Fidelis™ Cybersecurity

Talk Metadata to Me

How to Decode Your Network's Deepest and Darkest Secrets

www.fidelissecurity.com

Contents

Introduction

When you detect an attack you want to find out what's going on — fast! But to investigate it you need historical data. Logs and netflow data can provide a start, but they are not enough to trace threats back to their source. Full packet capture (PCAP) systems advertise themselves as a digital video recorder (DVR) for the network. But storage fees can run into the millions — and running analytics against them is clunky at best. They are poorly indexed and contain too much data.

What's the answer? *Metadata*.

Whilst it doesn't sound sexy, rich metadata gathered from your network can capture more than ninety percent of the useful data that a full packet capture system would at twenty percent of the cost. More importantly, you can actually store and analyse it in real time so you can find (and stop) attacks that you would never have been able to discover otherwise.

About This Paper

This paper explains what rich historical metadata is, how you get it and how the insights and analytics it enables can materially transform the way you detect and investigate critical security threats.

What Do You Need to Find...to Find the Attacker?

Believe it or not, finding attackers is not just about finding malware. Malware is *one* of the many tools that attackers use to break in and steal your data. But it's not the *only* tool.

When an attacker gets past your defences and steals your data, it's the result of a *series* of actions and multiple tactics. Perhaps they got their foothold by sending a spear phishing email disguised to look like it came from a trustworthy source. When you click on the link or the attachment, the malware executes. Boom!

The attacker has their initial foothold.

But that is just the beginning. From there, the attacker will execute a series of additional tactics in order to accomplish their mission — none of which necessarily needs to include malware. For example, the attacker might:

- Install web shells to maintain persistence
- Escalate privileges by stealing and cracking legitimate credentials from your domain server

Multiple Tactics = Multiple Opportunities to Find the Bad Guys

Whilst attackers' multiple tactics create challenges for malware-centric security solutions, they can create more opportunities to find them if you have got deep visibility into your network. Why? Because with that visibility comes the ability to see all of the other tactics and techniques that attackers use to entrench themselves, expand control and explore your network. When you find one clue, you can begin to piece the attack together, identify other impacted systems and stop it once and for all.

A few of their tactics that you will want to find — in addition to malware and command and control activity — include:

- SQL injection
- Exploitation of known vulnerabilities and zero-days
- Exploitation of weak authentication
- Use of web shells
- Data and credential theft
- Content staging
- Changing data
- Hijacking services
- Cross-site scripting (XSS)
- Malicious content and services

- Hijack legitimate applications or operating system services, such as your browser or word processor, to perform reconnaissance undetected
- Perform exhaustive searches to find the data targeted for exfiltration and stage content in hidden directories
- Obfuscate data by encapsulating, compressing, transforming or encrypting it in order to send it out of the network.

Whilst there are many complexities and nuances in any given attack, the basic *modus operandi* remains the same: infiltrate, establish command and control, move laterally and exfiltrate data.

This means that to find attackers that are burrowing deep into your network you need to look as deep as they are hiding their exploits. Unfortunately, most "advanced" threat solutions don't go that deep. They are narrowly focused on finding malware coming in and command and control traffic beaconing out. That is a risky strategy. Why? Because, if you are only looking for malware and the attacker isn't using malware (or they are and you miss it), you won't be able to find them.

Creating the Right Haystack to Find the Needle(s)

Increasingly, detection — and especially detecting attackers at each stage of the attack lifecycle — is becoming a data problem. Or, more to the point, it is becoming a big data problem. Organisations are drowning in data and vendors are promoting "capture everything" technologies as the solution. But "big data" does not always mean "smart data." In fact, for many incidents, organisations actually had the data or logs to detect an attack. They just weren't able to analyse or operationalise it effectively. That's a problem.

In order to find attackers who are working hard to stay out of sight, you need to create the right haystack of big data to find them.

Historically, there have been two options for creating that haystack:

1. Stitch a patchwork of security products together, feed events from them into a SIEM and hope that your security analysts can mentally piece together the information.

   or...

2. Invest in a full packet capture system to record every packet. Then, write a big check to your storage vendor to archive all of the information and hire an army of forensic analysts to sift through the packets.

"Big data" doesn't always mean "smart data." In fact, for many incidents, organisations actually had the data or logs to detect an attack. They just weren't able to analyse or operationalise it effectively.

Now, there's a third option.

3. With Fidelis Network's™ Deep Session Inspection® capability you can capture and analyse rich metadata at a fraction of the cost of other alternatives. More importantly, your Tier 1 security analysts can use the resulting intelligence to detect and stop attacks that you'd otherwise need a fully staffed SOC/CIRT team to handle.

Why Is Metadata Such a Big Deal?

Simply put, metadata is data that describes other data. Think about a phone conversation. If you had a recording of the conversation you could listen to every word that was said. But, if you had an easily searchable description of everything that was said you could get almost the same value in a format that was much easier to consume.

It is the same with your network traffic. Until recently, it simply was not possible to capture and store such rich metadata at scale. Now that it is, you can capture metadata about every document and communication protocol. Whilst other network inspection devices can collect *some* metadata, Fidelis Network is unique in its ability to go well beyond the high-level "stream" metadata and collect "rich metadata" from *inside* the session. For instance, for a web session, other vendors collect the source IP, destination IP, the URL and maybe some information about the headers. Fidelis collects all of that, plus rich metadata from *within the web session.*

And that is important, because, the richer the metadata you have, the richer the set of questions you can ask and answer quickly and without the aid of a PhD in forensics. And the richer set of questions you can answer, the better your chances are at detecting and stopping attacks on your network. For example:

The richer the metadata you have, the richer the set of questions you can ask and answer. And the richer set of questions you can answer, the better your chances are at detecting and stopping attacks on your network.

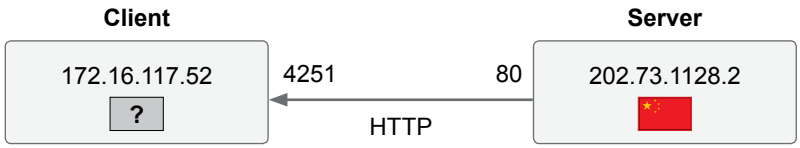
- Have we seen the document or executable being transmitted before?
- Who authored the document and when?
- Does the document have an attachment? If so, is it malicious?
- What is the document type?
- Does the document have tags that describe sensitive data?
- Who else in the enterprise has a copy of the document?
- Was there any personally identifiable information (PII) in the document?
- Who was logged onto the machine that sent the document?

What's Wrong With Netflow Data?

When you put "network" and "metadata" in the same sentence, the mind quickly makes the leap to netflow data. Nothing is "wrong" with netflow data. It provides some basic information about network sessions. However, it has limited use when investigating advanced attacks.

Typical netflow data will tell you that a conversation took place and the language that was spoken. It gives you the source and destination of a session along with the transport protocol, type of service and length of the session. Whilst that is helpful it does not tell you anything about what transpired during that session. So, whilst it is one type of metadata, it is by no means the most important or useful type of metadata that a security analyst could want when investigating a suspected incident.

Sample Netflow Data of Network Session



Duration	<1 second
Sensor	linux71-sen1
Session Start	2016-01-27 10:56:00
Timestamp	2016-01-27 10:56:00
Transport	TCP

Overview of Rich Metadata Captured by Fidelis Network

The rich network metadata that Fidelis Network's Deep Session Inspection technology collects is a combination of attributes that describe the network communication (transport and protocol), applications and file objects in transit on the wire. This detailed information contains all the necessary descriptors to quickly identify and react to malicious traffic and objects during an investigation.

Application- and Protocol-Level Metadata Collected by Fidelis Network

Web Applications

- Web pages and filenames accessed
- Referrer address, user agent string
- Host system IP address, port, geo location
- Connector server type, IP address, port, geo location, connection type
- Object hashes (javascript, images, php, html)
- File size and type transferred

Social Media

- Unique user ID
- Type of access (mail, post, app)
- Information about access (e.g. subject)
- Profile link

Email (SMTP, IMAP, Webmail, etc.)

- Sender's email and IP address
- Recipient's email and IP address
- Subject of email
- Attached filenames and file attributes (see below)

Encrypted Web Access

- Encryption type (SSL/TLS) and version
- Cipher used, bit strength, hash used
- Client/Server FQDN

Internal File Share

- Client accessing the share
- Fileshare, directory, filename accessed
- Type of action (read/write)

Other Attributes

- Instant messaging (Skype, Slack, etc.)
- Bittorrent
- IRC
- Telnet
- FTP and TFTP
- Database content (Oracle, MySQL, etc.)
- Remote connections (RDP, RFP [VNC], Poison Ivy)
- Websockets

Content-Level Metadata Collected by Fidelis Network

Documents (MS Office, PDF, etc.)

- Author
- Filename, file hash
- Header and footer information
- Creation date
- Encryption

Executable Files

- Filename and file hash
- Creation date
- ImpHash
- Operating system (Windows, Mac, Linux)

Archives (zip, rar, tar, gzip, etc.)

- Filename and file hash
- Compression method
- Creation and modification date
- Directories
- ImpHash
- Encryption

Certificates

- Type of certificate
- Subject name and issuer name of certificate
- Start and end date of certificate
- Key length

Embedded Objects

- Creation date
- Modification date
- File name and file hash
- ImpHash

Other Attributes

- Flash
- Java-class
- JavaScript
- Tnet, MIME
- XML

In addition to these standard attributes, users can define a vast array of 'tags' and apply them to stored transactions (e.g. "Based on reputation at the network layer does it look like a TOR session?" or "Based on content inspection do I see the phrase 'Non Disclosure Agreement' in the body?").

Four Deep Secrets Your Metadata Can Tell You

OK. So metadata is good. And rich metadata is even better when it comes to finding attackers. But what can you actually learn from it? Here are four deep secrets you can learn from your data along with specific examples of how organisations are using Fidelis Network to find them.

How, Why and When Were You Compromised?

Every incident responder or security analyst can tell you what happens when you get a "serious" alert. You swing into investigation mode. Inevitably that means pulling logs, triaging endpoints and piecing together disparate data. In many cases the data you want just isn't available.

Determining which systems are compromised and retrieving the data can take days. In a best-case scenario, you can triage the alert. But getting deeper visibility means pulling PCAP data, which often lives off-site. Then, once you've got it, you have to find the right "packets," decode them and analyse the traffic to get to the relevant data. That takes time and expertise — both of which are in short supply during an investigation.

Fidelis automates the collection, analysis and storage of your network data so it's ready for you to investigate immediately. By providing content-enriched metadata in near-real time, Fidelis Network removes much of this complexity. The rich metadata that we capture about every session on your network makes it possible to investigate suspected incidents in seconds and gives you answers to questions that were previously impossible to know.

CUSTOMER CASE STUDY: Turning Analysts Into Investigators

One large retail customer had a four-person security operations team. They managed eight different security solutions, including tools to look for malware, capture netflow and perform full packet capture. In their own words, the team said they "didn't really get to do any security work since they were too busy maintaining the existing security solutions." With Fidelis Network's ability to capture rich metadata they were able to accelerate their investigation and reduce the time required to resolve incidents from days to hours. With rich metadata at their fingertips, the team could quickly pivot from an alert to find the root cause and determine the severity. The context and history that the metadata provided enabled entry-level analysts to operate as if they were Tier 3 analysts.

In one case, they identified a spear phishing email. They tried to use Microsoft Exchange to find similar emails with the same subject line or hash but were unable to do so. Using Fidelis Network they were able to quickly locate and analyse the necessary metadata, identify related suspected spear phishing emails in seconds and determine that there had been unusual activity from some of the impacted endpoints. Those endpoints were quickly contained and emails were removed from the mailboxes of the others who hadn't opened them yet. The full extent of the attack was determined quickly and all the affected systems were remediated.

Have You Been Compromised in the Past?

Here is another familiar scenario. A high-profile data breach generates headlines and a new zero-day exploit, campaign or piece of malware is uncovered. The CEO and the board want to know, "Are we safe?" In most cases, the honest answer is, "I don't know." But that's not what the CEO wants to hear.

The reality is that most security tools are monitoring in real time to detect *future* threats. Whilst many mature security teams subscribe to threat intelligence feeds and participate in intelligence sharing organisations, they have a hard time applying that intel. When security teams get new intel — as they do daily — there is no easy way to look back in time to see if they have been compromised in the past.

How do you find something that you didn't know you should have been looking for? It sounds like a riddle. But the answer, again, is metadata. By capturing rich metadata, you can apply new threat intelligence and indicators of compromise to all traffic — including historical traffic — so you can look back in time and determine if you were affected by the threat. The ability to apply new threat intelligence to historical data is an incredibly unique capability. Not only will it enable you to confidently answer the question, "Are we safe," the next time the CEO asks, it will equip you to detect attacks other solutions can't even see.

How do you find something that you didn't know you should have been looking for? It sounds like a riddle. But the answer, again, is metadata.

CUSTOMER CASE STUDY: Investigating a New Zero-Day Exploit

When a new flash zero-day exploit was publicised a U.S.-based manufacturing company applied it to their environment and discovered that they were compromised and that nation state actors had been active in their environment for a week. Because Fidelis Network stores the hash of every single Flash (SWF) file that had traversed the network, within minutes the security team was able to determine the exploit had been served up in a targeted drive-by. With this knowledge, the incident response team was able to scope the extent of the incident and conduct a thorough response. Because they had observed and profiled the adversary's behaviour, they could then use Fidelis Network to determine that the explusion had been successful.

Are You a Victim of a Multi-Vector Attack?

As we have said, attacks are a series of actions. Once attackers get beyond the initial compromise, they establish a foothold, escalate privileges, move laterally and ultimately exfiltrate data. Unfortunately, most security tools have their sights narrowly trained on specific vectors and tactics, such as malware used in the initial stages of a compromise. When attackers execute a series of stealthy tactics, these tools have no way to connect the dots and see the full scope of the attack.

That all changes when you have rich metadata.

Because threat actors — and even malicious insiders — evolve their tactics in an attempt to defeat your defences, the best detection must focus not just on what is *currently* happening on your network. It must also look at the cumulative effect of attackers' *past* actions. Applying security rules and analytics to retrospective metadata allows you to correlate seemingly unrelated network activity and disparate events across *sessions* and *time* to detect techniques that span multiple vectors.

Applying security rules and analytics to retrospective metadata allows you to correlate seemingly unrelated network activity and disparate events across sessions and time to detect techniques that span multiple vectors.

Here, the use of big data analytics is key to detecting the advanced lateral movement that is often indicative of multi-year cyber espionage campaigns. By using network metadata and endpoint telemetry you can investigate activity going back months or years and find malicious behaviour — no matter how hard attackers work to cover their tracks.

CUSTOMER CASE STUDY: Putting the Pieces Together

In many cases, a single action may not necessarily seem suspicious until you put it in the context of a series of actions occurring over a longer period of time. Fidelis Network customers use the product on a daily basis to identify these types of multi-vector attacks.

In one example, a customer used Fidelis Network to detect ransomware. Most ransomware exhibits similar behaviour. In the case of fileshares, by looking for multiple reads of unencrypted files followed by writes of encrypted files, you can detect ransomware in the early stages.

In another example, Fidelis Network customers routinely use the product to detect multi-vector attacks such as the Angler Exploit Kit by correlating related activity across multiple sessions. Instead of just alerting on a visit to the landing page or the delivery of the exploit, alerts in Fidelis Network tie together each stage of the exploitation chain, so responders can understand how they were initially compromised, what malware was downloaded, and see its beaconing activity. This visibility enables rapid and comprehensive remediation.

What's Actually Going on in Your Network?

Whilst organisations spend millions to build and maintain the network, the dirty secret of most network security teams is that they often do not actually know what is happening on their network. Or, at least they don't know everything.

Whilst "visibility" is a vague and overused term it is exactly what you get when you capture rich metadata. And with that visibility comes the peace of mind that you can always find out what is going on (or has happened) in your network. Whilst that's critical when you are hunting attackers, it's equally useful when you look at how you can optimise your network architecture, enforce policies and achieve compliance.

CUSTOMER CASE STUDY: Our Users Are Doing What?

Every day Fidelis customers learn new things about their network from the rich metadata provided by Fidelis Network. Specific examples of discoveries customers have made include:

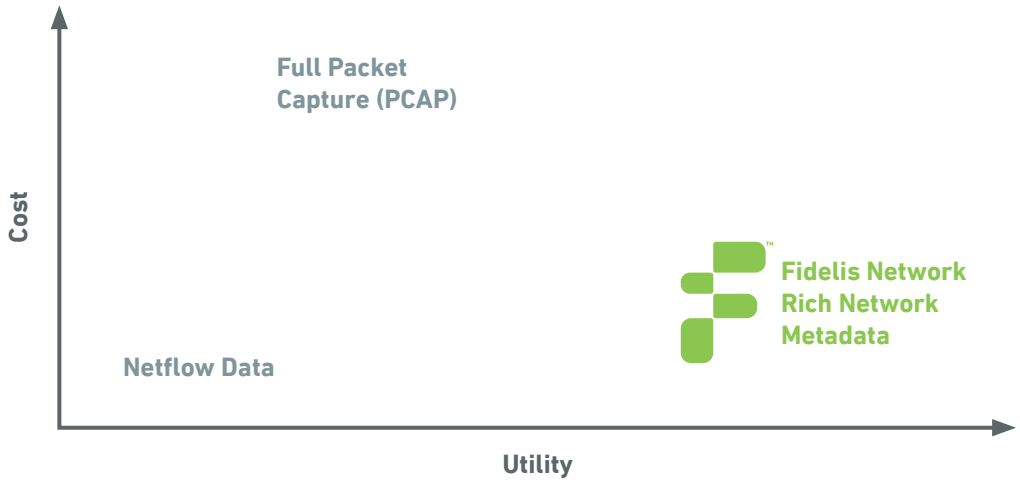
- Confidential information being shared with external partners
- Payment card information (PCI) being shared in an unsecure manner
- Unauthorised use of Remote Desktop Protocol (RDP) by attackers and insiders
- Use of unauthorised cloud services and other shadow IT
- Misconfigured proxies
- Unknown and unsecured network egress points
- New resources such as IPs or domains appearing inside the enterprise.

From Packets to Sessions — How Fidelis Captures Metadata

Historically, the only way organisations could come close to capturing high-fidelity data about what is happening on their network was to invest in a packet capture system. You can think of it as a DVR for your network. Unfortunately, just like a real DVR, because PCAP writes all packets to disk, the storage can quickly fill up (and add up). Think of how fast your DVR would fill up if you wanted to record every hour of every channel each day.

Full packet capture systems have their role in court cases where you need forensic data. But they were never designed to facilitate the detection or investigation of advanced threat actors. Packet inspection technologies get useful information out of the first packet, but that is where they stop. They don't correlate client and server data or decode the content any deeper. Just like a DVR, you have a recording. But the only way to understand what has happened is to "roll the tape" or, in this case, "pull the packets."

Comparison of Cost vs. Utility of Netflow, Full Packet Capture and Fidelis Network



Overview of Fidelis Deep Session Inspection Technology

Fidelis' patented Deep Session Inspection technology is what makes Fidelis Network unique from other network security technologies.

The Deep Session Inspection engine picks up every packet that traverses the network; reassembles those packets into session buffers in RAM; and recursively decodes and analyses the protocols, applications and content objects in those session buffers in real-time — while the sessions are occurring. This gives us the unique ability to "see deeper" into the applications and, in particular, *the content* that is flowing over the network — no matter how deeply or recursively encapsulated, embedded, encoded or compressed that content may be.

The ability to do deep, recursive content decoding and analysis in real time, across all ports and protocols, without depending on a proxy to stop the traffic, is unique to Fidelis. It enables us to see threats that are invisible to other network security systems. And that makes all the difference, because when it comes to looking for advanced threats, the vast majority of them are hidden deep within the content that is flowing over the network.

Finally, because Fidelis can see deeply into network traffic before the communication is completed, we can take decisive action on the traffic, whether alerting or preventing.

Fidelis Network takes a different approach. We monitor all of the packets flowing between each source and destination. But instead of just capturing them we reassemble them. We put the packets together in separate buffers for client and server. Then, we process the whole conversation as a single session. And we do this at line speed, in real time, across all ports and protocols. Next, we analyse the session using our patented Deep Session Inspection® technology (see sidebar) to generate rich metadata about every session.

The Fidelis Network Collector appliance then gathers all of this rich metadata so you get all of the useful attributes from every session. You don't just get the netflow data. You get metadata about every session that occurs on your network in a form that you can access and query. One of the big advantages is that since we are only storing the metadata — and not all of the data — we can cost-effectively scale to fit any network size. The table below outlines some of the key differences between our approach and a full packet capture system.

Comparison of Fidelis Network and Full Packet Capture (PCAP) Solutions

	Fidelis Network	Full Packet Capture (PCAP)
What data is collected?	Fidelis Network sensors collect all rich data about every network session and all embedded content and store it in a single location for aggregated analysis.	Every packet is captured. But because the volume is *huge*, you must be *selective* on what you monitor.
How is the data stored?	Metadata is compact and normalised. It has a much smaller data footprint so it is possible to cost effectively store a much larger set of data for a longer period of time.	PCAP data is large and unstructured. This makes it very expensive to store for long durations. Also, by storing a full copy, the data can be subject to regulation (e.g. PII, financial, ITAR, etc.).
How can analysts work with it?	Since the data is normalised, searches are much faster and it is easier to apply analytics across multiple sessions.	Searches are slower because analysts must reassemble the data and then process a much larger data set.
How can you apply threat intel?	Intel can easily be automatically applied to historical data. When you identify suspicious activity it is easy to get additional context.	It is not possible to apply threat intel to stored PCAP data.

Conclusion

If you've got this far, you will know what rich metadata is, why it's so valuable and what some of the deep secrets are that you can learn from it. The most common reaction we get from customers once they understand the capabilities of Fidelis Network and our Deep Session Inspection is: "I had no idea that was even possible." When we visit them after they have been using it for a few months they tell us "I don't remember what I did before I had this capability."

The reality is that there is currently no other solution on the market that can analyse and store the type of rich metadata about your network sessions as Fidelis Network.

Fidelis Cybersecurity is creating a world where attackers have no place left to hide. We reduce the time it takes to detect attacks and resolve security incidents. Our Fidelis Network™ and Fidelis Endpoint™ products look deep inside your traffic and content where attackers hide their exploits. Then, we pursue them out to your endpoints where your critical data lives. With Fidelis you'll know when you're being attacked, you can retrace attackers' footprints and prevent data theft at every stage of the attack lifecycle. To learn more about Fidelis Cybersecurity products and incident response services, please visit **www.fidelissecurity.com** and follow us on Twitter **@FidelisCyber**.