Flowmon
Driving Network Visibility

Handbook

HOW TO ANALYZE AND UNDERSTAND YOUR NETWORK

Part 2
Inside the Data: Full Packet Capture and Analysis

by Pavel Minarik, Chief Technology Officer
at Flowmon Networks

When it comes to network traffic monitoring, troubleshooting or threat detection, there are two options at our disposal. The first one is the NetFlow-based traffic monitoring we described last time. The other is called full packet capture and analysis that provides complete network visibility.

First of all, let's remind ourselves of the principle of flow-based (NetFlow, IPFIX) network traffic monitoring. Flow data represents an abstraction of the network traffic itself. Flow data statistics are created as an aggregation of the network traffic; using the source IP address, destination IP address, source port, destination port and protocol number as attributes that identify the individual flow records. The content of the communication is not stored, and the achievable aggregation rate is about 500:1.

> Thanks to flow data, we are able to analyse traffic structure, identify end-stations transferring large amounts of data or to troubleshoot network issues and wrong configurations.

One of the newest capabilities of flow data utilisation is a security analysis called Network Behaviour Analysis (NBA). NBA provides qualitatively different results than signature-based methods, therefore allowing to detect and respond to anomalies, undesirable behaviour or yet unknown or specific threats without the available signature.

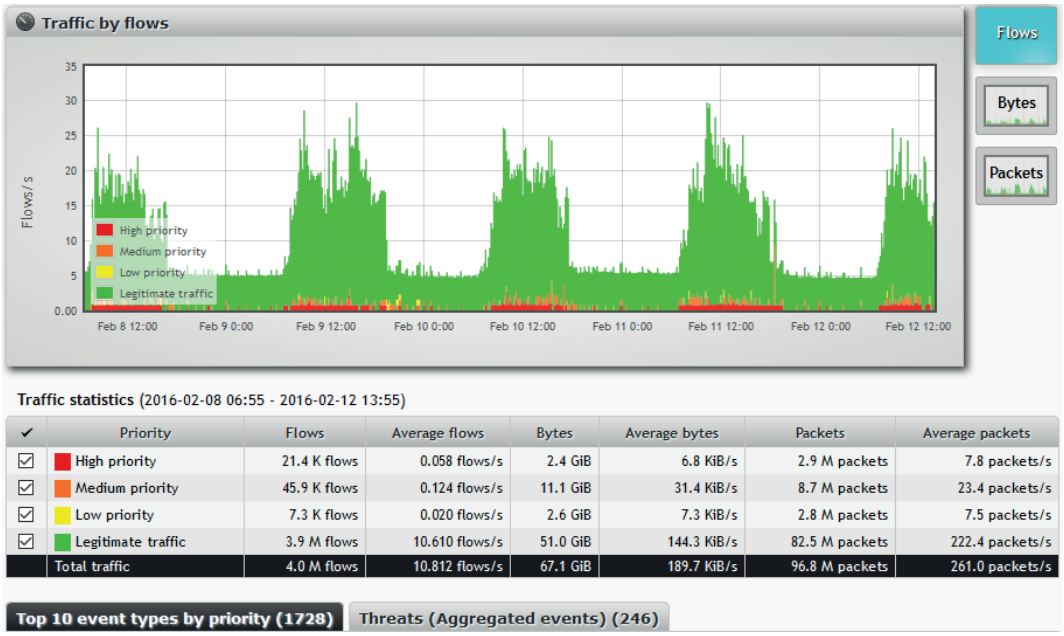

Figure 1: Network behaviour analysis in practice: The Graph shows volumes of legitimate and suspicious traffic.

1 Packet Analysis

Packet analysis looks inside communications to analyse its content. There is no aggregation, compression or trimming involved, and data is stored in its original size. Therefore, this method has extremely demanding performance and disc capacity requirements.

> Just imagine the capturing of a network with 250 Mbps traffic on average. It equals a data load with more than 31 MB per second, 1.8 GB per minute, 108 GB per hour and 2.6 TB a day. In case of 10Gbps networks we are reaching numbers which are hardly believable – it would be more than 100 TB of stored data per day.

However, large volumes of data are not the only drawback. The principal limitation of packet analysis is encrypted traffic. Without the encryption key, we are not able to understand the content of any transferred data, and often not even uncover the transfer protocol or application. Nevertheless, volumes of encrypted traffic constantly grow.

2 Two Approaches

There are two different approaches to packet analysis. The first one represents a continual, full-scale traffic recording (full packet capture). This demands appropriate technical equipment, especially high speed storage arrays with adequate capacity. Such an approach is very expensive, and is therefore suitable only for critical infrastructure and networks with a specific purpose. It must be underlined that storing such data may not be the only problem, as the effective analysis and "mining" of information is also very demanding.

The other approach is so called on-demand packet capture. When employing this approach, we are capturing packets only in cases of need - typically when we deal with system compatibility issues – upon discovering missing or damaged packets etc. On-demand packet capture is a very simple method and affordable to literally every network administrator, but it does have its pros and cons. The limitation of this approach is the fact that the administrator has to determine in advance which traffic should be stored. Therefore, there is no option to reach the traffic archive and get appropriate information for analysis in case of a security incident.

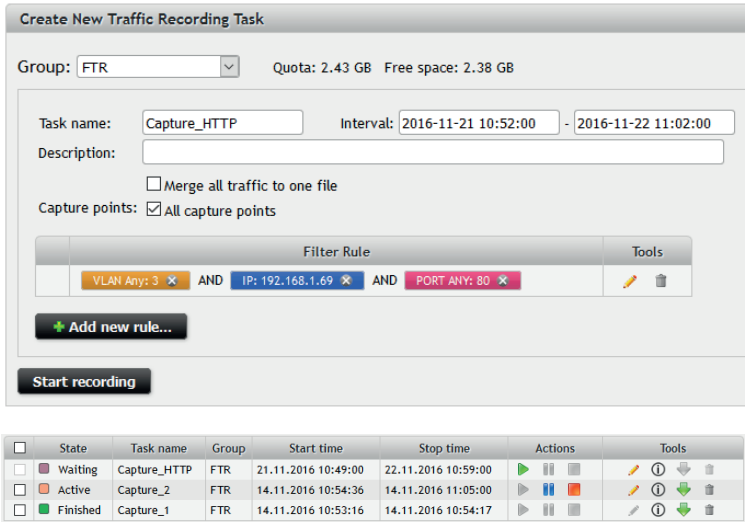

Figure 2: Creating traffic record in Flowmon Traffic Recorder GUI.

3 Packet Capture Tools

Two very well-known packet capture tools are tcpdump in the Linux operating system environment and WinPcap in the Microsoft Windows environment. Equipped with these tools, the network administrator usually arrives to a place with his notebook, connects the notebook to a mirror port or TAP and carries out the network traffic recording. Problems may arise in case of distant places, optical network interfaces or a 10Gpbs infrastructure – limitations that could hardly be overcome with a notebook.

Such situations can be solved by the installation of standalone probes across the network, which can be used as a platform for on-demand packet capture. By using the probes, the traffic can even be captured in high-speed networks (10Gbps or more) with different types of interfaces, and inquiries can be placed remotely. Professional probes for traffic recording often provide administrators with advanced packet filtering functions beyond L3 and L4 filter capabilities. As a result, network probes enable recording, i.e. of entire VoIP communication, or traffic based on the application protocol. We should also mention that full packet capture is also provided by modern firewalls – but they are typically restricted to perimeter network traffic.

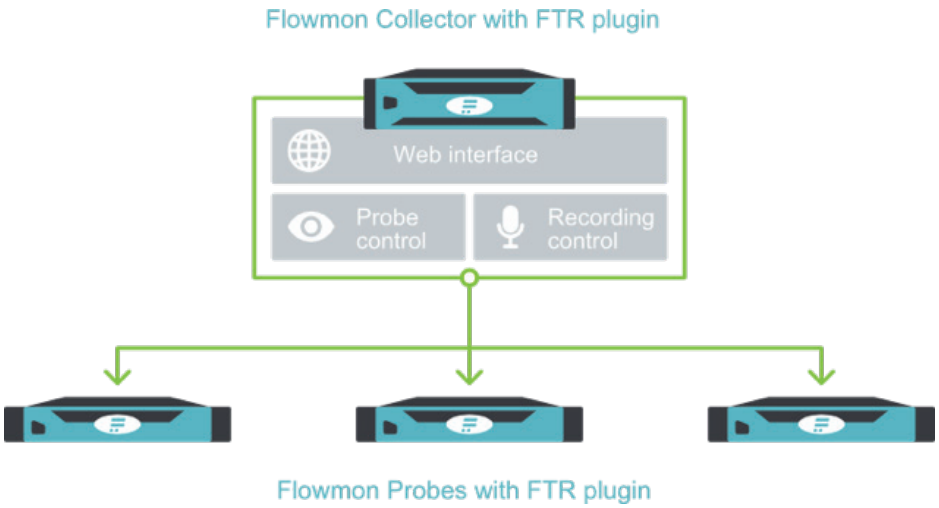

Figure 3: Capturing network traffic with standalone probes

4 Analysing the Traffic

Once we have the traffic stored in a file, we need an analytical tool. Which one to choose? There are plenty of commercial solutions available so you can decide which suits you best. Apart from those, there is a very popular open source tool called Wireshark. Wireshark is capable of recognising and decoding hundreds of protocols. Moreover, it is equipped with analytical functions for very deep traffic inspection such as filtering, reconstruction of TCP connections or phone calls, traffic decryption or data extraction. There are plenty of online courses, manuals and example data sets to be found on the Internet, so you can try these out on your own. On the other hand, Wireshark demands an advanced knowledge of TCP/IP protocols, data network principles and analytical skills from its user.

Let's take a look at a simple example of the benefits of packet analysis. A database administrator claims that everything is configured well. An administrator of the end-stations claims that application clients are properly installed and set up. Nevertheless, a user claims that the language character set is not displaying correctly and proves this by a screenshot. The truth lies inside the packets - packets don't lie. As a first step, it is necessary to set up a filter to capture communication between the clients and database server. Let's say that in our case the application is based on a MySQL database server running on default port 3306. The IP address of the client is 192.168.3.2. With this information in mind, we will set a filter for our traffic recording. Storing the traffic of a given client in full-scale and filtering particular packets afterwards is another option, but our approach allows us to downsize data volumes intended for analysis at the moment of recording.

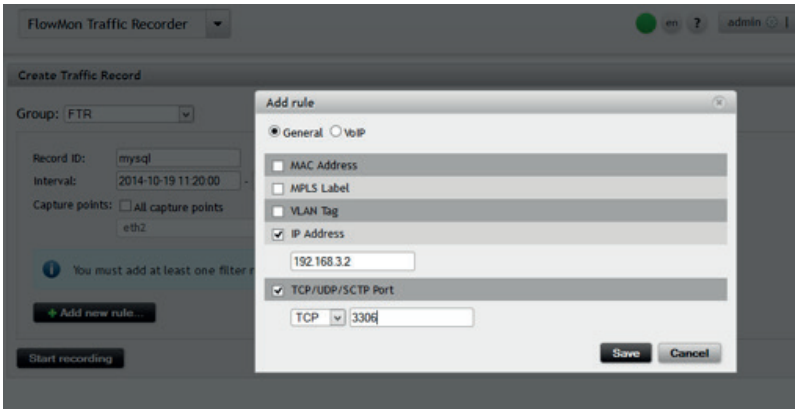

Figure 4: Setting up a filter to capture communication between one of the clients and MySQL server.

Our capture in PCAP format is then opened in Wireshark. We see the character set announced from the server to our client immediately. It should be a “Windows 1250” charset, but instead of it we get “LATIN2”. So it is proven that there is a problem with server configuration and that it is up to the administrator to make things right again.

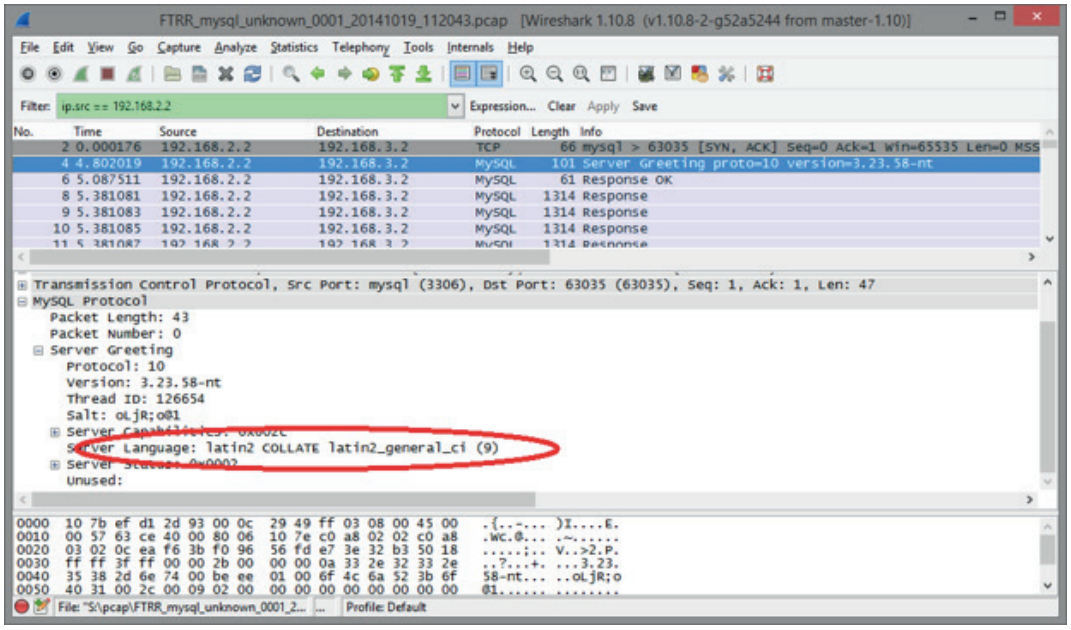

Figure 5: By using Wireshark we see the character set announced from the server immediately (highlighted in red).

5 Conclusion

In this simple example we have introduced capabilities for full packet capture and traffic analysis with an open source tool called Wireshark. Flow data-based monitoring provides answers to many questions and helps to detect the causes of network issues. Despite this, packet analysis is irreplaceable in situations when it is necessary to look inside the content of particular communications.

> When it comes to network traffic monitoring, analytics from Gartner estimate that flow analysis should be done 80 % of the time, and packet capture with probes should be done 20 % of the time.

The need for capturing packets isn't going away. However, with the improved insight provided by flow technologies the demand is certainly shrinking. Solutions for full-scale traffic recording and analysis are very expensive. Moreover, there are technology limits in a high-speed networking environment and restricted possibilities of utilisation when traffic is encrypted. Thus the roles have changed – complete flow-based infrastructure monitoring will be supported by on-demand deep packet analysis.

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