MASARYK UNIVERSITY
FACULTY OF INFORMATICS

Analysis and detection of Skype network traffic

DIPLOMA THESIS

Luboš Ptáček

Brno, autumn 2011

Declaration

I declare that this thesis is my own work and has not been submitted in any form for another degree or diploma at any university or other institution of tertiary education. Information derived from the published or unpublished work of others has been acknowledged in the text and a list of references is given.

Advisor: RNDr. Jan Vykopal

Acknowledgement

I would like to express my gratitude to RNDr. Jan Vykopal for supervising my work and thank Jan Hapala for advices.

Abstract

This thesis deals with traffic identification of the Skype application. Payload and flow based analysis of the standby traffic and voice calls is done. Skype flow patterns are used to create a plugin for NfSen to detect UDP voice calls in the network.

Keywords

Skype, voice calls, UDP, traffic, payload, analysis, flow, NetFlow, NfSen

Contents

Chapter 1

Introduction

Skype [20, 25] is a software application that allows users to start and receive voice and video calls and to chat over the Internet. Also voice or video conference, and file transfers are supported. Skype also offers application programming interface which can be used by third party applications to initiate communication in the Skype network. The Skype network is a P2P VoIP network.

Success of the application comes from user friendly operation as it can operate without manual user configuration. This user friendliness is largely due to ability to detect the current network configuration and use of mechanisms to circumvent many applied network restrictions.

Skype utilizes encryption to provide secure communication inside the whole Skype network. It also uses several techniques to conceal the traffic. This results to the fact that traditional port based or payload based methods of identification of Skype traffic cannot be applied.

Network administrators or operators are usually interested in the nature of traffic transfered through the network in order to optimize network performance, forecast future development or to have the traffic under control. Skype usage is also nowadays under the supervision of mobile operators. Amount of mobile phones with available WiFi or 3G connectivity increases and users have the option to choose between traditional phone calls and available VoIP services.

Unregulated Skype usage by employees for leisure and private purposes can lead to economic loss. Skype can often circumvent network restrictions like NAT and firewall by traversing them. Therefore enterprises are seeking solutions to regulate Skype activities over their networks. Skype establishes many concurrent connections in a rapid manner, which can be recognized as undesirable behaviour.

Skype traffic detection can be categorized into payload-oriented and into nonpayload-oriented approaches. One method based on Pearson's Chi Square test is used to detect Skype's fingerprints from the packet framing structure, exploiting the fact whether payload bytes are truly random or

not, a condition that has to be true for encrypted traffic [5]. Another method to classify Skype VoIP traffic is based on Naive Bayes Classifier using packet arrival rate and packet length [5]. There can also be used some machine learning algorithms like AdaBoost or C4.5 [3]. Machine learning involves the construction of a classifier that uses characteristics of the sequence of data packets to identify its class.

The limitations of port and payload based analysis have motivated the use of flow (sequence of packets) statistics for traffic classification. These techniques rely on the observation that different applications have distinct behavior patterns on the network. The goal of this thesis is the analysis of Skype traffic on the payload level and on the flow level, and subsequently to propose Skype detection method that utilizes flows. NetFlow protocol [6] reports aggregated information about traffic traversing the routers in the form of flow-records. While this kind of data is already effectively used for accounting, monitoring and anomaly detection, the limited amount of information it conveys has until now hindered its employment for traffic classification purposes.

Chapter 2 serves as an overview of the Skype history. Chapter 3 provides information about entities that we can distinguish in the Skype communication framework, and about application key components. Chapter 4 describes Skype application stages from the startup to the termination of the application. Also results of Skype traffic analysis will be presented. Method that detects voice calls is proposed based on measurements. NetFlow protocol, NfSen and NFDUMP tools are described in Chapter 5. Discussion related to the practical implementation of the proposed detection method is then given in Chapter 6. Plugin for NfSen was created and its backend and frontend part is presented here. Test calls were performed to test the plugin implementation and results are described.

Chapter 2

Skype history

Skype was founded in 2003 by Niklas Zennström from Sweden and Janus Friis from Denmark. The Skype software was developed by Estonians Ahti Heinla, Priit Kasesalu and Jaan Tallinn [13], who were also behind the peer-to-peer file sharing software Kazaa [23]. In April 2003, Skype.com and Skype.net domain names were registered. In August 2003, the first public beta version was released.

One of the initial names for the project was "Sky peer-to-peer", which was then abbreviated to "Skyper". However, some of the domain names associated with "Skyper" were already taken. Dropping the final letter left the current title "Skype", for which domain names were available. Calls to other users within the Skype service are free, while calls to both traditional landline telephones and mobile phones can be made for a fee using a debit-based user account system. Skype has also become popular for its additional features which include instant messaging, file transfer, and video conferencing. Skype has 663 million registered users as of 2010 [15]. The average number of users connected each month was 145 million in the fourth quarter of 2010, versus 105 million a year earlier, while paying customers rose over the same period to an average 8.8 million per month, from 7.3 million. Skype reached a record with 30 million simultaneous online users on 28 March 2011 [14]. The network is operated by Microsoft Skype Division, which has its headquarters in Luxembourg. Most of the development team [13] and 44% of the overall employees of Skype are situated in the offices of Tallinn and Tartu, Estonia. eBay acquired Skype Limited in September 2005 and in April 2009 announced plans to spin it off through an initial public offering in 2010. It was acquired by Silver Lake Partners in 2009. Microsoft agreed to purchase Skype for $8.5 billion on May 2011 and the company is to be incorporated as a division of Microsoft called Microsoft Skype Division.

Chapter 3

Skype

Following chapter will describe Skype entities and key components of the application.

3.1 Skype entities

According to [10], we can distinguish some entities in the communication framework.

Skype client (SC) End client which places voice calls. Each *SC* maintains a record *host cache* which contains IP addresses and port numbers of *super nodes*. Notation SC_A and SC_B denotes the *SC* of the Caller and Callee respectively.

Supernode (SN) Online nodes that maintain the skype overlaying network. As described in [7], a supernode performs routing tasks such as forwarding requests to appropriate destinations and answer to queries from other *SC*s or *SN*s. A supernode can also forward login requests in case the login server is not directly reachable from a *SC*. Any *SC* with a public IP address can be promoted to an *SN* without the awareness of the *SC* host. This behaviour can be switched off by changing the registry entries.

Skype HTTP Server (HS) The HTTP server of *ui.skype.com*.

Login server (LS) An *SN* that Skype uses to provide authentication services to *SC*s.

Neighbour supernode (NSN) *SN*s that are logically near to an *SC*. An *SC* must establish connection to some *SN*s for Skype communications. The *SC* locates and then binds to its *NSN* for such purpose. An *SN* can be the *NSN* of multiple *SC*s simultaneously.

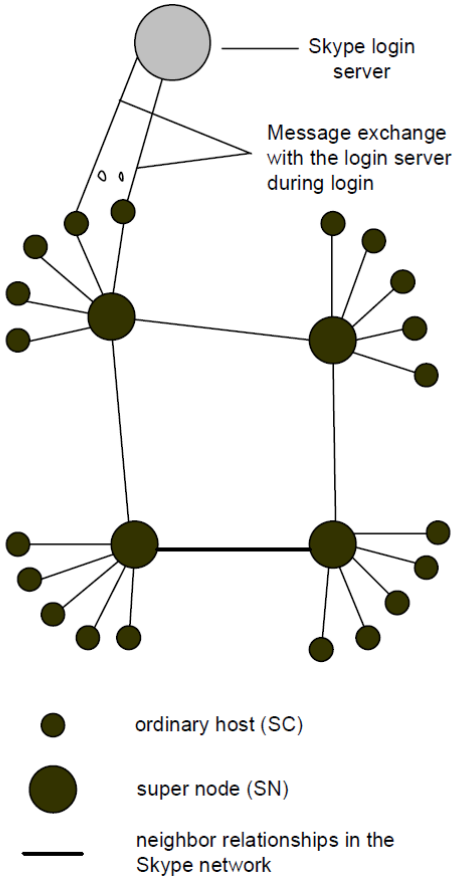


Figure 3.1: Basic Skype network, figure from [4]

3.2 Key components

Port numbers that Skype client utilizes for its communication in the Skype network and IP addresses of supernodes are important information that is stored by the application.

3.2.1 Ports

Skype uses UDP and TCP protocols for communication. *SC* has under common conditions three listening ports enabled. They are configured in the *Connection* dialog box. The first one is randomly chosen during the application installation and is higher then 1024 according to [16]. Then there are open listening ports 80 and 443 which can be disabled in the dialog box. *SC* is reachable by the first port number for UDP traffic. For TCP connection, Skype opens randomly chosen ports higher than 1024 to contact other hosts. If there are network restrictions applied, it tries again to this host on port 443. If this attempt fails, it will try on port 80. Connection on port 80 was seldom detected. This can be explained in a way that the *SC* tries concurrently to initiate connections with other hosts and there will be a successful connection earlier than the need of making an attempt on port 80. So the ports 443 and 80 serve as a fallback precaution.

3.2.2 Host cache

Skype network is an overlay network and thus each *SC* maintains a table of reachable supernodes. It is called *host cache (HC)* and is stored in the XML file "shared.xml" (in the case of Windows 7, it is stored in `C:\Users\<user name>\AppData\Roaming\Skype\`). The *HC* contains a maximum of 200 entries. The file *shared.xml* contains element `<HostCache>` with a hex string (`<HostCache>` with 200 entries has around 12800 characters). An example of the `<HostCache>` element follows. Part of the string is divided here in several separate substrings that store some information – description also follows.

```
<HostCache>
  41C80105004105020059A9FF2F9DC20001040002DCCDF0DE
  040003DCCDF0DE04000400050041050200
  ...
  04000400050041050200
  D5F0C7F6DE02
  0001010002
```

```
  A2A4E5DE
  040003
  D6E2E5DE
  04000400050041050200
  ...
  040004000500410502005D7B32BFF5930001040002 80CCF0DE
  04000380CCF0DE04000400
</HostCache>
```

From the knowledge that the string contains list of supernodes IP addresses and their corresponding ports, the string was analyzed and searched for some delimiters. The main delimiter is the substring `04000400050041050200`. It separates the list of supernodes and after it there are 12 characters. Decoding them as IP address and port number we can obtain address 213.240.199.246 and port 56834. Substring `0001010002` is the next delimiter (the middle character with value `1` can have different values). Substring `040003` separates two substrings – `A2A4E5DE` and `D6E2E5DE` – which have sometimes equal value. Values of these two substrings differ for each supernode, but their purpose is unknown.

3.2.3 Encryption

All Skype-to-Skype voice, video and instant message conversations are encrypted as described in [20]. Skype uses the Advanced Encryption Standard, also known as Rijndael, which is used by the US Government to protect sensitive information, and Skype uses the maximum 256-bit encryption. User public keys are certified by the Skype server at login using 1536 or 2048-bit RSA certificates.

Chapter 4

Skype traffic analysis

Following chapter will describe Skype application stages from the startup to the termination of the application. Also Skype traffic analysis and network measurements, that were performed, will be depicted. Measurements were performed on Windows Skype versions 4.1.0.136, 4.2.0.187 and 5.3.0.111 unless otherwise stated.

4.1 Stages in Skype network conversation

Wireshark [27] tool was used as the packet analyzer during the Skype traffic analysis and measurements. Skype stages and corresponding traffic from starting up the application till going offline will be described in this section.

4.1.1 Startup and UDP probing

The client application is started at SC_A. It sends UDP messages to multiple *SN*s saved in client's *host cache* HC_{SC_A} till *positive* response is received. Positively responding supernodes are important for the next stage (described in the subsection 4.1.2). *SC* initiates two-way UDP handshakes to the supernodes stored in *HC*. The source port is the one stored in *Connection* dialog box and destination ports are ports corresponding to the supernodes in the *HC*. The notation for messages in these handshake probes will be P1 and P4 (outgoing UDP probe and incoming response). The purpose of handshakes is to determine possible candidates that allow the *SC* to connect to the Skype network.

We can distinguish different payload sizes for messages – notation for the message X is s_X. If $s_{P4} = 18$ bytes, then we can call this handshake as *positive*. *SC* then tries to establish a TCP connection to the positively probed supernode. From the measurements, if $s_{P4} = 51$ or 53, then we can call this handshake as *negative*. No TCP connection will be initiated between Skype client and supernode. This UDP probing continues in rapid

and continuous manner until positive response is present and TCP connection is established.

Sometimes we can observe handshakes consisting of 2 more messages P2 and P3 at the start of the application. It holds true for the message sizes:

$$s_{P2} = 11,$$

$$s_{P3} = s_{P1} + 5.$$

Analysing the payload of the messages we can differentiate some byte sequences according to [7] – *session identifiers, function parameters, IP address exchange*.

Session identifiers P1 is an initiating message and first two payload bytes forms the session identifier. The important observation is that the session identifier is the same for P2 and P3, on the contrary it does not hold for P4. In fact, this identifier is a number that is increased by two on every new handshake.

Function parameter The third byte of UDP messages has particular values so we can consider it as some kind of function parameter. It holds that it contains value 0x02 for P1 and P4. The value is different for P2 and P3 but the lower nibble is always the same (a nibble is an aggregation of four bits so there are two nibbles in a byte – the higher one and the lower one). It is 0x7 for P2 and 0x3 for P3. There is one more value that is the same through the P3 payloads. The fourth byte possesses 0x01 value.

IP address exchange We can distinguish four four-byte sequences in the messages. Notation for the payload bytes x to y of the message X is X_{x-y}. $P2_{4-7}$ contains IP address of the *SC*, $P3_{9-12}$ contains IP address of the supernode. The *SC*'s IP address in this message is never private address. We have found out that it is the publicly visible IP address created by the NAT closest to the supernode. This is called a reflexive transport address according to [9]. Then sequences that are not IP addresses can be observed ($A_1A_2A_3A_4$ and $B_1B_2B_3B_4$). Our assumption is that they are some unique message identifiers as $P1_{8-11} = P3_{13-16}$ and $P2_{8-11} = P3_{5-8}$.

Byte sequences are depicted in the following tables. SC_1–SC_4 denotes *SC*'s IP address and N_1–N_4 denotes *SN*'s IP address.

$P1$	$SI\,SI$	02	$xx\,xx\,xx\,xx$	$A_1\,A_2\,A_3\,A_4$	$xx\,xx\,\ldots$

$P2$	$SI\,SI$	$x7$	$SC_1\,SC_2\,SC_3\,SC_4$	$B_1\,B_2\,B_3\,B_4$		
$P3$	$SI\,SI$	$x3$	01	$B_1\,B_2\,B_3\,B_4$	$N_1\,N_2\,N_3\,N_4$	$A_1\,A_2\,A_3\,A_4$	...
$P4$	$xx\,xx$	02	xx ...			

We can refer to a *full UDP probe* if P1, P2, P3 and P4 messages are present and to a *partial UDP probe* if only P1 and P4 messages are present. As we can detect particular IP addresses in the full UDP probes, we can consider these probes as some part of NAT detection algorithm which operates similar to STUN (Session Traversal Utilities for NAT) [9].

The *SC* continues in the partial UDP probing even after it has been logged into the Skype network.

4.1.2 TCP handshake with supernode

In this step Caller registers to the Skype network. A TCP request is sent to the positively responding supernode from the previous step (subsection 4.1.1). The TCP connection is established to the same port the UDP probe used. No payload patterns were observed when Skype utilizes outgoing ports different from port 443 as described further.

TCP problems If there is a problem with establishing a connection to the positively UDP probed supernode, then Skype will start initiating connection over ports 443 and 80. The tricky thing is that Skype does not follow communication protocols associated with these well-known ports. Skype employs some modification of TLS protocol [2] for port 443 as described further.

Port 443 It holds for the first message $R1$ sent to the supernode that $s_{R1} =$ 72 bytes. The main observation is that the message payload starts with 56 byte sequence:

80	46	01	03	01	00	$2d$	00
00	00	10	00	00	05	00	00
04	00	00	$0a$	00	00	09	00
00	64	00	00	62	00	00	08
00	00	03	00	00	06	01	00
80	07	00	$c0$	03	00	80	06
00	40	02	00	80	04	00	80

Using Wireshark to view details of captured network traffic we can decode this sequence as follows. It is TLS 1.0 `Client Hello` message in SSLv2 record layer.

Values {`80 46`} mean the length of the message (that is 70), {`01`} handshake message type (here Client Hello), {`03 01`} the version of the TLS protocol by which the client wishes to communicate during this session (here TLS 1.0), {`00 2d`} cipher specification length (here 45), {`00 00`} session ID length (here 0), {`00 10`} challenge length (here 16), {`00 00 05 00 00 04 00 00 0a 00 00 09 00 00 64 00 00 62 00 00 08 00 00 03 00 00 06 01 00 80 07 00 c0 03 00 80 06 00 40 02 00 80 04 00 80`} list of the cryptographic options supported by the client. R1 messages differ only in the last 16 bytes and it should represent the `challenge` attribute.

Response message R2 from the supernode contains always at the first 79 bytes these values:

16	03	01	00	$4a$	02	00	00
46	03	01	40	$1b$	$e4$	86	02
ad	$e0$	29	$e1$	77	74	$e5$	44
$b9$	$c9$	$9c$	$b4$	31	31	$5e$	02
dd	77	$9d$	15	$4a$	96	09	ba
$5d$	$a8$	70	20	$1c$	$a0$	$e4$	$f6$
$4c$	63	51	ae	$2f$	$8e$	$4e$	$e1$
$e6$	76	$6a$	$0a$	88	$d5$	$d8$	$c5$
$5c$	ae	98	$c5$	$e4$	81	$f2$	$2a$
69	bf	90	58	00	05	00	

This start of the message is similar to to the TLS `Server Hello`. Value {`16`} means content type (here `handshake`), {`03 01`} means TLS 1.0, {`00 4a`} means length (here 74), {`02`} handshake type (here `Server Hello`), {`00 00 46`} length (here 70), {`03 01`} means TLS 1.0.

Bytes $R2_{12-15}$ should be `gmt_unix_time`. In the TLS protocol description `gmt_unix_time` is "the current time and date in standard UNIX 32-bit format (seconds since the midnight starting Jan 1, 1970, GMT)". But the $R2_{12-15}$ bytes have fixed value {`40 1b e4 86`} and that is "Jan 31, 2004 18:23:18 Central Europe Standard Time".

Then there are 28 bytes in the `random_bytes` field, one byte of `session ID length`, 32 bytes of `session ID`. Field with bytes $R2_{77-78}$ has the value {`00 05`} and according to TLS it should be "the single cipher suite selected by the server from the list in Client Hello" – here it is `TLS_RSA_WITH_RC4_128_SHA`.

Port 80 In Skype version 2.0 there were noticed value recurrences in the higher nibbles of bytes of the R1 message as mentioned in [7]. These recurrences or any other were not confirmed by analysis in the messages to the port 80 in Skype versions 4.1 and 4.2.

4.1.3 Authentication

After connection to some supernode with destination port 33033, *SC* tries to authenticate itself to the Skype network. In previous versions of the Skype application there were four TCP messages exchanged in every connection to *LS* in general and after that the connection is closed as mentioned in [7]. All tested Skype versions were searched for some payload pattern but there was found a pattern only in communication with particular *LS*s like *LS* with IP address `195.46.253.219`. Our assumption is that the pattern depends on a particular login server or some unknown condition. When the pattern occurs then the first six messages M_1-M_6 exchanged between *SC* and *LS* look like depicted further.

M_1	$SC \rightarrow LS$	16 03 01 00 00
M_2	$SC \leftarrow LS$	17 03 01 00 00
M_3	$SC \rightarrow LS$	16 03 01 00 00
M_4	$SC \leftarrow LS$	17 03 01 00 00
M_5	$SC \rightarrow LS$	16 03 01 00 xx 42 cd ef e7 40 d7 2f 1d ...
M_6	$SC \leftarrow LS$	17 03 01 01 ...

Messages M_1-M_4 have fixed size of 5 bytes as opposite to M_5 and M_6. The fifth byte of the message M_5 does not contain always a fixed value. This holds also for the connections on outgoing ports 443 and 80 too if they are used.

4.1.4 Skype latest version check

In the next stage, if the application was installed for the first time and this is its first run, then it sends an HTTP request to *HS* (`ui.skype.com`). The request method is `GET` and we can find in the request URI following patterns: `/ui/` and `/installed`, see Table 4.1.

If Skype has already been installed, it will check with *HS* the latest version of the application. This check occurs every time Skype is started. Skype client sends an HTTP request and the patterns are `/ui/` and `/getlatestversion?ver=`, see Table 4.2.

/ui/0/5.3.0.111./en/installed ?info=google-toolbar:notoffered; ienotdefaultbrowser2,google-chrome:notoffered; alreadyoffered

Table 4.1: Example of the `GET` request after Skype installation

/ui/0/5.3.0.111./en/getlatestversion?ver=5.3.0.111 &uhash=1d1bb29ea1f2970757800d8e22b9ce8d6 &google-chrome:notoffered; alreadyoffered

Table 4.2: Example of the `GET` request after Skype start

This `GET` request was captured during measurements regularly every 4 hours and furthermore requests, but with the empty request URI, can occur between these regular `GET` requests. Under normal conditions (e.g. no TCP `RST` flag is present) 5 TCP packets are sent in both directions, otherwise at least 3 TCP packets are sent.

4.1.5 NAT and firewall determination

At the transport layer, Skype performs NAT and firewall detection. NAT traversal is an important function of Skype for determining what kind of NAT settings is the *SC* currently behind. Once determined, the client stores this kind of information in some manner in the "shared.xml" file as the element `<NatTracker>` is present. Each *SC* uses a variant of the STUN protocol to determine the type of NAT the client is behind. There is no global NAT and firewall traversal server. If there was one, the *SC* would have exchanged traffic with it at the start, but no IP address repetition or some specific behaviour at particular stages of the Skype session was observed.

4.1.6 Going online

Neighbour supernode (*NSN*) If the client starts only with the Offline status, no further traffic is recorded. Going Online, UDP packets are sent to multiple *SN*s rapidly and continuously. This kind of UDP pinging messages is detected during the whole Skype session (these UDP probes were described in the previous subsection 4.1.1). In this step, *SC* searches for

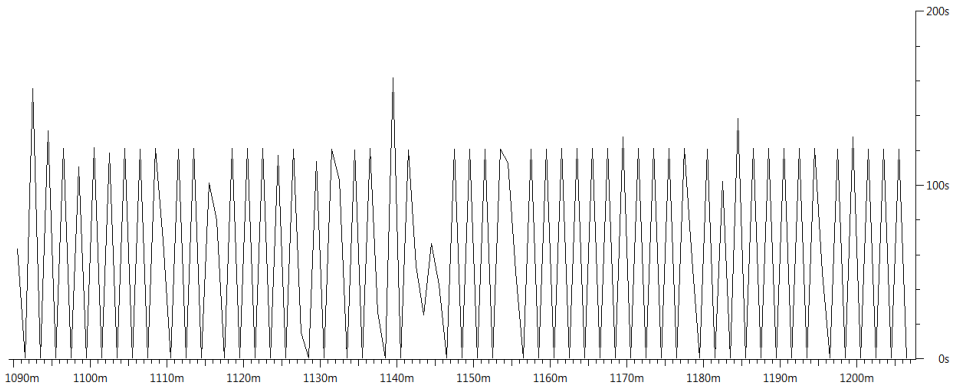


Figure 4.1: At least every 120 seconds *SC* sends TCP packet to its *NSN*.

available *neighbour supernode* (*NSN*) and binds to it with TCP connection.

After positive UDP handshake with some *SN* as described previously, *SC* initiates a TCP connection to this node. If the connection is successful, then *SC* and *SN* exchange several messages. If the connection is kept, then *SN* becomes *NSN* for the Skype client. This binding lasts as long as the *SC* is online but sometimes termination of the connection can occur (e.g. 39 hours lasting connection to *NSN* was captured during measurements till termination occurred). Our prediction is that it is caused by unavailability of the *NSN* or the supernode becomes a regular node. As the *SC* is always bound with some *NSN*, a new binding process will start after termination of the previous connection with *NSN*. It has not been found out any particular pattern that would be common for every startup of the *SC*. However, we have often observed packets with payload sizes 4, 8 and 27 (or 28) bytes initiated by *SC* between first 10 packets exchanged between *SC* and *NSN*.

There is probably some timeout counter used on the both sides as we can find regular traffic. We have found out that after 120 seconds lasting time window or if any of the parties sends packets containing some signalling information during the 120 seconds time window, then *SC* or *NSN* sends a packet with $s = 2$ and a sequence of messages M_1–M_4 begins. They can be some keep-alive messages. New timeout starts to count down for corresponding parties from the moment when M_1 and M_2 are sent.

M_1	$\rightarrow$	$s_{M_1} = 2$	$TCPPush, Ackflags$
M_2	$\leftarrow$	$s_{M_2} = 0$	$TCPAckflag$
M_3	$\leftarrow$	$s_{M_3} = 2$	$TCPPush, Ackflags$
M_4	$\rightarrow$	$s_{M_4} = 0$	$TCPAckflag$

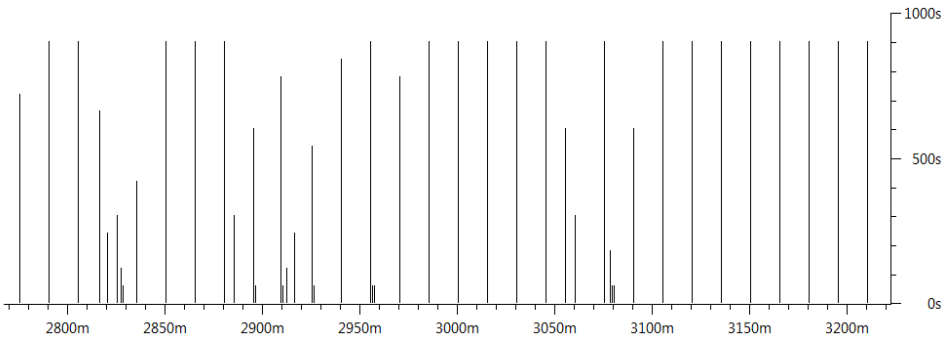


Figure 4.2: Inter packet gap of 900 seconds (i.e. 15 minutes) for outgoing TCP packets (with payload $>$ 200 bytes) to *NSN* was observed.

Periodicity in sending packets is shown in Figure 4.1. Almost 135 hours of standby Skype traffic with the longest continuously measured time slot of 67 hours were captured for analysis. Inter packet time gaps for outgoing packets in one minute time slots were counted up and that has exposed peaks of 120 seconds almost every two minutes. Moreover, it has been found out for the Skype standby mode that for most outgoing packets with payload greater than 200 bytes the inter packet time gap equals to 15 minutes, see calculated statistics in the Figure 4.2. This was detected only for Skype versions 4.1 and 4.2.

UDP probes From the results of measurements, *SC* establishes and quickly terminates in average 5 TCP connections with supernodes per hour. Related to UDP protocol, UDP probing as checking available peers in the Skype network occurs mainly in bursts. In average, 30 new peers are probed every hour (approximately 1900 per 67 hours) as shown in Figure 4.3, which shows us statistics for the 67 hours lasting Skype standby mode. The most peers are probed at the start of the application as we can see. Every peer is characterized by its ID in this figure so if a new peer is probed, then it gains his ID and corresponding dot is plotted to the graph. If some peer sends UDP packet to the *SC*, then corresponding ID is plotted on the negative y-axis. Lot of peers are probed repetitively. Very interesting is the fact that sometimes the *SC* sends UDP packets repetitively (e.g. 700 packets per 50 hours) even to the peers that have never responded.

Perl modul Geo::IPfree was used to determine the originating countries of the probed peers, see results in the Figure 4.4. The modul uses a local

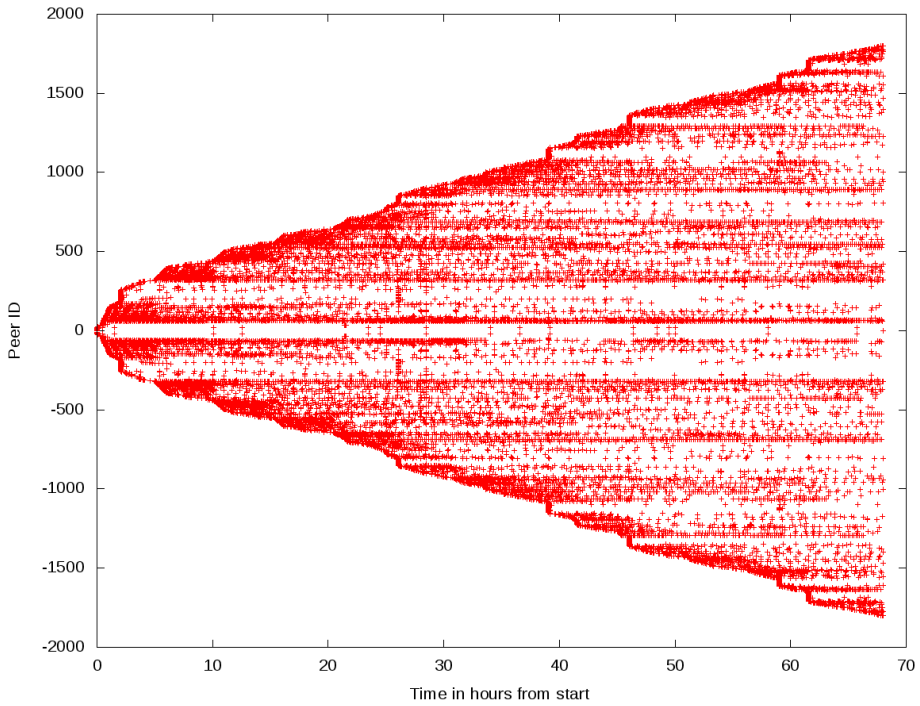


Figure 4.3: Each dot represents packet sent in a given time by *SC* to some other Skype peer whose corresponding ID is represented on the y-axis. Positive ID is for packet from *SC* to the peer, negative ID for packet to *SC*.

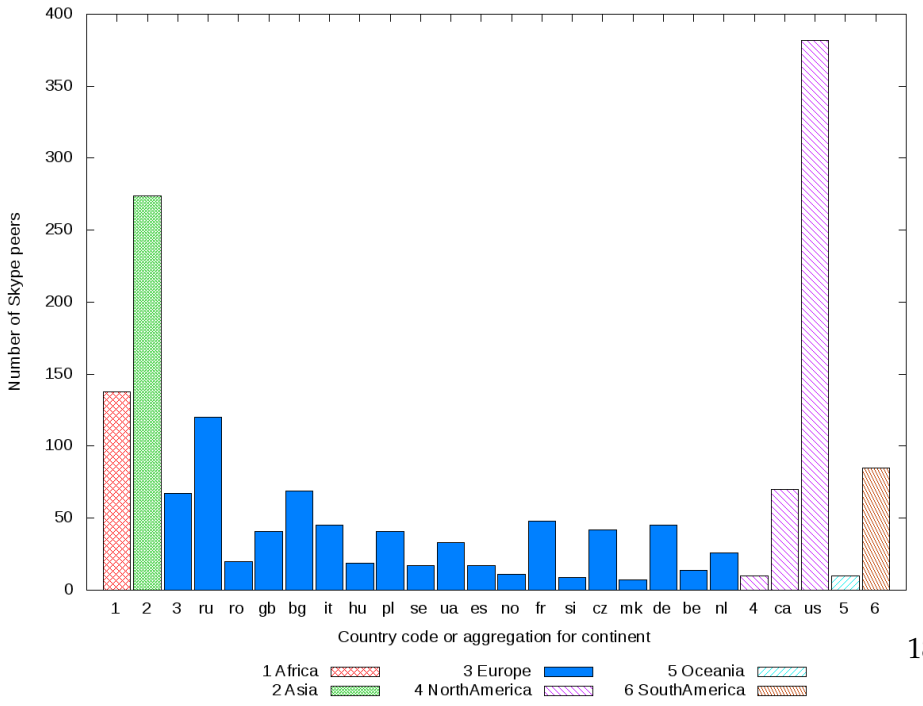


Figure 4.4: Country statistics for contacted peers.

file-based database to provide basic geolocation services. Countries were aggregated to continents if their country codes are not presented. It is visible that most contacted peers are located in the United States.

TCP probes If UDP protocol is restricted, *SC* utilizes TCP protocol for TCP probes but not in bursts like in UDP case and in lesser scope. Positively probed supernode becomes *SC's NSN*. TCP packets are exchanged continuously between *SC* and *NSN* with average packet rate 1 per 10 seconds.

4.1.7 Going offline

When the client is switched to Offline status, every open TCP connection is closed. TCP connection with the *NSN* is now ended. After sign out from the Skype application, TCP connection with IP address like `78.141.181.242` or `213.146.188.16` is established. This communication occurs sometimes after startup of the *SC* too. These addresses are associated with Skype Communications according to WHOIS database. Our assumption is that this connection can be somehow important for creating Skype global statistics about users. It is worthwhile to mention that destination ports 13392 or 12350 during all these captured connections were observed.

4.2 Voice call properties

In this section Skype voice call statistics will be presented.

4.2.1 Voice codec

SILK is a speech and audio codec developed internally at Skype which is used as the default codec for all Skype to Skype calls. It is highly scalable in terms of audio bandwidth, network bit rate, and complexity, making it the codec of choice for multiple modes and applications as described in [12, 8]. SILK is a replacement for the SVOPC codec [26] used firstly in Skype version 3.2. The SILK codec was a separate development branch from SVOPC and the final version was introduced in Skype version 4.0 (February 2009).

The SILK speech and audio codec is highly scalable in terms of audio bandwidth, network bit rate, and complexity. SILK supports four different audio bandwidths: 8000 Hz, 12000 Hz, 16000 Hz and 24000 Hz sampling frequency as shown in Figure 4.5.

	Sampling Rate (kHz)	Bit Rate (kbps)	CPU (MHz on x86 core)
Narrowband for PSTN gateways and low end devices	8	6 - 20	12 - 30
Mediumband for devices with limited wideband capability	12	7 - 25	16 - 40
Wideband for all-IP platforms	16	8 - 30	20 - 50
Super-Wideband, a new standard in speech quality	24	12 - 40	30 - 80
Key Advantage	Optimize clarity under hardware and network constraints	Adjust to degraded network conditions in real time	Match complexity to CPU resources in real time

Figure 4.5: The average network bit rate for corresponding audio bandwidths, figure from [19]

Narrowband mode should only be used to interface to PSTN (public switched telephone network) networks or on low end devices that do not support greater than 8000 Hz sampling frequency, mediumband mode for lower end devices that do not support greater than 12000 Hz sampling frequency or are under severe network bandwidth constrains (e.g. wireless devices). Wideband mode should be used for all-IP platforms that do not support greater than 16000 Hz sampling frequency. Super wideband mode should be used on all platforms that support 24000 Hz and greater sampling frequency.

The internal audio frame size of SILK is 20 ms. The SILK encoder can be set to join up to five internal frames into a single frame output. That means we can get 20, 40, 60, 80 or 100 ms frames of encoded speech or audio data. It is mentioned that SILK operates at a very low algorithmic delay, consisting of packetization delay, i.e. 20, 40, 60, 80 or 100 ms plus 5 ms lookahead delay.

The internal sampling frequency of the encoded speech or audio signal of SILK may change during the duration of a transmission. The average bit rate target can be adjusted on a per frame basis. This allows support for congestion control and network load management.

4.2.2 Call placement

The *SC* needs TCP connectivity to send call signalling information as described in [20]. It strongly prefers UDP connectivity for voice and video communication. If UDP is unavailable, it can utilize TCP for the media

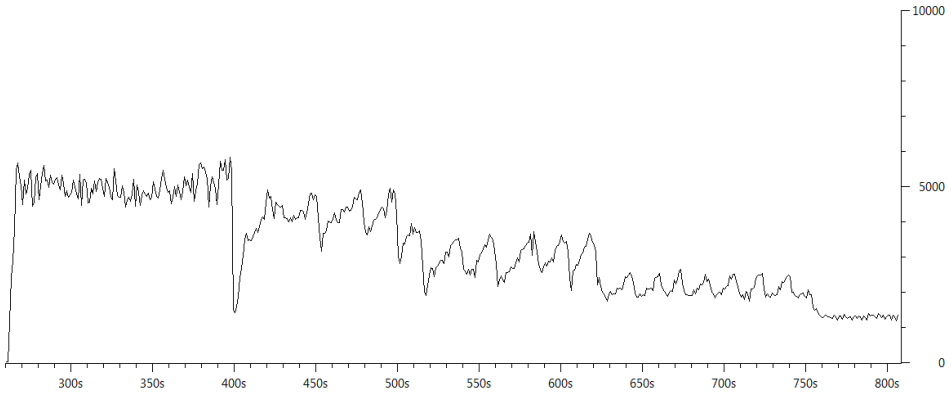


Figure 4.6: Bit rate during a voice call under decreasing available bandwidth every 120 seconds.

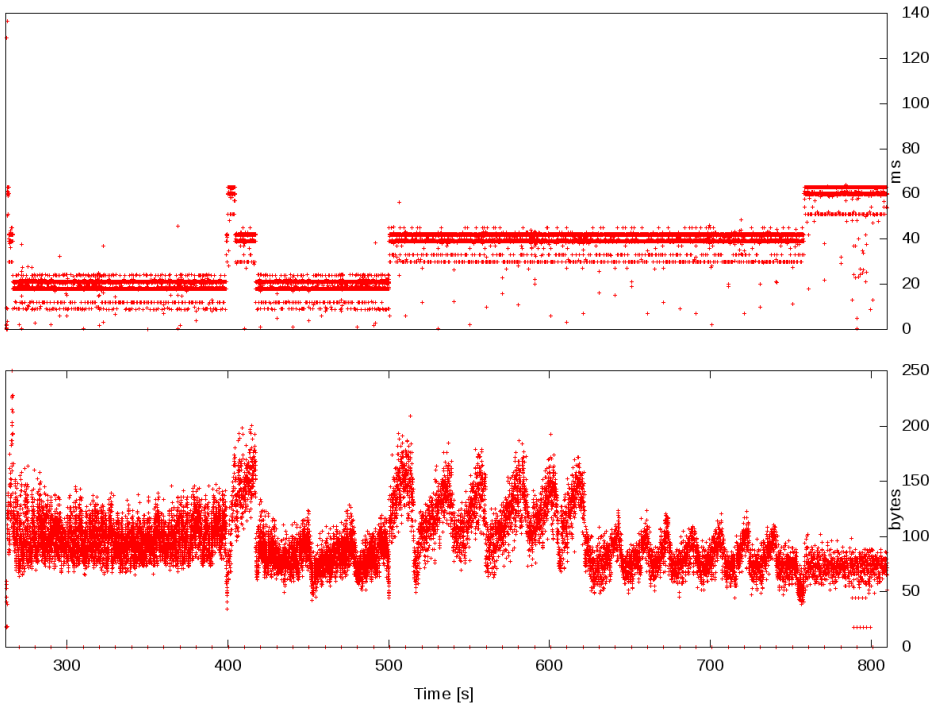


Figure 4.7: Inter packet gaps and bytes per packet during a voice call under decreasing available bandwidth every 120 seconds.

stream but with the additional overhead due to TCP being stateful. Before a user places his call, the client communicates with the peer network to test network connectivity. It checks whether the outgoing UDP port is available and the type of address translation used by network. Status checking and updating is also carried out through P2P architecture to identify online status of contacts on the contact list. We have observed that if there appears status change (e.g. some contact goes from *Offline* status to *Online* status), then *SC* is informed about this change from *NSN* by 2 up to 4 TCP messages exchange.

There are several possibilities how the voice or video communication is transfered through the network. If both caller (SC_A) and callee (SC_B) are on machines with public IP addresses, then upon pressing the Call button in the application interface the UDP probing to different peers occurs, SC_A establishes a TCP connection to the SC_B and some signalling information is exchanged. During the voice or video call UDP packets are then exchanged, TCP connection is alive during the whole call and is kept for sending some signalling information. It should be mentioned that for UDP data packets like voice packets the lower nibble of the third byte equals to 0xd.

IF SC_A is behind NAT, it is able often to traverse NAT and negotiate networking parameters (remote IP address and source port) through supernodes and then to initiate direct UDP connection. If SC_A is not able to communicate directly, then it will find the appropriate *relay* for the connection. Relays are supernodes that relay media traffic and signalling information between clients that are not able to reach each other directly. SC_A will then try to connect to some supernode that will be the relay node. Also SC_B connects to some other supernode that will serve as the relay node.

If SC_A is behind UDP restricted firewall, then TCP protocol is utilized for transfering data and relays are always used. However, not only one relay but several relays are used concurrently as we have detected by measurements. The assumption is that several relays are used for some fault tolerance or backup purposes.

How would voice call properties change, a UDP call was placed inside local network under decreasing available bandwidth, see Figures 4.6 and 4.7. Every 120 seconds the available bandwidth was decreased (the bandwidth was restricted by software tool). The bandwidth was unlimited for the first period, then it was changed to 4000, 3000, 2000 and 1000 bytes per second. We can see how 20, 40 and then 60 msec frames of encoded speech were sent and how the payload size of frames was changed. The payload size was always lower than 300 bytes. We can get the idea of voice call characteristic from the facts that Skype uses one particular voice codec (SILK)

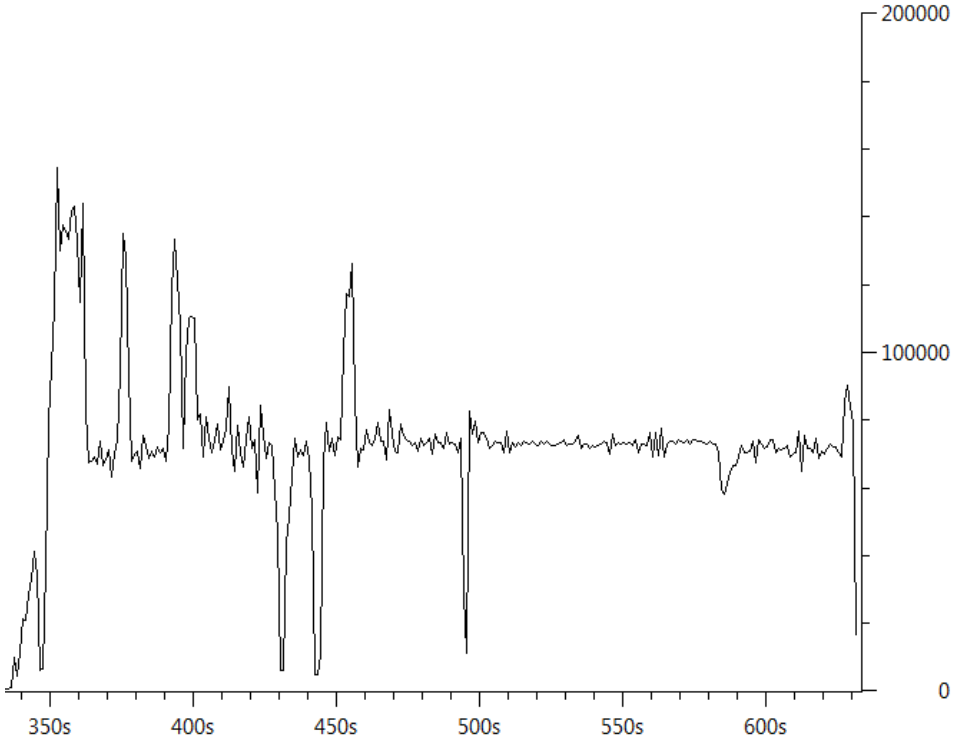


Figure 4.8: Bytes per second during a video call.

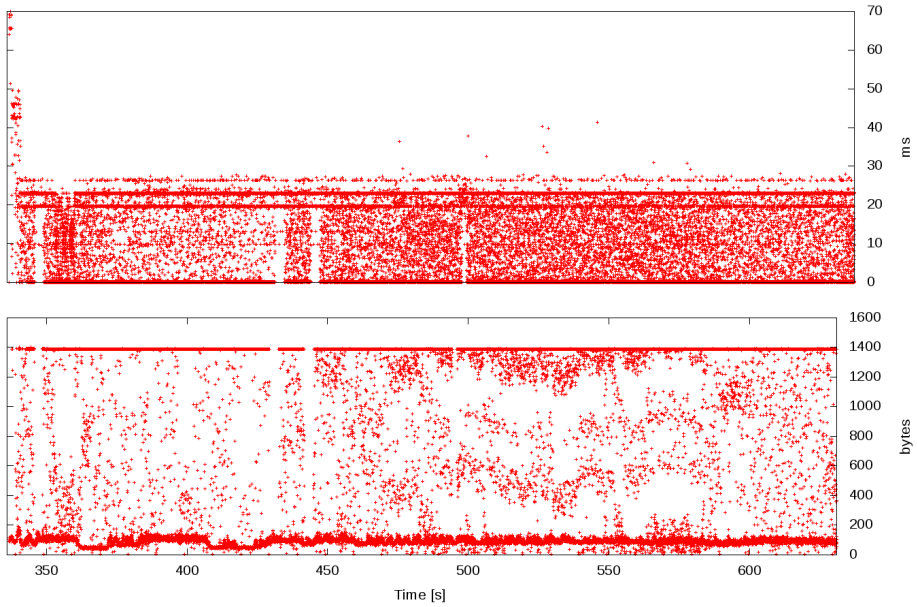


Figure 4.9: Inter packet gaps and bytes per packet during a video call.

for calls between Skype clients since application version 4 and that the parameters of the audio codec are available.

On the other hand, detailed video codec parameters are not publicly available. Skype uses a TrueMotion VP7 video codec developed by On2 Technologies as described in [11]. The description of the codec is not available. Worth to mention is that Skype can operate on bare minimum bandwidth including video as low as 4 kbps [11]. A video call was placed to analyze if there is particular pattern e.g. in inter packet time gaps. Figures 4.8 and 4.9 show a regular video call with 30 frames per second and resolution 320 x 240. As we can see, there is no clear distribution between inter packet time gaps. Dots around the value 20 ms shows us that not all voice packets are included in the video packets. Video packets are not sent continuously but mainly in bursts – this is visible by the distribution around the value 0 ms in the Inter packet gaps figure 4.9. Voice packets have payload size around the value of 150 bytes whereas video packets have payload size around 1380 bytes. As described in [22], Skype can utilize bandwidth from 100 kbps up to 1.5 Mbps in the case of HD video call (in group video calling up to 8 Mbps for download).

4.3 Proposed voice call detection method

This chapter has dealt with the analysis of Skype traffic. Patterns that could be used by packet filters and analyzers were described. The main purpose of analysis and measurements was to find out patterns that will reveal Skype traffic not from packet payload but from *flows* (next chapters deal with flows in detail). Flow is a sequence of packets with some common properties and gives us traffic statistics calculated from all packets in the flow. Therefore long lasting patterns are those that can be used.

Instant messaging and sending files were also under the scope of measurements. These activities do not reveal long lasting patterns. In the case of utilizing TCP protocol for voice or video calls Skype opens randomly chosen TCP ports and several concurrent connections to different relays are established. Traffic bandwidth is not equally divided between these connections. On the contrary, UDP protocol reveals characteristic that could be used for flow analysis. As detailed parameters of voice calls were presented, we can propose a detection method that will detect Skype voice calls.

Skype latest version check is performed at least every 4 hours so we can search for *SCs* that establish connection to `ui.skype.com`. UDP probes are sent throughout the whole Skype session and as they are sent from the

same source UDP port, we can search for the port number. We can search for it at two different occasions to increase the reliability of correctly detected port – once the latest version check is detected and once the voice call is detected. Voice calls have particular packets per second, bits per second and bytes per packet characteristics. Also calls have bidirectional patterns so the number of incoming connections to detected IP address and UDP port will be searched. The proposed method for detecting Skype voice calls follows, detailed explanation of values is presented in 6.1:

- Search for UDP voice calls.
  UDP connections where packets per seconds > 15 and < 70,
  bits per second > 13000 and < 99000,
  bytes per packet < 300.

- Search for Skype latest version check.
  TCP connections where destination is `ui.skype.com`,
  destination port is 80,
  number of packets > 3.

- Search for Skype UDP port at the time of voice call start and at the time of Skype latest version check.
  UDP connections established from source IP address,
  source port > 1024,
  bytes per packet > 40 and < 80,
  number of packets equals to 1.
  Port with the most connections is marked as the Skype UDP port.

- Count number of incoming calls to the particular IP address and UDP port.

- If the source UDP port of voice call equals to UDP ports detected during Skype latest version check and start of the voice call,
  number of incoming calls is higher than 0 and lesser than 25,
  Skype latest version check was detected during last 4 hours,
  then detected voice call is confirmed as voice call.

Chapter 5

NetFlow protocol

NetFlow is a network protocol developed by Cisco Systems for collecting IP traffic information. It has become an industry standard for network traffic monitoring and is widely used measurement solution today. NetFlow provides network administrators with acces to IP flow information from their data networks. A flow is defined as an unidirectional sequence of packets with some common properties that pass through a network device. Network elements (e.g. routers) export these collected flows to an external device – the NetFlow collector. Exported data is used for a variety of purposes, including ISP billing, network, user and application monitoring, capacity planning, security analysis.

During development of NetFlow there were several protocol versions presented. The first implementation was version v1 in 1996. It was restricted only to IPv4 and e.g. did not support IP masks. Versions v2, v3 and v4 were developed for internal CISCO purposes and have been never released. Version v5 came in 2009 and is the most common and is supported by different brands of network devices. Next versions till the version v9 are not widely used. So after v5 the next commonly used version on recent network devices is v9 [6]. IPv6 is supported and it uses templates to provide access to observations of IP packet flows in a flexible and extensible manner. A template defines a collection of fields, with corresponding descriptions of structure and semantics. The advantages are that it allows export of only required fields from the flows and that new fields can be added to the flow records without changing the structure of the export record format. Protocol IPFIX (Internet Protocol Flow Informatin eXport) becomes a successor to v9. It is IETF standardized NetFlow v9 with several extensions.

Netflow architecture consists of several NetFlow exporters. These devices monitor packets entering a particular location in the network and creates flows from these packets. The information from the flows is exported in the form of flow records to the NetFlow collector. Usually there is one collector which parses the incoming flow records and stores them. NetFlow monitoring using standalone NetFlow probes is an alternative to flow col-

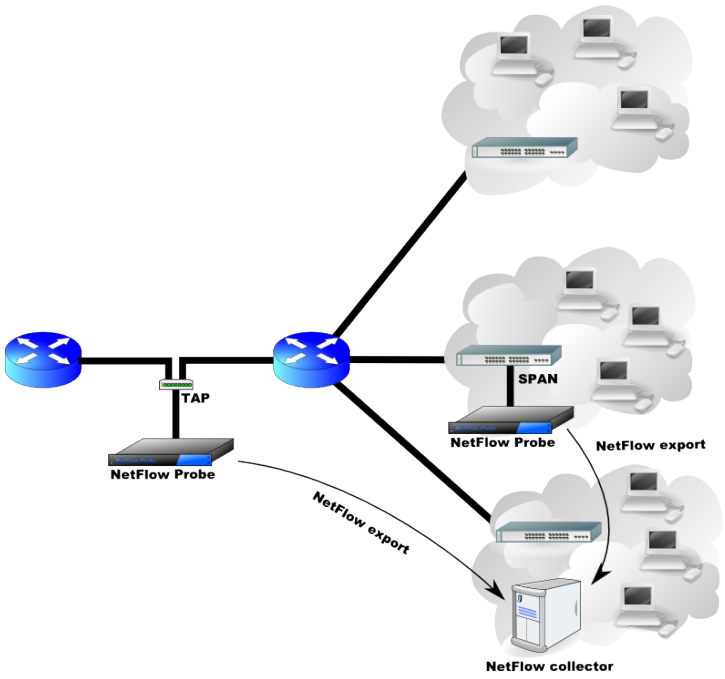


Figure 5.1: NetFlow architecture using standalone probes, figure from [24]

lection from routers or switches. This approach can overcome some limitations of router-based NetFlow monitoring, e.g. routers can be under heavy traffic load and must create flows simultaneously. The probes are transparently connected to the designated location and do not influence the packets passing this location as depicted in Figure 5.1.

The common properties that the sequence of packets in the flow share are (according to the traditional Cisco definition) these 7 values:

- Source IP address
- Destination IP address
- Source port
- Destination port
- IP protocol
- Input interface
- Type of Service

The exporter will output a flow record under several conditions:

- If the exporter can detect the end of the flow, e.g. the `FIN` and `RST` bits in a TCP connection.
- If the flow has been *inactive*, i.e. no packets belonging to the flow have been observed for a certain period of time (*inactive timeout*).
- Long-lasting flows are exported on a regular basis. That means that flow records are exported continuously after a certain period of time (*active timeout*) for the capturing flow.

The exported record can contain a wide variety of information about the traffic in a given flow. NetFlow v5 record contains version number, sequence number, input and output interface, timestamps of the flow start and end in milliseconds, number of bytes and packets associated with a flow, source and destination IP addresses and ports, IP protocol, Type Of Service, for TCP flows the aggregation of TCP flags, routing information like IP address of the next-hop along the route to the destination and source and destination IP masks.

5.1 NfSen and NFDUMP

NfSen [18] and NFDUMP [17] are tools that are distributed under the BSD licence. The NFDUMP tools collect, process and store NetFlow data on the command line, NfSen is a graphical web based front end for the NFDUMP.

5.1.1 NFDUMP

NFDUMP contains several tools that support NetFlow protocol in versions v5, v7 and v9.

- **nfcapd** (NetFlow capture daemon) – captures the NetFlow records from the exporter and stores them into files. Automatically rotate files every 5 minutes. There is one nfcapd process running for each exporter.
- **nfdump** (NetFlow dump) – reads and displays the NetFlow data from the files stored by nfcapd. It can create statistics, make flow data aggregations or filter stored data.
- **nfprofile** (NetFlow profiler) – reads the NetFlow data and filters it according to the profiles and stores it into files for later use.

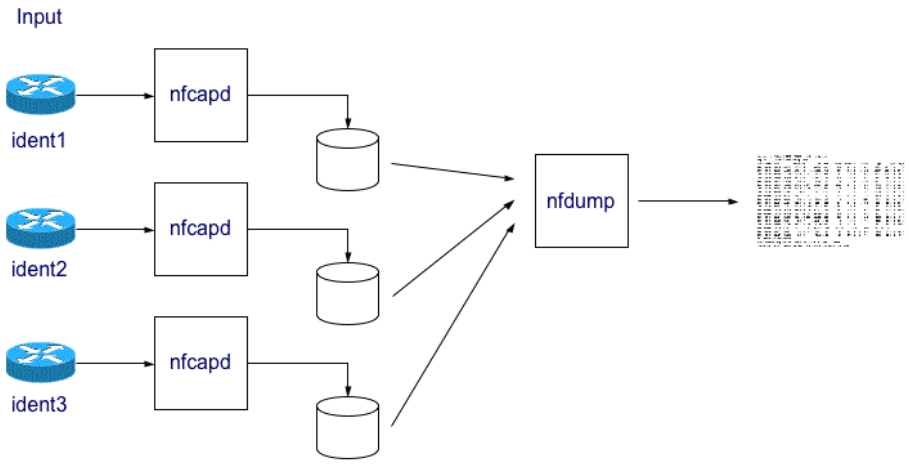


Figure 5.2: Capture daemons store NetFlow data that is read by nfdump, figure from [17]

- **nfreplay** (NetFlow replay) – reads the NetFlow data and sends it over the network to another host.

- **nfclean** (old data cleanup) and **ft2nfdump** (data convertor)

The goal of the design is to able to analyze netflow data from the past as well as to track interesting traffic patterns continuously. The amount of time back in the past is limited only by the disk space available for all the netflow data. The tools are optimized for speed for efficient filtering.

All data is stored to disk before analyzing. This separates the process of storing and analyzing the data (Figure 5.2). The data is stored in the *time slot* based fashion. Each file contains flow records captured during the configured 5 minutes time slot (this value is the actual value set on a particular NetFlow collector from Chapter 6). The output file format corresponding to the time slot is `nfcapd.YYYYMMddhhmm`, e.g. `nfcapd.201104271330`. If there are several exporters from which the NetFlow data are stored, then the data is organized in the corresponding directories. One can choose its own subdirectory structure, but in the main configuration it is year, then month and day.

Flows can be read either from a single file or from a sequence of files as depicted in Figure 5.3.

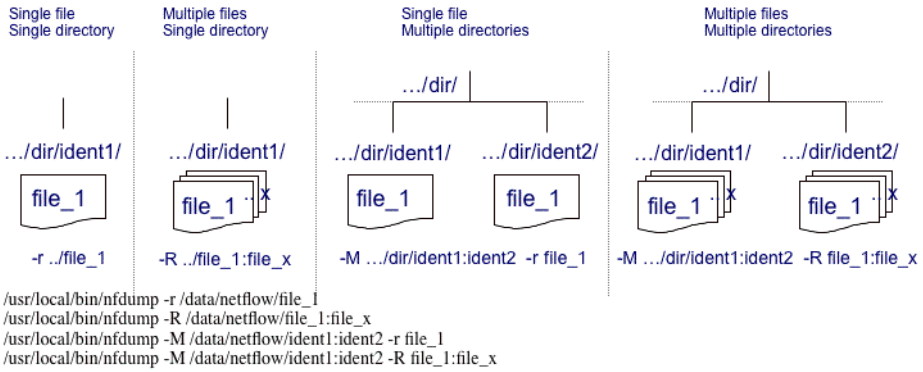


Figure 5.3: How the files are stored by nfcapd and read by nfdump, figure from [17]

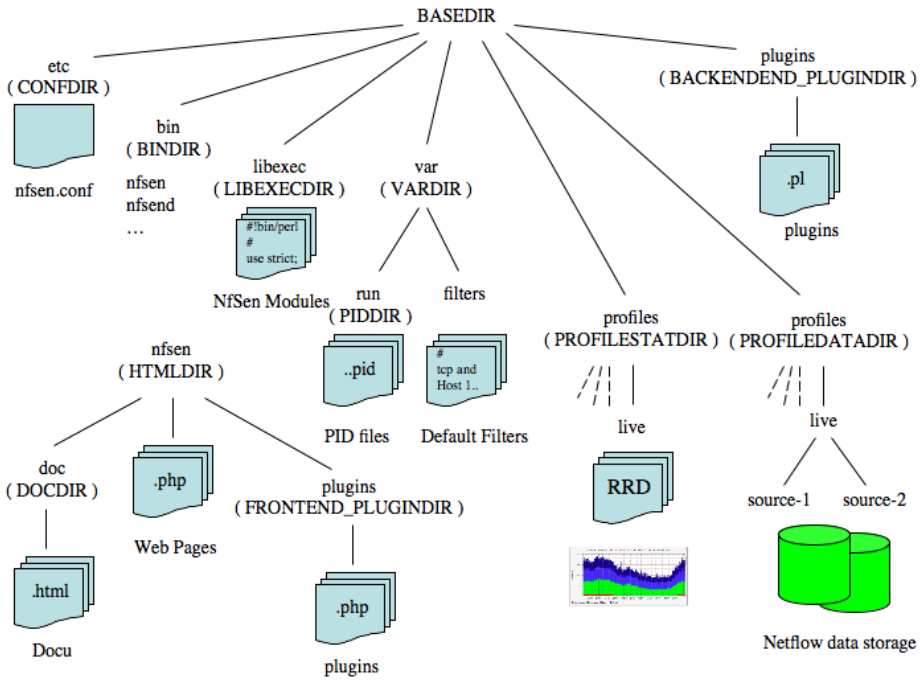


Figure 5.4: Directory structure, figure from [18]

5.1.2 NfSen

NfSen is graphical web based front end utilizes NFDUMP tools as mentioned before. It manages to display statistics about number of flows, packets and bytes in graphs using RRD (Round Robin Database) where graphs data is stored. As humans need different point of views on the world around us, they need also different interfaces to the stored data to view them in different ways. NfSen preserves benefits of the command line interface and brings concurrently graphical view. It allows easy navigation through stored NetFlow data, processing data from a single time slot or from a time window, creating history or continuous profiles to make specific view on the NetFlow data. NfSen also provides an option to execute specific actions (alerts) based on user defined conditions. One of the important functions is the option to extend NfSen with plugins. This affords us sufficient ways how to fit our additional needs and allows us to modify NfSen capabilities. Plugins may be selected from the NfSen navigation bar.

NfSen has a very flexible directory layout. That means that the administrator can configure NfSen to fit his own needs. Default layout is shown in Figure 5.4. All NetFlow data is stored under `PROFILEDATADIR`. NfSen is reachable by web interface after the installation, but there is also possibility to use the command line interface if it is needed.

5.1.3 Plugins for NfSen

Plugins allow to add additional functionality in the area of statistical processing and output presentation. Plugins structure is divided into two types, namely *frontend plugins* and *backend plugins*. Backend plugins are intended to process stored NetFlow data or to prepare them for output, or bring functionalities like alerting conditions and actions to the NfSen. They are run periodically or they can be run from the frontend plugin. Frontend plugins may display any kind of data resulting from backend plugin processing. The backend plugins are Perl modules. The frontend plugins are defined as PHP scripts with the same name as the backend plugin. Both plugins may exchange relevant data over nfsend.comm socket as shown in Figure 5.5. Over this communication channel any number of scalar and array values can be exchanged. Communication functions for the frontend plugin are defined in `nfsenutil.php` file and communication functions for the backend plugin are defined in the Perl module `Nfcomm.pm`. Both plugins must implement specific functions to be correctly integrated to NfSen.

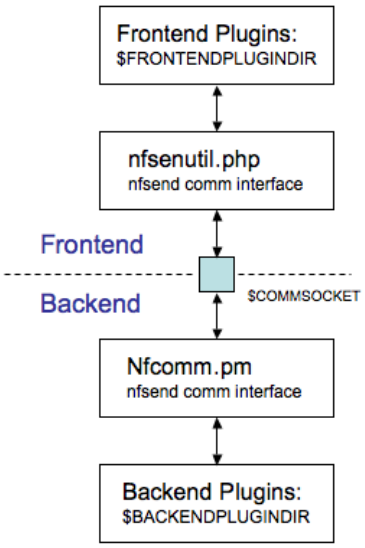


Figure 5.5: Communication concept for backend and frotend plugins, figure from [18]

Chapter 6

Creation of DetectSkype plugin for NfSen

As a result of measurements a plugin for NfSen was created. It was developed for NfSen collector situated on `nftest.ics.muni.cz`. Describing state as of July 2011, it processes flow records from 3 channels:

- 10GE connection between the Masaryk University network and CESNET
- 1GE from/to the Faculty od Informatics
- 1GE from/to the Vinařská dormitory

Real traffic is monitored but the data is anonymized on this collector. That means that IP addresses are anonymized but the characteristics of the flow records remains untouched.

The activity of the plugin can be divided into two ways. One is automatic and periodic processing of stored data and creation of statistics (described in the following section 6.1), the other one is processing user commands and presentation of results (described in the section 6.2).

6.1 Backend plugin

Backend plugin is written in Perl language. The main part constitutes of NfSen demanded function `run`. It is run periodically with relevant parameters (*profile, profilegroup, timeslot*) every 5 minutes by NfSen. Data processed by the plugin are stored in six database files. The used database is Berkeley DB, which is easily usable in Perl environment.

For creation of the plugin, it was necessary to define Skype activity characteristic from the flow point of view. The plugin utilizes the proposed method in 4.3. If Skype client is not able to utilize UDP protocol for placing the voice traffic, then TCP protocol is used. Several TCP connections are established with supernodes, dynamical source ports are present and

no constant traffic with some continuous pattern was detected. On the contrary, Skype UDP traffic demonstrates several patterns that could be used for voice call detection. Following subsections will cover implemented traffic patterns that are based on Skype traffic measurements described before.

6.1.1 TCP traffic to ui.skype.com

As Skype client initiates every startup a TCP connection to specific address, it was the first rule that was considered for implementation. The *SC* sends an HTTP request to the HTTP server `ui.skype.com` to check the latest version of the application. This request is also sent every four hour and under normal conditions (e.g. no TCP `RST` flag is present in the flow) 5 TCP packets are sent in both directions, otherwise at least 3 TCP packets are sent. Filter for nfdump that will gain flow records representing this rule follows:

```
proto tcp and $srcnet and dst ip $uiskypeaddress
and dst port 80 and packets > 3
```

`$srcnet` is set to `src net 147.250.0.0/16` as the university addresses are anonymized to this mask. `$uiskypeaddress` is set to 203.233.225.202 as this is the actual anonymized address for `ui.skype.com`. If anonymization is not used, then the parameter can be set to `ui.skype.com` because fully qualified hostnames are supported by nfdump.

IP addresses, that were in connection with Skype HTTP server for the last four hours and corresponding start time of the flow, will be stored in the database.

6.1.2 UDP voice call properties

Based on measurements and available documentation of SILK codec we can introduce traffic filter that deals with voice call properties. As we can filter flow records in respect to bit rate and other useful fields, the filter for obtaining flows with voice characteristic will be:

```
proto udp and duration > 296000 and pps > 15
and pps < 70 and bps > 13000
and bps < 99000 and bpp < 300
```

As TCP voice calls utilize several connections to supernodes as described before, the focus was placed on UDP calls. Since the introduced characteristic of the voice call does not hold true for short periods of time like at the beginning and at the end of calls (e.g. signalling packets are sent to the callee after the end of call and influence calculated fields of the flow

records), flow records that are spanned through the entire time slot (i.e. 5 minutes long flow records) were considered – therefore `duration > 296000 msec`. As the frames mainly of sizes 20, 40, 60 msec are generated by SILK voice codec in the rate 50, 25, 16 frames per second, records are filtered out with packets per second value in this manner with some overhead. Bytes per packet are set according to measurements i.e. `bpp < 300`. During processing of flows with given properties there is also check for the identical flow records. As outside trafic e.g. from CESNET to the Faculty of Informatics should be captured at two channels, there will be almost *identical* flow records for a given flow. They are eliminated by the check for almost identical start time field (threshold difference 3 seconds). Each flow with given properties is stored to the database. Every new flow in the next time slot is checked against these stored flows. If positive match is found i.e. source address, source port, destination address, destination port match, then these flows are merged together as continuous flow. If some flow is not merged with some new flow, then it is considered that the voice call has ended and this flow is exported to the history database with additional values like UDP port (values are explained further in the frontend plugin section 6.2).

6.1.3 UDP port

Next characteristic for the Skype client is the used UDP port. It is used as the source and incoming port for the UDP protocol and its value is always higher than 1024. It is set in the *Connection dialog* of the client so it is not dynamically changed during the run of application. Plugin searches for this port in two situations.

The first one is the moment when the connection to the HTTP server is detected as depicted in the subsection 6.1.1. Then in the next time slot IP addresses, from which the connections were established, are searched for the corresponding Skype UDP ports. Skype initiates UDP probes during the whole Skype session and, from the flow point of view, the flow consists of one packet because in the future generated probes to the same peer will belong to the new flow. Therefore the filter is:

```
  proto udp and src port > 1024 and bpp < 80
and bpp > 40 and packets = 1
and src ip in $filter_addresses
```

Bytes per packet are set in compliance with the packet length statistics gained from Skype standby measurements and measurements to find out how similar BitTorrent UDP probing activity is to the one of Skype. Dur-

ing testing the proposed plugin implementation BitTorrent activity was influencing this filter. Most of the BitTorrent UDP probe packets have sizes greater than 90 bytes. For UDP packets of Skype standby measurement (from 67 hours lasting log) we can present following packets statistics:

Packet lengths	*Count*	*Percent*
$40 - 79$	8472	52.98%
$80 - 159$	522	3.26%
$160 - 319$	6835	42.75%
$320 - 639$	161	1.01%

Therefore bytes per packet parameters are limited by the lower value `40` and upper value `80`. Detected flows are aggregated by IP source address and source port and the aggregated flows are sorted according to the aggregation number. Record with most flows is chosen as the representative one and corresponding source port is selected as possible UDP port of the Skype application.

Second situation, when the UDP port is searched, is when the start of potential voice call flow is detected (depicted in subsection 6.1.2). There is a burst of UDP probes always at the beginning of voice call.

As UDP ports are searched at two different times, we increase the probability that the port does not belong to some other network activity at a particular IP address.

6.1.4 Bidirectional property

Voice calls are bidirectional flows so the plugin calculates number of incoming flows to designated IP address and detected UDP port at the start of the potential call. As the plugin detects flows with the voice call characteristic, then in the case of incoming video call the incoming flow will be not detected and the bidirectional pattern will be missing. Skype can utilize up to 1.5 Mbps in the case of HD video call so the upper bound for `bits per second` parameter was set to 2000000 bits.

```
  proto udp and duration > 296000 and pps > 10
and pps < 210 and bps > 8000
and bps < 2000000 and $filter_addr_port
```

If the number of incoming flows to the designated IP address and UDP port is equal to `0`, then the bidirectional property is missing.

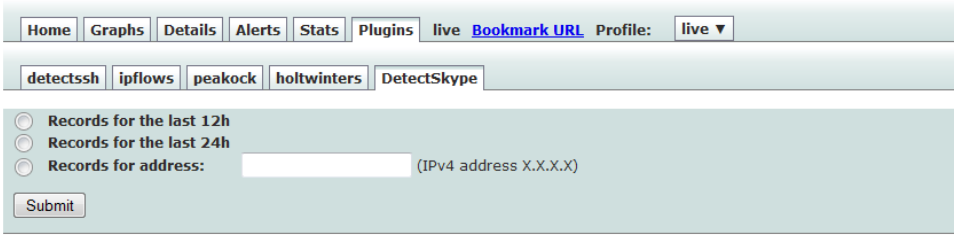


Figure 6.1: DetectSkype plugin frontend.

6.2 Frontend plugin

Frontend plugin is a component of the NfSen web interface. It implements two mandatory functions `DetectSkype_ParseInput` and `DetectSkype_Run`. `DetectSkype_ParseInput` is intended to parse possible form data and is called after selection in the plugins tab. `DetectSkype_Run` is the main function, it is up to it what will be displayed in the web browser and what parameters will be sent to the backend plugin.

There are several options to set how detected voice flows to display in the web browser, see Figure 6.1. User can choose to view stored records for the last 12 or 24 hours (these values are used to limit the number of lines in the output table). If selected, then a table with corresponding statistics for detected flows is displayed. Each line represents flow as described in subsection 6.1.2 UDP voice call properties. Continuously captured flow records from the same IP address and port to the same destination IP address and port are merged together – number of merged flow records is depicted by the column `Flows`. Description of table columns follows:

- **Start of the last flow** – Start time of the last record from the continuously captured flow records from particular IP address and port to particular IP address and port.
- **Duration** – Duration of the last captured flow record.
- **Proto** – Always UDP protocol as UDP voice calls are searched.
- **Src address** – Source address.
- **Src port** – Source port.
- **Destination address** – Destination address.

- **Dst port** – Destination port.
- **Avg packets** – Average number of packets (per flow record calculated from the merged records).
- **Avg pps** – Average packets per second value (per flow record calculated from the merged records).
- **Avg bps** – Average bits per second statistics (per flow record calculated from the merged records).
- **Avg bpp** – Average bytes per packet statistics (per flow record calculated from the merged records).
- **Flows** – Number of continuously captured and after that merged flows.
- **uiskype** – "Y" represents that connection from `Src address` to ui.skype.com was captured during the last 4 hours as described in subsection 6.1.1.
- **UDP uiPort** – UDP port that was detected around the time of above `uiskype` activity as described in subsection 6.1.3.
- **UDP 1stPort** – UDP port that was detected around the time the first flow record was captured as described in subsection 6.1.3.
- **Flows to 1stPort** – Number of incoming flows to `Src address` and `UDP 1stPort` as described in subsection 6.1.4.

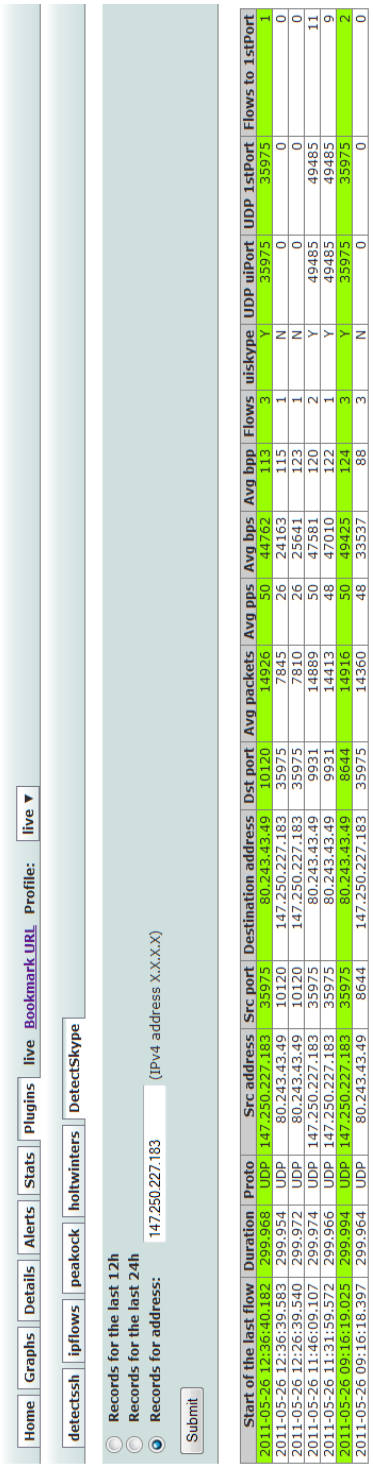
If `uiskype` equals "Y" and the `Src port` equals to `UDP uiPort` and `UDP 1stport`, then the line is marked as green as detected Skype voice call. If the `Flows to 1stPort` value is 0, then bidirectional pattern is missing and this flow is not Skype voice call. The value should be lower than 25 – Skype client can share a conference call with up to 25 people and group video calls can be between three or more clients (up to a maximum of 10) [20].

To display flows for particular IP address there is third option in the frontend plugin to display data statistics. Flow records are stored for two weeks and records in which given address figures as source address or destination address are displayed. In Figure 6.3, there are displayed flows in which IP address 147.250.227.183 figures. It is visible also how BitTorrent activity was filtered out by better choosing parameters for searching UDP ports. The most upper row shows us a voice call that has lasted for three

Records for the last 12h
Records for the last 24h
Records for address: (IPv4 address X.X.X.X)
Submit

Start of the last flow	Duration	Proto	Src address	Src port	Destination address	Dst port	Avg packets	Avg pps	Avg bps	Avg bpp	Flows	uiskype	UDP uiPort	UDP 1stPort	Flows to 1stPort
2011-05-27 08:34:46.535	299.984	UDP	61.169.82.175	41320	147.250.73.122	59258	15000	50	80004	200	1	N	0	0	0
2011-05-27 08:34:46.513	299.999	UDP	147.250.73.122	59258	61.169.82.175	41320	14979	49	79888	200	1	N	0	0	0
2011-05-27 08:33:14.097	299.953	UDP	84.123.165.208	4390	147.250.228.222	39478	14318	47	22443	58	1	N	0	0	0
2011-05-27 08:31:38.627	299.809	UDP	101.111.5.145	12316	147.250.95.25	32925	12370	41	17081	51	2	N	0	0	0
2011-05-27 08:27:06.560	299.971	UDP	76.209.107.78	33929	147.250.82.3	16065	15833	52	41587	98	1	N	0	0	0
2011-05-27 08:26:59.919	299.967	UDP	60.251.52.51	4367	147.250.215.92	4314	6459	21	20880	121	1	N	0	0	0
2011-05-27 08:24:46.528	299.983	UDP	61.169.82.175	41320	147.250.73.122	59258	15000	50	80004	200	1	N	0	0	0
2011-05-27 08:24:46.487	299.987	UDP	147.250.73.122	59258	61.169.82.175	41320	14495	48	77305	199	1	N	0	0	0
2011-05-27 08:20:32.918	299.961	UDP	147.250.74.81	60937	61.169.82.11	41656	15001	50	80007	200	5	N	0	0	0
2011-05-27 08:20:32.913	299.983	UDP	61.169.82.11	41656	147.250.74.81	60937	15000	50	80005	200	5	N	0	0	0
2011-05-27 08:17:06.538	299.953	UDP	76.209.107.78	33929	147.250.82.3	16065	14895	49	40137	101	1	N	0	0	0
2011-05-27 08:16:53.522	298.256	UDP	147.250.236.90	58089	83.56.188.53	30361	16058	53	22399	52	1	N	0	58089	251
2011-05-27 08:15:53.312	299.971	UDP	147.250.236.90	58089	91.92.216.180	53093	19994	66	27539	51	1	N	0	58089	251
2011-05-27 08:15:25.543	299.960	UDP	147.250.236.90	58089	88.237.207.49	57944	11880	39	16319	51	1	N	0	58089	251
2011-05-27 08:11:37.671	299.969	UDP	83.159.192.185	20410	147.250.95.25	52270	8628	28	21345	92	1	N	0	0	0
2011-05-27 08:08:08.515	299.949	UDP	147.250.236.90	58089	85.69.235.36	55034	14206	47	19739	52	1	N	0	58089	349
2011-05-27 08:07:53.740	299.985	UDP	61.169.82.175	24816	147.250.87.247	27726	15000	50	80004	200	5	N	0	0	0
2011-05-27 08:07:53.740	299.988	UDP	147.250.87.247	27726	61.169.82.175	24816	15000	50	80000	200	5	N	0	51216	0
2011-05-27 08:07:02.767	299.992	UDP	95.147.223.66	44576	147.250.95.25	32925	12429	41	17258	52	3	N	0	0	0
2011-05-27 08:06:52.499	297.229	UDP	147.250.236.90	58089	87.57.159.193	57014	13564	45	19000	52	1	N	0	58089	349
2011-05-27 08:06:02.361	296.382	UDP	147.250.236.90	58089	73.156.170.163	20877	19590	66	27153	51	1	N	0	58089	349
2011-05-27 08:06:01.961	299.557	UDP	13.151.139.169	17482	147.250.5.84	57190	15483	51	79889	193	2	N	0	0	0
2011-05-27 08:05:08.718	299.968	UDP	147.250.73.40	44866	86.78.160.41	16014	12064	40	16354	51	2	Y	0	0	0
2011-05-27 07:49:55.692	299.704	UDP	147.250.236.90	58089	150.105.157.48	44714	16644	55	23220	52	1	N	0	58089	58
2011-05-27 07:42:58.295	299.958	UDP	79.216.23.185	57882	147.250.87.99	55195	14999	50	80005	200	1	N	0	0	0
2011-05-27 07:42:58.267	299.997	UDP	147.250.87.99	55195	79.216.23.185	57882	15001	50	80001	199	1	N	0	0	0
2011-05-27 07:40:42.730	297.211	UDP	147.250.193.58	57337	85.177.66.28	44811	10604	35	34126	119	1	Y	57337	57337	3
2011-05-27 07:38:47.710	299.953	UDP	147.250.193.58	57337	37.220.46.115	19388	14736	49	75264	191	2	Y	57337	57337	2
2011-05-27 07:32:50.800	299.979	UDP	81.238.105.141	27841	147.250.95.25	32925	11704	39	16141	51	1	N	0	0	0

Figure 6.2: DetectSkype plugin frontend, first rows of statistics for the the last 24 hours.



Start of the last flow	Duration	Proto	Src address	Src port	Destination address	Dst port	Avg packets	Avg pps	Avg bps	Avg bpp	Flows	uiskype	UDP uiPort	UDP 1stPort	Flows to 1stPort
2011-05-26 12:36:40.182	299.968	UDP	147.250.227.183	35975	80.243.43.49	10120	14926	50	44762	113	3	Y	35975	35975	1
2011-05-26 12:36:39.583	299.954	UDP	80.243.43.49	10120	147.250.227.183	35975	7845	26	24163	115	1	N	0	0	0
2011-05-26 12:26:39.540	299.972	UDP	80.243.43.49	10120	147.250.227.183	35975	7810	26	25641	123	1	N	0	0	0
2011-05-26 11:46:09.107	299.974	UDP	147.250.227.183	35975	80.243.43.49	9931	14889	50	47581	120	2	Y	49485	49485	11
2011-05-26 11:31:59.572	299.966	UDP	147.250.227.183	35975	80.243.43.49	9931	14413	48	47010	122	1	Y	49485	49485	9
2011-05-26 09:16:19.025	299.994	UDP	147.250.227.183	35975	80.243.43.49	8644	14916	50	49425	124	3	Y	35975	35975	2
2011-05-26 09:16:18.397	299.964	UDP	80.243.43.49	8644	147.250.227.183	35975	14360	48	33537	88	3	N	0	0	0

Figure 6.3: DetectSkype plugin frontend, statistics for one IP address.

continuous flow records. That means that the start of the voice call was detected approximately at `12:26:40` (`Start of the last flow` minus 2 flow records back i.e. $2 * 5$ minutes). At that time we can observe incoming flow that indicates bidirectional character to the detected flow.

6.3 Testing the plugin implementation

Test calls were performed in July 2011 to test the proposed method for voice calls detection. As the plugin can detect voice calls longer than 5 minutes, duration of each test call was 17 minutes.

6.3.1 Voice call

In this test regular voice call was established. Computer *C1* was connected to the network of the Faculty of Informatics, Masaryk University. Configuration of *C1* was MS Windows 7, Skype version 5.3.0.111. Computer *C2* was located in Prague (Czech Republic), utilized 10/1 Mbps Internet connection and public IP address.

Call was established from *C1* to *C2* and lasted 17 minutes. The plugin detects 5 minutes lasting flow records with voice call characteristic, then 3 flow records should be detected as $3 * 5 < 17$. Voice call was successfully detected and average statistics are rewritten to the following table:

Source	C1
Src port	43029
Destination	C2
Dst port	35975
Avg pps	49
Avg bps	43725
Avg bpp	110
Flows	3
uiskype	Y
UDP uiPort	43029
UDP 1stPort	43029
Flows to 1stPort	1

6.3.2 Voice call from Android device

Voice call from the mobile phone (*M1*) with Android OS and Skype version 2.1.0.46 was established. Mobile phone was connected by WiFi to the

network of the Faculty of Informatics, Masaryk University.

Call was established from *M1* to *C2* and lasted 17 minutes. It should be detected as Skype for Android OS utilizes SILK voice codec in all Skype versions. Voice call was successfully detected and average statistics are rewritten to the following table:

Source	M1
Src port	38649
Destination	C2
Dst port	35975
Avg pps	48
Avg bps	33468
Avg bpp	87
Flows	3
uiskype	Y
UDP uiPort	38649
UDP 1stPort	38649
Flows to 1stPort	1

6.3.3 Conference call

Conference voice call from *C1* with 3 other Skype clients was established. Computer *C3* was located in Olomouc region (Czech Republic), utilized 2000/500 kbps Internet connection and had non-public IP address. Computer *C4* was located in Ostrava (Czech Republic), utilized 4/1 Mbps Internet connection and had public IP address.

Call was established from *C1* to *C2* at first. *C3* and *4* were immediately invited to the conference call after that. The call lasted 17 minutes. It should be detected three flows – to *C2*, *C3* and *C4*. Three voice calls were successfully detected and average statistics are rewritten to the following table:

Source	C1	C1	C1
Src port	43029	43029	43029
Destination	C2	C3	C4
Dst port	35975	12817	27323
Avg pps	48	50	47
Avg bps	44250	49047	47844
Avg bpp	114	124	121
Flows	3	3	3
uiskype	Y	Y	Y
UDP uiPort	43029	43029	43029
UDP 1stPort	43029	43029	43029
Flows to 1stPort	3	3	3

6.3.4 Voice call during BitTorrent activity

Call was established from *C1* to *C4* and lasted 17 minutes. During the call the torrent file with installation image of Ubuntu OS was downloaded by uTorrent BitTorrent client [21]. Torrent file was downloaded approximately from 30 seeders concurrently. The purpose was to test if the plugin correctly detects call if another application that utilizes UDP probes runs concurrently. Voice call was successfully detected and average statistics are rewritten to the following table:

Source	C1
Src port	43029
Destination	C4
Dst port	27323
Avg pps	48
Avg bps	46288
Avg bpp	120
Flows	3
uiskype	Y
UDP uiPort	43029
UDP 1stPort	43029
Flows to 1stPort	1

6.3.5 Video call

Video call with 20 frames per second and 640 x 480 pixels resolution from *C1* to *C2* was established and lasted 17 minutes. *C1* used integrated web-

cam with 1.3 MP resolution. It was awaited that the video call will not be detected because its traffic characteristic is different from regular voice call traffic. The video call was not detected but average statistics were manually gained from flow records stored by collector and rewritten to the following table:

Source	C1
Src port	43029
Destination	C2
Dst port	35975
Avg pps	97
Avg bps	537779
Avg bpp	694

6.3.6 Conclusion from test results

We can conclude from the performed tests that the plugin correctly detects UDP voice calls. As the Skype versions for Linux, Mac OS X, Android, Symbian and iOs also use SILK codec as the voice codec, we can rely on performed tests in a way that UDP voice calls established from these Skype versions will be correctly detected in the same manner. Test performed concurrently with BitTorrent activity pointed out that the detection method is reliable in this case.

Chapter 7

Conclusion

Network traffic of the Skype application has been studied in this thesis. Some Skype payload patterns were confirmed for newer versions of application as patterns were published only for earlier versions of Skype client, and additional patterns were newly presented. Our contribution was verification of proposed papers related to this field for newer versions of application – 4.1, 4.2 and 5.3 – and discovering new facts that were not published so far as we know. We have proposed e.g. new facts about TCP signalling traffic between Skype client and neighbour supernode, and about regular TCP traffic to `ui.skype.com`. From the presented facts it is possible to create data payload analyzer that will recognize Skype traffic quite efficient as similarly mentioned in [1].

Further in this thesis the plugin for NfSen with its backend and frontend part was presented. As NetFlow records do not offer data payload for inspection, approach different to payload analysis was used. Limitations are that it is quite difficult to recognize TCP relayed traffic as several connections with no constant traffic and with dynamical source ports are present, so it was not possible to consider this kind of detection. Characteristics of UDP voice calls were used for implementation. Test calls subsequently confirmed detection capabilities of the plugin.

Bibliography

[1] Davide Adami, Christian Callegari, Stefano Giordano, Michele Pagano, and Teresa Pepe. A Real-Time Algorithm for Skype Traffic Detection and Classification. In Sergey Balandin, Dmitri Moltchanov, and Yevgeni Koucheryavy, editors, *Smart Spaces and Next Generation Wired/Wireless Networking, 9th International Conference, NEW2AN 2009 and Second Conference on Smart Spaces, ruSMART 2009, St. Petersburg, Russia, September 15-18, 2009. Proceedings*, volume 5764 of *Lecture Notes in Computer Science*, pages 168–179. Springer, 2009.

[2] C. Allen and T. Dierks. RFC 2246 – The TLS Protocol Version 1.0. `http://tools.ietf.org/html/rfc2246`, January 1999.

[3] Duffy Angevine and A. Nur Zincir-Heywood. A preliminary investigation of skype traffic classification using a minimalist feature set. In *ARES*, pages 1075–1079. IEEE Computer Society, 2008.

[4] Salman Baset and Henning Schulzrinne. An analysis of the skype peer-to-peer internet telephony protocol. In *INFOCOM*. IEEE, 2006.

[5] Dario Bonfiglio, Marco Mellia, Michela Meo, Dario Rossi, and Paolo Tofanelli. Revealing skype traffic: when randomness plays with you. In Jun Murai and Kenjiro Cho, editors, *Proceedings of the ACM SIGCOMM 2007 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications, Kyoto, Japan, August 27-31, 2007*, pages 37–48. ACM, 2007.

[6] B. Claise. RFC 3954 – Cisco Systems NetFlow Services Export Version 9. `http://tools.ietf.org/html/rfc3954`, October 2004.

[7] Sven Ehlert and Sandrine Petgang. Analysis and Signature of Skype VoIP Session Traffic. In *Franunhofer FOKUS Technical Report NGNISKYPE-06b*, July 2006. Berlin, Germany.

[8] H. Astrom, J. Spittka and K. Vos. RTP Payload Format and File Storage Format for SILK Speech and Audio Codec. `http://developer.`

skype.com/resources/SILK_RTP_PayloadFormat.pdf, August 2010.

[9] J. Rosenberg, R. Mahy, P. Matthews and D. Wing. RFC 5389 – Session Traversal Utilities for NAT (STUN). http://tools.ietf.org/html/rfc5389, October 2008.

[10] Chun-Ming Leung and Yuen-Yan Chan. Network Forensic on Encrypted Peer-to-Peer VoIP Traffics and the Detection, Blocking, and Prioritization of Skype Traffics. In *WETICE*, pages 401–408. IEEE Computer Society, 2007.

[11] Jonathan Rosenberg. Skype and the Network, Technical Advisory Process Workshop on Broadband Network Management. http://www.openinternet.gov/workshops/docs/ws_tech_advisory_process/Skype-FCC.PPTX. Visited 27. 5. 2011.

[12] S. Jensen, K. Soerensen and K. Vos. SILK Speech Codec draft-vos-silk-01. http://developer.skype.com/resources/draft-vos-silk-01.txt, March 2010.

[13] Andreas Thomann. Skype - A Baltic Success Story. https://infocus.credit-suisse.com/app/article/index.cfm?fuseaction=OpenArticle&aoid=163167&coid=7805&lang=EN, September 2006. Visited 21. 2. 2011.

[14] Skype – The Big Blog – 30 million people online on Skype. http://blogs.skype.com/en/2011/03/30_million_people_online.html. Visited 27. 4. 2011.

[15] Skype grows FY revenues 20%, reaches 663 mln users. http://www.telecompaper.com/news/skype-grows-fy-revenues-20-reaches-663-mln-users, March 2011. Visited 27. 4. 2011.

[16] IT Administrators guide. http://download.skype.com/share/business/guides/skype-it-administrators-guide.pdf. Visited 21. 2. 2011.

[17] NFDUMP. http://nfdump.sourceforge.net. Visited 25. 5. 2011.

[18] NfSen – Netflow Sensor. http://nfsen.sourceforge.net. Visited 25. 5. 2011.

[19] SILK Data Sheet. http://developer.skype.com/resources/SILKDataSheet.pdf. Visited 21. 2. 2011.

[20] Free Skype calls and cheap calls to phones – Skype. http://www.skype.com. Visited 21. 2. 2011.

[21] uTorrent – A (very) tiny BitTorrent client. http://www.utorrent.com. Visited 12. 7. 2011.

[22] Help for Skype: How much bandwidth does Skype need? https://support.skype.com/en/faq/FA1417/How-much-bandwidth-does-Skype-need. Visited 7. 4. 2011.

[23] Kazaa – Wikipedia, the free encyclopedia. http://en.wikipedia.org/wiki/Kazaa. Visited 21. 2. 2011.

[24] NetFlow – Wikipedia, the free encyclopedia. http://en.wikipedia.org/wiki/Netflow. Visited 20. 5. 2011.

[25] Skype – Wikipedia, the free encyclopedia. http://en.wikipedia.org/wiki/Skype. Visited 12. 5. 2011.

[26] SVOPC – Wikipedia, the free encyclopedia. http://en.wikipedia.org/wiki/SVOPC. Visited 20. 5. 2011.

[27] Wireshark. Go deep. http://www.wireshark.org. Visited 12. 7. 2011.