2011

Examining Effects Of Virtual Machine Settings On Voice Over Internet Protocol In A Private Cloud Environment

Yuan Liao
Indiana State University

EXAMINING EFFECTS OF VIRTUAL MACHINE SETTINGS ON VOICE OVER

INTERNET PROTOCOL IN A PRIVATE CLOUD ENVIRONMENT

A dissertation

Presented to

The College of Graduate and Professional Studies

College of Technology

Indiana State University

Terre Haute, Indiana

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

---

by

Yuan Liao

May 2011

Keywords: Technology Management, Digital Communication, Cloud Computing, IP, Telephony

COMMITTEE MEMBERS

Committee Chair: Dr. Derrek Dunn

Professor

Savannah State University

Committee Co-Chair: Dr. Dewayne Brown

Professor

North Carolina A&T State University

Committee Member: Dr. Yuetong Lin

Assistant Professor

Indiana State University

Committee Member: Dr. Suhansa Rodchua

Assistant Professor

University of Central Missouri

Committee Member: Dr. Li-Shiang Tsay

Assistant Professor

North Carolina A&T State University

ABSTRACT

The virtualization of computing resources, as represented by the sustained growth of cloud computing, continues to thrive. Information Technology departments are building their private clouds due to the perception of significant cost savings by managing all physical computing resources from a single point and assigning them to applications or services as needed while remaining in control of their systems and information. As part of this trend, real time communication applications including IP telephony can be integrated with other software applications into one platform and deployed in private clouds to reduce capital expenditure and lower overall costs of daily based maintenance and real estate required for computer hardware.

As a critical component of private clouds, however, virtualization may adversely affect a real time communication application running in virtual machines as the layer of virtualization on the physical server adds system overhead and contributes to capacity lose. While real time communication services require a certain level of system performance and availability to address communication latency and overhead bottleneck, it is essential to investigate potential performance implications of private clouds on IP telephony applications.

The purpose of this study was to investigate how to apply cost benefits of private clouds to Voice over Internet Protocol applications without compromise on communication performance. Through the experimental study, the statistical technique ANCOVA was used to examine the effects of virtual machine settings on voice quality when network condition remained the same. Linear Regression analysis was used to test whether the voice quality can be

predicted from virtual computing resources and network bandwidth in private clouds. The results of this research provided a better understanding of the effects of virtual machine settings on voice quality of Voice over Internet Protocol applications in private clouds from the prospect of technology management.

ACKNOWLEDGMENTS

I want to express a deep gratitude to Dr. Derrek Dunn, my advisor and dissertation committee co-chair for continuous guidance and advice throughout the doctoral studies. I am also highly indebted to Dr. Dewayne Brown, dissertation committee co-chair, for his assistance and direction during the process of completing this dissertation.

Special thanks go to Dr. Yuetong Lin, dissertation committee member, who was always available for valuable counsel, direction, and support. Discussions with him about research in new fields of information technology not only broadened my perspective but also inspired me in many aspects of the dissertation research.

My appreciation is extended to Dr. Suhansa Rodchua, dissertation committee member, for her constructive criticism, numerous comments, and quick turnarounds on different versions of the dissertation. My acknowledgement also goes to Dr. Li-Shiang Tsay, dissertation committee member, for providing valuable suggestion and feedback to the dissertation writing.

I also want to thank Dr. George Maughan, director of the Ph.D. program, for reviewing the final manuscript and giving constructive comments that helped strengthen the dissertation.

Finally, I would like to thank my family, my wife Flora and my son Rick, for their patience, tolerance, and sacrifices throughout my coursework and dissertation process. Without their encouragement and support, I could not have this done.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF EQUATIONS

CHAPTER 1

INTRODUCTION

Chapter 1, Introduction, starts with the background of the research and defines the statement of the problem, and research questions. Additionally, the chapter presents the purpose and the significance of the study. Finally, Chapter 1 discusses the assumptions and limitations in the study, and provides definitions of terms.

Background

The evolution to cloud computing has advanced rapidly over the last few years. The trend has been driven by such technological innovations as collaboration software, service oriented architectures (SOAs), and data center virtualization (Rittinghouse & Ransome, 2010). It has shown that most computing resources within Information Technology (IT) data centers are underutilized (Linthicum, 2009). Cloud computing provides an end-to-end, unified solution that maximizes the ability to address the scalability, collaboration, virtualization, and performance requirements being driven by today's global business challenges and opportunities. Cloud computing is expected to be a significant growth driver in worldwide IT spending. According to International Data Corporation (IDC) (2008), cloud computing services are expected to grow at a compound annual growth rate (CAGR) of 27% and reach $42 billion by 2012; spending on non-cloud IT services is expected to grow at a CAGR of 5%. A properly designed and implemented cloud infrastructure provides the benefit of substantially lowering the total cost of hardware and software ownership over the traditional hosting environment though the utilization of virtualization and the use of open source software tools (Ou, 2009).

Despite all the benefits of cloud computing, there are some barriers to adoption. Security and privacy of data were highlighted as key concerns for organizations considering cloud computing (Foley, 2009). There is a great deal of uncertainty about how security at all levels from storage, server, data, operating system (OS), applications to network can be achieved in this new computing model cloud computing. Organizations today face numerous requirements of different scopes attempting to protect the privacy of individuals' information, and it is not clear whether the cloud computing model provides adequate protection of such information, or whether organizations will be found in violation of legal regulations because of this new model (Mather, Kumaraswamy, & Latif, 2009). In a public cloud, computing resources are offered as services on a subscription over the Internet from an off premise, third-party provider. As security management and daily based operation are relegated to the third party vendor, organizations have a low degree of control and oversight of the physical and logical security. These uncertainties have kept many organizations from embracing public clouds due to fear about security threats and loss of control of data and systems (Avanada, 2009).

Contrary to public clouds, private clouds have emerged as a response to the privacy and security concerns of many organizations. Private clouds leverage virtualization to allocate the workloads between all internal servers, storage and operating systems, thus improve the utilization of each computing resource while maintaining a high level of control over own data and systems. The key advantage of private clouds is control. In a private cloud, organizations own and control their infrastructure and applications behind a firewall, meanwhile gain some of the benefits of cloud computing without sacrificing on data security, corporate governance, and reliability concerns (Foley, 2008).

Virtual computing resources, however, may influence the performance of software applications as adding a layer of virtualization software on the computer server adds overhead to the overall system and contributes to capacity loss. There has been an increasing interest to examine the impact of virtualization on performance loss. The reason behind this increased interest is to do with the growing adoption of clouds computing by organizations which expect existing software applications to be able to run on virtual machines and to perform as good as on physical servers.

Statement of the Problem

Real time communication services require a certain level of system performance and availability to address communication latency and overhead bottleneck. Voice over Internet Protocol (VoIP) is not an exception. As virtualization can be a significant factor in performance loss of software applications because of interactions with the underlying virtual machine monitor and other virtual machines (Foster et al., 2008; Menon et al., 2005), this study examined the effects of virtual computing resources on performance of a VoIP application in private clouds.

Thus the basic research question in this research is: How do virtual machine settings in a private cloud environment impact on voice quality of VoIP applications? The findings from this research will answer the following substantive research questions:

1. Do virtual machine settings such as virtual RAM and virtual hard drive play a significant role in affecting voice quality after the effect of network condition is removed?

2. Do virtual machine settings and network bandwidth together significantly predict the expected voice quality?
3. Do virtual machine settings such as virtual RAM and virtual hard drive significantly predict the expected voice quality?

Statement of the Purpose

VoIP has become a popular alternative to traditional public switched telephone network (PSTN) for voice communication as it provides a benefit for lowering the cost of communication and infrastructures, increasing service flexibility, and converging voice and data networks to organizations of different sizes (Goods, 2002; Cherry, 2005). Flexibility is a major feature of the packet transmission in IP network. Multimedia streams and voice data files are disseminated into packets, and the upper application level details are unaware about the network routers between the source and the destination. It is costly for organizations to operate two separate networks: voice networks and data networks than a single network. Combining the two will significantly reduce operating cost such as technical support, maintenance, configuration, and upgrade. More importantly, VoIP provides many advanced features that may be more difficult to implement in the PSTN, such as number portability, integration with emails, multimedia features, voice portal, mobility, and so on. These benefits over traditional PSTN services have motivated the wide development and deployment of VoIP technology in the market place. Major research firms reported that a majority of Private Branch Exchange (PBX) systems sold in the first quarter of 2006 were IP-based (Poe, 2006).

As a part of trend of business process integration, organizations are unifying all forms of communications from real-time services such as VoIP, instant messaging, and video conferencing to non-real-time communication services such as unified voice mail in view to optimize business processes and reduce the response time, manage flows, and eliminate device and media dependencies through unified communications initiatives. Given private clouds are predicted to be the future of information technology (Brodkin, 2008), real time communications applications including computer-based VoIP applications can be integrated with IT software solutions over one platform and deployed to a virtual machine of private clouds due to the perception of big cost savings on overall hardware cost, maintenance cost, and real estate required for that hardware. However, the growing adoption of clouds computing by organizations expects the installed software applications to be able to run on virtual machines and to perform as well as on physical servers.

The purpose of this study is to investigate how to apply cost benefits of private clouds to VoIP without compromise on communication performance. Since real time communication services requires a certain level of system performance and availability to address communication latency and overhead bottleneck, it is essential to investigate potential performance implications of private clouds to allow VoIP applications to function properly in a private cloud environment.

Significance of the Study

In order for VoIP services to be widely adopted by organizations, they must offer the quality of service at the same level as the current Public Switched Telephone Network services in a variety of environments. As a result, VoIP must also be optimized to offer comparable quality voice services similarly to PSTN.

The field of IP telephony in the private clouds domain is relatively new. Procedures for quality control and management of VoIP applications in a private cloud environment, though developing very fast, have yet to evolve fully. While real time communication services such as

VoIP applications can run in private clouds, the performance of software applications in a virtualized environment can differ greatly from one that is within a non-virtualized environment (Menon et al., 2005). As a critical component of private clouds (Rittinghouse & Ransome, 2010), virtualization may influence the application and system performance (Rosenblum, 2004) because adding a layer of virtualization software on the computer severs adds overhead to the overall system and contributes to capacity loss.

Contributions made by this research are the evaluation of the performance impact of virtual computing resources on VoIP applications in private clouds. The findings are useful both to researchers and practitioners, and contribute to applied research literature, in particular IP telephony and cloud computing literature. Most of the research in the area of IP telephony so far has focused on non-virtualized environments. Current literature lacks a focused perspective on virtualized environments. This research intends to bridge this gap by utilizing quantitative techniques to test the relationship between a set of virtual computing resources in private clouds and VoIP performance from the perspective of cloud computing.

VoIP requires timely packet delivery with low latency, jitter, packet loss, and sufficient bandwidth. To achieve this, efficient deployment of VoIP must ensure that these real time traffic requirements can be guaranteed in a private cloud environment. For the practitioner, a better understanding of the impact of virtualization in private clouds would allow the decision makers to assess whether organizations would allow VoIP applications to run in private clouds for cost savings without compromise on system performance. This research provides valuable insight on what kind of performance loss and the effects of virtual computing resource in private clouds an organization should access before it considers deploying VoIP in private clouds.

Assumptions

Assumptions of the Study

This research assumes the implementation of private clouds is taken place in organizations that have significant existing IT investments in hardware and software, the cost of which cannot be recovered if moved to public clouds. Such an assumption is necessary to warrant the consideration of running a variety of real time communication applications including VoIP in a virtualized environment of private clouds. While private cloud can be possibly implemented in small organizations, however, the business scale may not justify the potential saving and the cost to invest on and maintain their own IT infrastructures.

In a private cloud, an origination sets up virtualized environments on its own servers within its own data centers. One assumption is that the investigated targets will be limited to the organizations who still feel they absolutely cannot host their data outside of their firewalls due to privacy and legal issues, but they want to take advantage of the cloud computing architectures.

Another assumption for this research is the assessment of virtual machine settings that affect performance of voice quality is possible. Virtual RAM and virtual hard disk would accurately translate to the degree of impact in the analysis.

Assumptions of Statistical Analysis

Statistical tests rely upon certain assumptions about the variables used in the analysis. When these assumptions are not met the results may not be trustworthy, resulting in a Type I or Type II error, or misestimating of significance or effect size (Osborne & Waters, 2002). Analysis of Covariance (ANCOVA) and Multiple Linear Regression are chosen statistical techniques used in the study. Based on Norusis (2005), the assumptions for the linear regression analysis include:

- Independence assumption: all of observations must be independent.
- Normality assumption: Samples are randomly drawn from the normally distributed propulsions with unknown population means
- Equal variance assumption: The variance of the distribution of the dependent variable must be the same for all values of the independent variable.
- Linear relationship assumption: the relationship between the dependent and the independent variable is linear.

Like any statistical procedure, the interpretation of ANCOVA depends on certain assumptions about the data entered into the model. ANCOVA has the similar assumptions as Analysis of Variance (ANOVA) with two additional assumptions apart from the assumptions made in ANOVA: 1) independence of the covariate and treatment effect, and (2) homogeneity of regression slopes (Norusis, 2006). The assumptions of ANCOVA include:

- A normal distribution requires that the underlying populations from which the samples are taken be normally distributed.
- Variances are equal (Homogeneity of Variance).
- The sample size is sufficient to obtain statistical significance.
- Data were randomly collected during the study.
- Independence of variance estimates.

Limitations

Limitations of the Study

There are certain limitations in this research. This research focuses specifically on the area of server virtualization in private clouds. Virtualization is a very broad topic. There are many types of virtualization in private clouds including server virtualization, application virtualization, desktop virtualization, operating system virtualization, presentation virtualization, etc. This research will limit the focus to server virtualization which allows different operating systems and different applications simultaneously to run on the same hardware.

Next, when the computer-based VoIP application runs over the virtual server in private clouds, many different factors are involved in the process. Not only virtual RAM and virtual hard disk, but also other computing resources such as network bandwidth, availability may impact performance of VoIP. Only virtual computing resources such as RAM and hard disk are considered in this study.

The measured experiments were conducted in a controlled LAN scenario. However, it is difficult to predict the total voice traffic that can be supported by a real network environment, taking into account important design and engineering factors, including VoIP flow and call distribution, future growth capacity, performance thresholds, the impact of VoIP on existing network services and applications, and the impact of background traffic on VoIP.

The use of a quantitative study approach provides objectiveness and depth of information. The research approach focused on virtual machine settings rather than all types of virtualizations in private clouds. This limits our ability to triangulate information from multiple perspectives. More quantitative and qualitative research will be expected to overcome these limitations.

Limitations of Statistical Analysis

The sample size may be small for generalization. If groups vary widely in sample size and smaller groups have larger variances, the chance of rejecting the null hypothesis when it is true, a Type I error, increases. If groups vary widely in sample size and larger groups have smaller variances, the Type II error or chance of rejecting the null hypothesis when it is false increases. Limitations of the statistical analysis were to have an adequate sample size. Based on Norusis (2006) and Garson (2002), ANCOVA analysis may include the following limitations:

- A covariate must be an interval or ratio variable.
- A limited number of covariates. Too many covariates can cause a decrease in statistical power because a degree of freedom is lost for each covariate added.
- The covariate must have a linear relationship to the dependent variable.
- Covariates must be independent of treatment.
- Data on covariates can be gathered before treatment is administered.

Definitions of Terms

Application Virtualization: Application virtualization describes the practice of running software from a remote server rather than on the user's computer. A virtualized application is redirected at runtime to interface with the virtual operating system and all related resources that are managed by it rather than an actual, physical implementation of that operating system. The virtualization layer must be installed on a machine to intercept file and registry operations performed by a virtualized application, where it can transparently redirect those operations to a virtualized destination. The application that performs file operations never knows that it is not directly accessing a physical resource. Using this approach, applications can be made portable by redirecting their I/O tasks to a single physical file, and traditionally incompatible applications can be executed side by side. Advantages of application virtualization include may include cost savings on hardware, software and OS licenses, ability to run multiple versions of an application program concurrently on a single computer, and ease of application management, upgrading and migration (Rittinghouse & Ransome, 2010).

Internet Protocol: Internet Protocol (IP) is a network layer protocol that contains addressing information and some control information that enables packets to be routed. IP is a connectionless protocol, which means that there is no continuing connection between the end points that are communicating. Each packet that travels through the Internet is treated as an independent unit of data without any relation to any other unit of data. IP is the primary network-layer protocol in the Internet protocol suite. Along with the Transmission Control Protocol (TCP), IP represents the heart of the Internet protocols. IP has two primary responsibilities: providing connectionless, best effort delivery of data grams through an inter network (Sakellari, 2010); and providing fragmentation and reassembly of data grams to support data links with different maximum-transmission unit (MTU) sizes (RFC 791, 1981).

Network Virtualization: Network virtualization is a method of combining the available resources in a network by splitting up the available bandwidth into channels, each of which is independent from the others, and each of which can be assigned or reassigned to a particular server or device in real time. Each channel is independently secured. With network virtualization, the network is “carved up” and can be used for multiple purposes such as running a protocol analyzer inside an Ethernet switch. Components of a virtual network could include NICs, switches, VLANs, network storage devices, virtual network containers, and network media. Network virtualization is intended to optimize network speed, reliability, flexibility, scalability, and security. Network virtualization is said to be especially effective in networks that experience sudden, large, and unforeseen surges in usage (Moreno & Reddy, 2006).

Public Clouds: also called external clouds. In public clouds, physical computing resources such as computer servers and storage, and software applications are offered as services on a subscription base. Infrastructure and system software are owned, operated, and managed by a cloud service provides from a remote place. The main benefit of using public clouds service is lower cost due to no initial investment on hardware and applications, billing based on usage of the shared resources, and scalability to meet needs. The down side of private clouds is the customer of the public cloud service offering has a low degree of control and oversight of the physical and logical security aspects of a private cloud (Mather et al., 2009; Rittinghouse & Ransome, 2010).

Private Clouds: also called internal clouds. A private cloud is a virtualized infrastructure including hardware and software resources that exist behind the firewall and within the confines of an organization network, providing cloud computing characteristics around the ability to better utilize hardware and software resources. Private clouds preserver some of the benefits of cloud computing including lower hardware costs and dynamic scalability, while maintaining some level of control over physical computing resources and data. However, organizations must own, build, and manage their infrastructure and application as virtualized services and, as such, do not benefit from lower upfront capital costs and less daily base management (Mather et al., 2009; Rittinghouse & Ransome, 2010).

Server Virtualization: Server Virtualization consolidates multiple physical servers into virtual servers that run on a single physical server. The resources of the server itself are hidden, or masked, from users, and software is used to divide the physical server into multiple virtual environments, the virtual environments are called virtual private servers, guests, instances, containers or emulations. There are three popular approaches to server virtualization: the virtual machine model, the paravirtual machine model, and the operating system level model (Mather et al., 2009; Rittinghouse & Ransome, 2010).

Skype: Skype was created by the entrepreneurs Niklas Zennström and Janus Friis in 2003. Skype is software to be used to make free video and voice calls, send instant messages and share files with other Skype users over the Internet. It can be also used to make calls to landlines and mobiles for a fee based on usage. Its software integrates functions of audio conference, instance message, file transfer, voice mail, call forwarding, and maintenance of buddy list, providing a variety of attractive services for Skype users. Skype had an average of 124 million connected users per month in the second quarter of 2010 (Skype, 2011).

Virtual Machine: Virtual machine is the virtualized representation of a physical machine that is run and maintained on a host by a software virtual machine monitor or hypervisor. The hypervisor implements the virtualization on the physical machine and can be one of two types. Type 1 hypervisors are sometimes referred to as native hypervisors as they run on “bare metal,” or directly on the host's hardware to control the hardware and to monitor guest operating-systems. Type 2 hypervisors are hosted hypervisors, meaning they are software applications running within a conventional operating system environment. All of a virtual machine’s hardware such as CPU, hard disks, RAM, etc, are emulated and managed through the use of a virtualization application. This virtualization application is installed on the host operating system, while guest operating systems are installed within virtual machines (Campbell & Jeronim, 2006).

Virtualization: Virtualization describes a virtual version of a device or resource, such as a server, storage device, network or even an operating system where the framework separates a resource or request for a service from the underlying physical delivery of that service. With virtual memory, for example, computer software gains access to more memory than is physically installed, via the background swapping of data to disk storage. Similarly, virtualization techniques can be applied to other IT infrastructure layers including networks, storage, laptop or server hardware, operating systems and applications. Devices, applications and human users are able to interact with the virtual resource as if it were a real single logical resource (Vmware, 2011).

VoIP: Voice over Internet Protocol is a technology to make voice calls directly from a computer, a special VoIP phone, or a traditional phone connected to a special adapter over a broadband Internet connection instead of a regular or analog phone line. VoIP services convert voice into a digital signal that travels over the Internet. When calling a regular phone number, the signal is converted to a regular telephone signal before it reaches the destination (FCC, 2009).

Summary

This study examined computing resources in private clouds with VoIP applications to determine if there were relationships between virtual machine settings and voice quality. Chapter 1 defined the purpose and significance for this study, as well as the statement of the problem, and research questions. Additionally, assumptions, limitations, and definitions of terms were provided.

CHAPTER 2

REVIEW OF LITERATURE

Circuit-switched to Packet-switched Telephony

PSTN is circuit-switched telephony which turns voice into analog signals and the signals are transmitted over a dedicated channel between two points for the duration of the call. Voice signals may pass through several switches before a connection is established. And during a call, no other network traffic can use those switches. These signals are carried over a separate data network known as Common Channel Signaling (CCS). The protocol used by CCS is Signaling System 7 (SS7). The PSTN applications and services include telephone call, voicemail, fax, and dial-up Internet connection (Garroppo et al., 1998). Advanced calling features such as calling waiting, call forwarding, three-way calling, call ID, call blocking, etc. are possible due the use of SS7 (Davidson et al., 2006).

The great advantage of circuit-switched telephony is that call quality is high because a dedicated line is being devoted to the call. The major drawback is that circuit-switched telephony is not cost efficient. Circuit-switched calls require a permanent 64 kbps dedicated circuit between the two telephones. That capacity cannot be used for any other purpose. Whether the caller or the person called is talking, the 64kbps connection cannot be used by any other party. The considerable capacity in the network is not being used most of the time. Convergence issue is another disadvantage of circuit-switched telephony. The circuit-switched network is not flexible to support broadband Internet, voice, and video on a dedicated end-to-end connection.

The alternative to circuit-switched network, packet-switched network divides up data into small packets based on the destination address in each packet and then sends them over the network by a variety of different routes, before being reassembled at the end into the format of the original message. The two core protocols are the Transmission Control Protocol and Internet Protocol. TCP is responsible for ensuring that the information is divided into packets that IP can manage and for reassembling the packets back into the complete information on the other end (Jacobs & Eleftheriadis, 1998). IP is in charge of routing packets across the Internet.

Packet switching allows for efficient use of the bandwidth available in a network, by dividing messages into packets to fit into the network capacity better than sending the entire message file intact, as well as routing packets along the least busy lines. Widmer et al. (2004) pointed out that many data networks have limited bandwidth, packet- switch based applications such as VoIP naturally prefer to trade off packet size for packet rate. Packet-switched data network has the capability to use bandwidth only when it is required. This capability, although seemingly small, is a major benefit of packet-switched based voice communication networking.

Circuit-switched networks and packet-switched networks have traditionally taken different responsibilities for communication services. Circuit-switched networks were used for real time communication services such as live voice and video when data must transmit and arrive in the same order. However, packet-switched networks handled data that can withstand certain transmission delays, such as Emails. Packet switching was not used for voice because the breaking up and reassembly of the packets would cause lower voice quality due to the variable delays during the transmission.

However, as the volume of data traffic overtook that of voice, data has become the primary traffic on many networks built for voice (Davidson et al., 2006). Packet-switched network has the capability to use bandwidth only when it is required. That benefit has made packet-switched networks increasingly being used for voice communications as well as data.

VoIP, in contrast to PSTN, is packet-switched telephony. The packet infrastructure of IP network achieves the flexibility in network transmission. Regular data files and multimedia streams are carried in packets, and the network routers between the source and destination are unaware of the upper application level details. VoIP services convert voice into a digital signal that travels over the Internet. If calling a regular phone number, the signal is converted to a regular telephone signal before it reaches the destination. VoIP can make a call directly from a computer, a special VoIP phone, or a traditional phone connected to a special adapter (Figure 1).

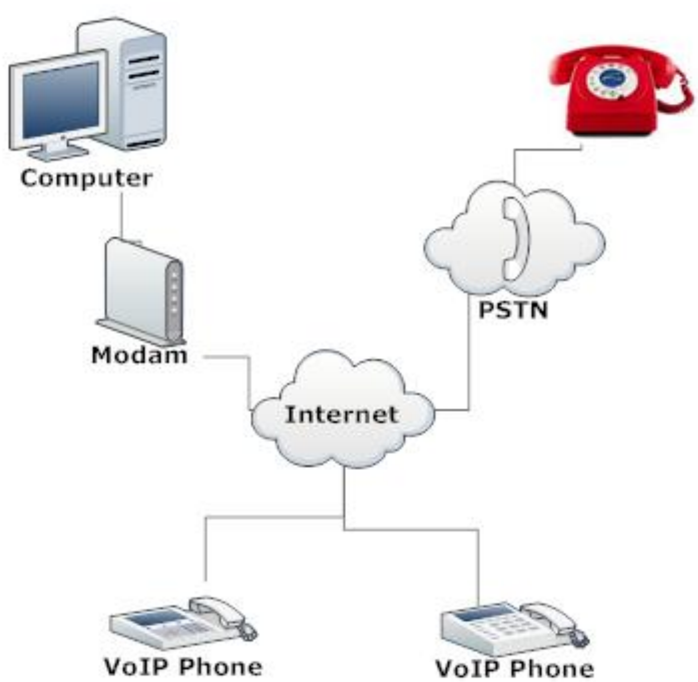

Figure 1. VoIP Network Service

Mehta and Udani (2001) pointed out that the VoIP data processing consists of the four steps: signaling, encoding, transport, and gateway control. 1) Signaling: The purpose of the signaling protocol is to create and manage connections or calls between endpoints. H.323 and the session initiation protocol (SIP) are two widely used signaling standards for call setup and management. 2) Encoding: Once a connection is setup, voice must be transmitted by converting the voice into digitized form, then segmenting the voice signal into a stream of packets. The first step in this process is converting analog voice signals to digital, using an analog-to-digital converter. 3) Transport: Voice samples are inserted into data packets to be carried on the Internet typically by the real time transport protocol (RTP). RTP packets have header fields that hold data needed to correctly reassemble the packets into a voice signal on the other end. Lastly, the encapsulated voice packets are carried as payload by the user datagram protocol (UDP) for ordinary data transmission. At the other end, the process is reversed: the packets are disassembled and put into the proper order, and then the digitized voice is processed by a digital-to-analog converter to render it into analog signals for the called party's handset speaker. 4) Gateway Control: The IP network must then ensure that the real time conversation transported across the telephony system is to be converted by a gateway to another format either because the call is being placed onto the PSTN or for interoperation with a different IP-based multimedia scheme.

The benefit of network resource sharing, together with the network convergence, motivates the wide development and deployment of VoIP. VoIP is an attractive choice of voice communication to organizations as it offers a variety of advantages over PSTN (Collins, 2010):

First of all, convergence of voice and data applications over a single IP based network. It is costly to operate two separate networks than a single one. There are significant financial savings when one infrastructure can be managed more efficiently than two separate networks. Combining two networks will significantly reduce operating cost such as technical support, maintenance, configuration, and upgrade.

Next, communication technologies and computer software applications can be intelligently linked to improve productivity of the working environment by taking advantage of the synergy of consolidation from two systems.

Thirdly, VoIP capitalizes on the benefits of unified communications (UC) to integrate real-time communication services such as instance messaging, telephone and data sharing with non-real-time communication services such as fax, Email and other applications. The UC system provides a consistent unified user interface and user experience across multiple devices and media types. Through the unified communications system, employees can see when their colleagues are logged on line or using the telephone.

Fourthly, VoIP can be used to optimize business processes and enhance human communications by eradicating unnecessary interruptions, managing flows, and eliminating device, media, and location dependencies. With an increasingly mobile workforce, organizations are rarely centralized in one location. VoIP facilitates flexible working practices, whereby members of staff work from home or in a distributed team environment. With improved bandwidth capabilities, VoIP makes video conferencing a viable and cost effective option for discussions between dispersed team workers.

Finally, an organization can integrate customer relationship management with VoIP. For example, converged call centers allow support stuff to answer all customer enquiry mediums, including telephone, Email, web chat, fax, web call back, and instant messaging. Customers appreciate the flexibility of interacting with an organization that can handle feedback from a range of different sources, and are even more inclined to do business with those who can offer an integrated response.

Quality Measurement of VoIP

VoIP has become a mainstream voice communication in many organizations with its advantages in both costs and integrated network architecture of voice and data applications. However that does not mean implementing VoIP without it challenges. The major hurdle remains as to whether the voice quality provided in packet-switched networks can meet the high bar set by the standard circuit-switched phone systems that users have used for a long time and would expect from any competing telephone services (Kondo & Matthews, 2004).

Unlike the circuit-switched network which has a dedicated circuit for the full duration of voice communication session, IP network is constructed with the best effort service model. The packets may not arrive at their destination in the original order or lost during transmission due to different paths that are taken by different packets. Therefore, IP networks have technical hurdles to support real time applications such as VoIP which is sensitive to underlying network performance such as queuing delays, congestion, serialization delays, and packet drops.

Quality of Service (QoS) refers to the capability to differentiate between types of traffic and types of services so that the different types of service and traffic can be treated differently and then guarantee a predicted level of performance to certain data flows in an IP network. The primary goal of QoS is to provide priority including dedicated bandwidth, controlled jitter and latency required by some real time and interactive traffic applications, and improved loss characteristics to different applications or data flows (Ni et al., 2004).

QoS is especially important for the new generation of Internet applications such as VoIP, as some core technologies such as IP were not designed to support prioritized traffic or guaranteed performance levels, making it much more difficult to implement QoS solutions across the Internet. QoS for VoIP is usually described in terms of service and voice quality. Service quality is associated with availability, post-dial delay, and call-completion rates (ITU-T, 2007). Voice quality is the user's experience. The voice quality perceived by the VoIP user is ultimately determined by parameters such as latency, jitter and packet loss (Wright, 2001).

Latency, the time it takes to get data across the network, is the primary indicator of the walkie-talkie effect. Humans are used to having conversations where they both talk at the same time. Most listeners notice when the delay is more that than about 150 ms; when it exceeds 200ms, they find it disturbing and describe the voice quality as poor based on the ITU-T Recommendation G.114 (1996). The latency is actually made up of four components (Kurose & Ross, 2005): 1) Propagation delay is the time to propagate end-to-end across the network. 2). Transmission delay is the time to get through the network devices along the path. 3) Queuing delay is the time that the packets queued in the buffer before being processed. 4) Processing delay is the time required to examine the packet's header and determine where to direct the packet.

Jitter is variance in inter-packet arrival times. Serialization delays from small packets which gets caught waiting for a large packets to transmit normally cause jitter. It can lead to the gaps in the playout of the voice stream. The jitter can be compensated by maintaining a playout buffer at the receiver side (Sreenan et al., 2000), which processes the incoming packets in such a way that early packets have 18 more delay and late packets have less delay. This means that the received voice stream can be recovered at a steady rate. In addition, arriving voice packets that exceed the maximum length of the jitter buffer will be discarded.

Packet loss is defined as the ratio of the packets transmitted successfully to the number of packets transmitted during the measurement interval expressed as a percentage. It includes the network packet loss and the late arrival packet loss dropped at the jitter buffer. During network transmission, packet loss is inevitable. Queuing in the IP routers induces much more unsteady delay during packet transmission. Packet loss can introduce audio distortion because of voice skips and clipping. Moreover, it can also introduce considerable impairment to voice signals. Typically, a packet loss rate of more than 5% is unacceptable for the VoIP users (Wang & Hu, 2004). In order to reach the equivalent level of voice quality in a PSTN, the threshold rate of packet loss should be set below 1% in VoIP networks.

The measurement methods of QoS can be categorized into subjective and objective. Subjective methods measure the quality of voice based on opinions of a group of people, therefore they are the benchmark for objective methods. Objective methods are divided into intrusive methods and non-intrusive methods, according to whether both the original signal and the transmitted signal are monitored to compare the degradation. The intrusive measurement typically uses two input signals: a reference signal (original signal) and a degraded or distorted signal taken from the network to generate output, whereas the non-intrusive measurement is based on observing parameters that permit the individuation of voice signal quality.

Subjective Voice Quality Measurement

The leading subjective measurement of voice quality is the Mean Opinion Score (MOS) described in the ITU-T Recommendation P.800 (1996). MOS is a subjective score of voice quality as perceived by a sizeable listener panel, at least 32 people, listening to speech over a communication system.

MOS scores are the values on a predefined scale from a set of standard and subjective tests, in which test sentences are read aloud by both male and female speakers over the communications medium. A number of listeners are then asked to actually rate the quality of voice. The testing considers a number of factors, including packet loss, circuit noise, talker echo, distortion, propagation time, end-to-end delay, and other transmission problems (Han et al., 2002). All individual scores from the tests are collected to calculate an arithmetic mean, resulting in a final MOS score.

MOS score is normally between 1 and 5, with 5 being the best and 1 the worst, as listed in Table 1 (ITU-T, 1996). MOS of 4.0 is considered toll quality within the telephone industry. Any MOS lower than 4.0 would then be below toll quality level. Toll quality is the telephone conversations quality level typically heard on a wired land line from a local telephone company.

The voice quality often depends on conditions of a given experiment. On one hand, the MOS scoring process follows up with certain requirements such as minimum number of naïve listener in order to obtain accurate and repeatable results. Research found that for a subjective voice quality test, the 95% confidence interval is approximately MOS score of 0.1(Rix et al., 2006), this small difference due to those strict guidelines does not represent an issue in practice. As the subjective test is often time consuming and costly to set up and execute for real time voice quality measurement, the comparable objective mechanisms are required for this handling. On the other hand, MOS gives a prime criterion for the perceptual quality assessments, and provides the base for many objective measurements.

Table 1.

Mean Opinion Score

Mean Opinion Score	Quality of the Speech	Impairment
5	Excellent	Imperceptible
4	Good	Perceptible, but not annoying
3	Fair	Slightly annoying
2	Poor	Annoying
1	Bad	Very annoying

In VoIP, the major factors that affect MOS are jitter, echo, speech compression, and packet loss rate. The variation in delay is known as jitter which can seriously affect the quality of streaming audio. Packets from the source reach the destination with different delays. The packet's delay varies with its position in the queues of the routers along the path between source and destination and this position can vary unpredictably. One method of compensating the variability of arrival rates is to use a jitter buffer between the network layer and the VoIP application. A jitter buffer hands the arriving packets to the processing application in order at a more consistent rate. If the jitter exceeds the size of the jitter buffer, the VoIP packet is discarded resulting in increased packet loss. If the jitter buffer size is increased to minimize the packet loss, then the end-to-end latency is increased due to the time spent in the jitter buffer resulting in a degradation of the user experience caused by high latency.

The second source of impairment is taker echo. Echo is the audible leak-through of a talker’s voice into her own receive path and is manifested in VoIP in two ways. The first manifestation is when the talker hears her own words a short period of time after she says something. However, as long as the delay between talking and hearing is less than 25 ms, the human brain compensates for the echo so that it is not noticeable. This type of echo is often called sidetone. The second manifestation of echo is the perceived echo and is caused when a talker's voice is returned to the talker from the leak-through at the distant end and is typically caused by the remote end instrument. The leak-through in digital portions of the connection does not occur. Because the delay associated with long distance calls is typically above 25 ms, perceived echo is noticeable and must be eliminated.

The third impairment affecting MOS is caused by speech compression. Speech compression is the use of predictive coding to reduce the bit rate needed for transmission of the voice bearer stream. The impairment caused by speech compression is usually referred to as the equipment impairment factor (Ie) and results from the distortion of the voice during the compression process. The codec G.711 introduces the least delay and gives the best voice quality as it does not use compression with an Ie value of 0. The Ie value generally increases proportional to the level of compression. High Ie value means consuming less bandwidth by doing compression. However, the compression reduces the clarity, introduces delay, and makes the voice quality sensitive to lost datagrams.

The last factor that affects MOS is packet loss rate. It is the ratio of the packets transmitted successfully to the number of packets transmitted during the measurement interval expressed as a percentage. Packet loss is the result of lost, corrupted, or discarded voice samples packets. Lost or dropped packets can result in highly noticeable performance issues or jitter with VoIP and will affect all other network applications to a degree (Kurose & Ross, 2005).

Subjective quality assessment quantifies user experience without estimation; thus, it is reliable to evaluate the voice quality of VoIP services. The results of subjective quality assessment can be used for definition of performance targets, appropriate product behavior, as well as benchmark of objective quality measures (Thorpe, 2002).The subjective quality assessment is a psycho-acoustic experiment in which the actual users are involved, however, it is time consuming and costly. Moreover, it is not applicable to real time and in-service quality management in principle. Therefore, an objective means for estimating subjective quality solely from objective characteristics of VoIP services is desired. Objective quality assessment is not a method for deriving a simple indicator of objective performance, but a method for deriving a good indicator of subjective quality in an objective manner.

Signal-Based Objective Voice Quality Measurement

Perceptual Evaluation of Speech Quality (PESQ) is one of the widely used objective methods for measuring VoIP quality of voice based on ITU-T Recommendation P.862. As the result of an integration of the Perceptual Analysis Measurement System (PAMS) and PSQM99, an enhanced version of Perceptual Speech Quality Measure (PSQM) of ITU-T Recommendation P.861, PESQ is a signal-based approach often used in different hardware solutions designed for objective voice quality testing.

PESQ takes into consideration elements such as coding distortions, errors, variable delay, noise, packet loss, and filtering in analog network components, as well as codec (Rix et al., 2002). There is a list of conditions for which the ITU-T recommendation is known to provide inaccurate predictions and not intended to be used. The effects of loudness loss, delay, sidetone, echo, and other impairments related to two-way interaction are not reflected in the PESQ scores. The resulting PESQ quality score can be converted to the subjective Mean Opinion Score using a mapping function described in the ITU-T Recommendation P.862.1.

PESQ estimates the delay impact of the network by using a time alignment algorithm.

Based on the set of delays, PESQ compares an original, unprocessed signal with a degraded version of this signal received on a destination side. This model transforms signals from both sides into one internal representation, which maintains the voice feature of perceptual frequency and loudness. Rango et al. (2006) recommended that the entire PESQ process can be structured in three steps: 1) Signaling preprocessing by taking signal frequency as input, and applies time alignment to the signals. 2) Perceptual modeling by transforming the signals from the sender and receiver sides separately into one internal representation. 3) Cognitive modeling by calculating a difference between reference signal and distorted signal for each time-frequency cell. These differences are considered to produce a MOS score prediction.

The benchmark experiments, performed by the ITU-T, have covered a wide range of conditions such as random and bursty loss, different codecs, etc. and have demonstrated a relatively high correlation with human testing results. The high correlation coefficient of 0.935 with subjective testing results means that the model is accurate under the same experimental conditions. The results can change from test to test but the PESQ score is a good indicator of voice quality within a given experiment.

However, the high correlation with subjective testing may not necessarily translate into high accuracy of the objective approach because correlations ignore absolute differences between subjective and objective scores. The results from the PESQ verification process by ITU-T show that there is only a 40% probability that the PESQ error is less than 0.25 MOS points, and that 30% of PESQ estimates have an error of 0.5 MOS points or larger. Thus, a significant amount of PESQ measurements deviate substantially from the actual MOS.

PESQ cannot be suitable for speech quality measurement in real time because it requires both original and degraded signals to estimate voice quality. The original signal is not available on a destination side. For dynamic voice quality, a computational model is required to estimate the level of voice quality at a given moment of time or during a given period of time.

Therefore, the signal-based measurements are not really well suited to assessing call quality on a data network, since they were specially designed for traditional telephone network. They require invasive hardware probes before beginning VoIP measurements. The models such as PESQ are not based on data network issues of delay, jitter, and datagram loss. Morrissey (2005) found that PESQ performs worse than expected in bursty packet loss scenarios compared to packet loss at a constant rate. The output from those models doesn't provide much information on how to improve the data network for the voice quality.

Computation-Based Objective Voice Quality Measurement

E-model is a computational objective measurement mechanism introduced by ITU-T Recommendation G.107 (2000), ITU-T Recommendation G.113 (2001). The E-model is a transmission planning tool that provides a prediction of the expected voice quality, as perceived by a typical telephone user, for a complete end-to-end (i.e. mouth-to-ear) telephone connection under conversational conditions. The E-model takes into account a wide range of telephony-band impairments, in particular the impairment due to low bit-rate coding devices and one way delay, as well as the classical telephony impairments of loss, noise and echo. It can be applied to assess the voice quality of wireline and wireless scenarios, based on circuit-switched and packet-switched technology.

Based on a mathematical algorithm, E-model transforms the individual transmission parameters into different individual impairment factors that are assumed to be additive on a psychological scale. The algorithm of the E-model also takes into account the combination effects for those impairments in the connection which occur simultaneously, as well as some masking effects.

E-Model characterizes damaging factors to estimate the quality degradation. It calculates a base value, then considering multiple network factors; the degradation is expressed as damage factors, and subtracted from this base. The primary output of the E-model calculations is a scalar quality rating value known as the "Transmission Rating Factor, R".

The R factor could be different under different codecs, and for wideband codecs the R factor may increase to above 100. For example, the R factor for an unimpaired connection may be equal to 110.The calculation of an R factor starts with the unadulterated signal. Basic formula for the E-model is

$$R = Ro - Is - Id - Ie,eff + A$$

- Ro the basic signal-to-noise ratio based on sender and receiver loudness ratings and the circuit and room noise
- Is the sum of real time or simultaneous speech transmission impairments, e.g. loudness levels, sidetone and PCM quantizing distortion. Is does not depend on network environment.
- Id the sum of delay impairments relative to the speech signal, e.g., talker echo, listener echo and absolute delay. Id is the Argument of Delay.
- Ie,eff the equipment impairment factor for special equipment, e.g., low bit-rate coding (determined subjectively for each codec and for each % packet loss). It mostly is affected by codec and packet loss. Values of Ie for specific codecs without packet loss are given in ITU-T Rec. G.113 App. I, and these values are transformed to Ie,eff in case of random packet loss, using the E-model algorithm.

– A the "advantage factor", which allows for an "advantage of access" for certain systems relative to conventional systems, trading voice quality for convenience. While all other impairment factors are subtracted from the basic signal-to-noise ratio Ro, A is added and thus compensates other impairments to a certain amount. It can be used to take into account the fact that the user will tolerate some decrease in transmission quality in exchange for the "advantage of access". Examples of such advantages are cordless and mobile systems or connections into hard-to-reach regions via multi satellite hops.

While the parameters mentioned above describe the individual factors affecting speech transmission quality, it does not imply that such factors are uncorrelated. It is the combined effect of all parameters together which leads to the overall level of speech transmission quality as perceived by the user (Cole & Rosenbluth, 2001). For transmission planning purposes, the E-model is a useful tool for assessing the combined effect of all parameters and hence differentiating between categories of speech transmission quality.

E-Model can be simplified. For example, Cole and Rosenbluth (2001) presented a simplified E-model with only two parameters: loss and delay. The loss rate and delay time are translated into delay impairment and equipment impairment factor by a mapping derived from experimental MOS score curve. The simplified E-model is presented as $R = 94.2 - Ie - Id$.

Cloud Computing Services

Organizations are gradually adopting cloud computing to quickly address key information technology demands in a cost effective way. Foster et al. (2008) list out three main factors contributing to the surge and interests in cloud computing:

1) Rapid decrease in hardware cost and increase in computing power and storage capacity, and the advent of multi-core architecture and modern supercomputers consisting of hundreds of thousands of cores;

2) The exponentially growing data size in scientific instrumentation or simulation and Internet publishing and archiving;

3) The wide-spread adoption of services computing and Web 2.0 applications.

Cloud computing provides organizations with the agility required to move quickly in highly competitive business environments. Combining IT services, cloud computing also provides users dynamic scalability of resources. Based on National Institute of Standards and Technology (NIST)'s Definition of Cloud Computing v15 (Mell & Grance, 2009), cloud computing is composed of five characteristics including on-demand self service, broad network access, resource pooling, rapid elasticity, and measured services (Figure 2).

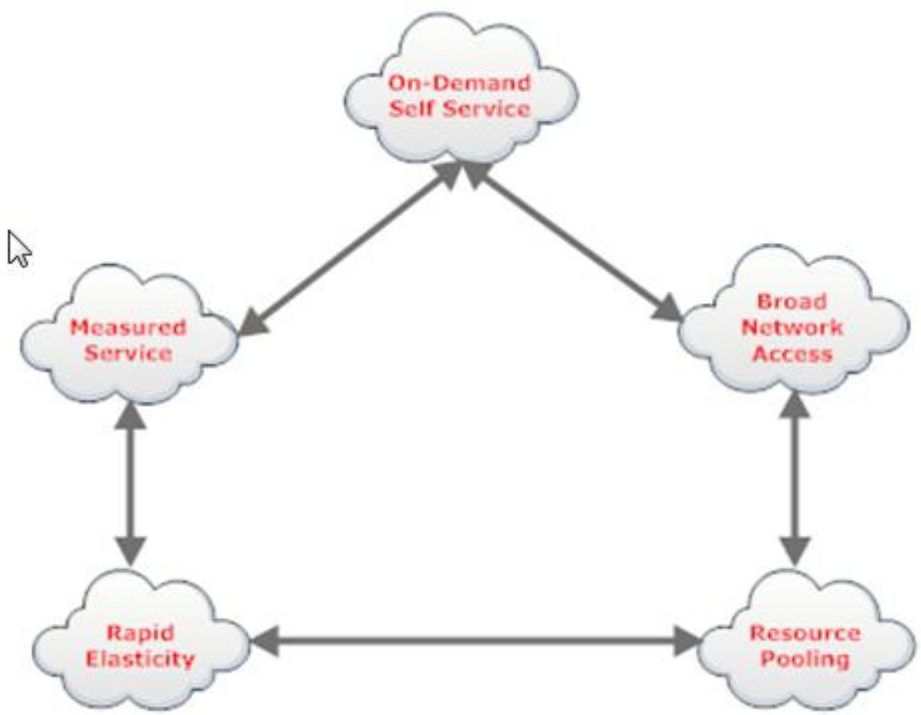

Figure 2. Cloud Computing Characteristics

On-demand self-service. The cloud computing enables a convenient, on-demand network access to a shared pool of computing resources from system software, servers, storage, applications to services than does traditional network computing. Organizations are able to provision, access or release their cloud resources, such as server capacity and network storage, as needed automatically without interacting with anyone from the cloud service providers.

Broad network access. With cloud computing, hardware doesn't have to be physically adjacent to an organization's office or data center. Cloud infrastructure can be located anywhere, especially areas with lower real estate costs. Normally the users have no control or knowledge over the exact location where those computing resources are located. The Internet based characteristic of cloud computing enables computing resources to be accessible from anywhere via the network and accessed through any standardized heterogeneous thin or thick client platforms such as desktop computers, mobile phones, tablet computers, etc., where users can access virtual servers and utilize the resources.

Resource pooling. The computing resources in the cloud are shared among multiple users. Cloud users may be using the same computing resources in the cluster at the same time. With the flexible, automated management, cloud computing can distribute the resources among the cloud's users, with different physical and virtual resources dynamically assigned and reassigned according to user demand. IT departments don't have to engineer for peak-load capacity as the peak load can be spread out among the available computing resources in the cloud.

Rapid elasticity. This attribute of cloud computing allows organizations to increase and decrease their computing resources instantly to application loads. "Cloud computing can offer a means to provide IT resources on demand and address spikes in usage." (Mather et al., 2009, Page 8). In times of growth, the organization can quickly expand its use of computing resources without investing in additional systems hardware and software. The costs related with capital expenditure required for the establishment of network and computer infrastructure tend to be high because organizations must account for spikes in demand for their services. Organizations no longer have to purchase assets for infrequent intensive computing tasks as additional cloud resources can be rapidly increased and released from the clouds with minimal management effort (Miller, 2009).

Measured Service. The usage of computing resources is based on the mechanism of pay-as-you-go charging. The utilized service can be monitored, controlled, and reported providing transparency for both the provider and the users. The cloud provider must essentially measure the amount of service provided and bill the users accordingly in term of the service utilized.

As information technology migrates from the traditional on-premise computing model to the off-site cloud computing model, computing service has evolved to a new level. Mehta et al. (2009) used the acronym “SPI” to stands for the three major services provided through the clouds: software-as-a-service (SaaS), platform-as-a-service (PaaS), and infrastructure-as-a-service (IaaS) (Figure 3).

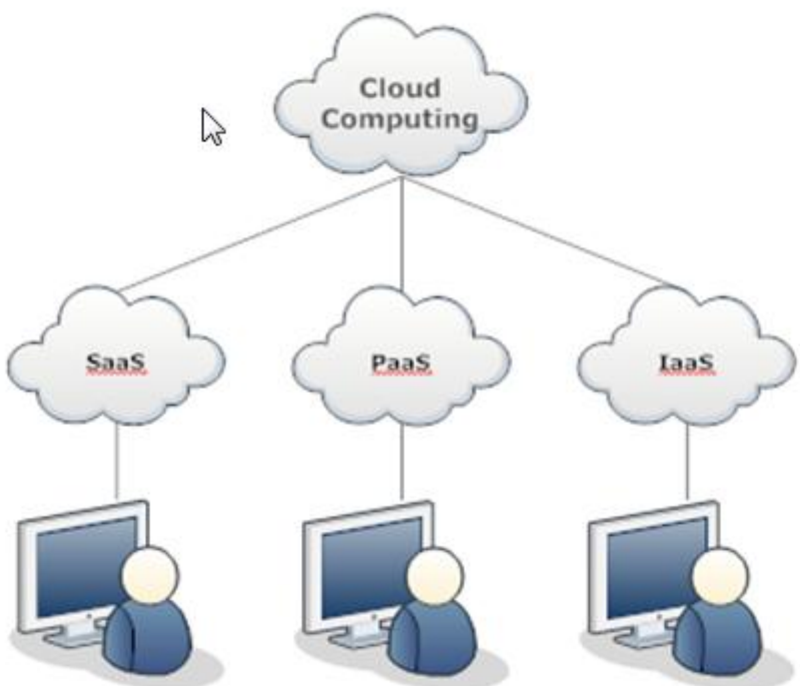

Figure 3. Cloud Computing Services

SaaS delivers special-purpose software that is remotely accessible by users through the Internet on demand. SaaS may be the most common Cloud technology that has been used in the past. Examples of Cloud SaaS deliverables include Web hosted software such as online Turbo Tax, Word processor and Web Email applications, as well as multimedia sharing and social networking solutions. Users simply interact with the software in the cloud via a computer. Nothing is actually stored on local computer and information can be accessed from any device. In the SaaS model, network-based management and access to available software from the cloud rather than at a local site, enable users to access applications remotely via the Internet. The centralized upgrade and patch updating simplifying maintenance and support. Other benefits of SaaS include accelerated deployments, data compatibility across the organization, usage-based pricing model, facilitated, organization-wide collaboration and global accessibility (Rittinghouse & Ransome, 2010).

PaaS is a variation of SaaS whereby a development platform is delivered as a service. PaaS solutions are development environments in the cloud that include computing platforms and solution stacks to support the complete lifecycle of developing and delivering applications and services available entirely through the Internet. An example of PaaS solutions is Salesforce.com where applications are developed on their platforms that interact with the core sales force application and data. These types of deliverables facilitate the deployment of applications for customers and eliminate the need, cost and complexity of buying and managing underlying hardware and software layers required for on-premise solutions. PaaS offerings may include workflow facilities for development life cycle of software design, development, testing, deployment, and hosting, as well as application services such as team collaboration, database integration, web service integration, security, storage, scalability, persistence, application versioning, state management, application instrumentation, and developer community facilitation. These services are provisioned as an integrated solution over the web.

IaaS is the delivery of computer infrastructure as a service. IaaS solutions provide a completed computing infrastructure, generally via virtualized environments that can host databases, servers, power, network storage and backup, and network services such as firewalls and broadband services available entirely through the Internet as a fully outsourced service. The service is typically billed on a utility computing basis, and the quantity of resources consumed typically reflects the level of activity. It is an evolution of web hosting and virtual private server offerings. The pay-as-you-go pricing model allows cloud users to purchase the exact amount of infrastructure required at any specific time and access to quality technology solutions and IT skill sets for in a cost effective way (Mather et al. 2009).

Benefits of Virtualization

Virtualization is a foundational technology platform fostering cloud computing (Mather, Kumaraswamy, & Latif, 2009). Virtualization creates a layer of abstraction as a virtual version between computer systems and software used to operate them such that the underlying computer hardware including physical computers, network resources, and storage can be unified as a pool of resources and resource overlays such as operating system and Web hosting environments can be built on top of them.

In the computer server virtualization, a single physical server can be partitioned to appear as multiple independent virtual machines (VM). The resources of the server itself are hidden, or masked, from users, and the physical server is divided into multiple virtual environments by the software. Multiple server environments such as Windows, Linux, etc can be housed on a single piece of physical hardware.

Server virtualization is the virtual version of server resources, including the number and identity of individual physical servers, CPUs, and OSs, to server users. The software application is used to divide one physical server into multiple isolated virtual environments by the server administrator. The virtual environments are also called virtual private servers, guests, instances, containers or emulations.

A virtual machine is a separate and independent software instance that includes its own operating system and application software. A virtual machine behaves exactly like a physical computer and contains it own virtual CPU, RAM, hard disk and network interface card. Virtual machines are based on the host and guest framework. Each guest runs on a virtual imitation of the hardware layer. The OS system can run without modifications. System administrators are able to create guests that use different OSs. The guest has no knowledge of the host's OS as it's not running on real hardware.

Virtual machines use a hypervisor called virtual machine monitor (VMM) to coordinate instructions to real computing resources such as CPU from the host. Goldberg (1973) classified two types of hypervisor:

Type 1 or native, bare metal hypervisors run directly on the host's hardware to control the hardware and to manage guest operating systems. A guest operating system thus runs on another level above the hypervisor. This model represents the classic implementation of virtual machine architectures.

Type 2 or hosted hypervisors run within a conventional operating system environment. With the hypervisor layer as a distinct second software level, guest operating systems run at the third level above the hardware.

Therefore, Type 1 hypervisor runs directly on the hardware; a Type 2 hypervisor runs on another operating system, such as Linux. As an OS inside a virtual machine is unaware that it’s been deprived of ring 0 privileges, it may attempt to perform restricted functions without permission. In these cases, the hypervisor validates all instructions of the guest OS and manage any executed code in a way that will cause no harm to other OSs running on the host through a "trap and emulate" technique.

An OS can't tell the difference between a physical machine and a virtual machine, nor can applications or other computers on a network. Even the virtual machine thinks it is a real computer. Nevertheless, a virtual machine is composed entirely of software and contains no hardware components whatsoever. As a result, virtual machines offer a number of distinct advantages over physical hardware.

Multiple virtual machines run on a single physical machine, with each virtual machine sharing the resources of that one physical computer across multiple environments. Different virtual machines can run different operating systems and multiple applications on the same physical computer. Virtualization also enables each application to be encapsulated such that they can be configured, deployed, started, migrated, suspended, resumed, stopped, etc., and thus provides better security, manageability, and isolation (Foster et al., 2008).

To support virtual environments with software approaches, a virtualization layer must be placed at certain levels along the machine stack. This virtualization layer thus partitions physical machine resources and maps the virtual requests from a VM to physical requests. The virtualization can take place at several different levels of abstractions, including the Instruction Set Architecture (ISA), Hardware Abstraction Layer (HAL), OS level and application level. At ISA level, virtualization emulates the entire instruction set architecture of a VM in software. At HAL level, virtualization exploits the similarity between the architectures of the guest and host machine, and directly executes certain instructions on the native CPU without emulation. At OS level, virtualization partitions the host OS by redirecting I/O requests, system calls or library function calls. Virtualization at application level is typically an implementation of a virtual environment that interprets binaries.

Different levels of virtualization can differ in isolation strength, resource requirement, performance overhead, scalability and flexibility. In general, when the virtualization layer is closer to the hardware, the created VMs are better isolated from one another and better separated from the host machine, but with more resource requirement and less flexibility. As the virtualization layer is moved up along the machine stack, the performance and scalability of created VMs can be improved.

The most compelling benefit of virtual machines is consolidation of physical servers. VM frees a workload from specific physical hardware devices so that the workload can run on any physical host with the proper computing resources available. Therefore, computing resources are used much more efficiently. Typical server hardware resources are not being used to their full capacity by applications. On average, a typical non-virtualized application server may reach just 5% to 15% utilization, but a physical server that hosts multiple VMs can may reach 50% to 80% utilization (Linthicum, 2009). As more VMs can be hosted on fewer physical servers, server consolidation translates into lower costs for hardware acquisition, maintenance, energy, and cooling system usage (Khnaser, 2010).

Dynamic resource management is another benefit of virtual machines. Whenever a business needs to expand its number of workstations or servers, it is often a lengthy and costly process. Virtual machines can be easily setup. There are no additional hardware costs, no need for extra physical space and no need to wait around. Virtualization also makes it easier for administrators to respond to project requests faster and meet sporadic utilization spikes.

Virtual machines reduce server downtime and improve system security through isolation. While virtual machines can share the physical resources of a single computer, they remain completely isolated from each other as if they were separate physical machines. When a physical machine fails, usually all of its software content becomes inaccessible. All the content of that server becomes unavailable and there is often some downtime to go along with this, until the problem is fixed. Virtual machines are separate entities from one another. Therefore if one of them fails or has a virus, they are completely isolated from all the other software on that physical machine, including other virtual machines. Isolation is an important reason why the availability and security of applications running in a virtual environment is improved comparing with applications running in a traditional, non-virtualized system.

Virtual Infrastructures of Private Clouds

Virtual machines are a fundamental component of virtual infrastructures. While a virtual machine represents the virtual version of the hardware resources of an entire computer, a virtual infrastructure represents the interconnected hardware resources of an entire IT infrastructure including computers, database, network devices and shared storage resources. Organizations of all sizes build virtual server and desktop infrastructures to improve the availability, security and manageability of mission critical applications through a virtual infrastructure.

A virtual infrastructure (Figure 4) shares physical resources of multiple machines across the entire infrastructure. A virtual machine shares the resources of a single physical computer across multiple virtual machines for maximum efficiency. Physical resources can be shared across multiple virtual machines and applications. The multiple servers of the underlying hardware infrastructure are aggregated along with networks and storage into a shared pool of IT resources that can be utilized by the applications when and where they're needed. This resource optimization drives greater flexibility in the organization and results in lower capital and operational costs.

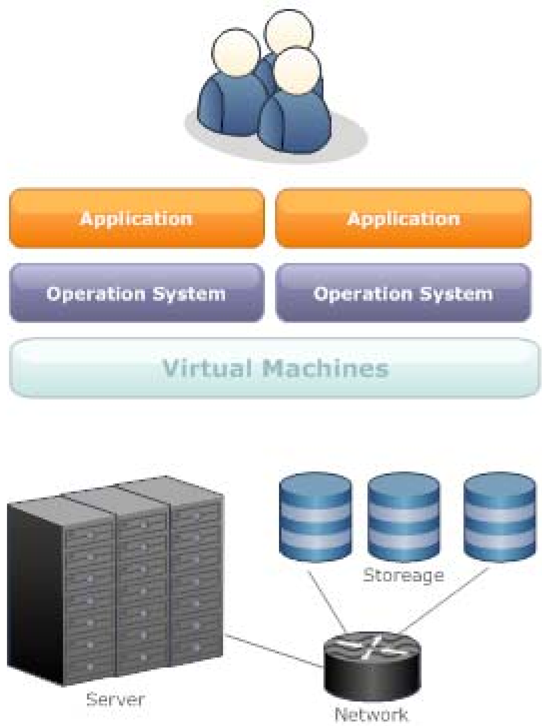

Figure 4. Virtual Infrastructure

In a public cloud, security management and daily base operations are relegated to the third party vendor, who is responsible for the public cloud service offering. Hence, the customer of the public cloud service offering has a low degree of control and oversight of the physical and logical security aspects of a private cloud. Contrary to public clouds, private clouds leverage virtualization to allocate the process load between all available servers, which improves the utilization of each computing resource. In a private cloud, an organization sets up a virtualized environment on its own servers, either in its own data centers or in those of a managed services provider. This structure is useful for organizations that either have significant existing IT investments or feel they absolutely must have total control over every aspect of their infrastructure (Reese, 2009).

The key advantage of private clouds is control. In a private cloud, organizations retain full control over their infrastructure, but also gain some benefits of cloud computing such as the ability to reduce costs without the pitfalls, capitalizing on data security, corporate governance, and reliability concerns (Foley, 2008). Organizations are interested in private clouds because, in many instances, they cannot host their data outside of their firewalls due to privacy and legal issues, but they want to take advantage of the cloud computing architecture. Many of them want to remain in control of their systems and information and have already invested in hardware and software, the cost of which cannot be recovered (Linthicum, 2009).

Private cloud architecture consists of two major components: a base hypervisor and a management server. A base hypervisor creates a layer of abstraction between virtual servers and the underlying hardware. The hypervisor are installed directly on every physical server planned for creating a host and partition it into multiple virtual machines that can run simultaneously, sharing the physical resources of the underlying server. Each virtual machine represents a complete system, with processors, memory, networking, storage and BIOS, and can run an unmodified operating system and applications. The management server is responsible for centralized management of private cloud hosts and their virtual machines.

Review of Similar Study and Previous Research

Mirzoev (2007) at the Indiana State University completed a Ph.D. dissertation entitled "A Hardware-Based Model for Reducing Field Loss by a Video Processing System". The objective of the study was to identify what computer hardware components, such as memory size and hard drive type, contribute to the loss of video fields in order to find an optimal computer hardware model that helps in reducing the number of factors that contribute to the loss of video fields. The study provided statistical analysis of the number of video streaming sources and how various hardware components such as RAM or hard drive affect the problem with lost video fields. Results showed that hard drive variable showed a significant main effect on the number of collected video field, but the amount of RAM did not have a significant main effect of the number of video fields transmitted. Future research with a larger sample size was suggested, particularly in computer hardware components such as RAM and hard drive with more various hardware settings in order to further investigate the influences of RAM and hard drive factors on the number of collected video fields. Additional research could include examining the effects of different CPU technologies on the number of collected video fields. Other areas of future research include looking at another computer component storage controller type that might affect the transmission of video.

Another related Ph.D. dissertation was “Application-specific benchmarking” by Zhang (2001) at the Harvard University. The purpose of the study was to introduce the approach application-specific benchmarking to performance evaluation in order to investigate how the operating system primitives affected the performance of OS dependent applications. The research identified the application performance as a dependent variable to be predicted and the operating system primitives as independent variables expected to influence the dependent variables. The experimental quantitative research was used to analyze performance bottlenecks of complex applications and predict performance on a variety of operating system primitive combinations. Results from the data analysis show one of the operating system primitives with the measurements showing that the memory system had a major influence on the application performance. The discovery that memory system is a significant factor that impacts the OS dependent application performance confirmed the hypothesis that there was a strong relationship between the operating system primitives and application performance. The results also shown that performance of commercial applications on a variety of Java Virtual Machine implementations was able to be predicted from a common optimization. Suggested future research was needed to explore profiling techniques instead of micro-benchmarks to gain a better understanding of the performance characteristics of OS dependent applications.

Shuf (2003) at the Princeton University completed a Ph.D. dissertation entitled “Improving the memory performance of Java workloads” .The study was to target the memory problem in the context of Java workloads with the purpose to identify opportunities for improving memory system performance using runtime techniques within the Java Virtual Machine. The research used a descriptive quantitative method to set memory load, memory store, cache hit, cache miss, heap load, and heap store instructions performance as dependent variables, and the Java workload using SPECjvm98 benchmarks as the independent variable. Results showed that small sections of a Java program normally account for most of the execution time and memory performance problems. The test-retest method was used to confirm the reliability of the data collected. The research results also provided new insight on memory system performance of Java workloads and how it compared with traditional and scientific application workloads.

Adabala (2004) at the Purdue University completed a thesis entitled “Execution of unmodified applications on distributed storage and compute resources”. The objective of the study was to present a virtual machine interface ISA that abstracts access to the computer resources, for the processing of applications on distributed storage and compute resources. The study used experimental methods and evaluated how the underlying computing resources should be virtualized and when unmodified applications developed for non-virtualized environment running and at which layer of system software should the virtualization be implemented. Using the application type as the independent variable, and the memory behavior as the dependent variable, the research provided new insight on how the operating system kernel enables computer systems to support memory-intensive workloads by leveraging virtualized memory and computing resources.

Menon et al. (2005) used the Xenoprof toolkit to analyze performance overheads incurred by networking applications running in Xen virtual machine. Their experimentations quantified virtual machine performance overheads for network I/O device virtualization in uni- and multi-processor systems as the dependent variable. As independent variables, the network I/O and CPU utilization were expected to influence the virtual machine's performance. Their experimental results identified that the network I/O load is the main factor on virtual machine overheads in single and multi-processor systems, while the CPU load was characterized as an extraneous variable in the study.

Many studies have been conducted on the measurement and analysis of VoIP’s voice quality. Utilizing objective measurement methodology, those studies quantified the factors that affect the quality in different working scenarios client-to-client audio conversation, group conference, and calls in mobile environments. Chiang et al. (2006) used an experimental measurement-based approach to evaluate how the performance of VoIP applications was affected by different network conditions. Their research results indicated that voice quality of VoIP is affected by not only bandwidth but also other potential pitfalls.

Summary

The field of real time communication in the private clouds domain is relatively new. Literature specific to virtual machine studies comparing the relationships between hardware settings of virtual machines and voice quality of VoIP was limited. Chapter 2, the Review of Literature, provided an introduction of packet-switched telephony and its quality measurement methods, as well as an overview of literature related to cloud computing services, virtual machine benefits, and virtual infrastructure in private clouds. Additionally, the chapter described reviews of similar studies and previous research.

CHAPTER 3

METHODOLOTY

Overview

Chapter 3, Methodology, provides the statement of problem, research questions, and hypotheses for this study. Additionally, the chapter discusses the Type I error selection, descriptions of variables, experimental design and data collection. Finally, Chapter 3 discusses the statistical analysis to be used and related assumptions.

Statement of Problem

The problem statement for this study was to ascertain whether there is a statistically significant correlation between VoIP voice quality and virtual machine settings such as RAM and hard drive in different network conditions.

Research Questions

The problem for this study was to determine the relationship between virtual machine settings and voice quality of VoIP in a private cloud system. Based on the review of literature and the statement of the problem, questions related to virtual machine settings in cloud computing fields. Research questions included:

1. Do virtual machine settings such as virtual RAM and virtual hard drive play a significant role in affecting voice quality of VoIP after the effect of network bandwidth is removed?

2. Do virtual machine settings and network bandwidth significantly predict the expected voice quality of VoIP?
3. Do virtual machine settings such as virtual RAM and virtual hard drive significantly predict the expected voice quality for VoIP?

Hypothesis Statements

Research questions were further developed into the hypotheses. Hypothesis 1 relates to the research question one, Hypothesis 2 relates to the research question two, and Hypothesis 3 relates to the research question three.

Hypothesis One: R Factor for Virtual RAM and Virtual Hard Drive

- H_{o1}: $\beta j = 0$. There is no statistically significant relationship between the dependent variable of R Factor and the independent variables of virtual RAM and virtual hard drive when network conditions are controlled as a covariate.
- H_{a1}: $\beta j \neq 0$. There is a correlation between the dependent variable of R Factor and the independent variables of virtual RAM and virtual hard drive when network conditions are controlled as a covariate.

Hypothesis Two: R Factor for Virtual RAM, Virtual Hard Drive, Bandwidth

- H_{o2}: $\beta j = 0$. Values of R factor cannot be significantly predicted from values of virtual RAM, virtual hard drive, and network bandwidth.
- H_{a2}: $\beta j \neq 0$. Values of R factor can be significantly predicted from values of virtual RAM, virtual hard drive, and network bandwidth.

Hypothesis Three: R Factor for Virtual RAM and Virtual Hard Drive

- H_{o3}: $\beta j = 0$. Values of R factor cannot be significantly predicted from values of virtual RAM, and virtual hard drive.

- H_{a3}: $\beta j \neq 0$. Values of R factor can be significantly predicted from values of virtual RAM, and virtual hard drive.

Type I Error (Alpha) Selection

The level of significance is denoted by the Greek letter (alpha) and is the probability of committing a Type I error. Alpha specifies how rare the sample results must be to support the rejection of null hypothesis as untenable. Rejecting null hypothesis when it is true is known as Type 1 error. Social scientists often use the 0.05 level as a cutoff: if there is 5% or less chance that a relationship is just due to chance, it can be concluded that the relationship is real, technically, it cannot be determined to accept the null hypothesis that the strength of the relationship is not different from zero. While an alpha level of 0.01 would be desired, it is understood that there is a cost implication in achieving this level. The lower the alpha level, the more the data must depart from the null hypothesis to be significant.

The voice quality results are important in terms of choosing to implement and improve VoIP applications. If the results of this study determined that virtual machine settings have an effect on voice quality (the dependent variable), an organization could use this information to make decisions on IP telephony deployment, network conditions, cloud computing service and VoIP selection. If the results indicated that virtual machine settings do not have an effect on voice quality (the dependent variable), then decisions about IP telephony deployment, network conditions, cloud computing service and VoIP selection would not be impacted.

If the null hypotheses stating that there is no difference in voice quality results for virtual machine settings in a private cloud environment is rejected, and indeed there is no difference, a Type I error would occur. As a result, a decision could be made by an organization to deploy VoIP applications over private cloud platforms. The cost to the organization in this situation would be the expense of deploying or improving VoIP in private clouds. Additionally, efforts towards with a realization and need to stop the efforts later in time could interrupt real communication service of employees. In this case, there would be a false confidence in VoIP deployment decisions as well as underlying platform selection and related deploying efforts.

If the null hypotheses stating that there is no difference in voice quality results for virtual machine settings in a private cloud environment is not rejected. Under this scenario, there is a difference, a Type II error would occur. As a result, a decision could be made by an organization not to deploy VoIP applications over private cloud platforms. The cost to the organization in this situation would be the expense of not redeploying VoIP in private clouds. Additionally, efforts towards with a realization and need to stop the efforts later in time could interrupt real communication service of employees. In this case, there would be a false confidence in VoIP deployment decisions as well as underlying platform selection and related deploying efforts. The organization could potentially lose the cost saving from server virtualization.

Since both types of errors could incur the costs, an alpha level of 0.05 is used for all research hypotheses in this study. However, it is desirable that the results from this study be practical, this would necessitate an increase in the sample size which may call for a lower alpha level. A larger sample size would lower the chances of a type II error, the error of accepting a null hypothesis when the alternative hypothesis is true.

Description of Variables

As discussed in the chapter of literature review, the signal-based objective voice quality measurement tools such as PESQ are not well suited to assessing call quality of VoIP on IP network, since they were specially designed for traditional telephone network. The computation-based objective voice quality measurement tool E-Model was originally developed as a transmission planning tool for telecommunication networks; however, it is widely used for VoIP service quality measurement as an objective measure of quality. The output of an E-model calculation is a single scalar, called an "R factor," derived from delays and equipment impairment factors. Once an R factor is obtained, it can be mapped to an estimated MOS on Table 2 (ITU-T, 2000). The typical range of R is between 50 and 94. The values below 50 are unacceptable, and the maximum rating for a typical telephone connection is 94. Thus, 94 is often taken as the base value for E-Model.

Table 2.

R Factor to MOS Mapping

Range of E-model Rating	Range of MOS Score	Speech transmission quality category	User satisfaction
90–100	4.3–5.0	Best	Very satisfied
80–90	4.0–4.3	High	Satisfied
70–80	3.6–4.0	Medium	Some users dissatisfied
60–70	3.1–3.6	Low	Many users dissatisfied
50–60	2.6–3.1	Poor	Nearly all users dissatisfied
0–50	1.0–2.6	Unacceptable	Not recommended

The voice quality in this research was assessed by measuring R factor, the output of E-model calculation. The dependent variable was R factor. As dependent variable selection is inherent in the research questions, the independent variables in this study were chosen on the basis of key virtual machine settings and their practical significance. The independent variables are: virtual RAM, virtual hard drive, and network bandwidth.

Experimental Design

Quantitative research is an inquiry into an identified problem, based on testing a theory, measured with numbers, and analyzed using statistical techniques, in order to determine whether the predictive generalizations of a theory hold true (Creswell, 2002). For the testing in this research, quantitative research methods and measurement-based experimental quantitative approaches were used in a controlled test environment to evaluate how the voice quality of VoIP applications was affected by different virtual machine settings in a private cloud environment.

The experimental private cloud environment was built up based on VMware vSphere 4 platform. The hypervisor VMware ESXi was installed on two physical Linux servers to create private cloud hosts planned for virtual machines. VMware vCenter Server was installed on two additional PCs to provide centralized control and visibility at every level of virtual infrastructure. vSphere Client, installed on the Windows XP Professional machines, was used to configure different sets of computing resource variables: virtual RAM and virtual hard drive for a virtual machine. Skype 5.1, one of widely used computer-based VoIP applications, was used to create a voice connection between two virtual machines. To simulate real world network environment, FreeBSD 8.1 DummyNet was used to generate different traffic conditions such as bottleneck bandwidth, add the network delay, and provide the specific packet loss rate. Meanwhile, VQManager was used to collect voice quality data for analysis by converting delay, jitter and packet loss rate of each call into R factor. To ensure consistency and accuracy, the same prerecorded streaming voice file repeats playing in the experiment.

Through the experiments, this study observed how the VoIP application reacted to different underlying virtual machine settings under a variety of network conditions, such as bottleneck bandwidth, required bandwidth, and unlimited bandwidth. 120 testing scenarios were established during the experimentation. Through analysis of the collected data, this research figured out which virtual computing resource had significant impact on the voice quality of VoIP calls.

Data Collection

In order to maintain internal validity consistent environmental conditions, the experiment should be conducted under similar prevailing conditions (Keppel, 1991). Access to an even larger selection of the VoIP applications running in different private clouds with different network conditions would have been desirable; however, due to the time limitation, the experiments were performed in a controlled test environment to simulate real world scenarios. The assumption in this instance was that this was a representative sample of all VoIP applications in private clouds. The control objectives for this study were to limit any biasing influences by conducting these tests in an environment that was close to the real world as possible (Levin, 1999).

The data for this study was collected by measuring and observing the outcome of 120 testing scenarios in the experiment. The data obtained from this experimental was recorded in Microsoft Excel and imported to SPSS 16.0 for this analysis. In this instance the independent variables were manipulated then the dependent variable was then measured. The independent variables virtual RAM and virtual hard drive data was collected through the virtual machine settings. Another independent variable Bandwidth was collected through FreeBSD 8.1

DummyNet. Virtual RAM was measured by MB as a scale variable. Virtual Hard drive was measured by gigabyte as a scale variable. Bandwidth was measured by Kbit/s as a scale variable. The R factor data was collected by converting the delay, jitter and packet loss rate of each Skype call in the testing through a voice quality monitoring and reporting tool VQManager (Figure 5). Raw data recorded was exported to the Microsoft Excel and SPSS 16.0 applications for analysis.

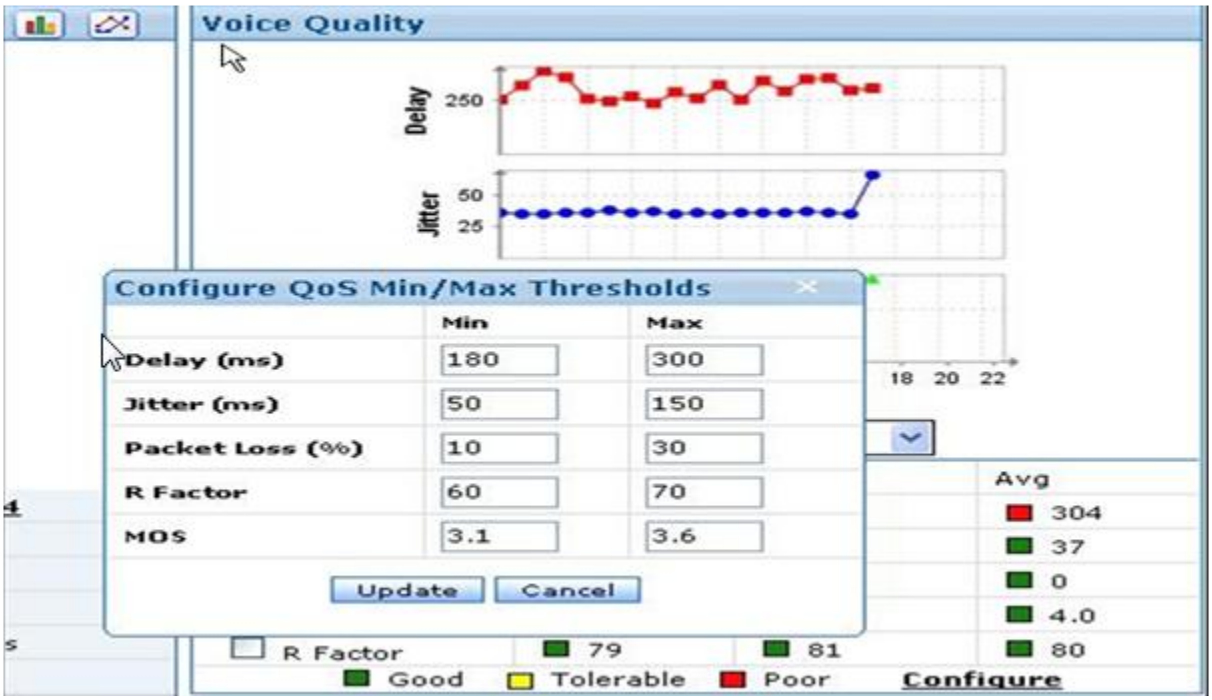

Figure 5. Voice Quality Data Collecting

Statistical Analysis and Assumptions

Descriptive Statistics

Descriptive statistics were used to describe voice quality and virtual machine setting data. The techniques were also used to detail other results as appropriate.

Analysis of Covariance (ANCOVA)

ANCOVA was utilized to test the relationships in Hypothesis One. ANCOVA helps remove the variability in the dependent variable due to the covariate for three types of experimental designs:

- In quasi-experimental (observational) designs, to remove the effects of variables which modify the relationship of the categorical independents to the interval dependents.
- In experimental designs, to control for factors which cannot be randomized but which can be measured on an interval scale.
- In regression models, to fit regressions where there are both categorical and interval independents. This third purpose has become displaced by logistic regression and other methods (Garson, 2002).

A good experimental design keeps all variables constant except the ones are tested but that randomization of subjects should, in the long-run, neutralize the effects of uncontrolled variables (Minium et al., 1999). The purpose of the study was to examine the relationships of virtual machine configurations and voice quality of VoIP in a private cloud environment. Since the experiments were designed to run in a variety of network conditions: bottleneck LAN, minimum bandwidth LAN, recommended bandwidth LAN, and so on. Certainly network bandwidth has main effect on voice quality as bandwidth correlates positively with R factor. Therefore the relationship between network bandwidth and voice quality was not investigated in the study.

ANCOVA was appropriate for the first research hypothesis as this research would like to see what the effect of virtual RAM and virtual hard drive would be if the network has the same bandwidth. That is what ANCOVA does; mathematically finagles the data so that the covariate bandwidth does not affect the relationship of the independent variables with the dependent variable voice quality.

Assumptions of ANCOVA include normal distribution of variables, sufficient sample size, data were randomly collected, independence of the covariate and treatments effect, and homogeneity of regression slop. ANCOVA also assumes that covariates must be an interval or ratio variables, a limited number of covariates, covariates must have a linear relationship to the dependent variable, covariates must be independent of treatment, and data on covariates can be gathered before treatment is administered.

Regression

Multiple Linear Regression was utilized to test the relationships in Hypotheses 2-3. Regression analysis helps to understand how independent variables are related to a dependent variable (Norusis, 2005). The null hypothesis is tested by using the F test and related significance value. If the overall test is not significant, there is no linear relationship between the dependent variable and all of the independent variables, so the null hypothesis cannot be rejected. Assumptions of regression include distribution of values for the dependent variable must be normal, independent or predictor variables must be independent of each other, the relationship between dependent and independent variables must be linear, and variance of the dependent variable distribution must be equal. Regression also assumes that the dependent variable is ratio or nominal or rank and independent variables are either interval or ratio.

Equation 1 was the general formula for all the hypotheses in the study, where $\hat{Y}$ is the dependent value and X_i is the *i*th independent variable. Regression equations were used for the hypotheses in the study.

Equation 1. Standard Regression Equation

$$\hat{Y} = \beta 0 + \beta 1X1 + \beta 2X2 + \ldots + \beta n\ Xn2$$

Equation 2 was the general formula to predict Voice Quality in Hypothesis 2, where Voice Quality was the dependent variable R factor, a ratio variable. RAM was the independent variable virtual RAM, a ratio variable. Hard Drive was the independent variable virtual hard drive, a ratio variable. Bandwidth was the independent variable network bandwidth, a ratio variable.

Equation 2. General Regression Equation of Hypothesis Two

$$\text{Voice Quality} = \beta 0 + \beta RAM_{(1,2,3,4,5)} + \beta \mathit{Hard\ Drive}_{(1,2,3,4)} + \beta \text{Bandwidth}_{(1,2,3,4,5,6)}$$

Equation 3 was the general formula to predict Voice Quality in Hypothesis 3, where Voice Quality was the dependent variable R factor, a ratio variable. RAM was the independent variable virtual RAM, a ratio variable, and Hard Drive was the independent variable virtual hard drive, a ratio variable.

Equation 3. General Regression Equation of Hypothesis Three

$$\text{Voice Quality} = \beta 0 + \beta RAM(1,2,3,4,5) + \beta \mathit{Hard\ Drive}\ (1,2,3,4)$$

Statistics for regression analysis include R, R square (R^2), and Adjusted R square (Adjusted R^2). R is the square root of R-Squared and is the correlation between the observed and predicted values of dependent variable and ranges from 0 to 1 (Norusis, 2006). R^2 is the proportion of variance in the dependent variable attributed to the regression equation and indicates the percentage of observed variability attributed to differences in the independent variables. The values of R^2 range from 0 to 1 and indicate that how the model fit the data well (Norusis, 2005). R^2 is an overall measure of the strength of association and does not reflect the extent to which any particular independent variable is associated with the dependent variable.

Adjusted R^2 attempts to correct R^2 to more closely estimate of how well the model fits and population (Norusis, 2006). Adjusted R-squared is computed using the formula 1 - ((1 - R^2)((N - 1) /(N - k - 1)) where k is the number of predictors.

Partial correlation and part correlation were also conducted as part of the regression tests. Partial and part correlation coefficient values can range from -1 to +1 (Norusis, 2006). The partial correlation remains between two variables after removing the correlation that is due to their mutual association with the other variables. The partial correlation coefficient is the coefficient between the independent and the dependent variable when the linear effects of the other independent variables in the model have been removed from both dependent and independent variables. The square of the partial correlation coefficients indicates what proportion of the unexplained variance in the dependent variable is explained by that variable. Related t-values and p-values can be used to evaluate the significance of the B weights, beta weights, part correlations, and partial correlations (Green & Salkind, 2008). The part correlation remains between the dependent variable and an independent variable when the linear effects of the other independent variables in the model have been removed from the independent variable. It is related to the change in R-squared when a variable is added to an equation. Large absolute values for the part coefficient values indicate that the variable provides unique information about the dependent variable not available from other independent variables in the regression equation.

In addition to normality, assumptions of regression are linearity, homogeneity or equality of variance, and independence (Minium, Clarke & Coladarci, 1999). To test linearity and equality of variance, scatter plots and plots of the residuals can be created for the predicted values of the dependent variable against each of the independent variables (Norusis, 2006). Levene’s test for equality of variances was also used to test the homogeneity of variance. If the p-value resulting from Levene’s test is less than some critical value such as .05, it is unlikely that the differences in sample variances have occurred based on random sampling (Norusis, 2005). As a result, the null hypothesis of equal variances is rejected and it can be concluded that there is a difference between variances.

Durbin-Watson tests were utilized to test for independence and correlation of adjacent residuals which are prediction errors from a regression analysis. The values of residuals range from 0 to 4 (Norusis, 2006). If residuals are not correlated, values are close to 2. Values less than 2 indicate adjacent residuals are positively correlated and values approaching 0 indicate positive correlation. Values greater than 2 indicate a negative correlation and values toward 4 indicate negative autocorrelation.

Summary

Chapter 3, Methodology, provided the statement of the problem, research questions, and hypotheses for this study. Additionally, the chapter discussed the Type I error selection, variables to be utilized, as well as experimental design and data collection. Finally, Chapter 3 discussed the statistical analysis to be used and related assumptions.

CHAPTER 4

RESULTS

This chapter presents results from the data analysis of the experiment which collected data on voice quality of VoIP from a private cloud system with a variety of virtual machine configuration over different network conditions in the controlled test environment. The purpose of the study was to examine the call performance of the VoIP application and analyze the relevance of the voice quality comparing virtual machine configurations in a private cloud environment. Voice quality was based on the output of the E-model calculation, R factor. In addition, the study examined the factor related to network conditions.

The data for this study was collected by measuring and observing the outcome of 120 testing scenarios in the experiment. The data obtained from this experimental was recorded to a spreadsheet, and imported to SPSS 16.0 for this analysis. One independent variable virtual RAM was manipulated in five different settings. Another virtual independent variable hard drive data was configured in four different settings. The controlled independent variable network bandwidth was set in 6 types of bandwidth by the FreeBSD 8.1 DummyNet. R factor data was collected from the reporting tool.

The findings are organized in three parts. Part 1 presents the results and statistical analysis for Hypothesis 1, Part 2 presents the results and statistical analysis for Hypothesis 2, and

Part 3 presents results and statistical analysis for Hypothesis 3. Research questions and related hypotheses are reinstated here:

1. Do virtual machine settings such as virtual RAM and virtual hard drive play a significant role in affecting voice quality of VoIP after the effect of network bandwidth is removed?
2. Do virtual machine settings and network bandwidth significantly predict the expected voice quality of VoIP?
3. Do virtual machine settings such as virtual RAM and virtual hard drive significantly predict the expected voice quality for VoIP?

Research questions were further developed into the hypotheses. Hypothesis 1 relates to the research question one, Hypothesis 2 relates to the research question two, and Hypothesis 3 relates to the research question three.

Hypothesis One: R Factor for Virtual RAM and Virtual Hard Drive

- H_{o1}: $\beta j = 0$. There is no statistically significant relationship between the dependent variable of R factor and the independent variables of virtual RAM and virtual hard drive when network conditions are controlled as a covariate.
- H_{a1}: $\beta j \neq 0$. There is a correlation between the dependent variable of R factor and the independent variables of virtual RAM and virtual hard drive when network conditions are controlled as a covariate.

Hypothesis Two: R Factor for Virtual RAM, Virtual Hard Drive, Bandwidth

- H_{o2}: $\beta j = 0$. Values of R factor cannot be significantly predicted from values of virtual RAM, virtual hard drive, and network bandwidth.

- H_{a2}: $\beta j \neq 0$. Values of R factor can be significantly predicted from values of virtual RAM, virtual hard drive, and network bandwidth.

Hypothesis Three: R Factor for Virtual RAM and Virtual Hard Drive

- H_{o3}: $\beta j = 0$. Values of R factor cannot be significantly predicted from values of virtual RAM, and virtual hard drive.
- H_{a3}: $\beta j \neq 0$. Values of R factor can be significantly predicted from values of virtual RAM, and virtual hard drive.

Part 1 - Results of Hypothesis One Testing

ANCOVA was performed to evaluate the relationship between voice quality and virtual machine settings when the effect of network bandwidth is removed. The independent variables were virtual RAM and virtual hard drive. The dependent variable was R factor.

In order to examine the variables, the descriptive statistics were derived in SPSS 16.0. These statistics are shown on Table 3 below. From this table the variables in this study have no disproportionate share of missing values. This is indicated by the last value labeled Valid N (listwise) which shows that there are no missing values. The skewness statistics for most variables fall within the -1 and +1 which suggests they are normally distributed. The Kurtosis statistics for most variables fall within -2 and +2, which suggest a normal distribution for those variables (Miles, & Shevlin, 2001).

Table 3.

Descriptive Statistics

		Bandwidth	Virtual Hard Drive	Virtual RAM	R Factor
N	Statistic	120	120	120	120
Minimum	Statistic	32	10	128	15.00
Maximum	Statistic	1014	40	2048	65.00
Mean	Statistic	334.33	25.00	793.60	42.5333
Std. Deviation	Statistic	344.637	11.227	701.203	16.29916
Variance	Statistic	118774.342	126.050	491685.217	265.663
Skewness	Statistic	1.098	.000	.900	-.169
	Std. Error	.221	.221	.221	.221
Kurtosis	Statistic	-.196	-1.367	-.651	-1.263
	Std. Error	.438	.438	.438	.438

The limitations of ANCOVA were tested from the correlation between of Bandwidth and R factor. The independent variable Bandwidth is a ratio variable. The correlation coefficient on Table 4 indicated that there is correlation between R factor and bandwidth. As the observed significance level is 0.000 (<.05), it can be determined that the covariate Bandwidth has a linear relationship with the dependent variable. Therefore the covariate bandwidth has met the limitations for ANCOVA analysis in this study.

Table 4.

Correlation Coefficient between R Factor and Bandwidth

		R Factor	Bandwidth
R Factor	Pearson Correlation	1	.843**
	Sig. (2-tailed)		.000
	N	120	120
Bandwidth	Pearson Correlation	.843**	1
	Sig. (2-tailed)	.000	
	N	120	120

**. Correlation is significant at the 0.01 level (2-tailed).

The assumption of homogeneity of regression slopes was tested from the interaction terms on Table 5. For the interaction terms, virtual RAM*Bandwidth, virtual hard drive*Bandwidth, their F values are 0.152 and 0.008, with the significance levels 0.962 and 0.999 separately. Since the significance levels are greater than 0.05, there was no evidence of violation of the equal slopes assumptions.

Table 5.

Homogeneity of Regressions Slopes Test for ANCOVA

Dependent Variable: R Factor

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	22288.005[a]	15	1485.867	24.647	.000	.780
Intercept	52101.840	1	52101.840	864.251	.000	.893
Memory * Bandwidth	36.633	4	9.158	.152	.962	.006
Hard Drive * Bandwidth	1.455	3	.485	.008	.999	.000
Memory	446.876	4	111.719	1.853	.124	.067
Hard Drive	2.436	3	.812	.013	.998	.000
Bandwidth	21238.483	1	21238.483	352.298	.000	.772
Error	6269.695	104	60.286			
Total	240742.000	120				
Corrected Total	28557.700	119				

a. R Squared = .780 (Adjusted R Squared = .749)

The assumption of homogeneity of variance was tested from the Levene's test. Results from Table 6 reveal no evidence of violation of the equal variances assumptions. The F value is .148, with an observed significance level which is greater than .05.

Table 6.

Levene's Test of Equality of Error Variances[a]

Dependent Variable: R Factor

F	df1	df2	Sig.
.148	19	100	1.000

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + memory * bandwidth + hard drive * bandwidth + memory + hard drive + bandwidth

From the results of ANCOVA on Table 7, the R Square which is the measure of how much of the variability in the outcome is accounted for by the predictors is 0.779. This suggests that the model accounts for 77.9 % of variance in the independent variables. In the F and Sig columns, there is evidence of the effects of one independent variable on the dependent variable when the network bandwidths are controlled as a covariate: the F value of virtual RAM is 3.966, with a significance of 0.005 (<.05), therefore, virtual RAM has significant effects on R factor when the effects of Bandwidth are removed. The F value of virtual hard drive is 0.006, with a significance of 0.999, greater than 0.05, thus virtual hard drive has no significant effect on R factor when the effects of bandwidth are removed.

Table 7.

ANCOVA Outputs

Dependent Variable: R Factor

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	22253.183[a]	20	1112.659	17.472	.000	.779
Intercept	52101.840	1	52101.840	818.157	.000	.892
Bandwidth	21238.483	1	21238.483	333.508	.000	.771
Virtual RAM	1010.200	4	252.550	3.966	.005	.138
Virtual Hard Drive	1.233	3	.411	.006	.999	.000
Virtual RAM * Virtual Hard Drive	3.267	12	.272	.004	1.000	.001
Error	6304.517	99	63.682			
Total	240742.000	120				
Corrected Total	28557.700	119				

a. R Squared = .779 (Adjusted R Squared = .735)

The null hypothesis states that there is no correlation between the dependent variable R factor and the independent variables virtual RAM and virtual hard drive when network bandwidths are controlled as a covariate. The Adjusted R Square =0.735 which suggests that the model accounts for the variance in the R factor. The results of the ANCOVA testing suggest that the null hypothesis is rejected as there is correlation between the dependent variable R factor and the independent variable virtual RAM when network bandwidth is controlled as a covariate. When the effect of network condition is removed, there is relationship between one virtual machine setting virtual RAM and voice quality. The voice quality is impacted by a variety of virtual RAM when network bandwidths are the same.

Part 2 - Results of Hypothesis Two Testing

Regression was performed to evaluate whether the expected voice quality can be significantly predicted by virtual machine settings and network bandwidth. The independent variables were virtual RAM, virtual hard drive and bandwidth. The dependent variable was R factor. The results were significant and the apparent unstandardized coefficients, Bs, are the weights to be used in Equation 4 for the predicted voice quality.

Equation 4. Estimated Regression Equation of Hypothesis Two

$$\text{Voice Quality} = 25.953 + 0.015\text{RAM} + 0.008\text{HardDrive} + 0.039\text{Bandwidth}$$

The correlation coefficients on Table 8 indicate that R factor has correlations with bandwidth and virtual RAM with the observed significance level smaller than 0.05, it can be determined that the independent variables bandwidth and virtual RAM have a linear relationship with the dependent variable. The correlation coefficient of virtual hard drive and the dependent variable has observed significance level greater than 0.05 suggests that there is no relationship between two variables.

Table 8.

Correlations between R Factor, Bandwidth, Virtual RAM, and Virtual Hard Drive

		R Factor	Bandwidth	Virtual Hard Drive	Virtual RAM
Pearson Correlation	R Factor	1.000	.862	.006	.168
	Bandwidth	.862	1.00	.000	.000
	Virtual Hard Drive	.006	.000	1.000	.000
	Virtual RAM	.168	.000	.000	1.000
Sig. (1-tailed)	R Factor	.	.000	.475	.034
	Bandwidth	.000	.	.500	.500
	Virtual Hard Drive	.475	.500	.	.500
	Virtual RAM	.034	.500	.500	.
N	R Factor	120	120	120	120
	Bandwidth	120	120	120	120
	Virtual Hard Drive	120	120	120	120
	Virtual RAM	120	120	120	120

The R square in Table 9 indicates that 77.2% of the observed variability in the R factor is attributable to differences in virtual RAM, virtual hard drive, and bandwidth. It suggests that three independents variables virtual RAM, virtual hard drive, and bandwidth together are linearly related to the dependent variable.

Table 9.

Hypothesis Two Regression Model Summary

		Model
		1
R		.879[a]
R Square		.772
Adjusted R Square		.766
Std. Error of the Estimate		7.49528
Change Statistics	R Square Change	.772
	F Change	130.777
	df1	3
	df2	116
	Sig. F Change	.000

a. Predictors: (Constant), Virtual RAM, Virtual Hard Drive, Bandwidth

The ratio of the regression mean square to the residual mean square, F, is 130.78 in Table 10, with the observed significance level smaller than 0.05, the null hypothesis that the values of R factor cannot be predicted from virtual RAM, virtual hard drive, and bandwidth is rejected.

Table 10.

Hypothesis Two ANOVA Table

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	22040.907	3	7346.969	130.777	.000[a]
	Residual	6516.793	116	56.179		
	Total	28557.700	119			

a. Predictors: (Constant), Virtual RAM, Virtual Hard Drive, Bandwidth
b. Dependent Variable: R Factor

Based on the regression coefficients on Table 11, the values of R factor can be predicted from Equation 4. R factor= 25.953 + (0.039 x bandwidth) + (0.008 x virtual hard drive) + (0.015 x virtual RAM). Since the observed significance level of two dependent variables is smaller than .05 for t, the coefficients for bandwidth and virtual RAM are not zero.

Table 11.

Hypothesis Two Regression Coefficients[a]

		Model			
		1			
		(Constant)	Bandwidth	Virtual Hard Drive	Virtual RAM
Unstandardized Coefficients	B	25.953	.039	.008	.015
	Std. Error	1.964	.002	.061	.004
Standardized Coefficients	Beta		.862	.006	.168
t		13.213	19.443	.131	3.777
Sig.		.000	.000	.896	.000
95% Confidence Interval for B	Lower Bound	22.063	.035	-.113	.007
	Upper Bound	29.843	.043	.129	.023
Correlations	Zero-order		.862	.006	.168
	Partial		.875	.012	.331
	Part		.862	.006	.168
Collinearity Statistics	Tolerance		1.000	1.000	1.000
	VIF		1.000	1.000	1.000

a. Dependent Variable: R Factor

Part 3 - Results of Hypothesis Three Testing

Regression was performed for the part to evaluate whether the expected voice quality can be significantly predicted by virtual machine settings. The independent variables were virtual RAM, and virtual hard drive. The dependent variable was R factor. The results were not significant and the apparent unstandardized coefficients, Bs, are the weights to be used in Equation 5 for predicted voice quality.

Equation 5. Estimated Regression Equation of Hypothesis Three

$$\text{Voice Quality} = 38.913 + 0.015\text{RAM} + 0.008\text{Hard Drive}$$

The correlation coefficients in Table 12 indicate that there is correlation between R factor and virtual RAM. The observed significance level is 0.034 (<0.05), it can be determined that the independent variable virtual RAM has a linear relationship with the dependent variable. The correlation coefficient of virtual hard drive and the dependent variable has observed significance level greater than 0.05 suggests that there is no relationship between two variables.

Table 12.

Correlation Coefficient between R Factor, Virtual RAM, and Virtual Hard Drive

		R Factor	Virtual Hard Drive	Virtual RAM
Pearson Correlation	R Factor	1.000	.006	.168
	Virtual Hard Drive	.006	1.000	.000
	Virtual RAM	.168	.000	1.000
Sig. (1-tailed)	R Factor	.	.475	.034
	Virtual Hard Drive	.475	.	.500
	Virtual RAM	.034	.500	.
N	R Factor	120	120	120
	Virtual Hard Drive	120	120	120
	Virtual RAM	120	120	120

The R square in Table 13 indicates that only 2.8% of the observed variability in the R factor is attributable to differences in virtual RAM and virtual hard drive. Therefore two independents variables virtual RAM and virtual hard drive are not significantly linearly related to the dependent variable R factor.

Table 13.

Hypothesis Three Regression Model Summary

		Model
		1
R		.168[a]
R Square		.028
Adjusted R Square		.011
Std. Error of the Estimate		15.40210
Change Statistics	R Square Change	.028
	F Change	1.691
	df1	2
	df2	117
	Sig. F Change	.189

a. Predictors: (Constant), Virtual RAM, Virtual Hard Drive

The ratio of the regression mean square to the residual mean square, F, is 1.691 in Table 14, with the observed significance level greater than 0.05, therefore, the null hypothesis that the value of R factor cannot be predicted from values of virtual RAM and virtual hard drive is not rejected.

Table 14.

Hypothesis Three ANOVA Table

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	802.424	2	401.212	1.691	.189[a]
	Residual	27755.276	117	237.225		
	Total	28557.700	119			

a. Predictors: (Constant), Virtual RAM, Virtual Hard Drive

b. Dependent Variable: R Factor

Based on the regression coefficients on Table 15, the values of R factor can be predicted from Equation 5. R factor = 38.913 + (0.008 x virtual hard drive) + (0.015 x virtual RAM). Since the observed significance level of two dependent variables is greater than .05 for t, the coefficients for bandwidth and virtual RAM are zero.

Table 15.

Hypothesis Three Regression Coefficients[a]

		Model		
		1		
		(Constant)	Virtual Hard Drive	Virtual RAM
Unstandardized Coefficients	B	38.913	.008	.015
	Std. Error	3.797	.126	.008
Standardized Coefficients	Beta		.006	.168
t		10.249	.064	1.838
Sig.		.000	.949	.069
95% Confidence Interval for B	Lower Bound	31.394	-.241	-.001
	Upper Bound	46.432	.257	.031
Correlations	Zero-order		.006	.168
	Partial		.006	.168
	Part		.006	.168
Collinearity Statistics	Tolerance		1.000	1.000
	VIF		1.000	1.000

a. Dependent Variable: R Factor

Summary

This chapter presented results from the data analysis of the experiment in a controlled test environment. Voice quality was collected based on virtual machine setting in a variety of network conditions. The information was presented in three parts. Part 1 provided results based for Hypothesis 1, Part 2 presented results based for Hypothesis 2 and Part 3 presented results based for Hypothesis 3.

CHAPTER 5

CONCLUSIONS AND RECOMMENDATIONS

Chapter 5 provides a summary, conclusions, and recommendations for further research. The first section provides a summary of the problem statement, the purpose of the study, research questions and hypotheses, and methodology. The second section presents conclusions for the results of findings. The third section offers recommendations for further research.

Summary

The evolution to cloud computing has advanced rapidly over the last few years. As a critical component of private clouds, however, virtualization may influence real time communication applications because adding a layer of virtualization software on the computer server adds overhead to the overall system. There has been an increasing interest to examine the impact of virtualization on performance loss. The reason behind this increased interest is to do with the growing adoption of clouds computing by organizations which expect existing software applications to be able to run on virtual machines and to perform as good as on physical servers. Since real time communication services require a certain level of system performance and availability to address communication latency and overhead bottleneck, it is essential to investigate potential performance implications of private clouds on VoIP applications.

Purpose of Study

The purpose of this study was to investigate the performance of VoIP applications in private clouds and analyze the relevance of the voice quality with virtual machine settings in a variety of network conditions.

Research Questions

Research questions included,

1. Do virtual machine settings such as virtual RAM and virtual hard drive play a significant role in affecting voice quality of VoIP after the effect of network bandwidth is removed?
2. Do virtual machine settings and network bandwidth significantly predict the expected voice quality value of VoIP?
3. Do virtual machine settings such as virtual RAM and virtual hard drive significantly predict the expected voice quality for VoIP?

Hypotheses

Research questions were further developed into the hypotheses. Hypothesis 1 relates to the research question one, Hypothesis 2 relates to the research question two, and Hypothesis 3 relates to the research question three.

Hypothesis One: R factor for Virtual RAM and Virtual Hard Drive

- H_{o1}: $\beta j = 0$. There is no statistically significant relationship between the dependent variable of R factor and the independent variables of virtual RAM and virtual hard drive when network conditions are controlled as a covariate.

- H_{a1}: $\beta j \neq 0$. There is a correlation between the dependent variable of R factor and the independent variables of virtual RAM and virtual hard drive when network conditions are controlled as a covariate.

Hypothesis Two: R factor for Virtual RAM, Virtual Hard Drive, Bandwidth

- H_{o2}: $\beta j = 0$. Values of R factor cannot be significantly predicted from values of virtual RAM, virtual hard drive, and network bandwidth.
- H_{a2}: $\beta j \neq 0$. Values of R factor can be significantly predicted from values of virtual RAM, virtual hard drive, and network bandwidth.

Hypothesis Three: R factor for Virtual RAM and Virtual Hard Drive

- H_{o3}: $\beta j = 0$. Values of R factor cannot be significantly predicted from values of virtual RAM, and virtual hard drive.
- H_{a3}: $\beta j \neq 0$. Values of R factor can be significantly predicted from values of virtual RAM, and virtual hard drive.

Methodology

The method of study included quantitative research methods and measurement-based experimental approaches. A comparative analysis of different virtual machine configurations with different network conditions was performed in a controlled test environment. To simulate real world network environment, DummyNet was used to generate different traffic conditions such as bottleneck bandwidth, add the network delay, and provide the specific packet loss rate. The data obtained from this experimental was recorded in Microsoft Excel and imported to SPSS 16.0 for this analysis. Through the experimental study, the statistical technique ANCOVA was used to examine the effects of virtual machine settings on voice quality. Multiple Linear

Regression analysis was used to test whether voice quality can be predicted from values of virtual computing resources in private clouds.

Conclusions

The data for this study was collected by measuring and observing the outcome of 120 testing scenarios in the experiment. The statistical techniques ANCOVA was used to examine the effects of virtual machine settings on voice quality. Multiple Linear Regression analysis was used to test whether voice quality can be predicted from values of virtual computing resources in private clouds.

The first null hypothesis that there is no statistically significant relationship between the dependent variable of R factor and the independent variables of virtual RAM and virtual hard drive when network conditions are controlled as a covariate was rejected; there was a significant relationship between voice quality and virtual machine settings when the network condition was the same. The ANCOVA analysis conducted in this paper leads to the conclusion that virtual RAM of the virtual machine in private clouds had significant effects on voice quality when the effect of network condition was removed, and virtual hard drive of the virtual machine in private clouds did not have a significant effect on voice quality when the effect of network condition was removed.

The second null hypothesis that values of R factor cannot be significantly predicted from values of virtual RAM, virtual hard drive, and network bandwidth was rejected; there was a significant relationship between values of R factor and virtual RAM, virtual hard drive, and network bandwidth. The regression analysis leads to the conclusion that the change in R factor value was significantly attributable to differences in virtual RAM, virtual hard drive, and bandwidth. The voice quality of VoIP applications in private clouds can be predicted from three factors including bandwidth of the network, virtual RAM, and virtual hard drive of virtual machine together in Equation 4.

The voice quality is always an important factor for organizations to make a decision on implementation of VoIP applications in private clouds. Without a solid quality communication service, the daily operations of an organization will be in jeopardy. As the results of this study determined that virtual machine settings and network bandwidth have an effect on voice quality (the dependent variable), an organization could use the Equation 4 to make decisions on IP telephony deployment based on the expected voice quality, available network capacity, and computing resource configurations in private cloud environments.

Equation 4 shows that the network bandwidth has more weight to statistically predict the voice quality than RAM and hard drive. With limited budget and resources, an organization can use this equation to improve the system to have the best quality of voice under limited budget. The organization may be a better off increasing network bandwidth over RAM and hard drive to improve voice quality in a cost effective way. When the scenario falls into the constraint of network capacity, the organization may increase RAM for improvement of voice quality as RAM has a relationship with voice quality.

The third null hypothesis that values of R factor cannot be significantly predicted from values of virtual RAM, and virtual hard drive was not rejected. The regression equation was changed as Equation 5 when the network bandwidth was factored into a constant.

The regression analysis leads to the conclusion that the change in R factor was not significantly attributable to differences in virtual machine setting on RAM and hard drive when the network bandwidth stays as constant. The data collected from the experimentation showed that under a typical network bandwidth, the increase on RAM had limited impact on voice quality, and the size of hard drive barely had influence on voice quality. For instance, when the bandwidth was set up as 128 Kb/s in the simulation testing, R factor had no change with the increase of hard drive; however R factor had a small increase from 42 to 44 when RAM was doubled to 256 Mb. The improvement on voice quality may not be able to justify the cost increase due to augment of RAM. Therefore the statistical analysis and verification of testing data leads to the conclusion that the voice quality of VoIP applications in private clouds may not be significantly improved from the increase in virtual machine settings when the factor of network bandwidth was taken out of the consideration.

Recommendations for Future Studies

The findings from this study were based on testing one VoIP application over the private cloud platform and the sample was collected through the experiments in a simulation environment. They should be reflective of the population from which the sample data was drawn. It may be possible to generalize these results to other VoIP applications in a private cloud environment. The analysis from ANCOVA can be used to plan virtual server capacity for VoIP in private clouds. The results from Multiple Linear Regression analysis can be used to improve the performance of VoIP applications.

This research provides a starting point for examine the effects of virtual machine settings on voice quality and developing a computing resource model in private cloud to improve voice performance. Clearly, opportunities exist to do a more in-depth quantitative and qualitative study using multiple perspectives on a given VoIP application in private clouds.

Future research can examine the basic formula for predicted voice quality. When VoIP application runs over the virtual server in private clouds, many different factors are involved in the process. The voice quality experienced by end users may be affected by not only virtual

RAM, virtual hard drive, and LAN bandwidth, but also other variables such as virtualization software, hardware integration, operating system, and CPU of host and client machines, etc. The findings will be more global if more related factors are taken into account.

The findings of this study indicate that the real time communication application suffered from performance loss with a lower R factor in private clouds than in a non-virtualized environment with comparative network capacity and computing resources. Server virtualization creates some significant challenges to real time communications in private clouds due to interactions between the underlying virtual machine monitor and other virtual machines. Adding a layer of virtualization software on the computer server adds overhead to the overall system. Therefore, virtual infrastructure desires further research to reduce system capacity loss and application performance degradation and bring more cost savings to adoption of private clouds.

REFERENCES

Adabala, S. (2004). Execution of unmodified applications on distributed storage and compute resources. Ph.D. dissertation, Purdue University, United States -- Indiana. Retrieved February 20, 2010, from Dissertations & Theses: Full Text. (Publication No. AAT 3154576).

Avanada (2009). Global Survey of Cloud Computing. Retrieved March 22, 2010, from http://www.avanade.com/_uploaded/pdf/avanadethoughtleadershipcloudsurveyexecutivesummary833173.pdf

Babcodk, C. (2009). "Why 'Private Cloud' Computing Is Real -- And Worth Considering," InformationWeek, Issue 1204.

Barham, P. et al. (2003). Xen and the Art of Virtualization. In *Proceedings of the nineteenth ACM symposium on Operating systems principles* (SOSP '03). ACM, New York, NY, USA, 164-177. DOI=10.1145/945445.945462 http://doi.acm.org/10.1145/945445.945462

Bittman, T. (2009). "Private Cloud Computing is Real – Get Over It," *Gartner Blog Network.*

Brodkin, J. (2008). "Private cloud networks are the future of corporate IT," *Network World,* Vol.12.

Campbell, S. & Jeronim, M. (2006). *Applied Virtualization Technology*, Hillsboro, OR: Intel Press, 2006.

Cherry, S. (2005). "Seven myths about voice over IP," *IEEE Spectrum*, vol. 42, pp. 52 – 57.

Chiang W., Wei-Cheng X., & Cheng Fu C. (2006). "A Performance Study of Voip Applications: MSN vs. Skype," *in Proceedings of MULTICOMM*, Turkey, pp. 13-18.

Cole, R. & Rosenbluth, J. (2001). "Voice over IP performance monitoring," *ACM SIG-COMM Computer Communication Review*, vol. 31, no. 2, pp. 9–24.

Collins, D. (2010). *Carrier Grade VoIP*. Columbus, OH: McGraw-Hil.

Cooper, K., & Xu, L. (2002). An efficient static analysis algorithm to detect redundant memory operations. In Proceedings of the 2002 workshop on memory system performance (pp. 92-107). Berlin, Germany: ACM Press.

Creswell, J.W. (2002). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches.* Thousand Oaks, CA: Sage Publications.

Davidson, J., Peters, J., Bhatia, M., Kalidindi,S., & Mukherjee, S. (2006). *Voice over IP Fundamentals (2nd Edition).* Cisco Press.

FCC (2009). "Voice over Internet Protocol (VoIP)," *Consumer Facts*. Retrieved March 14, 2010, from http://www.fcc.gov/cgb/consumerfacts/voip.pdf

Foley, J. (2008). "Private Clouds Take Shaper", *InformationWeek,* issue 1104.

Foley, J. (2009). "Survey: Fear Slows Cloud Computing Adoption," *InformationWeek*, issue 1203.

Foster, I., Zhao, Y., Raicu, I., & Lu, S. (2008). Cloud Computing and Grid Computing 360-Degree Compared. IEEE Grid Computing Environments (GCE08).

Garson, G. D. (2002.). *Guide to Writing Empirical Papers, Theses, and Dissertations*. Taylor & Francis, Inc.

Garroppo, R., Giordano, S., & Vaccaro, A. (1998). "A teletraffic analysis of dial-up connections over pstn," *Proceeding of Globecom*, Sydney, Australia.

Goldberg, R. (1973). Architecture of virtual machines. In *Proceedings of the workshop on virtual computer systems*. ACM, New York, NY, USA, 74-112.

Goode, B. (2002). "Voice Over Internet Protocol (VOIP)," *Proceeding of the IEEE*, vol. 90, pp. 1495-1517.

Green, S. B., & Salkind, N. J. (2008). *Using SPSS for Windows and Macintosh, Analyzing and Understanding Data, 5th ed.* Upper Saddle River, NJ: Pearson Education, Inc.

Han, J. S., Ahn, S. J., & Chung, J. W. (2002). "Study of Delay Patterns of Weighted Voice Traffic of End-to-End Users on the VoIP Network," *International Journal of Network Management*, vol. 12, no. 5, pp. 271-280.

IDC (2008). IT Cloud Services Forecast – 2008, 2012: A Key Driver of New Growth. Retrieved March 11, 2010, from http://blogs.idc.com/ie/?p=224

Infonetics Research (2006). Enterprise VoIP adoption in North America will more than double in 2010. Retrieved April 25, 2010, from http://www.infonetics.com/pr/2006/upna06.ipv.nr.asp

ITU-T Recommendation G.114 (1996). "*One-way transmission time.*"

ITU-T Recommendation P.800 (1996). *"Methods for subjective determination of transmission quality*."

ITU-T Recommendation P.861 (1998). *"Objective quality measurement of telephone-band (300-3400 Hz) speech codecs."*

ITU-T Recommendation G.108 (1999). "*Application of the E-model: A planning guide.*"

ITU-T Recommendation G.107 (2000). "*The E-model, a computational model for use in transmission planning.*"

ITU-T Recommendation G.113 (2001). *"Transmission impairments due to speech processing."*

ITU-T Recommendation P.862 (2001). “*Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs.”*

ITU-T Recommendation Y.1530 (2007). “Call processing performance for voice service in hybrid IP networks.”

Jacobs, S. & Eleftheriadis, A. (1998). “Streaming video using dynamic rate shaping and TCP congestion control," *Journal of Visual Communication and Image Representation*, vol. 9, no. 3, pp. 211-222.

Jennings, R. (2009). *Cloud Computing with the Windows Azure Platform.* Indianapolis, IN: Wiley Publishing.

Keppel, G. (1991). *Design and analysis a researcher’s handbook*. New Jersey: Prentice Hall.

Khnaser, E. (2010). “Virtualization lets you be ready for any disaster,” *InformationWeek*, Issue 1,262.

Kondo, Y. & Matthews, H.S. (2004). "Comparative analysis of traditional telephone and voice-over-Internet protocol (VoIP) systems," International Symposium on Electronics and the Environment (ISEE'04), pp.106-111,

Kurose, J., & Ross, K. (2005). *Computer Networking: A Top-Down Approach Featuring the Internet*, 3rd edition. Addison Wesley.

Levin, I. P. (1999). *Relating statistics and experimental design*. Sage Publications: Thousand Oaks.

Linthicum, D. (2009). *Cloud Computing and SOA Convergence in Your Enterprise: A Step-by-Step Guide*. Boston: Pearson Education.

Mather,T., Kumaraswamy, S., & Latif, S. (2009). *Cloud Security and Privacy*. Sebastopol, CA: O'Reilly Media.

Mehta, P., & Udani, S. (2001). "Overview of Voice over IP," *IEEE Potentials*, Vol.: 20 Issue: 4.

Mell, P., & Grance, T. (2009). *"The NIST Definition of Cloud Computing,"* Version 15, 10-7-09.

Menon, A., Santos, J., Turner, Y., Janakiraman, G., & Zwaenepoel, W. (2005). "Diagnosing Performance Overheads in the Xen Virtual Machine Environment," Proceedings of the 1st ACM/USENIX International Conference on Virtual Execution Environments (pp. 13-23).

Miles, J. & Shevlin, M. (2001). *Applying regression and correlation: a guide for students and researchers*. Thousand Oaks, CA: Sage Publications.

Miller, M. (2009). *Cloud Computing: Web-Based Applications That Change the Way You Work and Collaborate Online.* Indianapolis, IN: Que Publishing.

Minium, E. W., Clarke, R. C. & Coladarci, T. (1999). *Elements of statistical reasoning* (2nd ed.). New York: John Wiley & Sons, Inc.

Mirzoev, T. (2007). A hardware-based model for reducing field loss by a video processing system. Ph.D. dissertation, Indiana State University, United States -- Indiana. Retrieved February 27, 2010, from Dissertations & Theses: Full Text.(Publication No. AAT 3259538).

Moreno, V. & Reddy, K. (2006). *Network Virtualization*. Cisco Press.

Morrissey,P. (2005) "How to measure call quality," *Network Computing*.

Ni, Q., Romdhani, L. & Turletti, T. (2004). "A Survey of QoS Enhancements for IEEE 802.11 Wireless LAN," *Journal of Wireless Communications & Mobile Computing*, vol. 4, pp. 547-566.

Norusis, M. J. (2005). *SPSS 13.0 guide to data analysis*. Upper Saddle River, NJ: Prentice Hall.

Norusis, M. J. (2006). *SPSS 15.0 Statistical Procedures Companion*. Upper Saddle River, NJ:Prentice Hall.

Osborne, J. & Waters, E. (2002). Four assumptions of multiple regression that researchers should always test. *Practical Assessment, Research & Evaluation*, 8(2).

Ou, G (2006). Introduction to Server Virtualization. Retrieved April 22, 2010, from http://articles.techrepublic.com.com/5100-10878_11-6074941.html

Park, S., Lee, Y., & Shin, H. (2004). An experimental analysis of the effect of the operating system on memory performance in embedded multimedia computing. In Proceedings of the 4th ACM International Conference on Embedded Software (pp. 26-33). Pisa, Italy: ACM Press.

Poe, R. (2006). “IP PBXs: Features? What about The Price?” *VoIP News*. Retrieved April 14, 2010, from http://www.voip-news.com/feature/pbx-features-roi-072106/

Rango, F., Tropea, M., Fazio, P. & Marano, S. (2006). “Overview on VoIP: Subjective and objective measurement methods,” *IJCSNS International Journal of Computer Science and Network Security*, vol. 6, no. 1B.

Reese, G. (2009). *Cloud Application Architectures: Building Applications and Infrastructure in the Cloud.* Sebastopol, CA: O’Reilly Media.

RFC 791(1981). INTERNET PROTOCOL. DARPA INTERNET PROGRAM. PROTOCOL SPECIFICATION.

Rittinghouse, J., & Ransome, J. (2010). *Cloud Computing: Implementation, Management, and Security.* Boca Raton, FL: CRC Press.

Rix, A.W.,Hollier, M. P., Hekstra, A. P., and Beerends, J. G. (2002). "Perceptual evaluation of speech quality (PESQ): The new ITU standard for end-to-end speech quality assessment part I - time-delay compensation," *Journal of the Audio Engineering Society*, vol. 50, no. 10, pp. 755–764.

Rix, A.W., Beerends, J.G., Kim, D., Kroon, P., & Ghitza, O. (2006). "Objective Assessment of Speech and Audio Quality - Technology and Applications." *IEEE Transactions on Audio, Speech & Language Processing*, vol. 14, no. 6, pp. 1890-1901.

Rosenblum, M. (2004). The reincarnation of virtual machine. *Queue, Vol. 2, Issue 5. P* 36-40.

Sakellari, G. (2010) "The Cognitive Packet Network: A Survey," *The Computer Journal*, vol. 53, no. 3, pp. 268-279.

Shuf, Y. (2003). Improving the memory performance of Java workloads. Ph.D. dissertation, Princeton University, United States -- New Jersey. Retrieved January 27, 2010, from Dissertations & Theses: Full Text. (Publication No. AAT 3102223).

Sreenan, C., Chen, J., Agrawal, P., & Narendran, B. (2000). "Delay Reduction Techniques for Playout Buffering", *IEEE Transactions on Multimedia*, 2(2):88-100.

Skype, 2011. About Skype. Retrieved February 23, 2010, from http://about.skype.com/

Thorpe, L. (2002). "Subjective evaluation of speech compression codes and other non-linear voice-path devices for telephony applications," *International Journal of Speech Technology*, vol. 2, pp. 273–288.

Wang, R. & Hu, X. (2004). "VoIP Development in China", Computer, vol.37, issue 9, pp. 30 – 37.