DATA CENTER VIRTUALIZATION FRAMEWORK IN ETHIOPIAN HIGHER LEARNING INSTITUTIONS: THE CASE OF HAWASSA UNIVERSITY

ABDALLAH HUSEN

HAWASSA UNIVERSITY, HAWASSA, ETHIOPIA

November, 2022

DATA CENTER VIRTUALIZATION FRAMEWORK IN ETHIOPIAN HIGHER LEARNING INSTITUTIONS: THE CASE OF HAWASSA UNIVERSITY

ABDALLAH HUSEN

MAJOR ADVISOR: TESFAYE BAYU (PhD)

CO-ADVISOR: ZEMENU HAILE (MSc)

A THESIS SUBMITTED TO THE DEPARTMENT OF COMPUTER SCIENCE

HAWASSA UNIVERSITY

INSTITUTE OF TECHNOLOGY

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE IN COMPUTER SCIENCE

HAWASSA, ETHIOPIA

November, 2022

APPROVAL SHEET-I

This is to certify that the thesis entitled "**Data Center Virtualization Framework In Ethiopian Higher Learning Institutions: The Case Of Hawassa University**" submitted in partial fulfillment of the requirements for the degree of Master's with specialization in Computer Science, the Graduate Program of the Department/School of Informatics, and has been carried out by Abdallah Husen. Therefore we recommend that the student has fulfilled the requirements and hence hereby can submit the thesis to the department.

________________________	________________	________________
Name of major advisor	Signature	Date
________________________	________________	________________
Name of co-advisor	Signature	Date

APPROVAL SHEET-II

We, the undersigned, members of the Board of Examiners of the final open defense by Abdallah Husen have read and evaluated his thesis entitled "**Data Center Virtualization Framework In Ethiopian Higher Learning Institutions: The Case Of Hawassa University**", and examined the candidate. This is, therefore, to certify that the thesis has been accepted in partial fulfillment of the requirements for the degree.

______________________ ________________ ____________

Name of Major Advisor Signature Date

______________________ ________________ ____________

Name of Internal Examiner-I Signature Date

______________________ ________________ ____________

Name of Internal Examiner-II Signature Date

Sileshi D. (PHD) ___ 26/01/2023

Name of External examiner Signature Date

______________________ ________________ ____________

SGS Approval Signature Date

STATEMENT OF THE AUTHOUR

I hereby certify that I am the sole author of this thesis. All the used materials, references to the literature and the work of others have been referred to. This thesis has not been presented for examination anywhere else.

Name: ……………………………...……………………. Signature: ……….……………

Place: Institute of Technology, Hawassa University, Hawassa

Date of Submission: ………………………………………….

Dedication

I dedicate my study to my Mother, may Allah mercy be upon her, may her memories have given me the strength to continue.

Acknowledgments

First and foremost, I would like to praise and thank Allah , who have granted me countless blessings, knowledge, and opportunities so that I have finally been able to accomplish the thesis.

The success of this thesis depends largely on the support and guidance of many individuals. I would want to take this opportunity to thank everyone who has contributed to the successful completion of my thesis. First, my special and priceless thanks go to my advisor Dr Tesfaye Bayu for his endless support and kind encouragement during the entire work. His openness, guidance, and support taught me what kindness truly is. Equally and importantly I am grateful to appreciate Zemanu Haile for the support and encouragements. Also I'd want to extend my heartfelt gratitude to Dr Amir Teshome for his valuable, kind support and willingness to provide me the necessary input for this thesis and throughout my MSc program. I would also like to thank HU ICT department staff for giving me this golden opportunity to pursue my MSc.

Abstract

Ethiopian higher learning institutions(HLI) data center runs physical data center in which a one application for one server architecture is exercised. Such architecture leads to under-utilization and wastage of resources. Using virtualization technology enables lowering the size of the infrastructure, thus resulting in a huge savings on energy and other resources including management time. Cost reduction , energy efficiency and the reduction of a company's carbon footprint are some of the significant benefits of using virtualization but virtualization technology is a long way from Plug-and-Play to unlock those benefits effectively, an information technology (IT) expert requires the appropriate set of model to be followed. In most cases, organizations have enough resources to move to virtualization without requiring additional budget.

Certain researches have been conducted on virtualization especially client virtualization, application virtualization and network virtualization. But, little attention was given for server virtualization. The objective of this research is to explore the current traditional infrastructure practice and to propose data center virtualization framework using design science research methodology (DSRM). The goal is to look for option(s) that brings a preferred solution for University data centers that increases service availability and utilization of hardware. In this work a virtualization framework is proposed, and evaluated using the same services that are currently used in the university. Experiments are carried out to check and compare the resource utilization of both the physical machines and Virtual machines(VMs). By using both the experiment and the analysis result, well-matched virtualization framework was developed. The developed frameworks was tested on three services. Before virtualization each of these services were running in separate physical machine. In our experiment, by consolidating these services using virtualization, the study showed that it is possible to provide the same service using only a single server. The resource utilization of that single server increased as follows: Central processing system (CPU) usage increased from 4% to 17%, physical memory usage increased from 15% of 16GB to 62% of 16GB, and capacity of hard disk space in use was 67.4% up from 32%. The Universities data center will use the proposed well-matched framework to enable them increase their levels of utilization of the servers.

Keywords: Data center, virtualization, server virtualization , Design science.

List of Abbreviations

BC/DR	Business Continuity/Disaster Recovery
CDCs	Corporate data centers
CO2	Carbon dioxide
CPU	Central Processing Unit
DMZ	Demilitarized Zone
DS	Design Science
DSRM	Design Science Research Methodology
GUI	Graphical User Interfaces
HU	Hawassa University
HPC	High Performance Computing
Hyper-V	Microsoft Hypervisor
ICT	Information Communication Technology
IaaS	Infrastructure as a Service
IBM	International Business Machine
IDCs	Internet data centers
IS	Information systems
ISP	Internet Service Provider
IT	Information Technology
NIC	Network Interface Card
OS	Oprating System
OSs	Oprating Systems
PaaS	Platform as a Service
QDA	Qualitative Data Analysis
QoS	Quality of Service
RAM	Random Access Memory
SaaS	Software as a Service
SAN	Storage Area Networks
TCO	Total Cost of Ownership
UPS	Uninterruptible Power Supplies
VMM	Virtual machine monitor
Vms	virtual machines
vSwitch	Virtual Switch

Table of Contents

List of Figures

List of Tables

1. INTRODUCTION

1.1 Background

1.1.1 Data Center

A data center is a special facility conceived to house, manage, and support computing resources that are considered critical for one or more organizations. It provides the capabilities of centralized repositories for storage, management, networking, and dissemination of data. A particularly complex structure, a typical data center encompasses special building structures, power backup structures, cooling systems, special-purpose rooms (entrance and telecommunications, for example), equipment cabinets, structured cabling, network devices, storage systems, servers, mainframes, application software, physical security systems, monitoring centers, and many other support systems. All these resources and their interaction are (locally or remotely) managed by specialized personnel.

Higher education institutions employ a variety of data center types to suit their needs, though small data centers, enterprise data centers, and high-performance computing (HPC) are most common. The standard data center contains at least:

- **Power Room**: electricity room which provides power for the systems. It is one of the largest cost drivers for data centers. The worldwide energy consumption of data centers has increased dramatically, now accounting for about 1.3% of the world's electricity usage [1]
- **Server Room**: a room that devoted to store servers and related equipment's to manage air conditioning.
- **Network Operations Center (NOC)**: a control room for personnel to monitor the functioning of the equipment's. In addition to overseeing a network operations the control room may house many video screens to monitor the equipment's in case of any physical incident.

1.1.2 Types of Data Center

Data centers fall into two major categories including corporate data centers (CDCs) and Internet data centers (IDCs). Corporate data centers are owned and operated by private corporations, institutions or government agencies. Their prime purpose includes supporting data processing and Web-oriented services for their own organizations, business partners and customers. Equipment, applications, support and maintenance for the data center are typically supported by in-house IT departments or contract partners.

Internet data centers are primarily owned and operated by traditional telecommunications, unregulated competitive service providers or other types of commercial operators. Each operator, however, involves similar goals—to provide outsourced IT services accessed through Internet connectivity. Their business is to provide a menu of services to their clients. These services may include (but are not limited to) wide-area communications, Internet access, web or application hosting, colocation, managed servers, storage networks, content distribution and load sharing with new variations appearing almost daily. [2]

Data centers, On the other hand, can be classified into four categories based on availability (yearly uptime) and redundancy of electrical paths for power, IT equipment, and cooling systems. [2]

- **Tier 1:** Is the most fundamental type of data center tier, having no redundant infrastructure parts. It is expected to be up and running 99.671% of the time (annual downtime of 28.8 hours).
- **Tier 2:** A data center contains redundant components but a single source of power and cooling supplies. It is expected to be operational 99.741 percent of the time (with 22 hours of downtime per year).
- **Tier 3:** IT equipment are supplied by dual, and independent sources and they are always up and running. It's supposed to work 99.982 percent of the time (with only 1.6 hours of downtime per year).
- **Tier 4:** All of the infrastructures in data centers, both computing, and non-computing, are redundant. It should be up and running 99.995% of the time (with a 26.3-minute annual downtime).

1.1.3 Cloud Computing

Cloud computing is a computing paradigm for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction [3]. Cloud computing provides a number of advantages including flexibility, maximize hardware resource utilization, security, easy management, compliance with standards and many more. Virtualization is the main technology behind cloud computing. Without virtualization, an OS interacts directly with the physical machine and applications run on top of the OS. Virtualization introduces additional layer called virtual layer. A special software called hypervisor is used to create such layer. The machines created using virtual devices are called VMs.

There are three main types of cloud deployment models, these are public, private and hybrid.

Public Clouds: The cloud infrastructure is provisioned for open use by the general public. It may be owned, managed, and operated by a business, academic, or government organization, or some combination of them. It exists on the premises of the cloud provider.

Private Clouds: The cloud infrastructure is provisioned for exclusive use by a single organization comprising multiple consumers. It may be owned, managed, and operated by the organization, a third party, or some combination of them, and it may exist on or off premises

Hybrid Cloud: In this case the cloud infrastructure is a composition of two or more distinct cloud infrastructures (private, community, or public) that remain unique entities, but are bound together by standardized or proprietary technology that enables data and application portability (e.g., cloud bursting for load balancing between clouds) [4]

In addition cloud services are provided in different models. The three main types of models to deliver cloud computing services are:

Infrastructure as a Service (IaaS): can be used to satisfy the infrastructure needs of the students, faculties or researcher, globally or locally requiring some specific hardware configuration for a specific task. It provides a remote virtual hosting server for file storage, as it enables a user to save all of their file types in a virtual host and retrieve them from anywhere with an internet connection. Examples of IaaS include Google Compute Engine, Amazon Elastic Compute Cloud (EC2) and Microsoft Azure Digital Ocean.

Platform as a Service (PaaS): PaaS provides the entire infrastructure needed to run applications over the Internet. It allow customers to build their own application without the cost and complexity of buying and managing the underlying hardware and software layers. PaaS is based on a metering or subscription model so users only pay for what they use. Examples of PaaS are Google App Engine, Salesforce and Heroku.

Software as a Service (SaaS): the application service provider is hosting the application which runs and interacts through web browser, hosted desktop or remote client. It eliminates the need to install and run the application on customers own computer, thus simplifying maintenance and support services. Organizations that operate on SaaS are not burdened with the time-consuming and costly task of managing software updates, security patches and a host of other administrative duties for on premise software solutions. SaaS ensures that these tasks are managed quickly, efficiently and affordably on the back-end. [5]

A public cloud provider makes storage, VMs and other services available over the internet to anyone who wants them. Although workloads are isolated at the software level, they run on shared infrastructure. While A private cloud, by comparison, offers cloud-based services only to select users typically, those within a certain organization. Traditionally, private clouds run on infrastructure owned by that organization, although services such as Amazon Virtual Private Cloud now enable customers to build private clouds using public cloud data centers.

In general, public clouds and private clouds provide the same types of services. The crucial differences lie in who can access the services and who owns the infrastructure that hosts them.

1.2 Virtualization

Virtualization means creating something virtual instead of real, such as OS, a storage device or network entities, hardware. Virtualization is very popular with companies who have big IT infrastructures, because it enables running desktop computers, servers or storage devices in a visualized environment with almost identical functionality as real physical devices. additionally it enables running multiple VMs on the same physical host optimizing computing resources.

On the other hand virtualization describes the separation of resource or request for a service from the underlying physical delivery of that service. It allows for several application environments to run on the same machine in such a way that these environments are completely separated from each other. [6]. Virtualization enables many applications and OSs to run concurrently and in isolation on a single physical host machine. This facilitates several VMs to share the resources of the physical host machine, which in turn ensures improved utilization, optimization, and resource efficiency. On the other hand it is a key concept in cloud computing, without it then there is no cloud, this is what has enabled the emergency of a new industry that is and sustainable industry [7]

International Business Machine (IBM), through Massachusetts Institute of Technology (MIT), first implemented the concept of virtualization in the 1960s by Compatible Time

Sharing System (CTSS) in order to combat the rigidity and underutilization of the large mainframe computers. [8]. Mendel Rosenblum, a Stanford University researcher, found that this increase in servers became taxing on California's electrical grid because of the increase in the power required to operate, both powering and cooling. Rosenblum, in his research to find a better solution to this problem found out that by consolidating numerous servers into a larger one reduced the power and cooling requirements. Virtualization has two components; server virtualization which collapses multiple servers on to a single server, and desktop virtualization which reduces the size and support of the desktop system. [8]

The very rapid development in chip technology has brought huge increases in processing capacity. Most current computer hardware's are so powerful that in most cases only a small part of it, about 10 to 15 percent, is actually in use. So, most of the machine capacity available is not utilized and yet the machine continues to use space and electricity to function. This is an understandable waste that will become even more evident as the computing power of hardware continues to grow if not addressed using technologies like virtualization. Virtualization enables hardware to support more than one system and enables hardware utilization to go up to 75 or 80 percent which makes better sense for the capital being invested [9]

Although advances in server CPU enable much higher performance with less electricity consumption per CPU, overall server power consumption has continued to increase as more servers are installed with higher performance power-hungry processors which have more memory capacity. As more servers are deployed, they require more floor space. Through the packing of more servers in the same footprint the form factor of servers has become much smaller, in some cases shrinking by more than 70% in the course of the use of blade servers. This increase in packaging density has been matched by a major increase in the power density of data centers. Density has increased more than ten times from 300 watts per square foot in 1996 to over 4,000 watts per square foot in 2007, a trend that is expected to continue its increasing curve. [10]

Virtualization specifically server virtualization is a technology that helps us to install different OSs on the same hardware. To do this a hypervisor is needed which is a thin software layer that intercepts OS calls to the hardware. Virtualization hides the physical characteristics of computing resources from their users, their applications or end users. This includes making a single physical resource (such as a server, an OS, an application or a storage device) appear to function as multiple virtual resources.

1.2.1 Components of Virtualization

Before discussing the different types of virtualization, it is very compulsory to identify the different key components of virtualization. The following are the major components for virtualization to be accomplished.

Guest OS: is an OS that runs in a virtual environment independently of dedicated hardware resources on a hypervisor.

Hypervisor: also called Virtual Machine monitor (VMM) is the foundation of virtualization solutions. It allows multiple VMs to run on a single host to better utilize its hardware resources. The task of hypervisor is to handle resource and memory allocation for the VMs, ensuring they cannot disrupt each other.

Type of Hypervisor There are two types of hypervisor bare Metal hypervisor and hosted hypervisor (Figure 1 and Table 1). Type 1 hypervisor as the name bare metal indicates runs directly on the surface of hardware and includes hyper-v from Microsoft, VMware ESX/ESXI, and Citrix from xen, oracle VM server and amazon EC2. Hosted hypervisor on the other hand runs on the host OSs and includes Microsoft Virtual PC, VMWare workstation and Citrix Xen client. This research uses Microsoft Hyper-v for demonstration since it is incorporated into Windows Server OS and also it's easy and compatible with Microsoft products. [11]

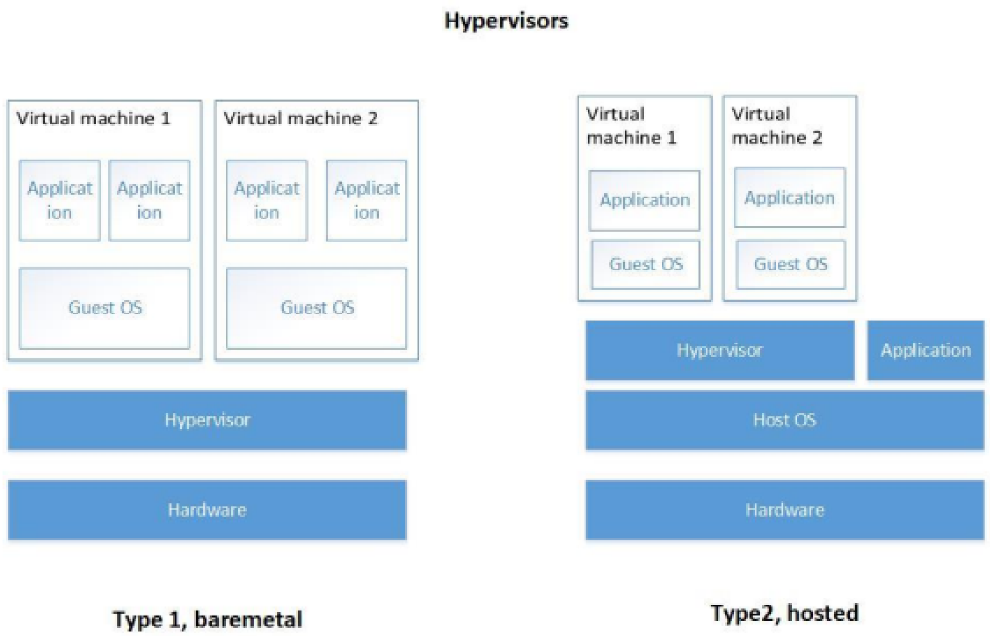

Figure 1. Types of VM

Hypervisor Type	Characteristics
Type 1 (Bare Metal)	▪ Runs directly on physical hardware ▪ Does not have OS running below it ▪ More secure than Type II ▪ Fully responsible for scheduling and allocating of the systems resources between VMs as Example: VMware ESX (Enterprise), Xen, Hyper-V
Type 2 (Hosted)	▪ VMM runs as an application in a normal OS ▪ The OS controls the real hardware resources called as Host OS ▪ Host OS has no knowledge of the type II VMM, which is treated like anyother process in the system ▪ The OS run inside of the Type II VMM is referred to as the Guest OS ▪ Lesssecure then Type 1 because any security Vulnerabilities that lead tothe compromise of the host ▪ OS will also give full control of the guest OS ▪ Host OS are heavyweight than Type II as Example : VMware GSX (workstation), UML (User- Mode Linux

Table 1. Types of hypervisor

Host machine The host machine is a physical machine intended to be virtualized and consists of the physical memory, hard disk, CPU etc.

VMs Sometimes called virtual instances is virtualization representatives runs on the hypervisor and implemented as a single file or folders on the host machine.

1.2.2 Advantages of Server Virtualisation for Education

Server virtualisation in education can have a number of advantages in that it can allow easier management of multiple server resources, less server hardware is required reducing capital expenditure and energy costs, expansion of further server resources is easier and less expensive and it is possible to run multiple OS on one server device.

The key advantages of Server Virtualisation for Education include:

- Significant savings possible through less need for hardware, maintenance, storage space and energy.
- Multiple OS including Windows or Linux can be run on a single machine.
- Both Mac and PC environments can be created to run on one machine
- Improved continuity of disaster recovery

- Reduced costs for additional servers.

1.2.3 Challenges of virtualization

According to Radhwan and Asmaa (2013) improper employment of server virtualization can result the following pitfalls

- Overloading the server utilization infrastructure which can introduce application latency.
- Increasing IT operational costs because of additional, time and resources required for extensive research efforts.
- Magnifying failures because a hardware failure could impact multiple virtual servers and the applications they host
- Introducing VM sprawl, which may substantially increase the overall number of server operating images that need to be managed by system administrators
- Enabling improper security processes because within the virtual server, the server administrator with access the root ID can alter or disable security settings; thereby , exposing servers to security vulnerabilities
- Exposing IT operations to network (traffic) uncertainties requiring enhanced IT skill sets to manage more environments at once [12]

1.2.4 Properties of Virtualization

Virtualization has three characteristics that make it ideal for cloud computing:

- **Partitioning:-** Many applications and OS are supported in a single physical system by partitioning (separating) the available resources.
- **Isolation:-** Each VM is isolated from its host physical system and other virtualized machines. Because of this isolation, if one virtual-instance crashes, it doesn‘t affect the other Vms. In addition, data isn‘t shared between one virtual container and another.
- **Encapsulation:-** A VM can be represented (and even stored) as a single file, so you can identify it easily based on the service it provides. In essence, the encapsulated process could be a business service. This encapsulated VM can be presented to an application as a complete entity. Therefore, encapsulation can protect each application so that it doesn‘t interfere with another application. [13]

1.2.5 Types of virtualization

There are five major types of virtualization each having its unique characteristics and significance. Virtualization can be network virtualization, storage virtualization , application virtualization, desktop/client virtualization and server virtualization, [11]

1.2.6 Storage virtualization

Storage virtualization is a system administration practice that allows decoupling the physical organization of the hardware from its logical representation. Using this technique, users do not have to be worried about the specific location of their data, which can be identified using a logical path.

Storage virtualization allows us to harness a wide range of storage facilities and represent them under a single logical file system. There are different techniques for storage virtualization, one of the most popular being network-based virtualization by means of storage area networks (SANs). SANs use a network-accessible device through a large bandwidth connection to provide storage facilities. [14]. If a host server goes down, another host can access that VM from the SAN and turn it on. This improves the high availability of these VM's by facilitating mobility between which host servers it will run on.

1.2.7 Network Virtualization

Network virtualization proposes decoupling of functionalities in a networking environment by separating the role of the traditional Internet Service Providers (ISPs) into two: infrastructure providers (InPs), who manage the physical infrastructure, and service providers (SPs), who create virtual networks by aggregating resources from multiple infrastructure providers and offer end-to-end network services. [15]. In a Network virtualization environment one or more VMs can access the local or external network using the physical network adaptor attached to physical machine. It can also be connected without physical network adaptor and it uses logical network adaptor.

1.2.8 Application Virtualization

Application virtualization is a process where applications get virtualized and are delivered from a server to the end user's device, such as laptops, smartphones, and tablets. The user can then access and use the applications from virtually anywhere. Any user actions are transmitted back to the hosting server. This type of virtualization is particularly popular for businesses that require the use of their applications on the go. [16]

1.2.9 Desktop/Client Virtualization

Similar to application virtualization mentioned above, desktop virtualization separates the desktop environment from the physical client device that is used to access it. The major advantages of desktop virtualization are that users are able to access all their personal files and applications from any location and on any computer. It also lowers the cost of licensing for installing software on desktops and maintenance and patch management is very simple. Virtual desktop infrastructure (VDI) is a type of desktop virtualization.

This type of technology makes the system administrator to virtually monitor and update the client machines like laptops, workstation desktop and mobile devices. It improves the client machines management and enhances the security to defend from hackers and cyber criminals. [17]

1.2.10 Server Virtualization

Virtualization has become popular in data centers since it provides an easy way to partition physical resources, allowing multiple applications to run in isolation on a single machine. The creation and management of VMs is called virtualization. It decouples software from hardware and splits multi-processor servers into more independent virtual hosts for better utilization of hardware resources, allowing services to be distributed one per processor. It promises to dramatically change how data centers operate by breaking the bond between physical servers and the resource shares granted to customers. It helps to consolidate the load of multiple underutilized volume servers onto fewer servers, thereby reducing costs, energy consumption and cooling requirements. [18]. There are many vendors providing virtualization like VMware, Cisco, Citrix, EMC, HewlettPackard, IBM, Microsoft, NetApp, Novell, Oracle, Quest, Sun, Symantec, Virtual Iron, and many more.

This sharing of resources is done by creating a specialized virtualization layers which converts a physical resources in to VMs. The possibilities of running multiple OS concurrently on the same machine strengthen the above fact. Figure 2 clearly revealed the difference between physical servers and visualized machines.

Virtualization Software Vendors

There are numerous vendors of virtualization software in the market. But now a day's three vendors are the most widely known.

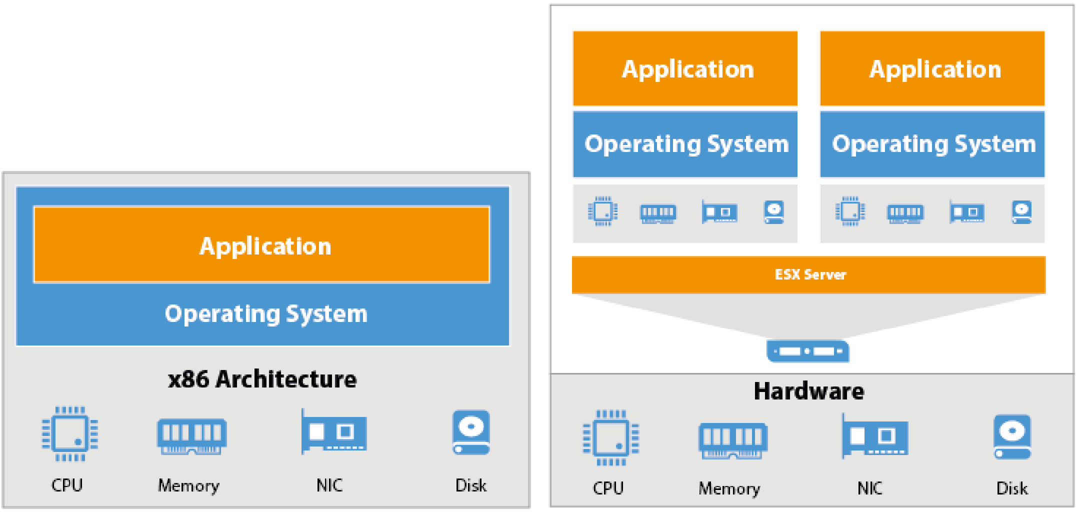

Figure 2. Physical Servers and visualized machines

Hyper-V: Microsoft Hyper-V is a hypervisor-based virtualization technology for x64 versions of Windows Server which is widely focus in virtualization business after the introduction of Microsoft server 2008.

VMware: VMware vSphere is a software suite that has many software components such as vCenter, ESXi, and vSphere client which is also a leading technology and innovation leader focusing on virtualization from the beginning. It is still dominant because it has great installation in enterprise environment, broad product range and technology advantage.

Citrix: Citrix also other virtualization software vendor cost-effective desktop, server, and cloud virtualization infrastructures.

Disadvantages of Server Virtualization: The disadvantages of virtualization are mostly those that would come with any technology transition. With careful planning and expert implementation, all of these drawbacks can be overcome. Server virtualization has the following disadvantages.

- Servers and applications compatibility issues: Only a few processors that support virtualization can be used to visualized servers.
- The resource allocation for each virtual system needs to be planned carefully. If very less resource is allocated, the application performance might be affected and if too much resource is allocated, it will result in under utilization

1.3 Conceptual Model of Server Virtualization

Server virtualization is a technique that changes server model architecture and establish a new layer called hypervisor between physical layer and OSs. This layer includes several blocks of activities which are controlled by a management system. Hypervisor is a software environment that supports communication between hardware system and virtualization main activities such as memory, I/O, CPU, File system, network virtualization and resource management blocks. On top of this environment, a virtual server pool including variety of OSs in different VMs can support variety of application software. Regarding this complexity, from a user perspective, server virtualization means a single server resources (including number and type of individual physical servers, processors) and OSs.[19]

Figure 3 shows Hypervisor/Bare metal virtualizes hardware resources for different VMs and their guest OS. In a virtual server, a physical server with all its resources becomes multiple virtual servers so that each virtual server has capabilities of the physical server and operates separately from other VMs. CPU virtualization means CPU processing power of physical server divided into several sections and each section is assigned to a virtual processor unit. Memory virtualization means dynamic allocation of memory to each of the VMs that are installed on the host server. In the virtualization, input/output gateways and bandwidth of physical server is divided between VMs. Virtual file system is a file cluster system with high performance that allows multiple virtual servers have read/write access to a file simultaneously. In virtual networking, when VMs are created on the host server, it assigned a virtual IP to each machine. Resource management has responsibility for managing real and virtual resources in the system. [19]

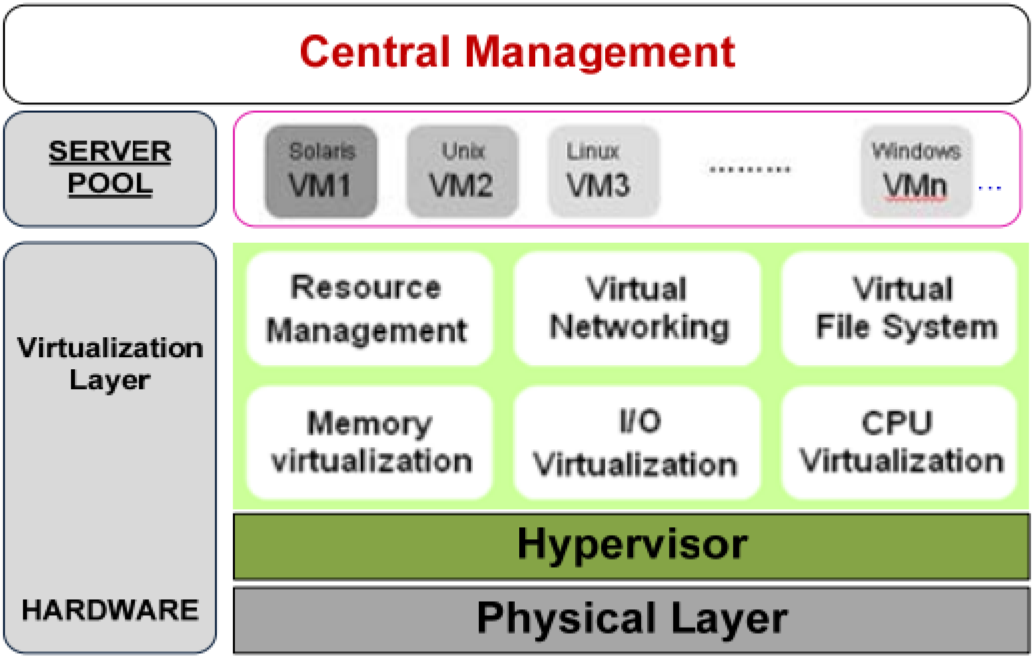

Figure 3. Conceptual Server Virtualization

Among 31 universities in Ethiopia the researcher take two of them from first generation universities, one from second generation universities, one from third generation universities and from other Ethiopian education and research network (Ethernet) are selected from data.

1.4 Overview of Selected Universities data center

Most of the current data center infrastructure in HLI was traditional one with no virtualization. And the data center running on components like servers, storage devices, chillers, switches, uninterruptible power supply (UPS), computer room air conditioners, direct expansion (DX) units, pumps, cooling tower, generators, power distribution units (PDUs), fire suppression, physical security devices and lightings and other. Generally data center consume large amount of energy consumption globally, data centers will become the world's largest users of energy consumption, with the ratio rising from 3% in 2017 to 4.5% in 2025. [20] Also data center become the nerve centres of our digital economy [21] even though HU runs physical data center in which a one application for one server architecture is exercised which have numerous barriers.

Learning institutions like universities spend a great deal of money each year to buys new servers and upgrade existing servers within their educational institutions. So that in educational institutions virtualization would allow them to offer students and faculty a variety of OSs and applications while reducing the number of servers and computers that are needed. Reducing the number of servers and computers could lower the equipment and support costs budgeted each year.

1.4.1 Devices of the Data Center of the University Institution

Since different organizations use data centers, their architectures and requirements can differ significantly. Regardless of classification, data centre safety is the number one priority. An effective data center operation is achieved through a balanced investment in the facility and the equipment it houses.

Below, you'll find a list of components that all data centers share, no matter what their specific purpose is.

Network

Network equipment is used to provide a connection between various external and internal IT devices, most commonly servers. There are several ports on this equipment, which are generally switches and routers.

Air conditioning cooling system:

Data centers typically feature precision air conditioning systems, which are either direct expansion type air conditioning systems or chilled water cooling systems. In HLI data center can handle the data center's maximum IT heat load, and function normally in extreme ambient conditions. Cooling system has to be able to achieve a full environmental control, including air temperature, humidity and pollution concentration.

The cooling is one of the three main infrastructures of a data center as it maintains IT equipment working in a safe and reliable manner and it can account for up to 40% of the total energy consumption in a data center.The complicated connection of chillers, compressors and air handlers create the best possible computing environment which ensures the long life of the servers and the strength of the organization they support.[22] [23]

Backup Generator:

Data center should be Provided with generator as generator provide power to the Data Center in the event of power failure. The generator should be checked on a regular basis to make sure that fuel is at an acceptable level for extended run time if needed.

Data Center Uninterruptible Power Supplies (UPS)

A UPS is, essentially, a battery backup that supplies power to your system in order to provide enough time to properly power down your equipment when there is a failure in utility power. You may have a facility that already has a UPS (Uninterruptible Power Supply) or even a generator.

UPS in the recent years are playing significant role in achieving protection against power interruptions and in improving efficiency and cost effectiveness and providing high secured data [24]

Fire Protection

Fire is the most devastating hassle in both traditional and modern data centers. Different fire protection systems can be used including active and passive design elements as well as implementation of fire protections programs.

Smoke detectors are the most important system installed to provide early warning of a developing fire by detecting particles generated by smoldering components prior to the development of flame. This excellent alert allows investigation, interruption of power, and manual fire suppression using hand held fire extinguishers before the fire grows to a large size.

Those fire protection elements can be active or passive. Active fire protection systems includes tasks like closing doors but, Passive fire protection elements include the installation of fire walls around the data center so a fire can be restricted to a portion of the facility for a limited time [25]

Computer Servers

Servers are the core of a data center system that generates useful output, and they consume a large amount of energy in a data center while also generating heat it is a load on the cooling system. This is primarily due to equipped with a large amount of processing power and computing capabilities.

The researcher take sample selection from (first generation Addis Ababa University and Hawassa University, from second generation Dilla University and from third generation Bule Hora University) and Ethernet data center also included in data collection.

In Addis Ababa University there are totally twenty six servers in the university data center among three are Huawei, six are Cisco servers, other six are dell servers and the rest eleven servers are HP servers but almost all serves of system servers are in virtualizes environments

Name of Servers	OS	Processor Speed	CPU Cores	RAM	Hard Disk	NIC	Installed Sys
Huawei(3)	Win	2.53 GHz	4	192GB	20TB	4	4
Cisco(6)	Linux	2.53 GHz	4	132GB	20TB	3	6
DEll(6)	Linux	2.53 GHz	2	132GB	4TB	4	1
Huawei(3)	Win	2.53 GHz	4	132GB	20TB	4	1

Table 2. AA University DC server list

The other data center the researcher try to observe is Ethiopian education and research network (Ethernet) data center which can offer the service software as a service, infrastructure as a service and platform as a service.

- Software as a service like National Academic Digital Repository of Ethiopia, National Academic Digital Library of Ethiopia and Ethiopian Journals Online are among the services.
- Platform as a Service VMware which is It is a robust, bare-metal hypervisor that installs directly onto your physical server. and ESXi cPanel which is cPanel is a web hosting control panel software.
- Infrastructure as a Service
- Also National High Performance Computing Center for Education and Research which is parallel processing for running advanced application programs efficiently, reliably and quickly

Name of Servers	OS	Processor Speed	CPU Cores	RAM	Hard Disk	NIC	Installed S
Hemis (5)	Win	2.53 GHz	4	192GB	132TB	4	5
Dell(6)	Linux	2.53 GHz	4	132GB	64TB	4	6
HPC(1)	Win	2865.26GHz	2	18593.7GB	51TB	4	1
National Exam	Win	285.26GHz	4	132GB	64TB	4	1
National SMS	Win	285.26GHz	4	132GB	64TB	4	1

Table 3. Ethernet DC server

In Hawassa University data center there are a lots of servers used to host and provide services are around seventeen (17) active servers like file server, Integrated Student Information System(SIS) server, SMS server, Store Management server, Library Management System (LMS) server, Domain Name System (DNS) server, proxy server, Balanced Score Card (BSC) server, digital library server, Research Management Server, Teeal Journal server, e-learning sever , Database Server, digital signage server, Web site server, Institutional Repository (IR) server and Hu-journals Server . All of the system services are running on separate physical servers, which is very inefficient and can be optimized using virtualization technologies.

Name of Servers	Operating System Type	Processo r Speed	CPU Core	Memory (RAM) capacity	Hard Drive Disk	No. of NIC	No. of System Installed
Hu-File	Window Server 2016	2.53 GHz	4	16GB	1TB	4	**1**
Hu-Student Info System	Window Server 2019	2.53 GHz	4	16GB	3TB	4	**2**
Hu-SMS	Window Server 2019	2.53 GHz	4	32GB	1TB	4	**1**
Hu-Store Management	Window Server 2016	2.53 GHz	4	8GB	1TB	4	**1**
Hu-Library Catalog	Ubuntu Desktop	2.53 GHz	4	8GB	1TB	2	**1**
Hu-DNS	Linux/Ubuntu/Kali)	2.53 GHz	4	8GB	1TB	4	**1**
Hu-Proxy	Ubuntu server	2.53 GHz	3	8GB	1TB	4	1
Hu-BSC	Window Server 2016	2.53 GHz	4	16GB	500GB	4	1
Hu-E-book	Window Server 2016	2.53 GHz	4	8GB	1TB	4	**1**
Hu-IR	Window Server 2016	2.53 GHz	4	32GB	1TB	4	**1**
Hu-Teeal Journal	Ubuntu server	2.53 GHz	**2**	2GB	500GB	1	**1**
Hu-E-learning	Ubuntu Desktop	2.53 GHz	3	16GB	1TB	4	**1**
Hu-Digital signage	Window Server 2016	2.53 GHz	**3**	4GB	500GB	4	**1**
Hu-Website	Window Server 2019	2.53 GHz	4	16GB	1TB	1	**1**
Hu-journals	Window Server 2016	2.53 GHz	4	16GB	1TB	4	**1**
Hu-Research Mgmt	Window Server 2016	2.53 GHz	4	16GB	1TB	4	**1**
Hu-Database	Window Server 2019	2.53 GHz	4	16GB	1TB	4	**1**

Figure 4. List of functional server in the Hawassa University data center

Figure 4 indicate list of current functional server in the HU data center with detail servers population sizing.

From second generation universities Dilla University is selected the university data center have totally eight physical servers among six of physical isolated and two are blade servers (HP blade servers and Huawei blade) servers each of the blade servers are virtualized and running deferment system application. HP blade servers runs DNS servers, Database servers, student back end and Huawei blade servers runs front-end (portal), active directory, e-learning, DU-journal servers, database system, and digital repository servers.

Name of Servers	OS	Processor Speed	CPU Cores	RAM	Hard Disk	NIC	Installed Sys
HP Blade	Win	2.53 GHz	4	57.6GB	20TB	2	4
Huawei Blade	Win	2.53 GHz	4	132GB	20TB	2	5
Video surveillance1	Win	2.53 GHz	4	32GB	1TB	2	1
Video surveillance2	Win	2.53 GHz	4	32GB	1TB	2	1
Access control	Win	2.53 GHz	4	16GB	1TB	4	1
School net	Win	2.53 GHz	4	16GB	1TB	4	1
Anti-virus	Win	2.53 GHz	4	8GB	1TB	4	1

Table 4. Dilla University DC server list

From third generation university Bule Hora university is one of them

Name of Servers	OS	Processor Speed	CPU Cores	RAM	Hard Disk	NIC	Installed Sys
SRR	Linux	2.53 GHz	4	16GB	1TB	2	1
BHU E-learning	Linux	2.53 GHz	4	16GB	2TB	4	1
LMS Server	Win	2.53 GHz	4	16GB	1TB	4	1
SIS	Win	2.53 GHz	4	32GB	4TB	2	1
Digital library	Linux	2.53 GHz	4	16GB	1TB	2	1

Table 5. Bule Hora University DC server list

1.5 Statement of the problem

In the last decades enterprise information system paradigm shifted from on premise dedicated infrastructure to virtual infrastructure. Different factors contribute for this paradigm shift. Mainly, the advantages provided by the virtual environments (e.g., cost reduction, reliability, mobility, quick deployment, scalability, security etc.) outmatch the benefits and the services provided by on dedicated infrastructure. Cloud computing is a computing paradigm which relies on virtualization technology and it is the standard approach for enterprises to host their information systems. Despite grows and popularity of the cloud computing paradigm, in Ethiopia the adoption is at a very slow pace. Specifically, universities in Ethiopia still build their data center infrastructure in a traditional approach which requires huge number of resources. In order to optimize business operation process with the growth of data on daily bases demands organizations to increase their infrastructure in storage, network and processes. As a result, universities are spending huge amount of money on data center infrastructure for computing (CPU, RAM, GUP ...), storage, cooling system, power, human resource etc... The rapid expansion of using automation in the Ethiopian higher education institutions and their data centers infrastructure faces a serious of risks associated with the resource, energy and cost effectiveness. The Universities data centers' host a variety of software applications running on heterogeneous systems running on physical servers. For example, in HU, more than seventeen independent servers are used to run different services which is costly on hardware as well as software, power concussion, cooler energy, network bandwidth and space and etc

Moreover, most servers in such data centers are idle (performing nothing) for more than 90% of the time while consuming huge power and also generate large amount of CO2 which is hazardous for environmental sustainability and global warming. [26]

All of these services can be run on limited resource using virtualization technologies that optimize the resource required. Adopting a virtualized infrastructure, specifically the cloud computing paradigm will increase the productivity of servers with no additional cost and at the same time reducing the energy consumption and future investments on the data centers and reduction of CO2 generation.

This research focuses on the development of a framework for migration of the current information system infrastructure, which is based on physical servers and network equipment, to a virtualized infrastructure where services are provided using VMs and virtual network equipment. To overcome some of the problems highlighted above that would allow to efficiently use available resources, reduce the power consumption and potentially reduce future investments in the data centers etc.

Thus, the paper investigates the data center practices that are employed in current data center, the efficiency, cost and availability of currently hosted services and how virtualization be adopted as an alternative and preferred solution for our universities.

1.6 Research Question

Universities spend a great deal of money each year to maintain, patch, and upgrade computer servers within their educational institutions ,and most Ethiopian Universities data center currently built their data center infrastructure traditionally as growth of data on daily bases leads the organization to increase their infrastructure in storage, network and processes. To face these challenges Universities are spending a huge amount of money on data center infrastructure which is storage, servers, networks cooling and power. In case of HU the main campus and the current data center infrastructure is the traditional one with no virtualization. The objective of this research is to develop the framework for migration of the current physical infrastructure to virtual infrastructure environment that brings a preferred solution for educational activities. Generally the following research question was to be answered in this research paper:

- Identify the problems of the current infrastructure systems of higher learning institutions? How could Data center virtualization be an alternative and preferred solution for University data centers?
- What data center practice is employed in HU Data center? What is the need for server virtualization in the University Data center?
- What are the problems that burdens for the implementation of virtualization in data center?
- What server virtualization frameworks do exist? and How can the framework be evaluated?
- How could a virtualized data center performed compared to the current system?

1.7 Objective of the study

1.7.1 General Objective

The general objective of the study is to develop a framework for migrating from physical infrastructure to virtual infrastructure and evaluate virtualize environment against non-virtualized environment.

1.7.2 Specific Objective

The Specific objective of this research is

1. To examine existing approaches to physical infrastructure deployment in higher learning institutions
2. To review literature's related with the general objective of the study
3. To identify components of the virtualized data center framework.
4. To develop a framework for migrating from physical infrastructure to virtual infrastructure for Ethiopian Universities.
5. Evaluate or test data centers that are deployed using virtualization technology with fundamental parameters CPU utilization, memory usage and disk usage.

1.8 Significance of the Study

This research is believed to produce a result that allows the HLI data center universities to implement a data center virtualization appropriately to effectively solve the existed problems in the current data center and making scalable, reliable and efficient data center. The result of this thesis will contribute to the ongoing researches in this domain area. Since virtualization is a step for cloud computing it helps organization to transform for this giant technology.

1.9 Scope of the Study

The thesis mainly focuses on developing a virtualization framework for HLI data center. It particularly focuses on full server virtualization. The study was to examine existing approaches to server deployment, determine the need for server virtualization, evaluate the existing server virtualization frameworks, develop a framework for server virtualization, and evaluate or validate the framework for server virtualization. Also the results of the thesis may not be applicable for other types of virtualization types like application, desktop/client and network virtualization.

1.10 Structure of the Thesis

Chapter 1: Is introduction which explains about data center, cloud computing, virtualization, statement of the problem, objectives, significance and scope of the study.

Chapter 2: Covers literature from different sources that support the work of the researcher. It covers main topics like current trends server Virtualization, virtualization environments, server virtualization assessment, comparative analysis of resources needed by physical server and virtual server, related works are discussed.

Chapter 3: Is generally about methodology like research design, sampling technique, data collection instruments, procedures etc. The researcher started by describing DSRM as the appropriate methodology for creation of artifacts. Also the thesis discuss about data analysis, data design and development, demonstration as well as evaluation.

Chapter 4: This chapter is about the designed proposed framework with it's various components.

Chapter 5: Is results and discussion: which defines the results after data is collected from the respected bodies and verified. It is also about experimentation which shows the importance of virtualized data center over physical data centers by using Hyper-V application software.

Chapter 6: Is the last chapter which is the section of concluding the researchers result and giving direction for future work.

2. LITERATURE REVIEW

2.1 Introduction

This chapter deals with the literature associated with virtualization and the researcher strongly finds related literature's that help to understand the theoretical background of data center virtualization using different search parameters. Google, Google scholar and IEEE were used to explore literature's. This is covered broadly under the following sub-headings: Current trends necessitating Server Virtualization, Server virtualization, forms of virtualization technologies, virtualization environments, server consolidation techniques, issues to consider with virtualization, benefits of server virtualization, factors driving the adoption of server virtualization, virtualization and resource management and overview of existing approaches to deploying server virtualization technology.

2.2 Current trends necessitating Server Virtualization

Ethiopia as a developing nation has suffered by limitation of educational budgets. Currently there are around 31 government owned universities and many private colleges. Each institutions invest huge amount of money for their IT infrastructure. This includes building data center and disaster recovery, network, storage, server and other related equipment for their Universities systems, the increasing application of IS into all areas of business and personal life has drawn attention to its environmental effects without quality education. Establishing educational institutions by itself can not give the solution for economic development and poverty reduction through education particularly if the education shouldn't be supported through up-to-date technologies and services.

Ethiopian government has been investing millions of dollars every year to support education in higher education institutions with technology. However it is not able to supply full ICT infrastructure requirements of all universities because the equipment cost of ICT like server and networking device is very high as a result the infrastructure requirements can be solved using implementing virtualization technology.

Major IT companies (for example, Facebook) who own large data centers have also reported to have oversubscribed their data center power to save data center infrastructure cost. A risk associated with data center power over subscription is that aggregate power could exceed the power limit due to simultaneous peaking of servers resulting into a power overload. [27]

The necessity for sustainable IS management derives from ever-growing power demands, electronic waste streams, data amounts and performance that are expected in future toward IT.In 2014, data centers in the U.S. consumed an estimated 70 billion kWh, representing approximately1.8% of the total U.S. electricity consumption [28]. Also electricity usage costs have become an increasing fraction of the total cost of ownership (TCO) for data centers. According to Lean & Green consortium, each kilowatt of energy consumed to power on a server requires an additional kilowatt to cool that machine. It is now to an extent where the costs of power consumed by a server exceed the cost of the server itself [29]. The ability to run multiple VMs on a single physical server has significant implications for the energy required to power an organization's IT services and can reduce power demand accordingly to reducing 50 servers in a certain data center over 3 years we can lessen 1,790,982 Lbs (895tones) CO2, 37 Lbs methane, 25 Lbs nitrous oxide, 7143 Lbs sulfur dioxide, 2239 Lbs nitrogen oxide that is equivalent to 177 passenger not driven for one year. So, it is risky for environment sustainability and leads to global warming. [30] It is also difficult for developing countries like Ethiopia to provide energy to these ever-growing data centers.

Virtualization solves efficiency problems by effectively providing a virtual, logical environment of computing resources rather than a rigid, physical environment by allowing multiple OS to run simultaneously on a single machine; and by tricking OS into thinking that a group of hardware components or servers is a single pool of computing resources. This can be achieved through the various ways of visualizing the computing environment which include but not limited to server virtualization, client virtualization, and etc. This study will mainly focus on server virtualization.

Organizations which have invested on many servers will realize the benefits of virtualization if it is on a many-to-one mapping, that is to say many applications running on different physical servers being consolidated to run on one physical server. Savings will be realized from purchasing and maintaining fewer servers and reduced power consumption. Even smaller organizations may reap form the benefits of virtualization, particularly when facing the necessity to upgrade their servers. One large server is less expensive to buy, operate and maintain than even a small population of servers [31]

In a virtualized environment the hardware represents a shared resource and is not explicitly dedicated to one application. Thus this allows the system to cope with demand easily, for instance an application which has a high CPU demand and requires more resources, the application can be replicated with a load balancing mechanism to handle the demand. In the same way, if an application is not highly demanded, its resources can be reduced and eventually some of the hardware resources could even be powered off. This ability to start services on-demand and switch them off, indicate that energy saving is achieved during off-peak time, when the servers are not extensively used [7]

The IT industry has in the recent years identified Green IT. and IT investors and consumers are beginning to look at the carbon footprint of an IT company and its products whereby a carbon footprint is a measure of the total set of greenhouse gas emissions caused directly and indirectly by an individual, an organization, an operation or a product [32].

IT efficiency can be improved by using energy efficient IT equipment, improved airflow management systems that reduce cooling requirements, and especially virtualization technology. Virtualization technology is a major strategy to reduce the energy consumption in data centers. Server virtualization aims at splitting hardware resources into several smaller VM, enabling the hosting of more than one VM on a single hardware, thereby increasing multitasking capability, increasing utilization of servers, and improving energy efficiency. In typical data centers, average utilization is only 20-30%.[33]

2.3 Virtualization environments

Technologies that have been widely tested such as Hyper-V and VMware are being used for this kind of virtualization [34]. Once the physical server is configured with the virtual operating system, the server is ready to begin hosting virtual servers. A virtual server is then configured with its operating system and support services. Virtual servers are combined based on CPU utilization, transaction processing, and hardware services, a combination which seeks to provide an optimal ratio of the number of virtual servers to a physical server. The virtual servers are loaded onto the physical server, turned on and at this point they operate as independent servers. Should a virtual server go down, it is quickly restored without impact to the other virtual severs on the same physical server [35]

2.4 Benefits of Virtualization

Virtualization promises to radically transform computing for the better utilization of resources available in the data center reducing overall costs and increasing agility. It reduces operational complexity, maintains flexibility in selecting software and hardware platforms and product vendors. It also increases agility in managing heterogeneous virtual environments. Some of the benefits of virtualization are:

2.4.1 Server Consolidation

Server consolidation is an approach to the efficient usage of computer server resources which involves consolidating data centers in order to reduce the total number of servers or servers locations required. The growth of server consolidation is due to virtualization which enables multiple VMs to share the physical resources of a computer.Server Consolidation provides cost savings in different ways to reduce the overall total cost of ownership of data center. These savings can be categorized as:

Staff Costs: Supporting staff is recruited on the basis of how many servers an organization has (one member of staff per twenty servers is often used as a benchmark). Most server consolidation projects aim to reduce costs by freeing staff from mundane server maintenance tasks. Gartner (2008) suggests that more than 70% of potential savings from a typical project will come from reduced staff requirements [36].

Hardware Costs: Consolidation can reduce costs through better server utilization, by reducing the total requirement for storage and by enabling the use of more cost effective back-up/restore mechanisms. Centralized purchasing may also enable better discounts to be negotiated with hardware suppliers

Software Costs: Consolidation may also reduce the total number of licenses needed while standardizing fewer applications, that may allow better deals to be negotiated with suppliers. With many (but not all) applications, the incremental cost of software licenses decreases as the number of user's increases [37]

Facilities Costs: Server consolidation reduces the amount of floor space needed for data centers. This is a particular benefit if one or more existing locations are full [37]

2.4.2 Simplified Management

A VM can be more easily managed, configured and controlled from outside than a physical one. While it is possible to switch off a physical host remotely, it is impossible to switch on a turned off physical machine. In the virtual environment, all the actions on the VM can be managed from a remote location, including the power on and power off. By implementing virtualization, the number of physical machines reduces automatically hence management becomes simpler and cost effective

2.4.3 Sand-boxing

Virtualization provides secure, quarantined sandboxes for running entrusted applications. Examples include address obfuscation. They also provide fault and error containment by isolating applications and services they execute, ultimately providing better behavior of these different faults.

2.4.4 Multiple Execution Environments

VMs are used to create operating systems and execution environments to guarantee resource management and using resource management schedulers with resource limitations. VMs provide the illusion of hardware configuration such as SCSI devices. Virtualization can also be used to simulate networks of independent computers. It enables to run multiple operating systems simultaneously having different versions, or even different vendors sharing and executing the workloads of different applications being processes.

2.4.5 Debugging and Performance

VMs allow powerful debugging and performance monitoring tools that can be installed in the VM monitor to debug operating systems without losing productivity.

2.4.6 Physical to Live Virtual (P2V) Live Migration

A VM can easily be moved from one physical machine and relocated to another, if needed [38] Typical examples of VM migration are: VM cloning (copy more instances of the same VM), VM relocation (to free the original machine hosting the VM). This scenario helps to increase the mobility of server machines in data center resulting in high availability and disaster recovery solutions.

2.5 Issues to consider with Virtualization

Before implementing virtualization solutions, due diligence needs to be performed to ensure that there are no obvious and high impact problems that result from virtualization. The following issues need to be noted:

1. The compatibility of existing software and systems with virtualization needs to be confirmed in a pilot rollout before proceeding with large scale deployment. Though most virtualization products claim to support unmodified applications, organizations

should not proceed before ensuring that this indeed is so. The costs of incompatibility can be much high than the benefits that virtualization brings.

2. The exiting licensing models could be unsuitable for virtualization. The existing models charge based on the number of instances of a particular software running, virtualization could end up being very costly if it leads to increases in the number of machines, and thereby the running instances of the software, in the system.
3. Every virtualization solution comes with a certain performance overhead. These overheads have to be measured and quantified during the evaluation stage to inform the virtualization decision. Only after ensuring that the loss, if any, in performance is within acceptable limits for the organization, should roll out begin.

2.5.1 Benefits of Server Consolidation

Server consolidation is accomplished by running multiple VM-based services on each physical server, reducing the number and total cost of physical servers required to support a given data center workload. Server consolidation is playing a growing role as data center operators seek to minimize costs, and as compute capacity rapidly increases in the many-core processor era.In order to run services with minimal slowdown compared to dedicated servers, co-located VMs must share the resources of a physical server without heavy contention. While CPU scheduling techniques can be used to dynamically multiplex CPU usage across VMs to closely match VM compute requirements over time, it is more challenging to efficiently share memory capacity. Each VM has a pre-set expectation of the physical memory capacity it can access [39]

Lower TCO: with improved utilization due to consolidation and lower maintenance cost by cutting down server proliferation, enterprises witness significant reduction in total cost of ownership of their IT resources. Enterprises are able to deploy newer applications without having to buy additional hardware [40]

High availability: features such as Live Migration with near zero downtime allow active VMs to be migrated during scheduled maintenance, thus allowing applications and services to continue running even during regular maintenance cycles. Also, it is possible to create highly available data center using VMs over fewer physical machines than was required earlier in active /passive modes [40]

Increased server utilization: The average server utilization in an enterprise environment currently ranges from 5 to 40 percent, leaving at least 60 percent of the available capacity unused. Virtualization enables one to consolidate 'complimentary' workloads on multiple physical servers and harness unused computing resources. This entails replacing several older servers dedicated to single applications with a single more powerful server running multiple applications [41]

2.6 Factors Driving the Adoption of Virtualization

The following trends are impacting data centers, hence driving the need for adoption of virtualization:

Increasing equipment power density: although advances in server CPUs have in some cases enabled higher performance with less power consumption per CPU, the general server power consumption has continued in the upward trend as more servers are installed with higher performance power consuming processors with more memory capacity. The more the servers are installed, the more the floor space required. To pack more servers in the same footprint the form factor of servers has become much smaller, in some cases decreasing by more than 70% through the use of blade servers. This increase in packaging density has been matched by a major increase in the power density of data centers. Density has increased more than ten times from 300 watts per square foot in 1996 to over 4,000 watts per square foot in 2007, a trend that is expected to continue its upward spiral [42]

The rapid growth of internet: the increasing dependence on electronic data is motivating the swift growth in the size and number of data centers. This growth results from the rapid adoption of Internet communications and media, the automation of business processes and applications, legal requirements for retention of records, and disaster recovery. The use of internet is growing at more than 10 percent per annum leading to an estimated 20% Compound Annual Growth Rate (CAGR) in data center demand. Video and music downloads, online gaming, social networks, e-commerce, and voice over the internet (VoIP) are key drivers. In addition, business use of the Internet has ramped up. Industries such as financial services investment, real estate, healthcare, retailing, manufacturing, and transportation are using IT for key business functions. The advent of the Sarbanes-Oxley Act with its requirement to retain electronic records has increased storage demand in some industries at 50 percent CAGR. Disaster recovery strategies that mandate duplicate records the increases demand further [42]

Increasing cooling requirements: the increase in server power density led to an associated increase in data center heat density. Servers require approximately between 1 to 1.5 watts of cooling for each watt of power used. The ratio of cooling power to server power consumption requirements is expected to continue increasing as data center server densities increase [42]

Increasing energy cost: data center expenditures for power and cooling can surpass that for equipment over the useful life of a server. For a typical $4,000 server rated at 500 watts, one study estimated that it would consume approximately $4,000 of electricity for power and cooling over a period of three years, at $0.08 per kilowatt-hour. For the case of Japan, this is estimated to double. The ratio of power and cooling expense to equipment expenses approximately increased from approximately a ratio of 0.1:1 in 2000 to that of 1:1 in 2007. With a likely increase in the number of data centers and servers and the advent of a carbon cap-and-trade scheme, the cost of energy for data center power and cooling is expected to continue increasing [42]

Low server utilization rates: data center efficiency is a major problem in terms of energy use. The server utilization rates average 5-10 percent for large data centers. Low server utilization means that companies are overpaying for energy, maintenance, operations support, while only using a small percentage of computing capacity [42]

Growing awareness of IT impacts on the environment: carbon emissions are estimated to be proportional to energy usage. In 2007, there were approximately 44 million servers worldwide consuming 0.5% of all electricity. Data centers in the server dense U.S. use more than 1% of all electricity. Their collective annual carbon emissions of 80 metric megatons of CO2 are approaching the carbon footprint of the Netherlands and Argentina. Carbon emissions from operations are expected to grow at more than 11% per year to 340 metric megatons by 2020. In addition, the carbon footprint of manufacturing the IT products is largely unaccounted for by IT organizations [42]

2.7 Virtualization and Resource Management

Server virtualization promises a dynamic reallocation of underlying computer and storage resources to meet spikes in demand on individual applications. This is the real key to increasing efficiency in resource allocation. One major reason that the non- virtualized environment is characterized by such low utilization is that each application requires sufficient computing power, memory and storage to meet its maximum load plus an overhead to accommodate future growth. It is often difficult to predict the speed of growth, which tends to follow a “hockey stick” graph rather than a smooth curve-usage is low while users learn a new application and then may jump quickly as they become comfortable with it and realize its potential for improving their work [31]

2.8 Server Virtualization Assessment

The servers were classified into two categories based on their capacities in terms of number of processors, speed of the processors, number of network ports, number, installed RAM, capacity of hard disk drives and number of systems installed on each server also server virtualization assessment involves problems such as capturing the current configuration, performance, and environment data for the infrastructure you are interested in virtualizing, accessing data for. With a set of requirements and restrictions for virtual infrastructure, creating key reports to support you build a detailed transition plan for the existing infrastructure. The final conclusion about evaluating server virtualization issues, you need to achieve the following issues:

- Server Workload (the amount of load in form of client requests, processing and communications resources expected in a specified time period) must be appropriate for virtualization Workload is not suitable for virtualization because of hardware, performance, or application compatibility limitations.
- Appropriate workload combinations of resources provide the optimal utilization of resources within the performance boundaries of physical server hardware and virtual software (eg Windows Server 20016 Hyper-V).
- VMs can be reconfigured into virtual hosts in the process of converting from existing infrastructure to virtual infrastructure.
- Preliminary assessment of the reduction in consumption on some aspects such as electricity, cooling and space.

There are three main data sets that you need to support to assess the server virtualization problem:

1. **Inventory of server software and hardware:-** Specifying which existing workload is appropriate for virtualization depends on selecting specific hardware and software information for each server within your project.
2. **Performance metric:-** include performance metrics for each server. In particular, you need to capture processor, memory, network, and disk performance parameters.
3. **The details of the environment:-** The benefits of cost reduction from migration to a facility virtual infrastructure. The main data need to collect is essentially about power and cooling issues, the cost of the space allocated to each physical server is the goal of virtualization. Once we have defined virtual candidates, the number of virtual hosts and their associated systems need to be added to the infrastructure, we can estimate the cost savings from reduced servers and related systems.

The implementation of server virtualization evaluation by manual method will be very difficult and is not possible to perform with a large number of physical workloads. However, in this regard, the Microsoft Assessment and Planning (MAP) toolkit can help you automatically collect performance, software, and hardware data for your existing environment, as well as access to data for virtual restrictions. predefined, create reports of virtualized candidates for workloads.

2.9 Comparative Analysis of Resources Needed by Physical Server and Virtual Server

An important feature of server virtualization is energy saving. In power enterprises, the number of original physical servers is a fixed value, which will not change under normal circumstances, but the number of virtual servers generated by network architecture is variable. This shows that the enterprise can adjust the virtual server according to its real-time situation. The more servers, the higher the power consumption. [43]

Table 6 shows the resource comparison between the physical server and the VM. From the comparison results, VM has obvious advantages in improving memory utilization, CPU utilization and space saving. [43]

	Physical Server	Virtual Server
Space utilization	3U-5U	unwanted
network connections	Exclusive use	Network sharing
CPU utilization	8%-20%	30%-60%
Memory utilization	15%-35%	60%-90%
Fault recovery	1-2 days	Less than 1 hour
Operation and maintenance cost	High cost	low cost
Physical optical disk drive	Exclusive use	Network sharing

Table 6. Resource comparison between physical server and virtual server

2.10 Existing server virtualization frameworks

There are many server virtualization frameworks are proposed such as I.T. Alliance framework, Green IT framework using virtualization, Nyamweya framework on virtualization but any of these frameworks couldn't be used in the implementation of server virtualization in the higher learning institutions, the study noted significant weaknesses which necessitated the development of a new framework. The study critically reviewed each of these frameworks to come up with an appropriate and more practical framework which the higher learning institutions that can adopt in the implementation of server virtualization projects.

2.10.1 I.T. Alliance Framework for Server Virtualization

I.T. Alliance (2010) has an established framework for server virtualization and transformation projects. According to the, server transformation through virtualization provides improved cost of ownership and can deliver cost savings through a reduction in servers and resources, such as space, power and cooling. The I.T Alliance Framework help customers to structure their virtualization program according to established industry best practices.

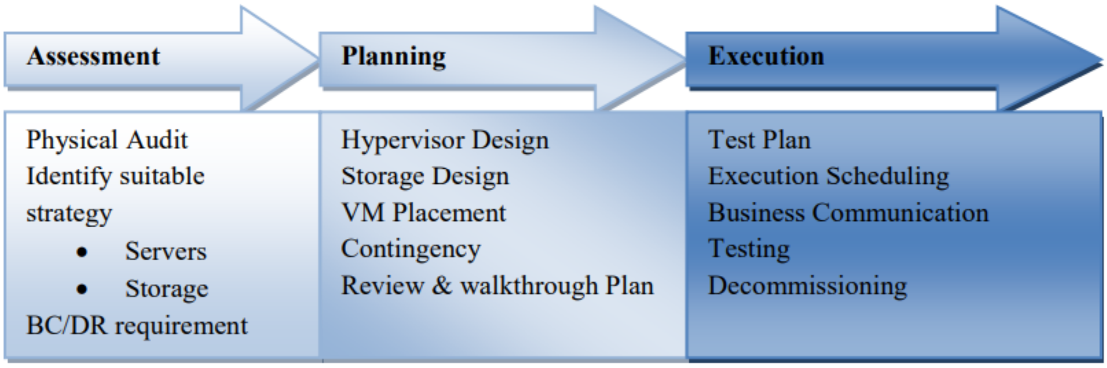

Figure 5. I.T. Alliance Framework for Server Virtualization (I.T. Alliance, 2010)

Generally, each server virtualization project has three distinct phases: Assessment, Planning and Execution. The actual planning and execution activities are dependent on the business needs and services each client requires. Assessment phase involves: Physical Audit, Identify suitable strategy in terms of Servers and Storage, Business Continuity/Disaster Recovery (BC/DR) requirement. Planning phase involves: Hypervisor Design, Storage Design, VM Placement, Contingency, and Review & walk through Plan. Finally, Execution involves: Test Plan, Execution Scheduling, Business Communication, Testing, and Decommissioning. The framework is as illustrated in (Figure 5) but the frameworks don't take into account the staffing of the project, milestones schedule as well as doesn't show the required components also looks a serious of steps than a framework.

2.10.2 Green IT framework using virtualization

The framework as developed on green IT framework using virtualization done at Pakistan. They designed a framework as a serious of steps from developing plan for green data center to measuring the performance in terms of energy efficiency, resource utilization and CO2 emissions. (Figure 6) The ultimate goal of the research is saving huge amount of energy, minimization of CO2 emissions and as a result reducing global warming and other consequences from it using virtualization technology. The researcher uses five steps for proposing the framework using virtualization [44]. The proposed Green IT framework using virtualization and green metrics should be used and followed by data center managers to implement green initiatives in their data center to make it more energy efficient and green, but this framework doesn't show components required for implementations.

The following steps also clearly show the main components included by the researcher:
Develop plan for green data center Under this the researcher shows that green data centers is the result of different challenges from theoretical to realistic manners. In order to develop a plan for green data centers advantage should be identified. The following is the advantage of IT framework using virtualization: lower server storage temperature, better system reliability, better uptime, lower total cost of ownership, maximizes hardware and software utilization, lower heating and cooling cost etc.

Categorize data center in to measurable components Data center is the name given for the collection of server, storage device, UPS, switch gear, computer room air conditioners, pumps, lightening, generator and many others. So measuring the performance of data center is the sum of measuring the real performance of individual components.

Identify green metrics and set components A good metric would be something that measures the efficiency, the sustainability and the cost of a green initiative. The biggest problem any metrics faces when applied for calculating energy efficiency is the lack of standardized system of categorizing different resources of data center.

Identify and implement Virtualization type The type of virtualization is as per the requirement of certain organizations. These can be desktop, server, network, storage, application and operating systems. The virtual machine monitor will be from many different vendors like Vmware, Cisco, Citrix, Hawllet Packard, IBM, Microsoft, NetApp, Novel, Oracle, Quest, Sun, Symantec and virtual Iron.

Measure the performance in terms of energy efficiency and CO2 emissions This stage is used to measure the performance of the data center from time to time using the selected metrics. The phase has the following two sub steps: i. Collect and categorize data (Facilities data, Infrastructure data, IT equipment data) ii. Analysis using energy and CO2 efficiency calculators Using the above five stages the researcher develop the following framework:

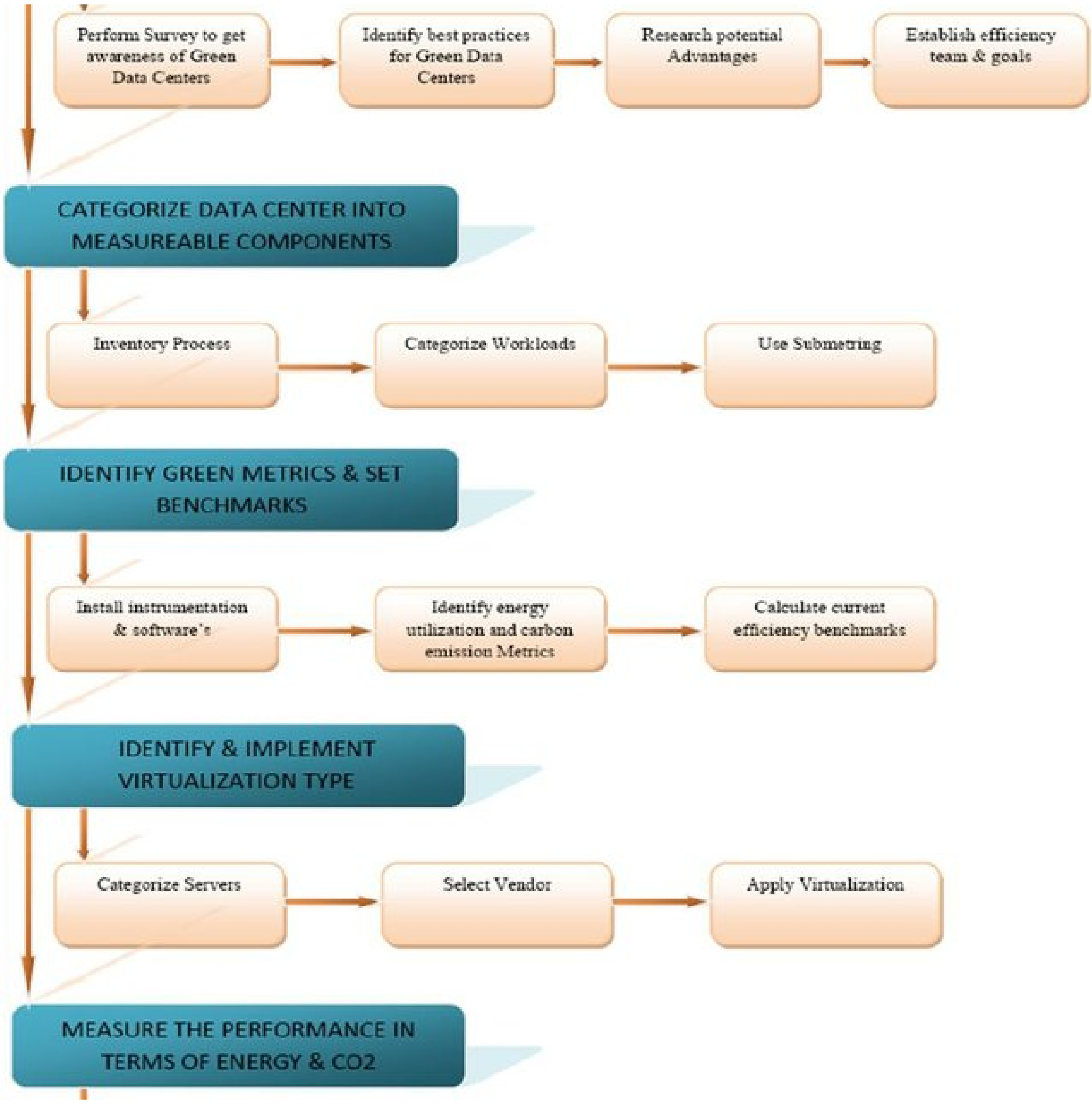

Figure 6. Green IT framework using virtualization

However the developed framework shortages of showing the specific components of a virtualized framework.

2.10.3 Nyamweya framework on virtualization

Nyamweya (2013) develops a framework on virtualization in selected government ministries of Kenya. Exploratory method was used in which the Ministries of Energy and Finance, and also the directorate of e-Government were used as cases. This research aimed at developing a framework to be used for implementation of server virtualization. Exploratory and applied a case study method where the Ministries of Energy and Finance, also the Directorate of e-Government were used as cases , She develops a framework as a series of steps from planning to support however the model is not mentioned the test phase as well as the framework is not self-descriptive. Further research recommended on impact of the Integrated server virtualization framework and assessment of customer satisfaction.

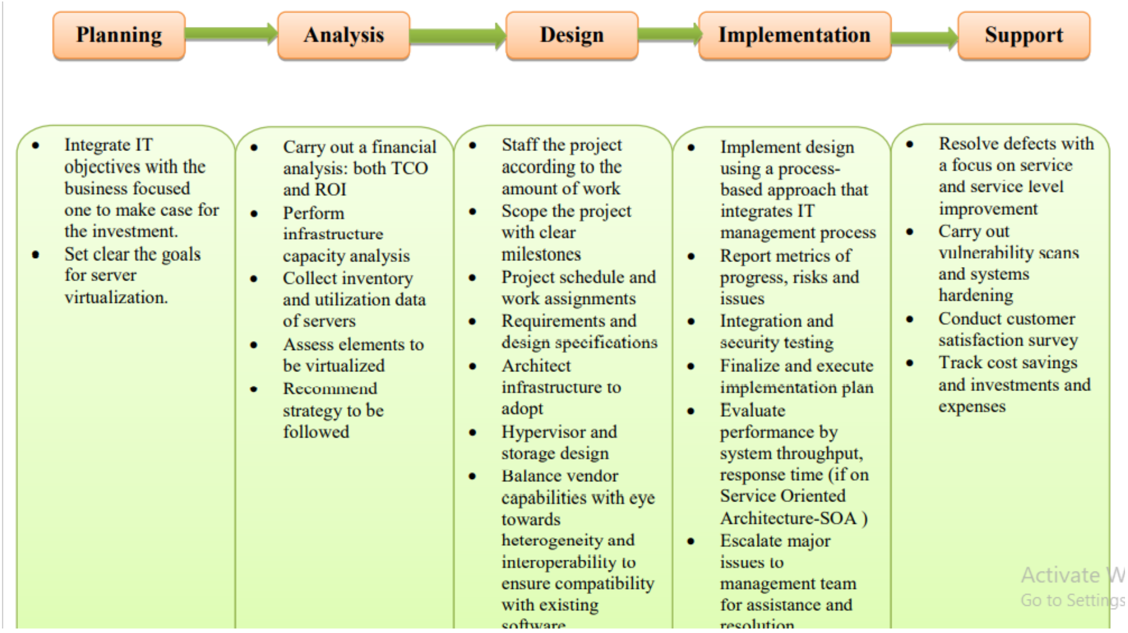

Figure 7. Nyamweya framework on virtualization

2.10.4 Virtualization framework in banking sector

A virtualization framework by Getenet Worku (2017), is developed for the banking sector specifically for Wogagen bank using design science methodology. By reviewing different literature's the researcher proposed a framework which shows different components that fit with the banking industry. The main intention of the researcher is resource sharing and isolation of data between users. His proposed virtualized framework contains five components (user, channel, web interface, what to virtualize/ to whom/ by what and virtualization monitoring). However scope of his research is client virtualization which is different with server virtualization.

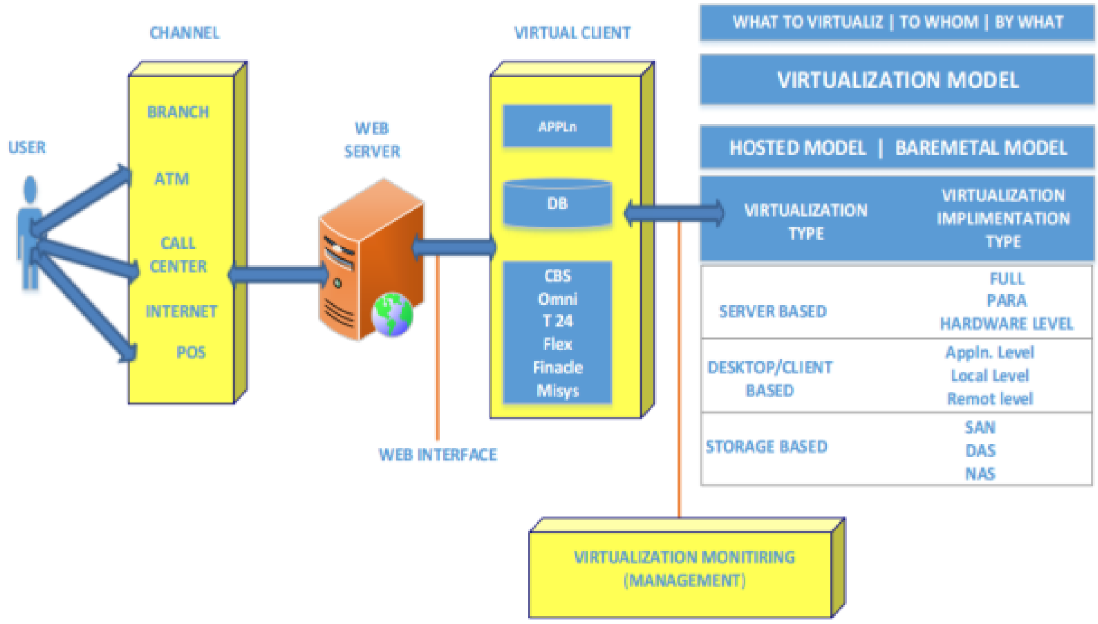

Figure 8. Virtualization framework from [Getenet Worku (2017)]

2.11 Conceptual Framework from Literature Review

A conceptual framework is an interconnected set of ideas (theories) about how a particular phenomenon functions or is related to its parts. The framework serves as the basis for understanding the causal or correlational patterns of interconnections across events, ideas, observations, concepts, knowledge, interpretations and other components of experience [45]. According to Getenet Worku (2017) the variables used to construct the frameworks have bi-directional of three interconnected flows. (Figure 9) Factors of data center virtualization, approaches of data center virtualization and results of data center virtualization.

- **Factors of data center virtualization:** The aim of this variables is to assess the opportunities and challenges of virtualization implementation in the data center.
- **Approaches of data center virtualization:** As data center virtualization encompasses broad ranges of technology and process that enable a data center to operate and provide services. The aim of this variable is used to determine the virtualization model, virtualization type, and virtualization implementation type in relation to the opportunities of virtualization for universities data center.
- **Results of data center virtualization:-** The aim of this variable is to identify the key benefits of Data center virtualization.

Generally the envisaged server virtualization framework borrows from the reviewed frame-

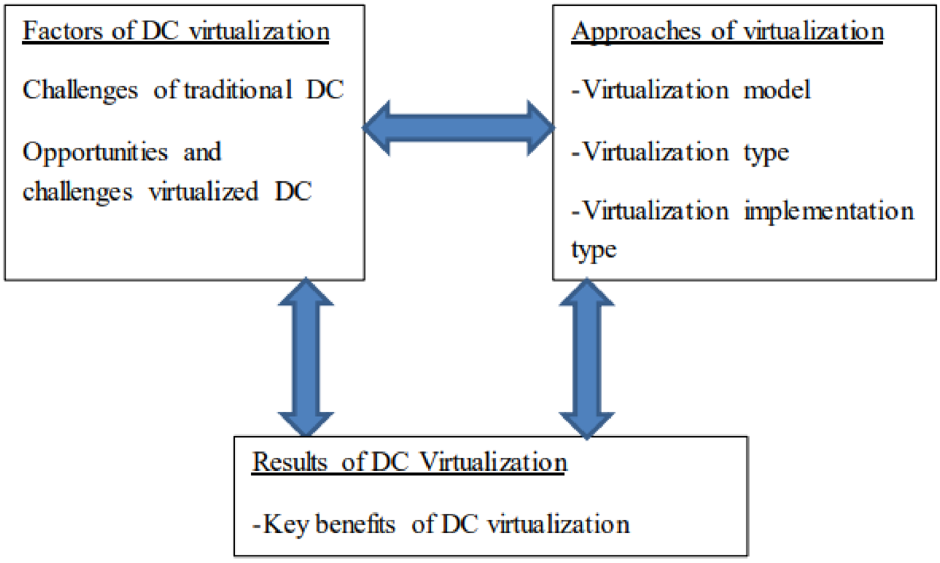

Figure 9. Conceptual Framework

works for server virtualization and IT project phases of analysis, design and implementation. A server virtualization project, like any IT project would be carried out in phases that include the following phases: planning, analysis, design, and implementation. [9]

2.12 Chapter Summary

As we have discussed in the previous sections, there are a number of frameworks that have been proposed in the past. While those frameworks have some common phases and procedures each of them also have unique features and shortcomings.

The IT Alliance framework for virtualization server does not take into consideration the business objectives together with the IT objectives in order to come up with a case why server virtualization should be implemented at the planning phase. It also does not provide for financial capacity analysis. At the planning phase, the framework does not take into account the staffing of the project; milestones, schedule and work assignments. Test and Support for the project is also not taken into account

The work of Getenet Worku on Data center virtualization framework in banking sectors (2017) emphasizes on desktop/client virtualization rather than server virtualization. The integrated server virtualization framework by Nyamweya (2013) fails to describe the implementation model and the framework is not self-descriptive. In addition it doesn't include a testing phase. The testing phase allows us to evaluate the planned implementation, fix bugs and avoid unnecessary disruption in a production environment and Simulate different use cases and evaluate the observed behaviour with the expected functionality.

3. RESEARCH METHODOLOGY

3.1 Overview

In this chapter issues related to the research approach, research design, research method, samples of the study with the selection mechanisms, instruments and procedures of data collection etc. are anticipated.

3.2 Design Science Research Methodology

The objective of this study is exploring the traditional or physical data center infrastructure and developing a virtualized data center framework for HLI data center. A framework is an interconnected set of ideas (theories) about how a particular phenomenon functions or is related to its parts. The framework serves as the basis for understanding the causal or correlational patterns of interconnections across events, ideas, observations, concepts, knowledge, interpretations and other components of experience. So that in order to achieve the specific objectives DSRM was used which is a system of principles, practices and procedures required to carry out a study. IS can draw advantage from DSRM by often using theories from diverse disciplines, such as social sciences, engineering, computer science, economics and philosophy to address problems at the intersection of IT and organizations [46] DS approach was used because; the researcher is intended to develop a framework that guided the HU data center to implement server virtualization.

DS research is highly used in disciplines that are concerned with creation of artifacts. Many IS researchers pioneered DS but little DS researches were conducted in more than a decades. Because of this fact many IS researches are mainly found on the engineering journals than IS journals. Lack of a methodology to serve as a commonly accepted framework for DS research and of template for its presentation may have contributed to its slow adoption. The DSRM incorporates principles, practices, and procedures required to carry out such research and should meet three objectives: it is consistent with prior literature, it provides a nominal process model for doing DS research, and it provides a mental model for presenting and evaluating DS research in ISs. The more widely accepted and also used DS process includes six steps: problem identification and motivation, definition of the objectives for a solution, design and development, demonstration, evaluation, and communication.[47]

Qualitative research methods was used which is an empirical research method which is widely used in social, behavioral, organizational and evaluative research. Most data in qualitative research are not normally in the form of numbers and to be collected from various sources, which involves many techniques such as data description, decoding, translation, etc. to understand their meaning in a natural setting. [48]

DS research is a prescription-driven and problem-solving paradigm that seeks to create viable artifacts in the form of a construct, a model, a method of an instantiation or design artifacts which provide solutions for management problems [49]. According to this authors DSRM comprises of six subsequent activities namely:

- Identify problem and motivate
- Define objectives for a solution
- Design and development
- Demonstration
- Evaluation
- Communication

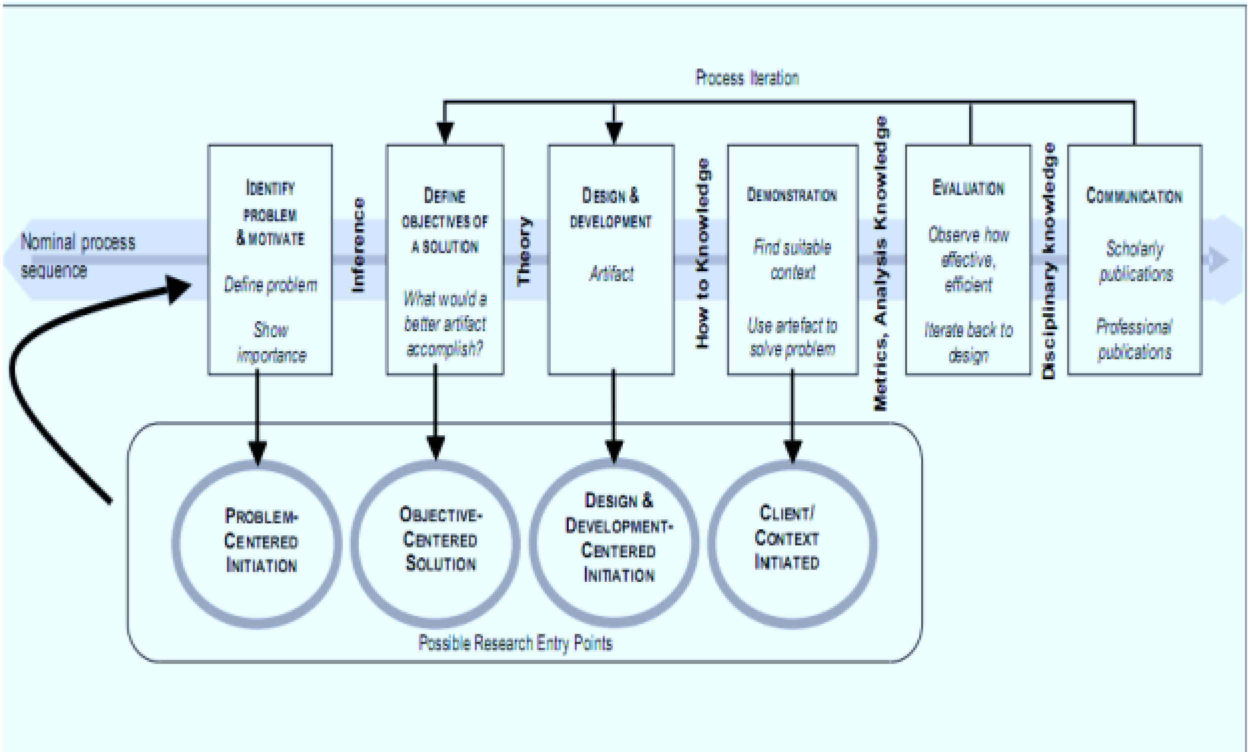

Figure 10. Design science research methodology process models

3.3 Problem Identification

This is the first step in the process model of DSRM. In this stage the specific problems in the traditional or physical data center is critically identified and the researcher strongly work to solve those problems. HU runs physical data center in which a one application for one server architecture is which is not cost effective, difficult for migration, difficult for backup and recovery, take more space, and also difficult for management. After the problems are clearly identified the researcher showed the specific importance of work to solve the existed problems. The researcher's objective is developing an artifact (in this case framework) that can effectively and efficiently provide a solution for the current problem in the data center. In order to come up with the solution, the researcher started with literature review to gather relevant requirements and aspects of existing data center as well as gather appropriate data from the respected bodies. Observation and interview (preferably semi structured interview) are among the secondary techniques. The following description shows how to use those instruments in this thesis.

The study aims to identify the requirements for developing the proposed data center virtualization framework from knowledge of the state of problems. Qualitative objectives are used, which described how a proposed framework is expected to support solutions to problems not previously addressed. After data was collected from respected bodies, it should be narrated and summarized. It involves data preparation, analysis, and finally data interpretation. Qualitative data analysis method is used to analyze the data collected from interview, observations.

Observation Observation is a technique that involves systematically selecting, watching, and recording behavior and characteristics of phenomena at the specific research area. It is one data collection instrument by looking the actual work conducted in the HLI data center. The researcher carefully watches selected (Addis Abeba, Dilla,Bule Hora University as well as Ethernet Data center) data center to gain information on how to use the different information technology infrastructures.

Interview Interview is a form of self-report which is relatively simple method but producing wealth information. [50] The different types of interview, includes structured interview, unstructured interview and semi structured interview. A researcher selected semi - structured interview because it provides information which assures the comparability of data. It also requires a few interviewing skills than does unstructured interview. Concentrated interview with information technology managers, network administrators, database administrators, system administrators, and ICT officers was made for better understanding of the current physical data center infrastructure architecture and trends using well prepared interview guide which is adopted from literatures and modified accordingly. Semi-structured interview and observation are used to gather rich and in depth understanding of practices in the HLI data centers. Qualitative data collection techniques are particularly effective at gaining insight into the processes and events that led up to the observed variation. [51]

Accordingly different IT staffs including IT experts and infrastructure managers in the University were selected for interviews. The researcher interviewed all members (100%) from different groups in the data center which are included in the survey sample because it is believed that those people have a good exposure with data center activities. So that purposive sampling was selected to gather appropriate data from seventeen (17) ICT staffs of the HLI (ICT Director(4), Software Expert(2), System Analysis(2), System Administrator(4), Network Administrator(4),data base administrator(1) (Table 7 shows the information technology staffs involved for gathering data.

Respondents	Total Population	Target Population	Total Selected
ICT Director	4	1	100%
Software Expert	2	2	100%
System Analysis	2	2	100%
System Administrator	4	4	100%
Network Administrator	4	4	100%
Data Base Administrator	1	1	100%

Table 7. Sampling for Interview

3.4 Methods of Data Analysis

The researcher use Qualitative Data Analysis (QDA) method to analysis the data which involves the process and procedures for analyzing data and providing some level of understanding, explanation, and interpretation of patterns and themes in textual data to measures along with searching for patterns of relationship that exist among data-groups. Data analysis is bodies of methods that help to describe facts, detect patterns, develop explanations, processed and analyzed. [52] Describes that data analysis involves the following three steps:

- Preparing the data for analysis,
- Analyzing the data, and
- Interpreting the data.

The researcher analyzed the data collected through expert interview with a focus on problem identification, and determination of type and category of data center services on data center virtualization and performance evaluation of physical (non-virtualized). After relevant information is collected data presentation and interpretation was the necessary next step. The information collected through interview with key informants was organized and narrated. Explanations were supplied for each question along with respondent's justification. Ultimately generalizations were made and presented accordingly for the qualitative data by way of narrating and interpreting the situations.

3.5 Server Performance Requirements:

In order to evaluate performance of server systems we need to take into account the interaction of the various processes involved in processing a request, for software processes and data structures, as well as contention for hardware components (e.g., processors, disks, and networks) and we have to evaluate various hardware resources that is used to support a (client/server) C/S system. [53].

As shown in Section 2.8 there are evaluation of virtualized data center or server performance is being evaluated according to:

CPU utilization: As response times grow exponentially for organizations commonly use processor utilization as a measurement for success, with 5% utilization considered to be inefficient (poor), 25% to be better, 50% to be good, 75% to be very good, and 100% to be best. However, this can be misleading because these are corporate-wide averages. Some percentage of servers will not be candidates for virtualization. These non-virtualized servers are still part of the overall computation for corporate-wide server utilization. High CPU usage, CPU utilization must not be exceeded than 65% as an average. [54]

Memory usage: When the computer is idle, it would use the amount of RAM required by system processes. So that depends on your operating system. For example it should be around 15-20% of your total memory for windows 10. Most linux distributions use less than what windows would use, so that would be around 10% of your total memory. [54]

Disk usage: The most important thing that slows down system is the hard disk usage, if the hard disk usage is 90%, the system will slow down Larger size hard disks have thousands of files which should be properly indexed. [54]

3.6 Design and Development

The conceptual models of DSR must produce a viable artifact in the form of models, methods or frameworks. This stage determines the desired functionality and its architecture and then creating the actual artifact. The knowledge of theory that can be brought to bear in a solution is required to move from objective of a solution to design and development of the proposed framework. Based on the required knowledge of a solution identified in stage one (problem identification and motivation) and stage two (defining the objective of the solution) with the literature review and seventeen (17) expert interviews analysis to develop the proposed framework with the process selected. After determining the data center virtualization framework architecture, the proposed well-mach data center virtualization framework for HU data center with the implementation plan to deploy virtualized data center could be developed.

A framework is a model artifact that provides a broad overview or skeleton of interlinked items which helps as a guide to achieve the specific objective. The researcher used the conceptual framework to inferred objectives rationally from the problem specification. In order to develop the study conceptual framework, researcher started with the literature review to gather relevant requirements and aspects of existing data center frameworks.

As discussed previously data gathering is designed on the basis of the reviewed literature and the questionnaire was distributed to the respondents in terms of interview. The data collection process was administrated by the researcher. Data collections through interview were conducted by speaking to the respondents face to face. Before conducting the interview, the researcher has tried to create conducive atmosphere and explain the factors of data center virtualization. As a result necessary information was collected, organized and processed separately for interpreting and summarizing purpose to produce the major findings. Finally, the researcher proposed a virtualized data center framework.

As shown in Section 2.9 there are a number of frameworks that have been proposed in the past, while those frameworks have some common phases and procedures each of them also have unique features and shortcomings so that the framework is developed based on best experiences from various literatures and possible modifications.

3.7 Demonstration

Demonstration is a part where the study demonstrates the use of the artifact to solve one or more instances of the problem using a prototyping tool. This includes experimentation, simulation and other activity. Resource required for the demonstration includes effective knowledge of how to use the artifact to solve the problem. Virtualization software can convert a single computer into multiple ones. From different types of virtualization software Microsoft Hyper-V is selected. The factors that have contributed to the choice of Microsoft Hyper-V platform are licensing infrastructure familiarity as well, and it was significantly less expensive and built-in to the Microsoft Server 2016 software package, to analyze for the usage and performance of process, memory, CPU, usage of servers for demonstration.

In a demonstration part of this thesis the utilization of CPU, RAM and other scenarios of both the physical servers and the virtualized machines (instances) were measured. After demonstration was completed the measuring values of the two approaches (physical & virtual) machines are compared. For the purpose of taking the results, both GUI and command line interface tools can be used. Non-GUI testing tools can be windows power shell, windows management instrumentation etc. On the contrary server manager, task manager, performance monitor and resource monitors are among the GUI tools to take and compare the utilization of physical and VMs on the required components explained above. The researcher selected the graphical user interface tools called resource monitor to capture the utilization of physical machines and the virtual instances. Experiment was done to check the resource utilization of the virtual instances and determine the number of instances to be deployed on a single physical server.

The researcher select one physical servers has RAM 16GB, 2.53 GHz Processor,storage capacity 1000GB which is running one of the application in the university data center. (Figure 18). The researcher deployed three VMs on the same machine running three distinct applications and each servers use 64GB Hard disk, 4000MB of memory with dynamic memory allocation is enabled which means the amount of memory available to this VM to be dynamically within the range between minimum 512MB and maximum 1048576MB and with shared physical servers network. (Figure 17).

3.8 Evaluation

The study aims to observe and measure how well the framework supports a solution to the problem. This activity involves comparing the objectives of a solution to actual observed results from use of the framework in the demonstration. It requires knowledge of relevant metrics and analysis techniques. Depending on the nature of the problem setting and the artifact, evaluation could take many forms, including comparison of artifact‘s functionality with the solution objectives this study took the qualitative evaluation analysis techniques which include a comparison of the proposed framework functionality with the solution objectives from activity two above, objective qualitative performance measures, such as business impacts, and business continuity extents of use and user satisfaction as the result the researcher present the developed framework to ICT staff and got good feedback from respected experts.

4. THE PROPOSED VIRTUALIZATION FRAMEWORK

4.1 Selected Virtualization type for University data center

Virtualization software allows multiple OS and applications to run on the same server at the same time, and, as a result, lowers costs and increases efficiency of a company's existing hardware. There are different types of virtualization which is selected according to the need of an organization. There are three known server virtualization methods. Full virtualization, para-virtualization and hardware-assisted virtualization [55].

Full virtualization is a common and cost-effective type of virtualization, which is basically a method by which computer service requests are separated from the physical hardware that facilitates them, even the guest OS did not aware that it is run as a virtual environment and did not require modification. But, in para-virtualization method the guest OS should be modified which creates portability and compatibility issues. Even non modified OS (like windows) cannot be supported by para virtualization.

To detect the required virtualization solutions that can be implemented for universities , the researcher started by interpreting the requirements of university data centers staff by means of interview to have better understanding on the idea, experience and best practice. The interpretation of the interviewee the researcher identifies the types of virtualization model which are server virtualization types as a result full virtualization is selected to be deployed in the HU data center due to its many advantage over para virtualization. The main advantages of the full virtualization includes isolating users from each other, ease of emulation of different architectures, sharing a computer system among multiple users and emulating new hardware to achieve improved reliability and security. [56]. Additionally it allowing procedures for migration and portability as the same guest OS that can run in the virtualized environment.

4.2 Proposed Server Virtualization Framework for HLI data center

After the collected data from interview and observation is analyzed proper data center virtualization framework should be designed for the HLI data center. As previously mentioned the objective is to use server virtualization in order to migrate from the current bare metal based services. Such installation of services consume huge energy without performing useful work. In an average server environment, 30% of the servers are "dead" only consuming energy, without being properly utilized [57]. By implementing server virtualization the University data center servers can share CPU, memory, storage, and networking capabilities, which are pulled from the hypervisor of the physical server. Additionally services can get scalability, availability, better up-time, easy migration, sustainable backup and recovery and many more.

Before implementing server virtualization in any firm it is important to seriously plan and consider virtualization risks associated with it. It is also important for the data center to check whether it has the necessary infrastructure capabilities to handle the virtualization tasks. In addition, failure of single consolidated server should be considered, because it is handling the workload of multiple applications since if virtualization is not planned properly, the effort will have a high probability of failure. In order to properly implement virtualization there is a need to answer some of these questions:

- Types of virtualization technologies exist?
- What is cost/benefit ratio of virtualization?
- What new challenges it will bring to business firms?
- Structure of virtualization solution being implemented?
- Which applications or services are good virtualization candidates?
- Server platforms best suited to support virtualization?

By considering the knowledge derived from the literatures and the universities intention from the interview result the researcher decided to propose a data center virtualization for educational institutions. The proposed model defines that, the process of virtualization should be structured and designed in such a way that it must fulfill the necessary requirements and should be within the scope & infrastructure domain already installed in the data center. It is therefore much more than simply loading a virtualization technology on different servers and transforming one or two workloads into VMs. Rather it is a complex and rigorous process that need to be implemented and monitored properly.

So the proposed server virtualization is well-matched server virtualization framework takes care of all the weaknesses identified in the review of the existing frameworks and integrates the general IT project phases. The phases of the proposed well-matched server virtualization framework will then include: planning, analysis, design, implementation, test, deployment and continuous support.

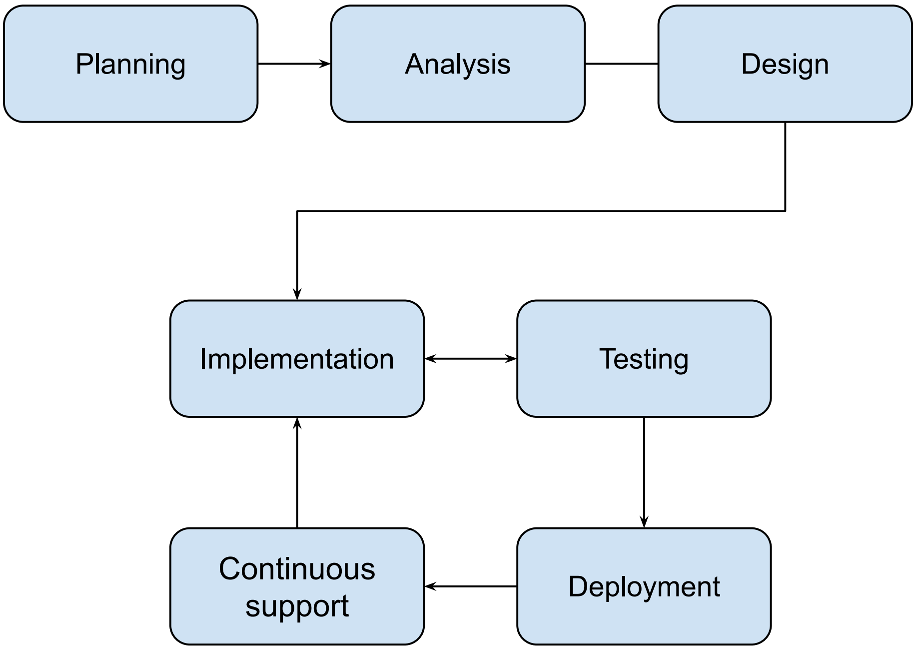

Figure 11. Proposed Well-matched Server Virtualization Framework for HLI Data center

Figure 11 shows the developed framework for HLI Data center. The developed framework is not generic but, can be implemented in similar sister Universities.

Planning:-

- Integrate IT objectives with the business focused one to make case for the investment
- Conduct a study to determine the current state
- Set clear the goals for server virtualization

Analysis:-

- Carry out a financial and infrastructure capacity analysis
- Collect inventory and utilization data
- Conduct Return on Investment (ROI) analysis
- Assess elements to be virtualized
- Collect inventory and utilization data of servers
- Recommend strategy to be followed

Design:-

- Staff the project
- Scope the project with clear milestones

- Project schedule and work assignments
- Requirements and design specifications
- Architect infrastructure to adopt
- Hypervisor and storage design
- Balance vendor capabilities with eye towards heterogeneity and interoperability to ensure compatibility with existing software

Implementation:-

- Implement design using a process-based approach that integrates IT management process
- Use automation tools to prepare reproducible deployment workflow
- Report metrics of progress, risks and issues
- Finalize and execute implementation plan
- Escalate major issues to management team for assistance and resolution
- Document virtualization process

Test:-

- Simulate different use cases and evaluate the observed behaviour with the expected functionality
- Integration, compliance and security testing
- Fix any observed issues

Deployment:-

- Execute the implementation procedure in a production setup
- Optimize services based on specific needs

Continuous Support:-

- Resolve defects with a focus on service and service level improvement
- Carry out vulnerability scans and systems hardening
- Conduct customer satisfaction survey
- Track cost savings and investments and expenses
- For upgrade, security patch or addition of new service, go back to implementation phase and repeat the process

4.3 Evaluation of the proposed framework for server virtualization

After the required framework is developed it is required to be evaluated by different ICT experts. The developed framework has evaluated in two phases. In the first phase, it was presented for ICT staffs of HU for further and in-depth evaluation. The second phase of evaluation is proceed by the researcher to test the framework. The researcher take one server from HU data center and build virtualized infrastructure to host three services. These services are HU-Ebook, HU-File Server and HU-Website. These are the same services used in the university digital infrastructure. The evaluation was performed by measuring and comparing the resource utilization (CPU, memory, and disk usage) of virtualized infrastructure with non-virtualized setup.

4.3.1 Planning:

Planning is the preliminary task for undertaking different activities and there were clear set goals for server virtualization that included Need assessment current physical data center infrastructure. The researcher plan to deploy three system service that will run on three virtual servers in single physical server.

4.3.2 Analysis:

The researcher make inventory of server with their utilization by providing server resources to the various universities data centres system services to supporting cost-saving initiatives, which was achieved by virtualizing server. The selected server capacity analysis that was performed indicated that the servers had four processors, RAM of 16GB, a processing speed of 2.53 GHz per processor, four network ports, and one hard disk drive with a capacity of 1000GB. Furthermore the servers had low levels of utilization, with average CPU usage at 5%, physical memory usage of 15%, 32% of hard disk space in use.

Selected Host server system information:

Using virtualization to create and run VMs with server OSs require powerful hardware with processor virtualization extensions. Used server met all necessary hardware requirements, as shown in Table 8

OS Name	Windows Server 2016
System Type	x86-64
Processor	Intel(R) Xeon(R) CPU E5630 @ 2.53 GHz 2.53GHz
Installed physical memory (RAM)	16GB
Hyper-V - VM Monitor Mode Extensions	Yes
Hyper-V - Second Level Address Translation Extensions	Yes
Hyper-V - Virtualization Enabled in Firmware	Yes
Hyper-V – Data Execution Protection	Yes

Table 8. Host Server specifications

4.3.3 Design:

The virtual infrastructure is built based on one server machines Dell PowerEdge R740 Rack Server, connected with 1G Ethernet network and running windows server 2016 OS. Over this OS the Microsoft 2016 Hyper-V server are installed. The study used one physical server to host the three systems . Three VMs were created using Hyper-v as the hypervisor. The VMs were then loaded with Microsoft Server 2016 OS.

4.3.4 Implementation:

Three web-based system were then hosted on the VMs just as they were on the physical servers. When the server was powered on and the systems were running, the utilization of the physical server after the implementation of server virtualization, CPU usage increased from 5% to 66%, physical memory usage increased from 4% of 16GB to 83.33% of 16GB, and capacity of hard disk space in use was 67.4% up from 32%.

4.3.5 Deployment:

As previously discussed the numbers of VMs per physical server was depends on number of workloads so that the optimize services based on specific needs of universities. Let's start from a physical machine running server 2016. Now follow the following to install Microsoft hyper-v as a window role.

1. Open Server Manager, this can be found in the start menu. If it's not there simply type "Server Manager" with the start menu open and it should be found in the search.
2. In the Roles Summary area of the Server Manager main window, click Add Roles.
3. On the Select Server Roles page, click Hyper-V.
4. On the Create Virtual Networks page, click one or more network adapters if you want to make their network connection available to VMs.
5. On the Confirm Installation Selections page, click Install.

6. The computer must be restarted to complete the installation. Click Close to finish the wizard, and then click yes to restart the computer.
7. After you restart the computer, log on with the same account you used to install the role. After the Resume Configuration Wizard completes the installation, click Close to finish the wizard.

Now, it is time for creating virtual instances using the following steps

1. Open Hyper-V Manager. Click Start, point to Administrative Tools, and then click Hyper-V Manager.
2. From the Action pane, click New, and then click VM.
3. From the New VM Wizard, click Next.
4. On the Specify Name and Location page, specify what you want to name the VM and where you want to store it.
5. On the Memory page, specify enough memory to run the guest OS you want to use on the VM.
6. On the Networking page, connect the network adapter to an existing virtual network if you want to establish network connectivity at this point.
7. On the Connect Virtual Hard Disk page, specify a name, location, and size to create a virtual hard disk so you can install an OS on it.
8. On the Installation Options page, choose the method you want to use to install the OS (Install an OS from a boot CD/DVD-ROM. You can use either physical media or an image file (.iso file).) After you create the VM, you can start the VM and install the OS.

In the final step of this process, you connect to the VM to setup the OS. As part of the setup, you install a software package that improves integration between the virtualization server and the VM and after you have completed the setup and integration services are installed, you can proceed to test the VM by customizing it to suit your testing goals. For example, you can view or modify the virtual hardware that is configured for the VM. From the VM pane, right-click the name of the VM you created and click Settings. From the Settings window, click the name of the hardware to view or change it.

Create a vSwitch for Hyper-V VMs : There is no pre-configuration of vSwitch during Hyper-V setup. If you attempt to create a VM right after the set-up process, you won't be able to connect it to a network. To set up a network environment, you'll need to select vSwitch Manager in the right pane of Hyper-V Manager:

1. Open Hyper-V Manager, select the Hyper-V host Server name.

2. Select Action > vSwitch Manager.

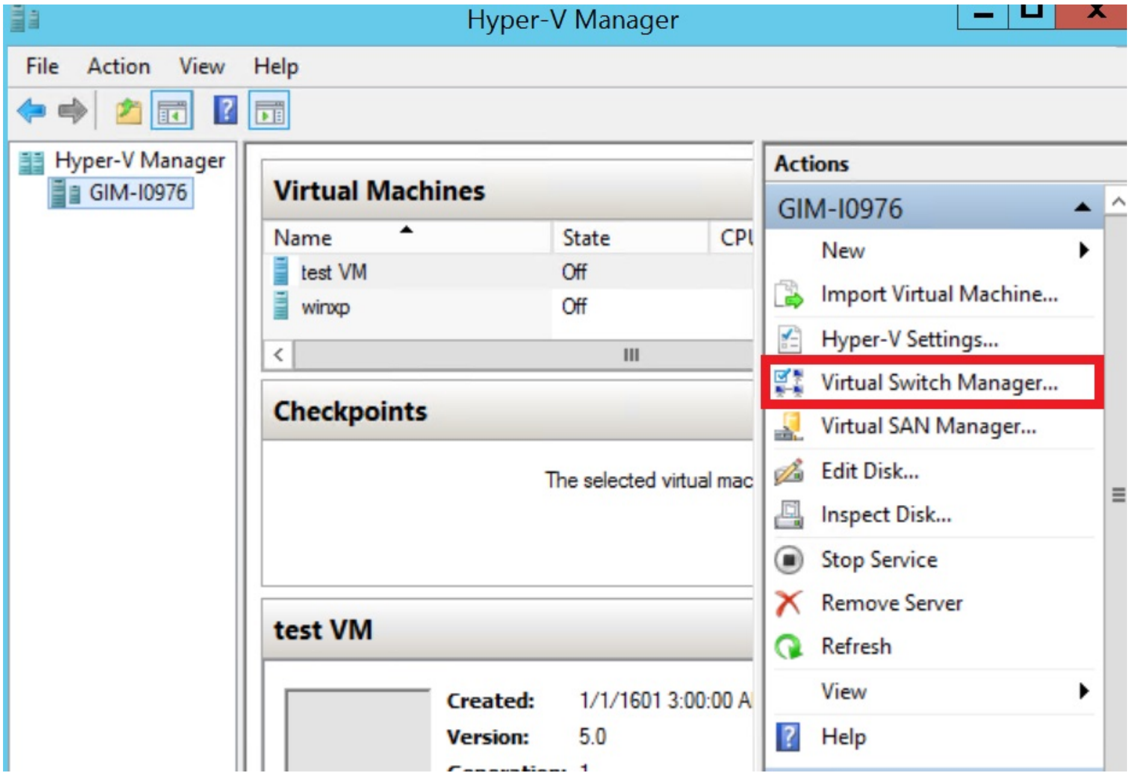

Figure 12. Hyper-V Manager

3. Choose the type of vSwitch you want and add a name for the vSwitch.

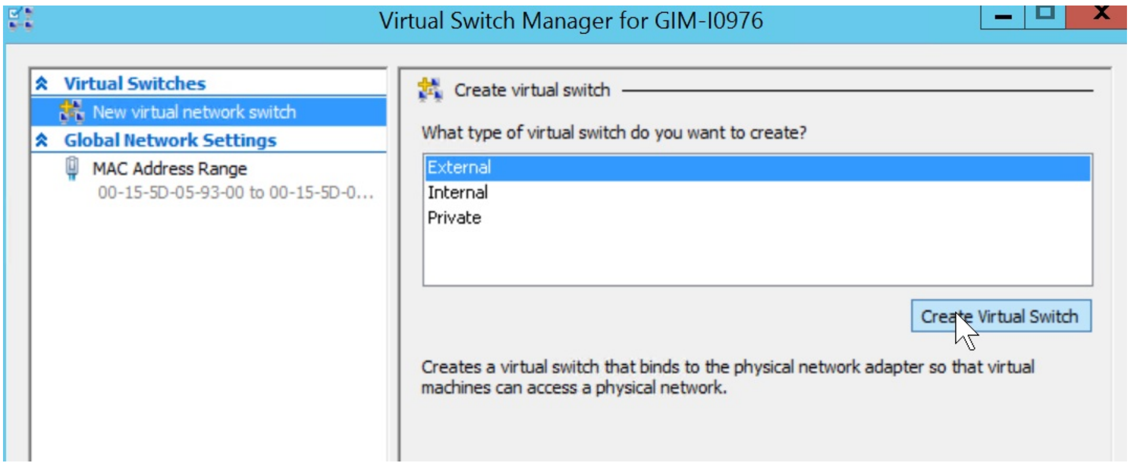

Figure 13. vSwitch Manager

There are three vSwitch types available External vSwitch, Internal vSwitch and Private vSwitch.

External vSwitch: will link a physical NIC of the Hyper-V host with a virtual one and then give your VMs access outside of the host, meaning your physical network and internet (if your physical network is connected to internet)

Internal vSwitch: should be used for building an independent virtual network when you need to connect VMs to each other and to a hypervisor as well.
Private vSwitch: will create a virtual network where all connected VMs will see each other, but not the Hyper-V host. This will completely isolate the VMs in that sandbox.

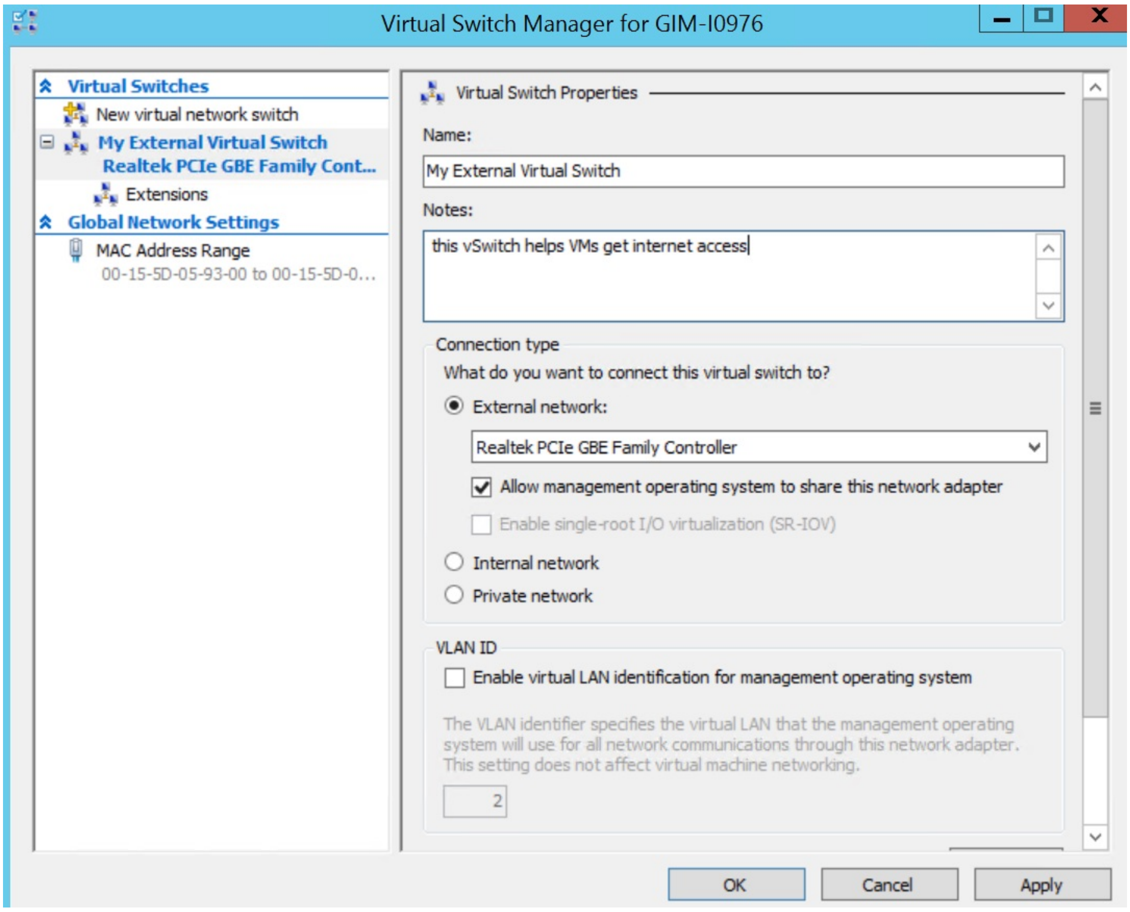

Figure 14. External vSwitch creation

Once you click the Apply button, be prepared to lose physical connectivity for a moment while Hyper-V needs to turn the physical NIC off, configure the vSwitch and turn both on.

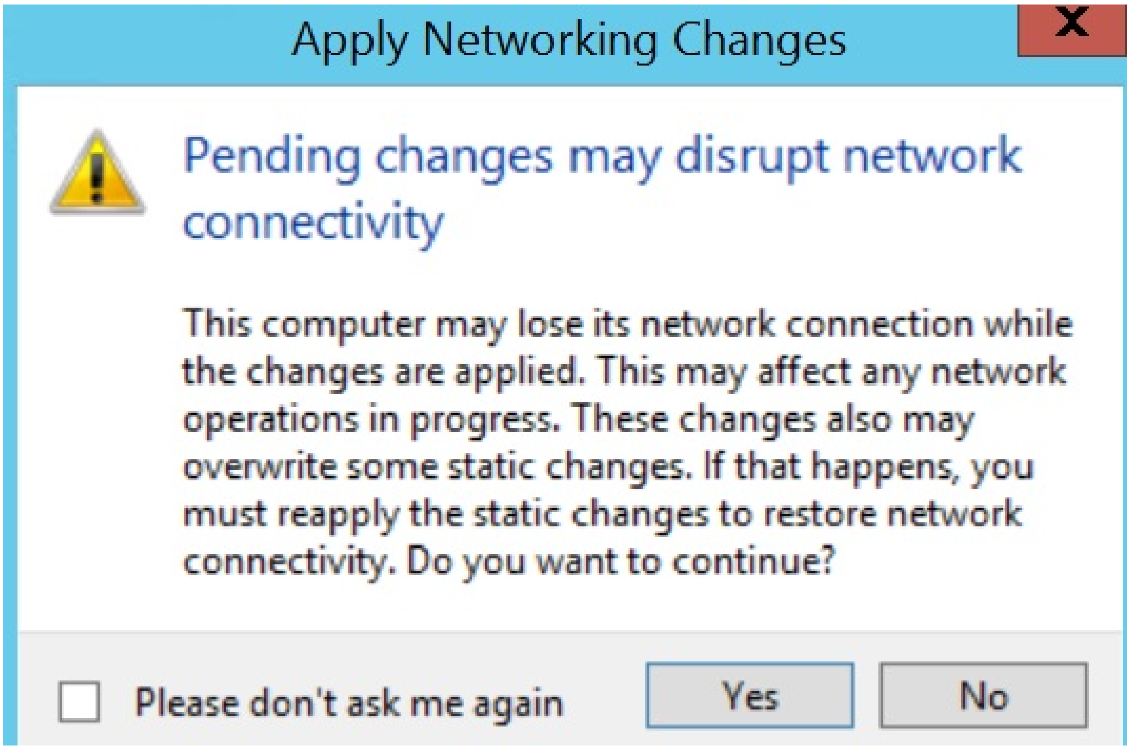

Figure 15. External vSwitch creation warning

4. VM network connectivity : Be informed, that VM connected to an internal vSwitch or private vSwitch will get IP address automatically if only DHCP server is present at the same virtual network. If there is no DHCP server, perform some post-configuration for VMs connected to a private vSwitch.
   Finally go to hypervisor OS Network Connections and find the connection related to your internal vSwitch. Configure the static IP address and sub net mask manually.

4.3.6 Test:

The test phase similar to deployment phase but it allows to check a system/program will behave in a product environment and safe places to find and fix environment-related bugs before moving into production environment. In this study the researcher use only production environment since there were no free servers in the university data center for testing, it was difficult to do the experiment in develop environment.

4.3.7 Continuous Support:

The defect still needs to be fixed, but during the implementation the researcher and the web administrator ensured that all the systems (HU-Ebook, HU-File Server and HU-Website) were running without unnecessary downtime.

4.4 Chapter Summary

In this chapter the virtualization framework is developed for the University data center which believes to solve the existing problems. The proposed well-matched server virtualization framework results discovered an increase in CPU usage, physical memory usage and capacity of hard disk space. The next chapter will be the researchers result and discussion of study.

5. RESULT AND DISCUSSION

5.1 Introduction

The preceding chapter discussed the proposed framework for this study. The intention of this chapter is to present, interpret and analyze data obtained from the sources. The respondents of the study have a direct relation with the data center as well as virtualization technologies (ICT director, Software Expert, System Analysis Expert, network administrators and System administrators, senior system administrators) so that they play key informant interviews with IT experts to propose the virtualized data center framework and to assess how this framework helps on the service delivery performance and effectiveness in the higher educational institution services. The researcher had direct involvement in working in universities data center as well as the researcher was able to develop close working relationships with IT professionals who represented in ICT departments. (Table 9) shows biographic information of respondent.

Educational level & work experience	**Frequency**	**Percent**
Educational level		
MSC	7	41.18%
BSC	10	58.82%
Total	17	100%
Work Experience		
> 7 years	15	88.24%
>= 5 years	2	11.76%
Total	17	100%

Table 9. Respondents Educational Level and Work Experience

What makes data center virtualization attractive?

In the physical infrastructure or physical server each server has its own hardware: Memory, network, processing and storage resources. On this hardware, the server OS is loaded. From the OS you can then run the applications and each physical server includes memory, processor, network connection, hard drive, and an OS for running programs and applications. But with a virtual infrastructure A virtual server operates in a “multi-tenant” environment, meaning that multiple VMs run on the same physical hardware), In this case you have the same physical server with all the resources, but instead of the server OS, there’s a hypervisor such as vSphere or Hyper-V loaded on it. The hypervisor is where you actually create your VMs.

Many IT organizations must buy and use multiple servers because of limited x86 servers. However, each uses storage capacity to compensate for the current high storage and support settings. Inefficiency and high operating costs occur. So, solutions with virtualization can be an option. Virtualization is needed in software to simulate device functions and create virtual computer systems. Virtualization takes place in a virtual version of compilation of something created instead of the actual version. The benefits generated include economies of scale and greater efficiency. So that virtualization is a key technology used in data center to optimize resources. As IT needs continue to evolve, virtualization can no longer be regarded as an isolated technology to solve a single problem. The factors that make data center virtualization attractive options for educational institutions:

Efficiency Cost of New Hardware: High requirements for hardware along with the large amount of data stored in a company / institution by using virtualization methods in cloud storage, it can be overcome. You can store more data without having to buy more hardware. In the end this will reduce the amount of costs you have to spend to buy new hardware. But the cost goes far beyond hardware – lack of downtime, easier maintenance, less electricity used. Over time, this all adds significant cost savings.

Reduce Cooling System: The large amount of hardware that works on data center facilities makes the room heat fast. Because the more hardware you need, the more cooling you need to keep the temperature stable. Virtualization can be used in the data center to reduce the amount of hardware usage. This will also impact on reducing air conditioning requirements.

Easier Migration to the Cloud: By moving to a virtual machine, you are closer to a cloud environment. You can even reach a point where you can use a VM to and from your data center to create a strong cloud-based infrastructure. But outside the actual virtual machine, virtual technology makes you closer to the cloud-based mindset. This will help you migrate more easily.

5.2 Interview Interpretation

Data centers are the heart of any organization so that the analysis of the collected data is done by the researcher. The target of this interpretation is to present the views and ideas of the respondents in regards of the data center virtualization. The main components that are needed to be analyzed and interpreted from the interview for the development of data center virtualization in higher learning institutions are:

- Data center architecture in HLI currently traditional (physical)

- Reason for implementation of virtualized data centers
- The role of IT infrastructure staffs in virtualized data center
- Benefits of virtualization

5.3 A need for server virtualization

The researcher established the server virtualization need by seeking to bring out the levels of server utilization, the trend of server hardware purchase in the universities will increase year to year since the university data center have more servers than need, for single system installation prior acquisition are takes place before the servers are deployed. This meant that whenever there was a new system to be deployed on a server, a new one was bought for the same purpose which can expenditure the universities.

Higher education institutions must adopt data center strategies that are flexible and scalable to support continued growth, while staying within budget however, Ethiopia higher education institutes including HU build data center traditionally so this does not protect the investment costs of ICT that affect the financial efficiency of institutes. Generally almost all respondents agree that the current data center is traditional and has various problems. Hence virtualization should be implemented in the institutions data center.

The researcher find that from interview the investments of data center virtualization, the higher education institutes important to deliver a solution that is cost effective, reliable and scalability of institutions. All respondents agreed that high investment cost on IT infrastructures, inefficient use of IT resources.

5.4 Existing server virtualization frameworks

All (100%) respondents indicated that there was no framework for server virtualization that was being used in the University data centers. This is also clearly indicated from the levels of server utilization are very low CPU usage.

In regards to the cost suffered on the existing traditional (physical) ICT infrastructure to support the educational services, many respondents agreed that very expensive. All of the respondents (100%) revealed that the trend of server hardware purchase in the universities is on the increase and all the respondents (100%) indicated that the universities will benefit from server virtualization. Generally the suggestions from all respondents on virtualized data center indicated that, compared with the existing IT infrastructure to support educational services, using virtualized data center in higher education institutes is efficient to perform its activities at lower costs. Two respondents of both Addis Abeba University and Ethernet data center IT expert stated their data center starting virtualizes server but this was made by personal motivation of staffs exert not in planed infrastructure. Also the respondent point the current data center have many challenge like power fluctuation, poor cooling system, security camera and access door device not functioning well, professional (system administrator) turnover, periodic maintenance, not well-organized servers like backup weak and other, additionally this respondent also mentioned that the university data center don't have a user management server. But all the respondents indicated that the universities will benefit from data center virtualization and a hundred percent of the respondent pointed out that none of the server virtualization frameworks such as I.T. Alliance framework, Green IT framework and others were not being used in the universities data centers.

Observation is also one data collection instrument by actually looking what is happened in the university data center. The data center is traditional and did not have any standard design. The data center did not have a free space for future expansions and still uses a normal key for security, which is not recommended in modern data cenetr standard.

The other issue the respondents gives two reasons on the implementation of data center virtualization technology, first by explaining the insufficient of traditional data center and secondly, by explaining the technical advantage of data center virtualization. In case of its insufficiency they try to express in terms of the traditional data center focus on business and application system stability, safety and reliability. But, the infrastructure ignores considering resource utilization rate which is too low, the poor business continuity and the high operational energy cost. The other issue in terms of its technical advantage is that the virtualization technology improves the efficiency of resource utilization and service levels, releases the potential of the data center, and brings great operational flexibility to the data center. Virtualization technology is not just a simple integration tool, but it introduces more advanced automation, system management is more agile in supporting the business development to the data centers.

5.5 Sizing server hardware for virtual machines

The types of applications that are running on VM determines the FAQ that is "how many VMs can you fit on a single physical host or how many VMs can handle host server". For example, servers that have very light resource requirements, such as web, file and print servers, and servers with medium to heavy resource requirements, such as SQL and Exchange servers are not equal in running VMs. Nowadays, almost all servers come with two or four cores per physical CPU. A good rule of thumb is that four single CPU VMs can be supported per CPU core. This can vary by as much as 1-2 per core, and up to 8-10 per core based on the average CPU utilisation of applications running on VMs. [58]

The three elements to consider when selecting virtualization hardware include the CPU, memory, and network I/O capacity. They're all critical for workload consolidation.

Processors. A processor can potentially operate anywhere from three to five VMs per core, though a conservative rule of thumb is two workloads per CPU core, so plan your server resources accordingly. For example, if a server will be running four workloads with headroom for another four, you can probably deploy the server with four dual-core or two quad-core processors. Existing processors need to be upgraded if they are too old to support the planned workloads, or you demand hardware-assist features tailored for virtualization, technologies or hardware acceleration for virtual workload migration. [59]

Memory. A server must provide ample memory space for the OS/hypervisor and every virtual machine. The amount of memory needed can vary dramatically between virtual workloads, so there is no guideline for memory. "I would be surprised to see a virtualized server going out with less than four gigs [GB] of memory," Sobel said. "That's not to say it's not possible.". Still, it is critically important to avoid memory shortfalls, otherwise the affected VMs will page excessively to disk – significantly reducing performance and risking application crashes. Memory upgrades are fast and easy, and require little more research than matching the current module type. However, memory upgrades are limited by the number of free DIMM (dual in-line memory module) slots on the server's motherboard. Large (high capacity) memory modules can be expensive, and it is far more costly to replace existing RAM modules when free memory slots are scarce. [59]

Disk I/O. Servers typically include some amount of local disk storage. As more virtual applications contend for the disk space, I/O activity (measured in IOPS or I/O Operations per Second) can quickly become a bottleneck resulting in poor application performance. "It [a virtual machine] is all now just a big file against the disk," Sobel said. "Disk I/O is by far the biggest constraining factor with server virtualization." A new disk controller and corresponding hard drives can usually mitigate disk I/O to a large extent. For example, a small contingent of basic SATA drives on the server can be replaced with a Wide Ultra-2 or Ultra-3 (Ultra160) SCSI controller and drives. The cost for such a drive upgrade is usually quite reasonable, but the effort and risk involved with migrating data from the old drive set to the new drives may prove problematic. Some organizations avoid use of local storage on virtual servers, instead committing all VM storage needs to iSCSI (Ethernet) or Fibre Channel SAN storage. [59]

Network connectivity. Applications also need to access the LAN, so there should be adequate bandwidth available for all VMs hosted on a server. Advanced connectivity such as 10 Gigabit Ethernet (10 GigE) is rarely necessary given the current state of technology, but it is possible to aggregate two or more common 1 Gigabit Ethernet (1 GigE) NIC ports. Noncritical VMs can leverage multi-port NICs and risk the card as a single point of failure, while critical workloads can utilize separate NICs to achieve redundancy at the cost of additional expansion slots on the server's motherboard. [59]

Generally determine what you want to achieve with your virtual infrastructure and the types of workloads you run when choosing a CPU for virtualization. More memory and more storage help with server consolidation as well, but size memory and storage according to your needs to prevent wasted physical resources. Finally, ensure you have enough network bandwidth to accommodate your virtual workloads. The current HU data center servers have almost the same CPU and Memory size so that the researcher take choose one of server from data center for lab simulation, but for virtualization its better to select the server that have high CPU,RAM and hard drives.

5.6 Performance Analysis Experiment

5.6.1 Configuration of VMs

Verifying a theory with an experiment is a good supporter for the study conducted. Related literature's shown that VM are created using abstracting software called hypervisor to analyze the usage and performance of CPU, process, memory usage of servers without virtualization and with virtualized servers Hyper-V from Microsoft was selected to the experiment for the purpose of creating different VMs or servers on one physical server. Before creating each VM first by installing Hyper V roles in windows server machine and then enable client Hyper V. From Hyper-V manger we have VM wizards that provide a simple and flexible way to create a VM. Microsoft Hyper-V is installed directly on the physical machine as a role of Microsoft windows server. Hyper-V can be used as a stand-alone product or an integrated part of windows server and it can be used efficiently to run multiple OSs, Windows, Linux and other in parallel on a single server. Proper Hard disk space and memory size should be allocated for the VMs. In this research the researcher configure Hyper-V as one role of the Microsoft server 2016 shown in (Figure 16)

As the researcher mentioned above proper attention should be given while configuration of hyper-V. Hence the researcher configures hyper-v properly and the three system

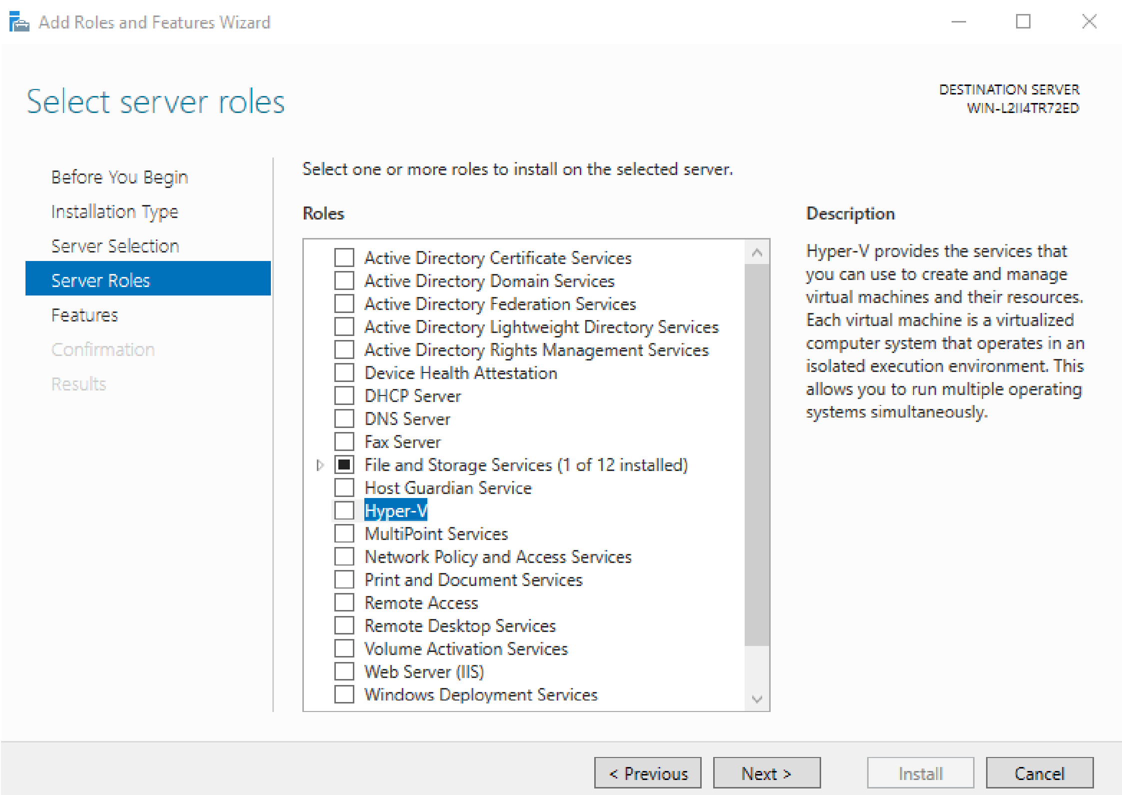

Figure 16. Hyper-V installation

applications used in the HU data center namely HU-Ebook, HU-Website and HU-File Server or samba server are created as servers with OS Windows Server 2016, Ubuntu 22.04.1 LTS and Windows Server 2016 are installed in each VMs respectively. In this experiment the physical memory, CPU and hard disk is given for the individual VMs but network is left to be shared. As indicated on above section the resource allocation for each virtual system needs to be planned carefully. If very less resource is allocated, the application performance might be affected and if too much resource is allocated, it will result in under-utilization. The allocation of the resources depends on the workload that the specific server performs. The user process communicates with the server which intern communicates with the actual database to implement what is required. Figure 17 shows the overview virtually created using the wizard that have the physical server with a capacity of 4 core processor, 16GB Memory and 1TB HD capacities. After VMs has created , it is necessary to configure the hard disk, memory and processor to provide the appropriate computing resources for the workload that is planned to run on the Vm.

5.6.2 Performance Comparison Test

In addition to interview the researcher assessed to analyze the usage and performance of CPU, disk and memory usage of servers with and without virtualization. Hyper-V, which is Microsoft virtualization platform, is used to build virtualization environment. Hyper-v

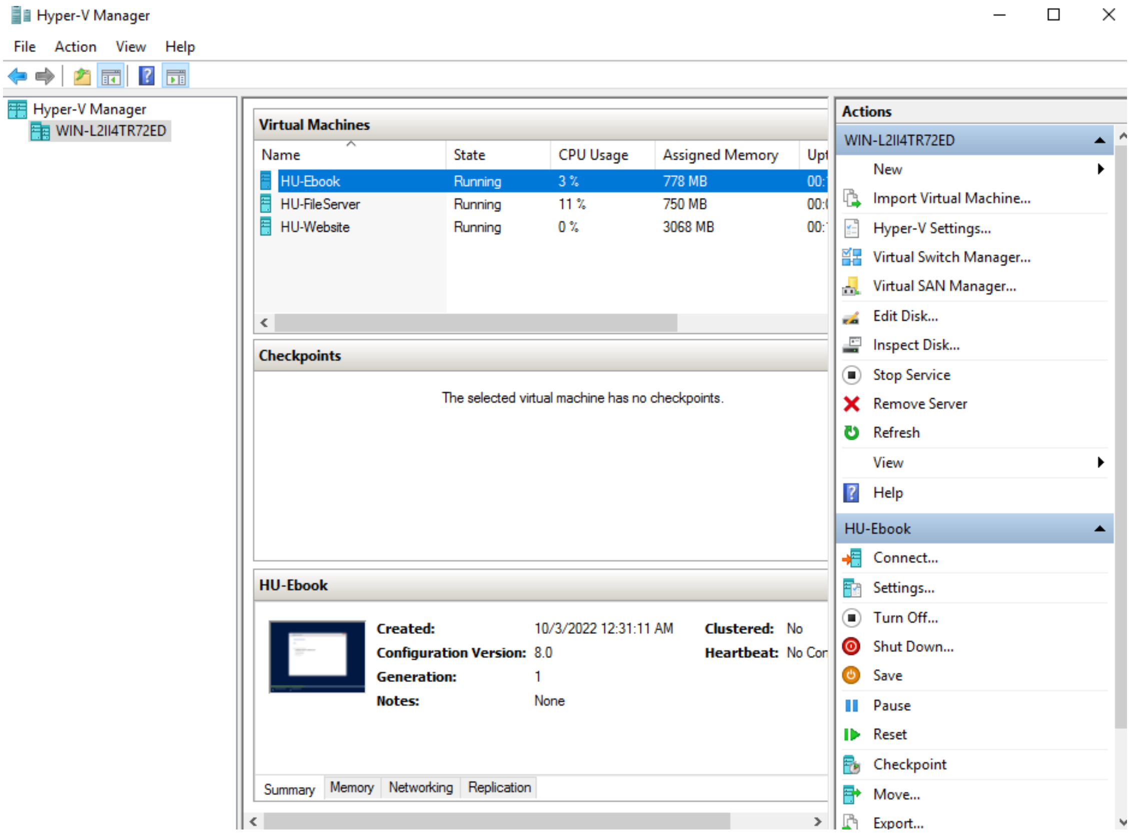

Figure 17. Created Virtual Machines

provides lots of configuration options including the number of virtual processors you should configure, the required disk and memory amounts. Hyper-V provides processor compatibility settings to make it easier to use an older OS and perform a live migration of a virtual machine to another physical computer with a different processor version. The researcher selected the graphical user interface tools (Windows resource monitor and task manger) that provides a real time information of CPU, memory, disk and network metrics. The following figures show performance of the server before and after virtualization.

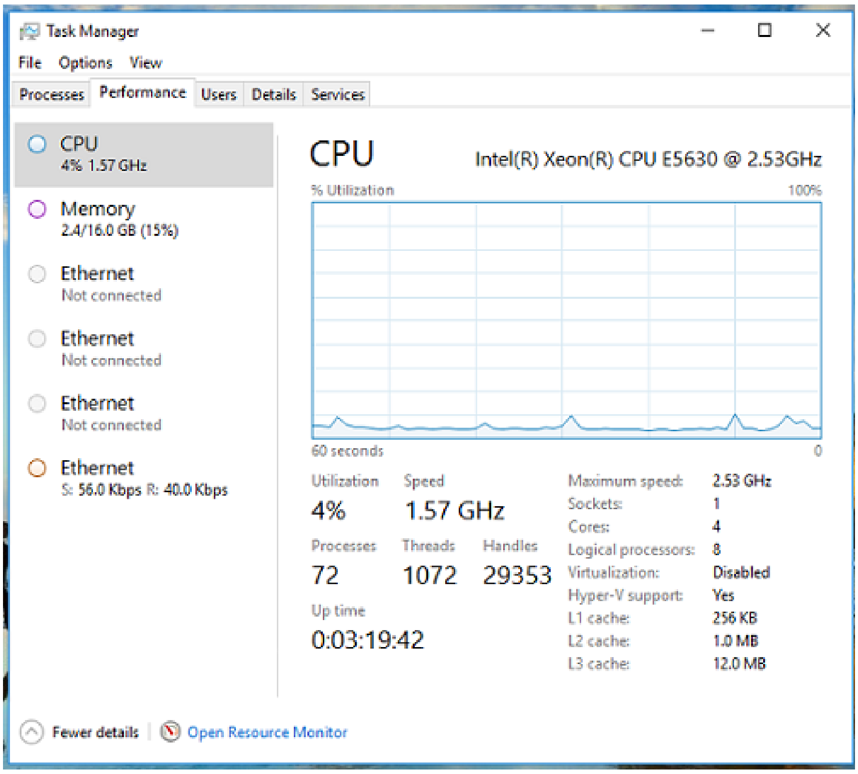

Figure 18. Performance of single physical servers before virtualization

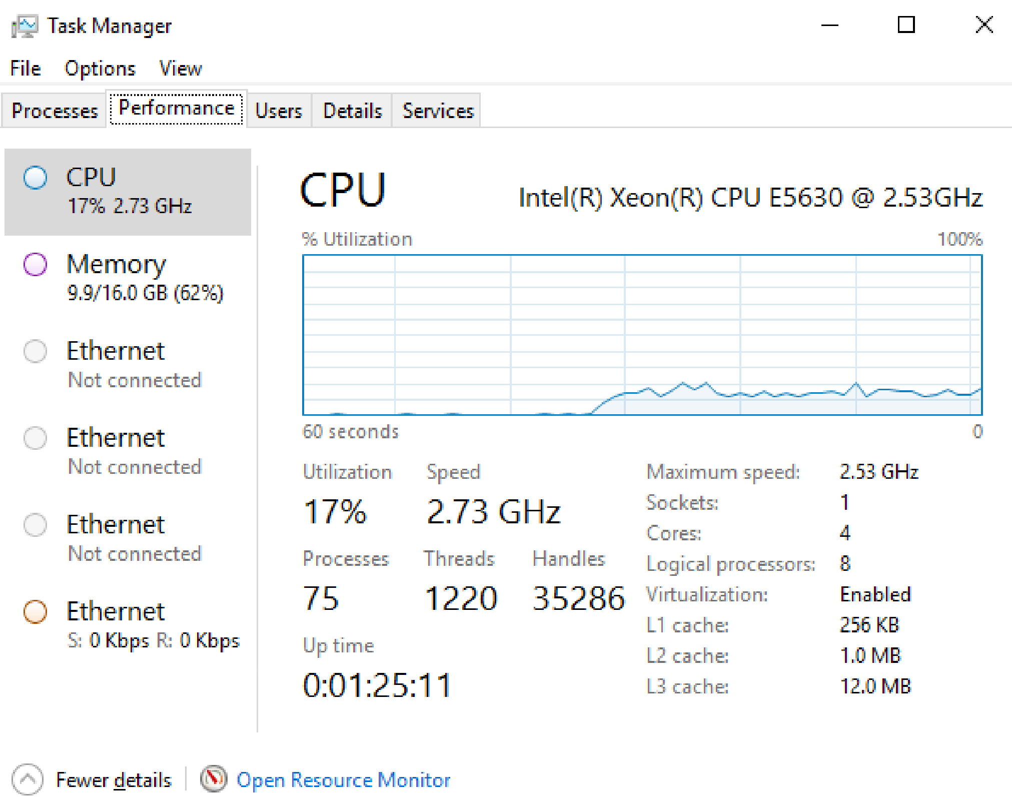

Figure 19. Performance of single physical servers after virtualization

The screen-shoot figures (Figure 19 and Figure 18) shows the CPU and memory utilization of the server. Additionally, the researcher measured the disk utilization of the server in both before and after virtualization. The measured disk usage before and after virtualization are is 32% and 67.4% consecutively.

A comparison of the two cases is shown in Figure 20. The bar graph shows an increase in CPU, memory and disk utilization of the system after virtualization. This is expected as installing and running virtual machines require more resource. It should be noted that using virtualization we are able to run three services in one server. Which required three different server in a previous setup. Thus, we are able to reduce the number of physical servers used to run similar number of services.

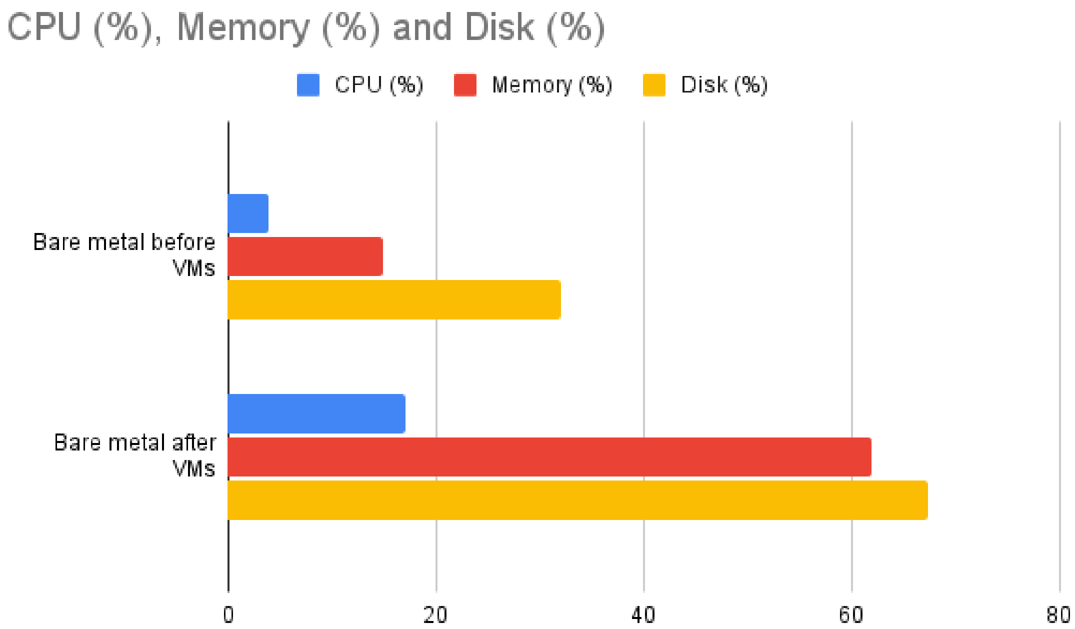

Figure 20. Bare metal before VMs and Bare metal after VMs

Figure 21 shows VM running HU-Ebook server uses 4% of CPU and 80% of memory which is very low. As indicated in the literature review part, the utilization of physical resources 25% to be better, 50% to be good, 75% to be very good, and 100% to be best. Here the experiment exposes that a one server to one application architecture is not business oriented. Hence, it is advisable to deploy many servers/ VMs on a single physical server.

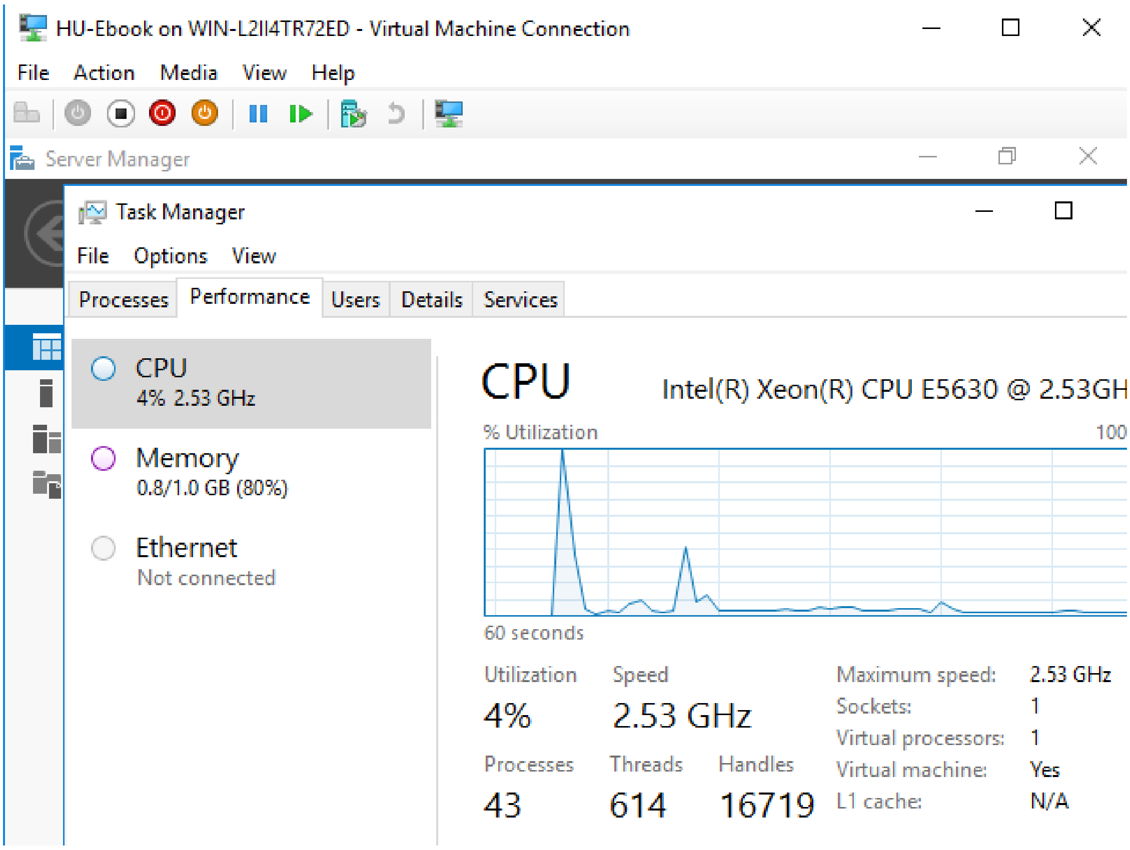

Figure 21. Performance of HU-Ebook virtual server

The other virtual server that the researcher deployed is HU- website servers which can running (Figure 22) uses 35% of CPU and 80% memory. But the CPU usage and memory of any system or server is dependent on the state. While the server is running and many users are interacted with the system the usage of resources are increasing. Servers with different workloads and computations will have different resource usages without ignoring other facts like the resources the original servers have and also the version of servers.

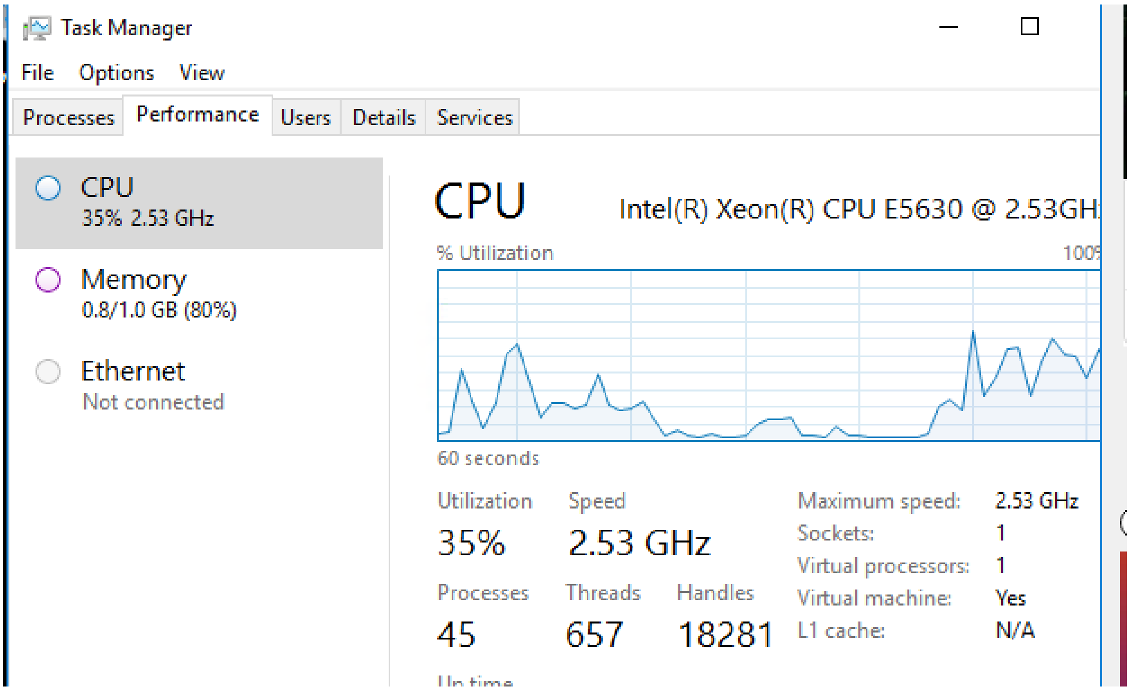

Figure 22. Performance of HU-Website virtual server

Finally Figure 23 shows virtual machine running HU-File Server it consumes 27% of CPU and 90% physical memory which is medium when compared with maximum utilization supported.

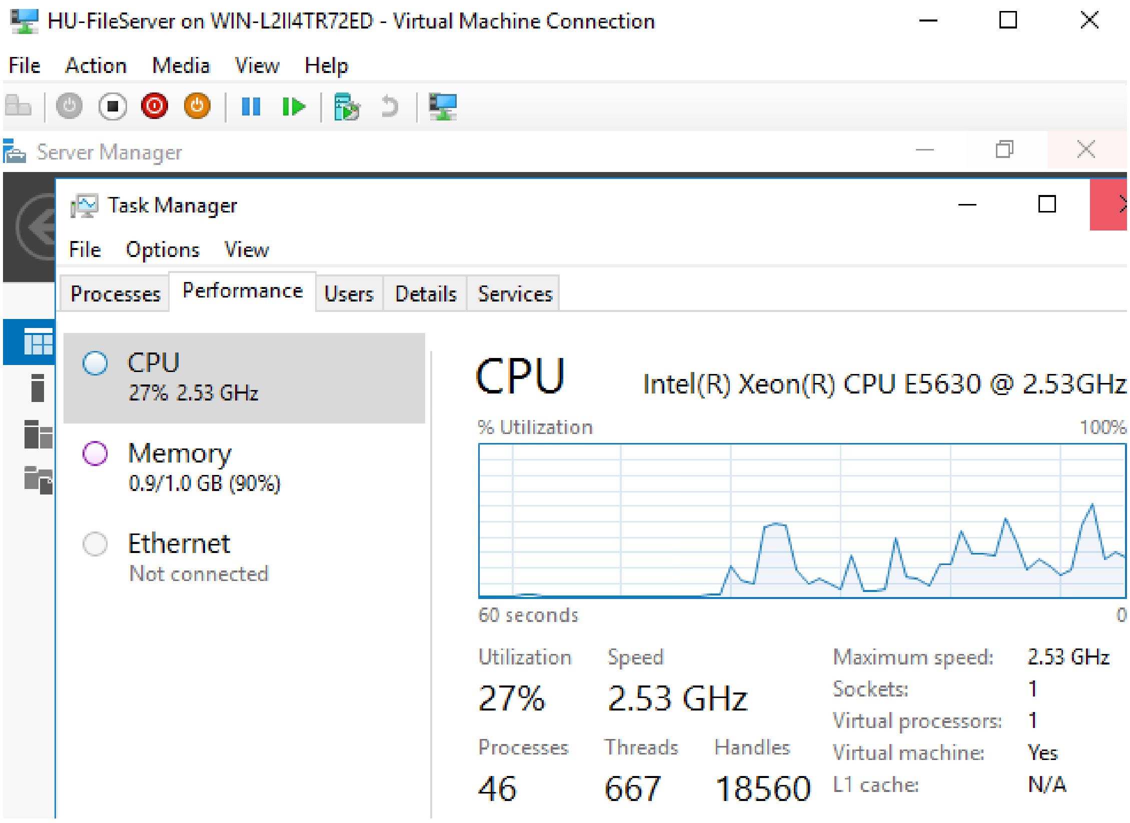

Figure 23. Performance of HU-File virtual server

5.7 Chapter summary

From the above experiment you can see the resource consolidation has certified, that to deploy more than one virtual server within a single physical machine to utilize the server appropriately. Even though the number of VM deployed on a single server depends on the amount of workload it had and the number of users interacted with the server. By applying this study the University data cenetr saves two physical server for other purpose of the services since it is possible to provide the same service using only a single server.

6. CONCLUSION AND RECOMMENDATION

6.1 Overview

In the previous chapter we presented results and discussions on the evaluation of the proposed framework. In this presents a conclusion for the study and indicate future direction for fellow researchers.

6.2 Conclusion

This study highlights the importance of virtualization technology being implemented in data centers to save the cost and maximize the efficiency of different resources available. To that end, it presents a framework to migrate from the physical device based infrastructure to a virtualized environment. The framework is implemented and evaluated on services that are being used in the university.

The first chapter presents introduction about basic concepts including data center, cloud computing and virtualization which is the core technology enabling the cloud. The chapter presents detail description and different aspects of virtualization, like components, types, benefits and challenges of virtualization. The chapter also presents the current existing IT infrastructure of HLI data centers. The problem statement, the objectives and scope of the study are also presented in this chapter.

The second chapter presents a literature review analysing previously conducted related works. The chapter presents a scientific work on the benefits of having virtualization in data centers to efficiently use available resources. Additionally, the chapter presents four existing server virtualization frameworks and describe their short comings to address the problem described in the previous chapter.

The third chapter presents the research methodology used for the study. Design Science Research Methodology (DSRM) was used which is a system of principles, practices and procedures required to carry out a study. The chapter presents the qualitative data analysis using data from observation and interviews. From this, the problem was formulated, server requirements, design, implementation and evaluation methods are described.

The proposed framework is presented in the forth chapter. The researcher proposed a seven phase framework to properly implement virtualization in university data center. It starts by planing, analysis, design, implementation, testing, deployment finally continuous support. The chapter describes the tasks that are needed to be done on each phases. The researcher demonstrated feasibility of the framework by implementing the phases in an experimental environment. A single server and Hyper-v virtualization software are used in the experiment . Three virtual machines were created to server three application (HU-Ebook, HU-File Server and HU-Website), that were previously in separate physical servers.

In the fifth chapter, result of the implementation and discussions on the result are presented. The result from interview and observation highly confirmed the previously described problems. Additionally, the challenges in the data center create obstacles for the mission of educational universities. The experimental results showed the feasibility of the framework in migrating to virtualized infrastructure. The evaluation showed that by using virtualization we can provide the same level of services by using less physical infrastructure. In other terms, implementing virtualization technologies in data center attains highest level resource utilization.

6.3 Recommendation

The study recommends that university data center use the proposed well-matched framework in the implementation of server virtualization in data center. This would enable the university to increase their levels of resource utilization in data centers, which will lead to a decrease on the server hardware purchase expenditure. The future studies can be carried out on exploring security models of virtualization implementation frameworks as well as studying other virtualization types like storage, network and application virtualization. Finally, as a future work the experiments preformed in this study can be further expanded to include more services and test different use cases.

References

[1] Jonathan Koomey et al. “Growth in data center electricity use 2005 to 2010”. In: *A report by Analytical Press, completed at the request of The New York Times* 9.2011 (2011), p. 161.

[2] Anixter International Inc. *Data Center Infrastructure Resource Guide*. accessed: 10.02.2022.

[3] Peter Mell and Timothy Grance. “The nist definition of cloud computing, September 2011”. In: *Accessed on May* (2014).

[4] Vaishali Pardeshi. “Architecture and adoption model for cloud in higher education: Indian perspective”. In: *Bonfring international journal of industrial engineering and management Science* 4.2 (2014), pp. 43–47.

[5] Peter Mell and Timothy Grance. “The NIST definition of cloud computing. National Institute of Standards and Technology (NIST)”. In: *Vers* 15 (2011), pp. 800–145.

[6] K Thirupathi Rao, P Sai Kiran, and L Siva Shanker Reddy. “Energy efficiency in datacenters through virtualization: A case study”. In: *Global Journal of Computer Science and Technology* (2010).

[7] P Buchanan, Vincent Yampolsky, and Bill Buchanan. “Comparison of the power consumption and carbon footprint of a cloud infrastructure against standard desktops”. In: *Computer* 55 (2011), p. 77.

[8] A Chawdhry and Charles Mance. “Virtualization: providing better computing to universities”. In: *Information Systems Educators Conferenc (ISECON) Nashville Tennessee, USA* 27.1401 (2010), pp. 1–7.

[9] Zipporah Nyamweya Buyeke. “Server virtualization framework: case of selected government ministries, Kenya”. PhD thesis. Strathmore University, 2013.

[10] Robert R Harmon and Nora Auseklis. “Sustainable IT services: Assessing the impact of green computing practices”. In: *PICMET’09-2009 Portland International Conference on Management of Engineering & Technology*. IEEE. 2009, pp. 1707–1717.

[11] Rabi Prasad Padhy, Manas Ranjan Patra, and Suresh Chandra Satapathy. “Available Online at www. jgrcs. info”. In: *Journal of Global Research in Computer Science* 2.12 (2011).

[12] Radhwan Y Ameen and Asmaa Y Hamo. “Survey of server virtualization”. In: *arXiv preprint arXiv:1304.3557* (2013).

[13] Pankaj Sareen. “Cloud computing: types, architecture, applications, concerns, virtualization and role of it governance in cloud”. In: *International Journal of Advanced Research in Computer Science and Software Engineering* 3.3 (2013).

[14] Rajkumar Buyya, Christian Vecchiola, and S. Thamarai Selvi. *Chapter 3 - Virtualization*. Ed. by Rajkumar Buyya, Christian Vecchiola, and S. Thamarai Selvi. Boston, 2013. DOI: https://doi.org/10.1016/B978-0-12-411454-8.00003-6. URL: https://www.sciencedirect.com/science/article/pii/B9780124114548000036.

[15] N.M. Mosharaf Kabir Chowdhury and Raouf Boutaba. “A survey of network virtualization”. In: *Computer Networks* 54.5 (2010), pp. 862–876. ISSN: 1389-1286. DOI: https://doi.org/10.1016/j.comnet.2009.10.017. URL: https://www.sciencedirect.com/science/article/pii/S1389128609003387.

[16] Sogand Shirinbab. *PERFORMANCE IMPLICATIONS OF VIRTUALIZATION*. accessed: 10.02.2022.

[17] Aaqib Rashid and Amit Chaturvedi. “Virtualization and its role in cloud computing environment”. In: *International Journal of Computer Sciences and Engineering* 7.4 (2019), pp. 1131–1136.

[18] Tom Brey and Larry Lamers. “Using virtualization to improve data center efficiency”. In: *the green grid, whitepaper* 19 (2009).

[19] Mohammad Reza Ahmadi and Davood Maleki. “Performance evaluation of server virtualization in data center applications”. In: *2010 5th International symposium on telecommunications*. IEEE. 2010, pp. 638–644.

[20] Yanan Liu et al. “Energy consumption and emission mitigation prediction based on data center traffic and PUE for global data centers”. In: *Global Energy Interconnection* 3.3 (2020), pp. 272–282.

[21] Ladina Gilly. “Data centre design standards and best practices for public research high performance computing centres”. In: (2016).

[22] Behzad Norouzi-Khangah et al. “Performance assessment of cooling systems in data centers; Methodology and application of a new thermal metric”. In: *Case Studies in Thermal Engineering* 8 (2016), pp. 152–163.

[23] Jinkyun Cho and Byungseon Sean Kim. “Evaluation of air management system’s thermal performance for superior cooling efficiency in high-density data centers”. In: *Energy and buildings* 43.9 (2011), pp. 2145–2155.

[24] Narinder Kumar and Ashwani Kumar. “Experimental assessment of UPS battery load model considering harmonics and its investigation in the distribution systems”. In: *Materials Today: Proceedings* 5.1 (2018), pp. 709–715.

[25] Rihards Balodis and Inara Opmane. “History of data centre development”. In: *Reflections on the History of Computing*. Springer, 2012, pp. 180–203.

[26] Mueen Uddin, Asadullah Shah, and Amjad Rehman. “Metrics for Computing Performance of Data Center for Instigating Energy Efficient Data Centers”. In: (2014).

[27] Qiang Wu et al. “Dynamo: Facebook’s data center-wide power management system”. In: *ACM SIGARCH Computer Architecture News* 44.3 (2016), pp. 469–480.

[28] Chaoqiang Jin et al. “A review of power consumption models of servers in data centers”. In: *applied energy* 265 (2020), p. 114806.

[29] Christian Dawson. *The environmental benefits of virtualization: The mere existence of virtualization is more important than the specific type of virtualization*. 2008.

[30] Md Sharif Hossain. “Panel estimation for CO2 emissions, energy consumption, economic growth, trade openness and urbanization of newly industrialized countries”. In: *Energy policy* 39.11 (2011), pp. 6991–6999.

[31] Dirk Paessler. “Server virtualization and network management”. In: *Database and network journal* 38.5 (2008), p. 13.

[32] Patrick Kurp. “Green computing”. In: *Communications of the ACM* 51.10 (2008), pp. 11–13.

[33] Jens Vykoukal, Martin Wolf, and Roman Beck. “Does green IT matter? Analysis of the relationship between green IT and grid technology from a resource-based view perspective”. In: *PACIS 2009 proceedings* (2009), p. 51.

[34] L Sgallari. “Reducing Infrastructure Cost Through Virtualization”. In: *The Architecture Journal* 20 (2009), p. 43.

[35] Joan K Pierson, SE Kruck, and Faye Teer. “Trends in names of undergraduate computer-related majors in AACSB-accredited schools of business in the USA”. In: *Journal of Computer Information Systems* 49.2 (2008), pp. 26–31.

[36] Andy Rowsell-Jones and Simon Mingay. “Going Green, the CIO’s Role in Enterprisewide Environmental Sustainability”. In: *executive summary), Gartner EXP Premier (Stamford, CT: Gartner Inc* (2008).

[37] Juan Luo. “Dynamical resource allocation algorithm research in network virtualization”. In: *Journal of China Institute of Communications* 32.7 (2011), pp. 64–70.

[38] Prashanta Kumar Das. “A Practical Approach on Virtual Machine Live Migration”. In: *Handbook of Research on Securing Cloud-Based Databases with Biometric Applications*. IGI Global, 2015, pp. 464–483.

[39] Kevin Lim et al. “Disaggregated memory benefits for server consolidation”. In: *HP Laboratories* (2011).

[40] Srikanth Sundarrajan and Hariprasad Nellitheertha. “Server Consolidation and virtualization”. In: *IFOSYS White Paper* (2006).

[41] Manogna Chebiyyam et al. “Server consolidation: Leveraging the benefits of virtualization”. In: *Infosys Research, SETLabs Briefings* 7.1 (2009), pp. 65–75.

[42] K Sudha, K Sujitha, and P Selvanayaki. “Green Networking”. In: *Journal of Computer Applications* 5 (2012), pp. 179–189.

[43] Hai Yu et al. “Application of server virtualization technology in power information construction”. In: *Journal of Physics: Conference Series*. Vol. 1744. 2. IOP Publishing. 2021, p. 022008.

[44] Mueen Uddin and Azizah Abdul Rahman. “Energy efficiency and low carbon enabler green IT framework for data centers considering green metrics”. In: *Renewable and Sustainable Energy Reviews* 16.6 (2012), pp. 4078–4094.

[45] Marilla D Svinicki. “A guidebook on conceptual frameworks for research in engineering education”. In: *Rigorous Research in Engineering Education* 7.13 (2010), pp. 1–53.

[46] Alan Hevner and Shirley Gregor. “Envisioning entrepreneurship and digital innovation through a design science research lens: A matrix approach”. In: *Information & Management* (2020), p. 103350.

[47] Ken Peffers et al. “A design science research methodology for information systems research”. In: *Journal of management information systems* 24.3 (2007), pp. 45–77.

[48] Shengnan Zhang, HT le Fever, and Fever HT le Zhang S. “An Examination of the Practicability of COBIT Framework and the Proposal of a COBIT-BSC Model”. In: *Journal of Economics, Business and Management* 1.4 (2013), pp. 391–395.

[49] Ken Peffers, Tuure Tuunanen, and Björn Niehaves. *Design science research genres: introduction to the special issue on exemplars and criteria for applicable design science research*. 2018.

[50] Jennifer Rowley. “Conducting research interviews”. In: *Management research review* (2012).

[51] Catherine M Castellan. “Quantitative and qualitative research: A view for clarity”. In: *International journal of education* 2.2 (2010), p. 1.

[52] R Lyman Ott and Micheal T Longnecker. *An introduction to statistical methods and data analysis*. Cengage Learning, 2015.

[53] Daniel A Menascé and Virgilio AF Almeida. "Performance of Client/Server Systems". In: *Performance Evaluation: Origins and Directions*. Springer, 2000, pp. 201–218.

[54] Manal A Abdel-Fattah, Yehia Helmy, and Alaa Mohamed El-Sayed. "A Proposed Framework for Evaluating Performance of Data Center". In: *International Journal of Computer Science and Information Security (IJCSIS)* 16.11 (2018).

[55] Hyungro Lee. "Virtualization basics: Understanding techniques and fundamentals". In: *School of Informatics and Computing Indiana University* 815 (2014).

[56] Sharvari Tamane. "A review on virtualization: A cloud technology". In: *International Journal on Recent and Innovation Trends in Computing and Communication* 3.7 (2015), pp. 4582–4585.

[57] Mueen Uddin and Azizah Abdul Rahman. "Virtualization implementation model for cost effective & efficient data centers". In: *arXiv preprint arXiv:1206.0988* (2012).

[58] Eric Siebert. *Sizing server hardware for virtual machines*. accessed: 10.13.2022.

[59] SearchWindowsServer.com. *Key considerations: Choosing server hardware for virtualization*. accessed: 10.02.2022.

A. Appendix

HAWASSA UNIVERSITY DEPARTMENT OF COMPUTER SCIENCE, SCHOOL OF GRADUATES STUDIES

Appendix: Interview Questions

Personal Information

Name (optional): ______________________________

Age: 20-30 ☐ 31-40 ☐ Above ☐

Work Experience: =5 ☐ =>5 ☐ >7 ☐ | Level of Education: MSC ☐ BSC ☐

1. Would you please mention which type of data center architecture in Hawassa University uses? Traditional (physical), Virtualized or Hyper converged.
2. Does the idea of virtualization started before in your organizations? If not what would be the reason? Which type? Server, Application, network, storage.
3. What kinds of benefits could you expect from virtualized infrastructure?
4. Can you confidently suggest your organization can use virtualized data center? Yes/ No Why
5. Do you believe most top management members and IT infrastructure decision makers are usually change resistant? Yes/ No Why? If so what would be the reason?
6. Do you believe virtualized data center solution can use unnecessary complication for the Universities and data center administrators? Yes/ No Why?
7. Do you believe the Infrastructures in the data center capable of supporting virtualization technology?
8. What is the trend of ICT server purchase in the data center? Increasing /Decreasing
9. Is there use of any server virtualization framework in your Universities? Yes/ No
10. In question number 9 if your answer is No which type of virtualization your data center need A. Full virtualization ☐ B. Para-virtualization ☐ or C. Hardware-assisted ☐
11. Data center server capacity information
    i. How many processors does each server have? What is the speed of the processors?
    ii. What is the capacity of the physical memory of the servers?
    iii. What is the capacity of the disk drives?
    iv. How many systems are installed on each server?

A. Appendix

ሀዋሳ ዩኒቨርሲቲ
ቴክኖሎጂ ኢንስቲትዩት
ኢንፎርማቲክስ ፋኩሊቲ
ኮምፒውተር ሳይንስ ት/ክፍል

Hawassa University
Institute of Technology
Faculty of Informatics
Department of Computer Science

ቁጥር፡ Ref. No: Info/CS/ 112 /15

ቀን፡ Date: 12 / 04 /2022

ለ፡ አዲስ አበባ ሳይንስ እና ቴክኖሎጂ ዩኒቨርሲቲ

አ.አ

ጉዳዩ፡- ትብብር እንዲደረግላቸዉ ስለመጠየቅ

ከላይበ ርዕሱ ለመጥቀስ እንደተሞከረዉ አብደላህ ሁሴን በሐዋሳ ዩኒቨርሲቲ ቴክኖሎጂ ኢንስቲትዩት የኮምፒዉተር ሳይንስ የመደበኛ የሁለተኛ ዲግሪ ተማሪ ሲሆን አሁን ላይ "Data Center virtualization Frame Work in Ethiopian Higher learining Institutions, the case of Hawassa University" በሚል ርዕስ የመመረቂያ ፅሁፉን በመስራት ላይ ይገኛል ። በመሆኑም የምርምሩ ውጤት ችግር ፈቺ እና ማሕበረሰብ አቀፍ አገልግሎት ሰጭ እንዲሆን ለማድረግ የሁሉም ባለድርሻ አካላት መልካም ትብብር ያስፈልጋል።

ስለሆነም የተመራቂ ተማሪያችንን የፈጠራ ውጤት እውን ለማድረግ በዘርፉ እንደየአስፈላጊነቱ ለስራው የሚረዳ እውነተኛ መረጃ መስብሰብ አስፈላጊ ነው ። በመሆኑም በመስሪያቤታችሁ ለምርምር ስራው መሳካት አስፈላጊዉን መረጃ በመስጠት እንድትተባበሩ ስንል በአክብሮት እንጠይቃለን።

ከሰላምታ ጋር

ግልባጭ፡-

- ለኢንፎርማቲክስ ፋኩልቲ

ቴ/ኢ.

"የተማረትውልድHIV/ADIS'ንበአግባቡይከላከል!"

(+) 251-(0)46-822-39-41 Fax: (+) 251-046-220-54-21E- mail: informatics@hu.edu.et
(+) 251-(0)46-220-94-35Fax: (+) 251-046-220-54-21 Website:www.hu.edu.et ✉5 Hawassa, Ethiopia