Flowmon
Driving Network Visibility

Handbook

HOW TO ANALYZE AND UNDERSTAND YOUR NETWORK

Part 3:
Network Traffic Monitoring or Packet Analysis? Get the Benefits of Both

by Pavel Minarik, Chief Technology Officer
at Flowmon Networks

In previous two articles, we took a look at two different approaches to network traffic monitoring and analysis. We described principles of flow-based traffic monitoring and the complete analysis of full packet traces. But why we should we enjoy the benefits of one or the other only? In this article, we will describe a modern solution that combines the benefits of both approaches.

> Flow data represents an abstraction of the network traffic, an aggregation based on the source IP address, destination IP address, source port, destination port and protocol number. The content of the communication is not stored, and the achievable aggregation rate reaches 500:1. Packet analysis is focused on recording and analysing full-scale network traffic, including the application layer. Therefore, it is a very performance and disc capacity demanding method.

1 Pros & Cons

Let's take a look at the main pros and cons of both approaches as presented in Table 1. It's obvious that flow data doesn't contain enough information for some tasks. By contrast, as a result of packet analysis, the IT department is usually overloaded with barely manageable volumes of detailed data. When we combine both perspectives and extend the traditional flow data by information from the application layer, we can get an appropriate detail which will provide us with an insight into data communication, flexible reporting, and effective troubleshooting of operational issues and the automatic detection of security incidents.

	Strong points	Weak points
Flow Data	• Works in high-speed networks • Resistant to encrypted traffic • Traffic visibility and reporting • Network Behaviour Analysis	• No application layer data • Not enough information for some tasks • Sampling (routers, switches)
Packet Analysis	• Full-scale network traffic capture • Provides sufficient detail for troubleshooting and analysis • Supports forensic analysis • Signature-based detection	• Useless for encrypted traffic • Very resource consuming • Too many details for the vast majority of tasks

This extension of flow data standard, by information from the application layer, is enabled by the international flow data standard named IPFIX. IPFIX has set up a lot of new attributes that are based on application layer information. To facilitate an expansion of this standard between vendors, IPFIX is equipped with so-called "enterprise extensions" that smoothly broaden the scale of the information provided. One of the most important innovations is the signature-based identification of an application. A part of the network traffic statistics is the application ID, which is determined by the application protocol classification mechanism. Thanks to the first few bytes of an application layer, it is possible to recognise hundreds of applications.

The best known implementation of this technology is Cisco's NBAR2 (Next Generation Network-Based Application Recognition). Flow data monitoring is combined with a continual packet analysis that extends the traffic statistics by the name of an application or application protocol. Based on this information, modern flow collectors can enable traffic reporting and analysis.

One of the most widespread communication protocols is HTTP, or its encrypted version HTTPS. Today it is used to provide access to websites, but this is not its sole function. The protocol is also the basis of communication between the components of business systems, or applications working with sensitive data (i.e. electronic banking). By identifying this transfer protocol, we are able to extend flow data statistics by fundamental HTTP request attributes – a hostname or URL information. Thanks to SNI (Server Name Indication), we are able to get host name information even when the HTTPS protocol is used. SNI is a mechanism by which a client indicates which hostname they are attempting to connect to at the start of the handshaking process.

Start Time - first seen	Duration	Source IP address	Destination IP address	HTTP hostname	HTTP URL	Source Port	Destination Port	Packets	Bytes
2014-09-05 18:45:02.153	0.181 s	192.168.0.76	bud02s02-in-f1.1e100.net	clients2.google.com		54633	https	13	1834
2014-09-05 18:45:38.539	0.071 s	192.168.0.58	bud03s03-in-f8.1e100.net	safebrowsing-cache.google.com		55356	https	5	404
2014-09-05 18:45:38.189	0.305 s	192.168.0.58	prg02s12-in-f1.1e100.net	safebrowsing.google.com		55354	https	12	1889
2014-09-05 18:45:38.594	0.494 s	192.168.0.58	bud02s03-in-f8.1e100.net	safebrowsing-cache.google.com		55357	https	24	4035
2014-09-05 18:45:38.162	0.053 s	192.168.0.58	prg02s12-in-f1.1e100.net	safebrowsing.google.com		55353	https	5	398
2014-09-05 18:46:03.406	0.275 s	192.168.0.112	de-in-f99.1e100.net	www.google.com		63783	https	11	3767
2014-09-05 18:46:03.345	0.318 s	192.168.0.112	de-in-f99.1e100.net	www.google.com		63782	https	5	729
2014-09-05 18:41:35.560	4 m, 55.424 s	192.168.0.112	bud02s02-in-f22.1e100.net	mail.google.com		63658	https	28	7658
2014-09-05 18:47:24.358	0.293 s	192.168.0.39	bud02s02-in-f6.1e100.net	safebrowsing-cache.google.com		59524	https	24	4025
2014-09-05 18:47:24.195	0.109 s	192.168.0.39	prg02s12-in-f1.1e100.net	safebrowsing.google.com		59522	https	10	2061
2014-09-05 18:47:24.323	0.070 s	192.168.0.39	bud02s02-in-f6.1e100.net	safebrowsing-cache.google.com		59523	https	5	729
2014-09-05 18:47:24.170	0.051 s	192.168.0.39	prg02s12-in-f1.1e100.net	safebrowsing.google.com		59521	https	5	729
2014-09-05 18:48:01.479	0.239 s	192.168.0.76	bud02s02-in-f1.1e100.net	clients2.google.com		54659	https	8	1543
2014-09-05 18:48:22.435	0.291 s	192.168.0.45	de-in-f99.1e100.net	www.google.com		55167	https	10	4676
2014-09-05 18:48:44.208	0.121 s	192.168.0.71	de-in-f104.1e100.net	www.google.com		59104	https	5	729
2014-09-05 18:48:44.268	0.256 s	192.168.0.71	de-in-f104.1e100.net	www.google.com		59105	https	11	3707
2014-09-05 18:50:33.624	0.050 s	192.168.0.24	prg02s12-in-f7.1e100.net	safebrowsing.google.com		55798	https	5	729
2014-09-05 18:50:33.773	0.059 s	192.168.0.24	bud02s02-in-f14.1e100.net	safebrowsing-cache.google.com		55800	https	5	404
2014-09-05 18:50:33.649	0.098 s	192.168.0.24	prg02s12-in-f7.1e100.net	safebrowsing.google.com		55799	https	10	3061
2014-09-05 18:50:33.803	0.234 s	192.168.0.24	bud02s02-in-f14.1e100.net	safebrowsing-cache.google.com		55801	https	22	3945
						Flows 20	Bytes 45.76 KB		Packets 221

Figure 1: The listing of flow data extended by HTTP host name information, even if traffic is encrypted. The URL is not provided since the request is encrypted.

Similarly, we can get other information from HTTP communication; for example the operating system and its version, the identification of a browser and its version or a device type in case of mobile phones. This information is a part of the attribute known as the User Agent. The User Agent is a textual string which is a part of the client's request. Thanks to this information, it is possible to detect the device type, devices that use a legacy operation system or web browser, or even to detect new devices in the network. This information can also be used in case of a security incident. When running DNS traffic, we can monitor the type of query and domain name or the DNS server response. The end-station which receives a significant number of "non-existing domain" answers is suspicious and warrants the attention of the system administrator. Moreover, the domains are related to reputation databases so that the efficient detection of suspicious communication is quite easy, i.e. communication with known botnet command and control centres. Integration with IP reputation databases and host name reputation databases should be included in every modern network traffic analysis solution. In addition, many quality databases are available for free, i.e. a database of known attackers is available at D-Shield.org.

2 Use Case: Protection Against Attacks on VoIP System

Voice over Internet Protocol (VoIP) is widespread in today's businesses. By analysing the Session Initial Protocol (SIP) which is responsible for signalling and controlling voice communication sessions, we can get an overview of real telephone calls. By analysing flow data itself, it is possible to measure the quality of the call. When monitoring a SIP signalisation we can control client registration requests - in particular VoIP gateways and detecting security threats. This means that an overview of telephone calls is part of regular network traffic monitoring and can be used when troubleshooting call quality issues.

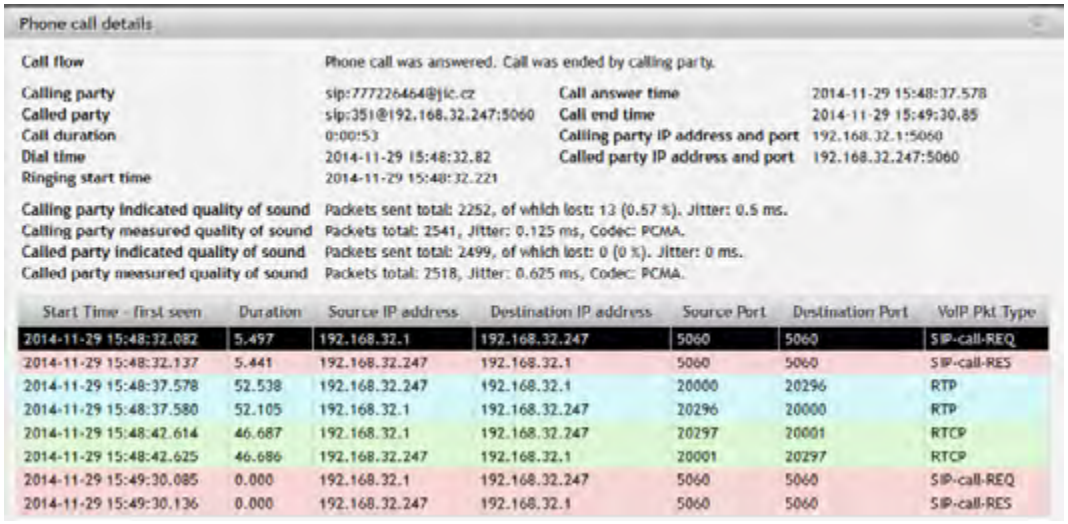

Figure 2: We can see details of the phone call, including the listing of flow data. Information provided by packet analysis includes calling parties, time stamps, an audio codec and information about audio quality.

3 Use Case n.2: SIP monitoring

Let's take a look at an example from the IP telephony security area where we utilise information from application analysis of the SIP protocol. One of the most popular attacks is toll fraud, which aims to get money by performing fraudulent phone calls. This type of fraud should not be underestimated, as total losses caused by this technique are estimated to be 72 billion dollars per year. What is the principle? An attacker establishes a company abroad and rents a premium service to run legitimate high-cost numbers. Afterwards, the attacker founds a poorly-configured SIP gateway in some organisation. Subsequently, the attacker utilises a compromised machine to establish a large amount of calls to his premium number via the poorly configured gateway. By the end of the month, the organisation receives a very large telephone bill. Since the service was delivered to the organisation by the operator, there is no way to avoid paying the bill.

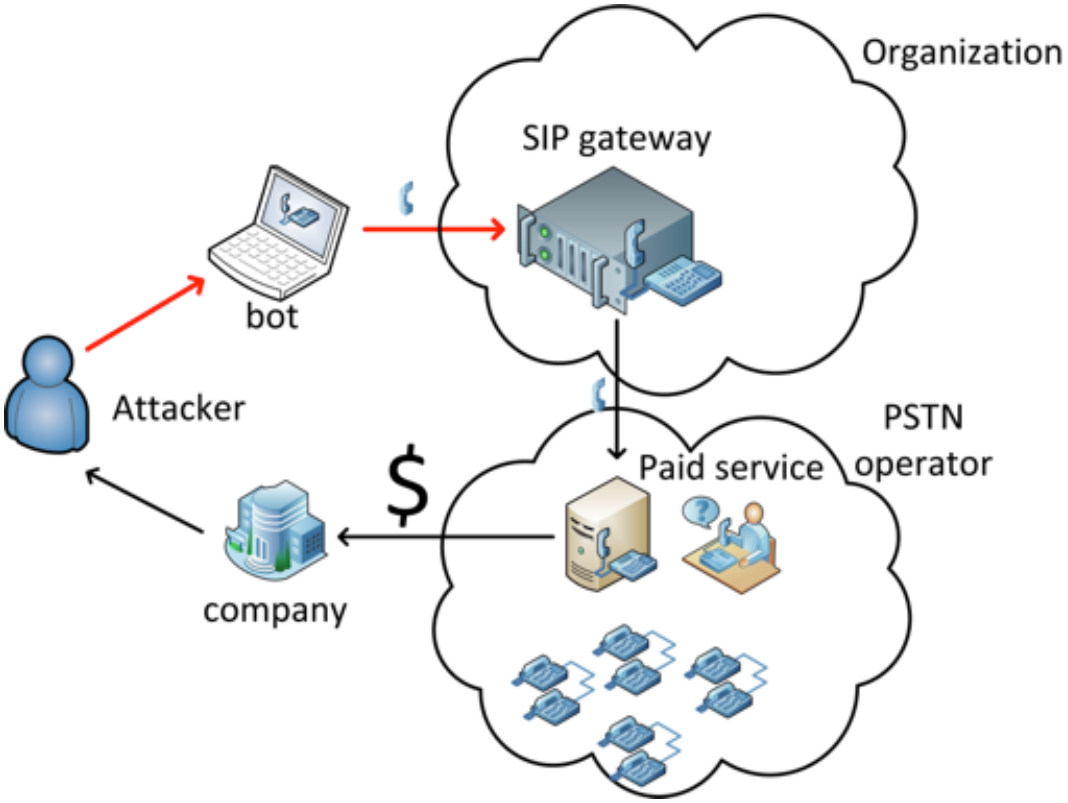

Figure 3: The principle of unsecured VoIP gateway abuse. The attacker performs fraudulent phone calls on a premium high-cost service number which is under his control.

An attack starts with a number of unsuccessful INVITE or REGISTER scans. These scans are part of the SIP protocol signalisation. When we monitor an occurrence of these incidents and set up a notification in case of their rapid increase, we are able to detect the attack even before an occurrence of financial loss or a minimisation of such losses. However, VoIP gateways are not the only targets of toll fraud. Alternatively, the attacker can take advantage of machines infected by botnets in the local network. Such end-stations are usually able to communicate with the gateway without restrictions so may attempt to gain access to its configuration by a SSH dictionary attack.

Detecting attacks which utilise SIP signalisation is based on the principles of behaviour analysis, i.e. by an automatic analysis of monitored information entropy (IP addresses which communicate with gateways, identification of calling parties, INVITE and REGISTER messages etc.) we are able to detect suspicious activities and unauthorised connection attempts.

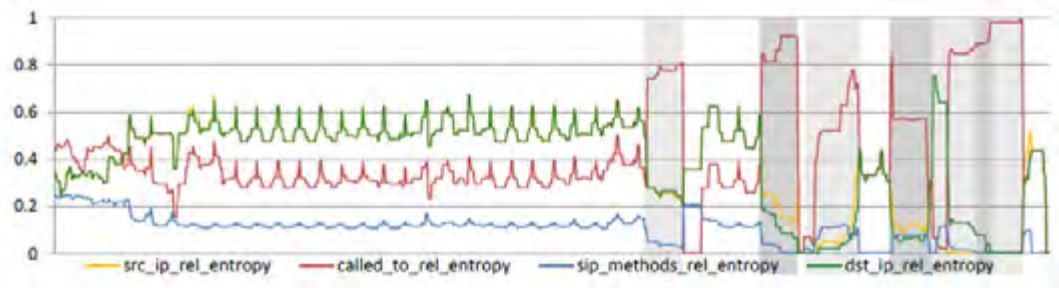

Figure 4: In this picture changes in entropy of SIP traffic characteristics during the running attack are shown. Attacks are marked by the grey fields.

4 Get the Benefits of Both Approaches

We have demonstrated how we can combine the benefits of flow data monitoring and application layer analysis. As a benefit we get more detailed information about data communication, better capabilities of traffic analysis and anomaly detection. At the same time we preserve the excellent compression/aggregation rate of network traffic statistics against original traffic volume. If necessary, it is always possible to carry out a full-scale traffic recording.

Author

RNDr. Pavel Minarik, PhD.

Pavel Minarik has worked in the area of cyber security since 2006. During this time he has participated in several research projects as a senior researcher at the Institute of Computer Science at Masaryk University. He is the author of more than ten publications in the domain of behaviour analysis and numerous algorithms for traffic processing and anomaly detection. As Chief Technology Officer at Flowmon Networks, Pavel is responsible for the technology roadmap, product design and development, as well as technical support and customer projects worldwide.

About Flowmon Networks

Flowmon Networks empowers businesses to manage and secure their computer networks confidently. Through our high performance network monitoring technology and lean-forward behavior analytics, IT pros worldwide benefit from absolute network traffic visibility to enhance network & application performance and deal with modern cyber threats. Driven by a passion for technology, we are leading the way of NetFlow/IPFIX network monitoring that is high performing, scalable and easy to use. The world's largest businesses, internet service providers, government entities or even small and midsize companies rely on our solutions to take control over their networks, keep order and overcome uncertainty. With our solution recognized by Gartner, recommended by Cisco, Check Point and IBM, we are one of the fastest growing companies in the industry.
www.flowmon.com