****

软 件 工 程 实 验

软件产品设计与实现

基于Scapy的分析与扩展

Version 0.1

实 验 小 组 B组

小 组 成 员 陈鸿超（SY1806214）

李铎坤（SY1806219）

刘 颖（SY1806418）

袁梦阳（BY1806157）

编制时间： 2019年4月

版本变更历史

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 版本 | 提交日期 | 编制人 | 修改说明 | 审核人 | 版本说明 |
| v0.1 | 2019/4/17 | 袁梦阳 | 初版 |  | 初版的2、4.1、4.2 |

目录

[版本变更历史 I](#_Toc6246039)

[1 引言 1](#_Toc6246040)

[2 总体设计 2](#_Toc6246041)

[3 系统架构 3](#_Toc6246042)

[4 扩展功能设计 4](#_Toc6246043)

[5 实现方案 6](#_Toc6246044)

[6 集成方案 7](#_Toc6246045)

# 引言

# 总体设计

## 需求规定

本项目基于Scapy开源项目进行扩展，面向人群为计算机网络方面的开发与测试人员，主要目的是便于用户快捷地进行报文的获取、解析和发送等操作。为了满足此需要，本项目考虑实现一个更全面、快速的报文操作工具。通过对Scapy工具的扩展，在Scapy开源项目中扩展出开发人员需求但却缺少的功能，使得Scapy更加易用与全面。

## 硬件运行环境

1. CPU：Intel i5-4590及以上，AMD Ryzen 3 1200及以上
2. 内存：4G内存及以上（系统最低2GB，Scapy报文解析较占用内存空间）
3. 硬盘：20G硬盘及以上（Windows 7 64位安装最低配置）

## 软件运行环境

1. 操作系统：
   * 1. Windows 7及以上： 已安装Npcap/WinPcap，
     2. Ubuntu 16.04及以上： 已安装tcpdump
2. Anaconda（可创建多版本Python虚拟环境），Python 2.7+，Python 3.4+
3. 第三方Python库：matplotlib，pyx，vpython，cryptography等

## 编程语言与工具

1. 编程语言：Python
2. 编程工具：pycharm

# 系统架构

# 扩展功能设计

## pcap文件切分功能设计

pcap文件的报文使用字节码存储，所占用的存储空间很小。但对于解析后的报文数据而言，每条报文都包含大量对象、字段和方法，使得解析后的数据占用的存储空间是原始报文大小的数十倍。

由于解析pcap文件对内存造成极大的负担，而且内存占用率过高会影响程序的运行效率，那么，对于较大的pcap文件，将其切分成若干个小文件，再进行批处理非常必要。

本项目扩展了Scapy对大pcap文件进行切分的功能。开发人员只需提供所要切分的pcap文件路径、切分后的存储目录以及切分批大小即可按照默认设定简便地对pcap文件进行切分。同时开发人员还可自行设定以下切分过程中的需求：

1. 开发人员可以通过更改切分模式设置来设定切分模式。本项目的pcap文件切分功能提供两种切分模式供开发人员选择。第一种为兼容模式，调取文件解析模块解析数据后再进行切分，切分后可存储为与源文件不同的格式，但效率较低。另一种为极速模式，无需数据解析，对源数据进行切分，速度较快，但不可转换为与源文件不同的格式。默认为兼容模式。
2. 开发人员可以通过更改命名前缀设定来设置切分后的文件命名。默认情况下，新生成的文件命名方式为：源文件名\_00000~源文件名\_99999。当用户设定了命名前缀的情况下，新生成的文件命名方式为：前缀\_00000~前缀\_99999。
3. 开发人员可以设定报文数据的链路类型。本设置只有在切分模式为兼容模式时有效。
4. 开发人员可以设定切分后的报文的存储方式。开发人员可以选择报文字节码的存储方式是大端存储还是小端存储。本设置只有在兼容模式下游侠。

## 获取报文协议类型功能设计

Scapy根据报文的协议层次将报文数据分层解析，解析的结果也按照协议层堆叠起来。解析结果的每一层是一个单独的对象，拥有一个payload变量，用于保存下一层协议的对象实例。

比如一条DNS报文pkt，其解析结果是一个物理层的Ether对象实例，pkt.payload是一个网络层的IP对象实例，pkt.payload.payload是一个传输层的UDP对象实例，依次类推。

在处理多协议流量数据时，用户通常需要对不同协议类型的报文进行不同的处理操作。但用户无法直接从Scapy解析的结果中获取报文的协议类型，必须自行设计一种判定逻辑来得到报文的协议类型。因此，为解析后的报文对象内置一个获取协议类型的函数可以为用户提供极大的方便。

本项目扩展了Scapy获取报文协议类型的功能。开发人员可以直接通过本功能来获得报文的协议类型。用户可以选择两种报文协议类型的返回方式：一种是返回单层协议类型，另一种是返回所有层协议类型。

同时当开发人员选择的返回类型为单层时，可以选择一个的网络层，表示需要返回该层协议的类型。默认为返回可识别的最高层的协议。

# 实现方案

# 集成方案