# Practical teaching of distributed systems: A scalable environment for on-demand remote experimentation

Nina Slamnik-Kriještorac, Henrique C. Carvalho de Resende, and Johann M. Marquez-Barja

University of Antwerp - imec, IDLab - Faculty of Applied Engineering,

Sint-Pietersvliet 7, 2000 Antwerp, Belgium.

 $E-mail:\ nina.slamnikkrijestorac@uantwerpen.be, Henrique.CarvalhoDeResende@uantwerpen.be, johann.marquez-barja@uantwerpen.be and and an anti-structure and a standard structure and a structur$ 

### ABSTRACT

A practical compound of education in computer science and electrical engineering, driven by increased availability and maturity of many emerging technologies, should be enriched by various laboratory resources in order to synchronize the paces between technology advancements and education. In particular, advancements in containerization as a virtualization technique pave the way towards allowing students to deploy their project applications with a lightweight resource footprint on top of the cloud. Being backed by a valuable feedback from 45 Bachelor students, in this paper we present the best practices on how virtualization can be leveraged to create a scalable environment for on-demand remote experimentation with distributed systems.

## **1 INTRODUCTION**

The work in a laboratory is an inevitable component of students' education, since it empowers and enables students to acquire practical knowledge as a compound of the overall expertise. Crocker et al. [1], and Hofstein and Lunetta [2] point at the importance of learning by inquiry, as well as the increased involvement of students throughout the learning process.

Nowadays, due to the increasing availability and popularity of many emerging technologies, various fields in computer science and electrical engineering education are enriched by plethora of laboratory resources that should be exploited in order to sync the paces between technological advancements and education. In the context of resources, we consider e.g., the low-cost devices (e.g., Raspberry Pis (RPis), Arduino boards, etc.), and cloud resources, that are financially more accessible than high-processing machines, having the same educational outcome in the areas of practical computer and electrical engineering education. As a result of significant advancements in virtualization techniques, containerization (e.g., Docker containers, Linux containers) opened the path for computer science/engineering students to develop applications with a lightweight resource footprint. Such lightweight design enables deploying applications within resource constrained platforms, sharing only the operating system level calls with the host and other virtualized environments. Thus, all of these resources should be



Figure 1: The design pillars of the remote and on-demand laboratory.

better exploited, making laboratory environment *scalable*, *available on-demand*, and *remotely accessible* (Fig. 1).

Concerning the aforementioned opportunities, Fig. 1 depicts the goal of this paper, which is: i) to present the *best practices* on how the virtualization techniques can be leveraged towards creating a scalable environment for remote experimentation, and development of distributed systems, and ii) to assess the social impact of such experimentation environment on the students as a target social group, thereby inspecting the applicability of such approach to experimentation opportunities in education for many low-income societies. Thus, the presentation of the best practices is backed by a valuable feedback that is collected from 45 Bachelor students within the course Distributed systems, three times during the semester.

Due to various unpredictable constraints that might result in a limited or no access to physical laboratory, a quick action and suitable response are inevitable. As such circumstances significantly affect both students and educators in terms of achieving the educational ultimate goals, i.e., effective teaching, and acquiring knowledge in the area of interest, there is a need for scalable laboratory that can be provided on-demand, and reached remotely. By exploiting the cloud resources, and virtualization techniques that allow students to generate a desired environment for running their programs, courses that fall into the engineering and computer science fields can be efficiently re-designed and customized to various circumstances. Therein, instead of providing each student with a high-performance computer, or a low-cost variant of devices with lower capabilities (e.g., RPis), but with a limited access to laboratory, cloud resources can be provided and released on-demand, and easily accessed by any internet connected computer.

Furthermore, there are usually numerous regulations and limitations imposed by faculties and universities, that do not allow students to bring the equipment out of the laboratories. On the other hand, the students' pace is different, therefore some of them cannot cope with the workload within the scheduled lab hours. Henceforth, the need to access laboratory resources in the *out-of-lab* hours should be recognized as of high importance. As we have ascertained in our previous educational-related research work [7], students raised concerns regarding limitation in accessibility of laboratory equipment, i.e., RPi devices that they used to design and develop a distributed file system. To tackle the challenge regarding limited access, the remote access nature of cloud computing was integrated to the Distributed Systems course where we designed a remote experimentation system for the students, allowing them to access containerized environment on the cloud, and work on their project out of laboratory session hours as well.

The first two months of the semester students were performing experimentation at the University premises. However, COVID-19 required the immediate reaction to move our teaching and experimentation remotely, and to embrace digital solutions for remote education. Due to such strong need to pursue an adequate substitutional environment for in-class teaching and physical laboratory, we present the best practices on remote teaching and experimentation within the academic course for the teaching of Distributed systems. First, we present how our physical laboratory, previously consisted of 50 RPi devices, can be efficiently replaced by virtual resources provided in the cloud, by mapping RPi functionalities to the pieces of cloud machines, i.e., containers. Second, we analyze the valuable feedback gathered from students, inspecting their experience with the new teaching and experimentation environment, and assessing their preference between in-lab physical equipment and remotely accessible cloud machines.

#### 2 RELATED WORK

During the last 15 years, the research community recognized a wide range of practices for modernizing laboratories, and facilitating the process of acquiring practical knowledge. As the digital world where we live in has also created opportunities to efficiently perform remote teaching, a reasonable need for unlimited, and ondemand access and work in the laboratories has arisen. Additionally, the same remote laboratory set-up can be reused, i.e., shared for a set of different experiments, increasing the level of use but also the revenue of the institution that developed the remote laboratories [9]. Furthermore, in a particular case of the Spanish University for Distance Education (UNED, Universidad Nacional de Educación a Distancia), which has more than 200,000 students enrolled, the remote experimentation is a well-recognized practice that deconstructs laboratories as sets of services that can be offered to the students, providing Laboratory as a Service (LaaS) [9]. The concept of LaaS supports students who lack the time or resources to attend classes within University campus [9].

As stated by Tobarra et al. [9], the idea to perform laboratory exercises remotely was born two decades ago, and so far it is actively exploited in multiple scientific fields, such as chemistry, physics, electronics, robotics, and nuclear reactor. Based on the remote laboriented research conducted by Matarrita and Concari [6], there is a distribution of around 130 articles based on the field to which remote experimentation is applied. The results show that most of the work tackle remote experimentation in physics, being followed by various engineering courses. According to Ionescu et al. [3], some of the important reasons for designing, and building remote laboratories in the aforementioned fields are: i) increased need for additional laboratory resources due to ever increasing number of students, ii) the high enrollment, iii) insufficient financial resources to cover increased needs, and last but not least iv) the amount of time that educators need to spend to prepare and perform laboratory exercises, which might exceed the time allocated to teaching duties.

As presented by Jourjon et al. [4], FORGE toolkit is an eLearning ecosystem that consists of teaching and educational materials, tools, and experiments, allowing educators and students from more than one hundred US and EU universities to access these resources under an open scheme and policies. The robustness of such platform is based on the deployment during a post-graduate course, where more than 6000 experiments were provisioned [4].

Vivar and Magna [10] present an interesting approach on how to bypass the barrier of a limited number of networking devices, by creating a remote network lab as a support to teaching computer networks. By connecting to a web server, students are able to interact remotely with commercial network devices (e.g., switches, routers, and firewall). Although students expressed a positive attitude towards such approach, the remote lab system presented by Vivar and Magna [10] is quite expensive, and allows access to only 16 students at the same time.

One of the examples on remote experimentation in electronics for undergraduate students is presented by Sousa et al. [8]. In their thorough illustration of the whole remote lab platform, so-called RemotElectLab, Sousa et al. [8] promote using a generic hardware platform with a generic access interface for the implementation of electronic circuits that replicate those used in the physical laboratory. The purpose of their RemotElecLab is not to completely replace the traditional hands-on approach, but rather to serve as a complementary laboratory set that can help students to efficiently use resources by relaxing the timetable and diminishing the lab occupation. Besides being a good example for remote teaching in electronics, the study follows the same approach as others by providing remote access to specialized hardware which can be expensive and non-scalable. Although the feedback from students was collected in the form of a short questionnaire, this work does not present any tangible results that reflect students' experience with new practices in experimentation, nor the prerequisites needed for this approach.

Despite the advantages of such approach in teaching and experimentation, the remote laboratory practice depends also on the user interface and its abstraction. To this end, Marquez-Barja et al. [5] analyze the impact of different types of user interfaces in remote telecommunications laboratories. The feedback they collected from students show the preference towards web-based rich interactive interface approaches. However, the high level of abstraction might cause lowering of the capabilities and operation of remote testbed facilities [5].

Unlike the previous practices that are built for a specific field in practical education, our scalable laboratory for experimentation on-demand is rather a generalized concept that can be beneficial to various engineering and computer science areas. Although designed for development, and work, on the distributed systems, our practices show how the transition from a physical laboratory consisted of personal computers and RPi devices can be transformed to a remote laboratory on the cloud, benefiting from the virtualization technique such as containerization. Finally, the valuable feedback gathered from 45 students justifies the feasibility and reasonability of creating such laboratory.

## **3 BEST PRACTICES**

#### 3.1 Course

As our best practices for the concept of remote experimentation are based on the Distributed systems module, this section provides general information about the course.

Distributed Systems is the course which belongs to the final year of Bachelor of Electronics and ICT Engineering Technology program, at Faculty of Applied Engineering, University of Antwerp, Belgium. Regarding general competences that students need to acquire to enroll for the course, a knowledge of use of computers and Internet is a must, while practical work within project implementation requires a minimum of basic programming skills in order to be able to create adequate software solutions. Throughout the semester, all laboratory sessions follow the lectures and the topic presented by lecturer previously, thereby preparing students with a satisfactory amount of information needed to properly understand tasks within practical exercises. The content of our course is carefully designed by following IEEE/ACM CS2016 joint curriculum for computer science engineering, as we presented in [7].

Furthermore, in order to portray both the best practices and the student feedback in a way comprehensible for the reader, here we provide a solid knowledge about the project background and the description of particular project stages. To build a remote laboratory for experimentation on-demand, we used the cloud resources that belong to Internet and Data Laboratory (IDLab), a research group within University of Antwerp. The containers organized in a Docker Portainer are used to realize the general idea of the student project, which is to build a comprehensive distributed file system (so-called system Y) in a ring topology (Fig. ??), with files being automatically replicated to nodes and naming server created to keep track of nodes' connections within the ring. As stated in Introduction, during the first two months of the semester students were working on the project within the in-lab environment at the University premises. This phase included the laboratory setup that consisted of 50 RPis, providing homogeneous set of five to each group of students, during all 18 laboratory exercises. Due to the lockdown caused by COVID-19, the immediate reaction to embrace digital solutions for remote education was necessary to enable students working on the project in an uninterrupted way.

#### 3.2 Scalable environment for experimentation

Now if we take a closer look to the Fig. 2, we can see what are the essential entities that are required in order to follow our practice on on-demand remote experimentation.

The left side of the Fig. 2 presents a physical laboratory environment for Distributed systems at the University premises. It consists of 50 RPi devices that are assigned to the student groups. All groups have five students, and each of them works on a single RPi device. Furthermore, all devices in one group are connected on the same



Figure 2: The transition from a physical laboratory at the University, to scalable and remote environment for experimentation on-demand.

network, so they can work as distributed nodes within distributed file system organized in a ring topology (as illustrated in Fig. ??), and communicate to each other in order to share files. Students are instructed to install Raspbian operating system on their RPi machines, to tweak the network settings, to maintain connectivity between different machines, and to prepare a java environment for running corresponding java applications. Apart from devices themselves, additional equipment such as keyboards, mice, monitors, and network switches is needed to configure RPis, and to make them work properly. After initial configuration, students are encouraged to use their own laptops to securely connect to RPis via Secure Shell (SSH). Due to the Faculty and University regulations, students are allowed to work on such physical laboratory setup only during the scheduled lab sessions. Tackling scalability, if any of devices does not work properly, or more students enroll the course, the set of redundant devices is limited, and it takes additional time and cost resources to order additional equipment.

Furthermore, on the right side of the Fig. 2 we illustrate the remote laboratory designed and created on top of the cloud resources. For each physical machine (i.e., RPi device) we instantiated a lightweight Docker container in a Docker swarm, using Raspbian Docker image to make a remote machine compatible with the operating system on the physical ones. Instead of five RPis, each student group gets a cluster of five Docker containers that share IP address, but receive the upcoming requests on different dedicated ports. The network and resource management of student clusters is performed by Portainer, allowing educators to monitor work of these remote machines. In case any container fails, another lightweight instance can be easily created, and set up, as same as in case more machines are needed. This approach saves a lot of resources, since for a regular Raspbian operating system image for RPi devices around 4.3GB of storage is needed, therefore a memory

card with at least 8GB storage capacity is required. On the other hand, containers use resources more efficiently, as a single Raspbian Docker container requires around 180MB of storage, and 4.47MB of RAM. With such light requirements, a large set of Docker containers can be instantiated on top of bare metal, cloud, or inside a Virtual Machine (VM).

The students can access the cloud machines, i.e., Docker containers, from any place where network connectivity is granted. However, they need to use University's Virtual Private Network (VPN) connection to securely access the University cloud. In this case, students are encouraged to work on their laboratory tasks and project in a flexible manner, and out of scheduled hours as well.

Our 45 students that enrolled the course Distributed systems had the opportunity to work in both types of laboratory, i.e., physical and remote. During the first project phase students were working on the RPi devices, developing the Naming server functionality, as described in Section 3.1. Due to the unforeseen circumstances that occurred in this semester, that significantly limited the access to physical laboratory and classrooms, the teaching and experimentation were both shifted to the remote mode. Therefore, from the Discovery part of the project until the end, students were using cloud resources to perform experimentation.

Such circumstances enabled students to gain different skills, and to acquire different experience throughout the semester. Thus, it motivated us to design a survey in the way presented in the following section, and to ask students to compare different environments and their experience with both.

#### 4 RESULTS AND DISCUSSION

#### 4.1 Survey

In this section we present the feedback that we collected from 45 students throughout the semester, during 18 laboratory sessions. First, we present the survey that we have carefully designed to tackle the essential parts of experimentation and students' experience. Second, we discuss the obtained results, and show the lessons learnt from our perspective that can be useful for different engineering courses, and anytime such laboratory needs to be created on demand.

In Table 1, we list the three following sets of questions: i) Set 1 consists of nine questions that refer to the differences between students' experience with hands-on (i.e., RPis and Personal Computers (PCs)), and with remote cloud machines, ii) Set 2 inspects the amount of time that students spend experimenting on the remote machines out of the scheduled laboratory sessions, while iii) Set 3 spans five questions that aim at investigating the students' awareness of mapping the features of a distributed system to a specific environment such as Docker container in the cloud, as well as students' satisfaction with such remote experimenting approach.

Furthermore, Table 2 presents the results that are calculated according to equation 1, which was applied to the collection of students' answers. The equation 1 calculates the average number of students that voted for a particular category (from Strongly Affirmative, to Strongly Negative) expressed as a percentage. In equation 1, M is the number of survey iterations that is our case three (M = 3),  $N_j$  is the number of responses in *j*-th iteration ( $0 \le N_j \le 45$ ), and  $a_{ij}$  is the student's answer for *i*-th category, in

#### Table 1: Survey questions

Set	Please evaluate the following statements using the scale from Strongly Affirmative to Strongly Negative.					
01	Do you feel this new experimentation environment					
χı	is a good replacement for the in-class laboratory?					
Q2	Do you feel more comfortable asking questions in in-class labs than in the remote ones?					
Q3	Do you find working on the remote cloud resources					
	as same as working on the physical Raspberry Pi					
	machines?					
Q4	Did your previous experience in working with					
	Linux-based systems help you to grasp the					
	practical work within Distributed systems?					
Q5	Do you find that is an advantage of remote access					
Q6 Q7	To have machines available 24/7?					
	nhysical machines (your lanton Raspherry Pi)?					
	Did you need to understand the container-based					
	virtualization to successfully finalize your exercises					
	on the remote					
	machines?					
Q8	Do you think that programming of the functionalities					
	was crucial to realize laboratory exercises in a					
	successful way?					
Q9 Set 2	Do you experience difficulties in working with no					
	graphical user interface?					
	vou spend working on the remote machines out of lab					
	you spend working on the remote machines out of tub					
	110415.					
	less than 2h per week					
01	less than 2h per week between 2h and 4h per week					
Q1	less than 2h per week between 2h and 4h per week between 4h and 8h per week					
Q1	less than 2h per week between 2h and 4h per week between 4h and 8h per week more than 8h per week					
Q1 Set	less than 2h per week between 2h and 4h per week between 4h and 8h per week more than 8h per week Please evaluate the following statements using the scale					
Q1 Set 3	less than 2h per week between 2h and 4h per week between 4h and 8h per week more than 8h per week <i>Please evaluate the following statements using the scale</i> <i>from Strongly Affirmative to Strongly Negative.</i>					
Q1 Set 3	less than 2h per week between 2h and 4h per week between 2h and 4h per week more than 8h per week <i>Please evaluate the following statements using the scale</i> <i>from Strongly Affirmative to Strongly Negative.</i> Did your previous knowledge on networking help you to understand the principles of connecting remote machines					
Q1 Set 3 Q1	less than 2h per week between 2h and 4h per week between 2h and 4h per week more than 8h per week <i>Please evaluate the following statements using the scale</i> <i>from Strongly Affirmative to Strongly Negative.</i> Did your previous knowledge on networking help you to understand the principles of connecting remote machines, and to set-up communication between them (unicast.					
Q1 Set 3 Q1	less than 2h per week between 2h and 4h per week between 2h and 4h per week more than 8h per week <i>Please evaluate the following statements using the scale</i> <i>from Strongly Affirmative to Strongly Negative.</i> Did your previous knowledge on networking help you to understand the principles of connecting remote machines, and to set-up communication between them (unicast, multicast, and broadcast)?					
Q1 Set 3 Q1	less than 2h per week between 2h and 4h per week between 2h and 4h per week more than 8h per week <i>Please evaluate the following statements using the scale</i> <i>from Strongly Affirmative to Strongly Negative.</i> Did your previous knowledge on networking help you to understand the principles of connecting remote machines, and to set-up communication between them (unicast, multicast, and broadcast)? Do you find experience with accessing, and working					
Q1 Set 3 Q1	less than 2h per week between 2h and 4h per week between 2h and 4h per week more than 8h per week <i>Please evaluate the following statements using the scale</i> <i>from Strongly Affirmative to Strongly Negative.</i> Did your previous knowledge on networking help you to understand the principles of connecting remote machines, and to set-up communication between them (unicast, multicast, and broadcast)? Do you find experience with accessing, and working on the remote machines within a cloud environment,					
Q1 Set 3 Q1 Q2	less than 2h per week between 2h and 4h per week between 2h and 4h per week more than 8h per week <i>Please evaluate the following statements using the scale</i> <i>from Strongly Affirmative to Strongly Negative.</i> Did your previous knowledge on networking help you to understand the principles of connecting remote machines, and to set-up communication between them (unicast, multicast, and broadcast)? Do you find experience with accessing, and working on the remote machines within a cloud environment, as useful for your general understanding of distributed					
Q1 Set 3 Q1 Q2	less than 2h per week between 2h and 4h per week between 2h and 4h per week more than 8h per week <i>Please evaluate the following statements using the scale</i> <i>from Strongly Affirmative to Strongly Negative.</i> Did your previous knowledge on networking help you to understand the principles of connecting remote machines, and to set-up communication between them (unicast, multicast, and broadcast)? Do you find experience with accessing, and working on the remote machines within a cloud environment, as useful for your general understanding of distributed systems?					
Q1 Set 3 Q1 Q2	less than 2h per week between 2h and 4h per week between 2h and 4h per week more than 8h per week <i>Please evaluate the following statements using the scale</i> <i>from Strongly Affirmative to Strongly Negative.</i> Did your previous knowledge on networking help you to understand the principles of connecting remote machines, and to set-up communication between them (unicast, multicast, and broadcast)? Do you find experience with accessing, and working on the remote machines within a cloud environment, as useful for your general understanding of distributed systems? Does your experience with accessing, and working on top of					
Q1 Set 3 Q1 Q2 Q2	less than 2h per week between 2h and 4h per week between 2h and 4h per week more than 8h per week <i>Please evaluate the following statements using the scale</i> <i>from Strongly Affirmative to Strongly Negative.</i> Did your previous knowledge on networking help you to understand the principles of connecting remote machines, and to set-up communication between them (unicast, multicast, and broadcast)? Do you find experience with accessing, and working on the remote machines within a cloud environment, as useful for your general understanding of distributed systems? Does your experience with accessing, and working on top of the cloud resources make you more comfortable to					
Q1 Set 3 Q1 Q2 Q3	less than 2h per week between 2h and 4h per week between 2h and 8h per week more than 8h per week <i>Please evaluate the following statements using the scale</i> <i>from Strongly Affirmative to Strongly Negative.</i> Did your previous knowledge on networking help you to understand the principles of connecting remote machines, and to set-up communication between them (unicast, multicast, and broadcast)? Do you find experience with accessing, and working on the remote machines within a cloud environment, as useful for your general understanding of distributed systems? Does your experience with accessing, and working on top of the cloud resources make you more comfortable to experiment with widely adopted cloud platforms such as					
Q1 Set 3 Q1 Q2 Q3	less than 2h per week between 2h and 4h per week between 2h and 4h per week more than 8h per week <i>Please evaluate the following statements using the scale</i> <i>from Strongly Affirmative to Strongly Negative.</i> Did your previous knowledge on networking help you to understand the principles of connecting remote machines, and to set-up communication between them (unicast, multicast, and broadcast)? Do you find experience with accessing, and working on the remote machines within a cloud environment, as useful for your general understanding of distributed systems? Does your experience with accessing, and working on top of the cloud resources make you more comfortable to experiment with widely adopted cloud platforms such as Amazon Web Services (AWS)?					
Q1 Set 3 Q1 Q2 Q3	less than 2h per week between 2h and 4h per week between 2h and 4h per week more than 8h per week <i>Please evaluate the following statements using the scale</i> <i>from Strongly Affirmative to Strongly Negative.</i> Did your previous knowledge on networking help you to understand the principles of connecting remote machines, and to set-up communication between them (unicast, multicast, and broadcast)? Do you find experience with accessing, and working on the remote machines within a cloud environment, as useful for your general understanding of distributed systems? Does your experience with accessing, and working on top of the cloud resources make you more comfortable to experiment with widely adopted cloud platforms such as Amazon Web Services (AWS)? Do you think that enabling connectivity between					
Q1 Set 3 Q1 Q2 Q3	Iterationless than 2h per weekbetween 2h and 4h per weekbetween 4h and 8h per weekmore than 8h per weekPlease evaluate the following statements using the scalefrom Strongly Affirmative to Strongly Negative.Did your previous knowledge on networking help you tounderstand the principles of connecting remote machines,and to set-up communication between them (unicast,multicast, and broadcast)?Do you find experience with accessing, and workingon the remote machines within a cloud environment,as useful for your general understanding of distributedsystems?Does your experience with accessing, and workingon top ofthe cloud resources make you more comfortable toexperiment with widely adopted cloud platforms such asAmazon Web Services (AWS)?Do you think that enabling connectivity betweencloud machines is less complex than with physical					
Q1 Set 3 Q1 Q2 Q3 Q4	less than 2h per week between 2h and 4h per week between 2h and 4h per week more than 8h per week <i>Please evaluate the following statements using the scale</i> <i>from Strongly Affirmative to Strongly Negative.</i> Did your previous knowledge on networking help you to understand the principles of connecting remote machines, and to set-up communication between them (unicast, multicast, and broadcast)? Do you find experience with accessing, and working on the remote machines within a cloud environment, as useful for your general understanding of distributed systems? Does your experience with accessing, and working on top of the cloud resources make you more comfortable to experiment with widely adopted cloud platforms such as Amazon Web Services (AWS)? Do you think that enabling connectivity between cloud machines is less complex than with physical machines such as Raspberry Pis (no need for					
Q1 Set 3 Q1 Q2 Q3 Q4	less than 2h per week between 2h and 4h per week between 2h and 4h per week more than 8h per week <i>Please evaluate the following statements using the scale</i> <i>from Strongly Affirmative to Strongly Negative.</i> Did your previous knowledge on networking help you to understand the principles of connecting remote machines, and to set-up communication between them (unicast, multicast, and broadcast)? Do you find experience with accessing, and working on the remote machines within a cloud environment, as useful for your general understanding of distributed systems? Does your experience with accessing, and working on top of the cloud resources make you more comfortable to experiment with widely adopted cloud platforms such as Amazon Web Services (AWS)? Do you think that enabling connectivity between cloud machines is less complex than with physical machines such as Raspberry Pis (no need for setting-up a network from scratch, using physical durines work ac write to Y2					
Q1 Set 3 Q1 Q2 Q3 Q4	less than 2h per week between 2h and 4h per week between 4h and 8h per week more than 8h per week <i>Please evaluate the following statements using the scale</i> <i>from Strongly Affirmative to Strongly Negative.</i> Did your previous knowledge on networking help you to understand the principles of connecting remote machines, and to set-up communication between them (unicast, multicast, and broadcast)? Do you find experience with accessing, and working on the remote machines within a cloud environment, as useful for your general understanding of distributed systems? Does your experience with accessing, and working on top of the cloud resources make you more comfortable to experiment with widely adopted cloud platforms such as Amazon Web Services (AWS)? Do you think that enabling connectivity between cloud machines is less complex than with physical machines such as Raspberry Pis (no need for setting-up a network from scratch, using physical devices such as switch, etc.)?					
Q1 Set 3 Q1 Q2 Q3 Q4	less than 2h per week between 2h and 4h per week between 4h and 8h per week more than 8h per week <i>Please evaluate the following statements using the scale</i> <i>from Strongly Affirmative to Strongly Negative.</i> Did your previous knowledge on networking help you to understand the principles of connecting remote machines, and to set-up communication between them (unicast, multicast, and broadcast)? Do you find experience with accessing, and working on the remote machines within a cloud environment, as useful for your general understanding of distributed systems? Does your experience with accessing, and working on top of the cloud resources make you more comfortable to experiment with widely adopted cloud platforms such as Amazon Web Services (AWS)? Do you think that enabling connectivity between cloud machines is less complex than with physical machines such as Raspberry Pis (no need for setting-up a network from scratch, using physical devices such as switch, etc.)? Do you think that difference between setting-up a java environment for running applications on each					
Q1 Set 3 Q1 Q2 Q3 Q4 Q5	Iterationless than 2h per weekbetween 2h and 4h per weekbetween 4h and 8h per weekmore than 8h per weekPlease evaluate the following statements using the scalefrom Strongly Affirmative to Strongly Negative.Did your previous knowledge on networking help you tounderstand the principles of connecting remote machines,and to set-up communication between them (unicast,multicast, and broadcast)?Do you find experience with accessing, and workingon the remote machines within a cloud environment,as useful for your general understanding of distributedsystems?Does your experience with accessing, and workingon top ofthe cloud resources make you more comfortable toexperiment with widely adopted cloud platforms such asAmazon Web Services (AWS)?Do you think that enabling connectivity betweencloud machines is less complex than with physicalmachines such as Raspberry Pis (no need forsetting-up a network from scratch, using physicaldevices such as switch, etc.)?Do you think that difference between setting-up a javaenvironment for running applications on eachdistributed node in case of physical machines					

Set 1	Strongly Affirmative	Affirmative	Neutral	Negative	Strongly Negative
Q1	14,13	63,65	21,27	0,95	0
Q2	13,49	25,56	38,62	17,57	4,76
Q3	4,66	32,06	29,79	28,31	5,19
Q4	31,48	57,09	11,43	0	0
Q5	56,03	28,41	15,56	0	0
Q6	11,75	23,49	41,59	19,84	3,33
Q7	2,8	10,26	24,13	42,49	20,32
Q8	21,59	52,8	20,85	4,76	0
Q9	9,37	9,42	29,89	34,92	16,4
Set 3	Strongly Affirmative	Affirmative	Neutral	Negative	Strongly Negative
Q1	31,59	55,61	12,8	0	0
Q2	20,26	55,19	24,55	0	0
Q3	4,66	25,45	61,8	8,1	0
Q4	6,08	32,12	39,21	18,84	3,76
Q5	11,75	44,71	32,22	9,47	1,85

Table 2: Average values of the answers

*j*-th survey iteration  $(a_{ij} = \{0, 1\})$ .

$$\frac{1}{M} \sum_{j=1}^{M} \frac{\sum_{i=1}^{N_j} a_{ij}}{N_j} \cdot 100\%$$
(1)

#### 4.2 Discussion

Here we discuss the main findings that we obtained based on the feedback collected from students.

• A general satisfaction with remote on-demand laboratory is expressed in answers on the first, second, sixth, and ninth question from Set 1. The table 2, and graphical representation of the results in Fig. 3 show that majority of students (strongly) agree that a new experimentation environment is a valid replacement for physical laboratory. Applying a statistical Student t-test on the collected sample, we obtained the result of  $p_{value} = 0.028$ , which means that the difference between positive and negative attitude towards new environment is also statistically significant. From an educator's perspective, this result is essential since it justifies the idea and the efforts to embrace new teaching/experimenting practices, i.e., to migrate a physical laboratory to the cloud. Furthermore, this result also illustrates that students are open to gain new practical skills that will lead them to competitive job and career opportunities. However, if we tackle answers on the question 2, students are mostly neutral towards physical presence when it comes to asking questions during classes. However, there are also students that rather ask questions in one environment over another. As we can see from the results, among these students more of them find physical in-classroom sessions more convenient for asking questions, but the difference is not statistically significant. Nevertheless, such result is expected as teaching during both lectures and laboratory sessions is traditionally performed in classrooms, for all courses and students are more used to this kind of interaction. This trend is now changing, and

practice presented in this paper is one of the useful ways for educators to embrace new teaching and experimentation methods, and to proceed with similar approaches.

- Question 5 from Set 1, and question 1 from Set 2 reflect students' experience with unlimited access, as one of the key characteristics of our remote laboratory (Fig. 1). Our results show that providing students with 24/7 access to laboratory resources proves to be a favorable practice. As presented in our previous research [7], students raised concerns regarding limited access. Therefore, this new approach reflects our effort to resolve this important issue by enabling students to access their working machines in an uninterrupted and unlimited way. With  $p_{value} = 0.012 < 0.05$  as a result of t-test, we can see that the result is statistically significant, i.e., the positive feedback is significantly more frequent than negative (no students voted for Negative/Strongly Negative). Furthermore, this is also supported by result in Fig. 6, which also shows that more than 50% of students work in the remote laboratory between 2 and 8 hours a week out of scheduled session hours. Taking into account that students work on their project in teams, and that they distribute the workload among team members, it means that not all of them work on the remote machines at the same time each week. Thus, based on this result we can expect that all groups work in the laboratory for a significant amount of time.
- Tackling students' previous experience in questions 4, 7, and 8, we can clearly see that their programming skills and work on the Linux-based systems significantly help during experimentation within our course. Such result justifies the course prerequisites, and shows the importance of a careful design of laboratory sessions. Furthermore, another result that supports our experimentation on-demand is given by answers on question 7, thereby having most of the students to claim that familiarity with containerization was not crucial to successfully finish laboratory tasks. Since studying virtualization techniques is not officially a part of our course, we recognize expressing indifference towards type of machines (either virtual or physical) as of high importance for this and similar courses.
- Regarding the goal to keep the existing functionalities of machines as in physical lab environment (Fig. 1), students have positively evaluated their experience with networking containers on the cloud ( $p_{value} = 0.039$ ), thereby stating that experience with remote experimentation helped them to understand the distributed systems in general ( $p_{value} = 0.043$ ). Furthermore, most of the students are neutral while expressing the attitude towards experimenting with public clouds (e.g., Amazon Web Services (AWS)), which might be due to lack of students' experience and knowledge on the public cloud platforms.

### 5 CONCLUSION

In order to exploit the cloud resources at the University of Antwerp, IDLab research group, as well as virtualization techniques, we created a scalable cloud-based environment for remote experimentation and development of distributed systems. In this paper we



Figure 3: Students' response on questions within set 1 (questions Q1, Q2, Q3, Q4).



Figure 4: Students' response on questions within set 1 (questions Q5, Q6, Q7, Q8, Q9).



Figure 5: Students' response on questions within set 3

presented the best practices on how virtualization can be leveraged to create a scalable environment for on-demand remote experimentation with distributed systems. Enabling students i) to add



Figure 6: Students' response on question 1 within set 2

more machines to their experiments, ii) to access these machines on-demand during 24/7, and iii) to keep the existing functionalities of machines in the physical laboratory environment, we carefully designed a survey to assess students' satisfaction and experience with this new environment. The valuable feedback collected from 45 Bachelor students reflects statistically significant satisfaction with experimentation on the containers deployed on top of the cloud.

#### **6** ACKNOWLEDGEMENT

We thank Thomas Huybrechts and Dr. Peter Hellinckx for their help and support in setting-up the remote laboratory for Distributed systems.

#### REFERENCES

- Karen Croker, Holger Andersson, David Lush, Rob Prince, and Stephen Gomez. 2010. Enhancing the student experience of laboratory practicals through digital video guides. *Bioscience Education* (2010). doi: http://dx.doi.org/10.3108/beej.16.2.
- [2] Avi Hofstein and Vincent N. Lunetta. 2004. The Laboratory in Science Education: Foundations for the Twenty-First Century. doi: http://dx.doi.org/10.1002/sce. 10106.
- [3] C. M. Ionescu, E. Fabregas, S. M. Cristescu, S. Dormido, and R. De Keyser. 2013. A Remote Laboratory as an Innovative Educational Tool for Practicing Control Engineering Concepts. *IEEE Transactions on Education* 56, 4 (2013), 436–442. doi: http://dx.doi.org/10.1109/TE.2013.2249516.
- [4] G. Jourjon, J. M. Marquez-Barja, T. Rakotoarivelo, A. Mikroyannidis, K. Lampropoulos, S. Denazis, C. Tranoris, D. Pareit, J. Domingue, L. A. Dasilva, and M. Ott. 2017. FORGE Toolkit: Leveraging Distributed Systems in eLearning Platforms. *IEEE Transactions on Emerging Topics in Computing* 5, 1 (2017), 7–19. doi: http://10.1109/TETC.2015.2511454.
- [5] J. M. Marquez-Barja, N. Kaminski, and L. A. Dasilva. 2018. Assessing the Impact of User Interface Abstraction on Online Telecommunications Course Laboratories. *IEEE Access* 6 (2018), 50394–50403. doi: http://dx.doi.org/10.1109/ACCESS.2018. 2868904.
- [6] C. A. Matarrita and S. Beatriz Concari. 2016. Remote laboratories used in physics teaching: A state of the art. In 2016 13th International Conference on Remote Engineering and Virtual Instrumentation (REV). 385–390. doi: http://dx.doi.org/10. 1109/REV.2016.7444509.
- [7] N. Slamnik-Krijestorac, S. Bosmans, P. Hellinckx, and J. M. Marquez-Barja. 2019. Enhancing Students' Learning Experience via Low-Cost Network Laboratories. *IEEE Communications Magazine* 57, 11 (2019), 34–40. doi: http://dx.doi.org/10. 1109/MCOM.001.1900233.
- [8] N. Sousa, G. R. Alves, and M. G. Gericota. 2010. An Integrated Reusable Remote Laboratory to Complement Electronics Teaching. *IEEE Transactions on Learning Technologies* 3, 3 (2010), 265–271. doi: http://dx.doi.org/10.1109/TLT.2009.51.
- [9] L. Tobarra, S. Ros, R. Hernández, A. Marcos-Barreiro, A. Robles-Gómez, A. C. Caminero, R. Pastor, and M. Castro. 2015. Creation of Customized Remote Laboratories Using Deconstruction. *IEEE Revista Iberoamericana de Tecnologias del Aprendizaje* 10, 2 (2015), 69–76. doi: http://dx.doi.org/10.1109/RUTA.2015.2418011.
- [10] M. A. Vivar and A. R. Magna. 2008. Design, implementation and use of a remote network lab as an aid to support teaching computer network. In 2008 Third International Conference on Digital Information Management. 905–909. doi: http://dx.doi.org/10.1109/ICDIM.2008.4746847.