corelight

White Paper

Why evidence-first analytics are the foundation of modern NDR

Today's security teams use powerful analytics to accelerate incident response across the SOC. It's increasingly vital to augment SIEM and EDR tools with analytics from Network Detection and Response (NDR) solutions to obtain the breadth of insights and detections necessary for effective defenses. NDR analytics fuel SOC activity across threat detection coverage, asset visibility, and threat hunting.

Corelight has built a suite of analytics rooted in decades of experience with world-class defenders who have contributed their practices and design patterns to the open source community. A key learning from these defenders is that evidence quality determines analytic outcomes and so evidence must come *first*. An evidence-first analytics approach delivers broad, deep and accurate insights based on three pillars:

1. **The best evidence sets the strongest foundation.** The best evidence enables the best analytics, accelerates alert investigation, and allows defenders to investigate attacks spanning today, yesterday and tomorrow using retrospective analysis, forensics, and threat hunting.

2. **Analytics need the right tool for the job - there is no silver bullet.** Machine learning, queries, behavioral detection, threat intelligence and traditional IDS signatures are each useful for different attack activity. We leverage them together for the most accurate analytics, broadest toolset consolidation and most effective alert aggregation.

3. **Threat hunting is core to modern detection.** Threat hunting requires unfettered access to evidence and can drive new detections and broader analytics coverage. In addition, hunting also reveals operational issues and accelerates routine incident response by understanding what “normal” is in the environment.

Let’s use two case studies to explore these themes and demonstrate how these three pillars are critical to leveraging the power of NDR. We will start with the detection and investigation of the Sliver command-and-control (C2) framework[1] in an environment. Afterwards, we’ll look at an example of a threat hunt investigating building automation traffic.

[1] https://github.com/BishopFox/sliver

Initial discovery

As a C2 framework, we can detect Sliver via behavioral analytics (finding the tool itself) or by detecting the activity of malware delivered using the framework. Corelight has built a range of detection techniques across attacker infrastructure (IOCs), toolkits (Sliver, Cobalt Strike, Manjusaka, etc), and techniques (DGA malware, DNS exfiltration, etc). We use a wide range of detection engineering techniques to increase both investigative efficiency and alert accuracy.

Regarding investigative efficiency, consider that in this example a separately deployed signature IDS, threat intelligence platform and SIEM machine learning toolkit would see different parts of the attacks but none of the tools would properly aggregate the resulting alerts for the analyst to prioritize the problem and drive an effective investigation. In addition, Corelight's integration of multiple NDR tools in a single platform removes the need for the security engineering team to maintain and tune each of those disparate tools, freeing up time for other analytics and automation initiatives.

Regarding alert accuracy, consider that different detection methods have a different balance of false positive vs. detection (or false negative) rates. While no one detection approach is well suited for all the known variations within a given attack scenario, by using the strengths of different techniques we can cover the broadest range possible with the highest accuracy:

Detection Scenario		IOCs	Signatures	Falcon LogScale rules	Behavioral Detection[1]	Supervised Machine Learning	Anomaly detection[2]
Known	Attacker infrastructure						
	Vulnerabilities						
	Attacker tools						
	Techniques						
New	Vulnerabilities						
	Attacker tools						
	Techniques						

Green color coding shows that the technique is useful for a meaningful range of attack types at a manageable FP rate. 1. Behavioral detection identifies network behavior patterns, for example using Zeek's scripting framework. 2. Anomaly detection is most commonly done with unsupervised machine learning.

For the Sliver C2 framework example we see that behavioral detection is the right tool for the job because the tool's activities are clear and identifiable. In this instance, signatures aren't the right fit because detecting the behaviors often requires analysis across multiple network connections. Likewise, we don't need to incur the FP rates of machine learning (ML) for a tool we can find well with behavioral detection. For all of its progress and earned popularity, ML is a probabilistic detection method. That means it will carry a higher FP rate than other techniques (despite tuning) so as a general rule we should use ML to tackle detections which can't be readily identified using simpler methods.

If we look downstream of the Sliver framework however, other techniques come into play. Attackers often re-use techniques for part of their campaigns so leveraging signatures and IOCs offers an easy and reliable way to find that subsequent activity. Likewise, supervised ML is highly effective for finding attackers hiding their exfiltration via DNS or detecting Tor usage. These situations are hard to detect well with signatures, rules or behavioral detection, but are a good fit for the probabilistic methods of ML.

Ultimately, using the right detection tool for the right job allows us to both provide the best alert accuracy and also the most effective alert aggregation and technology consolidation for defenders.

Investigation and Confirmation

For security analysts, detecting the Sliver framework is just the beginning. From there, analysts must:

1. **Verify the alert:** analysts can see additional confirming activity from the Sliver toolkit such as beaconing, telltale user agents or HTTP header ordering. The confirming evidence varies by attack of course, but analysts needing access to the right evidence is a constant.

2. **Investigate the scope of the attack:** as a C2 framework Sliver should carry out both upstream and downstream activity from the infected host. Analysts can follow the story laid out by network protocol logs to find the point of initial compromise as well as the lateral movement and any exfiltration attempts downstream of this specific detection. For example, connecting a DNS reply to a related HTTP session and subsequent file transfer can quickly take the analyst from an indicator to identifying exfiltration.

3. **Confirm both the extent of any exfiltration and ensure remediation**. Here network evidence can prove that either exfiltration didn't occur, or if it did then it can reveal its true scope. The difference between "we think" and "we know" becomes incredibly important here as businesses face significant fiscal and policy implications from breaches. Afterwards, that same network evidence can verify the attacker is truly removed from the environment through ongoing validation of successful containment and remediation.

Each of these three activities demands richly detailed, interconnected network evidence. Corelight's evidence stack is based on technologies derived from open source design patterns (observed across top SOCs spanning industries and continents) and new analytics from the Corelight Labs team:

Community Design Pattern	
Zeek logs	Interconnected protocol logs (for case investigation, downstream analysis and threat hunting)
Extracted files	Extracted and deduplicated files (for static or dynamic analysis, sandboxing or Virus Total lookups)
Policy-driven PCAP	Policy based filters to capture the small fraction of PCAP analysts need (for longer retention times or cost savings vs. traditional full PCAP)

New Corelight Labs Content	
Encrypted traffic analysis	Find both behaviors and attacks (e.g. custom encryption, auth bypass for SSH/RDP, cleartext inside of encrypted streams, and more)
Derived activity insights	Identify activity from protocols and applications (ex: VPN providers, encrypted DNS lookups, SSH/RDP file transfers)
Environmental data	Organization-specific information (ex: network / asset inventories, CMDB info, vulnerability scan results)

This evidence stack was developed by these defenders to accelerate investigations and enable the best network-centric analytics. Back to our Sliver C2 framework example, there are three key attributes of the evidence stack that analysts need to drive this investigation successfully:

1. **Rich depth:** unlike Netflow, analysts need rich insight into protocols and applications to follow the attacker movement through the environment. Corelight starts with Zeek: richly detailed and interlinked protocol logs built on decades of evolution, it gives a wealth of information distilled to provide what the analyst needs. Then we add local environment insight, because simple things (like knowing a Linux attack is targeting a Windows host) can accelerate routine investigations.

2. **Retention time:** while the detection of Sliver was quick, the incursion may have begun weeks or months prior. As a result the evidence stack must be concise and cost effective to retain for quarters or years, not months or days. There are many examples of this need, but the most well known is the Sunburst attack, where the hard question was not "am I infected" but instead "what happened three months ago?"

3. **Readily searchable:** the evidence is only as good as the analysts' ability to use it! Corelight addresses this by using Zeek (which links activity across logs) and CrowdStrike Falcon LogScale (for rapid, easily chainable search) as the analyst traverses the attack chain around the initial Sliver detection. Both benefit from a broadly available set of Zeek training for analysts, provided by the open source Zeek project and training organizations like SANS (as well as Corelight), and through Crowdstrike.

The right evidence stack not only enables analysts to process alerts quickly, but is critical for investigating the complete attack chain, confirming effective remediation, and revealing scope of impact.

Threat Hunting

Our second case study concerns threat hunting, which - in contrast to alert-driven workflows - works to find attacks that have evaded detection.

For example, starting with the question "what is in my environment?" a threat hunter looks not just at the assets on the network but the protocols as well. As a result, our hunter found BACnet (a building automation protocol) in use. This wasn't surprising, however looking at the activity of the various devices showed one was connecting to multiple external IP addresses and also exhibiting unusual connection behavior (communicating to Brazil when this organization had no assets or technology there) and larger data transfer volume than the other devices. Further investigation showed that the device had in fact been compromised.

Long practiced by seasoned defenders, threat hunting is now increasingly common in the enterprise as well. There are three main approaches to a modern threat hunting practice:

Technique	Definition
Hypothesis driven	Starts from an initial hypothesis on attacker TTPs and then delves into the record of enterprise activity that it seeds ("what are the least common HTTP User Agents, and what requests were made using them?")
Entity driven	Starts from a site-specific risk assessment (high value assets such as domain controllers or development managers) or behavior patterns ("has anyone from the Engineering subnet tried to connect to the Finance subnet?")
Retrospective detection	As new forms of attacks, attacker infrastructure, and tooling become known, threat hunters look back in time for previous instances of these newly discovered attacks ("are we exposed to the same attack that hit Uber?")

While the need for threat hunting is clear and even a modest investment can produce substantial results, it can remain daunting for many organizations. Corelight helps enable hunting in three key ways:

1. **Starting hypothesis:** Corelight provides a battery of over 70 hunting queries, each of which provides the start of a hypothesis-driven hunt. We developed our threat hunting guide from the practices of veteran defenders. Encoding their knowledge helps analysts effectively explore their networks.

2. **Rich evidence:** effective retrospective detection requires the right evidence. Building on an evidence stack that continues to evolve in partnership with defenders in the open source community ensures that hunters have comprehensive evidence to assess newly discovered threats.

3. **Rapid search:** we leverage the speed of Crowdstrike's Falcon LogScale platform for fast search and pivots through our evidence. Search speed is critical because if technology can't keep pace with the threat hunter's mind then not only is the analysis hampered but we risk losing the thread of the investigation itself.

Threat hunting is an integral part of a modern detection strategy because no defenders have perfect detection coverage. As a result, threat hunting finds new attacks and fuels creation of new detections to find those attacks in the future. In addition, threat hunting educates the security team about their own environment. This enables defenders to investigate subsequent alerts more quickly and helps them know their own network as well as attackers will from their reconnaissance.

Conclusion

We continue to expand our analytics coverage based on partnerships with our customers, fellow industry researchers (ex: Microsoft's MAPP program), and insight from the open source community. Regardless of the detection source, as we saw with both the lifecycle of a Sliver investigation and a BACnet threat hunt, Corelight's evidence-first analytics approach provides the most effective detection, accelerates the entire incident response cycle and enables threat hunting as well. The result is faster overall time to detection and response compared to "alert-first" approaches. The three pillars we have built Corelight's technology upon come from lessons learned by working with the expert defenders of the open source community. We take these roots to heart, and are proud to bring to enterprises the best evidence and analytics we can offer.

For more information, visit www.corelight.com.

Corelight transforms network and cloud activity into evidence and analytics to hunt for threats, accelerate incident response, gain complete network visibility and create powerful analytics. Corelight's global customers include Fortune 500 companies, major government agencies, and large universities.

info@corelight.com | 888-547-9497