

# Advanced Computer Networking Experiment Report

#### **Group Members:**

姓名 庄舒漫 学号 22\$15118 姓名 唐栋 学号 22\$151060

报告日期: 2022.06.09

## 1.实验概述

## 实验任务

只实现了 project1

write a program that can capture the coming packets and revert the webpage which you select to test by using the captured data.

## 实验步骤

Capture the packets

Reconstruct http message

Revert the webpage with captured data

## 开发平台

Os: win10

编译工具: mingw64,npcap,python3.9,cmake

可视化平台: QT5

## 2.实验过程

## 1.Npcap 配置

在 vscode 平台将相应的 npcap 动态链接库和函数头部文件放入工程文件夹使用语句

```
include_directories(Include)
include_directories(Lib/x64)
link_directories(Lib/x64)
enable_testing()
add_executable(NPCAP main.cpp)

target_link_libraries(NPCAP wpcap.lib Packet.lib)
target_link_libraries(NPCAP ws2_32.lib)
```

注意动态链接要在 add\_executable 语句后,而加载静态库 include/link 要在该语句之前

Make 后就会生成 win 平台的可执行文件

■ NPCAP.exe

## 2.根据 npcap 函数捕包

#### 1. pcap findalldevs

NpCap 提供了 pcap\_findalldevs\_ex 和 pcap\_findalldevs 函数来获取计算机上的网络接口设备的列表;此函数会为传入的 pcap\_if\_t 赋值——该类型是一个表示了设备列表的链表头;每一个这样的节点都包含了 name 和 description 域来描述设备。

除此之外,pcap\_if\_t 结构体还包含了一个 pcap\_addr 结构体;后者包含了一个地址列表、一个掩码列表、一个广播地址列表和一个目的地址的列表;此外,pcap\_findalldevs\_ex 还能返回远程适配器信息和一个位于所给的本地文件夹的 pcap 文件列表。

#### 2. pcap\_open

用来打开一个适配器,实际调用的是 pcap\_open live; 它接受五个参数:

name: 适配器的名称(GUID)

snaplen:制定要捕获数据包中的哪些部分。在一些操作系统中(比如 xBSD 和 Win32),驱动可以被配置成只捕获数据包的初始化部分:这样可以减少应用程序间复制数据的量,从而提高捕获效率;本次实验中,将值定为 65535,比能遇到的最大的 MTU 还要大,因此总能收到完整的数据包。

flags: 主要的意义是其中包含的混杂模式开关; 一般情况下,适配器只接收发给它自己的数据包, 而那些在其他机器之间通讯的数据包,将会被丢弃。但混杂模式将会捕获所有的数据包——因为我们需要捕获其他适配器的数据包,所以需要打开这个开关。

to\_ms: 指定读取数据的超时时间,以毫秒计; 在适配器上使用其他 API 进行读取操作的时候,这些函数会在这里设定的时间内响应——即使没有数据 包或者捕获失败了; 在统计模式下, to\_ms 还可以用来定义统计的时间间隔: 设置为 0 说明没有超时——如果没有数据包到达,则永远不返回; 对应的还有-1: 读操作立刻返回。

errbuf: 用于存储错误信息字符串的缓冲区 该函数返回一个 pcap t 类型的 handle。

#### 3. pcap loop

API 函数 pcap\_loop 和 pcap\_dispatch 都用来在打开的适配器中捕获数据包;但是前者会已知捕获直到捕获到的数据包数量达到要求数量,而后者在到达了前面 API 设定的超时时间之后就会返回(尽管这得不到保证);前者会在一小段时间内阻塞网络的应用,故一般项目都会使用后者作为读取数据包的函数;虽然在本次实验中,使用前者就够了。

这两个函数都有一个回调函数;这个回调函数会在这两个函数捕获到数据包的时候被调用,用来处理捕获到的数据包;这个回调函数需要遵项特定的格式。但是需要注意的是我们无法发现 CRC 冗余校验码——因为帧到达适配器之后,会经过校验确认的过程;这个过程成功,则适配器会删除 CRC;否则,大多数适配器会删除整个包,因此无法被 NpCap 确认到。

#### 4. pcap\_datalink

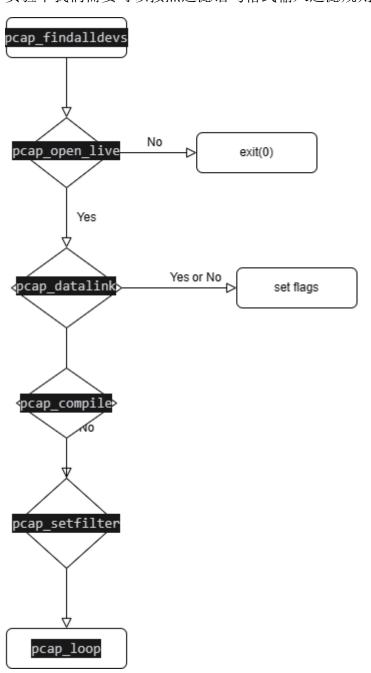
用于对 MAC 层进行了检测,以确保在处理一个以太网络,确保 MAC 首

部是 14 位的。IP 数据包的首部就位于 MAC 首部的后面,将从 IP 数据包的首部解析到源 IP 地址和目的 IP 地址。

本次实验主要抓取了以太网的包和本地回环网络的数据包,由于这两种包包头格式不一样,所以必须用该函数区分,设置标志

#### 5. pcap\_compile & pcap\_setfilter

用来设置过滤器,以避免处理一些无用的包,提高包处理的效率。在本次 实验中我们需要可以按照过滤语句格式输入过滤规则



## 3.定义网络层,传输层,应用层协议头部信息

报文定义如下所示:

```
typedef struct ip_header
       u_char ver_ihl;
      u_char ver_ihl; // Version (4 bits
u_char tos; // Type of service
u_short tlen; // Total Length
u_short identification; // Identification
      u_short identification; // Identification
u_short flags_fo; // Flags (3 bits) + Fragment offset (13 bits)
u_char ttl; // Time to Live
u_char proto; // Protocol
u_short crc; // Header checksum
ip_address saddr; // Source address
ip_address daddr; // Destination address
u_int op_pad; // Option + Padding
}ip_header;
typedef struct udp_header
       u_short sport; // Source port
u_short dport; // Destination port
u_short len; // Datagram Length
      u_short len; // Datagram
u_short crc; // Checksum
}udp_header;
typedef struct tcp_header
       unsigned short src_port; //源端口号
unsigned short dst_port; //目的端口号
unsigned int seq_no; //序列号
unsigned int ack no: //确以号
       unsigned int ack_no;
       unsigned char reserved_1; //保留6位中的4位首部长度
       unsigned char reseverd_2; //保留6位中的2位
       unsigned short wnd_size; //16位窗口大小
unsigned short chk_sum; //16位TCP检验和
unsigned short urgt_p; //16为紧急指针
}tcp_header;
```

#### 这里需要注意的要点:

- 1.超过一个字节字段定义要注意大小端,大小端是指字节排序方式,大端高位数据在低地址,小端低位数据在低地址,要注意使用 windows 特有函数(不是 mingw 编译器中带有),ntohs(短型), ntohl(长型整数)(win\_32),在一个字节内只要按位取需要的值即可
- 2. 关于数据报帧头,以太网是14字节,本地回环包是4字节。
- 3. 以太网最小帧大小是 60, 小于这个会用 0x00padding
- 4.以太网最大数据包长度是 1514 = 54 + 1460 (数据大小)

## 4.解析 http 数据包内容

```
①请求方法
②请求URL
③HTTP协议及版本

POST /chapter17/user.html HTTP/1.1
Accept: image/jpeg, application/x-ms-application, ..., */*
Referer: http://localhost:8088/chapter17/user/register.html?
code=100&time=123123
Accept-Language: zh-CN
User-Agent: Mozilla/4.0 (compatible; MSIE 8.0; Windows NT 6.1; Content-Type: application/x-www-form-urlencoded Host: localhost:8088
Content-Length: 112
Connection: Keep-Alive Cache-Control: no-cache Cookie: JSESSIONID=24DF2688E37EE4F66D9669D2542AC17B

name=tom&password=1234&realName=tomson
```

图 15-4 HTTP 请求报文

//blog.csdn.net/ulike MFY



需要注意的头部信息:

content\_type: 关心是一个 html/txt 文本页面或者是 png,jpeg 图片格式,或者是.css 网页渲染器

Content\_encoding:现在 http 网页为了加快响应速度,基本都会采用 gzip 压缩网页内容传输,所以需要判断是否为 gzip 压缩,并进行解压

Charset:主要为了后面解压后用什么编码二进制串 Content\_length:为了判断是否获取数据包全部内容 示例:

这里我抓取的是使用 ipv4 协议的 http 协议的一个网站,网址为 <u>www.weislank.com</u> 抓取 http 相应报文内容如下图所示:

```
> HTTP/1.1 200 OK\r\n
 Date: Sun, 18 Jun 2023 04:19:08 GMT\r\n
  Server: Apache/2\r\n
 Upgrade: h2,h2c\r\n
 Connection: Upgrade, Keep-Alive\r\
 Vary: Accept-Encoding, User-Agent\r\n
 Content-Encoding: gzip\r\n
> Content-Length: 7283\r\n
 Keep-Alive: timeout=2, max=100\r\n
 Content-Type: text/html; charset=utf-8\r
 \r\n
  [HTTP response 1/1]
  [Time since request: 0.104551000 seconds]
 [Request in frame: 1146]
 [Request URI: http://weislank.com/]
 Content-encoded entity body (gzip): 7283 bytes -> 30740 bytes
 File Data: 30740 bytes
```

可以看到这里使用 gzip 压缩,编码格式是 utf-8,数据包内容一共有 7283 个字节 这是我重组帧之后得到的数据内容,用 nodepad++打开:



此时 length 大小确实为 7283

使用 gzip 解压后,并使用 utf-8 编码,得到 html 文件如下所示:

## 4.重组数据报

虽然 http 数据报内容显然无法用一个窗口发送完毕,但我发现服务器并没有对 ip 数据报分片,而是采用了不同 tcp 包+fin 片标识来发送整个数据包 所以,程序实现逻辑如下:

1. 找到第一个相应数据包 以 HTTP 开头的数据包,去掉 http 协议头部信息,将数据部分内容写入二进制文件中

```
00 ed 06 3e 00 00 48 54
                          54 50 2f 31 2e 31 20 32
                                                       ···>··HT TP/1.1 2
30 30 20 4f 4b 0d 0a 44
                           61
                              74
                                 65 3a 20 53 75
                                                       00 OK⋅⋅D ate: Sun
2c 20 31 38 20 4a 75 6e
                           20 32 30 32 33 20 30
                                                        18 Jun
                                                                2023 04
3a 31 39 3a 30 34 20 47
                          4d 54 0d 0a 53 65 72 76
                                                       :19:04 G MT · · Serv
65 72 3a 20 41 70 61 63
                          68 65 2f 32 0d 0a 55 70
                                                       er: Apac he/2⋅⋅Up
67 72 61 64 65
                3a 20 68
                          32 2c 68 32 63 0d 0a 43
                                                       grade: h 2,h2c⋅⋅C
6f 6e 6e 65 63 74 69 6f
                          6e 3a 20 55 70 67 72 61
                                                       onnectio n: Upgra
64 65 2c 20 4b 65 65 70
                          2d 41
                                 6c 69 76 65 0d 0a
                                                       de, Keep -Alive·
56 61 72 79 3a 20 41 63
                          63 65 70 74 2d 45 6e 63
                                                       Vary: Ac cept-Enc
6f 64 69 6e 67 2c 55 73
                          65 72 2d 41 67 65 6e 74
                                                       oding,Us er-Agent
0d 0a 43 6f 6e 74 65 6e
                          74 2d 45 6e 63 6f 64 69

· · Conten t-Encodi

6e 67 3a 20 67 7a 69 70
                          0d 0a 43 6f 6e 74 65 6e
                                                       ng: gzip ∙∙Conten
74 2d 4c 65 6e 67 74 68
                          3a 20 37 32 38 33 0d 0a
                                                       t-Length : 7283 ·
4b 65 65 70 2d 41 6c 69
                          76 65 3a 20 74
                                          69 6d 65
                                                       Keep-Ali ve: time
6f 75 74 3d 32 2c 20 6d
                              78
                                 3d 31 30
                                                       out=2, m ax=100⋅⋅
                          61
                                          30 0d 0a
43 6f 6e 74 65 6e 74 2d
                          54 79 70 65 3a 20 74 65
                                                       Content- Type: te
78 74 2f 68 74 6d 6c 3b
                          20 63 68 61 72 73 65
                                                       xt/html;
                                                                charset
3d 75 74 66 2d 38 0d 0a
                          0d 0a 1f 8b 08 00 00 00
                                                       =utf-8··
00 00 00 03 ed 3d 6b 53
                          1b 47
                                b6 df 53 b5 ff
                                                       ····-=kS ·G··S··a
a2 a4 16 48 d0
               63 f4 00
                              00
                                 fb 48 f6
                                          de 9b
                                                       · · · H · c · · · · · · H · · · ·
                           e1
ad 4d ee 8d 77 f7 de eb
                          f5 a6 e6 d1 82 89 25 8d
                                                       ·M··w··· ....%
56 33 32 76 b2 54 09 13
                          0c d8 bc 6c 63 fc 00 02
                                                       V32v·T·· ···1c···
d8 c6 b0 4e 0c 7e 9b a7
                          f9 31 ab 1e 49 9f f2 17
                                                       · · · N · ~ · · · 1 · · I · ·
ee e9 e9 91 34 4f 49 80
                          13 ef
                                                       ····40I·
                                dd 8d ca 65 66 7a
4e 9f 3e 7d 5e 7d fa 4c
                                 fb ed 8f
                                                       N->}^}-L ---->-
                          eb
                              a8
                                           3e f9 f0
e4 ff 7c fa 1b a6 5f 4d
                          25 7b 7f f6 56 b7 f1 97
                                                       ..|... M %{..V...
                          21 95 63 84 7e 2e ab 20
5c 21 4e 84 2b 86 e9 4e
                                                       \!N·+··N !·c·~.·
b5 c7 f7 fb 93 ff e6 8f
                          fb 08 ac de de af
                                                19
                                             aa
                                                       . . . . . . . . . . . . . . . .
3f fa 4b 4e 3a db e3 fb
                          6f
                             ff ef 7f e5 ff 50 4e
                                                       ? · KN: · · · · o · · · · · PN
65 38 55 e2 93
                          72 5a 45 69 e8 24 a1 1e
               c8 c7 08
                                                       e8U····· rZEi·$∙
                          d4 8b d7 87 b4 b9 4d 3c
24 f6 21 bd 9b 2a a9 49
                                                       $.!..*.I .....M<
                                                       ..0.... ·vq·J··|
f2 08 4f ae e0 8d 9b da
                          dc 76 71 ff 4a f1 db 5b
dd 41 fa bc 82 3f cd a5
                          50 8f ef 0c 3a
                                          3f 20 67
                                                       ·A···?·· P···:?
                          55 44 8a 90 95 32 aa
45 c5 84 d2 0b 81 cf d6
                                                24
                                                          ····· UD····2·$
a7 3d 7a 7f bf 37 36 80
                           24 25 c9 a5 cf
                                          7c bf
                                                 37
                                                       ·=z··76· $%···|·7
ae dd dc 2c 6c 3d 28 ec
                           ac e1 6b 43 a5 83 85
                                                       ···,1=(· ··kC···
                                                      1 · · · · √8 · · · n ·
                          de 5c 38 14 c2 f7 6e c3
6c 1e cf ef e3 8b 23 f8
```

2. 判断接下来的序列号是否 == 上一个包序列号+数据大小(1460)

```
if(seq_pre!=0&& seq_no!= seq_pre+len_pre){
    if(remove("webpage_new.dat")==0) cout<<"remove success!"<<endl;
    printf("wrong sequence!\n");
    exit(0);
    // exit(0);
}</pre>
```

3. 判断 fin 标识是否为真,若为真,且数据大小==报文大小,退出程序

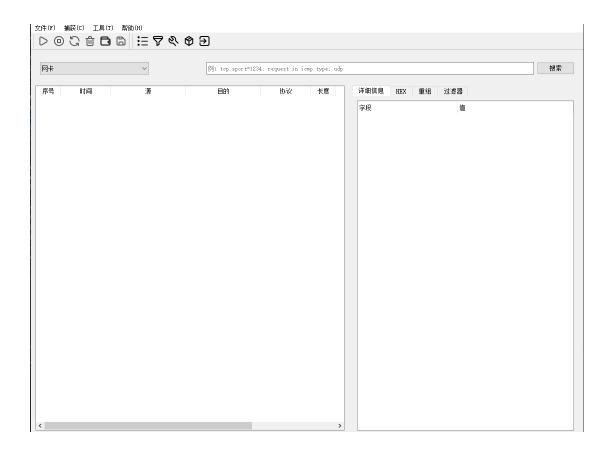
```
if(ack&&fin&&flag_pre){
    // deal_file();
    exit(0);
}
if(header->len <= 60) return;</pre>
```

#### 最终重组包信息为:

```
data len: 1460
HTTP/1.1 200 OK
Date: Sun, 18 Jun 2023 04:38:40 GMT
Server: Apache/2
Upgrade: h2,h2c
Connection: Upgrade, Keep-Alive
Vary: Accept-Encoding, User-Agent
Content-Encoding: gzip
Content-Length: 7283
Keep-Alive: timeout=2, max=100
Content-Type: text/html; charset=utf-8120.48.171.216:80 -> 10.249.43.35:50395
12:38:39.645125 len:1514
head len: 54
data len: 1460
120.48.171.216:80 -> 10.249.43.35:50395
12:38:39.645317 len:1514
head_len: 54
data_len: 1460
120.48.171.216:80 -> 10.249.43.35:50395
12:38:39.645317 len:1514
head len: 54
data len: 1460
120.48.171.216:80 -> 10.249.43.35:50395
12:38:39.645317 len:1514
head_len: 54
data_len: 1460
120.48.171.216:80 -> 10.249.43.35:50395
12:38:39.645317 len:313
head len: 54
data_len: 259
120.48.171.216:80 -> 10.249.43.35:50395
12:38:41.645864 len:60
```

## 5.可视化界面

可视化使用 python 编写,实现起来较为简单快捷,主要使用了 pyqt5 窗口如下所示:



## 实验结果

## 1.实验条件

## 1.1 运行条件:

访问公有网站: www.weislank.com

网卡: 以太网卡

过滤语句: src host 120.48.171.216

访问自己搭建网页:

```
WARNING: This is a development server. Do not use it in a production deployment. Use a production WSGI server instead.

* Running on all addresses (0.0.0.0)

* Running on <a href="http://127.0.0.1:8083">http://127.0.0.1:8083</a>

* Running on <a href="http://10.249.43.35:8083">http://10.249.43.35:8083</a>
```

网卡:本地回环抓包网卡 过滤语句: src port 8083

### 1.2 项目结构:

分为 c++文件和 python 文件 C++文件功能为捕包,重组包,分析网页内容

具体运行方法:

运行: C:\Users\18399\Desktop\project1\npcap\build 路径下的 NPCAP.exe 文件即可,注意网卡选择和过滤语句编写

Python 文件主要为可视化,页面搭载,以及数据包解压

具体运行方法:

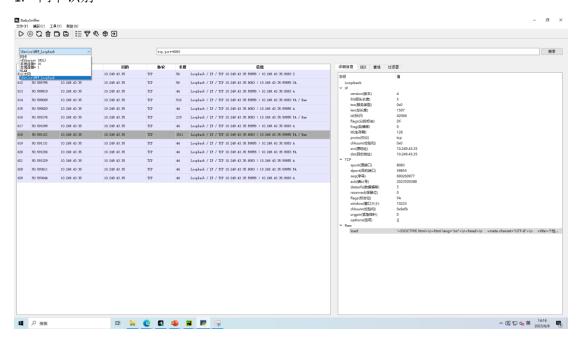
运行 C:\Users\18399\Desktop\project1\Sniffer-main\Sniffer-main\source 路径下的 main.py 即可实现可视化

运行 C:\Users\18399\Desktop\project1\Sniffer-main\Sniffer-main\utils 路径下的 webpage.py 即可搭建简单页面

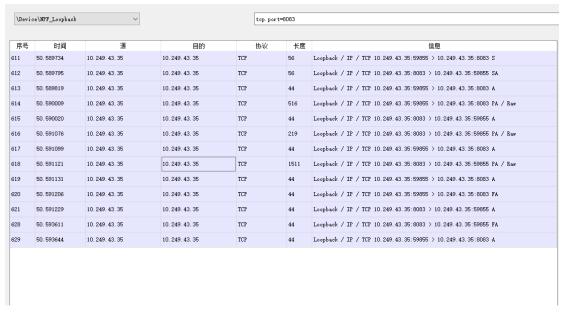
运行 C:\Users\18399\Desktop\project1\Sniffer-main\Sniffer-main\utils 路径下的 decode.py 对 二进制文件解压

## 2.可视化界面实现功能

#### 1. 网卡识别



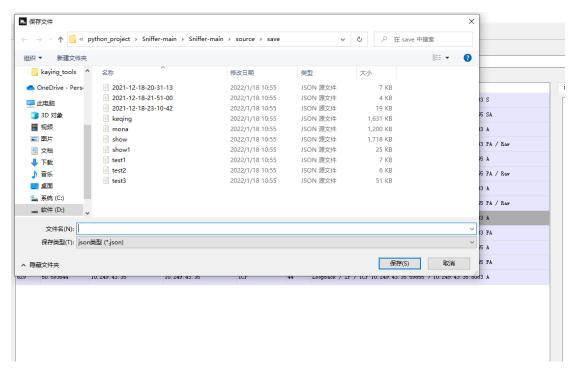
#### 2. 包过滤



#### 3. 协议头部解析



#### 4.保存包内容



## 3.还原网页

这一部分使用 c++编写 具体过程:

1. 运行结果

若抓取 www.weislank.com

NPCAP.exe 会生成一个二进制文件———》 webpage\_new.dat

将其复制到 decode.py 文件夹下,使用 decode.py 进行解压,编码后得到 html 网页-> Webpage.html

若抓取本地网页

直接在 NPCAP 路径下生成 webpage.html,打开即可

```
data_len: 1460
HTTP/1.1 200 OK
Date: Sun, 18 Jun 2023 04:38:40 GMT
Server: Apache/2
Upgrade: h2,h2c
Connection: Upgrade, Keep-Alive
Vary: Accept-Encoding, User-Agent
Content-Encoding: gzip
Content-Length: 7283
Keep-Alive: timeout=2, max=100
Content-Type: text/html; charset=utf-8120.48.171.216:80 -> 10.249.43.35:50395
12:38:39.645125 len:1514
head len: 54
data_len: 1460
120.48.171.216:80 -> 10.249.43.35:50395
12:38:39.645317 len:1514
head_len: 54
data_len: 1460
120.48.171.216:80 -> 10.249.43.35:50395
12:38:39.645317 len:1514
head_len: 54
data_len: 1460
120.48.171.216:80 -> 10.249.43.35:50395
12:38:39.645317 len:1514
head_len: 54
data_len: 1460
120.48.171.216:80 -> 10.249.43.35:50395
12:38:39.645317 len:313
head_len: 54
data_len: 259
120.48.171.216:80 -> 10.249.43.35:50395
12:38:41.645864 len:60
PS D:\cpp_project\npcap\build>

✓ build

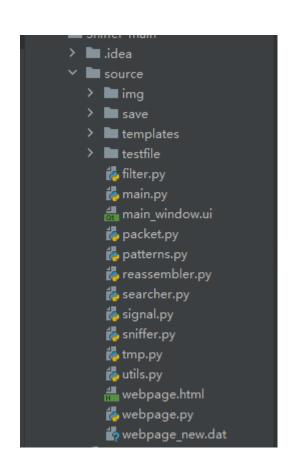
    > .cmake
    > CMakeFiles
    > Testing
   {} compile_commands.json
   ■ DartConfiguration.tcl
   file_deal.cpp

    file_deal.exe

   M Makefile
   ■ NPCAP.exe

■ webpage_new.dat
```

webpage.html



Address	0	1	2	3	4	5	6	7	8	9	a	b	С	d	е	f	Dump
00000000	1f	8b	08	00	00	00	00	00	00	03	ed	3d	6b	53	1b	47	.??kS.G□□
00000010	b6	df	53	b5	ff	61	a2	a4	16	48	d0	63	f4	00	e1	00	哆S?aiv.H衏??□□□□
00000020	fb	48	f6	de	9b	aa	ad	4d	ee	8d	77	f7	de	eb	f5	аб	鸋鲛洩璏?w鬂膈?□□
00000030	е6	d1	82	89	25	8d	56	33	32	76	b2	54	09	13	0c	d8	嫜倝%.V32v睺?[
00000040	bc	6с	63	fc	00	02	d8	С6	b0	4e	0c	7e	9b	a7	f9	31	糽c?.仄癗.~洤?□□□
00000050	ab	1e	49	9f	f2	17	ee	е9	е9	91	34	4f	49	80	13	ef	?I燆.铋閼40I€.?□[
00000060	dd	8d	ca	65	66	7a	4e	9f	Зе	7d	5e	7d	fa	4c	eb	<b>a</b> 8	?蔱fzN?}^}鶯毹□□[
00000070																	
00000080	7f	f6	56	b7	f1	97	5c	21	4e	84	2b	86	е9	4e	21	95	.鯲否梊!N?嗛N!?□□
00000090	63	84	7e	2e	ab	20	b5	с7	f7	fb	93	ff	е6	8f	fb	08	c剘.?登鼷???□□□□
000000a0	ac	de	de	af	aa	19	3f	fa	4b	4e	3a	db	e3	fb	6f	ff	薤??鶮N:坫鹢 [
000000b0	ef	7f	e5	ff	50	4e	65	38	55	e2	93	<b>c</b> 8	c7	80	72	5a	??PNe8U鈸惹.rZ□□[
																	Ei??\$?經*區詪□□□□
000000d0	d7	87	b4	b9	4d	3с	f2	08	4f	ae	e0	8d	9b	da	dc	76	讎垂M O .涄躹□</td
000000e0	71	ff	4a	f1	db	5b	dd	41	fa	bc	82	3f	cd	a5	50	8f	q J褊[軦 ?庭P.[
000000f0	ef	0c	3a	3f	20	67	45	c5	84	d2	0b	81	cf	d6	55	44	?:? gE艅?.现UD□□[
																	????z.?6€\$%□□□□□
00000110																	
																	kC 咈1.巷銒# \[
																	8.瞒n肧也瘼x鹝?□□
																	??s.\.煩~o uM□□[
																	{篰貦*11_+?梀□
																	.?E逖砚靭a飂?□□
																	??丹−蝖ѕ肤精隍□□□
																	^橘&捱掌?#罃GY□□□
																	錛 <z戶瓦 c泡[<="" td="" 来.=""></z戶瓦>
																	囆犯<嫥痟棶銘Umq[
																	蛋54億mz.O庫??□□□
																	?~第.?沮靨ⅰ 搾□□□
000001d0																	
																	₩緝瀟<€?遠盟?□□
																	R貆.4處/.汩.U^驭[
00000200	77	30	83	2f	ad	17	с7	1e	e3	d5	fd	52	7e	84	4c	69	w0???阏齊~凩i□□□□

## 🕏 威斯兰卡官方网站 威斯兰卡官方网站

- 网站首页

- 公司资讯联系我们

banner banner banner

防伪查询

product

#### 产品展示

乒乓球拍

乒乓球拍

**尼<u>乒乓球拍</u>** 

乒乓球拍

乒乓球拍

<u> 乒乓球拍</u>

