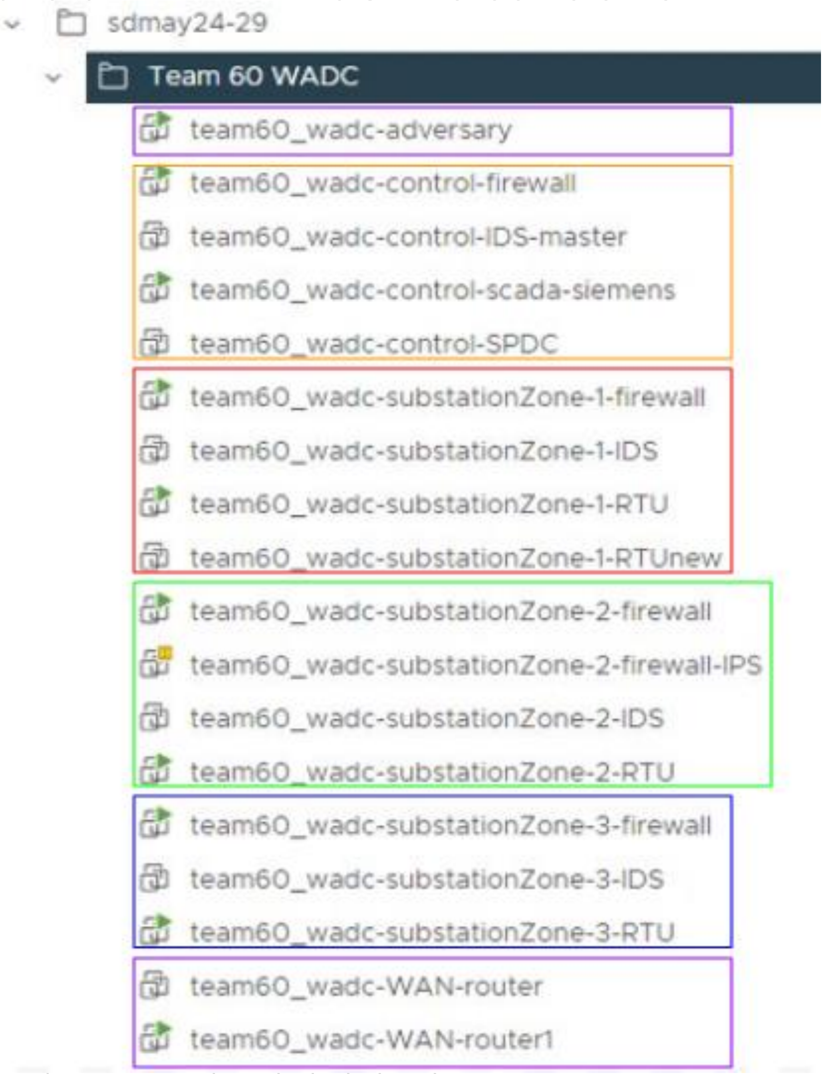
Grid-SIEM: Cybersecurity for Power Grid Using SIEM and Machine Learning Tools

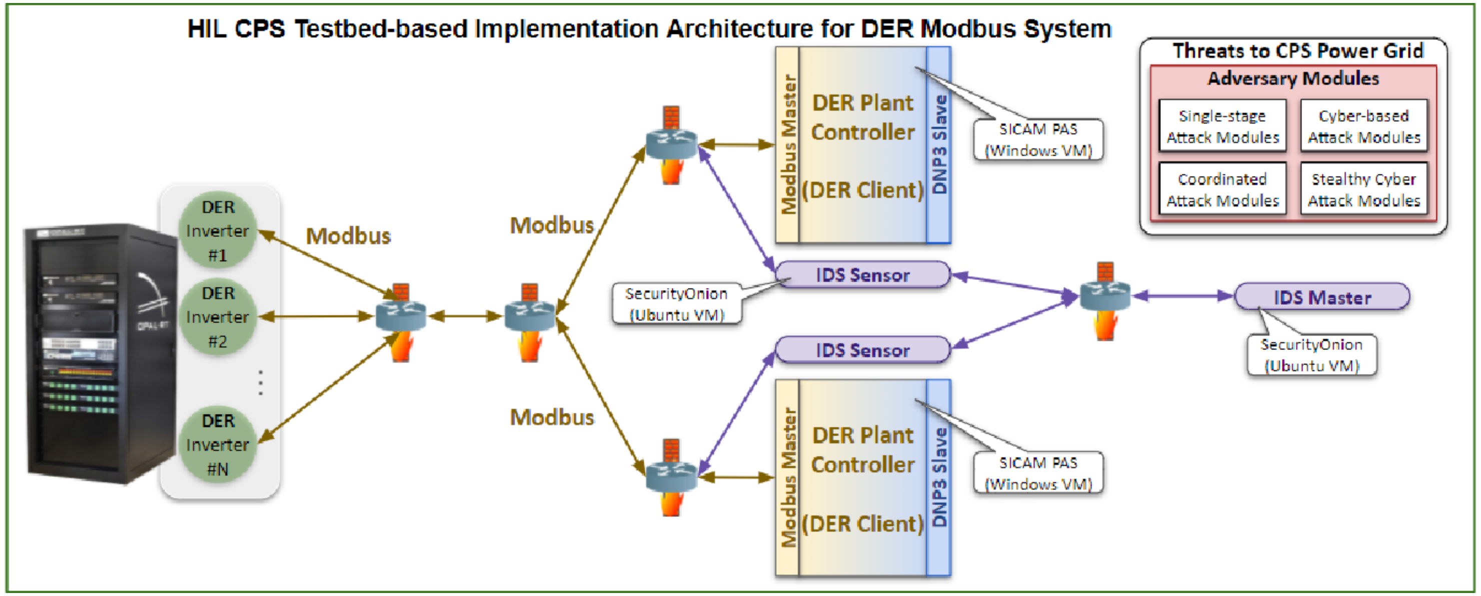
Trent Bickford, Westin Chamberlain, Ella Cook, Daniel Ocampo

Dr. Gelli Ravikumar

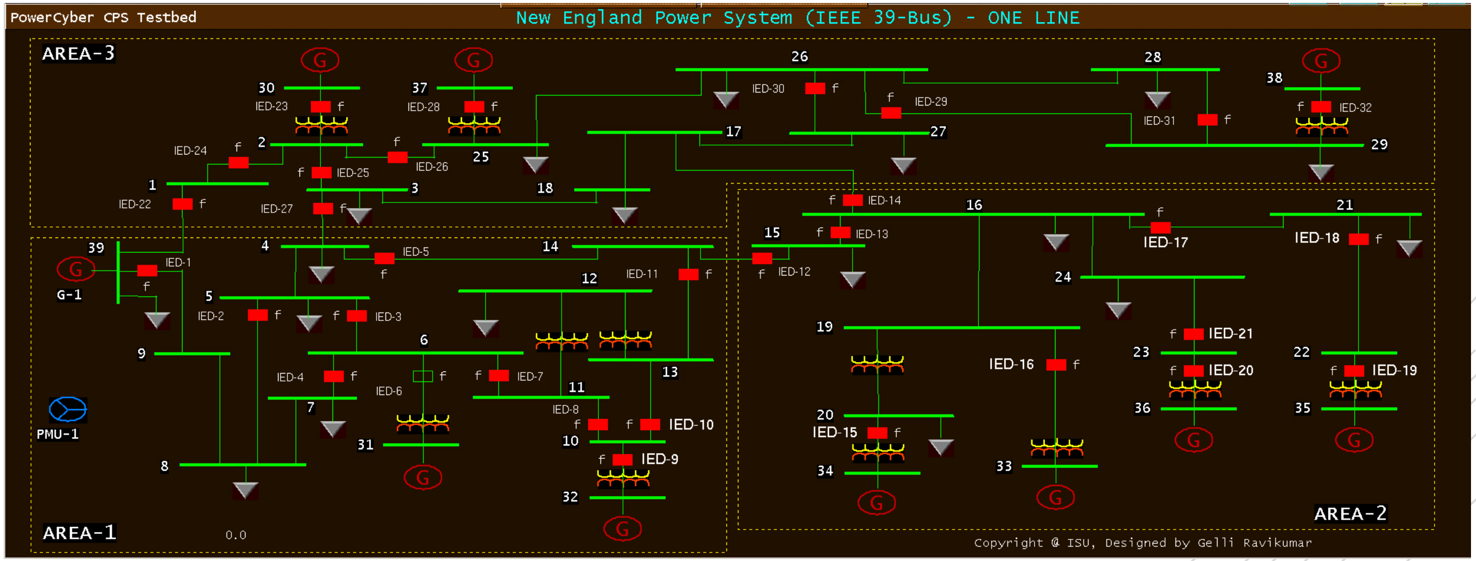
Grid-SIEM Project Context



- sdmay24-29
  - Team 60 WADC
    - team60_wadc-adversary
    - team60_wadc-control-firewall
    - team60_wadc-control-IDS-master
    - team60_wadc-control-scada-siemens
    - team60_wadc-control-SPDC
    - team60_wadc-substationZone-1-firewall
    - team60_wadc-substationZone-1-IDS
    - team60_wadc-substationZone-1-RTU
    - team60_wadc-substationZone-1-RTUnew
    - team60_wadc-substationZone-2-firewall
    - team60_wadc-substationZone-2-firewall-IPS
    - team60_wadc-substationZone-2-IDS
    - team60_wadc-substationZone-2-RTU
    - team60_wadc-substationZone-3-firewall
    - team60_wadc-substationZone-3-IDS
    - team60_wadc-substationZone-3-RTU
    - team60_wadc-WAN-router
    - team60_wadc-WAN-router1



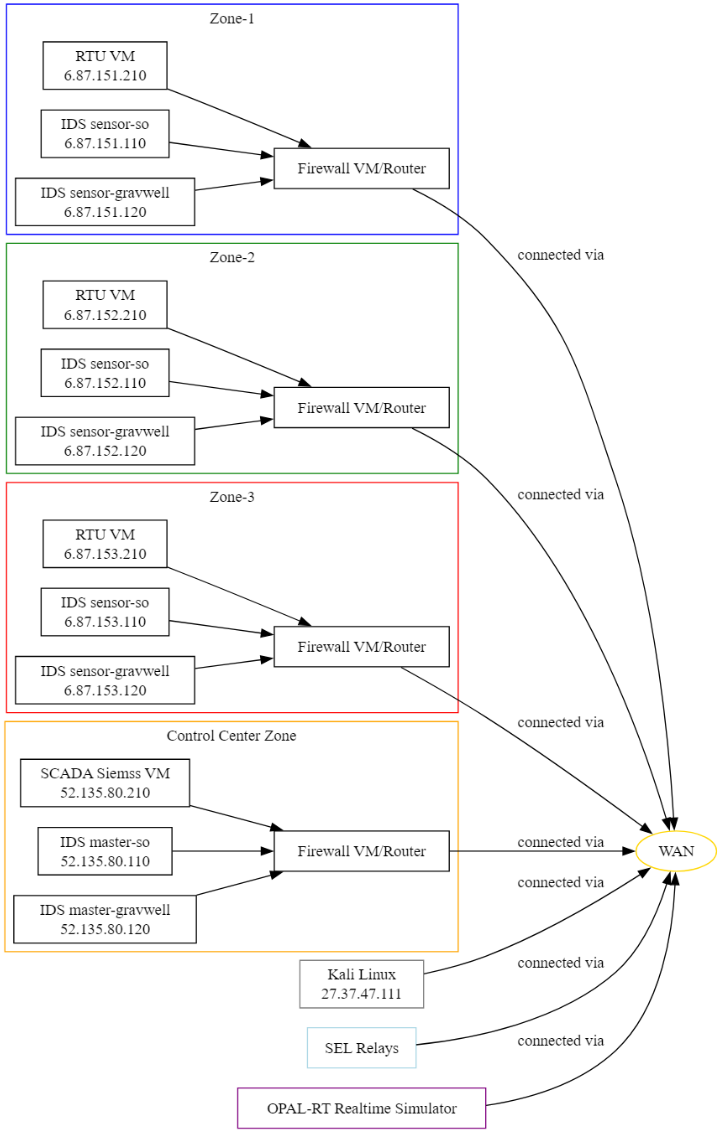
(G. Ravikumar)


PowerCyber CPS Testbed
New England Power System (IEEE 39-Bus) - ONE LINE
AREA-3
AREA-1
AREA-2
G-1
PMU-1
0.0
Copyright @ ISU, Designed by Gelli Ravikumar

Project Vision

- 3 main tasks
  - Integrate a Security Information Event Management (SIEM) platform into the existing virtual power system
  - Launch cyber-attacks against the SIEM implementation
  - Developing a Machine Learning (ML) component to further enhance the SIEM
- Who cares?
  - Developers of PowerCyber infrastructure at ISU
  - The IT community focused on securing industrial control systems (ICS)
  - Power grid management system operators
  - People who benefit from the use of power grids
- Use cases?
  - Further research with the PowerCyber testbed environment
  - Power grid systems looking to increase their security
- Important Definitions
  - OT: Operational Technology
  - APT: Advanced Persistent Threat
  - ICS: Industrial Control Systems

Conceptual/Visual Sketch



Zone-1

RTU VM 6.87.151.210

IDS sensor-so 6.87.151.110

IDS sensor-gravwell 6.87.151.120

Firewall VM/Router

Zone-2

RTU VM 6.87.152.210

IDS sensor-so 6.87.152.110

IDS sensor-gravwell 6.87.152.120

Firewall VM/Router

Zone-3

RTU VM 6.87.153.210

IDS sensor-so 6.87.153.110

IDS sensor-gravwell 6.87.153.120

Firewall VM/Router

Control Center Zone

SCADA Siemss VM 52.135.80.210

IDS master-so 52.135.80.110

IDS master-gravwell 52.135.80.120

Firewall VM/Router

Kali Linux 27.37.47.111

SEL Relays

OPAL-RT Realtime Simulator

connected via

connected via

connected via

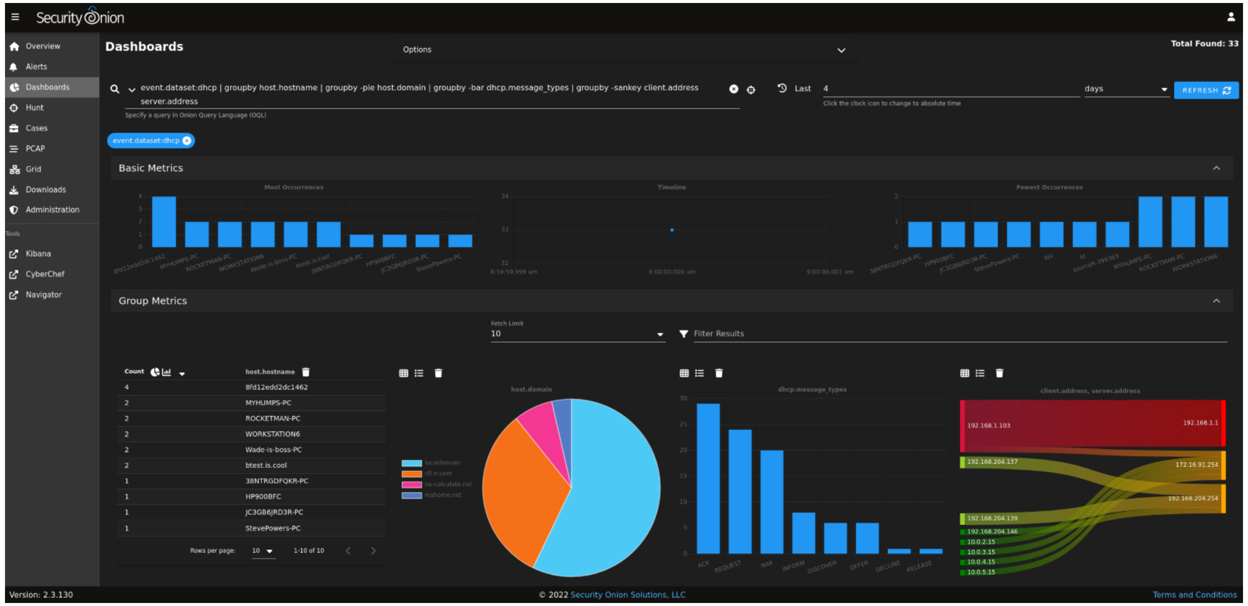
connected via

connected via

connected via

connected via

WAN



event.dataset:dhcp | groupby host.hostname | groupby -pie host.domain | groupby -bar dhcp.message_types | groupby -sankey client.address server.address

Specify a query in Onion Query Language (OQL)

Last 4 days

Click the clock icon to change to absolute time

Total Found: 33

Count	host.hostname
4	8fd12edd2dc1462
2	MYHUMPS-PC
2	ROCKETMAN-PC
2	WORKSTATION6
2	Wade-is-boss-PC
2	btest.is.cool
1	38NTRGDFQKR-PC
1	HP900BFC
1	JC3GB6JRD3R-PC
1	StevePowers-PC

Rows per page: 10 1-10 of 10

(blog.securityonion.net)

Functional Requirements

- SIEM functionalities
  - Able to detect attacks
  - Integrate machine learning for further detection
  - Forward nodes to collect information from PowerCyber infrastructure
- Analysis
  - PowerCyber system information displayed on Security Onion dashboard
  - Implementation should be able to detect launched Caldera attacks through SO
  - Machine learning should assist Security Onion in detecting unknown attacks
- Performance & Reliability
  - SIEM should have near 99.99% uptime
  - SCADA/ICS should have 99.99% availability
  - Machine learning should be capable of detecting incidents effectively

Non-Functional Requirements & Technical Constraints

- Non-Functional Requirements
    - Usability
        - Usability of SO at an administrator level must be user-friendly
    - Scalability
        - System should be able to efficiently handle increasing workloads without a decrease in performance
            - Accommodate higher levels of network traffic with ease
    - Maintainability
        - Should be able to accommodate future updates and maintenance
            - Clear documentation & use of standard technologies
- Technical Constraints
    - Resource limitations
        - VMware vSphere, PowerCyber infrastructure, storage space for logs
    - Uptime constraints
        - ICS must have an uptime of 99.999% availability

Technologies, Frameworks, & Standards

Technologies and Frameworks

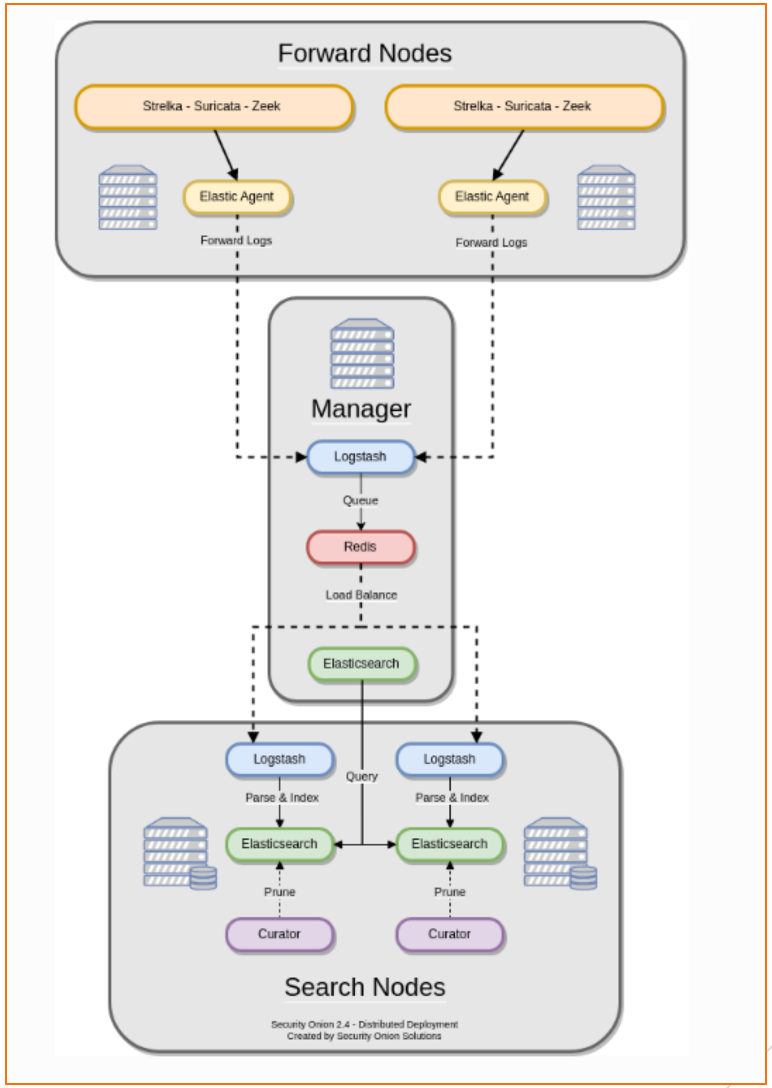
- Security Onion
- Gravwell
- Mitre Caldera
- VMware vSphere
- SciKit and Pandas
- Contagio

Standards

- ISO/IEC 27001: Provides pointers into managing cyber risk and resilience throughout project lifecycle.
- NIST Cyber Framework 2.0: Industry and government guidance to best follow modern cyber security practices.
- MITRE Attack/Defend Framework: will be used along with MITRE Caldera to identify and model threats and attacks against the power grid. In addition to assisting with defensive strategies.
- IEEE C37.2040: Cybersecurity Requirements for Substation Automation, Protection, and Control Systems - The automation of the power grid and security measures will follow this standard.
- IEEE P1402: Physical Security of Electrical Power Substations - The physical security of the PowerCyber environment will align with the IEEE P1402 standard to mitigate risk.
- NVD CVSS v3.0: Used to score the severity of the attacks we create and test.
- IEEE P2863: Recommended Practice for Organizational Governance of Artificial Intelligence - Specifies implementation and compliance with artificial intelligence.

SIEM Components - Security Onion

- Security Onion Forward Node
  - The SIEM sensors that will collect data from each of the respective zones
- Security Onion Manager Node
  - The SIEM master node where the logs and data collected by the sensors will aggregate
  - This is the node that user will be able to see the Security Onion Console (SOC)
  - Additions such as Kibana, CyberChef, and ATT&CK Navigator

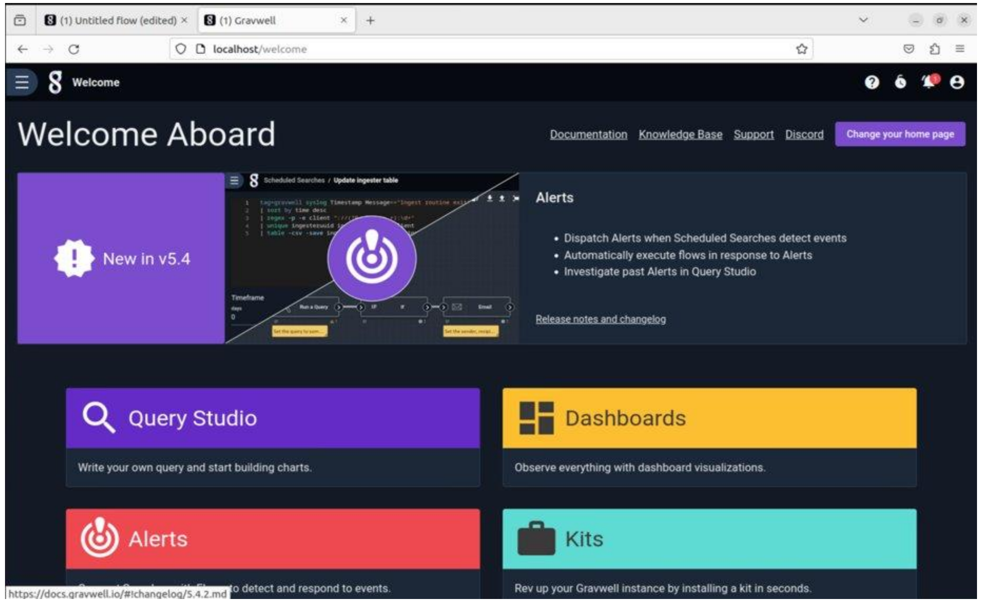


(securityonion.net/architecture)

Security Onion Implementation

SIEM Components - Gravwell

- Gravwell serves as an experimental tool for its data analysis capabilities.
- We do not expect to rely on it as we do Security Onion. Is not recommended as a standalone SIEM platform.
- Gravwell meeting Nov. 14.
- It has been challenging to funnel data into our Gravwell indexer. The best option is to feed it completed pcap files for attack analysis.
- Limited by Community Edition license.
- Workflow: automated search scripts that can be scheduled to detect malicious behavior.
- Gravwell has a feature called Backfill scheduling which can perform the script after an update is done. So, information from that time period is not lost.
- Playbooks and flows.



Welcome Aboard

Alerts

- Dispatch Alerts when Scheduled Searches detect events
- Automatically execute flows in response to Alerts
- Investigate past Alerts in Query Studio

Release notes and changelog

Write your own query and start building charts.

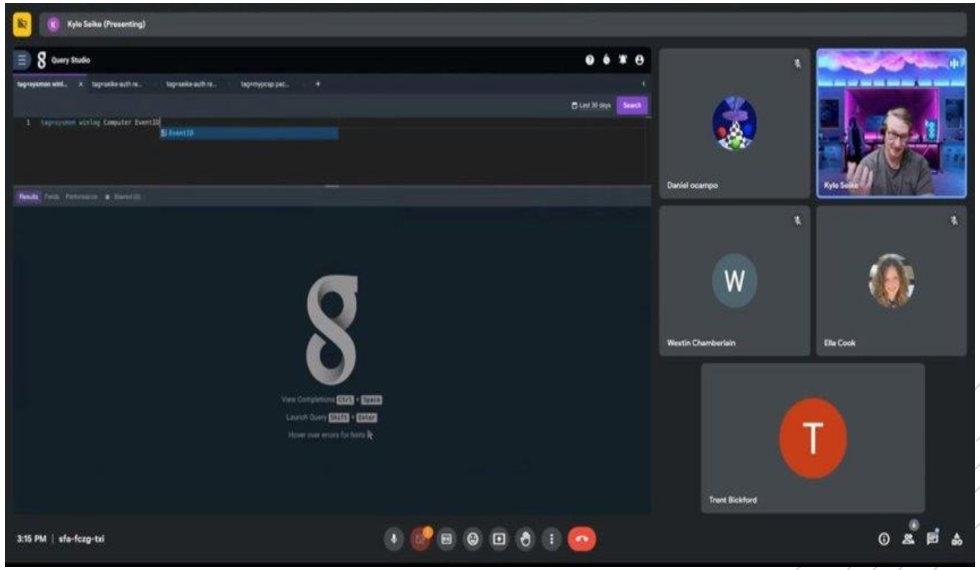
Observe everything with dashboard visualizations.

Alerts

to detect and respond to events.

Rev up your Gravwell instance by installing a kit in seconds.

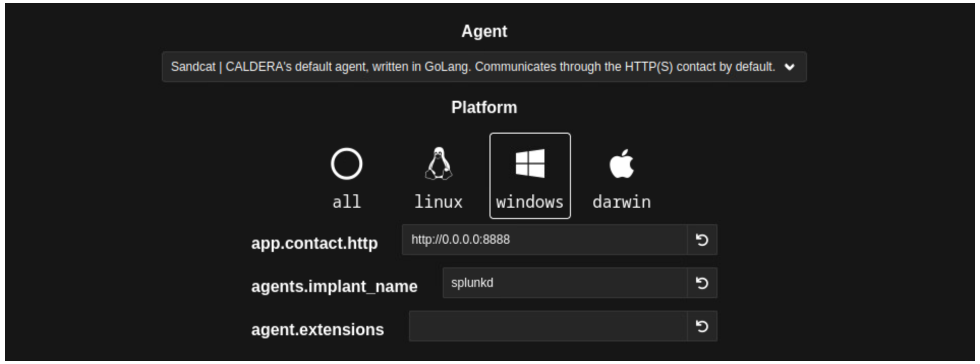
https://docs.gravwell.io/#!changelog/5.4.2.md



(Gravwell meeting. Our own dashboard)

Mitre Caldera

- Autonomous attacking
  - Deploy agents
  - Create adversary profiles
    - Where the autonomous comes from
    - Plugins
  - Begin operation
    - After completion, view logs
- Plugins
  - Adds customization
  - Modbus, Dnp3, bacnet
- Challenges
  - Requires PowerShell on target machine
  - Requires less than ideal firewall



Agent

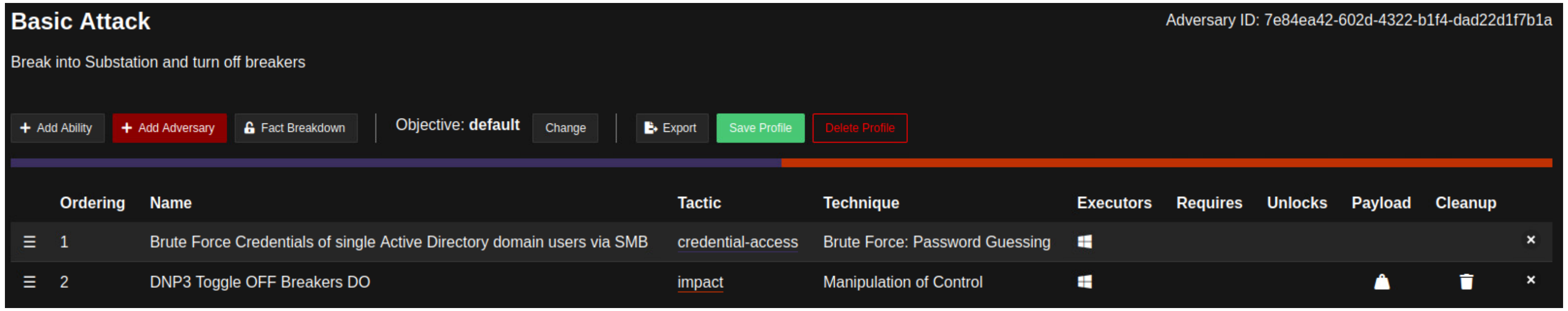
Sandcat | CALDERA's default agent, written in GoLang. Communicates through the HTTP(S) contact by default.

Platform

app.contact.http: http://0.0.0.0:8888

agents.implant_name: splunkd

agent.extensions:



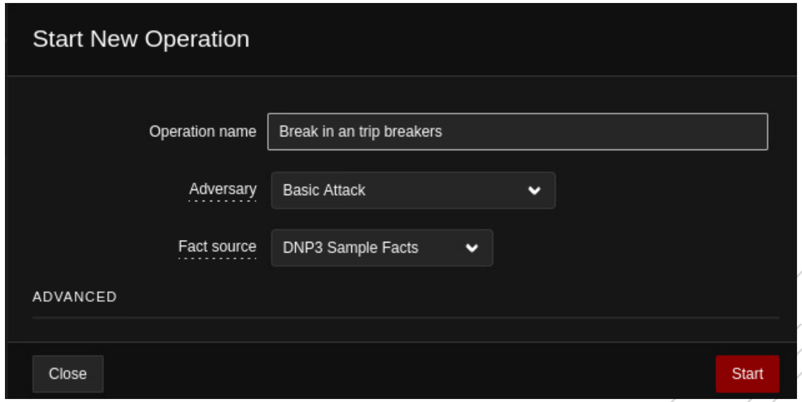
Basic Attack

Adversary ID: 7e84ea42-602d-4322-b1f4-dad22d1f7b1a

Break into Substation and turn off breakers

Objective: **default**

Ordering	Name	Tactic	Technique	Executors	Requires	Unlocks	Payload	Cleanup
1	Brute Force Credentials of single Active Directory domain users via SMB	credential-access	Brute Force: Password Guessing	⊞				
2	DNP3 Toggle OFF Breakers DO	impact	Manipulation of Control	⊞			✓	🗑



Start New Operation

Operation name: Break in an trip breakers

Adversary: Basic Attack

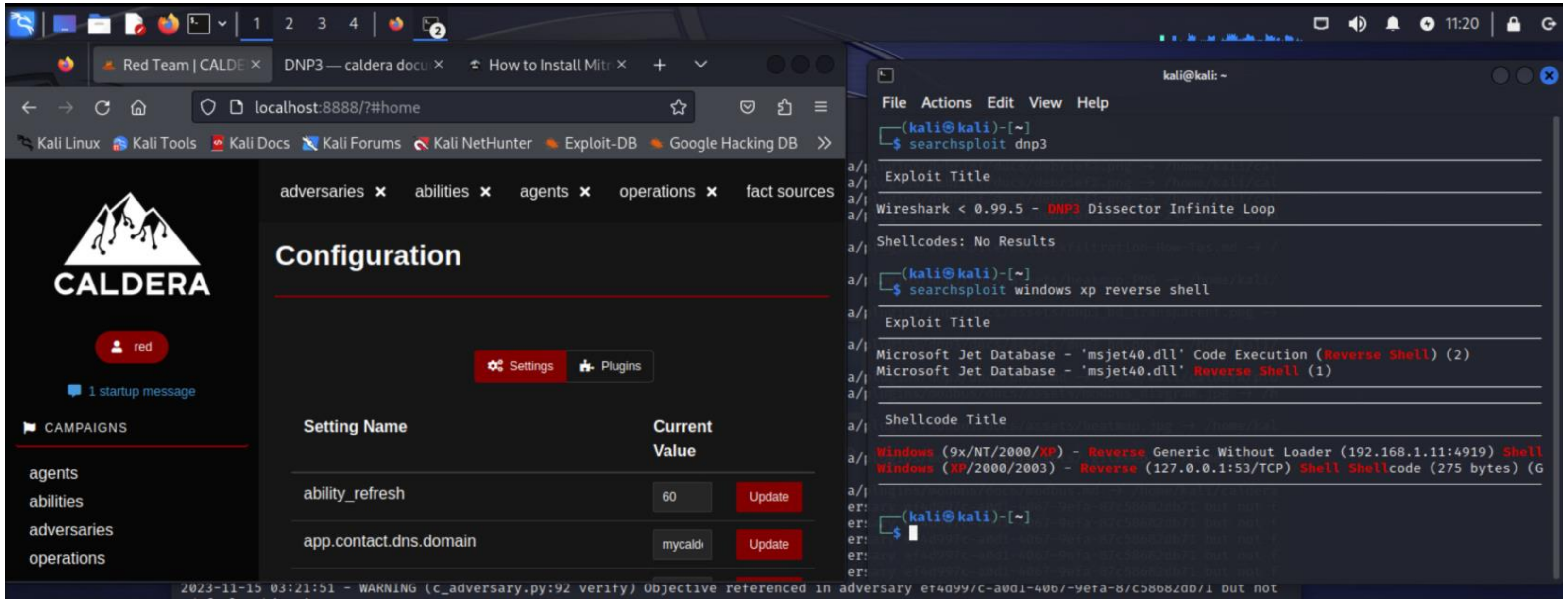
Fact source: DNP3 Sample Facts

ADVANCED

(our Caldera dashboard)

Kali Attack VM Implementation

- Mitre Caldera Implementation
  - Clone Repository
  - Install Pip and Go
    - Make sure to download pips requirements as well
  - Update path variable
    - PATH=$PATH:/usr/local/go/bin
  - Start server by running server.py
    - Server is hosted on the web
    - localhost:8888
  - Configuration
    - Set app.contact.http
- Metasploit and Searchsploit
  - Used in conjunction with Mitre Caldera
  - Reverse Shells
  - Any attacks we can't think of



Configuration

Setting Name	Current Value	
ability_refresh	60	Update
app.contact.dns.domain	mycald	Update

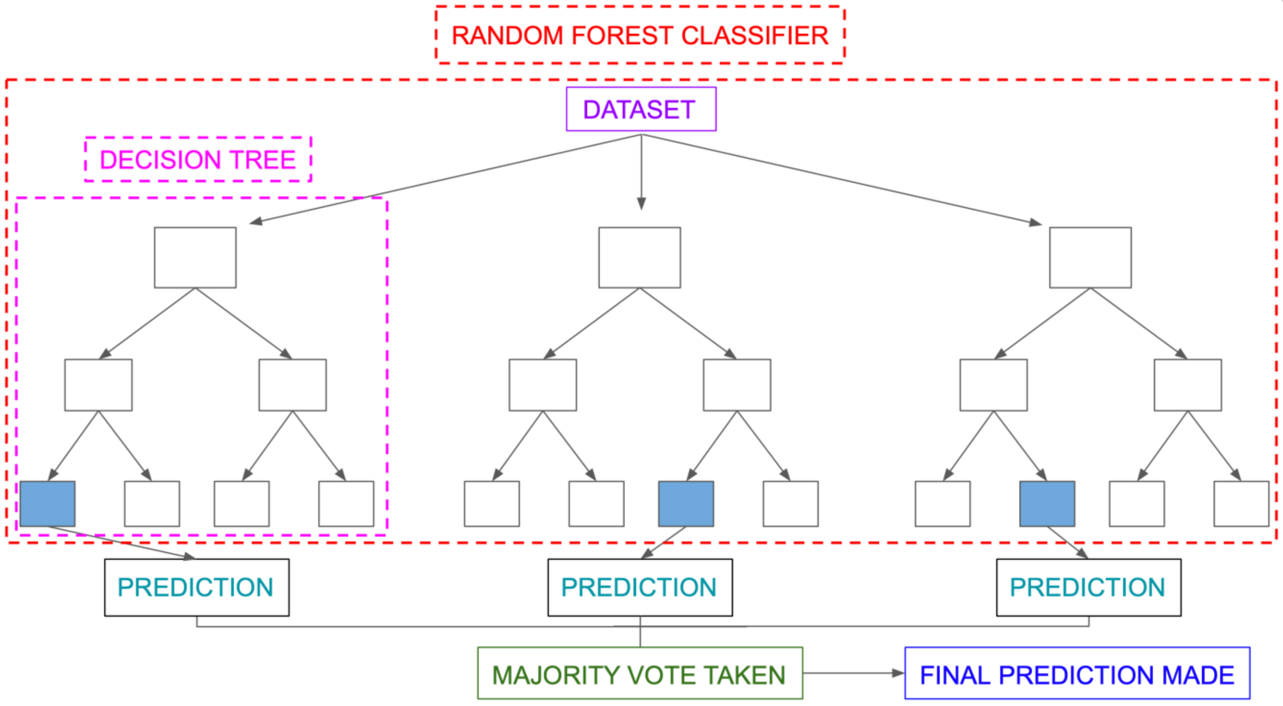
```
┌──(kali㉿kali)-[~]
└─$ searchsploit dnp3
────────────────────────────────────────────
 Exploit Title
────────────────────────────────────────────
Wireshark < 0.99.5 - DNP3 Dissector Infinite Loop
────────────────────────────────────────────
Shellcodes: No Results

┌──(kali㉿kali)-[~]
└─$ searchsploit windows xp reverse shell
────────────────────────────────────────────
 Exploit Title
────────────────────────────────────────────
Microsoft Jet Database - 'msjet40.dll' Code Execution (Reverse Shell) (2)
Microsoft Jet Database - 'msjet40.dll' Reverse Shell (1)
────────────────────────────────────────────
────────────────────────────────────────────
 Shellcode Title
────────────────────────────────────────────
Windows (9x/NT/2000/XP) - Reverse Generic Without Loader (192.168.1.11:4919) Shell
Windows (XP/2000/2003) - Reverse (127.0.0.1:53/TCP) Shell Shellcode (275 bytes) (G
────────────────────────────────────────────

┌──(kali㉿kali)-[~]
└─$ 
```

```
2023-11-15 03:21:51 - WARNING (c_adversary.py:92 verify) Objective referenced in adversary ef4d997c-a0d1-4067-9efa-87c58682db71 but not
```

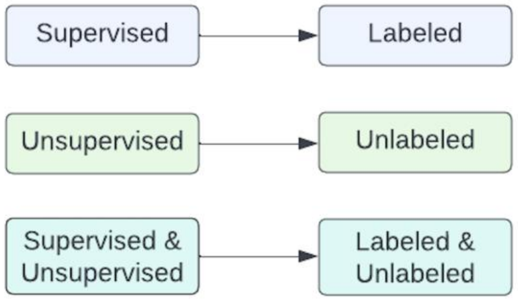
(Our Kali attack VM)

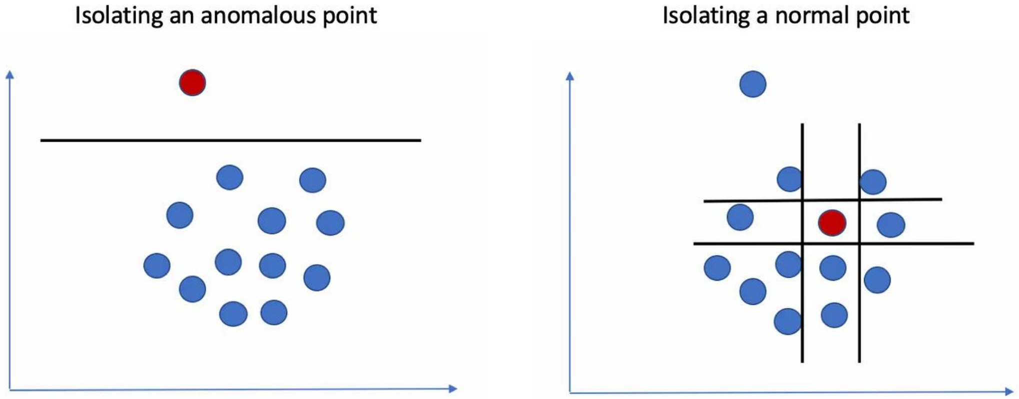
Machine Learning

- Scikit learn & Pandas
- Two-part approach
  - Supervised & unsupervised
- Binary Classification
  - Given malicious & normal labeled logs
  - Random forest to delineate
  - Random forest
    - Multiple decision trees
    - Majority vote
- Anomaly Detection
  - Isolation forest
    - Split data
    - Every data point isolated
    - Abnormal point less than a normal point



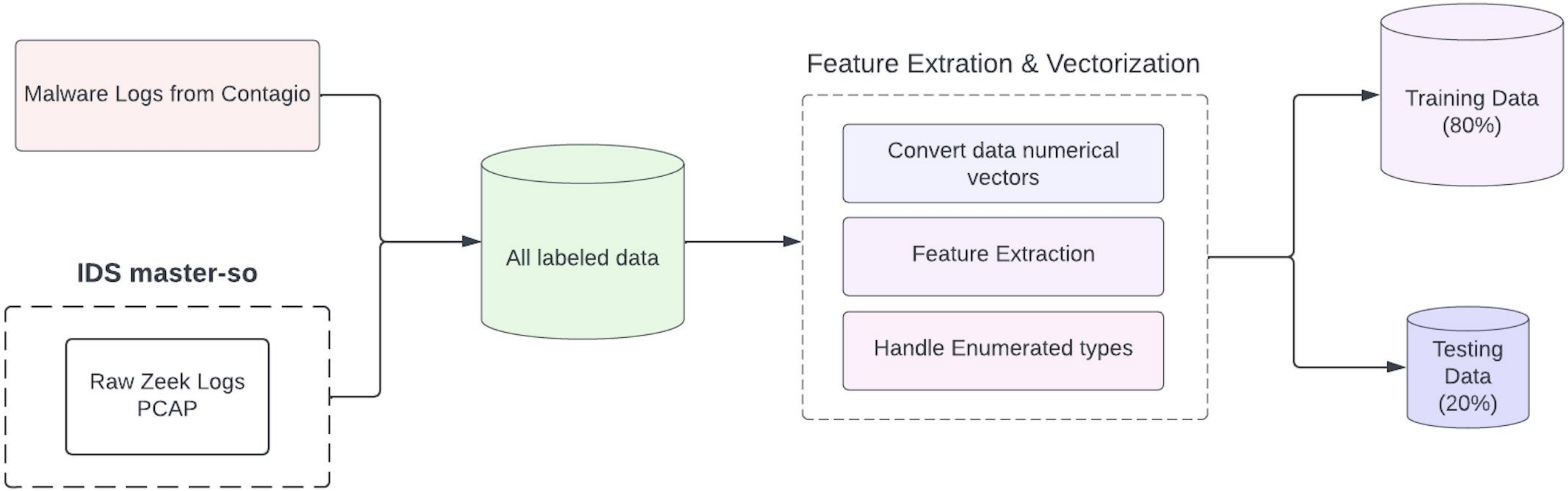
(Khushaktov)





(Mavuduru)

Machine Learning – Identifying Training & Test Data



Malware Logs from Contagio

IDS master-so

Raw Zeek Logs
PCAP

All labeled data

Feature Extration & Vectorization

Convert data numerical vectors

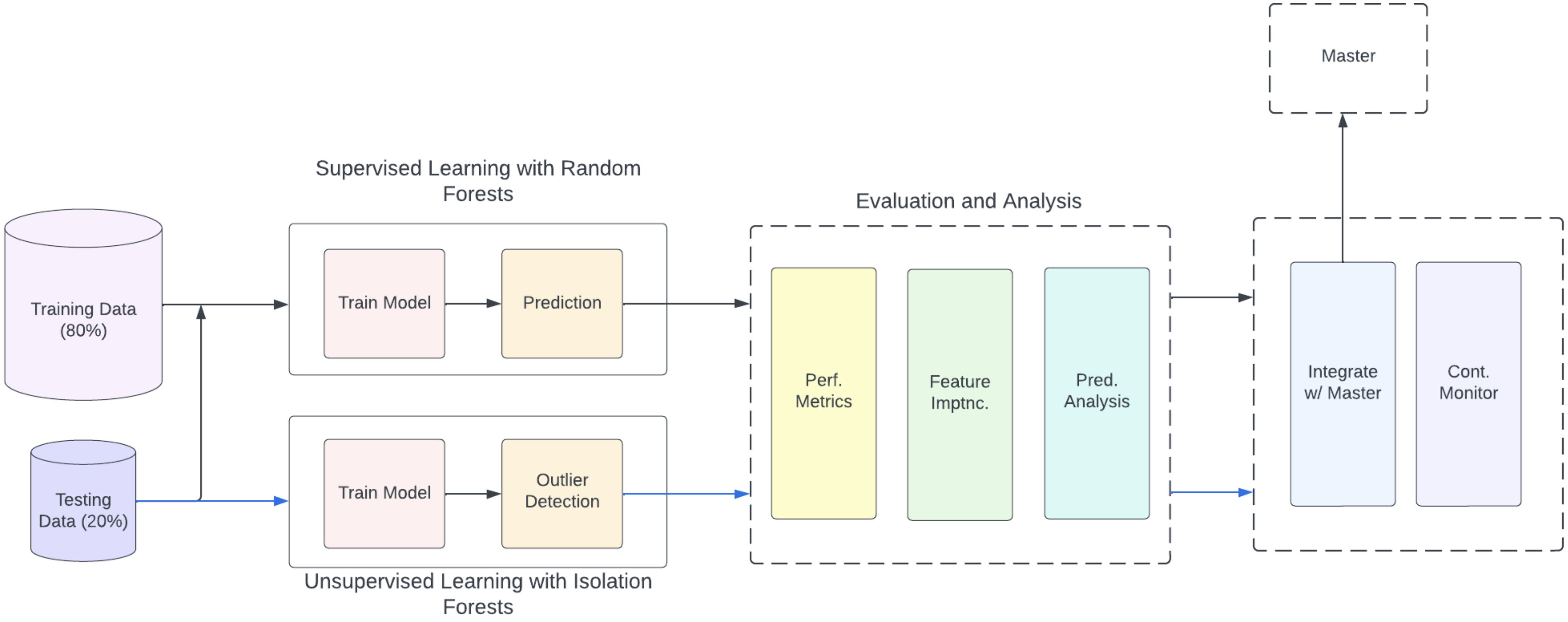
Feature Extraction

Handle Enumerated types

Training Data (80%)

Testing Data (20%)

Machine Learning – High Level Approach



Machine Learning - Output & Functionality

- Output display
  - Terminal
  - Only malicious logs

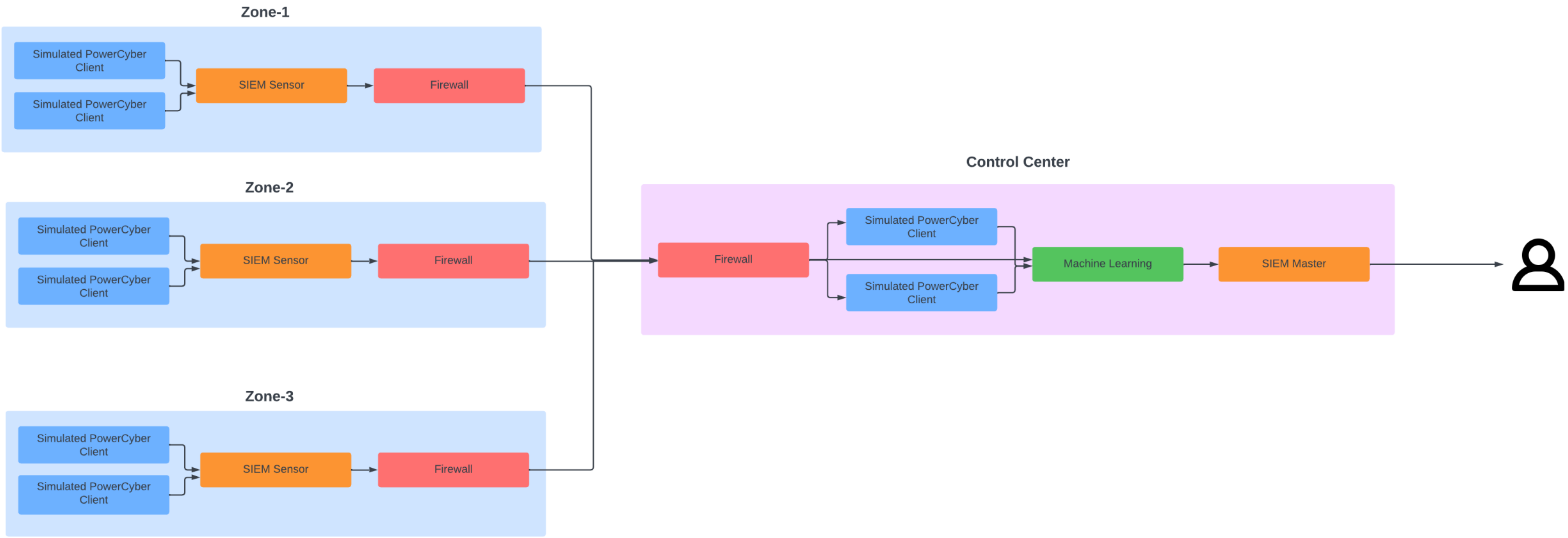
- For each log identified
  - List of features that contributed to identification
    - Accompanying percentage
      - Extent of contribution to decision
  - Final decision and percent probability of accuracy

- User interaction
  - Look through provided analysis as an aid

- Cron job
  - Email notification

- Based solely on analysis of logs pulled from Security Onion

Conceptual Final Design Diagram

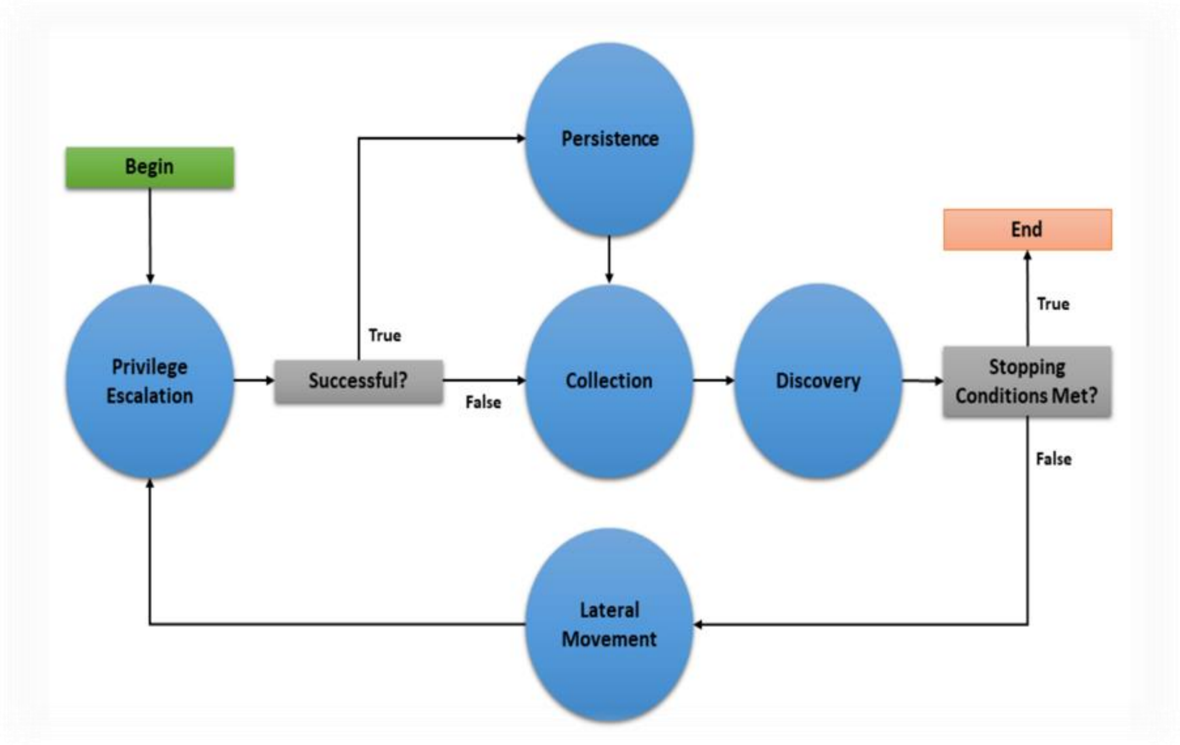

Zone-1
Simulated PowerCyber Client
Simulated PowerCyber Client
SIEM Sensor
Firewall
Zone-2
Simulated PowerCyber Client
Simulated PowerCyber Client
SIEM Sensor
Firewall
Zone-3
Simulated PowerCyber Client
Simulated PowerCyber Client
SIEM Sensor
Firewall
Control Center
Firewall
Simulated PowerCyber Client
Simulated PowerCyber Client
Machine Learning
SIEM Master

Design Complexity

Question	Response
What made the design difficult to implement?	• Understanding the complexity of the PowerCyber infrastructure, before integrating our own components. • Directing information from each of the zones/sensors into Security Onion for analysis. • Assessing how to properly train a ML model to act on our behalf and mitigate attacks. Supervised vs. Unsupervised. • Designing effective adversary emulation campaigns with Mitre Caldera to test the defense solutions put in place. • Working with older Windows operating system components within PowerCyber.
What kind of design iterations were needed?	• Researching SIEM frameworks to be used in conjunction or in place of Security Onion, like Gravwell and Splunk. • Integrating new protocols into the attack phase in addition to Modbus such as DNP3, bacnet. • Exploring an assortment of different vulnerabilities that affect OT systems. • Adjusting to defend from the attack approach used by different APTs.

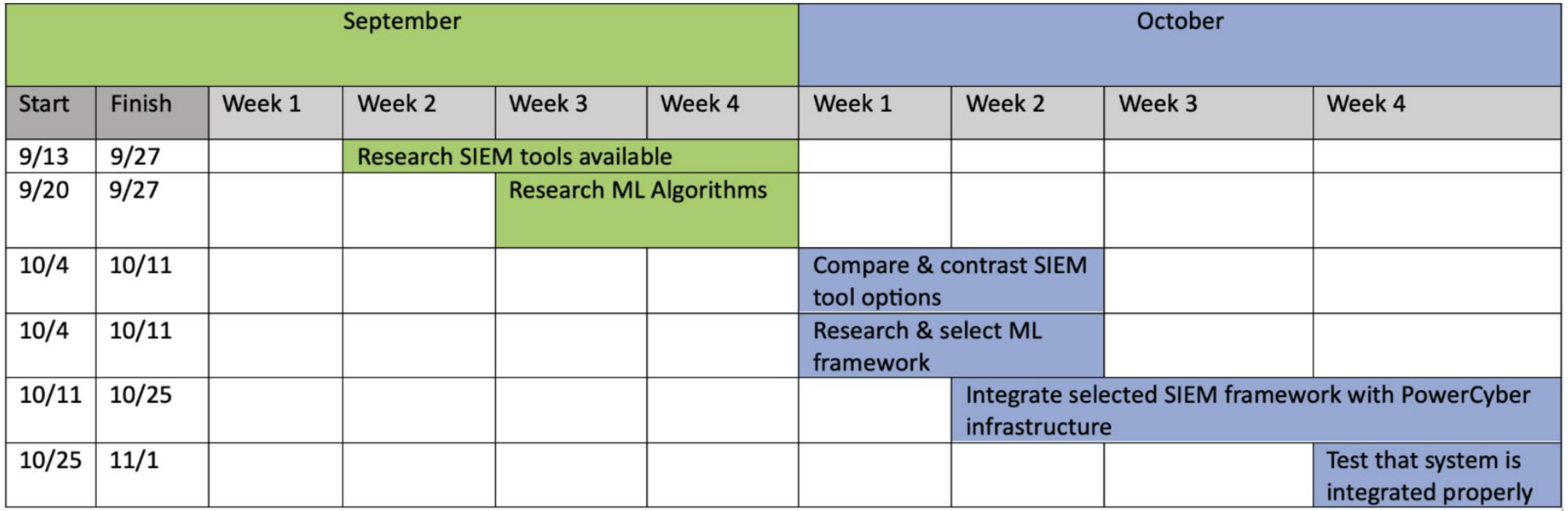
Test Plan

- Actual platform
- Includes
  - Interface/Integration testing
    - Security Onion and ML
    - Gravwell and Security Onion
  - System/Acceptance testing
    - Various attacks
    - Uptime, response, detection
    - Meet functional requirements
      - SIEM, ML, Attacks
  - Regression testing
    - Integration of ML
    - Snapshots
  - Security testing
    - run attacks, identify vulnerabilities

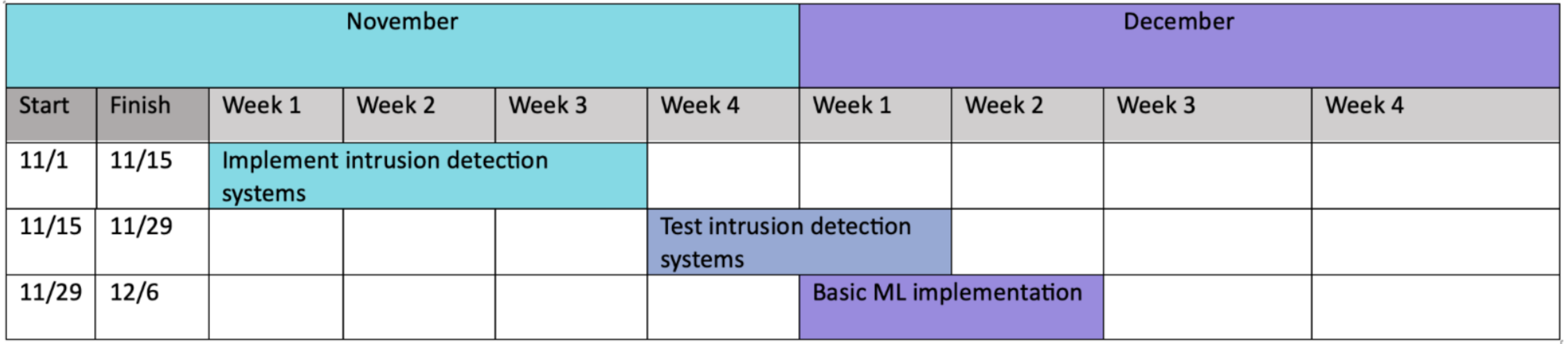


(caldera.readthedocs.io)

Project Plan - This Semester

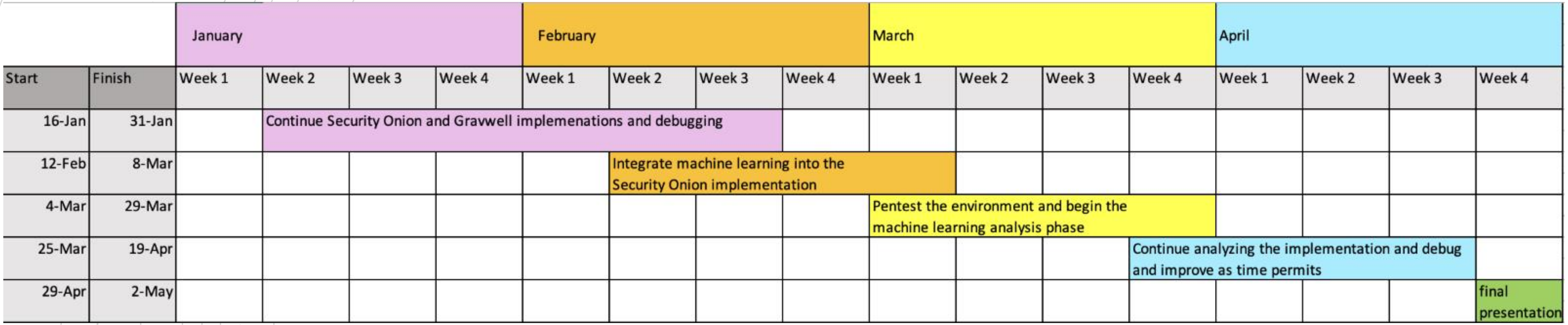


September						October			
Start	Finish	Week 1	Week 2	Week 3	Week 4	Week 1	Week 2	Week 3	Week 4
9/13	9/27		Research SIEM tools available						
9/20	9/27			Research ML Algorithms					
10/4	10/11					Compare & contrast SIEM tool options			
10/4	10/11					Research & select ML framework			
10/11	10/25						Integrate selected SIEM framework with PowerCyber infrastructure		
10/25	11/1								Test that system is integrated properly



November						December			
Start	Finish	Week 1	Week 2	Week 3	Week 4	Week 1	Week 2	Week 3	Week 4
11/1	11/15	Implement intrusion detection systems							
11/15	11/29				Test intrusion detection systems				
11/29	12/6					Basic ML implementation			

Project Plan - Next Semester



		January				February				March				April			
Start	Finish	Week 1	Week 2	Week 3	Week 4	Week 1	Week 2	Week 3	Week 4	Week 1	Week 2	Week 3	Week 4	Week 1	Week 2	Week 3	Week 4
16-Jan	31-Jan		Continue Security Onion and Gravwell implemenations and debugging														
12-Feb	8-Mar						Integrate machine learning into the Security Onion implementation										
4-Mar	29-Mar									Pentest the environment and begin the machine learning analysis phase							
25-Mar	19-Apr												Continue analyzing the implementation and debug and improve as time permits				
29-Apr	2-May																final presentation

Conclusion

- Currently we are in a stage that we have the SIEM and attack portion integrated, and the machine learning portion is being integrated
- Next semester steps include:
  - We will test the interconnectivity all the modules that we have made.
  - We will stage an attack from the Kali box and detecting it in our SIEM
  - We will verify our machine learning component by asking it to classify attacks

Member	Contributions
Trent Bickford	- Mitre Caldera Research/Setup - SIEM Research - Architecture Design
Daniel Ocampo	- Website, Gravwell setup - SIEM Research - Report writer.
Ella Cook	- Machine Learning Plan - SIEM Research
Westin Chamberlain	- Security Onion Setup - SIEM Research - Architecture Design

Q&A

Sources

G. Ravikumar, A. Singh, J. R. Babu, A. Moataz A and M. Govindarasu, "D-IDS for Cyber-Physical DER Modbus System - Architecture, Modeling, Testbed-based Evaluation," 2020 Resilience Week (RWS), Salt Lake City, UT, USA, 2020, pp. 153-159, doi: 10.1109/RWS50334.2020.9241259.

Khushaktov, Farkhod. *Random Forest Classifier*. 6 Aug. 2023. *Medium*, https://medium.com/@mrmaster907/introduction-random-forest-classification-by-example-6983d95c7b91.

M. Abdelkhalek, G. Ravikumar and M. Govindarasu, "ML-based Anomaly Detection System for DER Communication in Smart Grid," 2022 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT), New Orleans, LA, USA, 2022, pp. 1-5, doi: 10.1109/ISGT50606.2022.9817481.

Mavuduru, Amol. *Isolating an anomalous point versus a normal point*. 4 Nov. 2021. *Towards Data Science*, https://towardsdatascience.com/how-to-perform-anomaly-detection-with-the-isolation-forest-algorithm-e8c8372520bc.

S. N. Mohan, G. Ravikumar and M. Govindarasu, "Distributed Intrusion Detection System using Semantic-based Rules for SCADA in Smart Grid," 2020 IEEE/PES Transmission and Distribution Conference and Exposition (T&D), Chicago, IL, USA, 2020, pp. 1-5, doi: 10.1109/TD39804.2020.9299960.

G. Ravikumar, B. Hyder and M. Govindarasu, "Hardware-in-the-Loop CPS Security Architecture for DER Monitoring and Control Applications," 2020 IEEE Texas Power and Energy Conference (TPEC), College Station, TX, USA, 2020, pp. 1-5, doi: 10.1109/TPEC48276.2020.9042578.

“Architecture - Security Onion 2.4 Documentation.” Docs.securityonion.net, docs.securityonion.net/en/latest/architecture.html. Accessed 3 Dec. 2023

Community edition (no date) *Gravwell*. Available at: https://www.gravwell.io/community-edition (Accessed: 03 December 2023).