Advantage, Network Metadata

How to Enhance Efficiency in Incident Investigations Without PCAP

Security Challenges

Cybersecurity teams are busy. Really busy. Moreover, they have difficult tasks: securing a network while simultaneously judging how best to do so. The most critical of the many challenges confronting these teams and their ability to respond to an attack: quick and accurate data collection regarding the network. This is a critical and integral part of security investigations that look to validate whether potential criminal activities are (or were) present in the environment. Teams need quick and simple answers to the questions: who has been in the network, what was touched, what was the extent of the harm and how to we take precautions for future attacks.

To Find the Needle Faster, Shrink the Haystack

To investigate and validate a security incident, many teams know they need only a fraction of captured data. In fact, by extracting the traffic's metadata you can, in effect, shrink the size of the dataset that needs processing. To paraphrase an old security saying, "It's easier to find the needle in the haystack when the haystack is way smaller."

Analyzing Your Network's Security

Since the early days of networking, capturing direct copies of snippets of data from the flow of network traffic has been the go-to approach to collect data for incident investigation. These snippets, direct packet captures (PCAPs), are then taken offline and analyzed. Offline analysis ensures continuity of the main data flows and also provides analysts with the time needed for in-depth investigation, required to identify malicious behavior. PCAPs also provide a direct copy of the data, letting investigators know exactly what was compromised or stolen and how was it done. The advent of newer hardware and technologies helps automate the capture process, with larger data sets stored and indexed, to provide a more complete view into possible threats.

Key Points

- Cybersecurity teams need to investigate and validate data pertaining to security incidents
- Though complete, packet capture (PCAP) is inefficient and resource intensive
- By reducing the dataset and extracting only important information, metadata increases efficiencies in the response/investigation/validation cycle

There are, however, significant disadvantages to the PCAP approach. The full analysis of PCAPs, besides being manual and labor intensive (as well as difficult to store, manage and index for quick searches), requires that an investigator know the 'what, where and when' of what is meaningful in the data. Teams often filter large quantities of data to extract the same succinct nuggets of information that describe the data, while not necessarily needing the data itself (payload), for their investigative efforts.

The inefficiencies of full packet capture have been well known for some time. As such, the extraction of the needed meaningful nuggets (metadata) from the main data stream was deemed important early on, though not initially for network security purposes. By definition, metadata capture is the capture of the result of an analysis of the data seen on the wire, not of the data itself. As an example, the capture of key information, such as source and destination IP, protocol, source and destination TCP/UDP ports, service, time/date stamps and data quantity, is known to most network professionals, as it was popularized by NetFlow, a network protocol developed by Cisco. NetFlow was devised and used to identify problems in the network proper. Today, a variety of protocols (Argus, App Flow, JStream, JFLOW, as examples) captures similar metadata, which is readily available from most networking devices such as switches, packet brokers and routers.

Using Metadata for Incident Investigation

Security teams have used PCAP analysis as a primary means of investigation for some time. As such, they may have reservations about using metadata use because they're concerned that they do not have a full dataset for investigations. Or, they may view the use of PCAP versus metadata as a choice to be made, between one approach versus the other. Let's examine both concerns in more detail.

You can extra many forms and types of metadata information for analysis — and, good news, the extracted information can be purpose built and be summarized using a fraction of the file space needed for full data capture. Today, metadata capture is used more and more to help security incident investigations. It offers quick access to information relevant to a threat's point of entry, movement and affected ports and protocols, which helps teams determine what has happened and what systems were affected.

Metadata extraction is not constrained to specific information. Figure 1 illustrates the process of extraction, and, as shown, you can extract precisely the information needed for investigation. Data can then be enriched by comparing the extracted data against a third source. As an example, you can enrich HTTP/S data in the form of domain information by adding relevant domain information contained in WHOIS (such as when the domain was registered or taken down).

Figure 2 depicts the enrichment of extracted data to provide further details about specific protocols seen during an attack. Even with enriched metadata, this subset of data provides a wealth of information within a smaller storage footprint than PCAP. As seen in Figure 2, using rough calculations, this can be on the order of 2,000:1.

In real terms (using NetFlow as an example): "...the capture of hours of PCAPs would utilize the same amount of storage space as MONTHS of NetFlow data capture."1 The result? Security experts can parse through more devices, more communications and more security-specific data over a longer period of time than when recording and storing PCAPs. This allows for more in-depth analysis and visibility into different types of attack patterns (for example, low-and-slow attacks) and different tendencies exhibited by threat actor groups. In addition, with smaller space requirements, teams can afford to selectively packet capture assets or information based on need and maximize resource use. As you can see, it's not a matter of choice between the two approaches to investigation, but where to best use each approach and for what purposes.

Data Packet and Metadata Extraction

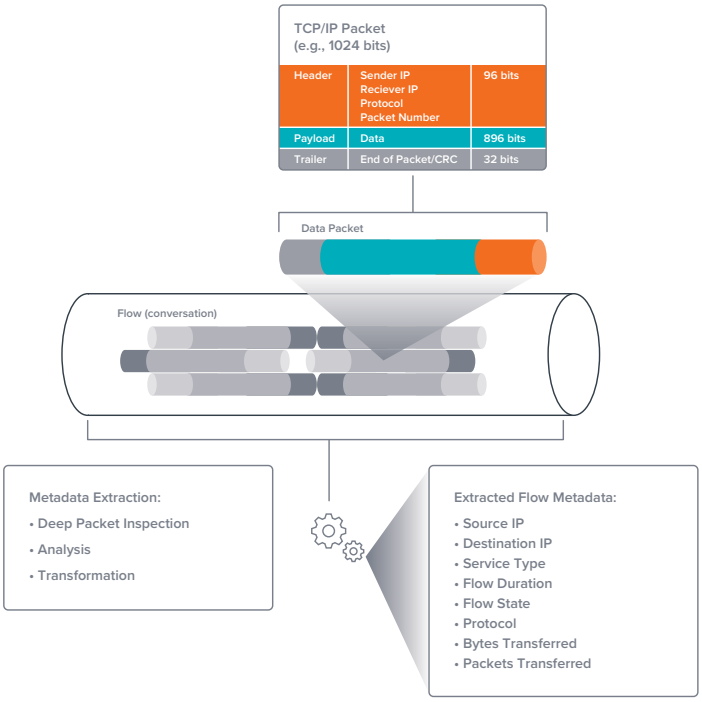
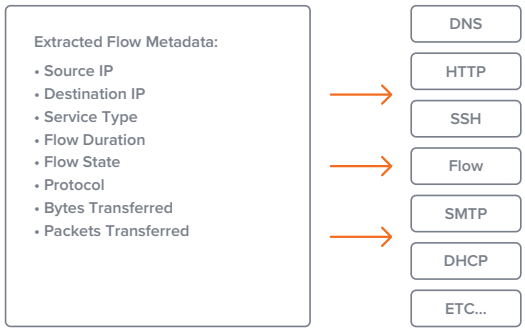


Figure 1: Metadata extraction

Metadata: Protocol Metadata Correlation



Storage Requirements: A comparison of data storage size needs

Figure 2: Metadata correlation and size comparison

When to Choose PCAP or Metadata?

When considering when to use metadata capture versus full PCAP, teams often fallback to familiar habits, which doesn't necessarily result in the quickest resolution. Metadata is useful to rapidly determine that something has happened and what was affected, which is ideal for containment and immediate response. PCAPs are generally required only when 1) a protocol was observed that you don't have a parser for or 2) you need to perform a full byte-level forensics analysis of the data flow that you otherwise can't get from a host and network live response. To achieve maximum efficiency in security detection, investigation and response, consider the following advantages:

Scenario	Description	Metadata	PCAP
Alert triage	Alerts provided by detection systems (firewalls, IDS/IPS, gateways, files analysis) that provide little context into related events and often require extensive review to determine what is real within the data.	**Advantage: Metadata analysis** Quickly correlate related events and parse over longer periods of time, providing the context needed for triage.	Determining what alerts are worthy of further scrutiny takes too long when parsing PCAP datasets. PCAPs must also be analyzed to extract important information to make determinations, which is time consuming.
Attack containment and device isolation	Quick determination of what devices and delivery mechanisms and/or protocols or communications were involved in the attack, as well as provide recommended next steps for threat containment and device isolation.	**Advantage: Metadata analysis** Quickly query large datasets to isolate which devices were involved in potential communication.	Determining what devices were involved in an attack requires lengthy queries across large datasets that take time away from quick response.
Historical tracking	Tracking of device activities over periods of times (weeks, months) that are more extensive than most preventative solutions to further quantify whether observed activity is an isolated event or part of a potentially larger campaign.	**Advantage: Metadata analysis** Quickly rewind to specific points in time and query large datasets to isolate patterns of behavior indicative of compromise.	Storage of the information needed for historical tracking is too costly when using PCAP.
Non-parsed protocol analysis	A device's network traffic uses protocols that are not specifically captured via metadata parsers, or are non-standard and require additional scrutiny.	Since Information regarding specific protocols is not captured by the available metadata parsers, visibility into potential threats is lost.	**Advantage: PCAP** In these scenarios, information regarding specific protocols can be directly captured via PCAP and full visibility established into the full data flow.
Unique-event recreation	This involves the full recreation of a network event and data exchanged during the event. For example, a full byte-level forensics analysis is required to determine if/when specific files have been deleted from a device.	Metadata capture, by definition, does not include the actual information exchanged in the communication, so full recreation of the event is impossible.	**Advantage: PCAP** In order to perform the full recreation of a network event (such as when files have been deleted), the event must be recreated using the full capture of the network communication and data exchanged, only available within a full PCAP.

In addition to the aforementioned scenarios, security professionals should also leverage their knowledge of the stages of a cyberattack . As an attack unfolds, security analysts need visibility into an attack at various points, including details regarding what occured prior to the exfiltration of data from an organization's crown jewels, such as database servers, financial data, patient data. Thus, the ability to provide more coverage and visibility via metadata capture provides more advanced warning, allowing teams to deploy the more costly PCAP technologies to surround only their most important assets.

A team's skillset is an often overlooked yet important aspect of investigation when considering security technologies. The security industry has fallen behind in the number of people who possess the skillsets required to rapidly assess what's important within ever increasing quantities of alerts. As such, teams must also realistically assess where their teams stand insofar as overall experience, and deploy detect-and-respond techniques commensurate with their present capacity.

To help match skillsets to techniques, the following table provides a comparison of skillsets required to manage some security technologies and methodologies used by today's teams.

Item	Skillset complexity	Reason
Protect and prevent solutions	Low to moderate	Technologies that fall within this space may provide an in-depth analysis of the specific attack artifact they are designed to detect. However, more often than not they provide little information as to why something was blocked. These are often point detections and do little for cross correlation of threat actor-related datasets.
PCAP	High	Security professionals must know what to look for, what patterns to discern and what anomalies look like within the larger PCAP dataset. This highly complex task requires a high-level skillset that is increasingly rare within the industry, and organizations often struggle to find, hire and retain individuals with these talents.
Metadata capture	Low to moderate	Metadata is ideal for first-level analysis and, in many instances, can be set to focus on specific data points that can easily be extracted. It requires less heavy lifting in the analysis to discern what is important (such as domain or DNS information). As such, the skillset required to derive important information from the data is lower than with PCAP analysis.

Recommendation: Metadata Capture as a Security Must-Have

We here at Gigamon recommend that teams look to implement metadata capture and storage as a must-have practice needed to secure a network. The security advantages are many and the cost savings substantial. PCAP acquisition and storage can be relegated to primary targets and should be deployed to selectively live capture information when required.

About Gigamon

Gigamon is the first company to deliver unified network visibility and analytics on all data-in-transit, from raw packets to apps, across physical, virtual and cloud infrastructure. We aggregate, transform and analyze network traffic to solve for critical performance and security needs, including rapid threat detection and response, freeing your organization to drive digital innovation. In short, we enable you to run fast, stay secure and innovate. Gigamon has been awarded over 75 technology patents and enjoys industry-leading customer satisfaction with more than 3,000 organizations, including 80 percent of the Fortune 100. Headquartered in Silicon Valley, Gigamon operates globally. For the full story on how Gigamon can help you, please visit www.gigamon.com.

Accelerate incident investigations using metadata. Learn more at gigamon.com/threatinsight

References:
1. Reid, Gavin. "NetFlow AND PCAP (Not or)." Cisco Blogs, July 8, 2016. https://blogs.cisco.com/security/netflow-and-pcap-not-or.