On the Necessity of Establishing a National Cybersecurity Testbed

Sufyan T. Faraj Al-Janabi
1College of Computer Science and IT
University of Anbar, Ramadi, Iraq
2College of Science and Technology
University of Human Development,
Sulaimani, KRG, Iraq
saljanabi@fulbrightmail.org

Abstract—Information security now is considered to be cross-disciplinary and comprehensive field. It integrates the accumulation of knowledge in many disciplines like computer science, mathematics, communications, electronics, physics, etc. Thus, there has been an ongoing effort to improve the experiences in information security experimentation. Many international institutions are investigating enhanced approaches to provide hands-on learning and research environments. However, academic institutions are facing the difficult challenge of providing lab infrastructures that meet the increasingly growing needs of cybersecurity training. In this paper, we report on the necessity and importance of building an effective national testbed for cybersecurity experimentation. We also present a general top-level architecture for this testbed emphasizing the most important enabling technologies.

Index Terms— community cloud; cybersecurity; information security; software defined networks; virtualization

I. INTRODUCTION

In the recent decades, there has been an explosive growth in adoption of cyberspace. Despite the fact that this growth has enhanced our capability to utilize our environment, it has also introduced new threats and challenges to our society. Dealing with this complex issue requires initiation of interdisciplinary scientific research and rigorous development and academic programs in the field of cybersecurity. These are considered to be highly-demanding programs as they are at the intersection of behavioral sciences, formal sciences, and the natural sciences [1]. Due to the severity of challenges imposed and their crucial social, economical, and political consequences, investment in information security has become a national priority for various countries around the world [2].

Moreover, information security complexity involves different tasks like the construction of information network infrastructure, information systems development, information security legislation, development of safety management systems, etc. Hence, information security can prudently be thought of as project-oriented and application-oriented profession [3]. In this direction, hands-on experiments are now essential for information security education. However, cybersecurity laboratory solutions typically require significant effort to build and maintain. Indeed, it is quite challenging to cybersecurity labs to keep pace with rapidly changing security issues to mimic real-world scenarios in a contained environment[4].

Academic institutions have no way but to follow an evolving trend of the increasing reliance on hands-on components to facilitate cybersecurity research and education. There is no doubt that cybersecurity programs in which students are engaged in practical applications are considered to be highly attractive and effective [5].

In order to experiment in information security, instructors and/or researchers can usually either use simulation tools or real testbeds. Simulation approaches provide a cost-effective and easy-to-reconfigure environment. However, the experiences gained from simulation tools can be far from real-world experiences. On the other hand, physical (real) testbeds usually provide real-world experience. However, using real testbeds involves important challenges like the need of a reasonable amount of investment, difficulties regarding flexibility and scalability, and the security threats that they can impose to the public network. In spite of the possibility of reducing these security threats by isolating the testbed from the public network, this can result in considerable difficulties in remote accessing of the testbed from the public network [6], [7].

In contrast with many other traditional courses such as Operating Systems and Compilers that have widely-adopted and effective lab exercises, information security
On the Necessity of Establishing a National Cybersecurity Testbed

Sufyan T. Faraj Al-Janabi

(especially cybersecurity) lab experimentation is still at its infancy. One can note that most information security labs usually have narrow coverage of security principles, concepts, innovative ideas, and real-life scenarios. Another important issue is that these existing labs are typically developed by different people and built upon different environments. This results in a steep learning curve to learn very new environment. Thus, it is necessary to consider information security lab environment from wider, more general, effective, and comprehensive scope [8].

Conducting effective experimentation in cybersecurity is a major activity in terms of the required cost, tools, and equipment. While large international institutes and organization can afford acquiring significant labs for such experiments; however, most (or all) national academic institutions lack the ability to have or afford such resources. Another situation is when researchers who are not affiliated with any academic or research institution want to conduct information security experiments. For such reasons, it is quite important to have a national cybersecurity testbed with distributed open architecture. It is believed that building such a testbed should one of the major national interests from both higher education and research perspectives [2]. Hence, this work represents a first step towards reaching this important national goal.

The remaining of the paper is organized as follows: Section 2 lists some motivations behind the proposal of the national cybersecurity lab. Some important enabling technologies for this proposal are highlighted in Section 3. Next, in Section 4 we report on the state-of-the-art in information security lab design. Some general considerations for the development of the proposal are explained in Section 5. Then, a top-level architecture for the proposed testbed is discussed in Section 6. Finally, Section 7 concludes the paper.

II. MOTIVATIONS

In this section we consider some of the most important points that have motivated this work. These can be summarized as follows:

1. Hacking activities are continuing to appear in almost all countries. These actions have serious economic impact. In addition, there is no doubt that some hacking activities are done for political purposes. This situation is expected to continue despite the fact that there are serious sanctions on such acts in most countries [2].

2. Despite that fact that the great development of globalization and ICT has brought significant economic opportunities for people, this has also presented serious security challenges. Thus, increased numbers of security talents are required. The security of information infrastructures is considered now as a national priority. Academic institutions have to respond to this by offering appropriate information security labs with perfect functions. In fact, information security education can be thought to be an engineering culture. Therefore, when efficient engineering practices are established in academic institutions, this would result in a significant improvement in students’ practical and cognition abilities of technology [3].

3. Most of available information security labs only cover a small portion of the fundamental security principles. Indeed, their underlying infrastructures are different, increasing the difficulty of integration of these labs. Therefore, effective information security labs are still in great demand in security education [8].

4. The lack of efficient cybersecurity components in computer science and IT curricula has been widely reported in many countries. A common difficulty in this direction is the integration of “real-world” labs into the academic courses. Without efficient hands-on real-world activities, it is not possible for the students to integrate the acquired security theoretical knowledge with up-to-date security technologies [9], [10].

5. As traditional undergraduate and graduate level programs do not typically provide the required in-depth training in information security, offering an open national cybersecurity testbed can offer the ability to develop effective certificate programs in cybersecurity targeted towards IT professionals. This is supposed to fill the reported gaps in academic education.

6. Establishment of a powerful national cybersecurity testbed would facilitate a strong cooperation and collaboration among academic and research institutions at both of the national and international levels. This can be beneficial for resources’ sharing, knowledge sharing, and dissemination [2].

7. Concerning the research side, the national open security testbed can provide researchers with the followings [5]:
   • A remotely accessible environment to conduct experiments.
   • An environment that can be isolated from outsiders and restrict public access to intellectual property during the development phase.
   • The ability to quickly deploy and configure IT and computing resources required in experimentation.
   • Supporting the capture of moments in time of the research environments for playback repeated experimentation.

III. POTENTIAL ENABLING TECHNOLOGIES

This section is dedicated to report on some potential technologies that can be used to enable the establishment of the proposed national cybersecurity
Virtualization enables a single physical host computer to simulate the hardware of a number of VMs. The host computer dedicates a portion of the hardware resources to each VM including processors, memory, disk storage, and input/output devices. There are some files in the host computer to describe the configuration of the VM including its processor type(s), allocated memory, installed operating system, and connected input/output devices. Indeed, it is possible to copy these VM configuration files and disk contents to create VM clones or to redeploy VMs across the networks. Some examples of virtualization software and technologies include VMware Workstation, VM Infrastructure, Xen, Virtual Box, and Microsoft Hyper-V [5].

These virtualization technologies have made it easier to set up virtualized information security testbeds compared to deploying physical (real) testbeds. Therefore, many institutions are developing testbeds based on virtualization for various purposes. Most of the currently available virtualization testbeds are built based on the concept of full-virtualization. In such case, VMs are complete systems with their kernels separated from the host. They run on top of software called a hypervisor or virtual machine monitor (VMM). Hence, they indirectly access the host resources via the hypervisor. However, full-virtualization has the disadvantage that VMs are heavy-weight and utilize system resources heavily. This puts constraints on the maximum number of VMs that a physical server can accommodate. Thus, the testbed environment might be restricted. On the other hand, it is also possible to develop a light-weight virtualized testbed to overcome the scalability issue of previously described testbeds. In this latter case, it would be possible to host multiple virtual networks over a single underlying network infrastructure. Developers should take care to prevent traffic and settings of these virtual networks from being interfered with each other. This approach is supposed to provide more cost-effective and scalable architecture [6].


c. Cloud Based Systems

“Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction” [12]. Virtualization is a key technology underpinning cloud computing. In cloud computing, there are three fundamental delivery models, which are (See Figure 1)[13]:

• Software-as-a-Service (SaaS): In this model, the application software is delivered as a service via the internet and is charged on a pay-per-use or subscription-based model.
• Platform-as-a-Service (PaaS): This model delivers a platform that supports the entire lifecycle of an application including runtime, test and development environments as a service via internet.
• Infrastructure-as-a-Service (IaaS): Infrastructure such as server or storage capacity is delivered as a service via the internet. Indeed, automated provisioning and virtualization technologies can be used to enable high scalability and flexibility of the resources.

Regarding the issues of external or internal deployment and the restriction of access to the services, cloud computing has the following deployment models [12], [13]:
• Public cloud: This type of cloud is available to the general public. The services are accessible to everyone using standard internet connection.
• Private cloud: This cloud is operated solely for one organization. It may be managed by the organization or a third party and may exist on premise or off premise.
• Hybrid cloud: This cloud model is a combination of different deployment models such as a public and a private cloud. In a hybrid cloud, users typically outsource non-business-critical information and processing to the public cloud, while keeping business critical services and data in-house.
• Community cloud: Community clouds are clouds that are tailored to the shared needs of a certain community. This model provides the capability to use cloud computing for realizing the required processes and simultaneously preserves high security by means of hybrid deployment models.
Figure 1: Elements of cloud computing [13].

Of special importance to our proposal could be the community cloud model. This model aspires to combine distributed resource provision from grid computing, distributed control from digital ecosystems, and sustainability from green computing. It also provides various use cases of cloud computing, while making greater use of self-management advances from autonomic computing. Some advantages of the community cloud model include [12]:

- Reducing the cost of setting up the cloud because of the division of costs among all participants.
- It is possible to outsource the management of the community cloud to a cloud provider. In this case, the provider would be an impartial third party that is bound by contract.
- Tools residing in the community cloud can be used to leverage the information stored to serve consumers.

C. Software Defined Networks

In general, Software Defined Networking (SDN) and its instance OpenFlow protocol might considered to major enabler technology to the development of open labs. SDN is a recent data networking paradigm that may drastically change the way of operation of current IP networks. It responds to the increasing need for flexible, open, and programmable networks. Significant applications of SDN now include Data Centers and corporate/campus scenarios. However, more investigation is still required for introducing SDN in large-scale IP provider networks. In this respect, different solutions have already been proposed [14].

In SDN, control is separated from data in network switches. Control is centrally allocated in a software controller. The software controller communicates with its switches for flow and access control rules. The communication between controller and its switches is defined by the OpenFlow protocol. Users and applications communicate with the network indirectly through the software controller[2].

Recently, several theoretical underpinnings on SDN have been established. Therefore, many novel network architectures have been deployed. Most of these architectures have individual but unified backbone and access network segments. The backbone segments rely on high performance wired connections to provide high reliability and availability, while the access segments exploit the benefits of the wireless medium. Thus, a major issue has been to bridge wireless and wired networking in order to build enhanced end-to-end systems with the use of SDN technologies. These testbeds enable experimentation on both wireless and wired networks using virtual and/or physical OpenFlow switches[15].

IV. STATE-OF-THE-ART IN INFORMATION SECURITY LABS

Achieving effective experimentation in cybersecurity is considered to be highly challenging to current network testbeds for a number of reasons, such as [16]:

- Scale: In order to be accurate and/or indicative, some cybersecurity experiments may need to be quite large and complex.
- Multi-party nature: Most interesting cybersecurity experiments involve more than one logical and/or physical party.
- Risk: Cybersecurity experiments may involve high risk when not properly contained and controlled.

Some typical challenges that can be encountered when setting up a realistic information security lab includes the following [9]:

- Requirement for protecting campus networks
- Requirement to access the Internet
- Difficulty in allocating required resources for different assignments
- Offering easy and secure access to the resources
- Incorporating latest development technologies
- Overhead of maintenance and configuration of the testbed

Therefore, simulation tools (such as Matlab, NS2, Labview, etc.) and labs are used as an alternative to real testbeds. The simulation tools are supposed to allow building virtual environments that simulate real world. However, the degree at which these simulations are close to real situations can vary from one tool to another or even from one experiment to another. Another experimentation strategy is emulation that can be considered to be in the middle between real (physical) labs and simulation. In this latter case, users can use remote access to physical equipment. Thus, these open labs looks to users as real labs, especially when sufficient resources exist and provided
Internet connections are reliable. Such labs also need effective management schemes to guarantee that time and scheduling schemes are in place [2].

Beside physical (real) and simulation labs, there are several virtual laboratory categories including [4]:
- Virtual Application Laboratories; which rely on desktop virtualization. In this type, the predefined algorithms of the underlying software restrict the simulation.
- Shared-Host Laboratories; which are built on a fixed pool of computers with remote desktop accesses.
- Single-VM Laboratories; which provide predefined VMs for users. This type usually does not have an management portal that creates user-customized virtual resources.
- Multi-VM Laboratories; which provide multiple VMs that can either run in the cloud or on a user’s PC. This type enables users to construct complex experiments. However, such labs may not provide flexible networking, sufficient isolation, or reconfiguration capacities.
- Multi-VM and Multinet Laboratory; which fully utilize the capabilities of cloud virtualization to provide sophisticated experimental environment with multiple VMs and multiple virtual networks.

In the recent years, the world has witnessed the establishment of many large scale national or regional labs for various purposes. Most important examples include GENI (https://www.geni.net) in the US and OFELIA (http://www.fp7-ofelia.eu/) in EU. The major aim of GENI is to host experiments for the future Internet. Some major US universities are forming the infrastructure of this open based on SDN technology. Branching from GENI, a number of customized open labs have been built for certain focuses. For example, some of these focused open labs are Emulab (https://www.emulab.net), CloudLab (https://www.cloudlab.us), AptLab (https://www.aptlab.net/) and DETERLab (https://www.isi.deterlab.net). Of special importance to us is DETERLab which focuses on information security experiments by emulating real world security complexity and experiments [2].

With a few exceptions, most of available large scale testbeds are proprietary. Several of these IT focused testbeds had advanced capabilities but were solely for military purposes. Some other IT focused testbeds (like DETERLab) are available for use by industry and academia. Another example is the Open Networking Lab (ON Lab) which is specialized for SDN. It is only available to members of the ON Lab community [17].

One more important issue to be discussed here is the management of risky experiments. Cybersecurity experiments are inherently risky. Such experiments might involve the release of dangerous malware code, operating a real botnet, and/or creating some other highly disruptive network conditions. Thus, it is necessary to implement suitable isolation capabilities within a testbed. These containment mechanisms may range from completed disconnection from the outside world to allowing narrowly controlled console access. However, this containment itself is highly limiting. In fact, full containment is not very useful. It is well known that powerful experiments are those that can interact with the larger environment. Meanwhile, this should be done in a carefully controlled and well understood ways [16].

In the remaining of this section, we try to give more focus on DETER Lab. The DETER Cybersecurity Testbed has been operated since 2004 to provide a US national resource for experimentation in cybersecurity. The DETER project is centered on experimental cybersecurity research, test, and evaluation. The DETER testbed is hosted at the University of Southern California’s Information Sciences Institute and at University of California at Berkeley. Users receive exclusive, hardware-level access to the number of machines they need, and may setup network topologies, operating systems, and applications of their choice. The testbed can be accessed from any machine that runs a web browser and has an SSH client. Under normal circumstances, no traffic is allowed to leave or enter an experiment except via the supported SSH tunnel [18].

The DETER’s contributions have already helped in better recognizing significant new challenges in cybersecurity field. Most of these challenges are related to the diversity of users and uses of network testbeds, the fast pace with which the cybersecurity field is moving, and the unpredictability and complexity of working with real hardware and software, especially at large scale. Based on these, it is now obvious that the best approach for cybersecurity experimentation can be the adoption of proactive development strategy instead of reactive response to specific risks [18].

DETERLab can be an inspiring example for open testbed to conduct cybersecurity experiments. The lab is open for all users, educators and students in US. Users can reserve required hardware resources and specify in details the type of operating system running on each machine, the software to install on each host, the connections between different reserved hosts, topology, etc. Then, DETERLab will reserve requested resources so that users next can remotely login to those resources and start conducting their experiments [2].

DETER Lab has integrated security education and research missions. This is necessary to transform cybersecurity research into a rigorous experimental science. The research and the educational outcomes of DETER Lab project are summarized in Figure 2 [19].
Figure 2: The synergy of research and educational used of DETER project outcomes [19]

V. PROPOSAL GENERAL CONSIDERATIONS

In this section, we mention some important considerations for the proposed national testbed for cybersecurity experimentation. These can be as follows:

A. Aspects of Services:

The lab should support the following three aspects of services [3]:
1. The lab should be research-oriented by providing good experimental environment for scientific research and validation.
2. The lab should also be teaching-oriented by helping instructors to conduct their courses and students to improve their practice activity.
3. Thirdly, the lab should support extension services via supporting socialization training services for information security majors.

B. Lab Apparatus

The construction of the lab (testbed) involves many new technologies such as to provide the students with the same work environment as they work after graduation. Hence, following principles need to be confirmed [3]:
- The lab apparatus should be practical, easy to use and apply.
- The technology should be advanced, function of the lab should be comprehensive, and devices should be compatible and scalable.
- The lab should provide services for objects with different levels.

C. Testbed Isolation

Due to the well known security concerns, system administrators are reluctant to allow cybersecurity testbeds to be deployed in the campus network. Therefore, the proposed testbed should somehow be disconnected from the outside world to provide an isolated network for experimentation. This is quite justified since the testbed might be used to test ‘dangerous’ computer viruses and worms. Some typical isolation techniques to investigated are: using physical isolation, enforcing lab isolation with firewall rules and settings, and enforced the isolation via the hypervisor configuration (virtualization based).

D. Distributed and Open Architecture

The proposed testbed architecture should be distributed across multiple sites such that to simulate how a real-world corporate network would be configured. Thus, the testbed can be used for projects related to Internet security. One important benefits of the proposed distributed architecture is enabling the share of computing and networking resources with smaller universities or colleges. Furthermore, the open model allows physical labs from different locations to integrate and share resources so that users from those labs and elsewhere can conduct experiments on resources that they do not have locally. This would enable users to emulate remotely computer and network resources.

E. Educational Experience from Instructor Perspective

Users need to become comfortable with the testbed environment. Thus, it is required to set up methods to train students and researchers in this environment. Such resources are needed to aid instructors and researchers in the creation of educational experience.

F. Educational Experience from Student Perspective

From the student perspective, lab models can be divided into the following three groups [7]:
- **Blended learning**: In this group, instruction is primarily face-to-face with lab experiences to supplement the classroom learning.
- **Online format**: In these labs, there is no face-to-face component and all learning takes place online.
- **Hybrid format**: Here the class meets face-to-face but the majority of the activities are online.

The choice of the learning format to be supported by the testbed can vary according to institutions involved. Indeed, the following two teaching philosophies need to be enforced [8]:
1. Information security education should focus on both the fundamental security principles and security-practice skills.
2. Information security education should be integrated into many other courses, including Operating Systems, Computer Networks, Software Engineering, Computer Architecture, etc.

G. Research Implications and Benefits

Besides supporting educational activities, the proposed national lab has also to be well-suited for supporting advanced research activities of national and international interest.
H. Testbed Management

The proposed testbed must have an automated system for management and monitoring. This would reduce the burden on administrators of the testbed to manage labnetworks for various usergroups.

I. Remote Access

The remote sharing capability is a basic concept for the proposed open architecture. Indeed, it is highly desirable for smaller institutions where resources are limited. However, the security aspects of this remote access capability must be enforced carefully. For example, it is possible for users to access the gateway of the testbed using an SSH client without providing them with any administrator privilege on the host. However, such settings need more investigation.

J. Software Tools

One final consideration to be mentioned here is the choice of suitable system, networking, and virtualization software tools. There is a wide range of such tools to choose from including various vendor and open source products.

VI. THE PROPOSED TESTBED TOP-LEVEL ARCHITECTURE

The proposed national cybersecurity lab architecture is distributed and open. Hence, it can be viewed as an aggregation of several labs hosted by different universities and institutions. The lab should have automated software for managing and control all these resources. The lab needs to be designed such that it has an efficient resource allocation system can be configured to allow a certain physical lab to simultaneously provide services to many experiments. Indeed, the lab configuration should enable one experiment to span more than one physical lab when needed. Each local lab need to be supported with at least two connections; one for management issues and the other for conducting experiments. Inspired by DETER Lab, the typical high-level data flow for doing experiments can be as follows:

- At first, experiment details and topology are sent to the lab as a request.
- The national lab should have a container allocation system that receives the request in order to evaluate required resources.
- The container then communicates with the resource allocation system for reserving the required resources.
- The physical resources are configured based on resource allocation information and logical resources are also reserved part of the physical resources.
- Finally, the experiment in conducted for the allocated time period. When the experiment (or allocated time) is finished, resources are released.

Basic experiment configuration should include the following standard three domains [7]:

- The internet domain; which is a generic descriptor for any system or network outside an organization’s control that does not reside in the enterprise domain. This domain represents where the home user, remote office, and the “un-trusted” internet reside.
- The enterprise domain; which represents the architecture of organization’s presence on the internet. This domain typically can be collection of servers providing web content, e-mail, and other network services. In other words, this domain is where the instructor would place servers that represent hypothetical internet presence.
- The administration domain; which is where the necessary support features reside. This domain is where the management, monitoring, and reporting software is installed and used.

Figure 3 shows a typical experiment configuration based on the above basic domains. Note that this configuration is not restricted to local labs. The open and distributed architecture of the proposed national lab can also emulate such configurations.

Virtualization technology will enable the national lab to offer multiple instances of the underlying network infrastructure. This should in the form of separate virtual networks to different user groups. Thus, the following requirements have to be met [6]:

1. Each user group should be oblivious with the presence of other virtual networks that might coexist on the same underlying infrastructure.
2. Traffic of each virtual network must be isolated from the other.
3. Firewall, intrusion-detection system, or other network configurations of any virtual network should not affect traffic of other networks.

The interface for accessing the virtual networks should be user-friendly and easily accessible by remote access tools (e.g., SSH).
The proposed lab architecture can be done on the basis of several phases. In early phases, simple virtualization and networking techniques can be considered. Next, suitable cloud computing configurations can be added. In advanced phases, more advanced and challenging technologies like SDN can be supported. Concerning experimentation domains, early phases experiments can be dedicated to standard Internet security experiments. Later on, various security aspects of wireless and mobile networks might be considered. In advanced phases of the project, it would be also possible to consider cybersecurity experimentation in some advanced application domains. These domains might include:

1. **Large-Scale Semi-Self-Organizing Systems:** The best examples of such systems are botnets. Botnets represent powerful and versatile platforms for attackers. They are characterized by the overall aggregate behavior of thousands or millions of elements [18].

2. **Critical Infrastructure Support:** These systems include power grids, water and gas distribution systems, and control systems for refineries and reactors. They are vulnerable to attacks in both the cyber and physical realms. The intersection of cybersecurity with critical infrastructures in cyber-physical systems [18].

3. **Quantum Cryptography:** The main problem of traditional secret-key cryptosystems is secure distribution of keys. Indeed, the security of public-key cryptosystems can be threatened by advances in technology and mathematics. The quantum cryptographic approach can provide “unconditional security” based on laws of physics. However, experimentation in this new field requires the adoption of new technologies.

### VII. Conclusion

In this paper, we have considered the necessity of building a national open testbed for cybersecurity experimentation. This testbed can have significant educational and research benefits. The establishment of such a testbed is prudently justified by both international trends and national security requirements. Basic motivations behind this proposal have been outlined. The proposed lab architecture can support instructor, students, and researchers all over the country or region regardless of their location. Top-level design considerations and major challenges have been addressed. Future works might include more detailed architecture design and more investigation of technology adoption choices.

### References


