PACKETS ANALYSIS

By Leonardo Cantarella

INTRODUCTION

This project aims to give more informations about network packets that pass through a device.
Thanks to the used technologies, packets can be processed in real-time and can be enriched with other details.

TSHARK (“THE SNIFFER”)

The terminal version of Wireshark.

This is the base of the “stack”.

Packets are our source to be analyzed, so Tshark “sniffs” them from a network interface, and writes them in a file using a JSON format.

This is the raw data tha will be taken by ...

LOGSTASH (“THE PIPE”)

Logstash can read from that file and can make usable that data by using some basic filtering and ensuring that JSON packets arrives to ...

KAFKA (“THE BROKER”)

Kafka stores our packets and make possible the streaming of them.
He is the intermediate from the source to the elaboration made by ...

SPARK (“THE CORE”)

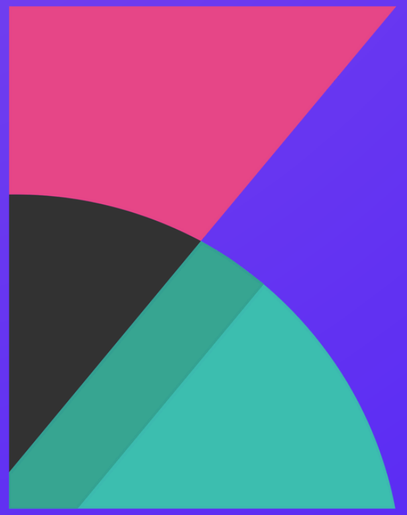
Spark is the core of the elaboration of the raw packets that arrives from tshark. Spark elaborates the packets by using web API and machine learnin model to enrich our data adding:

- Geolocation
- Application protocol detection (by port)
- Ip abuse
- Anomaly detection (using machine learning model)

ELASTICSEARCH (“THE INDEXER”)

Elasticsearch stores and indexes the enriched data processed by Spark. It allows for efficient querying and analysis of large datasets, supporting real-time search and analytics

KIBANA (“THE VIEWER”)



Allows to make graphics views in simple way, using a user-friendly interface, interacting with real-time data and helping to understand complex data in a more easy and understandable representation

PACKETS FLOW

THANK YOU!