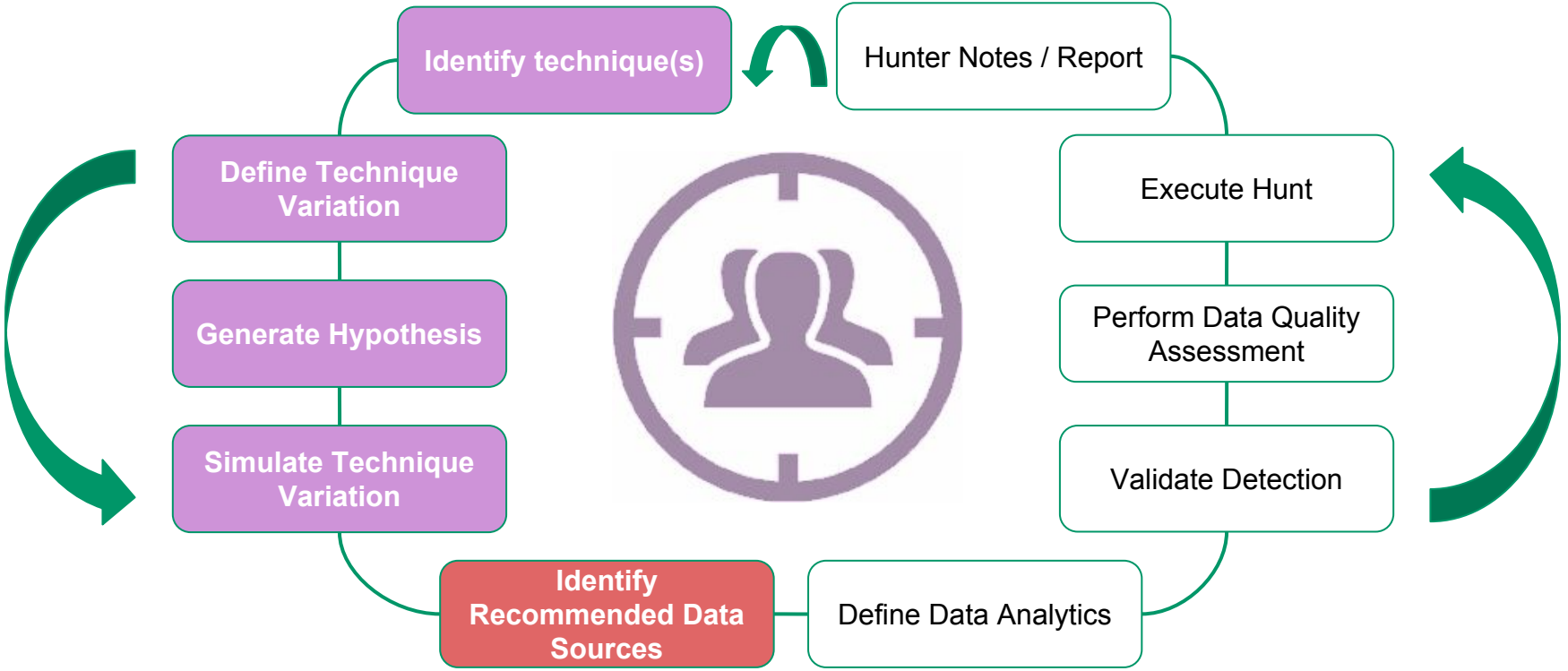
Threat Hunting Approach



Threat Hunting Approach



Threat Hunting Approach



Identify Technique: Pass the Hash

ID: T1075

Tactic: Lateral Movement

Platform: Windows

Data Sources: Authentication logs

Contributors: Travis Smith, Tripwire

Version: 1.0

"Pass the hash (PtH) is a method of authenticating as a user without having access to the user's cleartext password."

https://attack.mitre.org/wiki/Technique/T1075

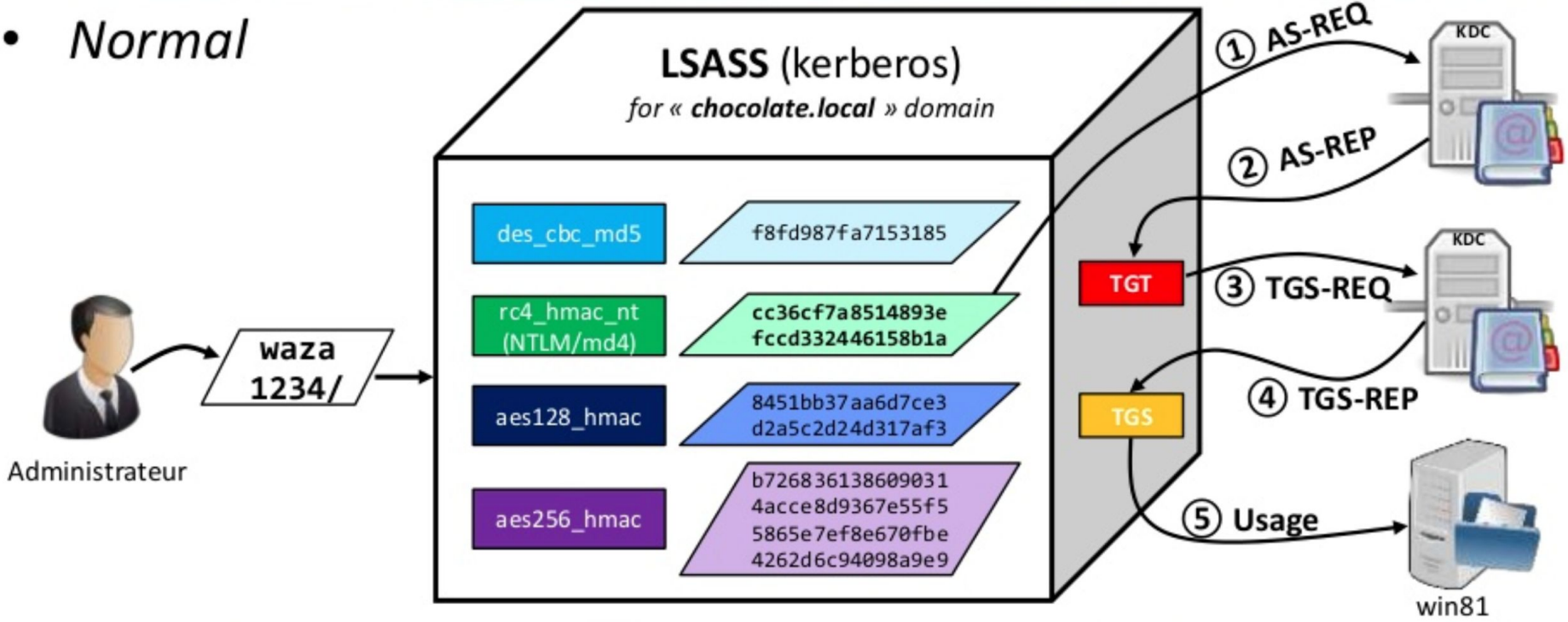
Define Technique Variation: Overpass-The-Hash

- Authentication via Kerberos
  - Authentication Protocol based on keys and tickets
- *“Upgrading a NT hash into a full Kerberos ticket”*
- It may require elevated privileges (privilege::debug or SYSTEM account)
  - Depends on how this attack is performed.
  - You might not need to be admin

https://www.blackhat.com/docs/us-14/materials/us-14-Duckwall-Abusing-Microsoft-Kerberos-Sorry-You-Guys-Don't-Get-It-wp.pdf

Technical Details: Normal kerberos Authentication

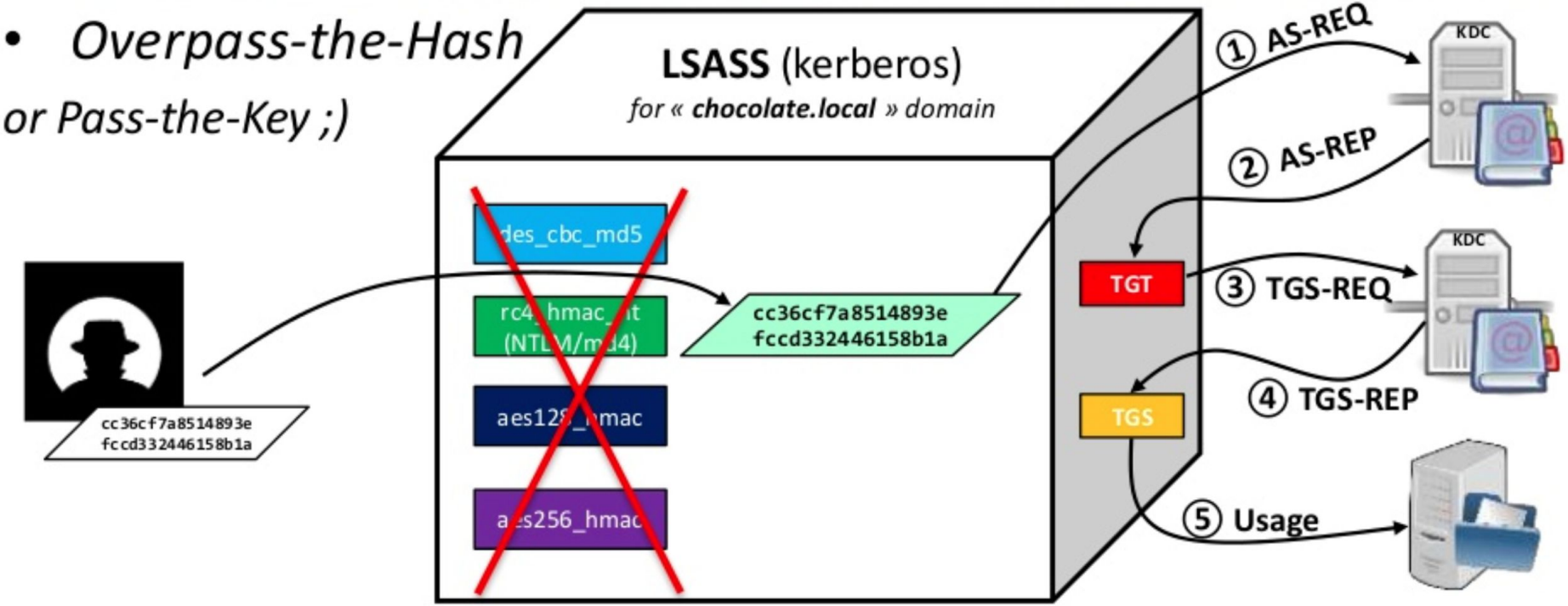
- *Normal*



Technical Details: Overpass-the-hash

- *Overpass-the-Hash*

or Pass-the-Key ;)

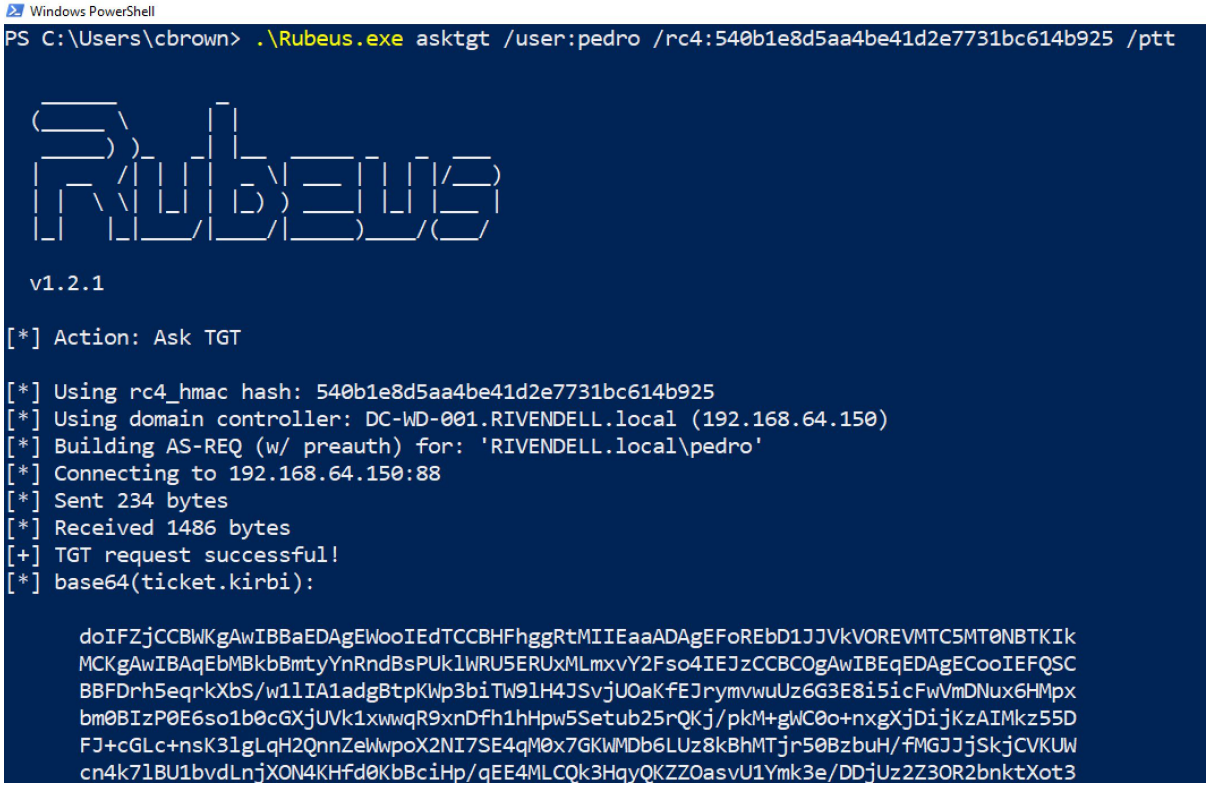


Technical Details: Rubeus - Overpass-the-hash

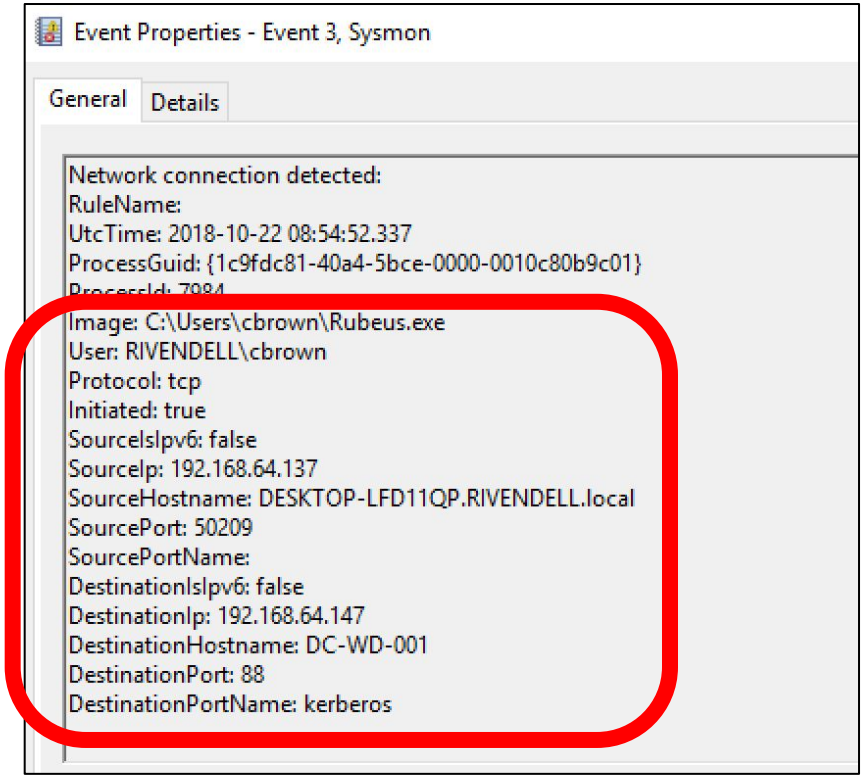
- Rubeus is a C# re-implementation of some of the functionality from Benjamin Delpy's Kekeo project
  - Kerberos structures built by hand…
  - Rubeus works nicely with execute-assembly
  - So why not use Kekeo? Because ASN.1!
    - Requires a commercial ASN.1 library to customize/rebuild the Kekeo codebase
- **Author: Will Schroeder @harmj0y @SpecterOps**
- **DEMO**

https://github.com/GhostPack/Rubeus

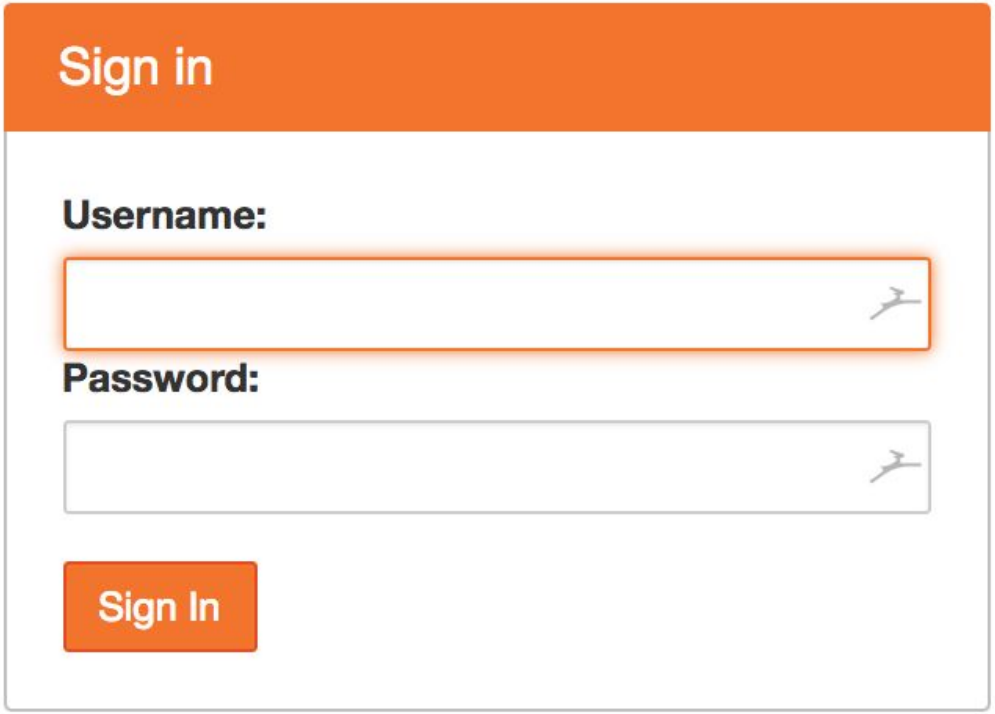
Technical Details: Rubeus - Overpass-the-hash



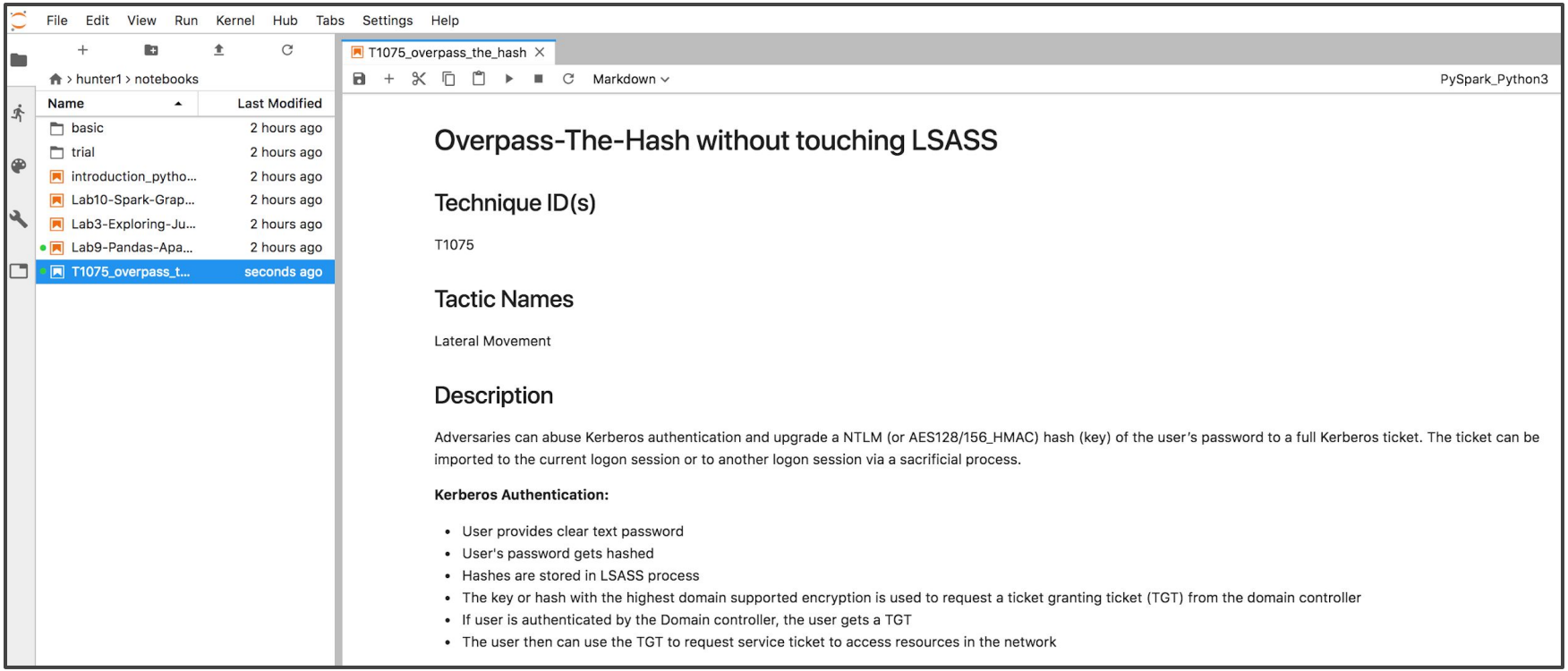
Mmmm… Rubeus?


Event Properties - Event 3, Sysmon
General
Details
Network connection detected:
RuleName:
UtcTime: 2018-10-22 08:54:52.337
ProcessGuid: {1c9fdc81-40a4-5bce-0000-0010c80b9c01}
Image: C:\Users\cbrown\Rubeus.exe
User: RIVENDELL\cbrown
Protocol: tcp
Initiated: true
SourceIsIpv6: false
SourceIp: 192.168.64.137
SourceHostname: DESKTOP-LFD11QP.RIVENDELL.local
SourcePort: 50209
SourcePortName:
DestinationIsIpv6: false
DestinationIp: 192.168.64.147
DestinationHostname: DC-WD-001
DestinationPort: 88
DestinationPortName: kerberos

Jupyter Notebooks -> Threat Hunter Playbooks



Creating a Threat Hunter Playbook



What's next for HELK?

HELKing the community...

Cypher for Apache Spark

```
val CITYFRIENDS_NA = SN_NA.cypher(
  """
    | MATCH (a:Person)-[:IS_LOCATED_IN]->(city:City)<-[:IS_LOCATED_IN]-(b:Person),
    |       (a)-[:KNOWS*1..2]->(b)
    | WHERE city.name = "New_York" OR city.name = "San_Francisco"
    | RETURN GRAPH result OF (a)-[r:SIMILAR_CIRCLE]->(b)
  """.stripMargin).graphs("result").cache

CITYFRIENDS_NA.asZeppelinGraph
```

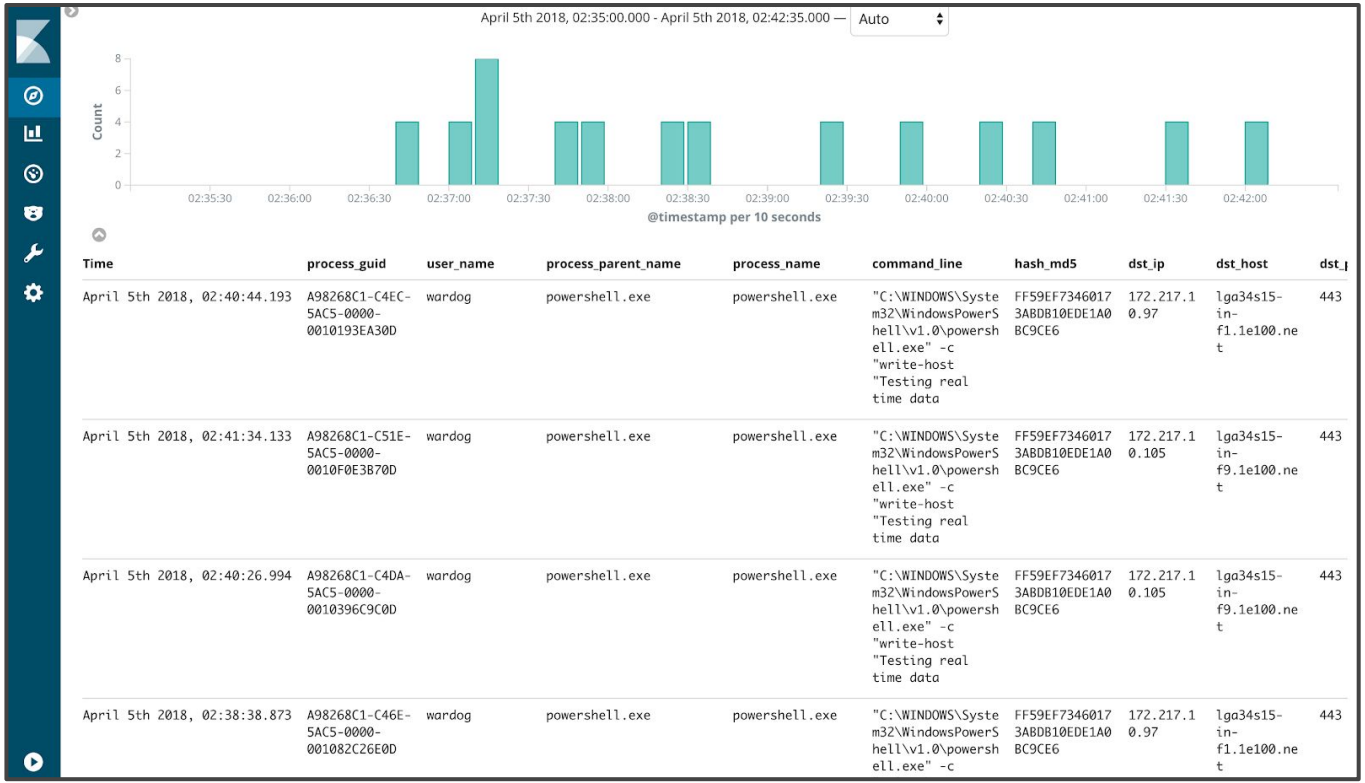
Nodes 37: Person

Relationships 99: SIMILAR_CIRCLE

Spark Structured Streaming

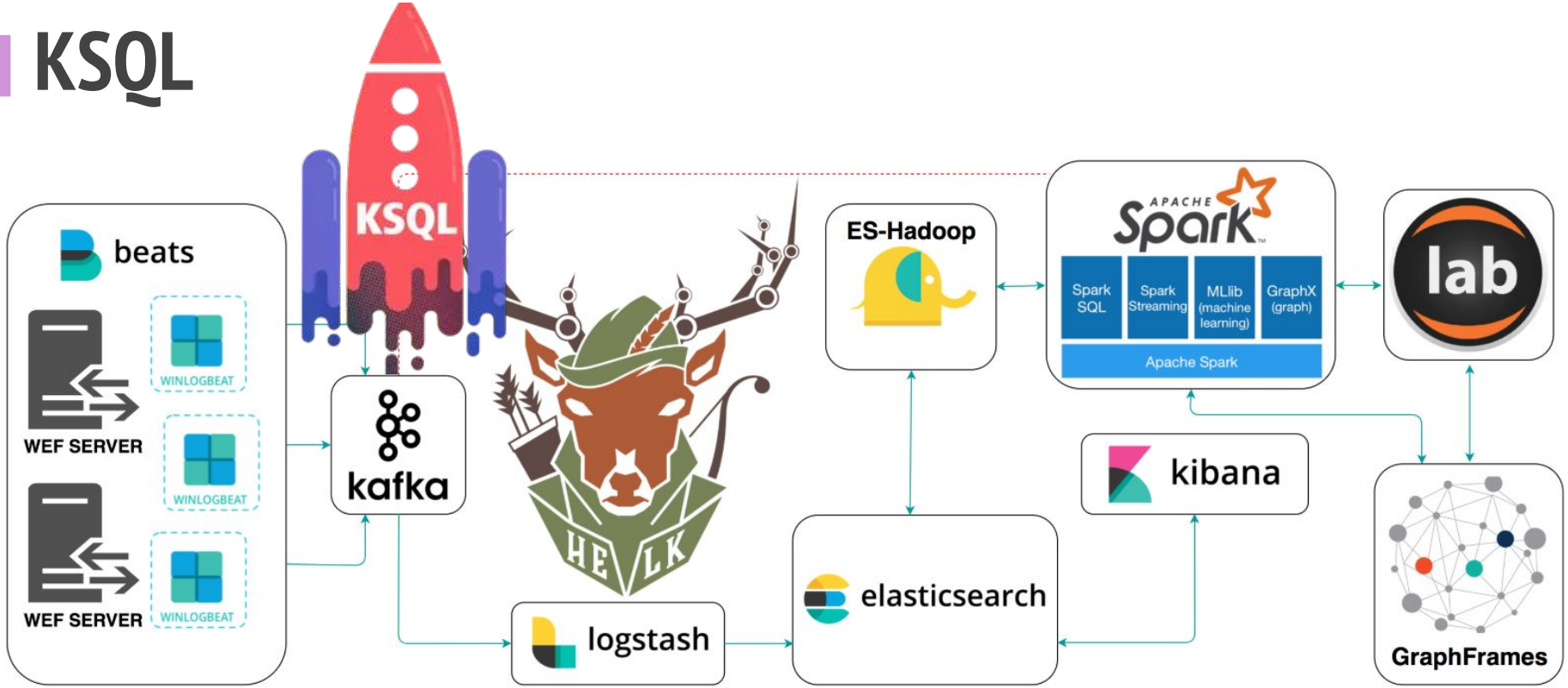
```
-------------------------------------------
Batch: 2
-------------------------------------------
+-------------------+--------------------+---------+--------------------+--------------+--------------------+--------------------+-------------+--------------------+---------------+
|           datetime|        process_guid|user_name|process_parent_name|  process_name|        command_line|            hash_md5|        dst_ip|            dst_host|dst_port_number|
+-------------------+--------------------+---------+--------------------+--------------+--------------------+--------------------+-------------+--------------------+---------------+
|2018-04-05 06:37:...|A98268C1-C419-5AC...|   wardog|     powershell.exe|powershell.exe|"C:\WINDOWS\Syste...|FF59EF73460173ABD...| 172.217.10.97|lga34s15-in-f1.1e...|            443|
|2018-04-05 06:37:...|A98268C1-C419-5AC...|   wardog|     powershell.exe|powershell.exe|"C:\WINDOWS\Syste...|FF59EF73460173ABD...|216.58.219.201|lga25s40-in-f201....|            443|
|2018-04-05 06:37:...|A98268C1-C41E-5AC...|   wardog|     powershell.exe|powershell.exe|"C:\WINDOWS\Syste...|FF59EF73460173ABD...| 172.217.10.97|lga34s15-in-f1.1e...|            443|
|2018-04-05 06:37:...|A98268C1-C41E-5AC...|   wardog|     powershell.exe|powershell.exe|"C:\WINDOWS\Syste...|FF59EF73460173ABD...|216.58.219.201|lga25s40-in-f201....|            443|
|2018-04-05 06:37:...|A98268C1-C414-5AC...|   wardog|     powershell.exe|powershell.exe|"C:\WINDOWS\Syste...|FF59EF73460173ABD...| 172.217.10.97|lga34s15-in-f1.1e...|            443|
|2018-04-05 06:37:...|A98268C1-C414-5AC...|   wardog|     powershell.exe|powershell.exe|"C:\WINDOWS\Syste...|FF59EF73460173ABD...|216.58.219.201|lga25s40-in-f201....|            443|
|2018-04-05 06:36:...|A98268C1-C408-5AC...|   wardog|     powershell.exe|powershell.exe|"C:\WINDOWS\Syste...|FF59EF73460173ABD...| 172.217.10.97|lga34s15-in-f1.1e...|            443|
|2018-04-05 06:36:...|A98268C1-C408-5AC...|   wardog|     powershell.exe|powershell.exe|"C:\WINDOWS\Syste...|FF59EF73460173ABD...|216.58.219.201|lga25s40-in-f201....|            443|
+-------------------+--------------------+---------+--------------------+--------------+--------------------+--------------------+-------------+--------------------+---------------+

-------------------------------------------
Batch: 3
-------------------------------------------
+-------------------+--------------------+---------+--------------------+--------------+--------------------+--------------------+-------------+--------------------+---------------+
|           datetime|        process_guid|user_name|process_parent_name|  process_name|        command_line|            hash_md5|        dst_ip|            dst_host|dst_port_number|
+-------------------+--------------------+---------+--------------------+--------------+--------------------+--------------------+-------------+--------------------+---------------+
|2018-04-05 06:37:...|A98268C1-C435-5AC...|   wardog|     powershell.exe|powershell.exe|"C:\WINDOWS\Syste...|FF59EF73460173ABD...| 172.217.10.97|lga34s15-in-f1.1e...|            443|
|2018-04-05 06:37:...|A98268C1-C435-5AC...|   wardog|     powershell.exe|powershell.exe|"C:\WINDOWS\Syste...|FF59EF73460173ABD...|216.58.219.201|lga25s40-in-f201....|            443|
+-------------------+--------------------+---------+--------------------+--------------+--------------------+--------------------+-------------+--------------------+---------------+

-------------------------------------------
Batch: 4
-------------------------------------------
+-------------------+--------------------+---------+--------------------+--------------+--------------------+--------------------+-------------+--------------------+---------------+
|           datetime|        process_guid|user_name|process_parent_name|  process_name|        command_line|            hash_md5|        dst_ip|            dst_host|dst_port_number|
+-------------------+--------------------+---------+--------------------+--------------+--------------------+--------------------+-------------+--------------------+---------------+
|2018-04-05 06:37:...|A98268C1-C43A-5AC...|   wardog|     powershell.exe|powershell.exe|"C:\WINDOWS\Syste...|FF59EF73460173ABD...| 172.217.10.97|lga34s15-in-f1.1e...|            443|
|2018-04-05 06:37:...|A98268C1-C43A-5AC...|   wardog|     powershell.exe|powershell.exe|"C:\WINDOWS\Syste...|FF59EF73460173ABD...|216.58.219.201|lga25s40-in-f201....|            443|
+-------------------+--------------------+---------+--------------------+--------------+--------------------+--------------------+-------------+--------------------+---------------+

-------------------------------------------
Batch: 5
-------------------------------------------
+-------------------+--------------------+---------+--------------------+--------------+--------------------+--------------------+-------------+--------------------+---------------+
|           datetime|        process_guid|user_name|process_parent_name|  process_name|        command_line|            hash_md5|        dst_ip|            dst_host|dst_port_number|
+-------------------+--------------------+---------+--------------------+--------------+--------------------+--------------------+-------------+--------------------+---------------+
|2018-04-05 06:37:...|A98268C1-C43F-5AC...|   wardog|     powershell.exe|powershell.exe|"C:\WINDOWS\Syste...|FF59EF73460173ABD...| 172.217.10.97|lga34s15-in-f1.1e...|            443|
|2018-04-05 06:37:...|A98268C1-C43F-5AC...|   wardog|     powershell.exe|powershell.exe|"C:\WINDOWS\Syste...|FF59EF73460173ABD...|216.58.219.201|lga25s40-in-f201....|            443|
|2018-04-05 06:37:...|A98268C1-C443-5AC...|   wardog|     powershell.exe|powershell.exe|"C:\WINDOWS\Syste...|FF59EF73460173ABD...| 172.217.10.97|lga34s15-in-f1.1e...|            443|
|2018-04-05 06:37:...|A98268C1-C443-5AC...|   wardog|     powershell.exe|powershell.exe|"C:\WINDOWS\Syste...|FF59EF73460173ABD...|216.58.219.201|lga25s40-in-f201....|            443|
+-------------------+--------------------+---------+--------------------+--------------+--------------------+--------------------+-------------+--------------------+---------------+
```

Spark Structured Streaming



Time	process_guid	user_name	process_parent_name	process_name	command_line	hash_md5	dst_ip	dst_host	dst_p
April 5th 2018, 02:40:44.193	A98268C1-C4EC-5AC5-0000-0010193EA30D	wardog	powershell.exe	powershell.exe	"C:\WINDOWS\System32\WindowsPowerShell\v1.0\powershell.exe" -c "write-host "Testing real time data	FF59EF73460173ABDB10EDE1A0BC9CE6	172.217.10.97	lga34s15-in-f1.1e100.net	443
April 5th 2018, 02:41:34.133	A98268C1-C51E-5AC5-0000-0010F0E3B70D	wardog	powershell.exe	powershell.exe	"C:\WINDOWS\System32\WindowsPowerShell\v1.0\powershell.exe" -c "write-host "Testing real time data	FF59EF73460173ABDB10EDE1A0BC9CE6	172.217.10.105	lga34s15-in-f9.1e100.net	443
April 5th 2018, 02:40:26.994	A98268C1-C4DA-5AC5-0000-0010396C9C0D	wardog	powershell.exe	powershell.exe	"C:\WINDOWS\System32\WindowsPowerShell\v1.0\powershell.exe" -c "write-host "Testing real time data	FF59EF73460173ABDB10EDE1A0BC9CE6	172.217.10.105	lga34s15-in-f9.1e100.net	443
April 5th 2018, 02:38:38.873	A98268C1-C46E-5AC5-0000-001082C26E0D	wardog	powershell.exe	powershell.exe	"C:\WINDOWS\System32\WindowsPowerShell\v1.0\powershell.exe" -c	FF59EF73460173ABDB10EDE1A0BC9CE6	172.217.10.97	lga34s15-in-f1.1e100.net	443

KSQL


KSQL
beats
WINLOGBEAT
WEF SERVER
WINLOGBEAT
WEF SERVER
WINLOGBEAT
kafka
HE LK
logstash
ES-Hadoop
elasticsearch
APACHE Spark
Spark SQL
Spark Streaming
MLlib (machine learning)
GraphX (graph)
Apache Spark
lab
kibana
GraphFrames

Thank You! Muchas Gracias!