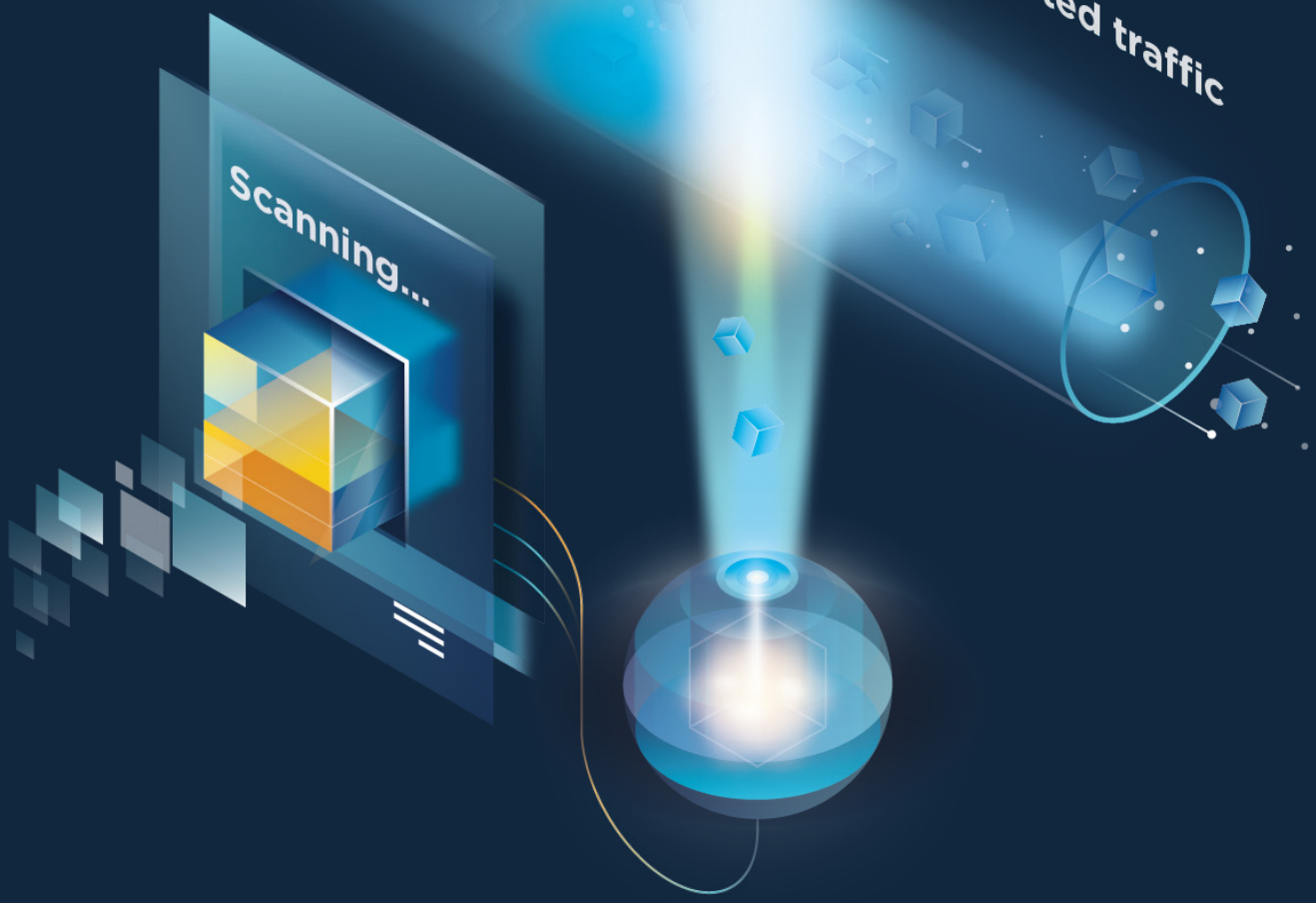
Progress® | Flowmon



Whitepaper

Encrypted Traffic Analysis

The data privacy-preserving way to regain visibility into encrypted communication

Executive Summary

Encryption is considered as security by design. It undoubtedly helps to avoid risks such as communication interception and misuse. Therefore it is natural that all responsible organizations adopt encryption as an important way of protecting business critical applications and services. According to Gartner 80 % of web traffic will be encrypted in 2019.

Ironically, encryption as a security measure created a grey zone of traffic with unlimited space for attackers to hide their activity. And when the volume of encrypted traffic grows year by year, this is a challenge for security professionals to keep their assets secure.

Unfortunately, traditional packet analysis based network measuring solutions for obvious reasons cannot understand what's inside such traffic. Consequently, effective troubleshooting, security monitoring and compliance enforcement are paralyzed.

Flowmon overcomes the inability of getting actionable network insights by introducing the concept of Encrypted Traffic Analysis, the only privacy-preserving and ultimately scalable way of understanding modern encrypted communication. Given such functionality of the Flowmon solution to automatically filter genuinely relevant data, tremendously streamlines malware and data exfiltration detection, vulnerability assessment and troubleshooting. This approach is much less privacy invasive and more cost efficient than the legacy solution of using SSL proxies to decrypt traffic, analyse it and then encrypt again.

50%

of all known cyber attacks use encryption to evade detection. In 2013 the number was below 5 %.

2/3

of all known cyber attacks use encryption to evade detection. In 2013 the number was below 5 %.

Decrypt or Analyse?

Although threats misusing SSL/TLS to mask their activities are on the rise, most businesses do not have capabilities to detect such behavior. When encryption is becoming standard in both web traffic and enterprise traffic, ignoring this trend poses serious risks to business security. Facing this challenge, is it better to use a decryption tool or traffic analysis? Let's compare both approches.

SSL/TLS Decryption Proxy	Encrypted Traffic Analysis
Decryption = Privacy intrusion	Privacy-preserving by working with metadata
Inline proxy is a potential point of failure	Non-intrusive monitoring
Affecting latency	Passive from network perspective
Demands higher throughput and related costs	Scales with your network
Demands time consuming key/certificate management	Use case: Real-time monitoring
Use case: blocking	

Distinguish Malicious from Legitimate

It is misleading to assess a website's credibility only by considering the presence of a signed HTTPS certificate as we were used to in the past. Although this is the very purpose of a certificate, HTTPS itself doesn't mean that the website isn't infected with malware that could mine cryptocurrencies using your resources or do other harm. There are many websites out there with a single purpose, to harvest your data, such as passwords or personal information.

The process of having a website scanned for malicious code and getting the certificate from responsible authorities is quite costly, which created the market of inexpensive (or free) and untrustworthy providers to emerge. And the result, as expected, is that even malicious websites give the illusion of being harmless. Successful risk classification of websites, because they present a certificate issued by a certification authority, is therefore not fully reliable. Within our internal network we are familiar with legitimate certificates and their issuers. Therefore, monitoring and detecting unusual certificates, which can indicate a man-in-the-middle attack, is very easy with ETA. However, that's not the case of external services, because of the reasons described above.

It is not typical that all the servers within a cluster face

Benefits of ETA

Compliance

- Monitoring expired and non-compliant SSL certificates
- Encryption strength (key length,...)
- Unwanted TLS versions that contain vulnerabilities
- Non-compliant clients

Cybersecurity

- Identification of malware infected stations
- dentification of C&C communication
- Discovering man-in-the-middle attacks
- Monitoring data exfiltration

issues at the same time, unless there is a systemic problem. Usually, a slowdown is caused by a specific instance within the ecosystem. Having detection methods to narrow down to that level of granularity

is key, and can be achieved at the architecting stage, when mechanisms should be put in place to be able to extract that information. They could be placed on the application ecosystem and the elements of the network that connect to it, be it via network performance or via the configuration of the load balancer.

Encrypted Traffic Analysis overcomes this problem as it can describe the characteristics of encrypted communication of a variety of services, such as Windows updates, secure shell, Firefox or Chrome and indicate unusual or unexpected source of communication from a specific computer. Each application is described as a set of SSL parameters. With the knowledge of those parameters we can track and measure those applications across the network for malware detection and other security, compliance and operational purposes. Definitions of such sets of parameters can be simplified within a single data field called a JA3 fingerprint, a signature of encrypted application traffic as seen on the network. Effectively, we can red flag malicious communication based on its characteristics, rather than relying on a certificate itself, or its issuing authority.

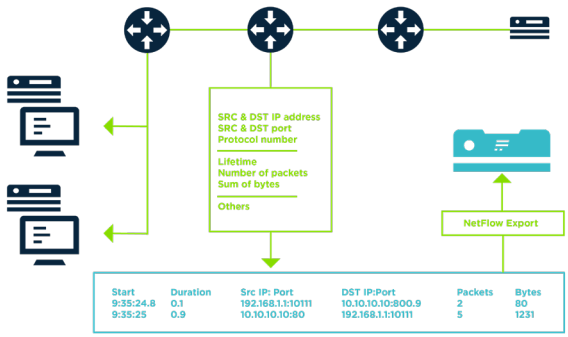
How Can Encrypted Traffic be Monitored?

To understand this, we need to explain how flow analysis works. Flowdata is a single packet flow in a network with the same identifying 5-tuple, composed of source IP address, destination IP address, source port, destination port and protocol. Based on this, packets are aggregated into flow records that accumulate the amount of transferred data, the number of packets and other information from the network and transport layer. However, in 2012 at Flowmon, we came up with the concept of flow data enriched with information from the application layer. This concept has recently been widely adopted by many other vendors. So, having detailed visibility into application protocols, such as HTTP, DNS and DHCP is no issue these days. In fact, this moves troubleshooting use cases to a completely different level.

Depending on what information you want to export, Flow is 100-500× more scalable than packet analysis. When enriched with packet level information, performance statistics and application monitoring, it provides a comparable level of detail to packet analysis for a fraction of the price. Support of 100Gbps line rate performance is not an obstacle and data retention can be weeks or months.

How Encrypted Traffic Analysis Works

Encryption by design ensures that the context of the communication is hidden. Therefore, it is impossible to apply security countermeasures onto the context of the communication and instead it needs to focus on the characteristics of such traffic. Any communication can be encrypted. There is no significant difference between legal and illegal communication from an encryption point of view. TLS (a successor to SSL) handshake is a non-encrypted session through which client and server negotiates the encryption rules. Only after a secure channel is established, the traffic becomes encrypted. By reading the handshake and its specific parameters we can identify unusual behaviour.



As we can see on the picture, Flowmon Probe analyzes packets and extracts important information from L2, L3-L4 and L7 layers. Then it generates IPFIX records, which is NetFlow enriched with additional data, in this scenario TLS details. The IPFIX records contain all the important packet information in an aggregated form without the necessity of storing entire packets. This ensures the reduction of massive storage requirements. By pre-filtering the relevant data for the user and presenting them in an organized way with dashboards, reports and drill downs, the time for analysis is reduced to a minimum. With such data, we can assess very easy certificates' validity and credibility at network level, the most effective privacy-preserving method.

The use cases of ETA can be loosely divided into two categories. The first being the cryptographic assessment where we investigate SSL/TLS protocol versions, cybersuites (encryption algorithms, key lengths) as well as certificates. The second category is monitoring and security. Here, we can leverage JA3 fingerprinting to pinpoint suspicious actors, use Application-Layer Protocol Negotiation (ALPN) to identify a protocol in encrypted communication, investigate Server Name Indication (SNI) and much more.

Benefits of Enriched Flow Data

The dynamics and diversity of today's networks challenge the prevailing network monitoring approach. Facing an increase in network speeds, visibility gaps caused by migration to cloud, IoT and software defined networking, packet capture solutions struggle to bring expected results promptly and at a reasonable price. Packet capture solutions were designed at the time when the dynamics of today's network environments would have been hard to imagine. Nowadays, they work well in specific use cases, but they cannot cope with the flexibility, scalability and ease of use of flow data in most of the everyday use cases that network engineers face.

We in Flowmon Networks believe that merging flow and packet level visibility into one versatile solution is the technology that will help you scale to future performance and capacity needs. So, let's do continuous flow monitoring and packet capture when needed. At the end of the day, you will most probably need to analyze PCAPs less than you expect.

Storage requirements for Flow vs. Packet Analysis	**Packet Analysis**	**Flow monitoring**
1 minute	75 GB	**0,15 GB**
1 hour	4 500 GB	**9 GB**
1 day	108 000 GB	**216 GB**

Shortlist of Flowmon exported TLS data

TLS server version
TLS cipher suite
TLS server name indication
TLS client version
TLS certificate issuer common name
TLS subject common name
TLS public key algorithm
TLS certificate validity until
TLS JA3 fingerprint

For the full list, please refer to our Flowmon Supported Flow Standards documentation (restricted access).

Demonstration and Use Cases

Cryptographic Compliance

Cryptographic assessment or compliance is when we audit our encryption stack to look for outdated protocols still in use, inadequate key lengths and expired certificates.

TLS protocol version

With Flowmon it's easy to inspect outdated TLS versions in your network. Take a look at this screenshot from Flowmon Monitoring Center:

TLS CLIENT VERSION	TLS SERVER VERSION	DEFAULT BYTES: INPUT	PACKETS	FLOWS
TLS 1.0	TLS 1.0	1.5 G	1.6 M	24
TLS 1.2	TLS 1.2	211.1 M	268.2 K	7060
TLS 1.2	TLS 1.0	28.5 K	139	6
		Flows 7.09 K	Bytes 1.7 G	Packets 1.9 M

TLS version 1.0 was released way back in 1999 (as a successor to SSL). This almost 20-year-old protocol is vulnerable to various attacks (such as POODLE), uses weak cryptography and is no longer compliant with the Payment Card Industry Data Security Standard (PCI-DSS).

Looks like a lot of traffic is still using TLS 1.0. We can dig deeper and find out which local machines / services need to be upgraded:

DESTINATION IP ADDRESS	TLS SERVER VERSION
192.168.70.3	TLS 1.0
192.168.47.90	TLS 1.0

Public key length and algorithm

Weak short keys and outdated algorithms are a serious security risk. Insufficient key length makes it easier for an attacker to perform brute force decryption. Outdated algorithms suffer from vulnerabilities malicious actors can exploit to break in (think Heartbleed). We should always check the key length and algorithm as one because different algorithms require different key lengths, for example, elliptic curve cryptography algorithms (ECC) have shorter keys while having equivalent key strength as RSA (RFC 4492).

Let's inspect public key lengths and algorithms in our network with Flowmon:

DEFAULT BYTES: INPUT	FLOWS	TLS PUBLIC KEY ALGORITHM	TLS PUBLIC KEY LENGTH
37.6 G	371715	rsaEncryption	2048
3.2 G	135570	id-ecPublicKey	256
175.5 M	18608	rsaEncryption	4096
1.8 M	750	rsaEncryption	1024
38.4 K	10	id-ecPublicKey	384
13.2 K	4	rsaEncryption	3072
	Flows 526.66 K	Bytes 41.0 G	Packets 44.2 M

There is an amount of traffic with only 1024-bit RSA keys. We can use filters to focus only on this traffic:

SOURCE IP ADDRESS	TLS ISSUER COMMON NAME	TLS SUBJECT COMMON NAME	TLS PUBLIC KEY ALGORITHM	TLS PUBLIC KEY LENGTH
161.69.169.29	TrustedSource_CA	TrustedSourceServer_IMQA01	rsaEncryption	1024
161.69.169.29	TrustedSource_CA	TrustedSourceServer_IMQA01	rsaEncryption	1024
161.69.169.27	TrustedSource_CA	TrustedSourceServer_IMQA01	rsaEncryption	1024

Such anomalies can be alerted, reported on and tracked on using the Dashboard.

Expired certificates

While checking for expired certificates is an obvious step, we might also want to check for soon-to-be-expired certificates to prepare in advance. An example of the expired certificate (notice the value in the last column):

DESTINATION IP ADDRESS	TLS ISSUER COMMON NAME	TLS SUBJECT COMMON NAME	TLS CERTIFICATE VALIDITY FROM	TLS CERTIFICATE VALIDITY TO
192.168.70.62	Go Daddy Secure Certificate Authority - G2	*.foxitsoftware.com	2015-02-07 08:14:38	2018-02-05 20:43:43

Now let's see the certificates that will expire between today and the endof the year:

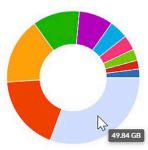
TLS ISSUER COMMON NAME	TLS SUBJECT COMMON NAME	TLS CERTIFICATE VALIDITY FROM	TLS CERTIFICATE VALIDITY TO
GeoTrust TLS RSA CA G1	demo.flowmon.com	2018-08-27 02:00:00	2018-12-28 13:00:00

Security and Monitoring

Malicious threats are increasing the adoption of SSL/TLS. According to Gartner, by 2019, more than 50% of new malware campaigns will use various forms of encryption and obfuscation to conceal delivery and ongoing communications, including data exfiltration. Encrypted traffic will carry over 70% of web malware by 2020.

JA3 fingerprinting

One of the easiest way to spot malicious threats or at least to provide an indicator of compromise (IoC) is JA3 fingerprinting. JA3 combines the five parameters of TLS communication (version, ciphers, extensions, elliptic curves and its formats) and produces a MD5 hash. This is our fingerprint. Interestingly, enough to identify various clients. For example, "e7d705a3286e19ea42f587b344ee6865" is the JA3 fingerprint for a standard TOR client. The reasoning behind this method is the idea that tools are more difficult to change than IP addresses or domain names. This method is very much in the field of signatures,



COLOR	START TIME - FIRST SEEN	DURATION	TLS JA3 FINGERPRINT	FLOWS	INPUT PACKETS	INPUT BYTES
1	2018-11-05 09:28:50	1 d, 5 m, 34.026 s	8c23d614aa018ed7bc6c88b545ece240	1.14 M (39.4%)	52 M (23.4%)	49.84 GB (24.2%)
2	2018-11-05 10:56:32	18 h, 39 m, 47.604 s	e6c5fac1007525a251453da8f7acceb8	152 (0.0%)	31.32 M (14.1%)	29.79 GB (14.5%)
3	2018-11-05 09:29:16	1 d, 5 m, 5.481 s	1885aa9927f99ed538ed895d9335995c	405.59 K (14.0%)	25.58 M (11.5%)	26.25 GB (12.8%)
4	2018-11-05 10:55:55	18 h, 36 m, 31.526 s	7cbc74d83663b6824ad5e41f0e3d4653	86 (0.0%)	18.14 M (8.2%)	18.62 GB (9.1%)
5	2018-11-05 09:31:49	1 d, 50.72 s	4a8ac798fe0f7d3c8ffd86b4a5f306d7	59.24 K (2.0%)	13.4 M (6.0%)	13.53 GB (6.6%)

and is highly reliant on available blacklists and whitelists. Nevertheless, you can leverage it to find outliers and other oddities in your network.

Protocol identification with ALPN

Application-Layer Protocol Negotiation (ALPN) is a TLS extension. With it, we can identify L7 layer protocols inside encrypted traffic. Supported protocols are several versions of HTTP, SPDY, NAT, WebRTC, FTP, IMAP, CoAP and other experimental protocols.

Server Name Indication (SNI)

Similarly to ALPN, SNI is a TLS extension. It allows for TLS-capable servers to host multiple services on the same IPs. Clients add this extension with the hostname of the website they want to connect to.

Looks like one of our employees was violating company policy by visiting file sharing websites:

TLS ISSUER COMMON NAME	TLS SUBJECT COMMON NAME	TLS SERVER NAME
COMODO RSA Domain Validation Secure Server CA	*.similarweb.com	www.similarweb.com
Let's Encrypt Authority X3	mega.nz	mega.nz
COMODO ECC Domain Validation Secure Server CA 2	ssl436500.cloudflaressl.com	openload.co
AlphaSSL CA - SHA256 - G2	*.zippyshare.com	zippyshare.com
Let's Encrypt Authority X3	filecrypt.cc	filecrypt.cc

Conclusion

Malware can now use encryption to hide initial infection when transporting payloads and when communicating with command and control centres. In fact, Gartner predicts that half of malware campaigns in 2019 will use some type of encryption to conceal malware delivery, command and control activity, or data exfiltration. And companies risk serious consequences when ignoring this trend.

In this situation, when it no longer makes sense to monitor your network with packet capture solutions, flowbased approach offers lightweight solution that enables inspection of encrypted traffic without the need to breach user privacy. Simply put, with Flowmon it doesn't matter whether the content of communication is encrypted or not. Flowmon leverages information and metrics from lower network layers, which are not encrypted. Thanks to this approach, our network and security monitoring solution works even on encrypted traffic.