DEPARTMENT OF INFORMATICS

TECHNISCHE UNIVERSITÄT MÜNCHEN

Master Seminar Thesis in Informatics

Virtualized Security Education Platform as a Service

Ahmet Turkmen

DEPARTMENT OF INFORMATICS

TECHNISCHE UNIVERSITÄT MÜNCHEN

Master Seminar Thesis in Informatics

Virtualized Security Education Platform as a Service

Author:	Ahmet Turkmen
Supervisor:	Dr.-Ing. Thomas Hüttl
Advisor:	Dr.-Ing. Thomas Hüttl
Submission Date:	22.08.2022

Abstract

Institutions and companies are seeking specialized way of learning in digitized education system. There are many online education platforms exists today. However, the gap between the demand and existing education solutions are increasing rapidly. Specialized set of tracks in online education is needed to address the growing demand. It is included in many online platforms such as Coursera, Udemy, Udacity and more. However, they are in general lack of practical exercises or cheating is relatively easy. Personalized Cyber security education in practice is generally not available in online education. There are platforms for practical cyber security education, such as Hackthebox, picoCTF, and more. However, either they do not provide personalised exercises for the courses at universities, colleges and high schools. This platform, named Haaukins, is a specialized virtual platform for cyber security education and exercises are free and specialized for the institutions and companies. The platform itself is open source and can be found at Github[1]. In this seminar thesis, I will be discussing about how Hauukins platform evolved in time to fulfill the demands of instructors at universities, colleges and high schools.

[1]https://github.com/aau-network-security/haaukins

Contents

1 Introduction

Cyber security education is in high demand in any field which has critical responsibilities in managing energy systems, monitoring airplanes, hospital management and many other places. Recent incidents in Cyber security field shows that attackers are targeting vital points of a country or a city. According to CSIS[1], twelve crucial cyber attacks occurred, to different countries around the world in June 2022, for instance one of them is targeted Lithuania's public infrastructures such as state railways, airports.[1] Individuals are target of cyber crime, today, everyone receives a millions of dollars from their distanced relatives with a mail to click. Surprisingly this phishing technique is still used and it works even though, not at targeted rate.

In this paper, an open source virtualised security education platform, named Haaukins, will be discussed in depth and its working mechanism will be investigated in detail. The main aim of the platform to increase self awareness of people about cyber security and provide practical challenges to the participants. The platform does not require any preliminary information and includes variety of challenges at different levels. Initially, the platform is built for high-school students to have very basic challenges and simple access to cyber security challenges. Since existing platforms usually require virtual private network (VPN) to access virtualised environment, which brings overhead for high school students. Additionally, the platform itself is completely open source and can be deployed to any Debian based machines at universities, institutions, high schools or even individual local computer. In time, Haaukins proved its success in Denmark, then advanced challenges are included for experienced users. The management of the platform was not at good condition due to the fact that administrators of the platform should install command line tool to manage the platform. As the platform usage increased and spread through universities, management of the platform should be easier and no one should require to install any tool to manage the platform. Therefore, administrator side as well as challenges inside the platform developed according to feedback received from users of the platform. The platform includes practical exercises which are in the format of Jeopardry Capture-The-Flag(CTF).[2] The format is also being used by other security education platforms such as Hackthebox[2], picoCTF[3]. Main idea of "CTF" style is to have a secret string, called 'flag', in a virtualized vulnerable environment and participants of the game should exploit the vulnerability to receive value of the 'flag'. The challenges in Haaukins platform is developed parallel to curricula of educational institutions. The content of the challenges in the platform shaped by lecture contents at universities or high schools, to consolidate theoretical information with real life practice.

[1] Center for Strategic International Studies
[2] https://www.hackthebox.com
[3] https://picoctf.org

2 Related Work

In cyber security community, there are already existing platforms to gain practical experience over browser. "Hack the box", is one of the well-known cyber security education platform however both working principle and its skeleton is different. It is commercial market player rather than non-profit organization, which brings some barriers to some people. It has both VPN and virtual machine on browser feature, yet, it is limited with two hours in free plan.

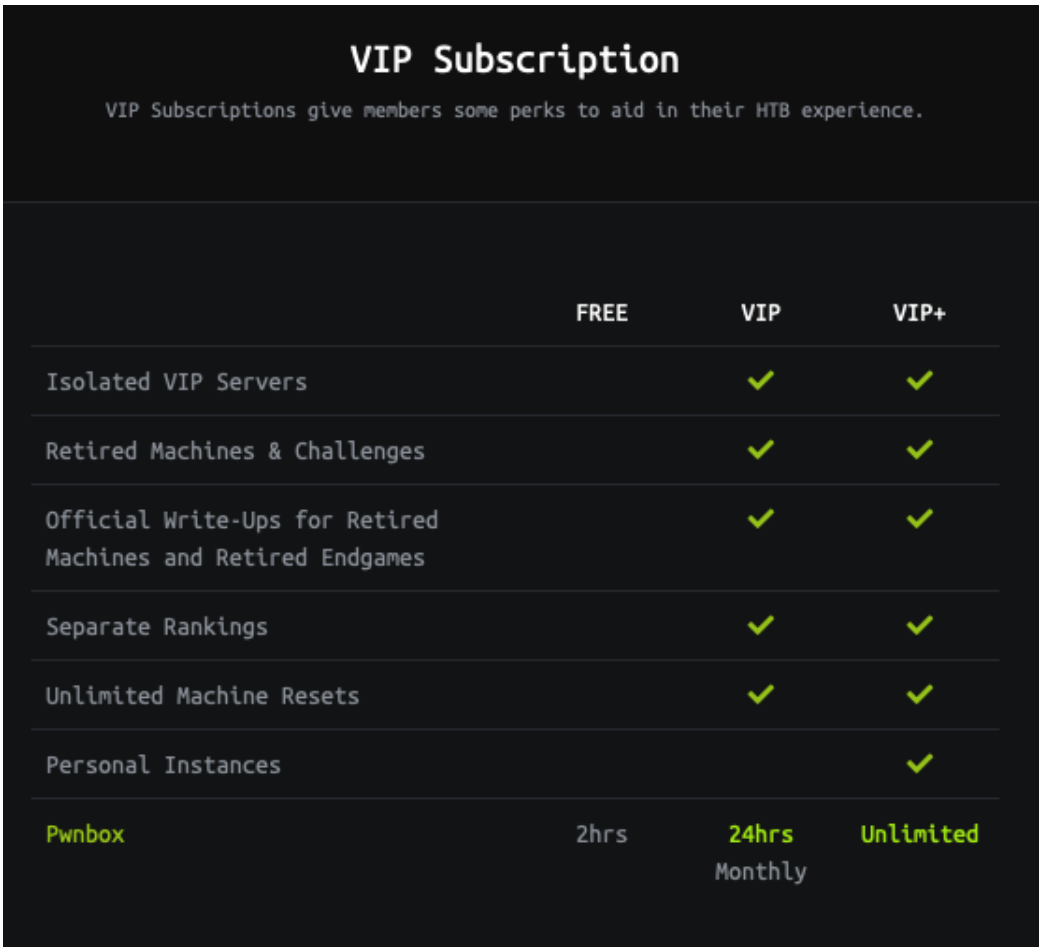

Figure 2.1: Taken from https://www.hackthebox.com/hacker/pricing

"Hack The Box" platform is great, extensive and well-known, but, it expanded in all kinds of fields from business to certification programs. It makes difficult for the people who want to have easy start with simple approach. Although the platform itself has great resources,

user interface, it lacks from simplicity. Here, the target group of "Hack The Box" differs, the argument can be discussed from variety of angle.

Another platform called "TryHackMe", it also contains all different challenges, exercises and virtual machine on browser feature, enables participants to install nothing on their computer. Similarly, "Hack The Box" platform, it has commercial purposes which are absolutely fair. However, considering some beginners who has no information about installing virtual private network to a computer, they require a subscription.

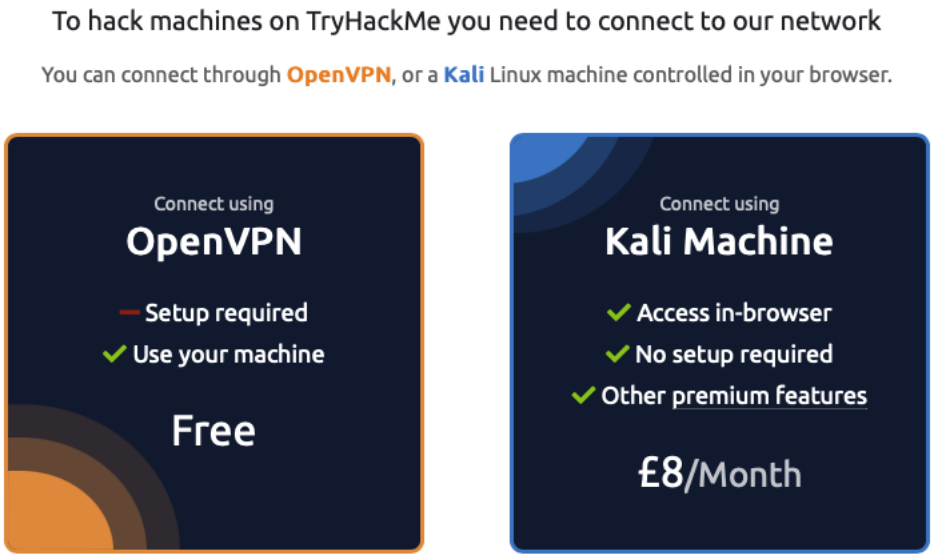

Figure 2.2: Taken from https://tryhackme.com/access

Even though they are affordable at some extent, the main problem comes when customized, curricula based exercises are required by universities. Furthermore, there is no full control of platform by universities in any of mentioned platforms, except Haaukins.

Another platform which is also developed by a university is PicoCTF[3]. It provides variety of challenges in different categories in a game based environment. It is mostly competition based platform where users has profiles and gain points in time. The points are permanent in the system and can be build on top of them similar to other commercial platforms. It is opposite of what Haaukins platform is providing. Although existing platforms are providing decent environments for cyber security practices, they have different target groups, approaches and policies in terms of their usage. Additionally, the requirements of an exercise in a lecture varies, ideally, the exercise should not be publicly available to prevent cheating. It is hard to achieve for any e-Learning platforms in particular when it is mostly practical exercises. Most of the solutions to exercises on well-known platforms are exists on the internet. Last but not least, challenges developed specifically for a lecture/organization stays private, only administrators can have an access to the solution of the challenges on Haaukins platform.

3 Architecture

In this section, technical and non-technical details of Haaukins platform will be discussed. The platform has main components which are driven by requirements of creating virtual labs in seconds with an automated way. The idea comes from a real life problem, imagining a group of students who competes with each other to get a flag in CTF (Capture The Flag) style event. Distributing computers to have same environment among participants of an event, is not feasible and fast. Running the event over the commercialized platforms could be problematic due to limited access to the platform and lack of curricula based challenges. Therefore, there is a need of automated, easy and highly accessible platform which can be managed entirely. Haaukins born with this information in mind, initially it had four main goals which are fully automated, transparent, highly accessible and realistic[4]. Fully automated, deployment of newly created exercises and assigning virtual labs are automated completely through continuous integration and deployment. It includes creating virtual networks for each participants and providing isolated set of exercises. Transparent, each participant for an event running Haaukins platform, assigned to have isolated vulnerable environment. It prevents other participants to interact with other team, to abuse their learning process. Highly accessible, the platform has buffering mechanism, it provides access to virtualized environment in seconds. In case of high load, it requires some processing time depending on size of an event. Realistic, another significant main goal of Haaukins is providing scenarios in exercises which are as close as possible to real life scenarios. It is achievable with help of virtualization methods used in Haaukins.

On top of the main goals, there are some additional goals included as education institutions used the platform in their lectures. Briefly, additional goals include:

- Easy management of the platform
- Clear and distinct explanation of exercises

The additional goals in mind, the platform management side is re-designed and built based on the feedback.

3.0.1 Main terms

The main components of Haaukins are; event, lab, challenge and management interface. Initially, management interface was only command line tool[1], which needs to be installed to

[1] A command-line interpreter or command-line processor uses a command-line interface (CLI) to receive commands from a user in the form of lines of text. (Taken from: https://en.wikipedia.org/wiki/Command-line$_{i}nterface$)

the computer who would like to manage the platform. However, it requires additional step and maintenance from the user. In order to make the management of the platform simple and easier web interface has been developed.[2].

Event

Event has two meaning in Haaukins platform, firstly it has its original meaning, special gathering for important times, secondly an event in technical meaning is to have multiple labs and teams. In Haaukins platform, events are disposable which means that they start, run and die as the other meaning of an event.

Lab

Lab is composed of virtualized Docker containers and virtual machines which includes vulnerable machines to hack and a virtual machine or virtual private network endpoint to access the machine.

Challenge

A challenge consists of one or more virtual images (container or virtual machines), they includes the vulnerable software to be hacked by participants. A challenge is shown in the platform as given in Figure 3.1

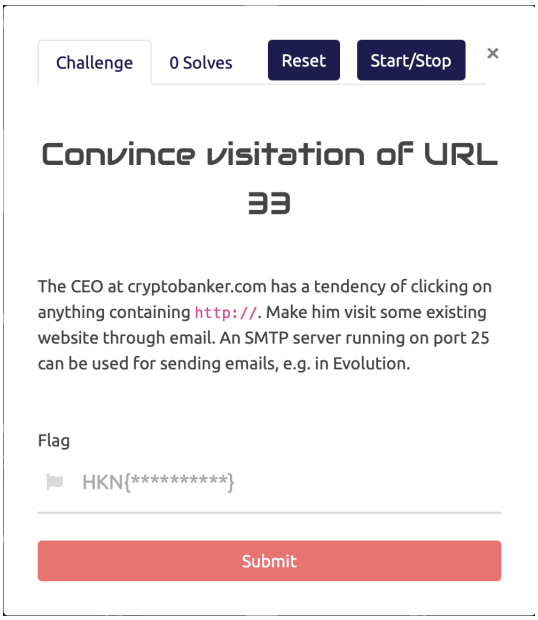

Figure 3.1: A challenge modal in Haaukins platform

[2]https://github.com/aau-network-security/haaukins-webclient

Flag

Flag is special secret string inserted into vulnerable application when building. Participants' goal is to reach hidden flags and submit them through web user interface shown given in Figure 3.1

3.0.2 Overview

In general, an event can contain one or more labs, a lab can contain at least one or more challenges. Overview relation between different components can be observed by Figure 3.2

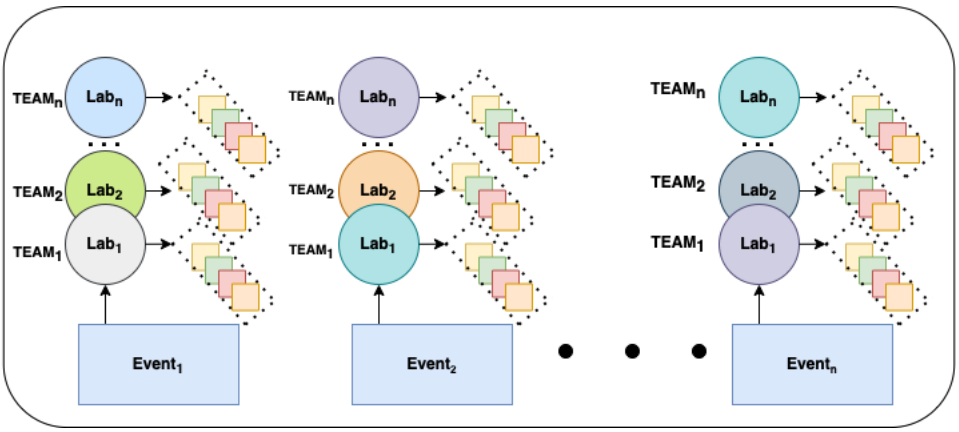

Figure 3.2: The relationship between different components in Haaukins platform

As shown in Figure 3.1, each lab consists of different challenges and a virtual machine for participants to use on browser[5]. The virtual machine is generally Kali or Parrot, since they both include wide range of cyber security tool. When a participant signs up on the event, they automatically assigned a lab which is completely isolated from other team's lab. All this automation does not require any extra step from the participant perspective. The automation of the environment is crucial to keep consistency and reliability at high level. A challenge can be built up from one or more containers or with virtual machines directly. Since challenges can be deployed as containers, they are lightweight compared to virtual machines. It brings great flexibility and scaling when there is high number of participants. Although containerization is great to use, sometimes it is not possible to emulate scenarios where the vulnerability is related to Windows Operating System core functionality for instance "Buffer Overrun In RPC Interface Could Allow Code Execution"[3]. In those situations, VMs can be used to emulate the challenge, there are some existing challenges on the platform which are using VMs.

[3]https://docs.microsoft.com/en-us/security-updates/securitybulletins/2003/ms03-026

Golang

Golang is a multi-threaded programming language developed by Google[4]. It provides built-in concurrency and fast processing capabilities. Haaukins platform is entirely written in Go programming language to benefit from built-in concurrency and huge community support. Golang is commonly used in system level automation tasks by large communities, some well-known projects written in Go are, Kubernetes, Docker, Prometheus, Terraform and many more. Due to its advantages compared to other languages and easy parallelism on programming tasks, the platform is build entirely with Go programming language.

Virtual Box

Virtual box is used as hypervisor to handle challenges on VM (Virtual Machine) and assigned VMs to the participants to the event. VMs are linked with Docker network interface to be able to attack vulnerable machines resides on same network.

Docker

Main core part of Haaukins platform is Docker containerization technology, the platform automates process of managing docker containers. Docker is providing virtual networking by creating macvlan network[5] interfaces when used with virtual machine on browser mode. In case of virtual private networking (VPN) mode, Docker bridge network[6] is used and network packets controlled with network tools such as iptables. In Figure 3.2, how a team is

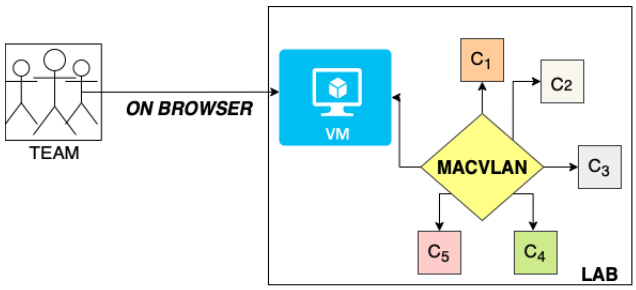

Figure 3.3: Lab internals, macvlan connection

connected isolated lab environment with macvlan networks is shown. All challenges and VM is connected to common macvlan network to reach vulnerable machines. Team/Participant is connected on browser to customized virtual machine which includes cyber security tools, for instance operating system Kali[7] or Parrot[8] The connection between user and virtual machine on browser is done by open source Apache Guacamole project.

[4]https://en.wikipedia.org/wiki/Go_(programming_language)
[5]https://docs.docker.com/network/macvlan/
[6]https://docs.docker.com/network/bridge/
[7]https://www.kali.org
[8]https://www.parrot.com/en

Apache Guacamole

[9] Apache Guacamole provides stable connection to a virtual machine through browser. It removes necessity of installing additional software to participants' computer when joining an event running on Haaukins platform. All the steps are automated with Go programming language and it enables users to have a custom set of challenges. Haaukins platform can manage as many events as possible, limited to server resources (CPU, Memory). Each event can include different sets of challenges to hack, at the end of hacking, participants will receive dynamically generated flag. It means although challenges are same the flags will be unique per participants, which is generally not supported by other platforms.

3.0.3 Virtual Private Network Access

Extensive usage of the platform extended its target group with professional security engineers. Even though professional security engineers are in seek of bug bounty programs, they interested to play with the platform. Professionals do not like to play with a virtual machine on a browser with no internet connection. Instead, their preference is mostly to use their customised environment for real tough challenges exists on the platform. Therefore, Wireguard[10] introduced to mitigate the problems of experienced users. It provides connectivity to challenges running on virtual network without using any virtual machine on browser.Compared to other existing platforms, virtual private network solution is different in Haaukins platform thanks to resilient Wireguard implementation. The platform is communicating with Wireguard service, which runs independently from the platform, to create requested clients from the platform. Additionally, VPN solution provides simultaneously access for a lab assigned to a team. A team up to four people can access and solve challenges at the same time. The default limit for a lab to be access at the same time is four.However, it is a configuration matter and can be adjusted according to need. Wireguard is special compared to OpenVPN, enormous improvements are introduced on ping time and throughput[6].Therefore, Haaukins platform outperforms compared to other commercial e-Learning platforms.

3.0.4 Challenges

Challenges in the platform is developed specifically for a course (e.g Network Security, Cryptography, Security Engineering) or company based training course to provide specialized learning experience. A challenge can be a docker container, a virtual machine or group of docker containers. Since challenges are created parallel to course syllabus, understanding conceptional topics becomes easy for participants. Therefore, the platform has been used at Aalborg University[11], Hogeschool Saxion[12] and Denmark Technical University (DTU)[13] in relative courses.

[9] https://guacamole.apache.org
[10] https://www.wireguard.com
[11] https://www.en.aau.dk/education/master/cyber-security/
[12] https://www.saxion.edu/programmes/bachelor/applied-computer-science
[13] https://www.compute.dtu.dk/english/continuing-education/master-of-cyber-security/study-program

Deployment of Challenges

Challenges, running in container(s), are deployed automatically to the platform in a micro service architecture approach.

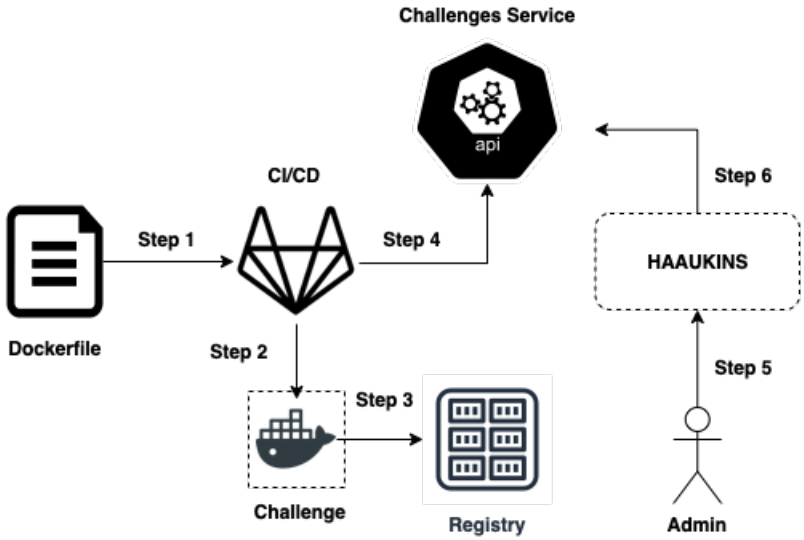

Figure 3.4: Challenge deployment architecture diagram

Figure 3.4 represents how a challenge is automatically deployed to the platform. The figure is drawn from a perspective of a person who deploys a challenge. Therefore, step numbers are created from this perspective. The process does not include any human interaction after Step 1.

In Step 1, challenge creator should define all components of a challenge to be running, which includes codes, framework, programming language and if required more information. Once, they have been defined in a simple Dockerfile[14] and pushed a repository (on Gitlab/Github), all other steps are automatically completed. It brings great opportunity for other institutions to contribute the platform. Any institution can create challenges based on their requirement and shipped to Haaukins platform in one manual step. Indeed, the collaboration, between the universities who use Haaukins platform, is adequately increased. They decided to create a cluster of point where Haaukins can be scaled and improved even more with innovative challenges.

In Step 2, Gitlab CI*[15] builds and creates a Docker image[16], it provides a packed application which can run anywhere as long as Docker daemon exists.

In Step 3, the created image is saved to a some kind of database with all content of the application. In case existence, it automatically updates existing image with new updates. In Step 4, a configuration file is pushed to a microservice which contains information about an

[14]https://docs.docker.com/engine/reference/builder/
[15]Continuous integration
[16]https://docs.docker.com/engine/reference/commandline/image/

exercise, for instance, image name, registry information, flag information and some metadata.

In Step 5, in case a challenge is requested to be included in an event, Haaukins can pull and run it as shown in the Figure 3.4

In Step 6, an event organizer or instructor creates an event on administration page.

3.0.5 Limitations

Haaukins platform is a project which is under development. Although it provides decent ways to interact with students, the platform lacks in some features. The known missing points are:

- Lack of multi network virtualization
- Dependency on Virtualbox
- Cloud incompatibility

Lack of multi network virtualization

Since the platform uses Docker networking mechanism, it is not possible to have a challenge which contains multiple network layers to be hacked. This issue discussed within the community and seems OpenVSwitch[17] is required to include more advanced network related challenges.

Dependency on Virtualbox

When the platform is deployed on a server, it runs smoothly, however in case of high load, Virtualbox is consuming high CPU on the server. It causes delays, freezing on screen, not able to complete the tasks on the VMs which are running on high loaded host machine.

Cloud incompatibility

Another pitfall is that, when the platform is deployed to cloud, it might not be feasible to run with the architecture described above due to requirement of high memory and CPU capabilities. Even though it is deployed a public cloud, it will not be cost effective solution.

The limitations may not be limited to given points above, however they are the main limitations known by the community. It is a huge advantage even though there are some limitations. There are bug reports happening time to time however they mostly not related to core functionalities of the platform. Nevertheless, the platform is used by the institutions and conferences in Denmark. Although it has some limitations, they are not related to main aim of the platform mentioned in the beginning of the paper.

[17] https://www.openvswitch.org

4 Management

In this section, management of the platform will be discussed. In earlier versions of the platform, it can only be managed with people who has understanding of command line interface[1]. It was blocking feature, limiting usage and management of the platform. As the platform used in lectures by instructors and contents of the challenges are customized based on curricula of the courses. The management of the platform must be informative, user friendly and self explained. Hence, web interface has been developed, it brought great convenience and safety. The command line tool was providing the actions given in the Figure 4.1

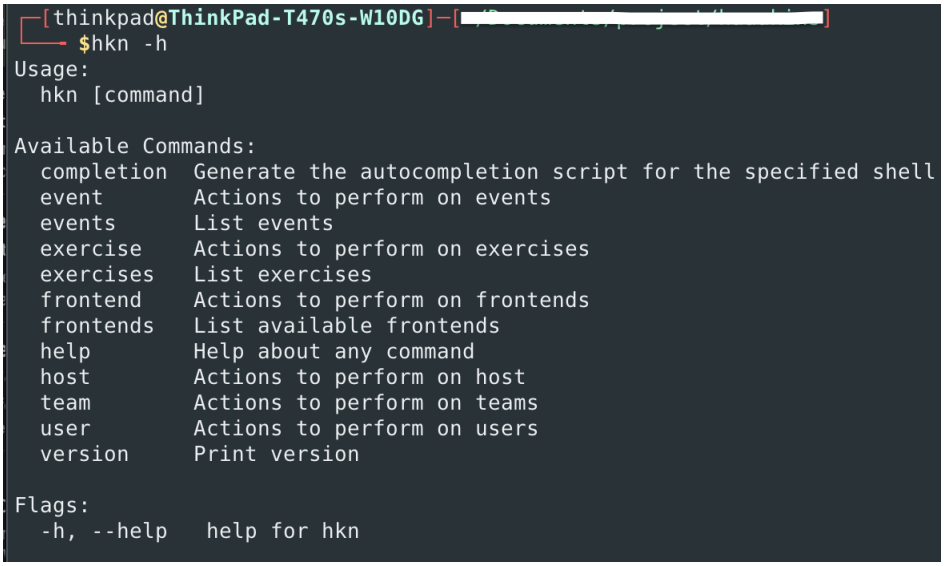

Figure 4.1: Command Line Tool Options of Haaukins

The CLI can be fun to play when you know otherwise it might be headache. Simple event creation command with CLI is shown in Figure 4.2

Figure 4.2: Creation of an event through CLI

[1]https://en.wikipedia.org/wiki/Command-line_interface

The problem should be observed from the Figure 4.2, the command does not explain itself, there are some challenges passed with "-e" flag. All the flags passed as argument to CLI tool needs to known by creator of an event. The command is creating an event with some group of challenges (scan, xss, mitm, csrf), with availability of five, capacity of 15, and name of the event is "e-Learning Techniques". However, all these information needs to known at first glance when start to manage the platform. Even more, all the challenge information should be given when creating an event. Otherwise, event creator might create an event blindly, randomly without a knowledge about what they want. In order to prevent such terrible situations, web based solution is introduced to the platform with all detailed information of challenges and fields.

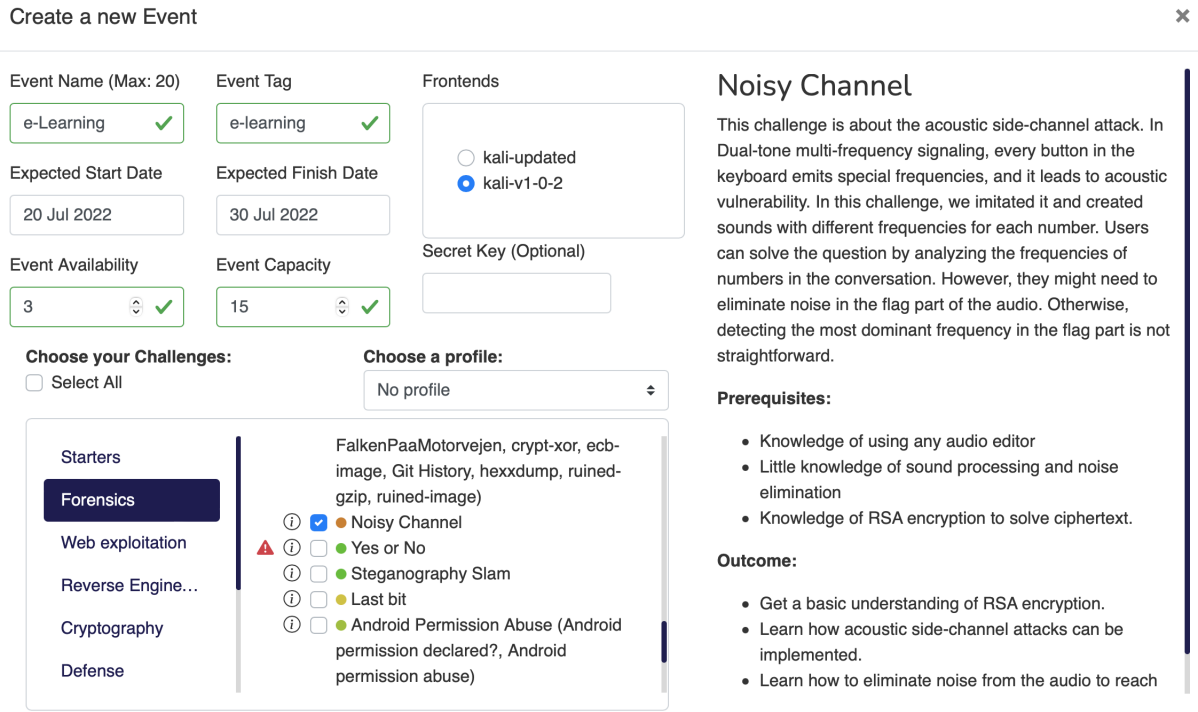

Figure 4.3: Event creation using web interface [2]

Figure 4.3 represents creation step of an event with detailed fields. Administrators and instructors can create events specifically on particular topic. In left corner side of the figure, categories of the challenges are listed, each category has at least thirty challenges and increases as time pass. When selecting set of challenges, all information about a challenge can be checked in right side of the figure. The fields (Event Name, Event Tag and all others) reveals extra information when hover on them. There are tons of new, user friendly and supportive features for leveraging the platform's capabilities. It substantially aids instructors to have valuable challenges in the event, parallel to their circular.

Administration of the platforms are people who are running the platform at Aalborg University. However, an instructor signs up for their usage, both administrators and instructors

can have similar management permissions except few admin permissions. It is possible to manage events on behalf of event creators or event creators can manage their events through web interface.

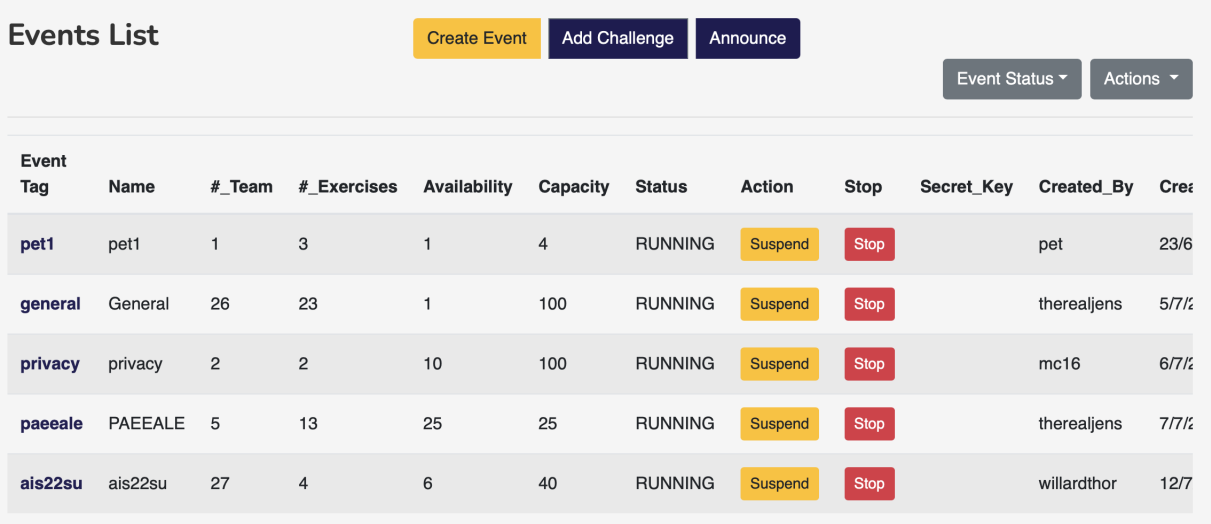

Figure 4.4: Listing events on web client of Haaukins platform[3]

Figure 4.4 represents possible actions that an authorized person can take either on their event only, or all events depending of their role.Event tag represents subdomain, for instance event called "general" runs at general.haaukins.com, "privacy" runs at privacy.haaukins.com

The access for management site of the platform is done only by invitation, otherwise it is not possible. In order to prevent spamming and good state of the platform, invitation only sign up exists. The invitation codes are generated from administrators upon checks on people who requests to have an access. Legitimate users are instructors who are providing mostly security related lectures, in some cases people at companies, who are providing security training to their employees.

An event can be suspended or stopped manually for some time through web interface. Similarly all students/teams can be managed, if a challenge is forgotten in case of event creation. It can be added afterwards and the platform will propagate resource allocation and assignment of the challenges to the event participants. All these steps and more are done self-explanatory web interface.

There are many other features and capabilities of the platform. The web client delivers all those features in a user-friendly, informative and easy to understand structure. All developments are done in a feedback loop from users of the platform. As usage of it increased in educational institutions, the platform became better and better.

Finally, there are still ongoing feature requests and improvements at Github[4].Since all Haaukins and its services are open source, it is possible to deploy and modify according to your requirements.There is high lack of having this flexibility on security education platforms.

[4]https://github.com/aau-network-security/haaukins-webclient

5 Conclusion

Haaukins platform initiated with support from non-profit organizations mainly Industriens Fond[1] in Denmark and developed by Aalborg University. Although existing platforms might give some flexibility at some extend, they do not provide full freedom in terms of usage. The existing platforms are also decent, has enormous amount of resources however they have different goals compared to non-profit platform, Haaukins.

The main aim is to help students who are struggling with understanding basics of lectures primarily by targeting high-school students. The platform showed itself with its success at high schools changed its target group to include university students. As the platform is evolved in time, the requirements of the platform changed according to user feedback comes from students and instructions. Participants of an event on the platform wanted to have more manageable environment by providing them an access to manage their environment. The management from students side is given and limited to their own virtualized environment. Even though they do not have any influence on any running other virtual environments, they can restart their virtual environment when there is a problem. The administration side of the platform is re-designed and built from scratch with user-friendly, easy to manage and informative features in mind, explained in the management chapter. All the improvements have been done on the platform to deliver pleasure time while learning both for students. The growing interest on security education is pushing the platform evolve more in the future, including more advanced features. The platform[2] is in active development at Github and generally available on event "General"[3]. Since it is open to anyone who is interested to do practical experience, anyone can sign up at no cost. Further requests and improvements are very welcome by the community, it is possible to initiate a discussion from main Github repository under discussions.[4] All in all, Haaukins platform provides fast, easy access and custom set of challenges to be solved for anyone who interested to up skill their security knowledge.

[1] https://industriensfond.dk
[2] https://github.com/aau-network-security/haaukins
[3] https://general.haaukins.com
[4] https://github.com/aau-network-security/haaukins/discussions

List of Figures

Bibliography

[1] CSIS. *Significant Cyber Incidents*. 2022. URL: https://www.csis.org/programs/strategic-technologies-program/significant-cyber-incidents (visited on 07/12/2022).

[2] T. Chothia and C. Novakovic. "An Offline Capture The Flag-Style Virtual Machine and an Assessment of Its Value for Cybersecurity Education". In: *2015 USENIX Summit on Gaming, Games, and Gamification in Security Education (3GSE 15)*. Washington, D.C.: USENIX Association, Aug. 2015. URL: https://www.usenix.org/conference/3gse15/summit-program/presentation/chothia.

[3] P. Chapman, J. Burket, and D. Brumley. "PicoCTF: A Game-Based Computer Security Competition for High School Students". In: *2014 USENIX Summit on Gaming, Games, and Gamification in Security Education (3GSE 14)*. San Diego, CA: USENIX Association, Aug. 2014. URL: https://www.usenix.org/conference/3gse14/summit-program/presentation/chapman.

[4] T. K. Panum, K. Hageman, J. M. Pedersen, and R. R. Hansen. "Haaukins: A Highly Accessible and Automated Virtualization Platform for Security Education". In: *2019 IEEE 19th International Conference on Advanced Learning Technologies (ICALT)*. Vol. 2161-377X. 2019, pp. 236–238. DOI: 10.1109/ICALT.2019.00073.

[5] G. M. Mennecozzi, K. Hageman, T. K. Panum, A. Türkmen, R.-V. Mahmoud, and J. M. Pedersen. "Bridging the Gap: Adapting a Security Education Platform to a New Audience". In: *2021 IEEE Global Engineering Education Conference (EDUCON)*. 2021, pp. 153–159. DOI: 10.1109/EDUCON46332.2021.9453985.

[6] J. Donenfeld. "WireGuard: Next Generation Kernel Network Tunnel". In: Jan. 2017. DOI: 10.14722/ndss.2017.23160.