Threat Hunting Architecture

Introduction

Threat hunting is an essential skill for organizations with mature security operations centers. In this blog I will lay out an essential framework for the two different classifications of threat hunting as well as several threat hunting models that you should become familiar with.

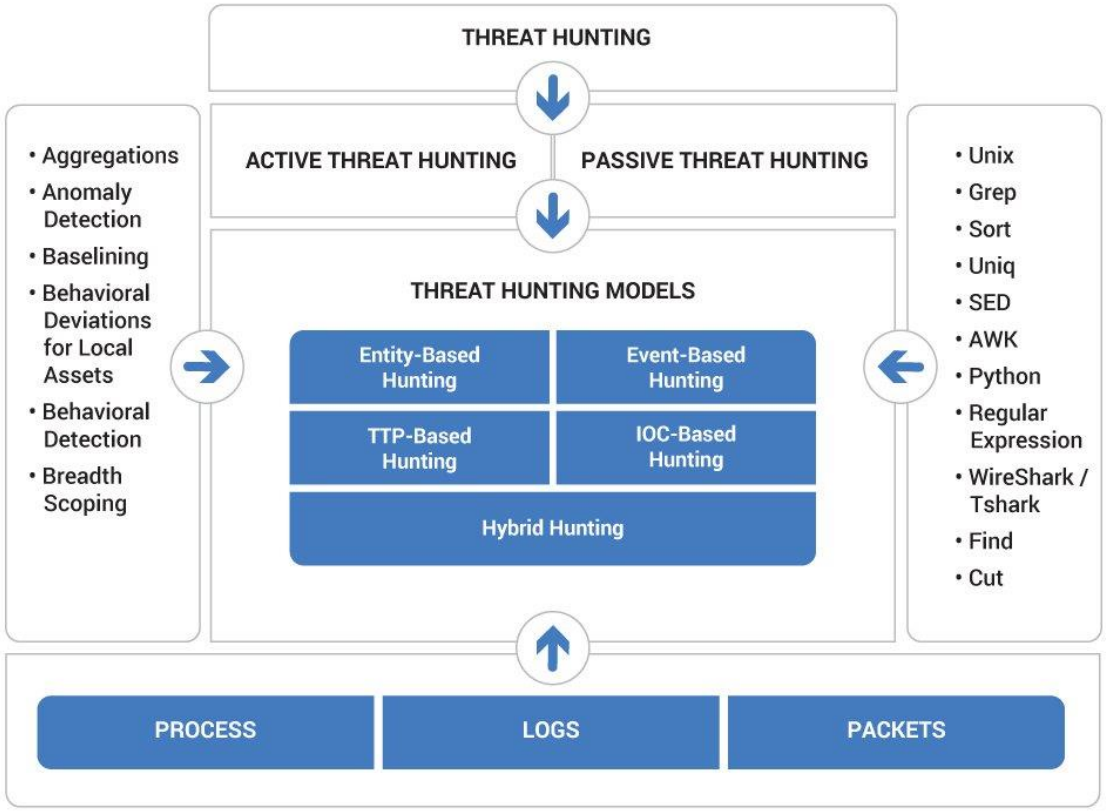

Figure 1: Diagram of Threat Hunting Architecture

Threat hunting is the step-by-step approach of proactively looking for signs of malicious activity within enterprise networks, without having initial knowledge of specific indications to look for, and subsequently ensuring that the malicious activity is removed from your systems and networks.

It's a technique that involves accumulating samples from different sources to assist in profiling malicious threat actors.

The Three Pillars of Threat Hunting

- Logs provide a record of events that have occurred in an enterprise. In them we can find a record of what actions were performed by an attacker in an enterprise.
- Packets are the data that flows through a network. In them we can find a record of what was communicated and how it was communicated in a network at a given point of time.
- Process provides us with the exact record of what effect an attacker had on an enterprise.

Classifying Threat Hunting

Passive Threat Hunting

Passive threat hunting is the process of analyzing logs/packets before they are normalized using an unstructured feed. The threat hunter needs to start with a very strong hypotheses, then use that hypothesis to extract specific indicators of compromise (IOC) from the feed. Then, IOCs are further verified using data from other threat intelligence feeds in order to confirm the accuracy of the extracted IOCs.

Threat hunters need to understand how passive threat hunting works because:

- A threat hunter may or may not have the tools available in order to hunt down particular threats using active hunting, so passive hunting may be the better option.

- If a threat hunter's hypothesis is very strong then sometimes they can find the threat very quickly using passive hunting. This reduces the time required to upload the data into an active hunting platform and normalize the data.

Limitations of Passive Threat Hunting

A threat hunter will not always be able to reach every corner of the organization, however, because of the limitations of passive hunting:

- Requires more effort to find IOCs
- Performing analysis across multiple log sources at once is very tough
- The hypothesis and assumptions regarding the threat have to be very strong to begin with
- Your ability to detect the process, or the exact record of what effect an attacker had on an enterprise, is limited

Active Threat Hunting

Active threat hunting is the process of analyzing logs, packets, and processes once the data has been normalized (i.e. converting all unstructured data to a structured format, and then making the data available for analysis). In active threat hunting it is easy to correlate different sets of log sources back to the IP address, user, and/or machine involved in order to identify what effect the threat had on the organization.

Active threat hunting allows threat hunters to hunt for complex scenarios of different sophisticated attacks such as:

- DNS Reconnaissance
- Domain Generation Algorithm
- Robotic Pattern Detection
- DNS Shadowing
- Fast Flux DNS
- Phishing

- Beaconing
- APTs
- Lateral Movement
- Browser Compromised
- DNS Amplification
- DNS Tunneling
- Skeleton Key Malware
- Low and Slow Attacks
- Composite Threat Detection
- DoS
- Intrusion Detection
- Cookie Visibility and Theft
- User Login Session Hijacking
- Broken Trust
- Session Fixation
- Honey Token Account Suspicious Activities
- Data Snooping / Data Aggregation
- Cross Channel Data Egress
- Banking Fraud
- Session Replay
- Watering Hole Attack

Favorite Hunting Techniques

Before we talk about threat hunting models, we need to understand hunting techniques. An essential technique is to first aggregate all feeds which will be required for hunting. Based on our input sources we can identify anomalies (i.e. an account is performing an activity which has never been seen before). In parallel, a threat hunter will create a baseline for each account address, asset, or even for specific sections of the enterprise. Once this information is available,

parallel analysis can be done on behavioral detection and breadth scoping (i.e. how far an attacker was able to penetrate).

- Aggregations
- Anomaly Detection
- Baselining
- Behavioral Deviations for Local Assets
- Behavioral Detection
- Breadth Scoping

Threat Hunting Models

Event-Based Hunting

All threat hunting models depend on how strong your hypotheses is. Every hypotheses should be based on event observations and a deep understanding of each log source: how it has been placed and how it is utilized in the enterprise.

DNS Tunnel

A threat hunter will primarily look into DNS logs for following scenarios:

- Volumetric Increase in DNS Requests
- Volumetric Increase in NXDOMAIN Requests
- Anomaly in DNS Response (Not an IP or NXDOMAIN)
- Rare Domain Request
- Check Length of Domain and Subdomain

Anomaly Detection in Proxy Logs

In the proxy logs a threat hunter can look for:

- Response Code (i.e. 200, 307, 403)

- Rare User Agent Used by Account
- Catalogue Behavior From Account (Account requesting multiple domains where PLD is constant but subdomains are changing.)

IOC-Based Hunting

An IOC is a bit of forensic data, primarily data found in system log transactions or files, that indicates potentially malicious activity on a system or on the network. It is like a fingerprint for a cyberthreat.

IOC-based hunting is one of the easiest ways to find a specific threat. The best way to describe IOC-based hunting is through the Pyramid of Pain.

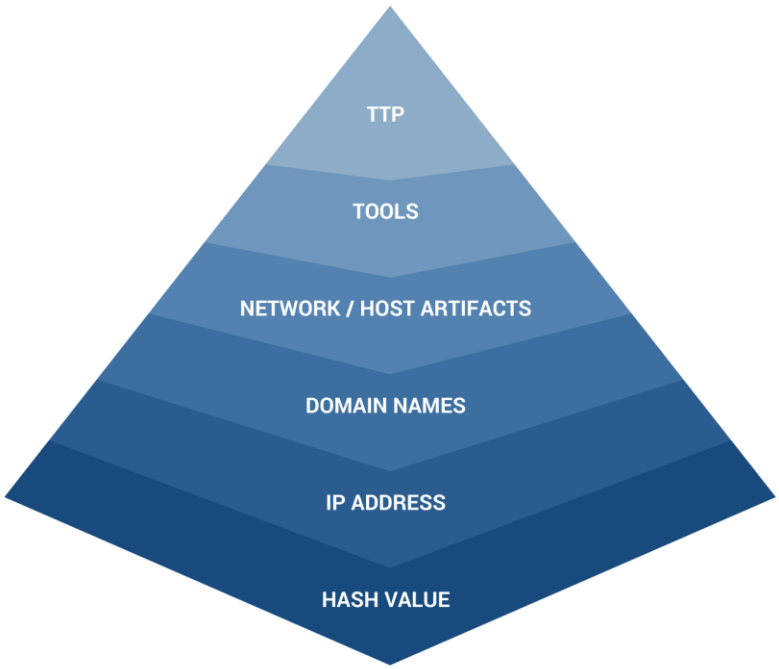

Figure 2: The Pyramid of Pain

As you identify an IOC, its location on the pyramid indicates how much pain that IOC will cause the attacker. The base of the pyramid, hash values, will cause a trivial amount of pain to an attacker. This includes hashes such as MD5, SHA1, and similar artifacts that identify specific suspicious or malicious files. IOCs are very useful to uniquely identify specific attack samples which have been detected in a security incident.

1. As we go from bottom to top it helps us to identify a specific threat.
2. If we have the IOCs for a specific threat then we can cause the attacker to change their methods.

Entity-Based Hunting

Understanding an organizations' network posture is a complex challenge. It does not matter how many resources you have, threat hunters always need to prioritize hunting network segments to get maximum throughput. Entity-based hunting is concentrated on high risk users (HRU) and high value assets (HVA).

- First, survey the network before you proceed
- Identify HRU and HVA
- Identify what applications are used by the business
- Gather risk assessment reports

Attackers will target HVA or HRU in order to access sensitive information or to act as base for lateral movement.
Threat hunters should prioritize investigating DMZ servers, application server LAN segments, and senior executives and their systems. Next, they should investigate domain controllers, container repositories, research and development systems, and centralized databases associated with each team.

Tactics, Techniques, and Procedures-Based Hunting

Threat hunters cannot always depend on IOCs. The useful lifespan of an IOC is relative to the associated lifespan of the related domains and IPs, which can be very short. If we really want to hamper an attacker's trajectory, we need to understand their tactics, techniques, and procedures (TTP). Being a threat hunter, we need to understand what technologies attackers use, how

effectively attackers are able to weaponize that technology against us, and what are the different ways attacker can adapt and change. We also need to understand what their primary goals are and how they would move laterally in an organization.

If a threat hunter can develop a 360-degree view of the attack, including artifacts, effects, measures, and propagations, then they can create a playbook for hunting.

Hybrid Threat Hunting

The hybrid threat hunting model uses a combination of multiple threat hunting models. This means that you need to be a subject matter expert in as many threat hunting models as possible. Hybrid hunting happens when you don't know how deep the attack has penetrated and how far it has spread laterally. In this model all above mentioned model can play individual role in each stage of the investigation.

One of the best examples of hybrid threat hunting is cross channel data egress. In this example a threat hunter is trying to find two independent sources; one from which source data was aggregated, and a second one from which source data was exfiltrated. In this scenario, a threat hunter can use a combination of event-based hunting, entity-based hunting, and IOC-based hunting.

Conclusion

Threat hunting involves several tradeoffs depending on what organizational data you have access to and what you know about possible attackers. Depending on your answer to each of these questions will inform which style or model of threat hunting will work best for your situation.

By Rohit D. Sadgune