Final Report

NoAH

A European Network of Affined Honeypots

Design Study

implemented as

Specific Support Action

Contract number: *011923*

Project coordinator: *FORTH*

Project website: *http://www.fp6-noah.org/*

Project Duration: *42* months from *01/04/2005* to *30/09/2008*

Project funded by the European Community
under the "Structuring the European Research Area" specific programme
Research Infrastructures Action

CONTENTS

A. ACTIVITY REPORT

A.1 *Project Summary*

A.1.1 Summary description of project objectives

Over the past decade, we have witnessed an increasing number of cyber-attacks on the Internet. Viruses, trojans, and other types of malicious software are discouraging the effective use of the Internet and are crippling the global IT infrastructure. In recent years, outbreaks have demonstrated that attackers already have the capability to compromise a large part of the Internet within minutes[1]. To make matters worse, laboratory studies suggest that it is possible to craft carefully designed attacks that can compromise tens of thousands of Internet-connected computers within seconds[2], and that the damage of such an attack could reach more than US$50 billion[3]. At these time scales, human reaction to an attack may be impossible. To successfully combat such threats, we need an infrastructure to assist in detecting and containing such attacks.

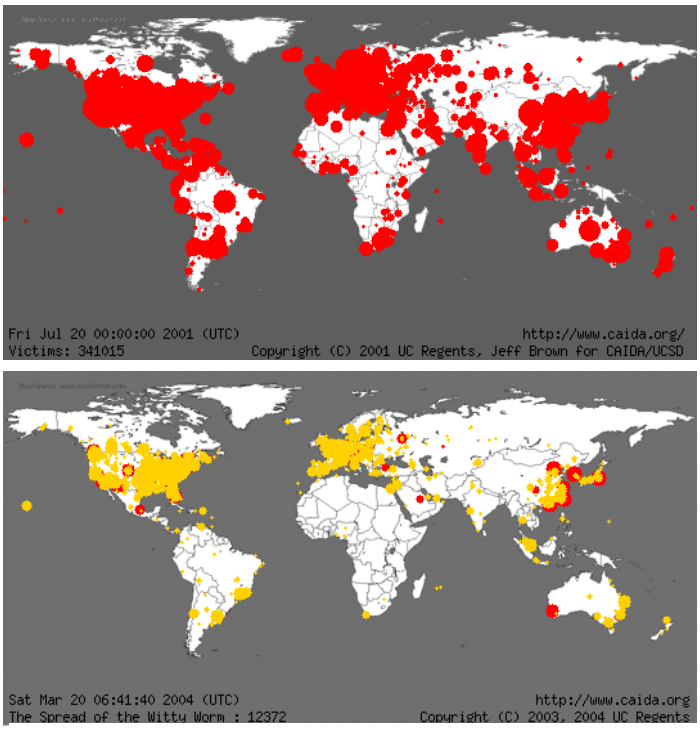


Figure 1 – Extent of the spread of the Code Red and Witty worms after 24 hours and 2 hours respectively

The main goal of the NoAH project has been to produce a design study and perform the necessary technical work towards the development of an infrastructure for security monitoring based on honeypot technology. Honeypots are computer systems that do not provide real production services. Instead, they are *intentionally vulnerable*, and at the same time *closely monitored systems*, that wait to be compromised by attackers. Once hit, honeypots can be used to analyze attacks: *where* did the attacker come from, *how* did he enter the system, *what* did he try to do after entering, etc. We expect that by gathering and correlating data from geographically dispersed honeypots, NoAH will be able to detect cyber-attacks before they have the chance to do any major damage. To achieve this, NoAH has explored the potential for automated generation of attack signatures or other containment-related information that may be used by reactive security systems. Additionally, NoAH has the goal of facilitating a distributed security analysis infrastructure for ISPs, NRENs and security organizations.

More specifically, NoAH had the following key objectives:

- **Design** an infrastructure of affined honeypots that will gather and correlate data about attackers, their methods, and actions on the Internet,
- **Develop techniques for the automatic identification** of novel attacks and for the automated generation of corresponding signatures, enabling the effective containment of the spread of an attack,

---

[1] http://www.caida.org/publications/papers/2002/codered/

[2] http://www.icir.org/vern/papers/topspeed-worm04.pdf

[3] http://www.icir.org/vern/papers/worst-case-worm.WEIS04.pdf

- Install and **operate a pilot NoAH infrastructure** to demonstrate the effectiveness and utility of a full-scale NoAH infrastructure,
- **Provide sanitised attack information** to the security research community on a **pilot basis**. Such a repository of information has the potential of boosting research and development in the area of attack detection and containment,
- **Extend the participation** in the infrastructure to ISPs, NRENs, and CERTs outside the NoAH consortium,
- **Disseminate** the results of the project to researchers and security analysts.

A.1.2 Contractors involved

The consortium of the NoAH project consisted of the following primary contractors: Vrije Universiteit (The Netherlands), TERENA (The Netherlands), FORTHnet (Greece), DFN-CERT (Germany), ETH Zürich (Switzerland), Virtual Trip LTD (Greece), Alcatel-Lucent (France). The project has been coordinated by FORTH-ICS (Greece).

Figure 2 – The NoAH consortium

The NoAH Project benefited greatly by the partner's complementary expertise and their deployed facilities. Indeed, the NoAH Consortium is a mixture of leading organizations – in their area of expertise – that each contributed to a different aspect of the project.

- *FORTH* is the largest Greek State R&D Center, consisting of seven Institutes in different disciplines. Its Institute of Computer Science has a relatively long history and recognized tradition in conducting basic and applied research, developing applications and products, providing services, and playing a leading role in Greece and internationally, in the fields of Information and Communication Technologies.
- *Vrije Universiteit* is a broad-based university in which education and research are closely interrelated. It enjoys a strong international reputation in all kinds of research areas. Security

is a prominent research area of the Computer Systems Group at Vrije Universiteit. The group has a track on security research, ranging from secure operating systems (i.e., Amoeba), via digital rights management to the design of secure agent platforms (i.e., Mansion).

- *ETH Zürich* is a science and technology university with an outstanding research record. Consisting of fifteen departments, they focus mainly on the engineering sciences and architecture, system-oriented sciences, mathematics and natural sciences areas and carry out research that is highly valued worldwide. There are twenty-one Nobel Laureates affiliated with ETH Zürich.
- *TERENA* is the association of the European research and education computer networks. Their aim is to foster the development of Internet technology, infrastructure and services to be used by the research and education community, by providing a forum of collaboration for their members.
- *FORTHnet* is the first commercial ISP in Greece one of the most dynamically developing Greek companies in the Telecommunications and Internet sector. With their continuous investment in new technologies and infrastructure they are able to develop integrated telecommunications proposals and solutions addressed to home users and businesses.
- *DFN-CERT* is a highly regarded competence center for Internet and computer security including infrastructure services. With their extended experience with both open source and software and hardware, they provide tailored insights and assist organisations in the selection of computer and networking security products.
- *Alcatel-Lucent* is the first truly global communication solutions provider for telecoms and enterprises. Their vision is to enrich people's lives by transforming the way the world communicates.
- *Virtual Trip LTD* is an emerging network-centric software development company founded in 2000. Their expertise ranges from hardware design and system software, to large-scale information systems.

A.2 *Summary of work performed*

A.2.1 Methodologies used

The NoAH project brought together a leading group of security researchers with forward-looking industry stakeholders with the common agenda of designing an infrastructure for addressing the growing challenge of cyber-attack detection, prevention, and containment. From a scientific point of view, the methodologies used are heavily rooted in experimental computer science, with the bulk of the research focusing on the design and analysis of experimental artefacts, and the production of experimental evidence in support of any hypothesis made during the course of the project.

One distinguishing feature of the NoAH strategy is that it went far beyond the typical approach of analyzing experimental artefacts in the form of "throw-away" prototypes. Significant effort was put in building robust systems, such as *Argos*[1] and *Honey@Home*[2], going through multiple design iterations and release cycles, to bring the systems to the maturity needed for the purposes of a research infrastructure design study. Similarly, the evaluation of the technology produced by NoAH was carried out in real-world settings thanks to the very same large-scale experimental infrastructure that NoAH was tasked to build. NoAH sensors are hosted by many organizations, both within and outside the project consortium, and the testbeds and data feeds had great influence on the overall design study. This is significantly different from the usual small-scale testbeds where systems are typically evaluated and the role of the industrial partners in this direction was crucial.

Furthermore, NoAH engaged aggressively with stakeholders in the cyber-security space, to refine the initial assumptions and work plan and obtain feedback during the later stages of the project. NoAH worked with various security infrastructure initiatives, including *Leurre.com*, the *Honeynet Alliance*, *ENISA*, and other European research groups working on similar problems such as the *Scriptgen* team.

Finally, NoAH was purposely aggressive in reaching out to the wider security community for disseminating the project results. Within the research world, NoAH **was instrumental in the early formation of a European experimental systems security community** for the first time, initially through the NoAH workshops and subsequently through the organization of new high-visibility workshops and conferences such as **ACM EuroSec**[3] and **EC2ND**[4]. In addition to visibility purely within the research world, NoAH was promoted through channels such as ENISA- and TERENA-sponsored events, as well as European-related magazines and newsletters. This was an important component for feedback on the research infrastructure envisioned as well as for ensuring uptake of project results.

EC2ND
European Conference on Computer Network Defense

[1] http://www.few.vu.nl/argos/

[2] http://www.honeyathome.org

[3] http://www.cs.vu.nl/eurosec08/

[4] http://2008.ec2nd.org/ec2nd/597-EE.html

A.2.2 WP0 Achievements and end results

Work-Package WP0 (*Requirements Analysis and State-of-the-Art*), which started on April 1st 2005, dealt with (i) **investigating the State-of-the-art** in honeypot technology and research and (ii) defining the requirements for the NoAH system. These two objectives were addressed in Task WP0.1 and Task WP0.2, which were summarized in deliverables D0.1 and D0.2 respectively.

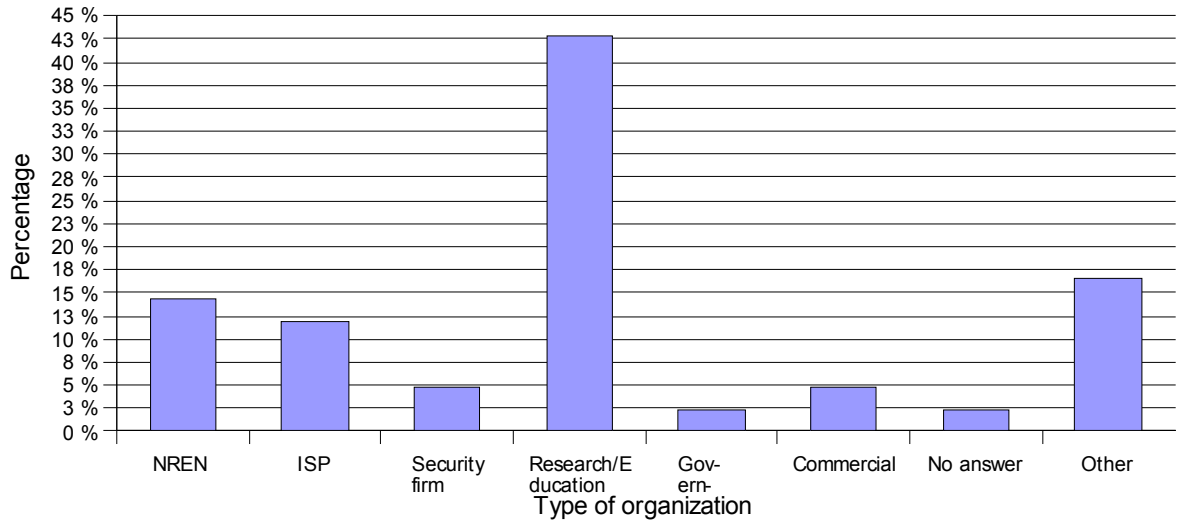


Figure 3 – Responders to the NoAH survey

In the context of Task WP0.2 (*Requirements Collection and Analysis*), the partners produced a questionnaire to gather the views of the security community on honeypot-related technologies and their future. While the survey was, in principle, open for everyone to answer, we proactively advertised it to selected stakeholders through mailing lists and the extended professional network. The responses suggested an increasing importance for honeypot technologies in the computer security arena and validated early assumptions and thoughts on the NoAH design.

More specifically, an invitation for participating in the NoAH survey was sent to NRNs (National Research Networks), Universities, ISPs, and other commercial and governmental organizations, as well as to individual persons involved in security monitoring issues. There were totally 42 responses for this questionnaire in a four-month period (June 2005 – September 2005). Figure 1 illustrates that out of the 42 organizations that responded to the questionnaire, 6 of them, or 14%, are NRNs, 5 were ISPs (12%), 18 are research organizations or universities (43%), one was a government organization (2%) and two were are commercial organizations (4%). There were also responses from one Telecommunications company, one RREN, one Open Source Security Solution Developer, one FFRDC, and one Technology Integrator. In addition to the questionnaire, there was a parallel discussion both via email and teleconference about what the NoAH project should be focusing on, and all the partners had the opportunity to state their position on this matter.

The following overall observations were drawn from the received responses on the questionnaire, from the discussion between consortium members, and from the broader survey on projects related to NoAH:

- The NoAH platform should be **easy to install, configure and update**. The installation and

configuration process should be as self–explanatory as possible and easy to follow by anyone. Expansions of the infrastructure with new features at low cost must be predicted. NoAH honeypots may be implemented to support various operating systems. This variety of operating systems should not cause any malfunction to the honeypot infrastructure.

- The NoAH honeypots **should not interfere with an organization's network resources** and should not in any case affect their network's bandwidth and introduce extra unnecessary traffic. The administrator of a honeypot should be able to define the level of interaction for the specific honeypot, as well as its security level. They should also not attack other honeypots or cause any malfunction to other network facilities within an organization or to an outsider. However they should be easy to install/disconnect to one's network (portability). In case of a hardware malfunction on the honeypot's resources, the honeypot should be isolated so as not to create problems for the network to which it is connected. In case attackers realize that they are in a honeypot, it should become difficult or even impossible for them to take over and control the honeypot.
- The NoAH infrastructure should be able to **detect various kinds of attacks**, such as virus/worm spread, intrusion attempts, bots attacks, etc. It should also be able to recognize new threats at the beginning of their spread, such as polymorphic and metamorphic viruses, or new critical vulnerabilities that come up that target various resources.
- The reaction of the NoAH infrastructure in case of an attack detected should be quick and accurate. The **number of false positives should be minimized**, if not completely eliminated, in order for the system to be trustworthy. Honeypots may need to communicate with each other in order to produce valid alerts quickly and to reduce false positives to the minimum.
- Once an attack has been detected, NoAH should send an inform **alert on the specific attack**. Also, a web site with information about those attacks should be developed. A mailing list might need to be developed, as well. Other potential actions that can be taken by the infrastructure should be described by a NoAH policy. The NoAH infrastructure should produce well-formatted signatures that are clear to various recipients (network technicians and administrators). Information that should be provided is: IP address or DNS name of the attacker, geographical information about them and information on the methods used by the attacker and the kind of the application/service exploited by the attacker.
- A **policy for the cooperation between NoAH and external partners should be clarified** (requirements for becoming a partner, the rights/obligations for each partner), as well as for the actions that should be taken in case of misuse or abuse of the infrastructure by any NoAH partner. Furthermore, a policy for the usage of the produced data should be clarified, and should also cover rights and duties of participating organizations that host a honeypot. The NoAH infrastructure should take into account issues about **data logging and user privacy**. Logging is risky with various data protection acts in different constituencies. It is extremely dangerous for any third parties to retain this data without tight restrictions on use and access. A policy for the data export of the honeypot nodes should exist, defining which data can be exported from the nodes and the various protection options for it. NoAH honeypots should not reveal themselves to the attackers. Mechanisms should exist for self-protection of a honeypot in case an intruder understands a honeypot.

The results of the survey were analyzed and are presented in more detail in deliverable D0.2. FORTHnet created a separate web page[1] on the NoAH web site featuring a selected subset of the survey results and stored a dump of the database containing the results for future use.

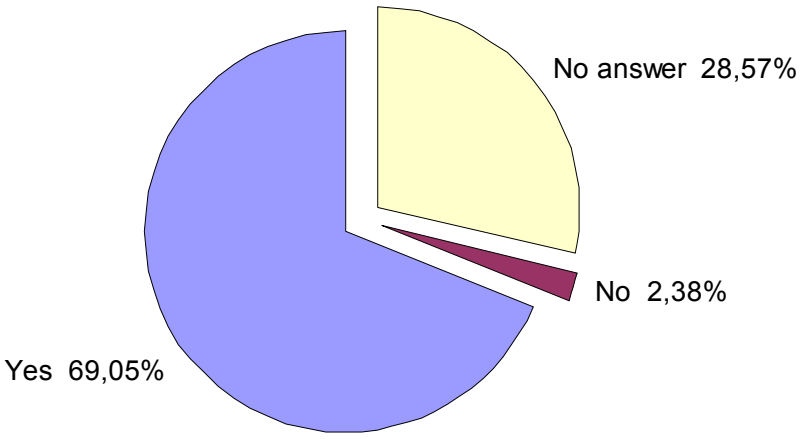


Figure 4 – 69% of the participants indicated that they are willing to share attack-related data

In addition to the questionnaire, during the first months of the project and within the context of task WP0.1, the project partners explored the state-of-the-art in honeypot technology and in cyber-attack detection more broadly. The results of this study were presented and evaluated in Deliverable D0.1, and helped in identifying the gap that NoAH aims to fill. This included research efforts, the results of which have been published in scientific conferences and journals, as well as existing honeypot initiatives and infrastructures from both the research community and the industry. Overall, nine projects devoted to the deployment of honeypot networks or networks of distributed intrusion detection systems (IDS) were analyzed.

Due to the variety of different approaches, the aims of the projects differ very much. The first major difference is whether the projects focus on the honeypot technology or on the architecture between the honeypots or IDS. The later applies to the DOMINO Overlay system that provides an open architecture for low-interaction honeypots and IDS. This architecture is comprised of a hierarchical structure of a large number of nodes that can host different types of low-interaction honeypots and IDS. Because of its structure it does not include single points of failure and is therefore resistant against denial-of-service attacks. In addition, low-interaction honeypots and IDS can easily be integrated into this architecture.

All other projects concentrate on honeypot technology. The eCSIRT and LEURRE.COM projects deploy broad networks of low-interaction honeypots. Primary objective is to get an overview over the current malicious activity. For the accuracy of the statistical data it is critical to deploy a broad distributed network of honeypots. A different low-interaction approach is proposed by the HoneyTank project. In contrast to the previous two projects, traffic is redirected by routers and relayed to an IDS that processes all connections. Target of this project is the detection of worms and automated attacks in ISP-scale networks. A major advantage of all three projects is the high scalability of the number of sensors. Thus, additional sensors can be easily integrated into the network without larger effort in maintenance. Although all projects include a single point of failure

[1] http://www.fp6-noah.org/survey/

given by a central database server or IDS, these components could be replaced by a distributed component.

The other projects deploy high-interaction honeypots. Aims of these projects vary from the capture of malicious software to the awareness of new worms or zero-day exploits. In addition, different solutions for high-interaction honeypots are proposed. Microsoft's Vigilante project modifies the operating system so that common types of attacks can be detected. The proposed approach is especially advantageous in the identification of attacks at a very early stage. In detail, intrusions are detected as early as they are able to inject untrusted data into the execution control flow of the honeypot in a specific way. A similar approach is introduced by the Minos project. Since the primary exploit vector is prevented, the attack is unable to succeed and it is very difficult for an attacker to evade this technique. As a consequence, even polymorphic worms are reliably detected which is in contrast to common network-based approaches. In addition, these approaches allow the generation of signatures that detect attacks very reliably and are not prone to false positives. Because of these advantages, the accuracy of this type of signature can be assumed to be superior to the common signatures of network based IDS. Moreover, host–based intrusion detection can be used to complement the application of widespread network based IDS or behavioural approaches to generate more accurate signatures and to exclude false warnings. However, all types of signatures allow the use of similar algorithms for an automatic generation. A complementary approach is the use of sequences of system calls for intrusion detection and for producing signatures. The fundamental idea is that most exploits use system calls to control the compromised system. Thus, malicious activity can be detected by tracking the system calls a process invokes for malicious activity (e.g., by *systrace*). In addition, this data can be used to produce complementary host–based signatures.

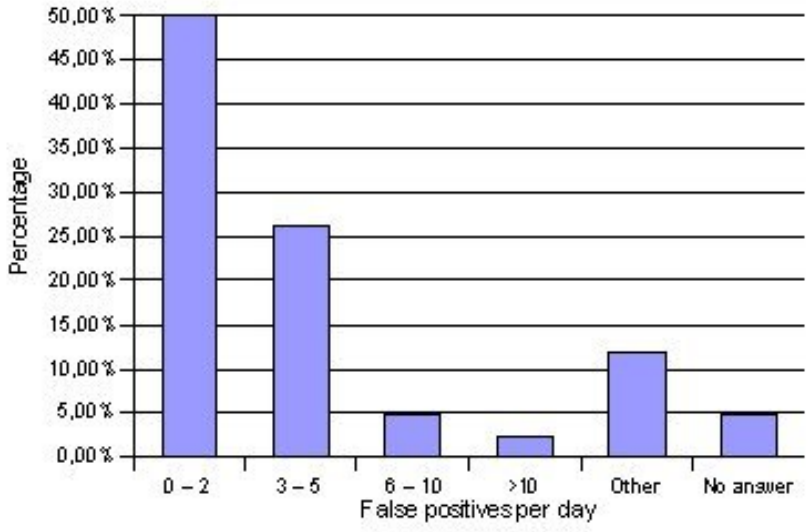


Figure 5 – Participants indicated that they would like to have a very small number of false positives in the attack detection mechanisms used.

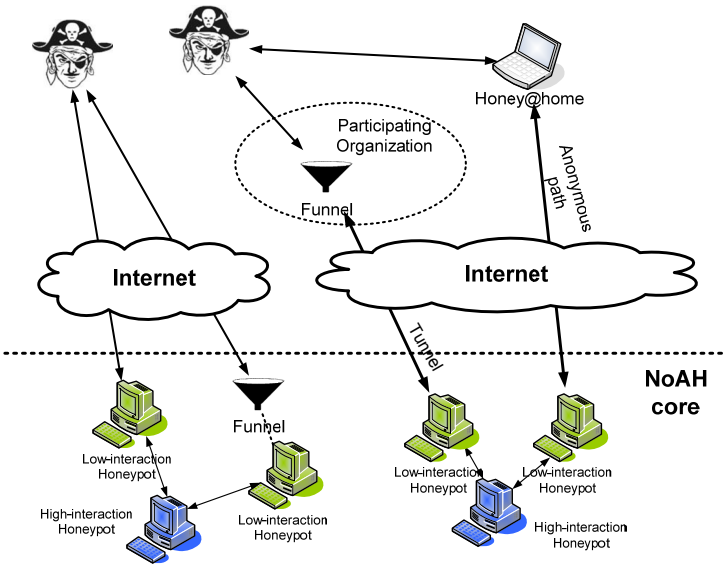
A principal disadvantage of the deployment of high-interaction honeypots is the low scalability of the number of honeypots. This is due to the nature of this type of honeypots which allows the attacker to compromise the system. As a consequence, significant effort has to be spent to limit the exposure of attacks that may originate from the compromised honeypots. An additional

drawback is that a compromised honeypot must be reinstalled in order to observe subsequent attacks that are independent of the initial attack. Even if the effort can be alleviated using virtual machines, the honeypot is still unusable while reinstalling. Therefore, if the locations of the honeypots are known, denial of service attacks against them can very easily be executed by continuous attacks. As a consequence, the placement of the honeypots must be carefully hidden and known surveillance methods should be taken into account.

However, it was identified that only high-interaction technologies would allow NoAH to achieve its goals. To cope with these challenges concerning the low scalability of high-interaction honeypots, hybrid approaches have been proposed that introduce a two-step structure by combination of low and high-interaction components. In the first step, all traffic is processed by low-interaction components (e.g. *honeyd*) that allow covering a broad range of IP addresses. In the second step, selected connections are relayed to high-interaction components where they are analysed. Thus, hybrid approaches try to combine the advantages of low- as well as high-interaction components. It was concluded that to establish a large sensor network following the aims of the NoAH project, it would be advantageous to distribute low-interaction components on a large scale relaying selected traffic to a smaller network of high-interaction honeypots where only unseen attacks have to be analysed.

A.2.3 WP1 Achievements and end results

WP1 officially started on July 1st 2005, and dealt with (i) defining the overall architecture of the NoAH infrastructure, including a containment environment for cyberattacks, and (ii) developing novel methods for attack detection and signature generation.



The design of the NoAH architecture is described in the deliverable D1.1. The NoAH architecture blends in readily-available technologies, such as *honeyd*[1], with customized components and new ones, such as *Honey@Home*. The main goal here was to combine low and high interaction honeypots, and to be inclusive at the technology and organization level. The architecture allows external organizations to participate in the NoAH infrastructure, enabling it to evolve to a Pan-European infrastructure in the future. More importantly, with the proposed "Honey@Home" component, the architecture allows individual (even home) users to participate in the infrastructure, and empowers ordinary citizens in the fight against cyber-attacks. The participation of individuals in the NoAH infrastructure can be proven valuable both in terms of attack detection efficiency, and in terms of social impact.

In terms of structure, NoAH relies on lightweight traffic redirectors that feed Argos, the NoAH containment environment, with potentially malicious traffic. To maximize coverage in terms of monitored IP address space beyond the IP addresses assigned to low- and high-interaction honeypots, NoAH developed two novel mechanisms: (i) *funneling and tunneling* and (ii) *Honey@Home*. These mechanisms enable small business and even home users to participate in tracking cyber-attacks by donating some of their unused IP addresses. Indeed, there is unmonitored black space that can be used for NoAH purposes at universities, institutions, public bodies and even homes. Being able to monitor this space will both broaden our view on what happens to the Internet and make it very difficult for attackers to identify NoAH honeypots and "blacklist" them.

The "funneling and tunneling" mechanism is geared towards organizations owning large batches of unused IP addresses. It enables them to participate in the NoAH project by redirecting traffic which reaches their unused IP address space to the NoAH core. It is able to redirect traffic from a whole batch of IP addresses with the installation of only one piece of software. On the other hand, Honey@Home targets smaller organizations and home users that only have a few spare unused IP addresses. Inspired from other "@home" projects, like SETI@home, Honey@Home is a simple "click and play" tool. It gives the ability to all kind of users, from naïve to experts, to install a honeypot at their personal computer with almost zero overhead and zero maintenance cost. Honey@Home listens to an unused IP address and forwards all its traffic to the NoAH core for further processing. It runs in the background and the user is able to view statistics about its activity

[1] http://www.honeyd.org/

and get a sense of what traffic is destined to his black space. Honey@Home offers the project the potential to have a *social impact*, since it allows individuals to get directly involved with NoAH and empowers them to help in the combat against future cyber-threats.

In terms of design, every Honey@Home client is responsible for a single dark IP address or the unused port space of the machine it is installed on. All the traffic received by the client is tunneled to the centralized honeypots of Honey@Home core through the TOR[1] anonymization network. The TOR network provides all the desired anonymity for both Honey@Home users and honeypots. Responses coming from the core are injected by Honey@Home to the network so as to reach the originators of the traffic. Honey@Home clients are connected through an SSL connection to the SSL server component. The SSL connection passes through the TOR anonymization network. The server component handles the client connections and is responsible for validating users and forwarding their packets to honeypots. Users are validated by supplying a key to the SSL server. The key is obtained after the user registers at the official website of Honey@Home. All registration information and user keys are stored in a MySQL database. After the user is validated, her packets are sent to honeypots and responses from honeypots are sent back to the user.

We run both low- and high-interaction honeypots to handle user's packets. For low-interaction honeypots, honeyd is used. Honeyd is a very popular and lightweight system with many interesting properties, such as network stack emulation. We use honeyd as a first filter to get rid of uninteresting traffic, such as TCP connections that do not complete the three-way handshake or attacks that can be easily emulated, for example SSH brute-force attacks. All the other traffic is forwarded to high-interaction honeypots. The forwarding is performed by a hand-off mechanism we implemented inside honeyd. Low-interaction honeypots create a new connection with the high-interaction one and send all the application content they receive. The handoff is based on destination port. For example, if an attacker wants to connect to port 445 or 139, the connection is forwarded to a high-interaction honeypot emulating the Windows XP operating system. As the choice of high-interaction honeypot is static, this means we may lose attacks for services that run on both platforms. Examples of such services are web servers (Linux Apache or Windows IIS), SMB sharing (Linux SAMBA or Windows sharing) and many more. However, our choice is currently made based on popularity of applications in terms of users and attack instances. For our prototype system, we emulate a limited number of services and applications and more specifically, Linux SMTPd, Microsoft Internet Information Server, MS-SQL server and the default Windows services.

The ideas behind Argos and its design are detailed in deliverable D1.3. Argos is capable of running a vulnerable operating system and its services and detecting attacks against it through a form of dynamic taint analysis. Because Argos does not require the traffic that arrives at it to be considered suspicious in the first place, it is capable of detecting zero-day attacks, achieving one of the primary objectives for the NoAH infrastructure. Argos relies on the QEMU emulator to enable it to detect remote attempts to compromise the emulated guest operating system. QEMU has the advantage of being an open source solution and therefore is easier in tweaking and extending than a closed proprietary solution. Also QEMU compares favourably to similar open source solutions like BOCHS, because it is significantly faster. Using dynamic taint analysis Argos tracks network data throughout the processor's execution and detects any attempts to use them in a malicious way. The Argos prototype source was publicly released in December 2005, with the launch of the Argos web

[1] http://www.torproject.org/

site: http://www.few.vu.nl/argos/. The public release of Argos was well received by the honeypot community, with references in several web pages and blogs.

Once a new attack has been identified, part of the NoAH design deals with generating appropriate containment signatures for it. The leader and also major contributor for this task was ETHZ. ETHZ started the work on this task with a comprehensive review of existing methods for attack detection and signature generation. Furthermore, ETHZ identified general design goals and the therewith linked challenges and initiated a discussion about the focus of NoAH. After discussions on what type of signatures NoAH should focus on, the conclusion was that NoAH should primarily focus on host-based attack detection because it is better suited for identifying polymorphic attacks and is more likely to provide us with zero false positive rates. Network-based signatures were not ignored though. Since there is no network-based signature type that can handle all kind of polymorphic attacks with an acceptable false positive rate, the concept of *meta-signatures* that combines multiple existing network-based signature types was proposed by ETHZ. Finally, two new attack detection and signature generation system concepts were developed. The first is based on network traffic analysis while the other is based on replaying potential attacks and analyzing their control-flow. The NoAH signature generator was designed and implemented as a *framework*, featuring a *plug-in structure* and a *template mechanism*. It also features a logger component and load balancing for efficient operation on multi-core systems. ETHZ started with the implementation of the *connection tracker* component of the framework and continued with the development of several plug-ins. The main purpose of the framework is to track the communication of a high interaction honeypot and to provide precise state and history information on demand (e.g. if Argos reports an attack). The level of detail in the information supplied by the framework depends on the presence of the appropriate plug-ins for the network/transport/application layer.

The integration activity (WP1.4) focused on putting the designed components together. Initially, the consortium discussed and decided on the exact interfaces between the developed components. However the work was not limited to components designed within the consortium: externally developed components and technologies were also discussed for integration with the architecture either as independent components (e.g. *Nepenthes*) or as services provided to the core components (e.g. *TOR*[1]).

The consortium also worked on the integration of the Shadow Honeypots technology with the NoAH architecture (WP1.5). While the initial plan was to investigate the use of *tightly coupled shadow honeypots*[2] in NoAH, the necessary tools to create tightly coupled shadow honeypots were not readily available. As a result, the consortium decided to investigate the integration of *loosely coupled shadow honeypots* into NoAH. By the end of the reporting period, a loosely coupled shadow honeypot prototype, called AID, was running on the NoAH containment environment. The first experience with the implemented prototype was positive. More specifically, the NoAH shadow honeypot instance implemented by FORTH considers the *Apache* web server as the production service to drive the prototyping effort, with two instances of Apache that serve the same site: one running on plain hardware and the second running on Argos. Because they share content but not runtime state, we can say that they are *loosely coupled*. A dispatcher based on the *Squid* web proxy forwards incoming requests to either of the two instances[3]. To achieve transparency for the Web site visitors, the Squid web caching system runs in reverse-proxy mode. Squid was modified so that

[1] http://www.torproject.org/

[2] In *tightly coupled shadow honeypots*, the production server and the honeypot share runtime state. This is in contrast with the *loosely coupled shadow honeypots*, where they run completely separately.

[3] http://www.squid-cache.org/

it forwards requests to the instrumented web server the first time they are observed. If the instrumented web server is able to reply, the request is marked safe and is forwarded to the production web server whenever it occurs again. Therefore, additional overhead is incurred only on the first request for a given URL. Subsequent requests for the same URL will be served with practically no additional overhead.

A second part of the shadow honeypot activity focused on exploring advanced *anomaly detectors for polymorphic attacks*. Shadow honeypots are dependent on reasonable-quality anomaly detectors in order to provide acceptable performance. In the case of AID, we used an application-level anomaly detector to determine with confidence if a request is safe for the production server. However, AID is specific to HTTP, which is a simple protocol. For more complex protocols it is more difficult to build similar application-level detectors. A solution to this problem is the use of generic network-level anomaly detectors, such as *network-level emulation*. Using the identified execution threshold, network-level emulation can be used as an anomaly detector for shadow honeypots. Any traffic that is classified as malicious or cannot be decided upon within the given execution threshold will be processed by the shadow server to allow us to learn more about the attack. However, almost all of the benign traffic will be handled by the production server. Thus, the total performance overhead incurred by the use of shadow honeypots is alleviated. Additionally, this second instance of shadow honeypots also demonstrates clearly how anomaly detection can benefit from honeypot feedback, suggesting that a promising direction for honeypot research is the exploration of how they could potentially interact with a fully passive monitoring infrastructure. Further work in this direction is, however, beyond the scope of the NoAH project.

A second design iteration was carried out taking into account the early experience from implementation and pilot deployment, as well as similar experiences with other research infrastructures such as Leurre.com, Planetlab, and EGEE. Based on these findings, FORTH proposed an updated design for the NoAH core that builds upon the design proposed by WP1.4. Using off-the-shelf components, the new design (which can be seen in Figure 6) enables the automatic bootstrapping of high-interaction honeypots that are added to the NoAH core. The distribution of host-OS updates and guest-OS images (Argos images) is also streamlined. Furthermore, DFN-CERT proposed how the NoAH core can be modified to allow balancing the traffic load between the available honeypots.

Finally, an important aspect that was addressed was the deployment and update of sensors in external organizations. FORTH worked on identifying the methods that can be used for automating the setup of NoAH sensors in the network of cooperating organizations. Currently the process of setting up and updating the remote NoAH sensor is manual. For deploying the sensors of the full-scale NoAH infrastructure, FORTH proposed using the production of custom installation CDs.

Two important issues for the smooth cooperation with external organizations are the procedure for joining the infrastructures and the terms of sharing the data generated by the deployed sensors. DFN-CERT had already outlined the requirements for the cooperation and data-sharing agreements between the NoAH consortium and the cooperating organization in the context of WP0.2 ("Requirements Collection and Analysis"). Based on these requirements, DFN-CERT proposed a non-disclosure and a cooperation agreement that address these issues. The proposed agreements define roles for the participating parties as well as the rights and responsibilities of each role.

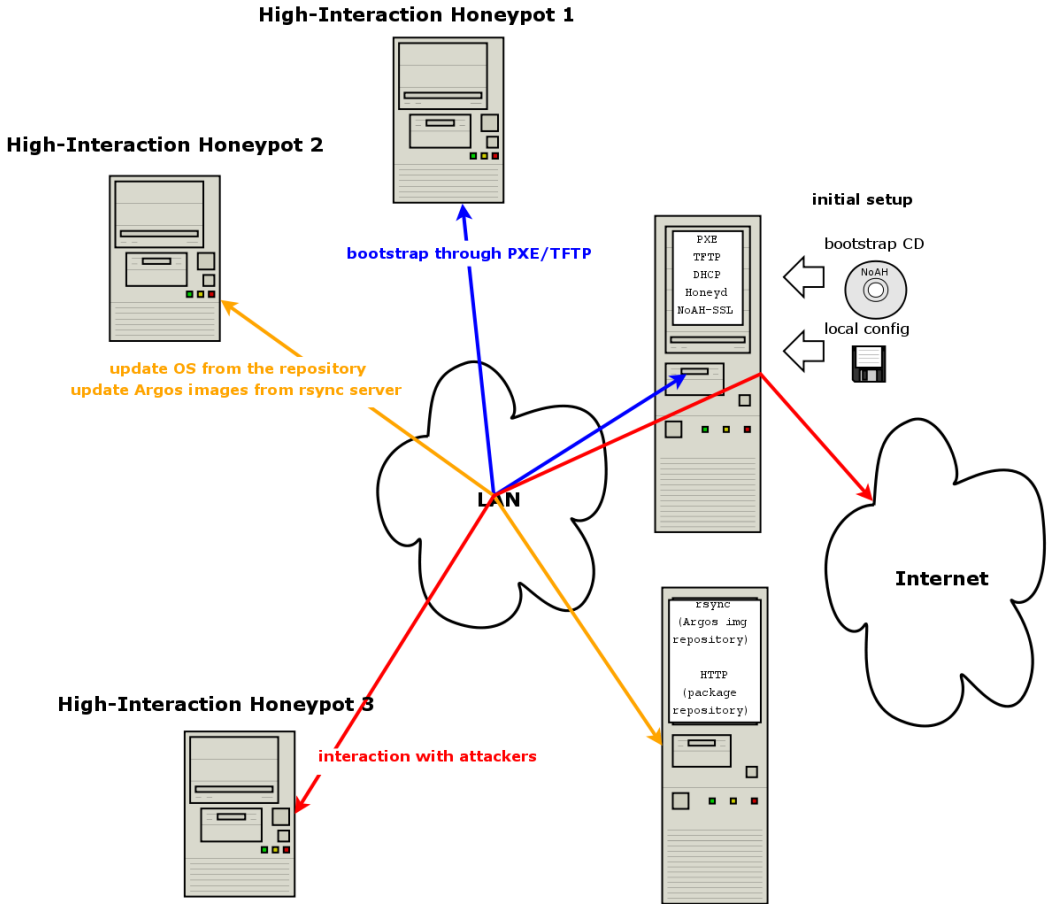


Figure 6 – The extended NoAH core design that resulted from WP1.6

Finally, solutions to potential scalability issues of a large-scale deployment of NoAH were explored. A potential bottleneck of such a deployment is the number of concurrent Honey@Home clients that can be served by the single Honey@Home server of the NoAH core. The use of multiple Honey@Home server instances and the use of load-balancing techniques can help to alleviate this problem. The editing of the deliverable D1.6 started in April 2008, with FORTH proposing a table of contents. Most of the deliverable was edited during summer and September 2008 by FORTH and DFN-CERT. At the end of September, the deliverable was handled to VU and Alcatel-Lucent for review. After incorporating their feedback into the document, FORTH submitted the deliverable D1.6 before the end of October 2008.

A.2.4 WP2 Achievements and end results

WP2 officially started on April 1st 2006 and had three main parts: (i) the implementation of the core NoAH components, (ii) their integration into a working prototype system, and (iii) the optimization of the core components in the critical path.

The consortium started with implementing the components designed in WP1 in the context of task WP2.1 ("*Core Components Implementation*"). As some early prototypes were already available, the goal of this task was to further enhance them so that they become fully functional. During the implementation process, there was feedback from Task WP1.4 ("*Architecture Integration*") which proceeded in parallel, so that the implemented components adhere to the interfaces defined in that task. The result was a smoother integration of the components during task WP2.2 ("*System Integration*"). In this task, the consortium put together the implemented components, thus creating a first *running prototype* of the NoAH infrastructure. Testing was an important part of this process to ensure the correctness of the setup and to provide guidance to subsequent optimizations.

After successfully testing the setup, the consortium worked on optimizing the infrastructure, as specified in task WP2.3 ("*System Optimization*"). Focus was given primarily to Argos because it is inherently the most resource-consuming component in the critical path of NoAH. Other components for which minor performance improvements were also possible were also explored. To address the Argos performance problem, VU implemented *lossy taint tracking* for Argos to reduce its memory footprint and CPU usage. In parallel, FORTH investigated the use of first-pass filters before potential attacks are fed to Argos. A first-pass filter would result in fewer connections being handled by Argos, thus reducing the number of Argos instances required to run in a NoAH core. These activities are covered in more depth in the *2nd NoAH Annual Report* along with some additional optimizations.

After this initial set on optimizations, it was concluded that Argos was still the component with the most room for further enhancements, so work continued along this direction with four individual optimizations explored. The first optimization was the completion of *Prospector*'s implementation by VU. The implementation of Prospector began in the context of task WP2.2 and is detailed in the corresponding section of the *2nd NoAH Annual Report*. Prospector extends Argos to improve the accuracy of pinpointing the bytes that contributed to a buffer overflow. This comes with an additional performance cost, but it enables the generation of more accurate signatures. During this reporting period VU did a rewrite of the Prospector code in order to clean-up the code and reduce the performance hit its incurs. They also worked on a signature generator that takes advantage of the improved logging capabilities of Prospector. Subsequently, the performance of Prospector was shown to incur an additional overhead of 16% compared to vanilla Argos. This was deemed as acceptable as it does not add much to the performance hit caused by Argos. We should note that the work on Prospector was also supported by the Dutch *DeWorm Project*.

The second optimization was the acceleration of the underlying Argos emulator. Argos runs programs approximately 15 times slower than they run natively on hardware. An acceleration module already exists for the QEMU emulator on which Argos is based, however using this module is not technically feasible. This is because the acceleration module runs most of the programs on the physical CPU, so it is not possible to hook instrumentation code. In their Eurosys2006 paper Alex Ho et al. proposed a new method for accelerating heavily instrumented environments like Argos. They proposed that instead of using a physical CPU to accelerate execution (like QEMU does), a fast *Xen*-based virtual CPU should be used.

VU adapted the proposed design to the needs of Argos and worked on its implementation. The prototype was also based on Xen and was named *Xargos*. By dynamically switching between a fast Xen-based virtual machine and a slow Argos instance, Xargos drastically reduces the overhead of tracking tainted data and protecting the system while retaining system-wide safety. Effectively, Xargos invokes Argos only when executing those regions of code that interact with tainted data. Otherwise, the applications run on the Xen-based virtualized CPU at (nearly) native speed. For this switch to take place VU had to modify Xen's and Argos' code. The guest operating system and the applications still run unchanged. At the end of the project, VU had a working prototype of Xargos, but since that prototype lacked support for Windows and recent versions of Linux, VU intends to continue working on it beyond NoAH. The in-depth evaluation of Xargos is also pending, but based on the results reported in Ho's Eurosys2006 paper, Xargos is expected to be an order of magnitude faster than Argos.

A similar direction for speeding-up Argos responses that was also explored was whether it is possible to have a fast main server whose actions are checked a posteriori by a slow Argos machine. Decoupling the analysis of logged data from the actual service that is being monitored can potentially lead to more performance gains. The investigation of this idea was in a very early stage at the time this task was officially completed.

The third optimization for Argos was *Hassle*, which was implemented by VU. Hassle is an enhanced version of Argos that enables detecting and fingerprinting of attacks carried by SSL-encrypted channels. While Argos is able to detect and block such attacks, the signatures it produces for them are meaningless. This is because the origin of the bytes that caused the security policy violation is traced back to an encrypted part of the logged network trace. The encrypted bytes of the same attack will be different each time it occurs, even if the attack is not polymorphic itself. Hassle solves this issue by restoring the meaningful correlation of the bytes that triggered the security alert with the actual unencrypted attack payload. Subsequently, signature generation can proceed as usual.

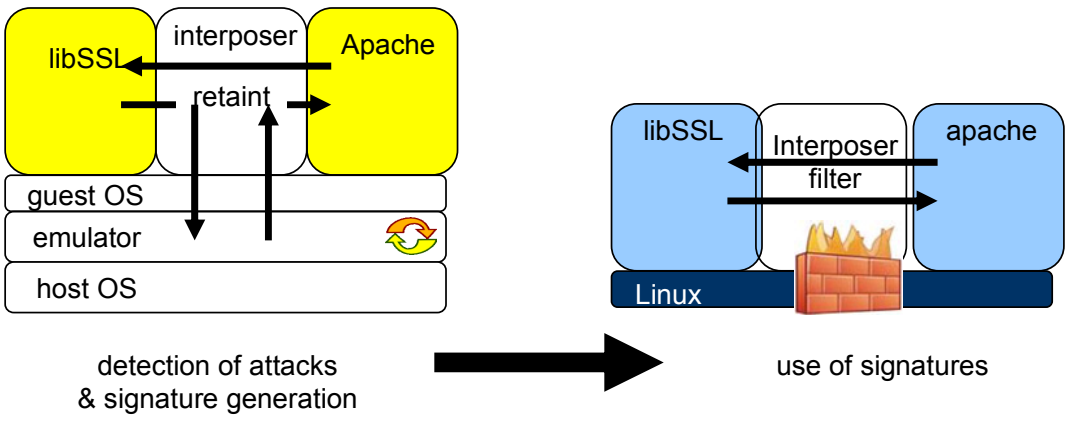


Figure 7 –The architecture of Hassle. By intercepting the OpenSSL library, the Argos instance can now have access to decrypted traffic.

To achieve this, Hassle uses library interposition to place a small amount of code between the application and the SSL encryption library. This code logs the decrypted attack payload and subsequently requests Argos to run the taint analysis using the offsets in the decrypted streams as tags. This process is called *retainting*. After retainting, the signature generator attached to Argos is

able to produce meaningful signature. However, this signature operates on higher-level protocol units. For this reason, VU implemented *interposer filters* that sit between the SSL library and the application and flag or drop all traffic towards the application that matches a signature. VU also evaluated the performance of Hassle using the Apache web server and the httperf benchmark. The evaluation confirmed that dynamic taint analysis of SSL encrypted channels is very expensive, incurring a slowdown of approximately a factor 100 over native code running SSL and a factor 70 over the dynamic taint analysis of non-encrypted channels. It was also observed that for the Apache web server the incurred overhead can be greatly reduced if it is configured to use symmetric encryption (https sessions) instead of asymmetric encryption.

The fourth optimization for Argos was the use of a first-pass filter to lower the number of connections that are eventually handled by Argos. A first-pass filter would drop connections that are deemed as uninteresting. As a result, fewer instances of Argos would need to run in the NoAH core, which means lower hardware costs. A primitive for of first-pass filtering is already performed by *Honeyd*. Honeyd filters out the non-established TCP connections. This is a typical case for scans that try to identify which ports are open in a network.

FORTH explored four additional first-pass filters for Argos that are capable of more elaborate filtering. The first is the use of the *Snort* IDS in *inline mode* which is able to drop connections that match certain signatures. However, Snort has to process several packets before being able to determine whether a connection should be dropped. Up to that point, Argos would have to serve the connection as usually. This means that using Snort would not yield significant performance improvements.

The second is the use of *payload caches*. Previous work had suggested that connections with the same payload on the first packet containing data are very likely to be duplicate requests, therefore uninteresting for processing by a high interaction honeypot. Again, this approach was judged as problematic because it is common for attacks to share the first payload packet with innocuous requests and only differentiate from them many packets later.

The third discussed option for offloading Argos was the use of *Nepenthes*. Nepenthes is a middle-interaction honeypot that was designed to automatically collect malware binaries. A *vulnerability module* is assigned to each monitored port. These are scripts that do not provide full service emulation (like e.g. Honeyd service scripts) but only emulate the necessary parts of the vulnerable service. The approach Nepenthes takes would be good for offloading Argos, but the specifics of its implementation have several drawbacks. Our main concern was that Nepenthes does not provide any feedback whether it was able to handle the incoming request. Even if we implement this mechanism, we will also have to implement a *replay mechanism* for handling connections that Nepenthes can't handle to Argos. More importantly, for each new attack that Argos identifies a new emulation module has to be manually written for Nepenthes.

The final approach that was considered was *Scriptgen*. Scriptgen is able to automatically produce service emulation scripts for honeyd using the output of Argos. The service emulation scripts produced by Scriptgen are based on finite state machines that are built incrementally. Each time that the script reaches at a point where it cannot proceed further, Argos is contacted and the connection is replayed to it. After the end of the connection, Scriptgen uses the output of Argos to add new states to the finite state machine of the service script and enable it to respond to similar requests in the future. Therefore, Scriptgen addresses the three major problems identified in Nepenthes: it can call Argos for help when it reaches a "dead-end", the connection replay mechanism already exists and the emulation scripts are generated automatically. All these make Scriptgen ideal for use as a first-pass filter for Argos.

The implementation of Honey@Home was performed by FORTH. Unused IP address space monitoring was implemented as follows. Each Honey@Home client is requesting an IP address from the local DHCP server (optionally it can be set to listen to a static IP address). Most broadband connection routers, like ADSL routers, organizations and institutes use DHCP servers to assign addresses. Every time Honey@Home client starts, it requests an address from the local DHCP server. The main advantage of this approach is that user does not need to statically set an IP address which may be even hard for him to find. Honey@Home client can be stopped or started any time without interrupting the normal operation of the user's network. Upon the client exit, the clients informs the DHCP server that the IP address is released. To obtain an IP address, the client first creates a pseudo-interface with a random MAC address and broadcasts a DHCP request (left part of the Figure). When the DHCP response is received, the client configures itself to wait for packets destined for the given address.

The Honey@Home client can be configured to claim an IP address statically or by using a BPF filter. Static allocation and BPF filter features are mainly targeted for more advanced users. For example, an administrator can setup a single Honey@Home client in an unused subnet and set the BPF filter to cover all the addresses of that subnet. In that way, the unused subnet becomes useful in a few steps. As long as the client runs, the subnet monitored by the Honey@Home client will contribute to the overall infrastructure.

One implementation challenge that we faced was that Honey@Home clients also receive legitimate traffic, such as broadcast ping or SMB queries. This traffic has to be white-listed and the decision can be made either locally or at the core. The Honey@Home client can be configured not to forward traffic sent to specified ports. Additionally, the Honey@Home client can be configured to forward traffic sent only to a specific set of ports. By default, all traffic destined to the unused address claimed by Honey@Home is sent to core. As Honey@Home captures traffic directed to an unused IP address and not the packets destined for the actual IP address of the host running the tool, there are no privacy concerns. All traffic destined to unused addresses is by default suspicious. The only legitimate traffic that is destined for unused addresses, according to our traffic traces, is attempts to connect to peer-to-peer ports. As peer-to-peer programs tend to have some sort of memory, like host caches of Gnutella, it is observed sometimes that some external hosts try to reconnect to an address that was used in the past, participated in a peer-to-peer network, but is not used anymore and thus claimed by Honey@Home. If the user is concerned about privacy issues, she can configure the client not to forward traffic directed to these ports.

Honey@Home can work even behind NAT. Hosts behind NAT cannot accept incoming connections from external hosts, only for ports that are explicitly forwarded to them through the router setup. For this case, Honey@Home clients can automatically configure the local router to forward specific ports to the physical machine on which client is running using the UPnP protocol. UPnP provides an API to configure the router to forward packets for specific address and ports and is supported by the majority of routers. However, modern routers have UPnP disabled by default for security issues as malware can also use it and make infected machines act like servers. For those users that are not privileged to change the router configuration, the Honey@Home client is limited to capture suspicious traffic that is generated by internal infected hosts, for example local scans.

FORTH also spent significant effort optimizing the Honey@Home code to resolve a performance bottleneck identified during the work on the integration task (WP2.2). The bottleneck appeared only in the Windows version of Honey@Home and caused transfer rates to drop to 400Kbps. At the end of this task the bottleneck persisted despite the thorough combing and fixes of the Honey@Home client code. However, FORTH continued the efforts to identify the cause of the problem with more intensity. Surprisingly, the root of the performance bottleneck was finally

identified in the Honeyd code and not the Honey@Home client code. The emulated TCP stack of Honeyd was not handling correctly delayed TCP-ACK packets. The bottleneck did not show up in Linux because its TCP implementation masks the Honeyd bug. FORTH contacted Niels Provos to let him know of the bug. Recognizing the acquired expertise on Honeyd, Niels added Spiros Antonatos to the list of Honeyd maintainers.

Finally, we should mention the conclusion of the investigation for the cause of a previously identified Argos misbehaviour. FORTH had identified that Argos uses a lot of the host operating system CPU even when the guest operating system is idle. This behaviour masks any insight that tools like *top* or *vmstat* could give us on the state of Argos. However, it was noticed that the bug was not Argos-specific. VU brought FORTH in contact with the QEMU developers, who acknowledged that this issue exists but stated that there is no trivial solution to it.

The most important of these optimizations were finished on time and are detailed in the D2.4 deliverable ("*Enhanced NoAH Implementation and Optimizations*"). The deliverable was successfully submitted on October 12th 2007.

A.2.5 WP3 Achievements and end results

The architecture of the pilot testbed (demonstrator) is shown in Figure 8. The sensors are either directly connected to the Internet or they receive data from a tunnel, funnel, Honey@Home sensor or a load balancer. The load balancer allows organizing honeypots in clusters. This is advantageous because the load can be balanced between all honeypots in this cluster. Additionally, the load balancer can identify when a honeypot is not responding and exclude it from its forwarding list. If an attacker connects to an IP address monitored by the demonstrator, the connection is dynamically redirected to a sensor node. In the pilot testbed all sensor nodes are equipped with a snort IDS and an Argos detector. All data gathered by the sensors are sent to a central NoAH database through a secure channel.

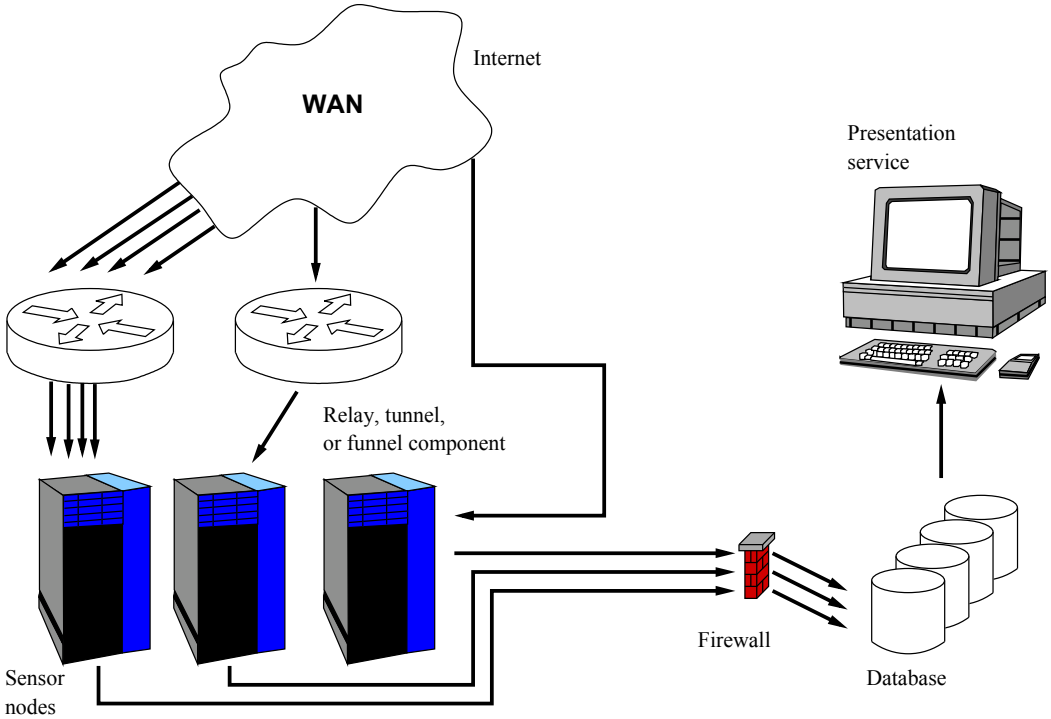
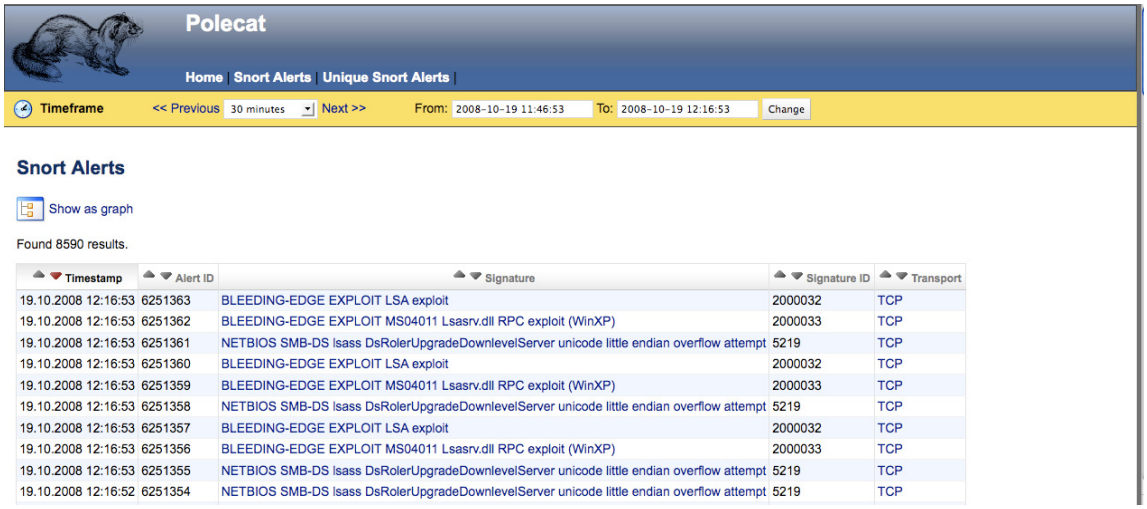


Figure 8 – The architecture of the NoAH demonstrator

In the pilot testbed deployed by DFN-CERT, several operating systems are deployed, such as SuSE linux 9.3, unpatched Windows XP, Windows XP SP2 fully patched as well as Windows 2000 Server SP4. Each sensor produces both Snort and Argos alerts. While Snort is able to identify known attacks, the generic detection engine of an Argos sensor lacks this property. However, Argos is able to detect previously unknown attacks that would not be captured by Snort but is not able to provide all the required information for the identification of the targeted vulnerability. Both types of alerts can be correlated based on the source and destination IP addresses as well as the timestamps. Each kind of alert is stored into an individual database. All sensitive information stored in the database is anonymized. More specifically, the IP addresses of the sensors and all the IP addresses contained in the home networks of the data suppliers are anonymized.

Polecat, a modified version of Weasel that is more scalable, does the presentation of the alerts. Polecat produces attack graphs in real time for a set of snort alerts. A screenshot of the Polecat application running for the demonstrator can be seen in Figure 9. On the top of the page, the user can select several timeframe options. She can select to view the alerts produced the last 30

minutes or provide a specific time range. The list of Snort alerts is presented either as list, as in the screenshot, or as a graph.



Timestamp	Alert ID	Signature	Signature ID	Transport
19.10.2008 12:16:53	6251363	BLEEDING-EDGE EXPLOIT LSA exploit	2000032	TCP
19.10.2008 12:16:53	6251362	BLEEDING-EDGE EXPLOIT MS04011 Lsasrv.dll RPC exploit (WinXP)	2000033	TCP
19.10.2008 12:16:53	6251361	NETBIOS SMB-DS lsass DsRolerUpgradeDownlevelServer unicode little endian overflow attempt	5219	TCP
19.10.2008 12:16:53	6251360	BLEEDING-EDGE EXPLOIT LSA exploit	2000032	TCP
19.10.2008 12:16:53	6251359	BLEEDING-EDGE EXPLOIT MS04011 Lsasrv.dll RPC exploit (WinXP)	2000033	TCP
19.10.2008 12:16:53	6251358	NETBIOS SMB-DS lsass DsRolerUpgradeDownlevelServer unicode little endian overflow attempt	5219	TCP
19.10.2008 12:16:53	6251357	BLEEDING-EDGE EXPLOIT LSA exploit	2000032	TCP
19.10.2008 12:16:53	6251356	BLEEDING-EDGE EXPLOIT MS04011 Lsasrv.dll RPC exploit (WinXP)	2000033	TCP
19.10.2008 12:16:53	6251355	NETBIOS SMB-DS lsass DsRolerUpgradeDownlevelServer unicode little endian overflow attempt	5219	TCP
19.10.2008 12:16:52	6251354	NETBIOS SMB-DS lsass DsRolerUpgradeDownlevelServer unicode little endian overflow attempt	5219	TCP

Figure 9 – A screenshot of the Polecat web-based application running for the pilot testbed

An example of such a graph is shown at Figure 10. We should note that for space reasons, only part of the graph is displayed. At the left side of the graph, we have the attackers, one eclipse box per attacker. At the middle of the graph, the vulnerabilities targeted by the attackers are shown. At the right side, each box represents a targeted IP address. For privacy reasons, all IP addresses are not shown in this graph. The visualization of the alerts allows us to make some very useful observations. First of all, each attacker sends several types of exploits, targeting different vulnerabilities. Second, we can identify the most popular exploits by observing the number of outgoing connections between the exploits and the targeted IP addresses. Finally, we can note any correlation between the exploits in terms of the targeted operating system and service. In our example graph, we can see that two of the most popular exploits are for Windows XP machines.

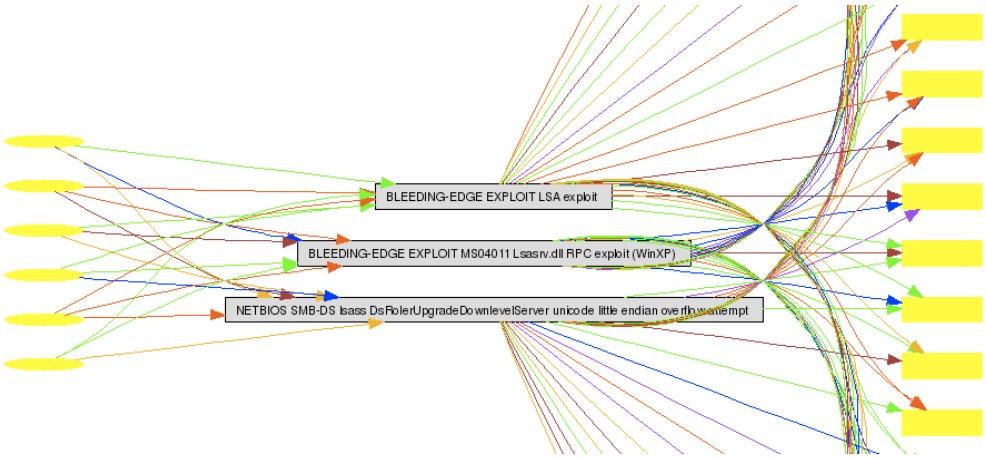


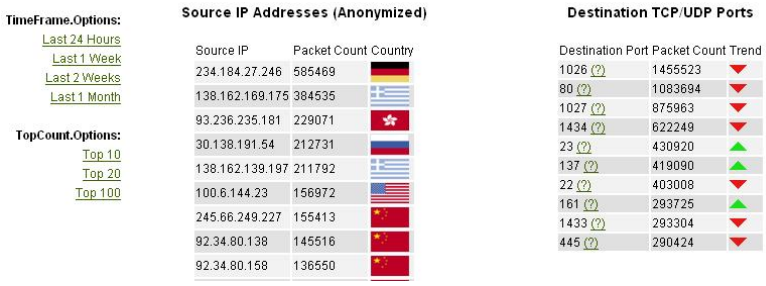
Figure 10 – A graph produced by the Polecat application. At the left side the attackers are listed. At the middle of the graph we can see the targeted vulnerabilities while at the right side a list of targeted IP addresses is shown

Visualization: At the time of this writing, the NoAH pilot infrastructure includes ten static sensors and several dozens of Honey@home clients monitoring more than 9,000 unused IP

addresses. The static sensors are geographically distributed and monitor unused addresses from diverse environments, ranging from universities and research institutions to ISPs and medical centers. On average, the high-interaction honeypots process around half a million packets per day. At this rate it becomes prohibitively expensive to inspect the data manually (e.g., through eyeballing samples), thus automated mechanisms are needed to summarize the data and display statistics and trends from the raw data. In the following paragraphs we present the types of summaries that are currently produced out of each sensor and how they are visualized.

Sensor statistics: Each sensor feeds three software components for the purpose of attack statistics collection. The first is a minimal daemon based on the *pcap* library that listens to a network interface and captures packets going to a given unused IP address space. Specific pieces of information for the captured packets are stored in a local Postgres database which is the second software component in the system. This information includes the protocol number, source and destination IP addresses, source and destination ports, flags in the case of a TCP packet and finally the packet timestamp. The last component is a collection of PHP files that drive the retrieval and rendering of various statistics for the traffic received. More specifically, these statistics are:

- Top source IP addresses. By default the top 10 source IP addresses that sent most packets for the last 2 hours is displayed. For each IP address the number of packets it sent and its geographic location are also displayed. The geographic location is retrieved by a local MaxMind database. Additionally, each IP address is clickable. By clicking it, a user is redirected to a webpage that displays all packets sent by that IP address for a configurable time period. The user is able to select that time period that varies from two hours up to the last month.



- Top destination ports. The top destination ports targeted by attackers are displayed. For each port the number of packets and a trend indication is also shown. The trend indication represents whether the sensor received more or less packets to that port in comparison with the previous time period. Again, the user can configure the time period up to the last month. By clicking a port number, a webpage containing all traffic sent to that port is displayed.
- Attack maps. This page includes two global maps. The first map displays the geographic distribution of distinct attack source IP addresses. Each country is colored based on how many attack IP addresses are hosted in that country. Countries that host no attackers are colored as white, low-activity countries are colored as green, while countries that host lots of attacking IP addresses are colored red. The second map is similar to the first one but is based on the number of packets sent by each country. The scale of both maps is calculated dynamically based on the traffic volume. Maps are generated once per day as they require significant processing power and it is not feasible to create them on demand or update in real-time.
- Attack graphs. This page includes three graphs. The first one is a breakdown of the TCP ports, the second one is a breakdown of UDP ports for the last two hours, and the third one is a breakdown of traffic type in terms of how much TCP, UDP and ICMP traffic was received during the last day.

- Backscatter traffic. Each sensor receives unsolicited traffic, that is, traffic that comes in response to a third party sending spoofed packets that happen to use the sensor's IP address. It is trivial to identify such traffic by inspecting the TCP flags of each packet. A plot for the number of backscatter packets received over the last week is displayed on a separate page.

Sensor and Honey@home monitoring: During the course of NoAH we realized that it is useful to have an internal infrastructure monitoring facility to track the status of NoAH sensors and Honey@home clients. For this purpose, FORTH designed a central site that monitors the availability of Honey@home clients, and the reachability of NoAH sensors. Honey@home clients were modified to send a heartbeat pulse to the NoAH core every minute as an indication of their uptime status. Using the heartbeat information, we know which clients are online and for how long. We visualize the heartbeat information in the familiar style of SNMP-like graphs which are published on a private web page.

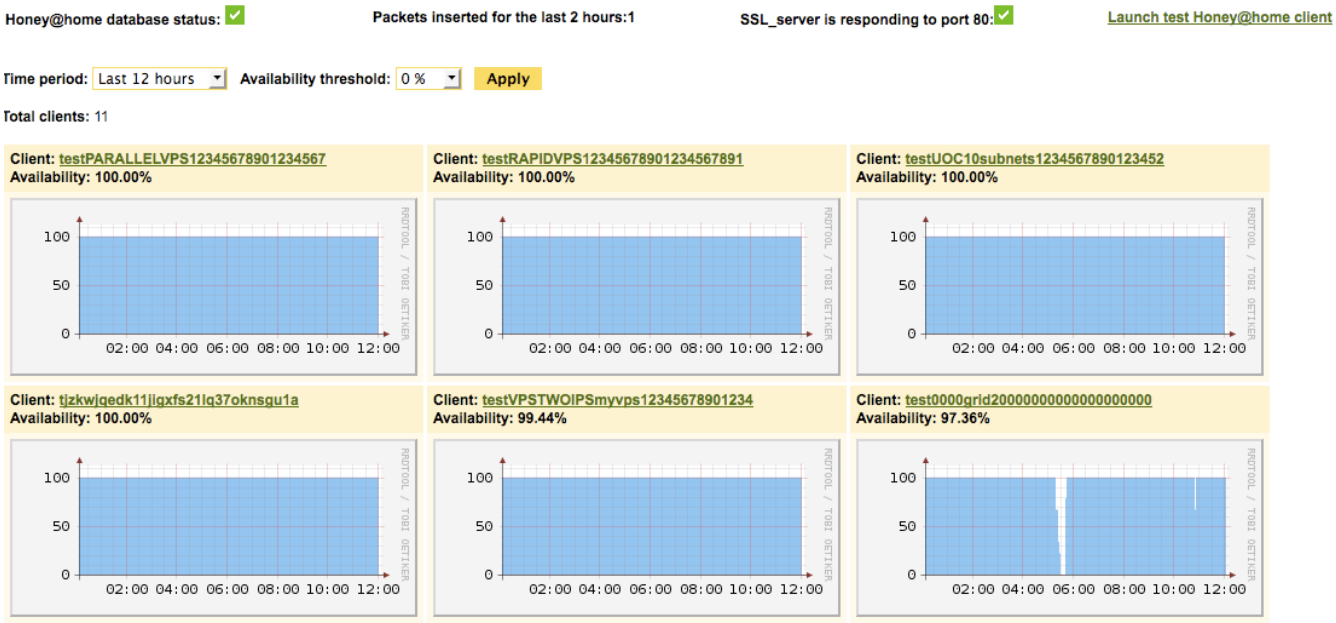


Figure 11 – Availability graphs of Honey@home clients

Figure 11 displays the uptime graphs for several clients we have deployed. By default, we plot only the graphs for those clients that were active at least for one minute during the last twelve hours. We have two parameters to customize the displayed graphs. The first one is the monitoring period. Changing this period, we can see which clients were active at least for one minute up to one year ago. The second one is the availability threshold. By default this threshold is 0%, meaning that clients that send at least one heartbeat are displayed. Setting this threshold to 100% covers only the clients that were up for the whole monitoring period (as defined by the first parameter).

The graphs consist of two parts. The first one is the key of the client. As the monitoring page is available only to NoAH administrators, there is no need to anonymize the client keys. Clicking on the client key the user is redirected to a page hosted under the honeyathome.org domain, which displays traffic statistics for the specific client. The page contains the top 10 source IP addresses that contacted the client, the top 10 destination ports that were contacted as well as protocol traffic breakdown. The second part is the SNMP-like graph that displays the availability of the client as

defined by the two parameters described before. The graph is also clickable, redirecting the user to a page containing the availability graphs for all monitoring periods. In that way, we can observe the behavior of the client for different time periods.

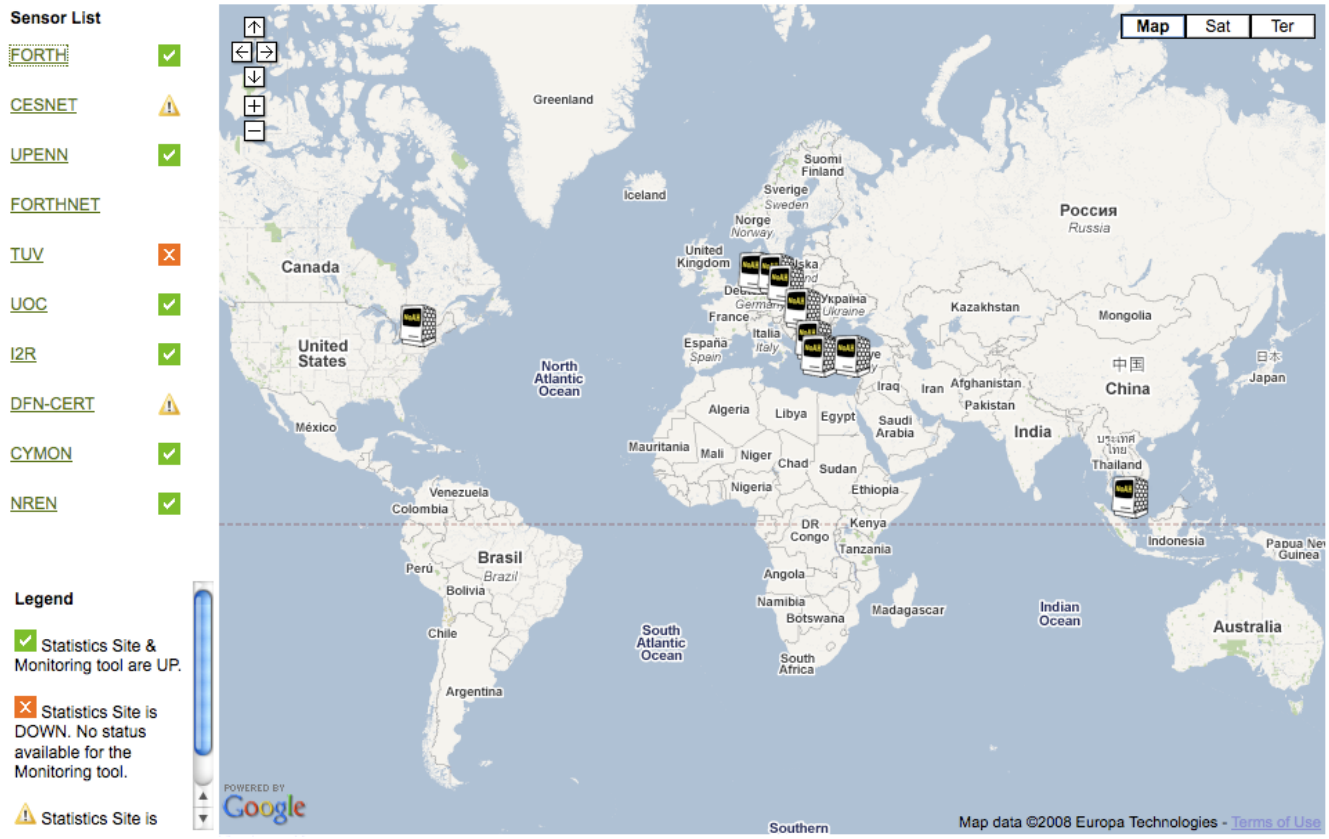


Figure 12 – NoAH sensor map

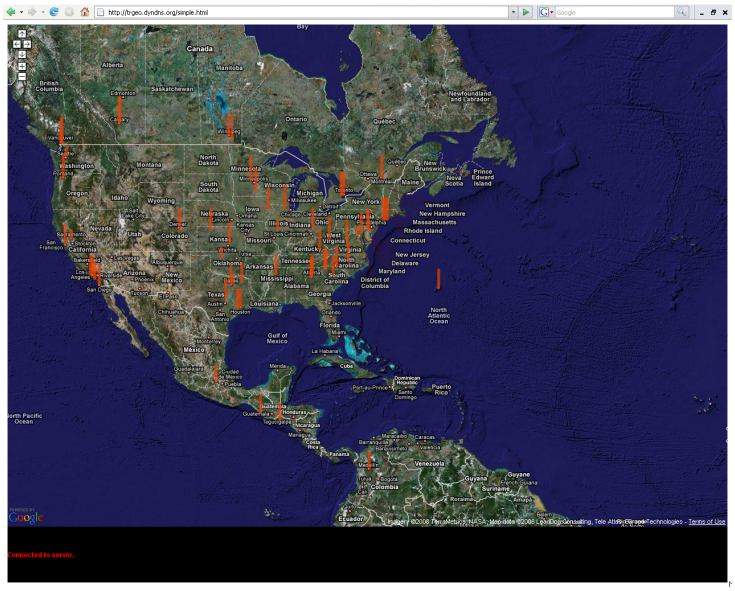
For the NoAH sensors, a web page that shows the sensors drawn on a Google map was constructed. A screenshot can be seen at Figure 12. Each NoAH logo represents a sensor. On the left of a map, a list of the sensors is displayed. Next to the name of each sensor there is a status icon. A green tick means that the sensor is up and displays statistics. A yellow exclamation mark means that the server is up but no statistics are displayed in its pages, indicating possibly the collection component is down. A red cross icon means that the server is either down or not accessible. The status of each sensor is inspected in the background by an external script.

We note that both Honey@home and sensor monitoring pages are password protected and accessible by selected IP addresses only. This security measure ensures the confidentiality of both sensor owners and Honey@home users.

TrGeo: The NoAH sensors are constantly hit by thousands of attacks that are geographically distributed. Although we present the source IP addresses of the attacks on the NoAH statistics website in a tabular form, it seems reasonable to have an easily understood visual overview of the attack origins. Towards this direction, FORTH has implemented TrGeo, a platform for geographic visualization of events captured by the NoAH infrastructure. Two versions of TrGeo

have been implemented. The first one is an offline version and the second one is a web-based application that can be accessible by everyone.

The basic concept behind TrGeo is to track locations of the attackers and display the traffic volume they send on the earth map. On each location a 3D bar is drawn that represents how much traffic the location sent in terms of bytes. As time passes, the height of these bars changes according to the traffic they sent. TrGeo is a client-server architecture. The server and the offline version of the client are customized components while the web-based client is implemented using standard Web gear.The TrGeo server is responsible for capturing and processing sensor traffic. Input traffic can be in tcpdump format, or in XML files describing the traffic, or live traffic captured at an interface. For every packet received, the TrGeo server extracts the source IP address and finds its geographic location through a lookup in a local copy of the Maxmind GeoLiteCity database. This database contains the mappings between the IP addresses and the latitude/longitude coordinates. Although this mapping may not be accurate for all possible IP addresses, we have found the Maxmind database accurate enough for the purposes of TrGeo. The server maintains a list of all different locations and the traffic volume they sent. This volume is calculated of the sum of packet sizes originating from that location. During the development and tesing of TrGeo, two major problems arose. The first one was that thousands of locations had to be visualized on a small map. To overcome this problem, an aggregation mechanism was introduced. All nearby locations inside a predefined radius are aggregated and their traffic volumes are summed up. The second one was that the state kept by the server becomes overwhelming after several hours. To alleviate this problem, older data are deleted according to an aging mechanism. Every time a packet is processed and its source location is identified, all other existing locations on the list will have their value reduced by a constant factor. This mechanism allows us to see locations that sent large volumes of traffic while inactive locations for a long time period will be erased from the map.



The offline version of TrGeo client is implemented in C#. The client connects to the server and periodically fetches a XML file containing all information processed by the server. This information is then visualized on the earth map. The client supports zoom functionality so that the user can focus on specific areas of the map. Furthermore, each bar on the map is clickable: by clicking a bar, a window is displayed with detailed information about the traffic originating from the represented locations. This information includes the name of the locations, the number of bytes sent and the source IP addresses bound to these locations. A pie chart is also displayed for better view of the information.

The web-based client consists of three basic parts: A GoogleMaps API JavaScript library, a Java Applet, and the HTML/JavaScript part that is the container of the first two. The GoogleMaps JavaScript library is provided by Google and enables a web developer to embed Google Maps in web page. It is responsible for loading the maps directly from Google and handling all events and

methods like zooming, drawing, moving, etc. The choice of the Google API was made for two reasons. First, it provides fine-grained control over coordinate ranges and, second, it offers zooming functionality with very good terrain detail. The earth map is the canvas on which we draw the bars of traffic sources. The Java Applet is responsible for communicating with the server. When the page that contains the client is loaded, the Applet is initialized and establishes a connection with the server. The approach we follow for data fetching is push-based. When a connection with a client is established, the server starts to send messages containing the coordinates and traffic volume of all sources. Every time a message from the TrGeo server is received, the applet informs the JavaScript to draw the specified data. As in the offline version, bars are also drawn on the earth canvas to represent traffic volume with each bar being clickable. The Java Applet is also responsible for sending requests to the server and receiving back the results, such as the name of a city when a user clicks on a bar. One might argue that sending request for each click might prove inefficient, the justification for this implementation decision is twofold. First, the size of the geographic location database is prohibitively large to embed at the client. Second, the number of clicks performed by a user is orders of magnitude less than the data displayed and thus having few clicks per second does not decrease the performance of our system.

Apart from the technical aspects of the NoAH demonstrator, the interaction among partners and cooperation with external entities must be considered. Cooperating partners have different interests resulting in different roles and tasks. For example, some partners may be more interested in the analysis of the observed attacks whereas others focus on more day-to-day operations. In addition, the NoAH project cooperates with sites which contribute data to the project, for example Institutes and Universities. For a smooth cooperation among partners and contributors and for the protection of sensitive information, formal regulations are needed. These rules are specified by a public NDA specifying the usage/access of all confidential data for the NoAH partners and the appropriate technical protection of the private data, a cooperation agreement between the data contributors and NoAH administrative role, a cooperation agreement between the scientific partners and the NoAH project, and a technical specification of the set-up and integration of new sensors into the architecture.

A.2.6 WP4 Achievements and end results

A.2.6.a Project Management

The efficient management of the project was crucial for its successful completion. Towards this goal, FORTH, the coordinating organization for NoAH, carried out a number of management activities that helped the project to run smoothly during its final 18 months. Such activities include, among others, the preparation of administrative forms, the intra-consortium communication regarding the progress of the project tasks, and the communication with the European Commission.

Internal consortium communication

Throughout the project FORTH facilitated regular contact with the NoAH partners. Given the geographical dispersion of the consortium, these contacts were mostly using email. Apart from the electronic communication, there were also several face-to-face meetings (approximately one every three months). In the sections below we will focus on the management activities carried out in order to coordinate the consortium actions towards the completion of the undertaken tasks.

Meetings preparations and chairing

FORTH was responsible for preparing the *agenda* and for chairing the NoAH face-to-face meetings. The meeting agendas were also communicated to the other partners before each meeting. This allowed the partners both to prepare better for the meeting and also to propose additional items that made the meetings more productive. The *meeting minutes* were also kept by FORTH. The minutes and the action points assigned during the meetings were put online in the private area of the NoAH website. A special webpage was maintained by FORTH for tracking the progress of the determined action points.

Preparation of the Periodic Reports, Review and Contract Amendment

FORTH coordinated the preparation of all the periodic reports, the midterm review and the Contract Amendment which took place after the review.

A.2.6.b Dissemination Activities Dissemination activities for April-June 2005

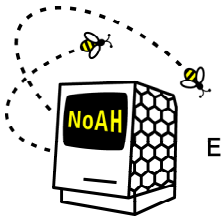
TERENA is the leader for this task. However for this task to be successful in its objective, the mobilization of the whole consortium was required. The consortium gave a high priority to the project dissemination activities and started them even before the official NoAH kickoff. In late 2004, Evangelos Markatos (FORTH) gave the following presentation:

- *Evangelos Markatos: "Writing a Successful proposal: the NoAH approach". Information Day on "Research Infrastructures: the revised workprogramme, the current call for proposals and the way forward". Athens 21/12/2004.*
  (http://www.ekt.gr/en/news/events/ekt/2004-12-20_21/eisigb.htm)

During the first months of the project, TERENA designed and launched the NoAH web site. The site consists of a public part that serves the dissemination of the project's activities and a private part for the collaboration among the partners as described in the objectives of task WP4.1. The private part became operational on 11/4/2005, and the public part became operational on 2/6/2005. TERENA continued to maintain the site on a day-to-day basis.

Also during the first months of the project, VU managed to have NoAH covered several times by the Dutch media and IT-press. The press coverage on the Dutch media continued throughout the first year of the project. More details are available in the "In the media" section of the NoAH web site: http://www.fp6-noah.org/media/ .

Finally, TERENA crafted the stylish logo of the project. While a mediocre logo only serves cosmetic purposes, a successful logo may be much more than that. On March 2006 the consortium received a request from the founder of the Mexican Honeynet Project to use a slightly altered version of the NoAH logo for it. The consortium decided to kindly refuse the request because we would like the logo for exclusive use in future NoAH branded honeypots. However the request itself shows the excellent potential of the NoAH logo and the usefulness of its use in further dissemination activities.



European Network of Affined Honeypots

Dissemination activities for July-September 2005

Result of these early dissemination activities were several requests for collaboration from researchers and members of honeypot related projects. Among them: Mr. Marc Dacier (Eurecom, LEURRE.COM), Mr. Thorsten Holz (Aachen University, German Honeynet Project), and Mr. Georg Wicherski (Mwcollect Alliance). To handle the collaboration requests, the consortium formed the **Project Liaison Committee** headed by Herbert Bos (VU). The Committee contacted the related project to further investigate their collaboration requests. There was significant interest in the NoAH projects and several stakeholders subscribed to the NoAH announcements letter. Also it was arranged with Mr. Marc Dacier and Mr. Thorsten Holz to attend the 2nd and 5th general NoAH meetings respectively.

Dissemination of the NoAH work was not confined only within Europe. In the summer of 2005, we had the following presentation on FORTH's work on shadow honeypots:

- *"Detecting Targeted Attacks Using Shadow Honeypots", 14th USENIX Security Symposium, Baltimore, USA, 4/8/2005.*

In the same period, VTRIP initiated the dissemination activities in the corporate world by distributing to their customer base an electronic newsletter regarding NoAH project. The newsletter included a list of the NoAH consortium members as well as the objectives and upcoming activities of the project. They also prepared a presentation of NoAH project to the user community and made information on NoAH available through their corporate web site.

Dissemination activities for October-December 2005

In October 2006, ERCIM News, a peer review newsletter of the European Research Consortium for Informatics and Mathematics, published a special on "Security and Trust Management", which included the following article authored by FORTH:

- *Kostas Anagnostakis and Evangelos Markatos: "Towards a European Malware Containment Infrastructure", ERCIM News 63, October 2005.*

This was not the only publication for the consortium in this period. VU authored and submitted to EUROSYS2006 a paper on the Argos containment environment. Also DFN-CERT submitted an abstract to the 18th FIRST conference.

- *Georgios Portokalidis, Asia Slowinska and Herbert Bos: "Argos: an Emulator for Fingerprinting Zero-Day Attacks", EUROSYS 2006, Leuven, Belgium, 18/4/2006 (to appear).*
- *Jan Kohlrausch and Jochen Schoenfelder: "The impact of Honeynets for CSIRTs" (abstract submitted to the 18th FIRST conference).*

The work of VU on Argos yielded significant publicity for the project in this period. The public release of the Argos source code helped a lot to have this publicity. Argos was mentioned in several online fora. Among them was the Internet Storm Center. Also the Nepenthes Honeypot team downloaded Argos, experimented with it, an installation guide and referred to it in their web page news section. Finally, Argos was deployed by SURFnet (Netherlands) in their SURF-IDS infrastructure.

Finally, VTRIP continued the dissemination activities in the corporate world, giving talks and presentations for the promotion of NoAH project to their clients. The talks aimed at informing and mobilizing the potential user community.

Dissemination activities for January-March 2006

For the final quarter of the first year of the project a lot of effort was spent on coordinating the organization of the two workshops that the project will hold. The first workshop was arranged to be co-located with TERENA Networking Conference 2006 in Catania, Italy on May 17th 2006, the **World Telecommunication Day 2006**, which this years celebrates "Promoting Global Cybersecurity".

Figure 13 – Catania: Venue of the first NoAH Workshop

The workshop's main goal was to present NoAH and its rationale to the community. To fulfil this goal, one session was allocated to NoAH speakers. Having already established contact with highly regarded honeypot researchers and projects, the consortium decided that it would be best to go for two additional sessions with external speakers instead of just one. The final result was a rather impressive list of speakers for the 1st NoAH workshop:

- Manuel Costa, **Microsoft**. Manuel has co-authored Vigilante, one of the most widely discussed honeypots papers at that time.
- Niels Provos, **Google**. Being the author of honeyd, the first widely used honeypot software, Niels Provos is one of the most influential persons in the area of honeypots.
- Christopher Kruegel, **University of Vienna**. Christopher is one of the leading experimental security scientists in Europe.

- Thorsten Holz, **University of Mannheim** and the German Honeynet project. Thorsten is one of the founders of mwcollect project, author of the honeyblog and instructor in honeypot classes worldwide.
- Ilias Chantzos of **Symantec** is responsible for Government Relations and Public Affairs for Europe, Middle East and Africa. Ilias represents Symantec before government bodies, national authorities and international organisations advising on public policy issues with particular regard to IT security and availability issues.

Framing these high-profile honeypot and security experts with speakers from the NoAH consortium resulted in a very strong workshop that would be worth attending. The expected high attendance of the workshop would help to bring about the greater exposure of the project to the community.

ETHZ initiated and largely handled the creation of a new workshop co-located with the SIGCOMM Conference 2006. It focused on Large Scale Attack Defence (LSAD). While we do not need to mention that SIGCOMM is considered by most to be the most prestigious networking conference in the world today, we should stress that they accepted to co-locate only two new workshops for 2006 and the LSAD workshop is the one of them. This is a great recognition for the NoAH consortium and its future potential in the area. The contents of the SIGCOMM LSAD workshop will be defined after a paper reviewing process which is on the way.

In parallel with the organization of the two workshops, the dissemination activities in Europe continued in this period. DFN-CERT exploited their contacts to present NoAH to the security professionals' community. In the 17th TF-CSIRT meeting there were two NoAH related presentations by Klaus Moeller (DFN-CERT) and Herbert Bos (VU):

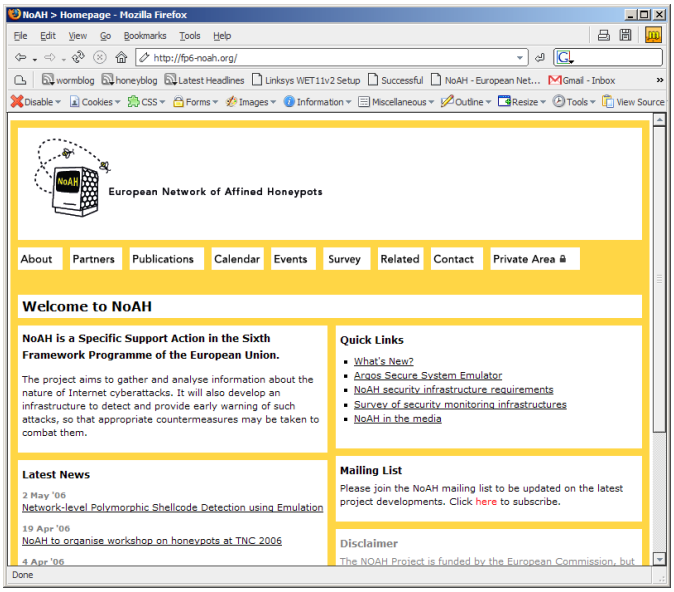
- Klaus Moeller: "NoAH Honeynet Project", 17th TF-CSIRT Meeting, Amsterdam, 24/1/2006.
- Herbert Bos: "Zero-Day Worm Containment", 17th TF-CSIRT Meeting, Amsterdam, the Netherlands, 24/1/2006.

Also, FORTHnet made a presentation on the NoAH project to their technical department. The presentation focused on the NoAH features, the information it can provide and its usefulness for an ISP. Also FORTHnet included the deployment of NoAH honeypots in a deliverable of a project funded by GSRT[1] as a solution for the provision of advanced services in security.

Finally, the project got a lot of attention from Thorsten's Holz popular honeyblog. Honeyblog is a collection of pointers and short reviews of honeypot related articles, tools and projects. NoAH was blogged in it on March 3rd 2006. Shortly after, on March 9th 2006, Argos was blogged too. The result was a boost on the number of document downloads from the NoAH website.

[1] GSRT stands for General Secretariat for Research and Technology , of the Ministry of Development in Greece

NoAH WebSite



The project's web site was setup by TERENA to consist of two parts: The public part (http://www.fp6-noah.org) servers as a dissemination vehicle of the project's created knowledge to the broad community. On the other hand, the private part (http://www.fp6-noah.org/private) serves as the partners' main collaboration tool. This integrates document sharing facilities, mailing list archives, contacts databases and remote management. The system is also extensible so that features can be added as necessary. The website and mailing lists continue to be maintained on a daily basis. A sophisticated web statistics package enabled the consortium to monitor the progress of its dissemination functions.

Dissemination activities for April-June 2006

Continuing on the success of the 1st NoAH workshop, the consortium carried out many other dissemination activities for NoAH in this reporting period. VU and DFN-CERT presented their work in EuroSys2006 and 18th FIRST conferences respectively. VU's presented their work on Argos, while DFN-CERT presented their view of the use of honeypots by CSIRTs:

- Georgios Portokalidis, Asia Slowinska and Herbert Bos: "*Argos: an Emulator for Fingerprinting Zero-Day Attacks*" (EuroSys2006).
- Jan Kohlrausch and Jochen Schoenfelder: "*The impact of Honeynets for CSIRTs*" (18th FIRST conference).

Also, FORTH authored an article on cyber-attacks which appeared in the monthly issue of the Greek edition of The Economist.

- April 2006: The Economist, Greek Edition: *A European platform for the detection and containment of cyberattacks*, by Spiros Antonatos, Kostas Anagnostakis and Evangelos Markatos.

Another important activity was the participation the consortium in the second *Survey on Research Infrastructures* conducted by the EC and the *European Science Foundation*. FORTH, as the project coordinator, was responsible for filling out the survey.

Finally, FORTHnet worked to make NoAH known to external organizations with which they collaborate regularly and also to different departments within their own organization. They informed *ComArch S.A.* (Poland) and *FORTHcrs S.A.* (Greece) for the project and its activities and invited them to participate on the deployment phase of NoAH. FORTHnet also promoted the future deployment of NoAH in their corporate environment by holding a meeting on the subject with the company's technical department in Athens. Finally, they initiated the promotion of the project from FORTHnet's corporate website by starting the compilation of an article regarding NoAH. The article will be forwarded to FORTHnet's marketing/e-press department for publication on the corporate portal, one of the most popular websites in Greece.

Dissemination activities for July-September 2006

In August 2006, the consortium was notified of the upcoming 2007 "*Today is the future*" exhibition, organized by the EC. The exhibition was announced to be opened by *German Chancellor Merkel* and *Commission President Barroso* and would be a high-profile event, showcasing the achievements of the EU-funded projects. The participation of NoAH in the event would be a great achievement by itself, so FORTH filed an application on behalf of the consortium for it.

For NoAH to have a noteworthy presence in such a prestigious event, FORTH held several brainstorming sessions to identify a suitable demo for the project. The demo that was chosen to work on was a real-time visualization of the attacks received by NoAH on a global map. Additionally, much effort was put into preparing supporting material for NoAH's presence in the exhibition. FORTH drafted a two-page information sheet for NoAH. However, because the result was judged as not very accessible to a non-technical public, TERENA started working from scratch on a technically-simplified NoAH leaflet. TERENA also started designing a large-sized NoAH poster. Of course the prepared supporting material is not strictly for use in the "Today is the future" exhibition and can be reused in any other similar events.

To further promote NoAH to the broadband users in Greece, FORTHnet made a proposal to their legal and marketing department. In their proposal, they asked permission to send NoAH informational material to FORTHnet's broadband subscribers. The proposal also included an analysis of the security and legal issues concerning a user who has deployed Honey@Home and his ISP. If their request is granted, it will be a major step towards the wider deployment of Honey@Home.

There were also presentations of the work on NoAH in conferences. VU presented results from their work at GovCERT Symposium in The Hague and the 19th TF-CSIRT meeting in Helsinki:

- Herbert Bos: "*The Worm – early detection and elimination*", GovCERT Symposium, The Hague, The Netherlands, 15/9/2006.
- Asia Slowinska: "*The NoAH approach to zero day worm detection*", 19th TF-CSIRT meeting, Helsinki, Finland, 22/9/2006.
- Herbert Bos: "*SafeCard: a Gigabit IPS on the network card*", RAID'06, Hamburg, Germany, 22/9/2006.

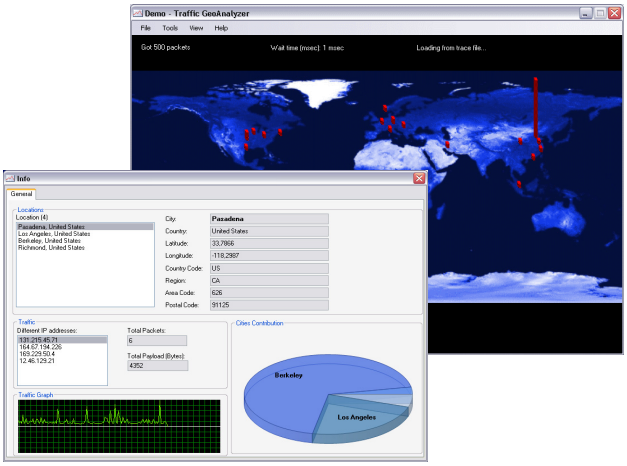
We should also mention that during this quarter of the reporting period, the workshop on *Large Scale Attack Detection* (LSAD) was held in

SIGCOMM 2006 in Pisa. The NoAH consortium was glad to be able to help with the organization of the workshop, with ETHZ leading the efforts. The level of involvement of the consortium is apparent from the make-up of the workshop's program committee, where two out of the three chairs are NoAH affiliates. However, in order not to clash with the SIGCOMM policies, the consortium decided to avoid publicly taking credit for the workshop.

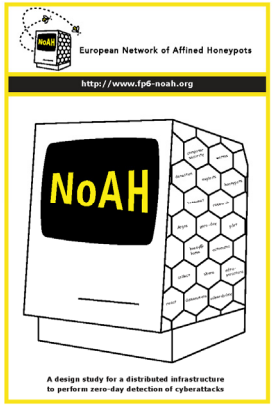
Finally, there was a contribution from FORTH to an ENISA report on Early Warning Systems across Europe. The report was proactively compiled by the agency as a first step towards a more concrete study on the feasibility of a European Early Warning System. FORTH did not have to include NoAH in its contribution, as it was already mentioned in the first draft of the report. This fact is an important acknowledgement for the project.

Dissemination activities for October-December 2006

Early in October 2006, the consortium received a notification that the "*Today is the future exhibition*" was cancelled due to unforeseen circumstances. However, the work on the envisioned exhibition demo and supporting material continued. The demo was nicked "*Traffic Geo-analyzer*" or simply "*TrGeo*". The main window of the demo visualizes the attacks received by the NoAH sensors on a globe map. The user can request more information about the attacks represented in a specific bar. This information is represented in the pop-up foreground window.



The work on supporting material included three different leaflets, aimed at different publics. TERENA finalized their design of a NoAH leaflet after receiving feedback from the partners during the 7th NoAH meeting.



While TrGeo was still under development, VTRIP was able to promote NoAH in exhibitions in Greece. They participated in Digital Technology Expo 2006, the major ICT show in Greece, held in Athens on October 20-22 2006. VTRIP advertised there their participation in NoAH as well as the project's main goals and achievements. NoAH was included in various promotional materials, such as the leaflet, posters and multimedia presentations. VTRIP also participated and promoted NoAH in the Money-Show Expo, held in Thessalonica on December 9-10 2006.

There were also activities in scientific conferences for this period. TERENA arranged to present NoAH in the *23rd Asia Pacific Advanced Network (APAN) meeting* in Manila, in January 2007. Also, the *10th IFIP Conference on Multimedia Security* (CMS 2006) was organized by FORTH, the coordinating organization of the project. The conference was held in Heraklion, Greece and also included the presentation of the *Animated CAPTCHAs* paper.

- E. Athanasopoulos and S. Antonatos: "*Enhanced CAPTCHAs: Using Animation to Tell Humans and Computers Apart*", Proceedings of CMS'06, Heraklion, Greece, 20/10/2006.

During this reporting period, there were also contacts with stakeholders in the area of information-security, to promote the use of NoAH in the future. Liaising with them is expected to increase the impact of NoAH in the long run. FORTH was contacted by Jean-Christophe Le Toquin, attorney for *Microsoft EMEA*. The two parties discussed a possible cooperation on tracking Europe-based bot-herders in which NoAH would play a major role. Also, FORTHnet joined the *Greek e-Business Forum IA-4 working group.* The Greek e-Business Forum is a permanent consultation mechanism between the State, the business sector and the academic community, aiming at establishing useful measures for promoting electronic business in Greece. Their IA-4 working group aims to establish the GR-CERT and FORTHnet plans to present them the NoAH infrastructure as an alternative security solution.

Finally, to enhance the presence of NoAH in popular media, FORTHnet investigated potential informational publications regarding the project on some of the most popular Greek websites. They also studied how NoAH can be integrated with their corporate network tools in the future and started preparing a proposal for their security department on the subject. Also FORTH contributed a paragraph on Distributed Honeypots to Wikipedia. The paragraph mentions NoAH and Honey@Home.

Dissemination activities for January-March 2007

In the final quarter of this reporting period the dissemination the consortium made talks in three different venues.

- Catalin Meirosu: "*NoAH: A European Infrastructure for Cyberattack detection*", 23rd APAN Meeting, Manila, 25/1/2007.
- Spiros Antonatos: "*Honey@Home*", 20th TF-CSIRT, Budapest, 30/1/2007.
- Catalin Meirosu: "*NoAH: A European Infrastructure for Cyberattack detection*", IUCC-TERENA, Tel Aviv, 16/3/2007.

It is important to note that these venues were geographically dispersed, with two of them lying beyond the European borders. Being accepted for making talks outside Europe is an acknowledgement for the project and for the quality of the work performed in it.

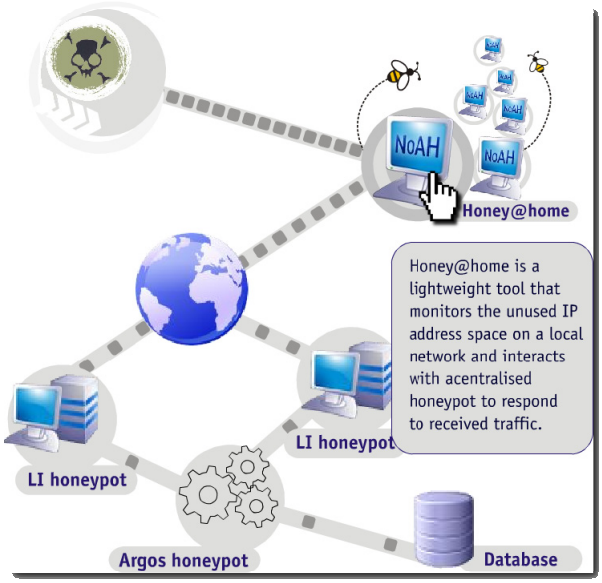
Also, TERENA liaised with the TNC2007 program committee to arrange an invited NoAH talk in the conference. FORTH undertook the responsibility to prepare the talk. Also TERENA finished working on a NoAH poster, which will first be exposed in TNC2007. The consortium also put effort to creating a flash-based demo that aims to explain how NoAH works to the tech-savvy public. The idea first came up in the 3rd quarter of the reporting period, during the 7th general NoAH meeting, which took place on 17/10/2006 in Heraklion. FORTHnet was the leading partner in this effort. Early in the reporting period, they sent to the consortium a draft sketch of the demo, which included a description of the demo functionality and also outlined its aesthetics.

After receiving comments from the consortium, FORTHnet entrusted the creation of the demo to a specialized graphic design company. Near the end of the reporting period a first version of the

demo was ready. The demo shows suspect packets forwarded by Honey@Home first to NoAH's low interaction honeypots and subsequently to NoAH's high interaction honeypots. If an attack is detected, alerts are derived, and a signature for the attack is produced and stored in a database.

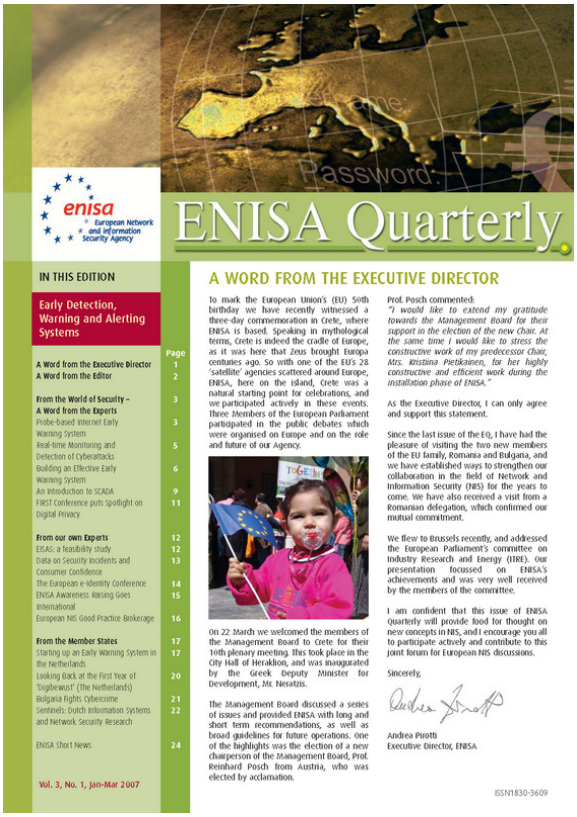
The user interacts with the demo by placing the cursor over the various NoAH components. When he does so, a window pops-up to provide more information on the specific component, an information box for Honey@Home has appeared because the cursor is over a Honey@Home sensor. In addition to aesthetic improvements, there are plans to create a non-interactive loop mode for the demo. In this mode, the informational pop-ups will appear one after another, following an attack packet through out its course in the NoAH infrastructure.

During this quarter of the reporting period VTRIP and FORTHnet continued promoting NoAH to the industry. FORTHnet continued their contacts on the establishment of GR-CERT with the IA-4 working group of the Greek e-business forum. VTRIP included information related to NoAH in their corporate newsletter, which aims to inform their customers and collaborators on VTRIP's R&D activities, accomplishments and services. The VTRIP newsletter is updated and distributed at regular intervals. Also, VTRIP continued distributing the NoAH leaflet on any appropriate occasion (e.g., meetings).



Another addition to the online presence of NoAH, was the creation of a mailing list open to the public for subscription. The new mailing list is named *announce@fp6-noah.org* and intends to make it easier for interested parties to be notified about the latest NoAH news. One can subscribe by clicking the subscription link in the front page of the NoAH website. The mailing list was made available to the public for subscription early in the reporting period there were four announcements sent to this list during it. Currently the list has 16 members that are not members of the NoAH consortium.

Finally, we should mention that TERENA upgraded the web analytics package used by the NoAH website. The new tool is more elaborate than the old one, allowing one to gain better insight into the site's popularity and visitor actions.



Dissemination activities for April-June 2007

During the first three months of this reporting period several dissemination activities took place, both on national and international level. On 21-24 May 2007, the *TERENA Networking Conference 2007* (TNC2007) took place in Lyngby, Denmark. Every year, TNC brings together NRENs from across Europe, to discuss the latest advancements in networking. NoAH had a strong presence in the conference, participating with an invited talk and a poster. TERENA had previously liaised with the TNC2007 committee in or order to reserve a slot for the NoAH talk. FORTH was responsible for preparing and making the talk. The NoAH poster was jointly prepared by TERENA, ETHZ, FORTH and VU. The poster uses the same shades of yellow used for all NoAH material in order to be easily identifiable as belonging to NoAH. The full details for the NoAH presence in TNC2007 are:

- Spiros Antonatos: *"Network of Affined Honeypots - More Than an Infrastructure"*, TERENA Networking Conference 2007, Lyngby, Denmark, 23/5/2007.
- Spiros Antonatos, Daniela Brauckhof, Bernhard Tellenbach, Asia Slowinska: *"The NoAH Project - poster presentation"*, TERENA Networking Conference 2007, Lyngby, Denmark, 20/5/2007-24/5/2007.

On the national level, there were NoAH presentations in Switzerland and The Netherlands. ETHZ gave a presentation at the *Zürcher Tagung symposium*[1] which is organized annually by the *Information Security Society Switzerland*. In 2007 the symposium took place on May 3 in Zurich and brought together the Swiss IT security stakeholders with people from industry, research and education attending it:

- Bernhard Tellenbach: *"Intrusion Detection Systeme: Trends und Herausforderungen"*, Zuercher Tagung 2007, Zurich, Switzerland, 3/5/2007.

In the Netherlands, VU made a presentation to *The Netherlands Organization for Applied Scientific Research* (TNO) and one to the *NLUUG Conference*. Both presentations were about the latest advancements in Argos (Prospector etc):

- Herbert Bos: *"Fingerprinting Intruders"*, TNO Security Knowledge Network Meeting, Groningen, The Netherlands, 31/5/2007.
- Asia Slowinska: *"Prospector: Analysis of Heap and Stack Overflows using Emulated Hardware"*, NLUUG 2007 Conference, Ede, The Netherlands, 10/5/2007.

Other efforts relating to dissemination was the continuation of the development efforts to produce demos and material that can be used to promote NoAH. TERENA and FORTH created a NoAH-branded presentation template for Power Point. This template will give to the NoAH related presentation the uniform look that exists in all other NoAH produced material. FORTHnet finalized the NoAH flash demo that they had started working on during the previous reporting period. The demo shows suspect packets forwarded by Honey@Home first to NoAH's low interaction honeypots and subsequently to NoAH's high interaction honeypots. If an attack is detected, alerts are derived, and a signature for the attack is produced and stored in a database. Another demo was produced by FORTH to display aggregate statistics for the attacks intercepted by the NoAH sensors.

Finally Virtual Trip included once more information on the latest NoAH advances in their corporate newsletter. The newsletter is distributed to the company's clients and collaborators around the world, informing them on the company news, accomplishments and R&D activities.

Dissemination activities for July-September 2007

During the 3rd quarter of 2007 members of the consortium gave two presentations addressing different audiences. ETHZ presented the signature generation features of the NoAH infrastructure in the 22nd *TF-CSIRT* meeting. TF-CSIRT is a forum for the exchange of experiences and knowledge between the members of the CSIRT[2] community across Europe. This community can benefit directly from the NoAH infrastructure and therefore it is very important to keep them up to date with the work in the project:

- Bernhard Tellenbach: *"Automated Signature Generation: Overview and the NoAH Approach"*, 22nd TF-CSIRT meeting, Porto, Portugal, 21/9/2007.

The second presentation addressed the intrusion detection research community:

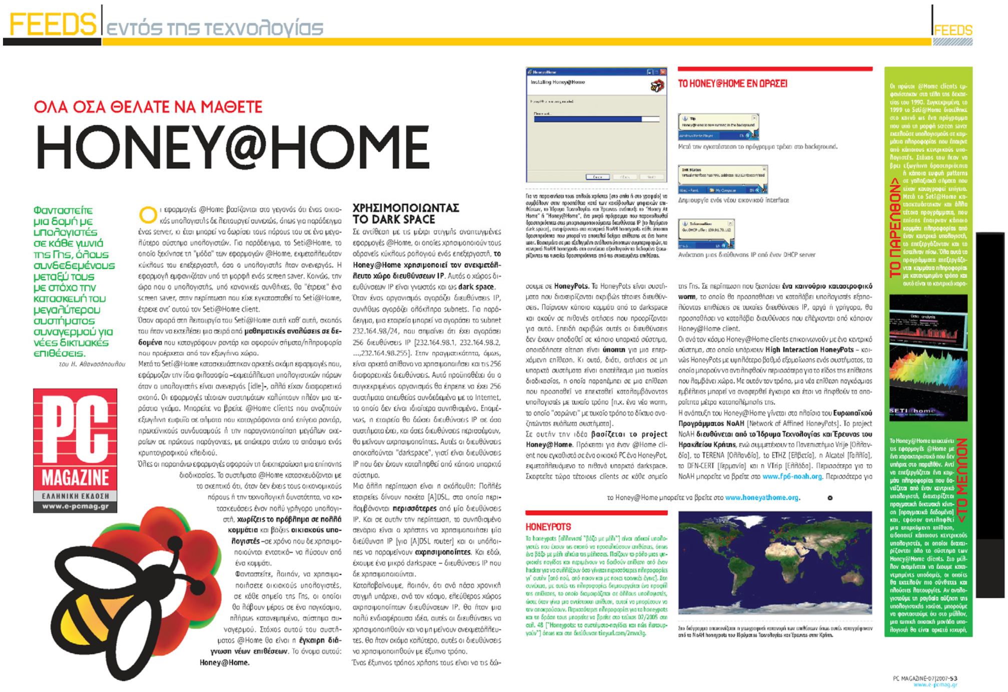
- M. Polychronakis, K. G. Anagnostakis, and E. P. Markatos: *"Emulation-based Detection of Non-self-contained Polymorphic Shellcode"*, Proceedings of the 10th International Symposium on Recent Advances in Intrusion Detection (RAID), Gold Coast, Australia, 6/9/2007.

[1] http://www.zuerchertagung.ch/

[2] CSIRT is an acronym for *"Computer Security Incident Response Team"*.

The consortium also addressed the general public with two different articles about Honey@Home in popular Greek media. The first article was published in the Greek edition of the *PC Magazine* and aimed to promote the use of Honey@Home to the Greek home users. The second article focused on explaining how Honey@Home can help to battle cyber-attacks:

- July 2007: *"Honey@Home"* by Elias Athanasopoulos, PC Magazine (Greek Edition), Nr. 7, 2007.
- September 2007: *"Honey@Home: Trapping Attacks on the Internet"* by Spiros Antonatos, Elias Athanasopoulos, Kostas Anagnostakis and Evangelos Markatos, The Economist (Greek Edition)


FEEDS εντός της τεχνολογίας

ΟΛΑ ΟΣΑ ΘΕΛΑΤΕ ΝΑ ΜΑΘΕΤΕ

HONEY@HOME

Φανταστείτε μια δομή με υπολογιστές σε κάθε γωνιά της Γης, όλους συνδεδεμένους μεταξύ τους με στόχο την κατασκευή του μεγαλύτερου συστήματος συναγερμού για νέες δικτυακές επιθέσεις.

ΧΡΗΣΙΜΟΠΟΙΩΝΤΑΣ ΤΟ DARK SPACE

TO HONEY@HOME ΣΕ ΔΡΑΣΗ

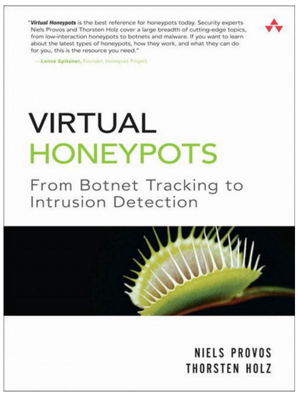
HONEYPOTS

ΤΟ ΠΑΡΕΛΘΟΝ

ΤΟ ΜΕΛΛΟΝ

PC MAGAZINE ΕΛΛΗΝΙΚΗ ΕΚΔΟΣΗ www.e-pcmag.gr

During this period, the project also received significant attention from the community around IT security. Niels Provos and Thorsten Holz included many details about Argos in their newly published book on *Virtual Honeypots*. This is an important recognition for the work in NoAH, given that the authors are two of the most recognizable honeypot researchers and the book is the most up-to-date book on the subject. Provos and Holz also mentioned Argos when they were later interviewed[1] on their new book:



- Niels Provos and Thorsten Holz: *"Virtual Honeypots: From Botnet Tracking to Intrusion Detection"*, Addison-Wesley Professional, July 26, 2007

[1] http://www.networkworld.com/news/2007/081007-virtualhoneypots.html?page=1

Additionally, NoAH was blogged once more, this time by *Prof. Urs E. Gattiker*[1], founder and CTO of *CyTRAP Labs GmbH*. This blog post is an acknowledgment that NoAH can also have an impact in companies in the IT security industry:

- CyTRAP Labs blog: *"Research that matters: more insights into NoAH the European Network of Affined Honeypots research project"*, September 30, 2007.

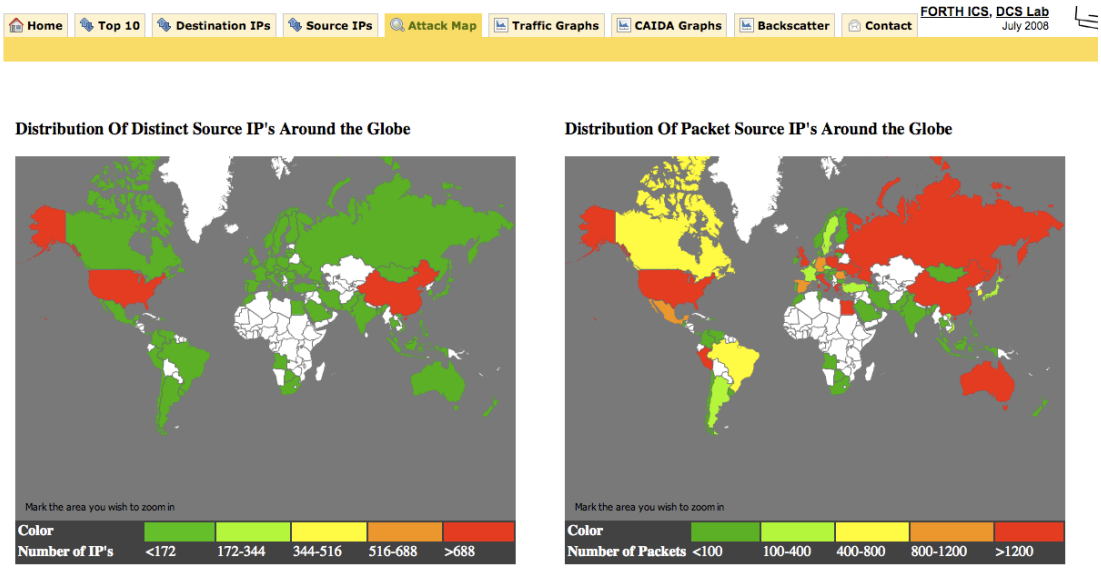


Figure 14 – Geographic distribution of attackers as observed by the NoAH statistics site

An important event for disseminating the project results was the launch of the *NoAH Statistics Site*[2]. The website was implemented by FORTH. The site provides *real-time statistics* on the data that the deployed NoAH infrastructure collects. These statistics include the top attackers, top destination ports that are targeted as well as the geographic distribution of the attackers and the volume of backscatter traffic. These statistics will help the NoAH partners to see what the most popular applications under attack are and fine-tune the high-interaction honeypots of the pilot testbed accordingly. More important, the site aimed to attract interest for the project in two different levels:

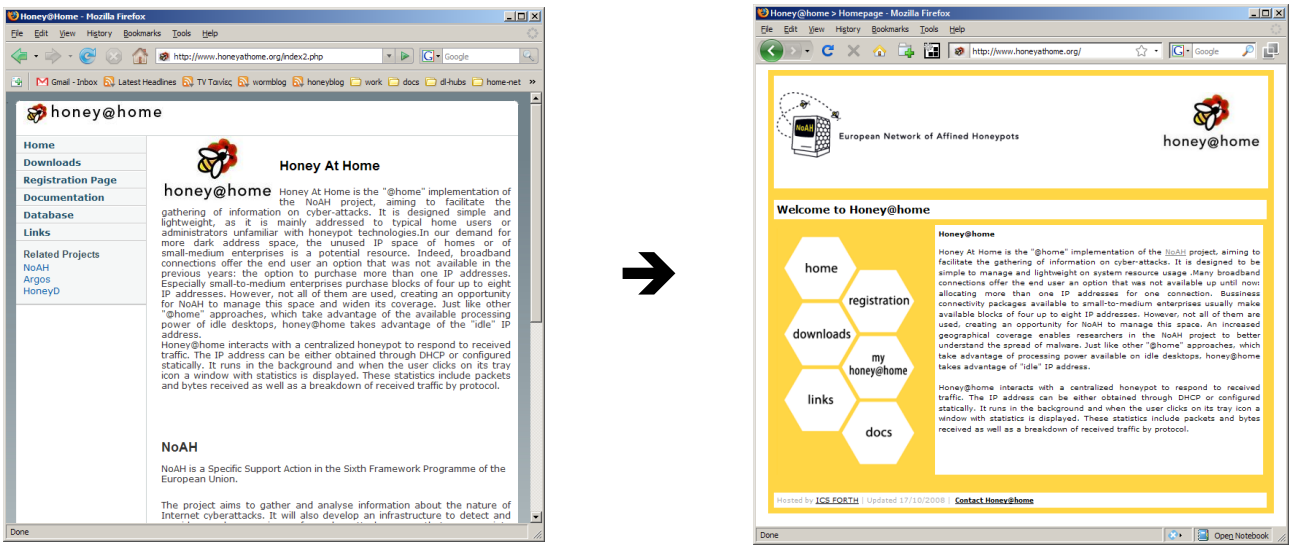
- Showcase the data collected by NoAH to the community and *raise the awareness on NoAH* and related technology by pointing interested parties to the main project website.
- Attract organizations to *join the NoAH infrastructure* and have a copy of the site deployed in their local network as an *added value benefit*.

To further promote the second aim of the site, FORTH drafted a teaser paragraph. This paragraph was communicated to the members of the SEEREN2 project and also to the mailing lists maintained by TERENA.

Finally, we should mention that FORTH redesigned the Honey@Home website in order to match the look & feel of the main NoAH website. The changes can be observed below.

[1] Professor Urs E. Gattiker was also reviewer of the project during its midterm review.

[2] http://stats.fp6-noah.org



Dissemination activities for October-December 2007

From October to December 2007 the consortium disseminated the work in the project in several scientific conferences and workshops. The consortium reached the malware and applied IT security research communities with two talks:

- S. Antonatos, K. G. Anagnostakis and E P. Markatos: *"Honey@Home: A New Approach to Large-Scale Threat Monitor"*, Proceedings of the 5th ACM Workshop on Recurring Malcode (WORM 2007), Alexandria, USA, 2/11/2007.
- A. Slowinska and H. Bos: *"The Age of Data: pinpointing guilty bytes in polymorphic buffer overflows on heap or stack"*, Proceedings of the 23rd Annual Computer Security Applications Conference (ACSAC'07), Miami, USA, 14/12/2007.

EC2ND
European Conference on Computer Network Defense

enisa

Figure 15 – EC2ND 2007 was organized by members of the NoAH consortium in collaboration with ENISA. Herbert Bos (from VU) delivers a talk on attack signature generation.

More important, members of the consortium did not just confine themselves to participating in existing events but actually *pioneered* efforts to establish new events that promote the exchange of

knowledge between IT-security researchers at a *European* level. Towards this end, in October 2007 the *3rd European Conference on Computer Network Defense* was organized in Heraklion in collaboration with the *European Network and Information Security Agency* (ENISA). FORTH led the organization of the conference and there was also a presentation by VU:

- M. Valkering, A. Slowinska and H. Bos: *"Tales from the Crypt: fingerprinting attacks on encrypted channels by way of retainting"*; 3rd European Conference on Computer Network Defense (EC2ND 2007), Heraklion, Greece, 4/10/2007.

There were also many dissemination activities on a regional level. ETHZ published an article about NoAH in the quarterly issue of the *"ETH Globe" magazine*. The magazine is issued by ETH Zurich and the November 2007 issue had a focus on *Risk and Security*:

- November 2007: *"Mit Honig gegen Hacker"* by Bernhard Tellenbach, ETH Globe Magazine, Nr. 4, 2007.

Also, Virtual Trip presented the benefits of NoAH compared to traditional security solutions in the *3rd Regional Electronic Security Forum* which was organized by the *SE Europe Telecommunications and Informatics Research Institute*:

- Konstantinos Xinidis : *"Emerging Ways to Protect your Network: From Vulnerability Scanning to Real-time Monitoring and Detection of Cyberattacks"*, 3rd Regional Electronic Security Forum, Thessaloniki, 11/10/2007.

They also distributed NoAH material and displayed a presentation of the project at their corporate booth in the *"Infosystem 2007 – HiTech Innovators Partenariat"*, held in Thessaloniki in October 2007.

During this period, the first preparations for the *2nd NoAH Workshop* were made. The consortium discussed the format of the workshop and drafted a list of candidate invited speakers. It was agreed to format the workshop as two sessions with 3 talks in each session. Half of the talks would be by members of the consortium and the rest by invited speakers. It was also agreed to hold the workshop during the *2008 TERENA Networking Conference* which would take place in May 2008 in Bruges, Belgium. A list of candidate speakers was also drafted by the consortium. TERENA made the initial planning and the first arrangements for the workshop.

Moreover, FORTH further enhanced the *NoAH Statistics Site*, which was launched in the previous quarter of the year. They added new type of plots aimed to make the data collected by the infrastructure directly comparable to other publicly available data so that they become more useful to security researchers and analysts. Such a source of publicly available data is the *CAIDA Network Telescope*.

Finally, in the October issue of the *USENIX ;Login:'* magazine, Sam Stover wrote quite positively about Argos.

Dissemination activities for January-March 2008

As the 2nd NoAH workshop approached the consortium focused their efforts to its preparation in order to match the success of its precursor. TERENA made the necessary arrangements with the organization committee of TNC 2008 in order to reserve the required slots. The consortium had already drafted a list of candidate speakers for the workshop. Subsequently, FORTH contacted the candidate speakers and by the end of January 2008 the names of the speakers were finalized.

Additionally, FORTH undertook the task to organize a *NoAH booth* at TNC 2008. The booth would help to attract more attendees to the 2nd NoAH workshop, therefore amplifying its overall impact. The booth would showcase the NoAH demos that had been previously developed. It would also give out promotional NoAH-branded items and Honey@Home installation CDs. TERENA made the arrangements for the space required for the booth. FORTH was responsible for ordering the promotional items and the equipment required to setup the booth. Also, FORTH created a stand-up banner that would be placed outside the booth in order to make it more visible from a distance.

While there were no NoAH talks during this period, FORTH published an article on the project in *The Parliament Magazine*:

- February 2008: *"NoAH: A European Network of Affined Honeypots for Cyber-Attack Tracking and Alerting"*, by Evangelos Markatos and Kostas Anagnostakis, The Parliament Magazine, Issue 261, 18/2/2008.

To promote the use of a "NoAH router" in real life environment, Alcatel-Lucent initiated several contacts during this period. They made arrangements for promoting the NoAH router to clients of the company involved in law enforcement and defence. The offer will be made through the specialized Police and Defence division of Alcatel-Lucent that has regular collaboration with these clients. Alcatel-Lucent also discussed the incorporation of the NoAH Router to a mobile security solution developed by *Alcatel-Lucent Ventures*[1]. One other promising contact was with *Renater*, the French NREN. Unfortunately, a lack of resources on the part of Renater prohibited the contacts from yielding tangible results. Alcatel-Lucent also established contact with Eurecom, the operators of the Leurre.com honeypot platform. The aim was to explain to them how the NoAH router can be beneficial for the Leurre.com honeypot platform.

Again in this period, the project received significant attention from third party electronic articles. This time it was *Shelia*, the specialized client-side honeypot developed by VU, that was blogged and mentioned in several articles. The references to Shelia that were identified are the following:

- Jamie Riden and Christian Seifert: *"A Guide to Different Kinds of Honeypots"*. SecurityFocus Infocus magazine, February 2008.
- SecGuru blog: *"Shelia - Client-side honeypot for attack detection"*.
- Wikipedia's *Client honeypot* article.
- The Mitre honeyclient project site refers Shelia.

[1] *Alcatel-Lucent Ventures* is Alcatel-Lucent's internal incubator that supports teams striving to turn innovative technologies and business ideas into new products.

Recognizing the proliferation of blogs as a dissemination medium for scientific activities, the consortium decided to launch a *NoAH blog* that will be used to post the latest news of the project. The NoAH blog is meant as a complement to the existing project mailing lists. Compared to the mailing lists, a blog has the advantage of being less volatile. i.e. the news will stay online much longer and they will be searchable through the existing web search engines. It was decided that except from project news, the blog post could also include real-world experiences, event announcements and opinions on related research. The NoAH blog was launched on February 5 2008. It is hosted by FORTH.

Figure 16 – Oud Sint Jan Conference Centre in Bruges: Venue of the 2nd NoAH Workshop.

Dissemination activities for April-June 2008

During this reporting period the consortium successfully organized the *2nd NoAH Workshop*. It took place in Bruges, Belgium on May 20 and was co-located with *2008 TERENA Networking Conference*. The speakers coming from the NoAH consortium were framed by a list of acknowledged invited speakers: *Stefano Zanero* (Politecnico di Milano), *Melanie Rieback* (Vrije Universiteit Amsterdam) and *Marc Dacier* (Symantec Research Labs, Europe). TERENA made most of the local arrangements for the workshop. After the end of the workshop, they also added a new section to the NoAH website with the presentations from the workshop.

While the *1st NoAH Workshop* aimed at presenting the

rationale of the project to the scientific community, its successor aimed to showcase the achievements of the implementation phase of the project. Also, the invited speakers were able to provide insights for novel applications of honeypot technology (e.g. in RFID based systems). The workshop was well-received by the community, being able to attract over 60 attendees. As we already mentioned, NoAH's presence in TNC 2008 was not confined to the workshop talks. FORTH took the initiative to setup and staff a promotional booth for the project. The booth aimed to inform the participants of TNC about the project and encourage them to participate in the infrastructure either by setting up a NoAH sensor in their network or individually setting up Honey@Home.

Figure 17 – The NoAH booth at TNC 2008.

Since there were approximately 10 more booths in the same space, it was crucial for the success of the NoAH booth to be able to stand out in order to attract visitors. More important, FORTH setup two PCs with large LCD panels that displayed three different NoAH demos:

- The attacks received by the deployed sensors during the last day were visualized by *TrGeo*
- Aggregate statistics about the attacks were displayed using the *NoAH statistics ticker*
- A *Google-Earth* visualization of the locations of the deployed NoAH sensors.

Of course, the NoAH publicity articles that were distributed from the booth contributed to the overall interest of the audience. They included NoAH T-shirts, NoAH anti-stress dices, Honey@Home key chains, and of course the NoAH leaflet and the Honey@Home installation CDs. Last, but not least, the success of the NoAH booth would not be possible without the efforts of the booth staff. They actively engaged in conversation with the visitors and in the end managed to make the NoAH booth the most popular in TNC 2008.

Apart from the NoAH workshop, in this period there was also a presentation on Eudaemon at the EUROSYS conference:

- G. Portokalidis and H. Bos; *"Eudaemon: Involuntary and On-Demand Emulation Against Zero-Day Exploits"*, Proceedings of ACM SIGOPS/EUROSYS 2008, Glasgow, UK, 4/4/2008.

Other dissemination activities in this reporting period included further efforts of Alcatel-Lucent to promote the NoAH router. They continued the contacts with the Police and Defence division of the company promoting the NoAH router as a solution that would allow their customers to easily

take advantage of honeypot technologies without relying on non-European providers. Also, FORTHnet presented NoAH to their shareholders.

Dissemination activities for July-September 2008

Perhaps the most significant achievement during this reporting period was the inclusion of two NoAH-centric tutorials in the upcoming EC2ND conference in Dublin in December 2008. In part due to the heavy involvement of NoAH partners in the EC2ND programme committee, the committee chairs decided to invite two tutorial speakers from NoAH partners to the EC2ND conference. The first invited tutorial will be given by VU on Argos and the underlying principles, focusing primarily on teaching attendees how to deploy and modify Argos to suite their needs. The second tutorial will be given by FORTH on polymorphic attacks and how the combination of honeypots and network level emulation, as detailed in one of the NoAH shadow honeypot instances, can help address the polymorphism problem. Given the visibility of the conference, its European scope and the involvement of ENISA in its organization, this is an excellent opportunity to evangelize the benefits and promote the use of NoAH-derivative technologies. Although the actual tutorials are going to take place after the end of the project, significant effort was put during the final months of the project towards negotiating and producing the content of this dissemination action.

Additionally, after the end of this reporting period, during the 1st *FORWARD Smart Environments Working Group* meeting, VU was able to present much of the work they performed in the context of NoAH and provide insights on how it can be adapted for use in mobile devices[1].

- G. Portokalidis; *"Security & Dependability @ VU"*, 1st FORWARD Smart Environments Working Group meeting, Vienna, Austria, 6-8/11/2008.

We consider this activity important because it shows that technology developed within NoAH has applications to areas beyond the initially foreseen.

[1] In addition to NoAH, this talk covered other VU activities on Smart Environments as well.

A.2.7 Summary of work performed

NoAH successfully achieved the primary objective of carrying out a design study towards building an infrastructure for European security research. More specifically, NoAH designed, implemented and carried out trials for an infrastructure of affined honeypots that gathers and correlates data about attackers, their methods, and actions on the Internet. NoAH has **successfully developed** a complete pilot honeypot-based security research infrastructure, addressing the technical and organizational challenges along the way. The NoAH design also provisioned for externally developed components like Nepenthes and Scriptgen

As part of this infrastructure, NoAH also produced and validated novel components for the automatic identification of novel attacks and for the automated generation of corresponding signatures. Most of these components, including the prominent Argos and Honey@Home systems, are openly available for download and use by researchers and end-users. Argos is already deployed as part of SURFnet's SURF-IDS infrastructure, and **is widely used by researchers in Europe**, the US, and Asia, having reached its 4th release and with **more than 2,700 downloads**, and it has become a key component in the *Scriptgen*[1] effort from *Eurecom*.

Honey@Home is a new approach, which in the form of a "screensaver-like" application will empower normal home users and small organizations to participate in the NoAH infrastructure and make a contribution towards fighting cyber-attackers. The *Honey@Home* initiative opens up NoAH and the security infrastructure space to people that are not familiar with honeypot technologies. Empowering non-expert users to contribute is likely to impact along two dimensions. First, the NoAH project will gain visibility outside the research community, while also increasing awareness on security issues and attack trends. By **empowering people** to directly contribute to improved situation awareness and security, the sense that attacks can be detected and prevented fast is likely to appeal to common users. By helping security researchers, we contribute to the effort of making life on the Internet safer. Second, it will allow security monitoring initiatives to significantly increase the fidelity and accuracy of the data collected.

Compared to legacy, previous-generation honeypot systems, significant enhancements have been achieved to increase efficiency through means such as the Xargos on-demand taint tracking system, aggressive filtering of uninteresting data, optimizations of the Honey@Home framework, and forensic tools such as Prospector that help pinpoint the attack vector on malicious payloads.

Furthermore, NoAH has achieved improved attack coverage compared to legacy honeypot technology through the implementation of the AID shadow honeypot instance and the Shelia client-side honeypot system. The former demonstrates how NoAH can cooperate with inline passive monitoring infrastructures, while the latter complements the server-centric focus of traditional honeypots with additional coverage on attacks directed to client software.

Finally, two different approaches were explored for the incorporation of client-side honeypots to NoAH. The first approach was the use of specialized instrumented applications that target specific malicious activities. A case study on this idea was performed and a prototype (named *Shelia*) was implemented. Shelia targets malware that is propagated by means of email by scanning a local spam email folder for items (links, attachments, etc) that could lead to a security breach. The second approach was *Eudaemon*, which brings the taint-tracking technology of Argos to desktop

[1] http://www.acsa-admin.org/2005/papers/118.pdf

applications. In order to not sacrifice the performance of client-side desktops for the sake of security, Eudamon can be turned on or off *on-demand*.

NoAH was able demonstrate the effectiveness and utility of a full-scale NoAH infrastructure, through a preliminary pilot operation. At the end of the project, the NoAH pilot was up and running with more than 8 sites currently contributing address space, including parties external to the project, thanks to mutual NDAs. A live stream of alerts is provided to the community by means of the Polecat web application. NoAH is currently providing open access to **sanitised attack information** to the security research community on demand. Such a repository of information has the potential of boosting research and development in the area of attack detection and containment.

Throughout this process, NoAH has worked closely with ISPs, NRENs, and CERTs outside the NoAH consortium, to ensure relevance and increase potential impact, and has aggressively publicized the project results to the security research community and European industry. NoAH has aggressively promoted a collaborative approach to security data collection and sharing, and has **contributed to mobilizing the European systems security research community** by organizing or having a leading role in the organization of workshops collocated with the TERENA Networking Conference 2006 (TNC), the ACM SIGCOMM'06 Conference, the IFIP Conference on Multimedia Security (CMS'06), the ACM CCS Workshop on Network Data Anonymization, and the two European Conferences on Computer Network Defence (EC2ND'07 and EC2ND'08). Furthermore, insights from the project have been disseminated through papers and publications in high-profile media. Some of the project's deliverables have been repeatedly downloaded from hundreds of users searching for more information on honeypots and how to combat cyberattacks. Finally, the members of NoAH have been instrumental in creating EuroSec: the European Workshop on System Security: the first of its kind in Europe.

A.3 *Project impact*

A.3.1 Impact on the state-of-the-art

Within the specific context of large-scale cyber-attack detection and containment using honeypot technology, NoAH has made significant contributions, both technical and non-technical. Perhaps one of the most important contributions to the state of the art is that NoAH is one of the very few projects that is both *open* and *inclusive*. Previous honeypot initiatives have been privately owned and operated and researchers had not been able to obtain the data they needed for research purposes. Other initiatives operated by volunteers were similarly restricted to industry insiders and trusted parties and the data feeds were pre-computed summaries that are of limited use to researchers. In contrast, NoAH explored the construction of a more open and inclusive architecture with the goal of boosting European security research.

In addition to being research-focused, NoAH also opened the architecture to normal citizens and end users, which can contribute data to the infrastructure through components such as Honey@Home. The *Honey@Home* initiative opens up NoAH and the security infrastructure space to people that are not familiar with honeypot technologies. Empowering non-expert users to contribute is likely to impact along two dimensions. First, the NoAH project will gain visibility outside the research community, while also increasing awareness on security issues and attack trends. By **empowering people** to directly contribute to improved situation awareness and security, the sense that attacks can be detected and prevented fast is likely to appeal to common users. By helping security researchers, we contribute to the effort of making life on the Internet safer. Second, it will allow security monitoring initiatives to significantly increase the fidelity and accuracy of the data collected. The research community will directly benefit from studying a more diverse set of attacks taking place globally and not only by the limited number of honeypots maintained by NoAH partners. This is particularly important as attackers get smarter and start evading legacy honeypots through their own internal blacklist mechanisms.

Similar expectations hold for the other NoAH components as well as the architectural "glue" that binds them together. For example, **Argos is widely used by researchers in Europe**, the US, and Asia, having reached its 4th release and with **more than 2,700 downloads**, and it has become a key component in the *Scriptgen*[1] effort from *Eurecom*. The overall NoAH architecture is also likely to be useful to other research efforts in Europe and beyond. Among the most important technical contributions that improve upon the state of the art are the first in-depth exploration of taint analysis as a method of attack detection, the wider exposure and resolution of the privacy issues surrounding honeypot technologies, and potentially the largest deployment of privacy-enhancing overlay technology known to date.

A.3.2 Impact on the European scientific community

Being a research infrastructure project, NoAH aims to serve the European scientific community, and specifically the emerging community of security researchers in both European Universities and Research Labs, as well as the growing number of European security industry players.

[1] http://www.acsa-admin.org/2005/papers/118.pdf

The most important contribution was in demonstrating the feasibility of a research infrastructure to support the European scientific community. With an open architecture and low operational costs, NoAH is expected to continue operation after the end of the project, providing infrastructure access and valuable datasets for security research in Europe. Furthermore, the lessons learned from the NoAH experience are likely to help guide future efforts, including those of other projects in this space such as FORWARD and WOMBAT.

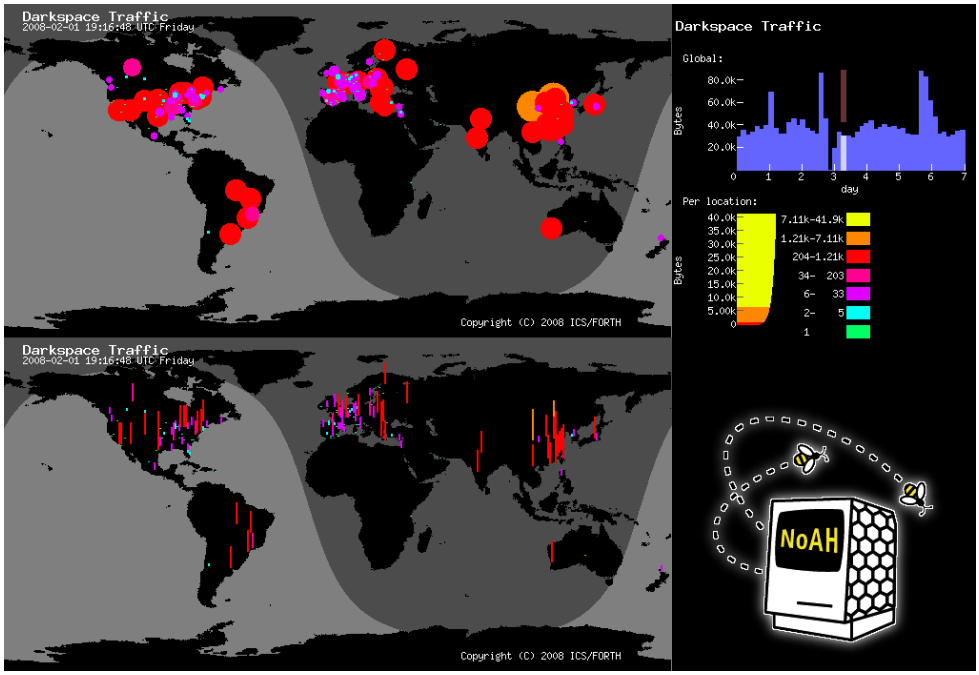


Figure 18: Geographic Origin and volume of attack traffic received by NoAH

Another important contribution of NoAH was in helping towards **bringing this scientific community together**, initially through the need to address a large set of common technical and non-technical challenges. The process for this was jumpstarted through the two NoAH sponsored workshops co-located with the NREN's TERENA conference and the networking research community's SIGCOMM conference, and the effort continues today through the European systems community's ACM EuroSec workshop[1] in 2008 and 2009, and the deep involvement in the ENISA-sponsored EC2ND conference in 2007 and 2008. NoAH had a leading role in all those efforts with NoAH partners chairing or co-chairing those events.

NoAH also helped the European scientific community achieve greater visibility of European research worldwide. In addition to the SIGCOMM co-located workshop, NoAH was also involved in the creation of the ACM workshop on Network Data Anonymization, motivated by the anonymization challenges the project

[1] http://www.cs.vu.nl/eurosec08/

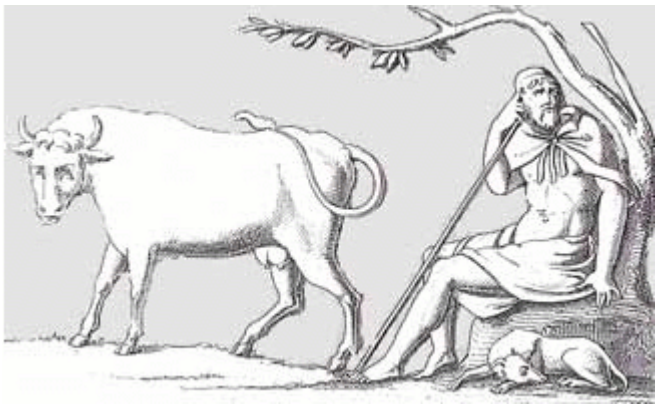
faced towards achieving greater openness and security. This serves the European scientific community well, as it enables higher visibility and closer cooperation with non-European organizations.

A.4 *Dissemination and use*

Summary of exploitable results

A.4.1 Argos

One of the most popular components developed by NoAH is the Argos attack detection system. It is the first practical system to implement information flow tracking for the purposes of detecting code injection attacks, and its value has been already appreciated in the research community. Given its proven value, the visibility in the security community and the relative maturity it has reached it is a natural candidate for potential exploitation.



There are several possible paths for exploiting Argos further. At the moment, Argos is actively being developed as a standard open-source project. One of the apparent advantages of open source projects is that they can spread development cost for systems that would otherwise not be sustainable if developed by a single party. With the right amount of effort from the "owning" partner, Argos could very well continue evolving and be useful to the research community and the security industry, much like other prominent open-source security projects such as honeyd, metasploit, and nessus. Building a larger Argos development community would benefit from wider promotion of the technology and its potential benefits. Potential partners in this activity could be research universities, infrastructure providers, software vendors wishing to protect instances of their own software, security consulting companies, managed security service providers, and intrusion detection companies that already use or are open to the use of open source software as part of their products or backend analysis infrastructure.

Commercial exploitation would require significant investment to drive Argos towards a product-ready state and combine it with other technologies necessary for tackling a subset of the current security challenges. The most promising exploitation route here would be teaming with a large security vendor willing to co-fund further development and license the technology afterwards. A new, standalone company built around Argos would require a lot more effort, but is likely to be commercially viable given the gaps in attack analysis tools on the market, the proven value of Argos, and the technical barriers which the Argos team has already managed to overcome.

A hybrid model is also possible, with an open source version available to the public, and accelerators or deeper analysis modules available on a commercial basis. As with the commercial

option, this could be achieved more easily through means of licensing out to a resourceful security vendor.

A.4.2 Honey@Home

One of the innovations in the NoAH project was the Honey@Home component. Honey@Home enables regular end-users to participate in a honeypot network and contribute security-related data. This allows honeypot infrastructures to draw from a larger and more diverse pool of resources, making the infrastructure more scalable, information-rich, and resilient compared to monolithic designs that often sit on a single-prefix of dark space. There are several ways in which this can be exploited further:

1. End-user/consumer software security product vendors: Security companies can license out this system to enhance their existing monitoring infrastructures. The client-side component is a mature prototype and needs modest effort to be turned into a full product. NoAH can license out the implementation and provide consulting to any company wishing to adopt this technology. The target market includes both small security companies as well as larger and well-established players in the space of end-user security (AV vendors, UTM vendors, security suite vendors). The value proposition here is that small companies may not be able to afford a full honeypot infrastructure and Honey@Home could offer them a low-cost alternative, while large companies may see this as a way to improve the quality of their data feeds. The IP rights for Honey@Home are with NoAH partners, who can engage in R&D contracts, license out technology, and provide consultancy to any company wishing to commercialize the technology along this direction.
2. Managed Security Service Providers (MSSPs): Honey@Home can very easily adapted to use as sensors by MSSPs. MSSPs serve small and medium-sized businesses that cannot afford to operate their own security monitoring infrastructure, in essence remotely managing IDS and IPS systems on behalf of their customers. The modified honey@work design could complement existing managed IDS/IPS monitoring with observations of an organizations dark address space and dark port space, something that is not done to date. The IP rights for Honey@Home are with NoAH partners, who can engage in R&D contracts (e.g., to further develop Honey@Home and produce a honey@work version), license out technology, and provide consultancy to any company wishing to commercialize the technology along this direction.
3. Independent security data company: Data from a large number of Honey@Home sensors can be highly valuable to a diverse set of users: security companies, security researchers, CERTs, law enforcement agencies, etc. An independent and perhaps even standalone company can take up the task of developing Honey@Home, recruiting sensors, and collecting and post-processing the data, and then sell a value-added data feed to interested parties. This could be achieved by a new company, or be taken up as a task by one of the large number of existing security consulting companies that carry out tasks such as vulnerability research, penetration testing, etc. The challenge in this direction is to provide incentives to users to install the Honey@Home sensor. Some of the options would be some free or subsidized service in exchange (e.g., free security alerts, subsidized anti-virus, visibility of “top contributing teams”, etc.) or even the originally envisioned attack-visualising screensaver. The IP rights for Honey@Home are with NoAH partners, who can

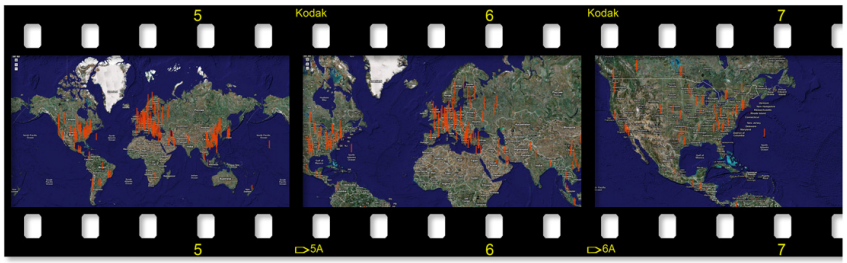
engage in R&D contracts, license out technology, and provide consultancy to any company wishing to commercialize the technology along this direction.

4. A community initiative: similar to SETI@home and Folding@home, one could envision a free-for-all community initiative around Honey@Home. Other projects have shown that this is feasible using minimal funding (usually as a side-activity in a research project or funded internally by the hosting institution). Although this may not be attractive from a revenue point of view, the community nature of such an effort is likely to attract significantly higher number of contributors compared to a similar effort run by someone for-profit. The indirect benefits to the community are, however, significant, and the organization that will take this effort up is likely to enjoy great visibility in the security community. The IP rights are with the original Honey@Home team, which can also participate in such an effort through a suitable partnership.

A.4.3 Operational NoAH infrastructure

The NoAH project successfully carried out a design study towards a security research infrastructure, demonstrating that it is feasible to provide a public and shared system for obtaining security related data feeds and carrying out security-related experiments. One possible exploitation direction would be to operate this infrastructure at a larger scale.

Such an effort could potentially be economically viable on its own, but given the primary goal to serve the research community in addition to industry, it is natural to first draw upon public funds to build the full infrastructure and start operations. The ultimate goal would be to make this activity pay for itself through membership fees similar to other European-scale organizations such as RIPE-NCC and TERENA. Given the scope of the work, it also seems reasonable to explore whether publicly-funded CERTs at the national level see this as an activity that could serve their constituencies.



Through the SSA action, the NoAH infrastructure has reached pilot-grade maturity. This is beyond the demonstrator phase, but not at industrial grade maturity yet. Therefore, it will be necessary for any interested party to engage further with R&D organizations including, but not limited to, NoAH partners.

Because the NoAH infrastructure involves multiple components, the IP rights are covered by the consortium agreement. For this potential exploitation line, the NoAH partners can be involved in further development through R&D contracts, license out technology, provide consulting and training, and even participate as early customers.

B. FINAL MANAGEMENT REPORT (FINANCIAL INFORMATION)

This section is intentionally left blank. The contents of this section will be separately submitted to the EC.

C. FINAL REPORT ON THE DISTRIBUTION OF THE COMMUNITY FINANCIAL CONTRIBUTION

This section is intentionally left blank. The contents of this section will be separately submitted to the EC.

D. QUESTIONNAIRES

This section is intentionally left blank. The contents of this section will be submitted using the on-line SESAM tool.