

Article

# Can a Network Attack Be Simulated in an Emulated Environment for Network Security Training?

Samuel Chapman <sup>†</sup>, Richard Smith <sup>†</sup>, Leandros Maglaras <sup>\*,†</sup> and Helge Janicke <sup>†</sup>

School of Computer Science and Informatics, De Montfort University, The Gateway, Leicester LE1 9BH, UK; P13188544@my365.dmu.ac.uk (S.C.); rgs@dmu.ac.uk (R.S.); heljanic@dmu.ac.uk (H.J.)

\* Correspondence: leandros.maglaras@dmu.ac.uk; Tel.: +44-116-207-8483

<sup>†</sup> These authors contributed equally to this work.

Received: 29 May 2017; Accepted: 8 August 2017; Published: 11 August 2017

**Abstract:** This paper outlines a tool developed with the purpose of creating a simple configurable emulated network environment that can be used in cyber defence exercises. Research has been conducted into the various related subject areas: cyber defence exercises, network threats, network emulation, network traffic replay, network topologies, and common network services. From this research a requirements specification was produced to encapsulate the features required to create this tool. A network, containing many of the aspects researched, was designed and implemented using Netkit-NG to act as a blueprint for the tool and to further knowledge in the construction of an emulated network. Following this the tool was developed and tested to ensure requirements were met.

**Keywords:** emulation; network threat; network attack; network services; network topology; cyber defence exercises

---

## 1. Introduction

Computer security is an essential part of keeping an organisation's data safe from malicious affect. One aspect of computer security is security of a network which is used to ensure networked computers and servers are reliably available to users, and data on these devices is secure. To ensure a network is secure network security experts are required to respond to threats as they appear. These experts are required to be trained in the various aspects of network security and have a strong knowledge base of various network services and how these can be secured [1]. One method of training is creating cyber defence exercises. These exercises utilise a blue team and a red team. The blue team is tasked with defending and securing a network, whereas the red team's objective is to attack the network. These exercises do not take place on live networks due to their nature, instead, a network is created with this purpose in mind [2].

Virtualised cyber security training and facilities are offered by several organisations, among these are the SysAdmin, Audit, Network, and Security (SANS) Institute [3], and Offensive Security [4]. The SANS Institute does not go into specifics regarding the training only that it offers hands on experience and utilises virtual lab environments which are capable of simulating "the most prevalent tactics and technologies". Offensive Security on the other hand offer more insight into their training stating that they utilise a virtual network of more than forty systems, ranging "from locked down modern servers to classic configurations still found in networks everywhere". Both offer network security training, however, neither states whether they allow customisation to their virtual networks. This work offers the ability to customise the structure of the network with ease. The creation of a network and simulation of traffic, both neutral and otherwise, can be time consuming; this work allows a network, containing a number of services, to be produced to a specification quickly and traffic to be subsequently simulated on it. The CyRIS (Cyber Range Instantiation System) [5] is a tool

similar in function to that proposed. The proposed tool does, however, obfuscate more of the cyber range's creation by only asking the types of devices the user would like. CyRIS requires the user to have knowledge enough regarding networks to create a configuration file which it will use to create the virtual machines and install other software to start the range; the proposed tool will require no such knowledge and will produce simulated attacks and traffic at random intervals. Another commercial offering in Raytheon's cyber range [6] which like CyRIS is fully customisable, on their website they quote that a customised cyber range can be created in a matter of hours, whereas the tool created by this work may be less complex but can ultimately create a full network in a matter of minutes. The Cyberbit range is yet another designed with the intention of training computer security professionals [7]. As with Offensive Security's offering this range is large and extensive but is not customisable for the end user. The various cyber ranges looked at here show that there are a number of different offerings with various features. The tool designed in this work is unique in that it is capable of constructing a full network quickly and without extensive knowledge, which will ultimately allow focus to be shifted more towards cyber security training rather than maintaining the cyber range.

Network emulation tools are used to mimic a real network environment, and can therefore be used for cyber defence exercises. These tools have to be configured to specifications created by an end-user. This can be a long process given the possible number of network services required to create a network. This article presents a research that was conducted into cyber defence exercises, network threats, network emulation, network traffic replay, network topologies, and network services. A tool that is capable of creating an emulated network containing the various services found from the research is presented. This tool is capable of mimicking network traffic produced by normal usage of network services, also known as white traffic, in an attempt to obfuscate any malicious network traffic that might be produced in a cyber defence exercise.

The contributions of this article are:

1. Design and implement a virtual network utilising common network services (see Section 3).
2. Add the ability to configure the network based on user requirements (see Section 5).
3. Generate and capture network traffic that simulates normal network usage (see Sections 4.1 and 4.2).
4. Automatically replay captured network traffic (see Section 4.3).
5. Test the system (see Appendix B).
6. Randomise the replay of captured network traffic, to further obfuscate any traffic of a malicious nature (see Section 4).

The tool produced by this work allows professionals to receive training in the security of sensor/actuator networks in a variety of user definable network scenarios. This tool also allows the production of a complex virtual network, and the simulation of network traffic, without requiring an extensive background in that area by creating the virtual network based solely on user input. This allows more focus to be placed on training professionals than setting up and maintaining the network, the attack simulation element of this tool also means common threats can be simulated with ease and without intimate knowledge of the tools required to attack a network, thus shifting focus more on defence and protection of a network.

#### *Paper Discovery Method*

The procurement of information regarding this subject was conducted in a specific manner. Firstly, the heading of a section was searched, initially this was done within specific parameters. Initially the date was considered, this was confined to papers written after 2005 to ensure current validity of data, the most recent papers were considered first. Also the number of citations was looked at in an effort to determine impact within the research area. These search terms were not binding however as this was sometimes inappropriate. For example the use of the Request for Comments (RFC) pages in Section 2.6 was deemed acceptable, although these pages could be considered dated, as these were the documents

that defined the protocols making them a very relevant source. The title of a document was then looked at to determine relevance to the subject area, the abstract of the document was looked at in the same manner. Once this had been done, the papers were read in more detail and the content used where appropriate.

## 2. Related Work

In this section we present all related work that is related to our work including, network emulators, common network threats, the construction of a network, and services that are common throughout networked systems. It takes into account the various aspects required for this system, and aids in making informed decisions regarding the tools and configuration required.

### 2.1. Cyber Defence Exercises

Cyber defence exercises usually consist of at least two teams, one defending a computer network (blue team) and the other attacking it (red team). This type of exercise is often used for training students and security teams. It also offers an opportunity to see how a team would react if under an actual cyber-attack. The United States National Security Agency (NSA) has been conducting these exercises since 2001 [8]. Both Mullins, and Adams et al. [9] agree that these exercises offer invaluable hands on experience and are an effective method of training people in proper cyber defence. Both agree this is because trainees are put in a pressurised environment during the exercise. However Adams takes this further suggesting it is also the element of teamwork inherent to these exercises that makes them truly effective methods of training. The cited authors come from military backgrounds, which could bring into question the reason they strongly approve of cyber defence exercises. However, most of the evidence gathered is from user experience and there are multiple reviewers at these events. Although this type of exercise has been going on for many years there has been little or no previous research into simulated red teams attacking a network.

### 2.2. Network Threats

A threat is something that has the capability of affecting any part of an organisation in a negative way [10].

#### 2.2.1. Threat Actors

A threat actor is *an individual or group posing a threat* [10]. Threats can come from a number of different actors, each with different motives and levels of funding. National Institute of Standards and Technology (NIST) has produced several scenarios where different threat actors are present. However, the SysAdmin, Audit, Network, and Security (SANS) institute and the National Cyber Security Strategy explicitly state what they believe are key threat actors:

- Cyber Criminals—Individuals or groups that attack systems for financial gain, for reputation, for performing malicious activities, for carrying out “conventional crime” or for erasing data.
- States or State Sponsored—Individual or groups funded by governments to attack systems typically used for cyber espionage or to cause damage.
- Hacktivists—Individuals or groups who attack systems for political or social reasons.
- Terrorists—Individuals or groups that attack systems to cause damage and destroy.
- High Privileged Insiders—Individuals or groups with privileged user access seeking financial gain.
- Partner—Trusted partner of an organisation seeking financial gain or a competitive advantage.

Although not all documents agree on each actor they all pose a threat to an organisation or nation. Also, the NIST and United Kingdom’s documentation holds governmental authority, making them viable sources of information even if they do not directly agree on all matters.

### 2.2.2. Attack Taxonomy

Network attacks can take many forms and several taxonomies have been produced by various groups in order to express these forms. Most of these taxonomies share similarities and it is the common attacks that will be discussed. However, attacks not pertaining to networked systems and attacks utilising human-computer interaction have been deemed irrelevant for the given context. The McAfee threats report [11] has been used in conjunction with other various taxonomies to ascertain how prevalent the types of attack are. This report outlines various statistics of threats detected by McAfee's various security applications.

Denial of Service (DOS) is discussed by a number of papers, each with very similar definitions. This type of attack aims to prevent legitimate users accessing network resources [12] by overwhelming the target system with non-legitimate requests. Chapman, [13], goes on to discuss a Distributed DOS (DDOS) which utilises a number of computers to overwhelm the target. This is somewhat more sophisticated and is harder to defend as there are many sources of the attack. Although the paper written by Hansman and Hunt [12] is several years old, Chapman's paper shows that the make-up of a DOS attack has not changed in a significant way. In the third quarter of 2016 DOS made up 33% of network attacks which is the largest proportion in the McAfee report. This shows that this type of attack is likely to be seen a lot by network security practitioners.

Brute force attacks made up 18 % of attacks in the third quarter of 2016 showing that they are a significant threat to networks. Chapman describes this type of attack as *attempting all possible password combinations*. While this type of attack is not particularly sophisticated it poses a clear threat to networked systems. A Domain Name Service (DNS) attack is an attack that aims to send legitimate web users to illegitimate websites. The attacker will pose as a DNS server and then direct users to a fake web page rather than the intended one. This is usually a website which the user has to log in to [14]. This type of attack gives the attacker various user credentials that can be used in a legitimate system. This type of attack made up 4% of the attacks in the third quarter 2016 [15]. The information found clearly shows the prevalence of the different types of network attack. This data will be used to produce the attacks featured in the proposed system.

### 2.3. Network Emulation

Network emulators offer the ability to mimic a network without the need for the requisite equipment, such as cables, routers, server computers, etc. There are many different emulators available, each offering various features to aid the environment creation process. The work of Fuertes and Vergara [16] attempts to measure the resource consumption of several emulators. The emulators chosen for review may have changed functionally since the date of the aforementioned publication, however, this is taken into consideration.

#### 2.3.1. Netkit

Netkit is a free, open source emulator which includes several real world tools to enable the creation of servers. Each device started via Netkit uses a kernel based on the Debian Operating System (OS) which is an OS in use today [17]. The use of a real-world OS allows the installation of software that is not natively available on the virtual machines. The work of Wolny and Szoltysik [17] may be biased, as they discussed Netkit in much greater detail than other software in their overview. They have also stated that Netkit is an available feature in a teaching operating system. Netkit labs can be set up using a set of text files that describe the virtual machines and the desired configuration meaning the labs do not require much storage space per machine which in turn makes it easy to transfer the labs to different machines [18]. It should be noted that Pizzonia and Rimondini are on the team that created Netkit so they are unlikely to have objective views on the subject, that being said, they would be well versed on Netkit's features. As mentioned previously, Netkit machines use the Debian kernel however, the kernel in use is Debian Lenny which has been deprecated since 2012. It should also be

noted that the most recent stable version of Netkit was released in 2011. There is a variation of Netkit called Netkit-NG, this uses a more recent version of the Debian kernel, Debian Wheezy, and was last updated in 2015. Netkit-NG also offers several different features, whilst still being functionally similar to Netkit.

### 2.3.2. Virtual Network User-Mode Linux (VNUML)

Similar to Netkit, VNUML is an open source emulator, however rather than using a specific kernel VNUML uses the current OSs kernel. Because of this the user has to have elevated privileges for some of the configuration [19]. In order to create a VNUML lab the user must write a file using a variant of eXtended Markup Language (XML) specific to VNUML. The whole lab is contained in a single file, split into two parts; virtual machines (VMs) and virtual networks [20] unlike Netkit which has several files describing each VM. The referenced document was written by the developers of the application. The segment of interest regards features rather than performance statistics or a comparison. VNUML has been superseded by a newer version of the tool called Virtual Networks over linux (VNX). VNX offers some new features but essentially works in the same way as VNUML, development and maintenance of VNUML has been abandoned.

### 2.3.3. VMware Server

VMware server is different in comparison to the aforementioned emulation software packages. VMware server can support multiple operating systems in a single network which means you are able to have the Windows OS running on one network node and Linux on another. This feature allows the user a wider range of network capabilities [21]. Since writing the paper, Li [21] has been cited several times in other papers, however VMware server has been updated since then and some of its features have changed in that period. To create a VMware server lab, the user must install the desired operating system on a Virtual Machine (VM) and manually set it up as though it were a stand-alone machine. They must then configure virtual network adapters on the host machine to allow several machines to communicate. When compared to VNUML or Netkit, this could be considered a lengthier setup process. The VM will also be contained within a single folder containing files which cannot be edited while the VM is running. Unlike the previous emulators, VMware has not been changed developers or names and is still maintained today.

Although each emulator is similar they all offer varying functions. Except from the aforementioned emulators there exist a number of others that can be used for conducting similar tasks e.g., GNS [22], Dynagen [23], Velnet [24], ModelNet [25], Vagrant [26], Vnet [27] among others. Between the three emulators that we described more analytically in this section, VMware Server is somewhat more difficult to setup and configure than the other options although it is able to offer more flexibility in terms of its functions. VNUML and Netkit allow the user to setup VMs using text files or folder structures stored on the host machine. This means that changes can be done without starting the lab. Netkit structures lab configuration using machine specific folders and start-up scripts allowing the user to see at a glance what is going to be emulated. This may be a more appropriate option for creating a pre-configurable network lab.

## 2.4. Replay of Network Traffic

In order to make the proposed environment more realistic, network traffic is captured and replayed across the networked system. This will help to obfuscate any *red team* traffic that will occur on the network creating more of a challenge for the *blue team*. After some research, the most appropriate tool for this task is Tcpreplay. Tcpreplay gives a user the ability to replay captured traffic onto a network. It also allows the user to edit the packets to a certain degree, meaning the source and destination addresses could be changed to suit the network. This information was taken directly from the maintainer's website, although they may be biased regarding the utility the description and features are the only items taken from this rather than data regarding its effectiveness.



Although the aforementioned website does not discuss Tcpreplay's performance, tests conducted by Hillstead et al. [28] indicate that the tool is able to reproduce accurate traffic that is produced as and when the user requires.

### 2.5. Network Topologies

The topology of a network refers to how network nodes are both physically and logically linked. The topology can affect the cost of a network and its integrity, so it should be something carefully considered. To effectively emulate a realistic network, an understating must be established regarding the most common topologies that are currently in use. Network topologies are rarely changed or redesigned therefore there does not seem to be a lot of research in this specific area. From several books, that cover basic network topics, the works selected have been cross-referenced with each other in an attempt to ensure information is correct and unbiased. There are four key network topologies, mesh, star, bus, and ring.

### 2.6. Common Network Services

Network services give a network various features such as file storage, the ability to produce web pages, email, etc. It is imperative that common services are used to properly emulate a corporate network and the attacks associated with them. A list of widely used services was generated by cross referencing the works of Dean [29], Duck and Read [30], Kurose and Ross [31], and White [32]. To get information about these protocols, their Request For Comment (RFC) pages have been used. An RFC is where a protocol has been defined and informs people in how communications utilising them should look. These are produced by the Internet Engineering Task Force (IETF) who are responsible for producing technical documents regarding internet protocols.

### 2.7. Email

Email is widely used around the world to send messages. There are three key protocols associated with email, Post Office Protocol—version 3 (POP3), Internet Message Access Protocol (IMAP), and Simple Mail Transfer Protocol (SMTP). The SMTP protocol is designed to transfer and deliver mail between two nodes, this is one of the most common protocols for sending an email. IMAP and POP3 are both protocols designed for retrieving mail from a server. Accessing mail on a server is often seen as a cheap solution as it means user's PCs do not have to have SMTP capabilities in order to receive mail. POP3 is used to download mail from a server directly to a PC. Once it is downloaded, the user is able to read and delete the mail, however, the mail will no longer exist on the server. The IMAP protocol allows users to access the server on which their mail is stored, from there they are able to manipulate it as though it were on their own device. This gives users the ability to access their mail from any device as long as they can access the server. IMAP would be the preferred protocol as it gives users more flexibility with the device they use. Except from the above, webmail that is any email client implemented as a web application running on a web server is an option widely used nowadays.

### 2.8. File Transfer

The transfer of files is required by most organisations as a method of sharing and storing data. This process is facilitated by the File Transfer Protocol (FTP). This protocol allows users to share and view data on a remote server. There are variants of FTP which allow a degree of security for file transfers, Secure File Transfer Program (SFTP) and FTP over SSL (FTPS). SFTP is very similar to FTP however it completes all tasks over an encrypted Secure Socket Shell (SSH) channel. The user will connect and then log in to the SFTP server then commence communication. FTP over SSL (FTPS) offers encryption for the FTP protocol using Secure Socket Layers (SSL) or Transport Layer Security (TLS). The use of this encryption must be specified when connecting to the server [33].

### 2.9. Webpages

The Hypertext Transfer Protocol (HTTP) is used for data transfer across the internet, and is commonly used for webpages. A variant of HTTP is Hypertext Transfer Protocol Secure (HTTPS) which uses SSL or TLS to encrypt traffic (Rescorla, 2000) in a similar way to FTPS. DNS is a method used to resolve human readable Uniform Resource Locators (URL) into an Internet Protocol (IP) address. This system is used throughout the internet for web pages and connecting to servers remotely. DNS is also used for addressing email. A database is used that contains both the URL and the IP it refers to, from which the URL can be resolved. This database is updated manually. A similar protocol, known as Dynamic Domain Name System is used for the same end. The key difference is that this system allows for updates to the database that contain the URLs and IP addresses dynamically from the network. Dynamic Host Configuration Protocol (DHCP) is a protocol that allows the automatic assigning of IP addresses. This is used so new devices do not have to be manually configured to be seen on the network. This is used by organisations as it makes it easier to add new devices to their networks.

### 2.10. Remote Access

There are two protocols that are commonly used for remote system access. The telnet protocol allows a user to connect to a networked device and use various system commands for general use of that system. The telnet protocol does not include any user authentication steps which leaves it open to attackers. The SSH protocol works in a similar way to Telnet however it requires users to provide credentials to gain access to the system making it somewhat more secure.

The literature surrounding the subject suggests that an automated system able to simulate network attacks has not before been devised. However, given the benefits of cyber defence exercises, an automated system would be beneficial as it should enable teams to train or learn cyber defence techniques without having to create two different teams: attack, and defence. The literature reviewed has also given indication as to the more prevalent attacks in recent times as well as various protocols that are used in modern day networks. These are taken into consideration when constructing the proposed system.

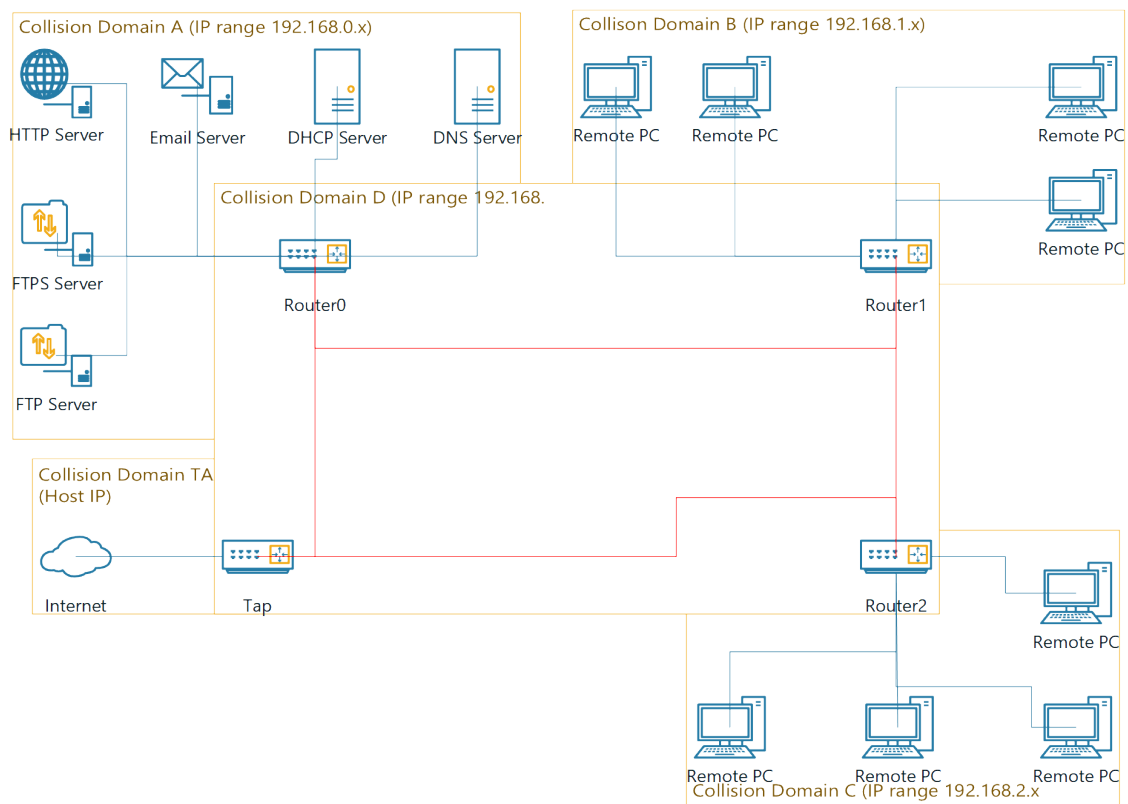
### 2.11. Key References

There were a number of important sources found during the review of the various literature, these have been reduced to the three considered most influential to the writings. The paper written by Adams et al. (2009) regarding cyber defence exercises is very important as it validates the reasons that such exercises have started, for the training of network security teams. Dean's book (2012) regarding networking was another influential piece of work for this review. It shows what topologies are available to choose from and states, in the author's opinion, which is best. It also goes on to explain the various networking protocols that are in common use today which was useful as a stepping stone to analysing these protocols. Fuertes and Vergara's work (2007) outlines the various network emulators that are available. Although this work was written several years ago, it is one of the few that offer a direct comparison of available emulators without discussing possible matters of opinion.

## 3. Network Overview

The network has been implemented in a Netkit environment and includes many of the services discussed in the literature review (see Section 2.6). A summary of this network is provided however a more in depth technical explanation, of both the network setup and Netkit-NG setup, has been provided. The topology used mimics that of a star-wired bus, as it was decided from the various literature that this was the most suitable for a network attempting to mimic that used in a corporate environment. Figure (see Figure 1) shows the network in this topology with the various nodes that

appear on it. This network has been split into several collision domains, which act as the different areas of the topology.



**Figure 1.** A diagram of the implemented network.

### 3.1. Collision Domain: TAP

The diagram depicts five collision domains, A, B, C, D, and TAP. The tap collision domain is used to provide access to the wider internet via the host computer, it will also allow the user to interact with the emulated network via the host machine, a more in depth view of the TAP node. It was deemed that this feature should be implemented to allow the network to better produce traffic by interacting with the internet.

### 3.2. Collision Domain: D

Collision domain D contains the routers of the network. These routers allow network traffic to be transported throughout the other collision domains. As well as routing, the nodes on this collision domain act as DHCP relays, by utilising the ISC-DHCP-Relay tool. This allows them to forward DHCP packets from the DHCP server on collision domain A to the nodes on their subnet.

### 3.3. Collision Domains: B & C

Collision Domains B and C contain remote PCs. These nodes do not utilise any servers and are used to simulate standard network nodes without special function. Netkit-NG does not have a web browser installed by default so this has to be done when the remote PCs initially start-up. This process is scripted so that the user does not have to do this, nor does the user require an internet connection for this. The web browser chosen was Links as this is an easy to use browser that can operate in the Linux terminal. Pine, an email client, is also used on these nodes. This is a simple text based mail client which allows users to view and send emails. This was chosen as it requires little configuration to function and it is possible to use without a graphical interface. To give users individual mailboxes,



several accounts were created on the email server, which are logged into when Pine is started on the remote PC.

#### 3.4. Collision Domain: A

Collision domain A contains all of the servers for the network, documentation for the setup procedures of this collision domain has been provided. The servers were placed on the same collision domain to make the setup process easier and to make the network simpler to understand to those that need to use it.

##### 3.4.1. File Transfer Protocol (FTP) Server

The FTP server daemon used is ProFTPD, this was chosen for a variety of reasons. ProFTPD is available on Netkit-NG, this means that there is no requirement to script its installation therefore reducing its time to start-up. It is also possible to configure FTP, SFTP, and SFTP, using this server daemon which offers a level of convenience during network setup. Several user accounts have been set up on this server to allow login when a connection is established.

##### 3.4.2. SSH File Transfer Protocol (SFTP) Server

The SFTP server is set up in the same way as the FTP server; however, it uses SFTP to ensure a higher level of security on the server. After some testing it was found that ProFTPD was unable to run FTP and SFTP server concurrently, it was therefore decided that they should run on separate nodes.

##### 3.4.3. Hypertext Transfer Protocol (HTTP) Server

The HTTP protocol has been implemented via the apache daemon. Like ProFTPD this was already installed on Netkit-NG so using this was the preferred choice to reduce network start-up time. This daemon also offers HTTPS functionality which has been implemented on the same node. Whether to have a separate HTTPS node was considered, however, to include both HTTP and HTTPS on the same node had no foreseeable drawbacks and helped to consolidate services making the network as a whole easier to understand.

##### 3.4.4. Email Server

The email server makes use of the IMAP, POP3, and SFTP protocols. To do this, it uses two server daemons, exim4, and OpenBSD-inetd. Initially, there were no daemons capable of handling email, so these are installed via script when the email server starts-up. Exim4 is used as the Mail Transfer Agent (MTA), which is used to handle emails and ensure they get to their proper destination. This server was chosen as it was found to be easy to set up and capable of handling both the POP3 and IMAP protocols. It was found in the literature review that IMAP was considered more appropriate for the given scenario. POP3 capabilities did not take any more time to implement in exim4 and allow the user to use this if it is required. OpenBSD-inetd has been used to direct any traffic, that uses POP3 or IMAP, coming to the user to the imapd and ipop3d programs. These programs are open source and enable access to the mailbox directory on the email server. OpenBSD-inetd was chosen as it is simple to direct specific traffic to an application. The imapd and ipop3d programs were used as they do not require installation to use, thus cutting start-up time. They also do not require much configuration.

##### 3.4.5. DHCP Server

The daemon chosen to implement DHCP was ISC-DHCP-Server. This daemon was chosen as the configuration files are easy to read and understand, provided there is already knowledge of the DHCP protocol. This daemon needed to be installed via a script as it did not come pre-installed with Netkit-NG. There is a daemon already installed on Netkit-NG that is capable of handling DHCP, dnsmasq (see Section 3.4.7).

### 3.4.6. Domain Name System (DNS) Server

Bind9 was implemented as the DNS server for the network. This server comes installed on Netkit-NG so there was no need to install this on start-up. Rather than utilising standard DNS, where each device would have to be manually configured within the chosen daemon, the network uses Dynamic DNS (DDNS). DDNS allows for address assignment as nodes appear on the network. It also allows devices with non static IP addresses to be assigned host names and communicate. This feature reduces the need for remembering a large number of IP addresses.

### 3.4.7. Critical Analysis

Although most of the services described in the literature are available in the designed network, some were not included. SFTP is not available to the network. This was due to difficulties in its initial setup. It was possible to configure an HTTPS server which uses technology similar SFTP it was more difficult to implement this with ProFTPD.

As mentioned in Section 3.4.5 the DHCP daemon native to Netkit-NG was not used. ISC-DHCP-Server was chosen over dnsmasq as it requires less configuration to enable DDNS whilst using separate DHCP and DNS servers. Using separate DHCP and DNS servers offers users more options when using the network as they can choose not to use the DNS server entirely, this was an option not found when using dnsmasq. ISC-DHCP-Server also works slightly better with the ISC-DHCP-Relay daemons utilised in the routers in collision domain D.

## 4. Network Traffic Replay

This section gives an overview of how network traffic was produced and replayed across the network. Once the network was built, network traffic could then be produced. This traffic would then be captured and used to create white traffic on the network. This is done to make it more difficult to distinguish normal network traffic from that of a threat actor.

### 4.1. Creation of Network Traffic

Creating network traffic involved using the servers created previously (see Section 3). These services were used on the *remote* nodes, as they are designed to mimic a non-server device which would use the available services. Creating network traffic involved using the various services available on the network. For example, to generate email traffic, several emails were sent using the *remote* nodes (see Section 3.3). This process was repeated in a similar manner for FTP, SFTP, SSH, HTTP, and HTTPS in that the services were simply used to generate the traffic. The DHCP service will automatically function when it is needed by the network, as this is one of its key functions. Equally, the DNS service would be used as required by the network, and when such traffic was produced for DNS as the above services were used. Similar to the DNS service, the routers direct traffic based on network usage and, as such, were used in the production of traffic.

As well as producing network traffic from the services available to the constructed network, traffic was also produced using websites available on the internet. This utilised the tap node (see Section 3.1). This traffic was generated by using the Links web browser to browse the internet. Several websites were selected to browse and these were selected based on several criteria. The first criteria was where they ranked in three different websites that list most used websites, these lists are generated by alexa.com, moz.com, and similarweb.com (n.d.). Three different lists were used as each claim to use several techniques to gather website usage data. The second criteria was whether log in credentials were required to access the website. This simply ensures that no private details would appear when the traffic was captured. The final criteria was whether the website could be displayed properly in the Links browser. Links is a text based browser and as such is not capable of displaying images and other more advanced web features. If a website met these criteria, it was cross-examined against the other lists. The websites selected were amazon.com, google.com, reddit.com, and wikipedia.com.

To generate traffic, these websites were simply browsed, links were followed, and in some cases generic search terms were entered, if a website had that functionality.

#### 4.2. Capture of Network Traffic

The tool used to capture the network traffic was `tcpdump`. This tool was used as it is able to capture all network traffic going to and from a node to a file, in the PCAP format and is included in the Netkit-NG software. The traffic was produced by the remote nodes (see Section 3.3) and captured on the router nodes as these nodes are used whenever the remote nodes communicate with the wider network (see Section 3.2). Separate files were produced based on the service used. For example email traffic has been stored in a file called `email1.pcap`. This was done so that traffic could be replayed at random intervals. Several files were produced for each service. This meant that the services could be used differently. For example two FTP files were produced, the first contains traffic of the remote node sending a file to the FTP server, and the second contains traffic of the remote node pulling a file from the FTP server. Again this was done to allow for some level of unpredictability when the traffic is replayed.

#### 4.3. Traffic Replay

For replaying network traffic, tool `TCPReplay` was used (see Section 2.4). This tool allows the editing of captured traffic which was required to make traffic look as though it is being created by a legitimate source, this is what is meant by white traffic. After editing the traffic it could then be replayed. Several bash scripts were created which perform randomization of traffic function. These scripts ensure that a random traffic capture file is replayed at random intervals in the space of one hour. This time frame can be altered if required. The capture and replay of traffic was a relatively simple procedure with few complications. The decision to utilise the available services allows a wider range of traffic to be produced by the network which would help to obscure any malicious traffic.

### 5. Program Overview

This section presents the various aspects of the program created to give users the ability to create a network environment to their own specification. This program was developed in the Python programming language as it offers the required functionality and can do so in an uncomplicated manner.

The user interface is a Command Line Interface (CLI). When the program is started, it prints basic information about the tool and then asks the user a series of questions regarding the network configuration they would like produced, such as how many instances of a particular node they would like Figure 2, shows the program in use and the questions asked.

There is room for improvement in the user interface as this program uses a CLI. While the program works without issue, the interface used is not as easy to use as a Graphical User Interface (GUI), which would offer a better display of the various options available to the user. The method used to garner the users subnet choice is also an area that could be improved.

```

Network Setup Tool
This tool should be used to configure a network to your specifications
The network will use a DHCP server to address the various machines
Would you like to delete the previous Lab? (Y/N)
y
Would you like to configure your own network? (Y/N)
y
How many Subnets would you like?
3
You have created 3 subnet/s, named 0-2
Would you like to implement DNS? (Y/N)
y
On which subnet would you like your dnsServer? Choose a number between 0 and 2
0
The DHCP Server will be placed on the same subnet as your DNS Server:
Subnet 0
Would you like a tap machine?
(This allows your network to talk to the wider internet)
You will require an active internet connection for this node to setup correctly
Y/N
y
The tap machine will be on subnet 254
How many email servers would you like? Maximum of 254
1
On which subnet would you like your emailServer? Choose a number between 0 and 2
0
How many FTP servers would you like? Maximum of 254
1
That was not a valid number. Try again...
1
On which subnet would you like your ftpServer? Choose a number between 0 and 2
0
How many ftps servers would you like? Maximum of 254
0
How many HTTP servers would you like? Maximum of 254. N.B. The HTTP Server also implements HTTPS
0
How many non-server nodes would you like?
1
On which subnet would you like your non-server node? Choose a number between 0 and 2
1

```

**Figure 2.** A screen capture of the program user interface, with various options completed.

## 6. Discussion

Research was conducted into common network services and cyber defence exercises in the literature review. A networked system was planned and created (see Section 3), this network can also be configured to a users specifications (see Section 5) which gives multiple learning opportunities in the practical implementation of networks and programming. The creation of the network puts into practice the services researched in the literature review. Extensive tests have been implemented. These tests are used to test that the functions work as described (see Appendix B).

Following the work conducted, there are aspects that could be improved or added. Adding features to the system could offer a better user experience. It could also give the system a wider user audience. The ability to emulate hostile traffic being one of them. This would increase the functionality of the system and allow a single team to conduct a cyber defence exercise rather than a red and blue team. Another feature that could be added is a graphical user interface (GUI). A GUI would make the program much easier to use. The command line interface (CLI) is useable, however, on a large network it would be tedious to configure the subnet for each device individually. A GUI could also offer a network diagram while the user is creating the network which would make their configuration much easier to visualise. Adding the ability to choose from a range of server daemons would also be a useful feature for the network creation program. This may be useful as different daemons might present different vulnerabilities or may react differently to different attack types. More research would have to be conducted in this area to gain a better understanding of how different daemons operate.

Currently, it is possible to use the program to create more nodes on a given subnet than is possible to address, the maximum being two-hundred and fifty-five. Doing so would likely create problems

within the network. The improvement would involve a calculation to check how many nodes are already assigned to a subnet and then give an error message if a user attempts to add more than two-hundred and fifty-five. Improving the way the DNS and DHCP servers function would also be desirable. Currently the DNS and DHCP servers have to be on the same subnet; if they are not, the DNS server will not function correctly. It would be useful if this was not the case as it would allow more user configuration. Implementing this would require changes to the program to make this option configurable, it would likely also require a reconfiguration of the two servers to allow them to communicate from separate subnets. More research would be required to make this improvement as it may not be possible to implement this.

This article describes the procedure of creating an emulated network containing various services. The produced tool is capable of mimicking network traffic produced by normal usage of network services, also known as white traffic, in an attempt to obfuscate any malicious network traffic that might be produced in a cyber defence exercise. The environment that is presented on this paper is the basis for feature development of a fully automated cyber defence environment that can be used for training. In this environment we will integrate already deployed IDS systems [34] in order to facilitate automatic detection of attackers. Also the datasets that will be collected will be publicly available for helping other scholars conduct research on cyber security.

**Author Contributions:** Samuel Chapman and Helge Janicke conducted requirement specification; Samuel Chapman implemented the virtual network; Samuel Chapman and Richard Smith generated and captured network traffic; Samuel Chapman and Leandros Maglaras performed system testing.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A. Technical Network Report

A general overview of the network configuration and daemons used has been provided. This appendix aims to provide a more detailed explanation of the network and how it was setup. A network created in Netkit-NG is made up of several different files and folders. *lab.conf* and *lab.dep* files are used by Netkit-NG for lab configuration detail. Netkit-NG requires each node to have a folder and a *.startup* file with the node's name, both of which can be empty, however, in most cases in this network are not. The node folders contain files and folders the user would like in the specific node, for example if the file *test.txt* was required for *pc1* this would be placed in its folder. The *.startup* file contains a bash script the user would like to run when a particular node is started, for example if the user required the *ls* command run when *pc1* was started this command would appear in *pc1.startup*.

### Appendix A.1. Lab.conf

The *lab.conf* file contains details regarding the hardware configuration for individual network nodes. It is possible to configure the amount of memory each node has as well as how many network interfaces they have and on which collision domain these interfaces sit. It was not deemed necessary to configure the amount of memory each node has as Netkit-NG nodes do not require a significant amount for the purpose intended. However if required for systems with scarce resources this can be set. The network interfaces are configured as such:

```
router0[0]=A
router0[1]=D
```

Router0 has been used as an example as this uses two network interfaces on different collision domains. The first section refers to the machine name. The section in square brackets refers to the network interface, in this example this refers to interface zero and one, when the node is started up the interface names will be prefixed with the word *eth*". The letter after the equals symbol refers to the collision domain the interface will sit on. Both the interface name and collision domain can be numbers, letters, or words, the letters and numbers chosen for this environment were chosen to make

remembering them simple. It should also be noted that if the collision domain is not already present in the network it will be created.

#### Appendix A.2. Lab.dep

Similar to the *lab.conf* file *lab.dep* provides configuration information for the network, however, this file is used to allow parallel start-up of Netkit-NG nodes, and to define an order regarding when each node should start-up. Simply providing a *lab.dep* file in the lab directory is enough to enable parallel start-up, which allows several nodes to start-up at once. The start-up order of nodes is defined using the following syntax:

```
httpServer: dhcpServer dnsServer
```

The node before the colon is the node the rule is intended for. After the colon are the nodes you would like started before the one in question. Simply put the rule states that *httpServer* will not start-up until *dhcpServer* and *dnsServer* have started. In the created network, all nodes must wait until the DHCP server has started and most must wait until at least one router and the DNS server has started. The DHCP server is started first to ensure all other nodes are able to retrieve IP addresses from it. The routers should be started before other nodes to ensure all are able to contact the DHCP server for these addresses. Because the network uses DDNS the DNS server is required to start-up before other nodes so that the DHCP server is able to update its records.

#### Appendix A.3. Collision Domain: TAP

The TAP collision domain contains the TAP node on the network. This node acts as a gateway to the internet however does not differ greatly from the router nodes on the network (see Appendix A.4), it uses commands placed in the *tap.startup* file to add routes to the other collision domains:

```
route add -net 192.168.0.0 netmask 255.255.255.0 gw 192.168.4.10 eth1
route add -net 192.168.1.0 netmask 255.255.255.0 gw 192.168.4.20 eth1
route add -net 192.168.2.0 netmask 255.255.255.0 gw 192.168.4.30 eth1
```

The first route command tells the node that any traffic with an IP address in the range 192.168.0.0–192.168.0.255 should be sent to the address 192.168.4.10. The other commands work in a similar manner. The *lab.conf* file for the TAP node is not like other nodes. The rule for the TAP machine is:

```
tap[0]=tap,192.168.44.130,192.168.44.150
tap[1]=D
```

Like a router this node has a network interface that sits on collision domain D. However, it has a second interface that sits on the TAP collision domain, the name *tap* is special in Netkit and allows the machine to communicate to the internet. Of the two IP addresses, the first is that of the host machine, and the second IP is the IP you would like to give to the tap nodes interface zero.

The tap node uses a slightly different *resolv.conf* file to the other machines. The *resolv.conf* file contains a list of DNS servers a node is able to contact, the tap node's *resolv.conf* looks like:

```
nameserver 192.168.0.5
nameserver 8.8.8.8
```



The first address is that of *dnsServer* (see Appendix A.6.6), which all other nodes use for DNS queries. The second is the address of Google's domain server, setting this allows the tap machine to make DNS queries outside of the network.

When the tap node starts it also installs the text-based web browser, Links. This is achieved using the command:

```
apt-get install links
```

The tap machine connects to the internet and therefore does not require anything more to install this program.

#### Appendix A.4. Collision Domain: D

As stated in Appendix A.3 this collision domain contains the routers for the network. The routers of this network are setup in the same way, the only differences being that the IP addresses differ between routers, as such only *router0* will be explained in detail. The *lab.conf* file assigns two interfaces to the router, 1 and 2. ), and the *lab.dep* file ensures the router cannot start until *dhcpServer* has started. *router0.startup* contains a number of commands for different purposes. The first command segment sets the IP addresses for both interfaces:

```
ifconfig eth0 192.168.0.55 up
ifconfig eth1 192.168.4.10 up
```

The IP address for the router was assigned statically to ensure it cannot be changed by the DHCP server, and the router always has the same IP. The next command segment assigns the router its routes:

```
route add -net 192.168.2.0 netmask 255.255.255.0 gw 192.168.4.30 eth1
route add -net 192.168.1.0 netmask 255.255.255.0 gw 192.168.4.20 eth1
route add default gw 192.168.4.40
```

These routes are assigned in the same way as the *tap* node, however a default route has been added. This was added so that any traffic intended for an IP not listed would go through the *tap* node and thus the internet. The final command segment consists of the setup and installation of ISC-DHCP-Relay (see Section 3.2). The command segment is as follows:

```
echo "isc-dhcp-relay isc-dhcp-relay/servers string dhcpServer.test.sc.com"
| debconf-set-selections
echo "isc-dhcp-relay isc-dhcp-relay/interfaces string" | debconf-set-selections
echo "isc-dhcp-relay isc-dhcp-relay/options string" | debconf-set-selections
apt-get update
apt-get -force-yes install isc-dhcp-relay -y
```

The first three *echo* statements contain setup information for the DHCP relay. This information is stored by the node and used to set options when ISC-DHCP-Relay is installed. The first command sets the DHCP server location to *dhcpServer.test.sc.com*, this is where DHCP requests will be forwarded when received by the router. The second leaves the interface field blank, leaving it blank allows the relay to detect which interface should forward DHCP requests. The final command is blank, to stop ISC-DHCP-Relay asking for additional options. These options are set to ensure the scripted installation requires no additional user setup. The *apt-get update* command is used to update where *apt-get* will look when installing new packages, this is used because the relevant files are installed from files local

to the node. The installation uses local files so an internet connection isn't required for the lab to function (see Appendix A.4.1).

A *resolv.conf* is also used by the routers and uses the following line:

```
domain test.sc.com
nameserver 192.168.0.5
```

The first line tells the node what the local domain is, this means that when the node attempts to communicate with another node on the network it does not have to use the fully qualified domain name (i.e., *remote1.test.sc.com*), it can instead use just the hostname of that node (i.e., *remote1*). The next line points the node to the local DNS server.

#### Appendix A.4.1. Installing from Local Files

Installing a program using local files requires several steps. The required files were downloaded, in the *.deb* format on a host machine, this included any dependencies. These files were then placed in the *etc/debPackages* folder on the node that required them. Once the files were downloaded the following command was run on the host machine.

```
dpkg-scanpackages . /dev/null | gzip -c9 > Packages.gz
```

This command creates a zip file with information regarding the *.deb* files.

The host's *etc/apt/sources.list* file was then changed so that it contained the single line:

```
deb file:///etc/debPackages /
```

This line tells the node to look in the */etc/debPackages* folder for packages. When the *apt-get install* command is run on the node it will now install without using an internet connection. This process has been repeated throughout the network as it is a fast and simple method of installing required applications, without an active connection to the internet.

#### Appendix A.5. Collision Domains: B & C

These subdomains do not contain any server daemons (see Section 3.3) and therefore require minimal setup, these are referred to as remote nodes. Each node has only one interface that sits on either collision domain B or C, the *lab.dep* file contains a variation of the line:

```
remote1: dhcpServer dnsServer router0 router1
```

The name of each node will change as will the name of the last router, depending on the machine and collision domain it will be placed on.

The node's *.startup* files are very similar, each contain the command:

```
dhclient eth0
```

This command makes the node query for a DHCP server when it starts up. If a DHCP server is found it will send and request several pieces of information for its eth0 interface, if one is not found the device will not obtain any information and the interface will not be able to communicate with the network which is why it is important that these devices start after both the dhcpServer and the routers.

The information sent and requested by the *dhclient* command is defined in *etc/dhcp/dhclient.conf*. The information sent is done so via the below segment:

```
send host-name = gethostname());
```

This will send the host name of the node to the DHCP server, which is used for DDNS (see Appendix A.6.6). The next segment in *dhclient.conf* is:

```
request subnet-mask, broadcast-address, time-offset, routers,
domain-name, domain-name-servers, domain-search, host-name,
dhcp6.name-servers, dhcp6.domain-search,
netbios-name-servers, netbios-scope, interface-mtu,
ntp-servers;
```

This will request various information from the DHCP server that will allow/aid the node in communicating with the network.

A default route is added on start-up, however these nodes will use the IP address of 192.168.1.11 or 192.168.2.22, this is collision domain dependant. The final section in the *.startup* files is the installation of the Links web browser, this is installed offline (see Appendix A.4.1).

The Pine program is installed on these nodes as an email client (see Section 3.3). This requires several files to work, one of which is the *root/mail/sent-mail* file, this contains information used by Pine for its mail folder and is used for no other purpose. The Pine program is held in *usr/local/bin/pine*, and the Pine manual page is held in *usr/share/man/man1/pine.1*. It also requires the hidden file *root/mail/.pinerc*, this file contains a large amount of configuration information for the Pine application. Pine requires a line to be added in this file. This line is shown below and is added at line 29.

```
inbox-path=imap.test.sc.com/user=remote1inbox
```

This line makes Pine look for the inbox of *remote1* on the IMAP server in the */usr* directory. The final section in the remote nodes *.startup* file is the following lines:

```
bash/hosthome/Documents/netkitLabs/replayScripts/masterReplayScript.sh & disown
```

The first line will start the bash script located at the specified file path, which will then run in the background. This script is used to replay network traffic (see Section 4). The second line is used because the nodes would hang during the start-up procedure, and this was found to stop this.

Like the routers (see Appendix A.4) the nodes in this collision domain use the *resolv.conf* file. The configuration adds a third line. This line is a second *nameserver* and is used to allow the node to query a non-local DNS server, which is capable of resolving internet requests.

#### Appendix A.6. Collision Domain: A

This subnet contains all of the servers required by the network. Each server is assigned one network interface by the *lab.conf* file, these interfaces all sit on collision domain A. The *lab.dep* file ensures that *dhcpServer* starts before the other servers, after this *dnsServer* will start, only then will the others start. When the other nodes on the network start, they will query the DHCP server using the *dhclient* command (see Appendix A.5). Each server node on this subnet will also have a default route of 192.168.0.55 which corresponds to router0. Another commonality between all servers is the use of an SSH daemon which allows remote command execution via a secure connection, this is started using:

```
/etc/init.d/ssh start
```

Once started a user is added, to allow access to the node via SSH. A user is added using the command:

```
useradd -d /home/toor -m -s /bin/bash toor
echo 'toor:root' | chpasswd
```

This adds a user called root and the appropriate directories. The user is then given the password *toor*. While not being a very secure password this is done for the purpose of this network scenario as password cracking will not be a focus of this environment.

Like the routers (see Appendix A.4), the server nodes in this collision domain use the *resolv.conf* file. The configuration used is the same as used by the router nodes.

#### Appendix A.6.1. FTP Server

All configuration for *ftpServer* is contained in its *.startup* file. Firstly, it will query the DHCP server for an IP address (see Appendix A.6.5). *ftpServer* will then add users (remote0–remote8), this is a simple operation but it is required to establish a connection with the *proftpd* daemon. Once the users have been added the FTP daemon is started:

```
/etc/init.d/proftpd start
```

#### Appendix A.6.2. SFTP Server

The *sftpServer.startup* file is exactly the same as *ftpServer.startup*, however *sftpServer* contains the file */etc/proftpd/conf.d/sftp.conf*. This file instructs the *proftpd* daemon to use the SFTP protocol rather than standard FTP. It also states which port to use and that it should use a public key. This is the only difference between the two server nodes.

#### Appendix A.6.3. HTTP Server

In addition to the commands discussed in Appendix A.6 *httpServer.startup* contains the following lines:

```
#enable ssl site
a2ensite default-ssl
a2enmod ssl
#Start apache server
/etc/init.d/apache2 start
```

The first two lines enable the apache2 SSL modules. These modules allow apache2 to use HTTPS as well as standard HTTP. The last line starts the apache server daemon.

HTTP works on this server without any additional configuration. HTTP webpages are stored in the *httpServer/var/www* directory, two very basic webpages were placed in this directory, *corp\_home.html* and *an\_page\_of\_VSTART.html*, these pages simply provide proof that *httpServer* works and can be accessed via the Links program using the address *http://httpServer.test.sc.com/* and then the name of the file.

Implementing HTTPS requires more configuration. In order for SSL to work a certificate must be generated, this certificate is used to confirm the identity of the server to the client node, for the purposes of this network it was decided to sign this certificate myself to ensure the network does not directly require an internet connection to function. To self-sign a certificate a tutorial was followed provided by Zahn and Zahn (n.d.), this goes through the entire process. Once the correct files were generated a small number of changes were made. These changes took place in the *httpServer/etc/apache2/sites-available* file.

The document root, on line five, was changed from `/var/www` to `/var/ssl` as was the directory, on line ten. These two changes ensure that `httpServer` is capable of using both the HTTP and HTTPS protocols.

Another change made to the same file was the addition of the following lines:

```
SSLCertificateFile /etc/apache2/ssl/https.crt
SSLCertificateKeyFile /etc/apache2/ssl/https.key
```

These lines enforce the directory where apache will look for the `.crt` and `.key` files, the names of these files were also changed, these changes helped to make the files easier to locate if further changes were required.

As with the standard HTTP webpages a very basic page was created in the `httpServer/var/ssl` directory. When `httpServer` is started the page that implements HTTPS can be reached using the address `https://httpServer.test.sc.com/ssl.html`.

#### Appendix A.6.4. Email Server

The `emailServer.startup` file is similar to the other nodes on this collision domain, however like the FTP server it adds several users. The email server also includes some installs that have been scripted (see Appendix A.4.1). The commands used to script the installation are:

```
apt-get update
apt-get -force-yes install exim4 -y
chmod 644 /etc/exim4/exim4.conf
chown root:root /etc/exim4/exim4.conf
/etc/init.d/exim4 start
/etc/init.d/opensbsd-inetd start
```

In addition to the installation commands are the `chmod` and `chown` commands. These are used to change the permissions of the `exim4.conf` file, the permissions had to be changed as there were errors encountered when starting `exim4` initially which were diagnosed as permission related. Similar to the nodes in collision domains B and C this node contains the Pine application as well as the `imapd` and `pop3d` daemons (see Section 3.4.4).

The `exim4` daemon is configured using the `emailServer/etc/exim4/exim4.conf` and `emailServer/etc/exim4/update-exim4.conf.conf` files. The `exim4.conf` file should not be changed directly by the user, which is why the `update-exim4.conf.conf` file is required. `update-exim4.conf.conf` contains information about the network so `exim4` can function properly and only receive traffic from a certain range of IP addresses.

Like the standard nodes (see Appendix A.5) the email server uses Pine. It is used here to allow emails to be stored in a remote inbox. The only change required in the hidden file `.pinerc` is following line eighty where the following is added:

```
customized-hdrs=Reply-To: remote1@test.sc.com
```

This allows the `Reply-To` field to be automatically filled in, while not strictly necessary this makes it easier for the nodes to communicate via email.

#### Appendix A.6.5. DHCP Server

`dhcpServer` uses most of the commands used by the other server nodes (see Appendix A.4), however it does not use `dhclient`, instead it uses the following:

```
ifconfig eth0 hw ether 00:00:00:00:00:03 up
```

```
ifconfig eth0 192.168.0.7 up
```

These commands set a hardware address for eth0, and subsequently assign an IP address for eth0. A hardware address is assigned to ensure the DHCP server will not attempt to give its assigned address to another node. This is done by defining a static lease in the *dhcpcd.conf* file.

After eth0 has been configured the following commands are run:

```
apt-get update
apt-get -force-yes install isc-dhcp-server -y
cp /etc/dhcp/myConf/dhcpd.conf /etc/dhcp
service isc-dhcp-server restart
```

This installs the server daemon (see Section 3.4.5), a configuration file is then copied and the server daemon is restarted. The configuration file is copied as there is a *dhcpcd.conf* file already in the *dhcpServer/etc/dhcp* directory. Under normal circumstances this would be replaced on start-up if the file was present in the lab directory structure, however when ISC-DHCP-Server is installed this file is replaced, because of this the file has to be copied into the directory following installation of the server daemon. After copying the new configuration file the server daemon has to be restarted to ensure the daemon uses the new file. Following this is the installation and configuration of the ISC-DHCP-Relay daemon, this is the same as the routers.

ISC-DHCP-Server uses two files for its configuration, *dhcpServer/etc/default/isc-dhcp-server* and *dhcpServer/etc/dhcp/myconf/dhcpd.conf*. The first file is used, in this lab, to configure what interface the DHCP server should serve requests from, the following line was used:

```
INTERFACES="eth0"
```

This line specifies eth0, this is the only interface *dhcpServer* has.

The *dhcpcd.conf* file is used to specify how the DHCP server allocates addresses and communicates with the network, and in this case how it updates the DNS server with the addresses it distributes. It should be noted that this file is divided into sections, which are denoted by curly braces and a semi colon . In order to update the DNS server the following lines are used:

```
ddns-update-style interim;
update-static-leases on;
key ddns.key
algorithm hmac-md5;
secret NvoL1uVTOeWOmWQPbip1Rg==;
;
zone test.sc.com
primary 192.168.0.5;
key ddns.key;
zone reverse.in-addr.arpa
primary 192.168.0.5;
key ddns.key;
option domain-name "test.sc.com";
option domain-name-servers dnsServer.test.sc.com;
```

The first line enables the DHCP server to send updates to the DNS server, and the second ensures that it will send updates about any static IP addresses that are assigned later in the file. The next section, starting *key ddns.key*, defines a key that should be sent by the DHCP server daemon when it updates



the DNS server, this is used as a method of authentication so the DNS server does not accept updates unless the key is provided. The two sections that begin *zone* denote the two DNS zones that *dhcpServer* can update (see Appendix A.6.6), you will also notice that the aforementioned key is referred to in these sections, this is provided when the DHCP server updates the zone. The final lines inform the DHCP server of the domain name and the location of the DNS server it will be updating.

The next four lines in *dhcpd.conf* are as follows:

```
default-lease-time 600;
max-lease-time 7200;
authoritative;
log-facility local7;
```

The first two dictate the lease time allowed on IP address allocations, this is measured in seconds. When a device queries the DHCP server for an address it will be given this information and should send another request when this time is over, if it does not the DHCP server is able to give the allocated address to another device. The next line states that this particular DHCP server is authoritative for the network. The final line makes the DHCP server create logs using the local7 utility, this is used for logging various information, which can be used for debugging errors.

Following the aforementioned sections are three sections that define subnets and the addresses the DHCP server can assign within those subnets. The sections are of a similar structure and look like:

```
subnet 192.168.1.0 netmask 255.255.255.0
range 192.168.1.50 192.168.1.250;
default-lease-time 600;
max-lease-time 7200;
option routers 192.168.1.11;
option ip-forwarding off;
option domain-name-servers 192.168.0.5;
option broadcast-address 192.168.1.255;
option subnet-mask 255.255.255.0;
```

The first line defines the subnet as 192.168.1.0 with a netmask of 255.255.255.0. The netmask is the range of IP addresses available to the subnet in this case the available addresses are in the range 192.168.1.0–192.168.1.255. The next line sets the range of IP addresses the DHCP server is actually able to assign. The two lines following this reiterate the lease times assigned earlier. The fifth line simply states the IP address of the router for a particular subnet. Following that line is one that turns off ip-forwarding, this is normally turned on if there is a router being assigned an address. The final three lines are information regarding the network, and is sent to the client node.

The final section in *dhcpd.conf* defines any client nodes which should have fixed IP addresses:

```
host dnsServer
hardware ethernet 00:00:00:00:00:01;
fixed-address 192.168.0.5;
```

The first line simply denotes the node the rule applies to, this name does not have to be the name of the node, however, doing so makes the rule easier to read/understand. The second line is the hardware address of the node the rule applies to, this information is supplied by each node when they query the DHCP server. The final line is the IP address that will be given to the specified device. Fixed addresses are used for the DHCP and DNS server nodes, because if they were to change address the network would not function properly.

### Appendix A.6.6. DNS Server

Like the other server nodes on this collision domain this node use the same *.startup* file, with the addition of the following line:

```
/etc/init.d/bind9 start
```

This line starts the Bind9 server daemon which is used for DNS. A key difference between *dnsServer* and other nodes on this collision domain is the omission of the line referencing the *nameserver* in the *resolv.conf* file, this is because it is the DNS server and does not need to reference itself. This server node uses the Bind9 server daemon to implement the DNS protocol (see Section 3.4.6), the configuration files for this are contained within the directory *dnsServer/etc/bind/*.

The *named.conf.local* file is read by Bind9 for information regarding domains it controls. This file contains the *ddns.key*, this is the same as found in the DHCP server and is what the DHCP servers key is compared against. The rest of the file contains information about the various zones under the control of this DNS server, there are two zones in this particular configuration file. The first zone is defined as:

```
zone "test.sc.com"
type master;
file "/etc/bind/db.com.sc.test";
allow-update key ddns.key; ; ;
```

The first section defines a new zone, *test.sc.com*, and states that this DNS server is the *master* of that zone, this means that it is the authoritative name server for the particular zone. Following this is a line with the location of a file, this file will be explained in more detail in the following paragraph. The final line states that the server should allow updates to this zone if the key in *ddns.key* is provided. The second zone is very similar to this one however it is called *reverse.in-addr.arpa* and the related file is located at */etc/bind/zones/reverse.db.com.sc.test*, the contents of which will be explained in a subsequent paragraph. This zone is used for reverse lookups, these are like normal DNS lookups (i.e., using domain names to find IP addresses) except the opposite way around.

The file *dnsServer/etc/bind/db.com.sc.test* contains information about the *test.sc.com* zone. The first section of this files relates to refresh rates of the DNS records, and when they expire. The next section is where records are denoted, records are used by the DNS server to locate a node based on its domain name. Typically, a record looks like:

```
dhcpServer.test.sc.com. IN A 192.168.0.7
```

This denotes that the domain name *dhcpServer.test.sc.com*. refers to the node with the IP address at the end of the line. There are few records in this file given that most of the network's nodes are given records using Dynamic DNS. Some exceptions exist and this is mostly for email protocol as emails are not addressed to a specific node's host name, the lines used for email are shown below.

```
mail.test.sc.com. IN A 192.168.0.8
imap.test.sc.com. IN A 192.168.0.8
pop.test.sc.com. IN A 192.168.0.8
```

These all relate to the location of the email server.

The second file used by Bind9 was */etc/bind/zones/reverse.db.com.sc.test*. The contents of which look very similar to that described above, however the lines denoting domain names are in reverse:

7.0.168.192 IN PTR dhcpServer.test.sc.com.

This allows the DNS server to resolve queries that look for the domain name of a given IP address.

## Appendix B. System Testing

Extensive have been implemented and the results are displayed in this appendix. These tests are used to test that the functions work as described.

### Appendix B.1. Function 1 Tests

The tests conducted in this set of tests regard the nodes on both the initial network and a network built to the same specification using the network creation program (see Section 5). Tests 1–3 use the remote nodes (see Section 3.3), these nodes are used as these nodes are the only ones that require use of the tested services. For each of the tests five nodes are selected, this number is chosen as it represents over 25% of the nodes in the initial network which is felt to be a representation of the network, it also allows a node from each collision domain to be tested it test three and four.

**Table A1.** Tests results for Function 1.

Name		Function 1 Test Results	
Description	Expected Results	Status (Pass/ Fail)	Notes
<b>Test 1</b>			
Send five emails using the email client on remote nodes.	The remote nodes will be able to send emails without error.	Pass	Initially this test failed. This was due to configuration changes to the network to implement DDNS (see Section 3.4.6), this was resolved by ensuring the email server updated the DNS server properly.
Check emails are received in the remote nodes' inbox.	The remote nodes are able to view emails sent to them.	Pass	
<b>Test 2</b>			
Send the FTP server 5 files from five nodes on the network.	The remote nodes are able to send files to the FTP server.	Pass	
Use the <i>get</i> command to download files to five remote nodes from the FTP server.	The files will be downloaded by the host nodes.	Pass	
Send the SFTP server 5 files from five nodes on the network.	The remote nodes are able to send files to the SFTP server.	Pass	
Use the <i>get</i> command to download files to five remote nodes from the SFTP server.	The files will be downloaded by the host node.	Pass	
Send the FTPS server 5 files from five nodes on the network.	N/A	Fail	This test is not possible as there is not an FTPS server on this network (see Section 3.4.7)
Use the <i>get</i> command to download files to five remote nodes from the FTPS server.	N/A	Fail	This test is not possible as there is not an FTPS server on this network (see Section 3.4.7)
<b>Test 3</b>			
Use the links web browser on five different remote nodes to go to the two HTTP webpages provided by the HTTP server.	The remote nodes are able to view the webpages.	Pass	
Use the links web browser on five of the remote nodes to go to the HTTPS webpage on the HTTP server.	The remote nodes are able to view the webpages.	Pass	
<b>Test 4</b>			
Use the ping tool on five different nodes to ping five other nodes via their DNS names.	All tested nodes will reply when pinged via their DNS names.	Pass	
Use the ifconfig tool on five different nodes to ensure they are assigned IP addresses from the DHCP server.	All tested nodes will have valid IP addresses.	Pass	

The following tests were conducted on a network created using the network creation program (see Section 5) this was designed to the same specification as the initial network (see Section 3).

**Table A2.** Tests results for Function 1 using the network creation program.

Name		Function 1 Test Results		
Description	Expected Results	Status (Pass/ Fail)	Notes	
<b>Test 1</b>				
Send five emails using the email client on remote nodes.	The remote nodes will be able to send emails without error.	Pass	Initially this test failed. This was due to a misconfiguration when the program edits the <i>.pinerc</i> on the <i>standardNode</i> nodes. The line edited cannot contain capital letters, if it does the email client will be unable to send emails	
Check emails are received in the remote nodes' inbox.	The remote nodes are able to view emails sent to them.	Pass		
<b>Test 2</b>				
Send the FTP server 5 files from five nodes on the network.	The remote nodes are able to send files to the FTP server.	Pass		
Use the <i>get</i> command to download files to five remote nodes from the FTP server.	The remote nodes are able to download files from the FTP server.	Pass		
Send the SFTP server 5 files from five nodes on the network.	The remote nodes are able to send files to the SFTP server.	Pass		
Use the <i>get</i> command to download files to five remote nodes from the SFTP server.	The remote nodes are able to download files from the SFTP server.	Pass		
Send the FTPS server 5 files from five nodes on the network.	N/A	Fail	This test is not possible as there is not an FTPS server on this network (see Section 3.4.7)	
Use the <i>get</i> command to download files to five remote nodes from the FTPS server.	N/A	Fail	This test is not possible as there is not an FTPS server on this network (see Section 3.4.7)	
<b>Test 3</b>				
Use the links web browser on five different remote nodes to go to the two HTTP webpages provided by the HTTP server.	The remote nodes are able to view the webpages.	Pass		
Use the links web browser on five of the remote nodes to go to the HTTPS webpage on the HTTP server.	The remote nodes are able to view the webpages.	Pass		
<b>Test 4</b>				
Use the ping tool on five different nodes to ping five other nodes via their DNS names.	All tested nodes will reply when pinged via their DNS names.	Pass		
Use the ifconfig tool on five different nodes to ensure they are assigned IP addresses from the DHCP server.	All tested nodes will have valid IP addresses.	Pass		

### Appendix B.2. Function 2 Tests

The following tests were conducted on the network creation tool (see Section 5).

**Table A3.** Tests results for Function 2.

Name	Function 2 Test Results		
Description	Expected Results	Status (Pass/ Fail)	Notes
<b>Test 1</b>			
Run through the network creation program and ensure values can be input where asked. This will be done 5 times, to ensure errors do not occur inconsistently	All areas that ask for user input will allow it.	Pass	
<b>Test 2</b>			
Run through the program and input special characters and letters where the input expects a number. This will be done using ten characters that are not numerical, and it will be a mix of special and alphabetical characters. Ten have been selected as it provides a large enough sample size to test the different available characters.	When entered special characters return an error message.	Pass	
Run through the program and input characters that are not Y, y, N, or n where these are the expected response. This will be done using ten characters and it will be a mix of special and alphanumeric characters. Ten have been selected as it provides a large enough sample size to test the different available characters.	When entered, characters that are not Y, y, N, or n will return an error message.	Pass	
Run through the program and where it expects numbers 0–254, enter the numbers –1 and 255. This will test that the boundaries implemented are effective.	When entered the numbers should return an error message.	Pass	

*Appendix B.3. Function 3 Tests*

The following tests were conducted on the output of the network creation tool (see Section 5).

**Table A4.** Tests results for Function 3.

Name	Function 3 Test Results		
Description	Expected Results	Status (Pass/ Fail)	Notes
<b>Test 1</b>			
Run through the program and ensure files and folders are created. The program will be used five times, each time different values will be entered for the different fields.	Files and folders will be output to the <i>createdLab</i> folder.	Pass	
<b>Test 2</b>			
Run through the program and ensure files and folders are created with the names expected, given the values input. The program will be used five times, each time different values will be entered for the different fields.	The expected files and folders will be output to the <i>createdLab</i> folder.	Pass	

**Table A4.** Cont.

Name	Function 3 Test Results		
Description	Expected Results	Status (Pass/ Fail)	Notes
<b>Test 3</b>			
Run through the program and ensure files and folders are created with the names expected, given the values input. The program will be used five times, each time different values will be entered for the different fields.	The expected files and folders will be output to the <i>createdLab</i> folder.	Pass	
Run through the program and ensure the files and folders created have expected names, given the values input. The program will be used five times, each time different values will be entered for the different fields, and ten different files will be checked each time, this will ensure a large number of the files output are checked for errors or discrepancies.	The expected files and folders will be output to the <i>createdLab</i> folder with the expected contents.	Pass	

*Appendix B.4. Function 4 Tests*

The tests conducted in this section are used to test whether network traffic is replayed successfully. (see Section 4) These test have been conducted on both the initial lab and a lab created using the network creation program to the same specification.

**Table A5.** Tests results for Function 4.

Name	Function 4 Test Results		
Description	Expected Results	Status (Pass/ Fail)	Notes
<b>Test 1</b>			
Start the network lab and commence traffic captures on the router nodes, using the tcpdump tool. The router nodes will be used as these will see the most network traffic on the network. Leave these traffic captures running for ten minutes, this should be enough time for the traffic to be generated.	The traffic captures will contain data that represents the expected network traffic.	Pass	

The following tests were conducted on a network created by the network creation program (see Section 5). This network was created to the specification of the initial network. (see Section 3).

**Table A6.** Tests results for Function 4 using the network creation program.

Name	Function 4 Test Results		
Description	Expected Results	Status (Pass/ Fail)	Notes
<b>Test 1</b>			
Start the network lab and commence traffic captures on the router nodes, using the tcpdump tool. The router nodes will be used as these will see the most network traffic on the network. Leave these traffic captures running for ten minutes, this should be enough time for the traffic to be generated.	The traffic captures will contain data that represents the expected network traffic.	Pass	



## References

1. Cook, A.; Smith, R.; Maglaras, L.; Janicke, H. Measuring the risk of cyber attack in industrial control systems. In Proceedings of the 4th International Symposium for ICS & SCADA Cyber Security Research 2016, Belfast, UK, 23–25 August 2016.
2. Hallaq, B.; Nicholson, A.; Smith, R.; Maglaras, L.; Janicke, H.; Jones, K. CYRAN: A Hybrid Cyber Range for Testing. *Secur. Solut. Appl. Cryptogr. Smart Grid Commun.* **2016**, doi:10.4018/978-1-5225-1829-7.ch012.
3. Cyber Security Training Delivery Options SANS Institute. Training Delivery Options. Available online: <https://uk.sans.org/training-delivery-options/> (accessed on 20 July 2017).
4. Hosted Penetration Testing Virtual Labs by Offensive Security Offensive Security. Available online: <https://www.offensive-security.com/offensive-security-solutions/virtual-penetration-testing-labs/> (accessed on 20 July 2017).
5. Pham, C.; Tang, D.; Chinen, K.I.; Beuran, R. CyRIS: A cyber range instantiation system for facilitating security training. In Proceedings of the Seventh Symposium on Information and Communication Technology, Ho Chi Minh City, Vietnam, 8–9 December 2016; pp. 251–258.
6. The Ultimate Cyber Proving Ground for Critical Systems and Technologies Raytheon Cyber. Available online: <http://www.raytheoncyber.com/capabilities/range/> (accessed on 20 July 2017).
7. Cyberbit Range. Available online: <https://www.cyberbit.com/solutions/cyber-range/> (accessed on 20 July 2017).
8. Mullins, B.E.; Lacey, T.H.; Mills, R.F.; Trechter, J.E.; Bass, S.D. How the cyber defense exercise shaped an information-assurance curriculum. *IEEE Secur. Priv.* **2007**, *5*, doi:10.1109/MSP.2007.111.
9. Adams, W.J.; Gavas, E.; Lacey, T.H.; Leblanc, S.P. Collective Views of the NSA/CSS Cyber Defense Exercise on Curricula and Learning Objectives. In Proceedings of the 2nd Conference on Cyber Security Experimentation and Test (CSET'09), Montreal, QC, Canada, 10–14 August 2009.
10. Johnson, C.; Badger, L.; Waltermire, D.; Snyder, J.; Skorupka, C. Guide to cyber threat information sharing. *NIST Spec. Publ.* **2016**, *800*, 150.
11. Cruz, B.; Gupta, D.; Kapoor, A.; Haifei, L.; McLean, D.; Moreno, F. *McAfee Labs Threats Report*; McAfee Inc.: Santa Clara, CA, USA, 2014. Available online: <http://www.mcafee.com/us/resources/reports/rp-quarterlythreat-q1-2014.pdf> (accessed on 20 July 2017)
12. Hansman, S.; Hunt, R. A taxonomy of network and computer attacks. *Comput. Secur.* **2005**, *24*, 31–43.
13. Chapman, I.M.; Leblanc, S.P.; Partington, A. Taxonomy of cyber attacks and simulation of their effects. In Proceedings of the Military Modeling & Simulation Symposium, Boston, MA, USA, 3–7 April 2011; pp. 73–80.
14. Janbeglou, M.; Zamani, M.; Ibrahim, S. Redirecting network traffic toward a fake DNS server on a LAN. In Proceedings of the 3rd IEEE International Conference on Computer Science and Information Technology (ICCSIT), Chengdu, China, 9–11 July 2010; Volume 2, pp. 429–433.
15. McAfee Labs Threats Report: September 2016 McAfee. Available online: <https://www.mcafee.com/us/security-awareness/articles/mcafee-labs-threats-report-sep-2016.aspx> (accessed on 30 March 2017).
16. Fuertes, W.M.; de Vergara, J.E.L. A quantitative comparison of virtual network environments based on performance measurements. In Proceedings of the 14th HP Software University Association Workshop, Munich, Germany, 2 July 2007; pp. 8–11.
17. Wolny, W.; Szoltysik, M. Overview of Existing Computer Network Environments Virtualization for Computer Network Learning. *Stud. Ekon.* **2014**, *188*, 250–264.
18. Pizzonia, M.; Rimondini, M. Netkit: Easy emulation of complex networks on inexpensive hardware. In Proceedings of the 4th International Conference on Testbeds and Research Infrastructures for the Development of Networks & Communities, ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), Innsbruck, Austria, 18–20 March 2008; p. 7.
19. Ruiz-Martinez, A.; Pereniguez-Garcia, F.; Marin-Lopez, R.; Ruiz-Martínez, P.M.; Skarmeta-Gomez, A.F. Teaching advanced concepts in computer networks: Vnuml-um virtualization tool. *IEEE Trans. Learn. Technol.* **2013**, *6*, 85–96.
20. Márquez, F.G.; Cambronero, D.F. Distributed Virtualization Scenarios Using VNUML. In Proceedings of the 1st International DMTF Academic Alliance Workshop on Systems and Virtualization Management: Standards and New Technologies, Toulouse, France, 23 October 2007; Volume 2007.

21. Li, P. Selecting and using virtualization solutions: Our experiences with VMware and VirtualBox. *J. Comput. Sci. Coll.* **2010**, *25*, 11–17.
22. Simulator, G.N. GNS 3. Available online: <http://paginapessoal.utfpr.edu.br/lincoln/palestras/ftsl/index.html/ApresentacaoGNS3.pdf> (accessed on 9 August 2017).
23. Anuzelli, G.; Files, N.; Emulation, P.; Optimizations, M.U.; Emulated, H.C. Dynamips/Dynagen Tutorial. Available online: <http://dynagen.org/tutorial.htm> (accessed on 10 July 2017).
24. Kneale, B.; De Horta, A.Y.; Box, I. Velnet: Virtual environment for learning networking. In Proceedings of the Sixth Australasian Conference on Computing Education-Volume 30, Dunedin, New Zealand, 10 January 2004; pp. 161–168.
25. Vahdat, A.; Yocum, K.; Walsh, K.; Mahadevan, P.; Kostić, D.; Chase, J.; Becker, D. Scalability and accuracy in a large-scale network emulator. *ACM SIGOPS Oper. Syst. Rev.* **2002**, *36*, 271–284.
26. Hashimoto, M. *Vagrant: Up and Running: Create and Manage Virtualized Development Environments*; O'Reilly Media, Inc.: Sebastopol, CA, USA, 2013.
27. Sundararaj, A.I.; Gupta, A.; Dinda, P.A. Dynamic topology adaptation of virtual networks of virtual machines. In Proceedings of the 7th workshop on Workshop on Languages, Compilers, and Run-Time Support for Scalable Systems, Houston, TX, USA, 21–23 October 2004; pp. 1–8.
28. Hillestad, O.I.; Libak, B.; Perkis, A. Performance evaluation of multimedia services over ip networks. In Proceedings of the IEEE International Conference on Multimedia and Expo (ICME), Amsterdam, The Netherlands, 6 July 2005; pp. 1464–1467.
29. Dean, T. *Network+ Guide to Networks*; Cengage Learning: Boston, MA, USA, 2012.
30. Duck, M.; Read, R. *Data Communications and Computer Networks: For Computer Scientists and Engineers*; Pearson Education: London, UK, 2003.
31. Kurose, J.F.; Ross, K.W. *Computer Networking: A Top-down Approach*; Addison-Wesley Reading: Boston, MA, USA, 2010; Volume 5.
32. White, C. *Data Communications and Computer Networks: A Business User's Approach*; Cengage Learning: Boston, MA, USA, 2015.
33. Ford-Hutchinson, P. Securing FTP with TLS. Available online: <https://tools.ietf.org/html/rfc4217> (accessed on 9 August 2017).
34. Maglaras, L.A.; Jiang, J.; Cruz, T.J. Combining ensemble methods and social network metrics for improving accuracy of OCSVM on intrusion detection in SCADA systems. *J. Inf. Secur. Appl.* **2016**, *30*, 15–26.



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).