Documentation

MeLi Challenge Dev

# Requirements

## Motivation

We need to find someone that is capable to make complex scenarios and analyze its behavior. For this purpose, we have several tools such as Virtual Computers and Traffic samplers. We know that sometimes the behavior is not the expected, so we want to analyze and make graphs of that to have a better understanding of it.

## Goal

Build a virtual scenario with at least 3 virtual computers and install at least 2 script in each to make some traffic between them. Install the traffic sampler [SFlow](https://github.com/sflow/sflowtool) or use TCPDump to monitor the activity in the network. Finally, with the output of SFlow or TCPDump make a program that with some algorithms make an analysis of the data and show the script patterns in different graphs and/or tables inside a static html.

# Project Architecture

## Overview

The project consists of four main modules:

* Virtual scenario
* Network traffic generator scripts
* Data collection script
* Data analysis program

### Virtual scenario

The virtual scenario consists of four virtual machines (terminals) which will send data to each other with the Network traffic generator scripts, and a central virtual machine in charge of collecting and processing the data generated by the terminals.

### Network traffic generator scripts

Each terminal has three scripts in charge of generating traffic over the network (only to the other three terminals, not to the central computer). Each script sends different type of data at a different rate.

### Data collection scripts

Each terminal is in charge of collecting the information of the incoming traffic, and saving it into a file.

The main computer has a script in charge of collecting the data generated by each terminal and merging it into a single file for the Data analysis program.

### Data analysis program

The Data analysis program receives a file containing information about all of the network traffic, processes it and generates an HTML file, with useful information about the network traffic.

## Used software

* OS: Ubuntu Server 14.04.4 LTS
* Virtualization Software: Virtual Box 4.3.28
* IDE: PyCharm Community Edition 5.0.4
* Programming Language: Python 3.4.3
* Utilities:
  + TCPDump
  + Mergecap (Wireshark Common package)
  + Pip3 (Python 3 Package Manager)
  + Yattag (Python Library)