KYPO: A Tool for Collaborative Study of Cyberattacks in Safe Cloud Environment

Zdenek Eichler, Radek Ošlejšek, Dalibor Toth
xeichler|oslejsek|xtoth2@fi.muni.cz

Masaryk University,
Faculty of Informatics,
602 00, Brno, Czech Republic,

Abstract. This paper introduces the KYPO – a cloud-based virtual environment faithfully simulating real networks and enabling users to study cyber attacks as well as to train users in isolated and controlled environment. Particularly, the paper focuses on the user environment and visualizations, providing views and interactions improving the understanding of processes emerged during experiments. Web user interface of the KYPO system supports several collaboration modes enabling the participants to experiment and replay different types of security related tasks.

Keywords: Human-Computer Interaction, Collaboration, KYPO, cyber security

1 Introduction

Cyber attacks become more and more sophisticated and frequent. Internet users face cyber attacks on everyday basis in the form of phishing e-mails, infected attachments or intrusion attempts. A viable option to study attacks and to train users is the simulation of cyber threats in isolated, controlled, scalable and flexible cloud-based environment enabling participants to experience and replay various scenarios in order to understand the impact of the attack on users and devices involved in the infrastructure.

There are many testbed solutions intended to support cyber security-related simulations and training programs in various manners. Some of them, namely DETER [2] and TWISC [3], employ the generic and publicly available *Emulab/Netbed* [13] infrastructure solution, which provides them with basic functionality for virtual appliances' deployment, flexible network topologies configuration, various network characteristics emulation, etc.

In contrast, several security-related testbeds require their own infrastructure solution to be established, which cannot be used for other purposes. For example, ViSe [1], LVC [11], and V-NetLab [8] testbeds employ the VMware virtualization, while the hypervisor-based security testbed [4] requires a KVM-based infrastructure. All these cases require to purchase and establish a dedicated infrastructure, which brings both strengths and weaknesses by itself – while the full control over

the infrastructure can lead to easier deployment of testbed's features, it also leads to high initial costs and limited growth-flexibility. The flexibility and scalability of this lowest layer represent the key factors for possibility to create as many computer networks as needed for specific exercise scenario from the perspective of collaboration.

As another perspective can be considered integrated user environment for specific user roles and use cases. The main goal is to provide access to specific device or computer in testbed. Next important functionality is based on special visualization approaches and analytical tools, usually narrowly focused on particular aspects of network monitoring and utilized by network administrators or security analysts. The level of user interfaces (UI) differs from project to project according to its main purpose, but the majority provides only basic administration of virtual networks and users operate via traditional ways, typically SSH connections to every machine.

Next section describes the KYPO platform, which is used for management of environments for cyber security scenarios described in the paper. Third chapter briefly presents visualizations used by exercise participants for better imagination and understanding. Following chapters discuss collaboration cases of training programs, which are used in KYPO scenarios and provides user experience evaluation.

2 KYPO Architecture

KYPO testbed platform depicted in Figure 1 provides the environment for modeling and running virtual computer networks. These networks serve as isolated environments for controlled analysis of various cyber attacks as well as for cyber security training programs [7].

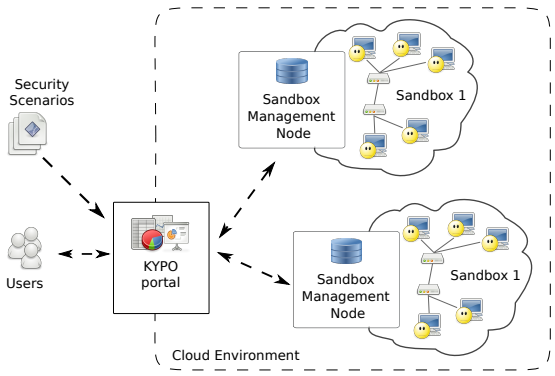

Fig. 1. KYPO Architecture

Security Scenarios are employed in the whole life-cycle of cyber experiments or training programs. They represent a basic document describing the plan and necessary details similarly to screenplays in movie production. Its well-structured JSON format encodes participant roles (e.g. attacker versus defender), their goals, detail instructions, roles of network nodes (e.g. mobile phone of attacker versus server to be compromised), network topology, characteristics of network links and nodes, etc. KYPO provides several predefined templates covering various security interests and domains like DDoS attack simulation, phishing, or simple hacking game. An example of a simple security scenario focused on DDoS attack simulation can be found in [6].

Network-related data encoded in a scenario are used by administrator who is responsible for the preparation of concrete training session. The scenario is uploaded to the administration interface of **KYPO portal**, which mediates access to the KYPO infrastructure for both administrators and participants. The network-related data are processed by the KYPO virtualization subsystem, which automatically allocates so called sandboxes.

Sandbox represents isolated computer network where users can safely perform their tasks. Network infrastructure of sandboxes is fully virtualized. Both nodes and links are build on top of a cloud managed by OpenNebula [9]. This approach provides scalable and flexible solution. Sandboxes can be allocated on demand and accessed remotely without the necessity to maintain hardware devices for each individual security experiment. The abstract network layers simulated by the cloud are transparent for running applications which are hardly able to detect the fact that they are not running on a physical network. The illusion of a real hard-wired network is therefore nearly perfect for both running software and users.

Once a sandbox is allocated in the cloud it can be accessed by authorized participants via KYPO portal. The portal provides users with instructions, various views on network state and also allows them to interact with the network. For example, users can connect to individual computers via VNC and then launch programs and commands on them, everything via web browser.

Activities within a sandbox are monitored by probes [5, 12]. Measured data, e.g. network traffic, CPU load or security events, are stored in a database deployed in so called **Sandbox Management Node**, SMN. Every sandbox has its own SMN serving as a data repository for experiments performed in the sandbox. These data are used to provide comprehensible visual feedback to the users via interactive visualizations running at KYPO portal.

Since our tool is designed for students, sandboxes must be easy remotely accessible. Accessibility was ensured by employing the concept of Web applications with minimal requirements on web browsers. As the most fitting approach was chosen an unifying environment of enterprise portal according to its component based architecture.

3 Visualizations

The system provides various visualizations developed specially for educational purposes, where tutor defines which visualizations should be accessible, depending on particular scenario. All visualizations are interactive and follow the Shneiderman's visualization mantra[10]: *Overview first, zoom and filter, then details-on-demand*. One of provided visualization developed for education is a 3D sequenced radar chart, which visually compares multiple variables in time. The visualization is implemented in WebGL in order to deliver accelerated visualization in Web environment. The surface of the solid figure (Figure 2) is a result of the composition of ordinary radar charts along a time scale.

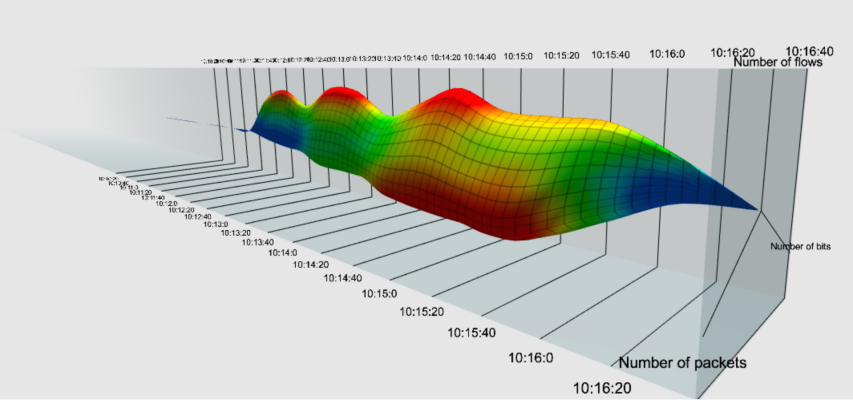

Fig. 2. 3D sequences radar chart

A network topology visualization is presented in Figure 3. Subjects of visualization are routers, links, computers and servers. Every node in the topology can be accompanied by a small sign, which represents the role of the node in the running scenario (e.g. attacker or victim). Supported is also visualization of data flow on particular links. This visualization also enables students to open e.g. a VNC (remote access) connection to the computer in sandbox and share the screen of remote computer with other students or lecturer.

4 Collaborative Environment

Our main focus is the security training programs, where the main advantage of our approach is the authenticity. Instead of describing the key principles of cyber attacks theoretically, we rather let students to try to perform a real cyber attacks or let them e.g. to become victims of a phishing attack, all in safe virtual environment. The system provides easy to use user environment, where a lecturer

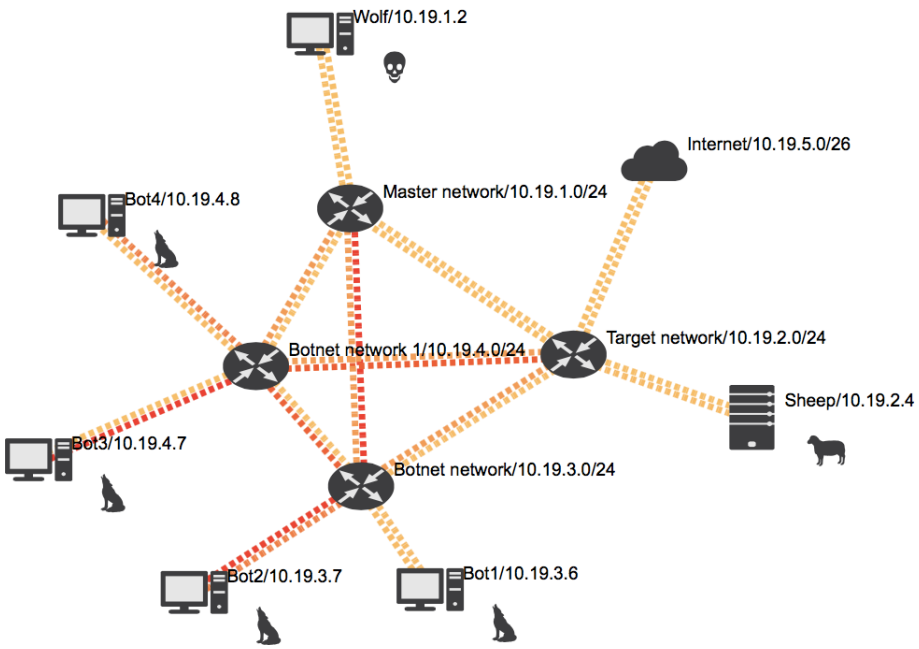

Fig. 3. Visualization of network topology

is able to easily define a huge amount of attributes which will be measured in the network during cyber experiments and then presented to students either in real-time or after the experiment.

Security scenarios can significantly differ in the way how the users collaborate. There can be many sandboxes and many training programs prepared or running in KYPO at the same time. In what follows, we discuss collaboration modes of students involved in the same training session. These modes are schematically suggested in Figure 4.

4.1 Individual views on shared data

Imagine DDoS security scenario. It aims to illustrate principles and impacts of several variants of distributed denial-of-service attacks. The attack is driven by the lecturer who runs appropriate commands in particular sandbox. The state of the sandbox is monitored, measured and recorded in the database running on the Sandbox management node. The DDoS attack can be performed online during the training session or in advance.

Students, each of them sitting at his or her own computer, share the sandbox and the measured data. They have typically the same set of visualizations at their disposal. In the case of DDoS scenario, the most useful are the visualization of network topology emphasizing computer roles (attacker, bots, sheep) together

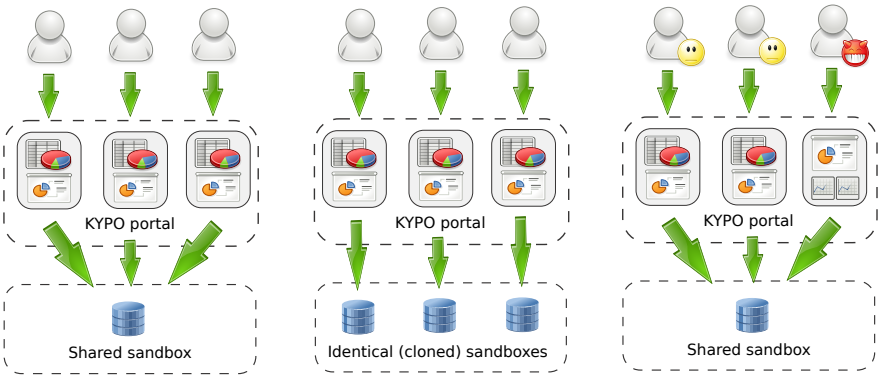

Fig. 4. Collaboration modes: Individual views on shared data (left), individual sandboxes (middle) and role-based collaboration (right)

with the utilization of links, as shown in Figure 3, and also analytical tools like 3D sequence radar chart depicted in Figure 2, which shows detailed link parameters on a single selected line. Although the visualizations are common for all the students involved in the training session they provide individual views on shared data. Students can focus on different links or return back in time without affecting the views of other students.

4.2 Individual sandboxes

Imagine a different scenario, where the users should try to compromise a computer. In this case, every participant should have its own private sandbox, in order to handle the attack on its own. Thanks to the cloud-based infrastructure of KYPO it is easy for the lecturer to allocate many identical sandboxes for individual users on demand. The sandboxes have the same network topology, network parameters, software running on nodes, and other aspects. Events and developments of the scenario caused by users in their sandbox are measured and stored inside this sandbox. Therefore, the KYPO is able to provide per-user data after the training session.

4.3 Role-based collaboration

Also mixed approach is supported, where students share particular sandbox and every student has it's own role in the scenario, operating different computers. A typical example are so called "capture the flag" games, where groups of participants have access to different vulnerable computers and the goal is to compromise computers of the other groups. Another popular variant of this game defines a group of defenders protecting a vulnerable network and a group of attackers

trying to compromise the network. In both the cases, students must cooperate within their groups although they are sitting at their own computers.

Security scenarios of KYPO system enables to define arbitrary roles. They also define which computer is accessible by which role. During the preparation of a training session, the lecturer of the session assigns roles to individual user accounts. The access to the computers inside the sandbox is protected by authentication. Therefore, during the session the KYPO portal provides users with the authentication data with respect to their role and the level achieved in the game.

4.4 Face-to-face Collaboration

Another use case, instead of the above mentioned remote collaboration, is a face-to-face collaboration. Remote collaboration through web portal enables collaboration of participants through network disregarding the geographic location. On the contrary, local face-to-face collaboration enables participants to collaborate during discussion. For this purpose, we are using the Leap Motion device, which helps participants to interactively collaborate when sharing the same computer without e.g. exchanging a mouse. The Leap Motion controller is a small USB device which captures movements of a hand performed above the device. The software recognizes particular gestures, which are then send to our visualizations. Currently, all visualizations described in Section 3 can be controlled by the device.

5 Evaluation

Our system was already presented at several cyber security workshops and conferences. Online demo at NOMS 2014 conference was focused on a DDoS scenario simulated in KYPO platform [6]. More complex variant of the attack with 40 virtual machines divided into 6 sub-networks was demonstrated during the tutorial named *Cybernetic Proving Ground: a Cloud-based Security Research Testbed*[1] attached to AIMS 2014 conference. The UI was used for the overview over the network topology traffic during the simulation and for replays of the whole scenario during the presentation for workshop members.

The second part of the AIMS tutorial was focused on a hands-on training session prepared in a form of game. The main goal was to compromise a server in a company network and to abuse it as an attacker in a DDoS attack. All 20 participants had heir own sandbox with prepared environment (several machines for this scenario) and several tasks to reach the goal (win the game). This game was successfully repeated at FIRST/TF-CSIRT Technical Colloquium[2] in 2015 with 25 involved participants.

[1] http://www.aims-conference.org/2014/labs.html
[2] https://www.first.org/events/colloquia/laspalmas2015

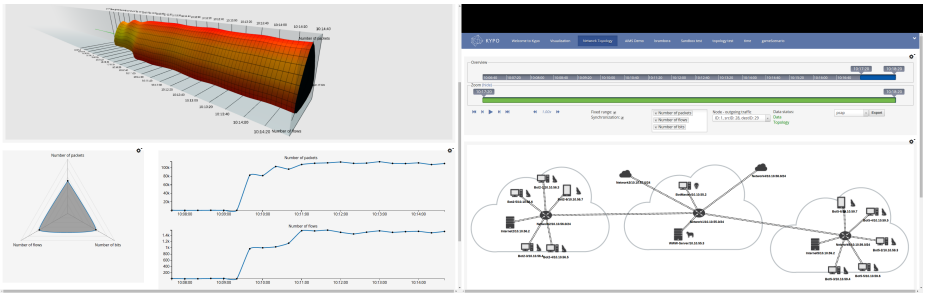

Fig. 5. Screen shot of the KYPO portal during the hands-on training (dual display mode)

During these workshops, no significant issues were detected. Unfortunately, no formal qualitative or quantitative feedback was collected. The formal evaluation of the KYPO system was therefore conduced on 10 university students of FI MU[3]. This preliminary evaluation brings promising results, as discussed in what follows.

5.1 Evaluation Process

At the beginning, subjects were asked to evaluate their knowledge about hacking (infiltration to the system) and DDoS attack. Then, subjects logged in to the system and every three or two students shared a sandbox. Instructions were provided to subjects in a form of a level based game very similar to that presented at the AIMS and TF/CSIRT Technical Colloquium, which led the students through the scenario. The goal was to compromise target system and then run DDoS attack from the compromised system. Every subject had its own computer and they were able to collaborate by sharing the screen of the attacker's computer through our web portal (VNC) and view various visualizations described in Section 3. When all subjects finished the game, the subjects were asked to the same questions again (same as before the course).

5.2 Results

The subjects evaluated their knowledge about DDoS and hacking on five-point Lickert scale (1 for *I don't know nothing about that*, 5 for *I'm able to perform such an attack*). The difference between before and after the course showed increased knowledge in all subjects. Comparison of DDoS knowledge and hacking knowledge is depicted in Figure 6. Subjects also evaluated the course itself, also on five-point Lickert scale (1 for *Strongly Disagree*, 5 for *Strongly Agree*) on following statements (mode is a value that appears most often):

[3] http://www.fi.muni.cz

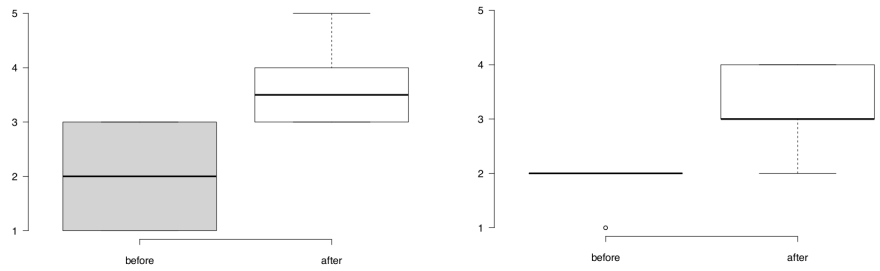

Fig. 6. Tukey boxplot displaying knowledge before and after the course: DDoS (left) and hacking (right)

- I enjoyed the course. (mode = 4)
- I learned something new. (mode = 4)
- I enjoyed the ability to perform real attack in safe and collaborative environment. (mode = 5)

6 Conclusion

In this paper we have presented a cloud-based research testbed for the simulation and visualization of network attacks, focused on education and practical exercise. Chosen web-based portal technology presents a flexible and scalable solution which allows users to collaborate through various interconnected visualizations in provided web portal satisfying the requirements of broader range of training programs. Usability of our solution was verified by practical demonstrations focused on DDoS attacks and a "hacking game".

Practical evaluation and subsequent survey indicate that the proposed collaborative virtual environment equipped with user friendly interactions could be beneficial for efficient understanding of security threats as well as for the safe forensic analysis of suspicious code or devices. Our next work is therefore aimed to enhancing collaborative tactics supported by smart and intuitive interactions and visualizations.

Acknowledgments. This work has been supported by the project "Cybernetic Proving Ground" (VG20132015103) funded by the Ministry of the Interior of the Czech Republic. We appreciate the access to computing facilities *(a)* owned by parties and projects contributing to the National Grid Infrastructure MetaCentrum, provided under the program "Projects of Large Infrastructure for Research, Development, and Innovations" (LM2010005), and *(b)* provided under the programme Center CERIT Scientific Cloud, part of the Operational Program Research and Development for Innovations, reg. no. CZ. 1.05/3.2.00/08.0144.

References

1. A. Arnes, P. Haas, G. Vigna, and R. A. Kemmerer. Using a virtual security testbed for digital forensic reconstruction. *Journal in Computer Virology*, 2(4):275–289, 2007.
2. T. Benzel. The science of cyber security experimentation: The deter project. In *Proceedings of the 27th Annual Computer Security Applications Conference*, ACSAC '11, pages 137–148, New York, NY, USA, 2011. ACM.
3. L. Chen. Construction of the New Generation Network Security Testbed-Testbed@ TWISC: Integration and Implementation on Software Aspect, 2008. Institute of Computer & Communication, National Cheng Kung University, Tainan, Taiwan.
4. D. Duchamp and G. De Angelis. A hypervisor based security testbed. In *Proceedings of the DETER Community Workshop on Cyber Security Experimentation and Test on DETER Community Workshop on Cyber Security Experimentation and Test 2007*, DETER, Berkeley, CA, USA, 2007. USENIX Association.
5. R. Hofstede, P. Celeda, B. Trammell, I. Drago, R. Sadre, A. Sperotto, and A. Pras. Flow monitoring explained: From packet capture to data analysis with netflow and ipfix. *Communications Surveys Tutorials, IEEE*, PP(99):1–1, 2014.
6. T. Jirsík, M. Husák, P. Čeleda, and Z. Eichler. Cloud-based security research testbed: A ddos use case. In H. Lutfiyya and P. Cholda, editors, *Proceedings of the Network Operations and Management Symposium (NOMS 2014)*, Krakow, Poland, 2014. IEEE Xplore Digital Library.
7. D. Kouřil, T. Rebok, T. Jirsík, J. Čegan, M. Drašar, M. Vizvăry, and J. Vykopal. Cloud-based testbed for simulation of cyber attacks. In H. Lutfiyya and P. Cholda, editors, *Proceedings of the Network Operations and Management Symposium (NOMS 2014)*, Krakow, Poland, 2014. IEEE Xplore Digital Library.
8. K. Krishna, W. Sun, P. Rana, T. Li, and R. Sekar. V-NetLab: a cost-effective platform to support course projects in computer security. In *Proceedings of 9th Colloquium for Information Systems Security Education*, 2005.
9. D. Milojicic, I. M. Llorente, and R. S. Montero. OpenNebula A Cloud Management Tool. *IEEE INTERNET COMPUTING*, 15(2):11–14, MAR-APR 2011.
10. B. Shneiderman. The eyes have it: A task by data type taxonomy for information visualizations. In *Proceedings of the 1996 IEEE Symposium on Visual Languages*, VL '96, pages 336–, Washington, DC, USA, 1996. IEEE Computer Society.
11. B. Van Leeuwen, V. Urias, J. Eldridge, C. Villamarin, and R. Olsberg. Performing cyber security analysis using a live, virtual, and constructive (LVC) testbed. In *Military Communications Conference 2010 - MILCOM 2010*, pages 1806–1811, 2010.
12. P. Velan and R. Krejčí. Flow information storage assessment using ipfixcol. In *Dependable Networks and Services*, volume 7279 of *Lecture Notes in Computer Science*, pages 155–158. Springer Berlin Heidelberg, 2012.
13. B. White, J. Lepreau, L. Stoller, R. Ricci, S. Guruprasad, M. Newbold, M. Hibler, C. Barb, and A. Joglekar. An Integrated Experimental Environment for Distributed Systems and Networks. In *OSDI02*, pages 255–270, Boston, MA, Dec. 2002.