2018

Strategies to Manage Cloud Computing Operational Costs

Frankie Nii A Sackey

Walden University

Walden University

College of Management and Technology

This is to certify that the doctoral study by

Frankie Nii A. Sackey

has been found to be complete and satisfactory in all respects,
and that any and all revisions required by
the review committee have been made.

Review Committee
Dr. Irene Williams, Committee Chairperson, Doctor of Business Administration Faculty

Dr. Tim Truitt, Committee Member, Doctor of Business Administration Faculty

Dr. Judith Blando, University Reviewer, Doctor of Business Administration Faculty

Chief Academic Officer
Eric Riedel, Ph.D.

Walden University
2018

Abstract

Strategies to Manage Cloud Computing Operational Costs

by

Frankie Nii A. Sackey

MSc, University of Texas at Dallas, 2000

MBA, Texas A&M University—Commerce, 1991

Masters of Infrastructure Planning, University of Stuttgart, 1989

BS, Kwame Nkrumah University of Science and Technology, 1984

Doctoral Study Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Business Administration

Walden University

October 2018

Abstract

Information technology (IT) managers worldwide have adopted cloud computing because of its potential to improve reliability, scalability, security, business agility, and cost savings; however, the rapid adoption of cloud computing has created challenges for IT managers, who have reported an estimated 30% wastage of cloud resources. The purpose of this single case study was to explore successful strategies and processes for managing infrastructure operations costs in cloud computing. The sociotechnical systems (STS) approach was the conceptual framework for the study. Semistructured interviews were conducted with 6 IT managers directly involved in cloud cost management. The data were analyzed using a qualitative data-analysis software to identify initial categories and emerging themes, which were refined in alignment with the STS framework. The key themes from the analysis indicated that successful cloud cost management began with assessing the current environment and architecting applications and systems to fit cloud services, using tools for monitoring and reporting, and actively managing costs in alignment with medium- and long-term goals. Findings also indicated that social considerations such as fostering collaboration among all stakeholders, employee training, and skills development were critical for success. The implications for positive social change that derive from effectively managing operational costs include improved financial posture, job stability, and environmental sustainability.

Strategies to Manage Cloud Computing Operational Costs

By

Frankie Nii A. Sackey

MSc, University of Texas at Dallas, 2000

MBA, Texas A&M University-Commerce, 1991

Masters of Infrastructure Planning, University of Stuttgart, 1989

BS, Kwame Nkrumah University of Science, and Technology, 1984

Doctoral Study Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Business Administration

Walden University

October 2018

Dedication

I would like to dedicate this study to God, for His grace, everlasting love, and the strength that He provided me to complete this study. That strength to persevere and focus on accomplishing my goals only comes from His grace. I also want to dedicate this to my dad, who taught me that education is the only thing no one can take away from me, so I must never stop educating myself; and my mom, whose love I can still feel in my life. I know both of you would be proud of me. I also want to dedicate this to my sister Felicia, who has been everything for me, always there for me. Always. For the rest of my siblings, thanks for your love and support. However strong we are as individuals, we are stronger when we stand together. I am always strengthened to know that any one of you is a call away. Last but not least, I want to dedicate this to my sons, Ashton and Kasey, for their love and understanding, especially at those times when I needed to get away to advance my study, and to Taliyah and Trinity—if Papa can do it, so can you.

Acknowledgments

I wish to use this opportunity to thank my committee members at Walden University—my second chair, Dr. Tim Truitt, and Dr. Judy Blando, URR, for their helpful suggestions and support. I also want to especially thank my chair, Dr. Irene Williams, who has created a supportive student environment in her classes, and whose direct support, encouragement, timely feedback, and guidance through the process have been vital to my successful completion of this study. I would also like to thank my lifelong friend Dr. Fred Ohene, who convinced me the effort would be worth the returns. I also want to thank my business partner Kapil Khanna; my intellectual discussions with him on this and other subjects have broadened my perspectives on both life and technology and fueled my desire for more inquiry. Finally, I also want to acknowledge the support of my sister Tina, who was kind enough to review this work and to offer useful advice on how to make it better. Thanks to all of you for your support over the last couple of years, and for enabling me to capitalize on this wonderful learning opportunity. I will be forever grateful.

Table of Contents

List of Tables

List of Figures

Section 1: Foundation of the Study

Until 2006, most information technology (IT) managers acquired and consumed computing resources through a process of purchasing, installing, configuring, operating, and managing computing resources in traditional on-premise or colocation facilities. Since 2010, the availability of cloud computing has made it possible for any IT manager, irrespective of the size of the organization, to acquire and consume such computing resources as a service. Cloud computing's promises of ease of use, convenience, on-demand access, provisioning agility, and scalability present an attractive alternative to information technologists, with potential as a platform for addressing the rapid growth and use of data by organizations (Hashem et al., 2015). This technological promise, coupled with the potential for cost savings, have fueled global spending on cloud-based products and services at an exponential rate. According to International Data Corporation (IDC) market research, estimated cloud-service spending expanded from about $16 billion in 2008 to $122.5 billion by the end of 2017, with this amount expected to rise to about $203 billion by 2020 (IDC, 2017a).

The proliferation of mobile devices and the ongoing need for data analytics are also driving organizations to adapt the ways in which they collect, store, process, manage, and consume information (Dinh, Lee, Niyato, & Wang, 2013). Cloud computing is one approach to meeting the rapidly growing requirements for such data systems because service providers have structured models for rapid growth. With cloud computing, this growth can be achieved without the addition of new infrastructure through purchasing, installation, and configuration of systems needed to support the growth of such data within the enterprise (Hashem et al., 2015). The ease of acquisition and use of cloud-based technologies also account in part for the assertion by Abdelmaboud, Jawawi, Ghani, and Elsafi (2015) that cloud computing is the future of computing. Abdelmaboud et al. identified cost as a key motivator in the decision to adopt cloud computing by IT managers. However, Tak, Urgaonkar, and Sivasubramaniam (2013) challenged the notion that applications migrated to cloud computing can easily claim cost advantages, questioning how any such advantage leads to tangible cost savings. If cost is a primary motivator for cloud adoption, it is unclear that organizations that transform to cloud computing realize their potential for cost savings.

I explored the strategies used by IT managers to manage their IT infrastructure operating costs following transformation to cloud computing environments. The goal is to help organizations using cloud computing to understand how best to capitalize on the cost-saving potential of cloud computing. In this section, I provide a brief background, a concise definition of the business problem, and the purpose of the study. I follow this with a review of the nature of the study and a presentation of the conceptual framework and significance of the study. I complete the section with summary findings from a review of the academic literature. In subsequent sections of this document, I detail the study’s methods, design, and data collection and analysis. The document concludes with a final section in which I discuss my findings, the applicability of the findings to professional practice, and areas for further research.

Background of the Problem

IT leaders face constant challenges as they seek new and innovative ways to compete in a global and dynamic marketplace. Two approaches to meeting this challenge and sustaining organizational financial performance and value are cost leadership and differentiation (Banker, Mashruwala, & Tripathy, 2014; Ibrahim, 2015). Since 2010, IT costs have been steadily growing as organizations have sought to sustain capabilities in innovation, product development, production, delivery, marketing, and other supporting processes. Business productivity and cost reduction have been the two top information technology and management issues for IT leaders in the last several years (Luftman et al., 2015). IT budget accounts for about 3.5% to 6.5% of revenue, of which about 44% is spent on IT infrastructure and related activities (Luftman et al., 2015). IT costs are an issue that IT managers will need to continue to manage to be successful.

Cloud computing has been an influential technological advancement since 2010. Garrison, Wakefield, and Kim (2015) stated that cloud computing has the unlimited and untapped potential for driving business innovation, with the capability to fundamentally shift competitive landscapes by providing a new platform for creating and delivering business value. These attributes of cloud computing indicate potential cost-savings opportunities for organizations adopting one or more of the service models, but a preliminary review of the literature on posttransformation cost savings indicated that not to always be the case (Tak et al., 2013). My review of the academic literature indicated a lack of a clear and consistent strategy and processes for IT leaders to use to capitalize on potential cost-savings opportunities in cloud computing. The goal of this study was to explore the strategies and operational processes used by organizations that have succeeded in managing infrastructure operating costs in cloud computing so that those practices can form the basis to inform other IT managers adopting cloud computing to accomplish the same objectives.

Problem Statement

IT cost reduction has been a major management concern for organizational leaders in the last few years (Luftman et al., 2015). In 2014, businesses spent about 40% of their IT budgets on IT infrastructure alone (Luftman et al., 2015). Cloud computing is one progressive strategy to reduce IT infrastructure operational costs (Oliveira, Thomas, & Espadanal, 2014), but organizations transforming to cloud computing are not consistently realizing the promise of cost savings in cloud environments (Rana, 2014). The general business problem was that IT infrastructure managers do not have the operational strategies and processes they need to manage their operational costs in cloud computing. The specific business problem was that some IT managers lack knowledge of strategies and processes for managing IT infrastructure operational costs in cloud computing environments.

Purpose Statement

The purpose of this qualitative multiple case study was to explore strategies and processes that IT managers use for managing infrastructure operations costs in cloud computing environments. The target population was IT operations managers from an IT organization in the Tampa, Florida area with a record of successfully managing IT infrastructure operational costs in a cloud computing environment. IT managers may benefit from the findings of this study because the strategies addressed may help them understand the pathways for successfully managing their infrastructure costs through cloud computing. By managing and reducing costs, organizations can focus on their core business, improve innovation, and sustain competitive advantage (Oliveira et al., 2014). The economic viability of these organizations can lead to more job opportunities within the local communities where they operate and lead to positive social change. Cloud computing service providers use energy-efficient and environmentally friendly technologies for their facilities (Liu et al., 2015). Thus, operational cost savings in cloud computing, which can have the long-term effect of increasing cloud adoption, also have the potential to reduce an organization's environmental impact and carbon footprint through reduced energy use.

Nature of the Study

I considered three research methods for the study: (a) qualitative, (b) quantitative, and (c) mixed method. Based on inductive reasoning, qualitative methods are applicable when the goal of the study is to explore a phenomenon (e.g., a single or collective case) with the aim of gaining an objective understanding of the meaning that participants assign to their observations and experiences of that phenomenon (Pietkiewicz & Smith, 2014). In contrast, a quantitative method is a deductive approach in which researchers use statistical methods to quantify and examine relationships or differences among variables (Yilmaz, 2013). A mixed method approach is a hybrid of the other two methods and is used by researchers when there is a need in the same study to both explore a phenomenon and examine quantitative relationships or differences among variables (Turner, Cardinal, & Burton, 2017). The objective of this study was to explore strategies and processes used by organizations for managing cost savings in cloud computing, so quantitative and mixed method methods were not applicable, and I selected a qualitative method for the study.

Of the several design options available to researchers using the qualitative research method, I considered four for use in this study—narrative, phenomenological, ethnography, and case study. Researchers use narrative research design when the objective of the study is to provide a detailed account of an event or outline a series of chronological actions pertaining to the person or persons who are under study (Percy, Kostere, & Kostere, 2015). Researchers use the phenomenological design to explore and understand the lived experiences of participants, and when they seek to uncover the essence of a phenomenon through a study of experiences relative to that phenomenon (Percy et al., 2015). Ethnography is applicable when the researcher seeks to explore and identify behaviors, beliefs, language, and interactions of a group's members sharing a common culture (Percy et al., 2015). Case study research is applicable when the object of the study is to explore a real-life situation in a specific setting (which also bounds the study), as a means to understand what is distinctive about that situation and setting (Hyett, Kenny, & Dickson-Swift, 2014). I selected the case study design as the one most applicable to explore the strategies and processes that the participating organizations used to manage their IT operational costs in cloud computing.

Research Question

What strategies and processes do IT managers use for managing IT infrastructure operational costs in cloud computing environments?

Interview Questions

I used the following interview questions to answer the overarching research question on strategies and processes that IT managers use for managing IT infrastructure operational costs in cloud computing.

1. What strategies have you used for managing IT infrastructure operational costs in cloud computing?
2. What processes worked the best for developing and implementing IT infrastructure operational costs in cloud computing?
3. What tools did you find to be effective to use for managing IT operational costs?
4. What types of staff training, if any, did you find to be effective for managing IT operational costs in cloud computing?
5. What organizational changes, if any, did you find to be effective for managing IT infrastructure operational costs in cloud computing?
6. What ongoing adjustments to IT operational procedures have you found to be necessary for managing costs on a month-to-month or an ongoing basis?
7. What else would you like to add that we have not yet addressed?

Conceptual Framework

I selected the sociotechnical systems (STS) approach as the conceptual framework for this study. Carayon et al. (2015) attributed the development of STS theory to a group of researchers and consultants associated with the London Tavistock Clinic around 1970, attributing further enhancements to the design of information systems to Trist and Bamforth. Based on the general systems theory first postulated by Von Bertalanffy in 1972, the key principle of STS theory is that the successful attainment of an organizational goal entails a linear and nonlinear interaction of social and technical factors. Optimization of only one factor and not both may adversely impact the attainment of the goal (Carayon et al., 2015). A key tenet of STS theory is the decomposition of complex systems into two primary subsystems: (a) the technology subsystem—equipment, machines, tools, technical environment, and related processes; and (b) the social subsystems—individuals, teams, management and governance structures (Carayon et al., 2015; Cherns, 1987). The goal of the theory is the joint optimization of both subsystems to improve organizational effectiveness and the life of people in the workplace.

Successful IT operations management entails the management of a complex set of interrelated activities encompassing the use of different types of IT resources with different capacity and performance objectives, as well as the IT staff and management responsible for the use of the systems to support business processes. Any strategy to comprehensively manage IT operational costs must address both technical and social aspects. The IT manager is responsible for balancing the joint management of these subsystems against the ever-changing needs of the business, as well as new sets of underlying contracts, operating rules, and protocols from service providers and their vendors. I expected STS theory to be an appropriate lens for exploring strategies and processes for managing IT infrastructure operational costs in cloud computing.

Operational Definitions

This section, I provide brief definitions of common terminology used in IT and cloud computing literature and throughout this document. Several definitions of *cloud computing* exist throughout the academic literature (Sultan, 2014), but the National Institute of Standards and Technology (NIST) definition is the one commonly referenced by most researchers. The NIST defined cloud computing as on-demand access to shared computing resources such as networks, servers, storage, database systems, and applications that subscribers can rapidly provision with minimal management effort or service provider interaction (Mell & Grance, 2014).

Cloud (services) provider: A cloud provider is a person or an organization responsible for making a cloud service available to subscribers. A cloud provider acquires and manages the computing infrastructure required for providing the services, runs the cloud software that provides the services, and delivers the cloud services to the cloud consumers through network access (Chou, 2015).

Cloud native: Cloud native is a term used to reference applications that are designed, developed, and deployed to capitalize on the nature and benefits of cloud services to provide application managers with features such as global availability, easy scalability, and failure tolerance (Gannon, Barga, & Sundaresan, 2017). A cloud-native organization, on the other hand, is an organization whose IT infrastructure is based entirely on cloud-native applications.

Compute clusters: Clustering refers to the use of multiple computing resources acting together to form a single highly available or high-capacity system (Manvi & Shyam, 2014). The cluster may consist of multiple compute nodes, storage systems, and/or redundant networking systems. A *compute resource* refers to a cluster with compute nodes to support high-capacity, high-availability computing needs.

Community cloud: A community cloud serves a group of cloud consumers with shared concerns such as mission objectives, security, privacy, and compliance policy, rather than serving a single organization as does a private cloud. A community cloud is managed by a third party or a group of organizations that. participate in that community of users. The community cloud infrastructure is implemented on-premise or outsourced to a hosting company (Goyal, 2014).

Elasticity: Elasticity is a key advertised benefit of cloud computing and is defined as the capability of a given system or service in the cloud environment to add or remove resources as needed, to match system load requirements (Coutinho, de Carvalho Sousa, Rego, Gomes, & de Souza, 2015).

Hybrid cloud: A hybrid cloud is a composition of two or more clouds (onsite private, onsite community, offsite private, offsite community, or public) that remain as distinct entities but are bound together by standardized or proprietary technology that enables data and application portability (Goyal, 2014).

Infrastructure as a service (IaaS): IaaS is the most basic of services offered by cloud service providers and consists of the provision of a server or servers or storage capacity in the cloud (Manvi & Shyam, 2014). Subscribers of IaaS have the need for computing, power, database, and storage systems but have or want no role in the installation and maintenance of the hardware and other supporting infrastructure.

Platform as service (PaaS): PaaS is a cloud-based platform that service providers use to offer software-development tools and application deployment and execution environment for on-demand use (Ananthakrishnan, Chard, Foster, & Tuecke, 2015). Organizations use the platform to quickly develop their custom applications or write software that integrates with existing applications.

Private cloud: A private cloud gives a single cloud consumer's organization exclusive access to and use of the computational resources in the environment. Private clouds are managed either by the cloud consumer organization or by a third party and with the hosting on the organization's premises or outsourced to a hosting company (Goyal, 2014).

Public cloud: A public cloud is one in which the cloud infrastructure and computing resources are made available to the public over a public network. A cloud service provider typically owns the public cloud and uses it to serve a diverse pool of clients (Goyal, 2014).

Software as a service (SaaS): SaaS is an application or suite of applications that reside in the cloud. The service provider delivers this on demand, on a single- or multitenant basis (Seethamraju, 2015). Vendors of applications such as Salesforce.com, Microsoft 365, and Google Suite offer them as SaaS solutions.

Utility computing: Utility computing refers to the procurement of computing services from a vendor on a pay-per-use basis (Rani, Rani, & Babu, 2015). The concept of utility computing is similar to that of traditional utility services, with a subscriber getting bills that commensurate with the number of services used over the billing periods.

Assumptions, Limitations, and Delimitations

Assumptions

An assumption, as used in the context of this study, is a supposition that an underlying statement is a fact, even though there is no proof of such fact in general or provided in this study (Yin, 2017). Several factors can affect an organization's ability to attain a specific outcome from a series of actions or the adoption of technology. Though some of these factors are quantifiable and measurable, others, such as organizational culture, competency, the experience level of staff, and organizational capabilities, are not easily quantifiable or measurable but can also have a considerable effect on the success of technology adoption. I did not assess the role of these intangible factors in this study but assumed that they had the same effect on outcomes for the case in this study; therefore, I assumed that the role of these intangible organizational assets in facilitating cost reductions for the organizations was equal for all organizations.

Any qualitative research relies on the collection of accurate data from participants selected for the study. In this study, I assumed that participants were truthful and accurate in their responses to questions, that participants responded without bias or prejudice, and that any supporting information provided by the IT managers as part of data collection were reflective of actual records from operational activities and not selected to exaggerate, represent, or favor any particular point of view or position.

Limitations

Limitations of a study refer to an area of potential weakness because of a shortcoming in design, methods, data collection, or analysis of the data in the study (Yin, 2017). For this study, I identified an appropriate organization and selected qualified individuals within the organization to interview to collect data. One limitation of this study was the selection of the Tampa, Florida area as the geographical scope of the study. The use of a single case and colocation of the participants in such a geographical setting may limit the extent to which a researcher can draw causal conclusions from a study. However, specifically in the case of information system research, Tsang (2014) contended that researchers can still infer a limited set of generalizations from small sample sets in a specific geographical location. Even with that, additional research may be necessary to establish that the findings from this study would generalize to other organizations.

I selected a qualitative case study design for this study. Yin (2017) noted a lack of any real mechanism to address complexity and contextual conditions in such a study design because of the small number of participants. From the perspective of critical realism, researchers can still use case study for identification of specific characteristics in a population and leverage what they learn from those characteristics as a basis for subsequent research or theory development (Gibson, 2017). Even with these known limitations, the qualitative case study was still a valid design for this study.

Delimitations

Delimitations, as used in this study, refer to the bounds or scope specified for a study (Ganapathy, 2016). The major cloud computing services are Amazon Web Services (AWS), Microsoft Azure, and Google Cloud. The providers offer services broadly categorized as software as a service (SaaS), platform as a service (PaaS), and infrastructure as a service (IaaS). I limited this study about strategies used by organizations to manage infrastructure costs in cloud computing to organizations using the IaaS service offering from the current leading cloud service provider (AWS).

A further limitation to my selection of IaaS service on AWS was the choice of a single geographical location (the Tampa area) for selecting the case for the study. Within this area, I selected an organization with IT infrastructure on AWS, with an established record of realizing operational cost savings. These boundaries or scope of the study limited the extent to which study findings could potentially be generalized for professional practice.

Significance of the Study

IT cost reduction is a top management concern for IT leaders, with IT infrastructure costs alone accounting for about 40% of total IT budgets (Luftman et al., 2015). The findings from this qualitative study could be of value to organizations and professional practice in understanding what strategies and processes successful organizations are using to manage IT infrastructure operational costs through cloud computing. The study's contribution to professional practice also extends to cloud service providers, who could use or adopt the findings to help refine and enhance their tools and processes and align services to better support their clients as they transition to and consume cloud services. Success in realizing promised cost savings can be a powerful incentive for other organizations to adopt cloud computing.

The study's implications for positive social change include the potential to affect organizations' profitability, innovation, and competitiveness through the realization of IT operations cost savings. The combined positive effect of increased innovation and a sustained competitive posture has the potential to increase employment and economic opportunities in the communities where organizations operate. Transforming and consolidating IT infrastructure activities in private/public cloud environments also has the potential to reduce energy use in organizations. Balasooriya, Wibowo, and Wells (2016) stated that cloud service providers are engaging in initiatives considered adaptive, energy efficient, and environmentally responsible in the way that they use resources to reduce waste. Cost savings that lead to more organizations adopting cloud computing have the potential to contribute indirectly to positive social change and environmental sustainability.

A Review of the Professional and Academic Literature

The literature review in an important part of any doctoral study. Webster and Watson (2002) stated that an effective literature review creates a solid foundation for advancing knowledge, facilitates theory development, bridges gaps in research, and exposes areas where additional research may be necessary. My goal in this section is to establish the progression of research work leading to the inquiry in this study and identify the gaps in the literature that I attempted to bridge upon completion of the study. After outlining the approach and the search strategy for the section, I provide a brief introduction to cloud computing, including its history, characteristics, and service and deployment models. I conclude by establishing the need for resource management in cloud computing.

The goal of this qualitative multiple case study was to explore strategies and processes that IT managers use for managing infrastructure operations costs in cloud computing environments. After the introductory section, I organize the rest of the review under five topics. I use the first topic area to discuss the conceptual framework I selected for the study, identifying elements of the framework and how the elements guided this study. In the second section, I discuss the key drivers for cloud computing adoption and the classic adoption approaches. I use sections on subsequent topics to discuss economic implications, pricing, and cost factors for cloud systems and follow that with a discussion of cost management and its applicability to IT operational activities. The final topic involves IT operations in cloud environments and ongoing challenges with cost management of IT infrastructure. In each of these sections, I present a comprehensive analysis of the existing published literature in the area, providing both syntheses and critical analysis where necessary. The sources of information for most of the content for the literature review came from searches of electronic research databases. Cloud computing as a technology has only matured within the last decade, and its related services and systems are still undergoing rapid changes. With such a history and current state, the traditional sources of research information in peer-reviewed magazines lacked quality articles about cloud computing, and the information in printed books was dated compared to the current state of the technologies. For this reason, even though the review process included some books, I only cite a few that I deemed necessary for the insight they provided.

For the academic articles, the primary data source was EBSCOhost's Computers and Applied Sciences Complete and Business Source Complete/Premier, ProQuest's ABI/INFORM Complete and Business Dissertations and Theses, and Emerald Management Journals. The keywords that I used for the electronic database search for the study included (a) *cloud computing*, (b) *cloud computing adoption*, (c) *cloud computing resource utilization*, (d) *cloud computing resource optimization*, and (f) *infrastructure as a service costs*, (e) *cost management*, and (f) *sociotechnical theory*. As part of the review process, I examined and qualified over 252 resources, seven of which were government publications and nine of which were book references. The rest of the sources were journal articles, about 90% of which were from peer-reviewed periodicals.

I often limited search results to articles published no earlier than 2014, even though I included a few articles published before 2014 because of their direct relevance, or in some instances because of lack of updated studies in specific subject areas. Of this list of articles, this study includes 170 references. In accordance with the requirements of Walden University, 85% of the references were peer-reviewed, and 85.3% of the references were within 5 years of my expected graduation date in the calendar year 2018.

I used a combination of Mendeley and Qippa and Microsoft Excel as tools to manage articles and references of the articles downloaded electronically from the libraries.

Cloud Computing Overview

In this worldwide era of innovative IT, the availability of substantial computing power has been an enabler for the development and production of goods and services. The use of computing resources in support of business processes is widespread and, in some situations, is an essential factor in competitive advantage. The traditional way by which enterprises acquire computing capabilities involves procuring and managing the needed computing resources in their in-house data centers (Phaphoom, Wang, Samuel, Helmer, & Abrahamsson, 2015). In the last several years, there has been a significant shift from this approach to the use of computing resources owned and managed by a service provider. Several definitions of *cloud computing* exist in the literature, but the most popular is the one given by the NIST. NIST defined cloud computing as a type of computing where extensive scalable IT-related capabilities are available as a resource across the Internet to multiple external customers (Mell & Grance, 2014). Other authors have defined cloud computing as a set of shared resources presented as a standardized computing utilityrunning on a server, which caters to multiple groups of users through multitenancy (Botta, De Donato, Persico, & Pescapé, 2016; Hashem et al., 2015). The same themes occur across definitions—the acquisition and use of easily accessible, shared, and scalable computing resources in a multitenanted environment.

IT managers can procure and use cloud computing services without directly owning the infrastructure that delivers them. In this respect, cloud computing service providers offer an alternative and efficient Internet-based computing service in scalable ranges that customers can access, preferably through high-bandwidth and low-latency network connections. The set of core services available from cloud service providers includes SaaS, PaaS, and IaaS. SaaS capabilities first became broadly available in late 2006, when key technology companies including Google, Amazon, Yahoo, and Microsoft started launching these services in the commercial marketplace. The advantages of on-demand computing power, with a fast and efficient application and chargeback capability, low-level maintenance, and fewer IT labor costs, made the services attractive from a cost perspective (Reddy, Mudali, & Roy, 2017). The result of these operational benefits has been the ongoing growth of enterprise spending on cloud computing.

According to IDC market research, estimated cloud-service spending grew from approximately $16 billion in 2008 to about $122.5 billion in 2017 (IDC, 2017a). This projected growth in spending reflects the attractiveness of a computing model that promises deployment and operational agility at a lower cost point than most organizations can achieve from owning their infrastructure and related applications. Cloud service providers have evolved their services around a set of core services to align with the core needs of most IT managers, leading some authors to refer to cloud computing as the *fifth utility*, on the same level as other traditional utility services such as electricity, gas, and telephone (Koehler, 2014; Ruan, Chan, Zhu, Wang, & Yang, 2016; Whaiduzzaman, Haque, Rejaul, & Gani, 2014). The use of the term *utility computing* for cloud computing was initially attributed to John McCarthy, though others (Alkhalil, Sahandi, & John, 2017) associated the use of the term with Leonard Kleinrock, a principal scientist at the Advanced Research Projects Agency Network (ARPANET). The prediction that cloud computing would become a common utility has come to fruition.

Evolution of cloud computing. As with other types of technological advancements, scientists have evolved cloud computing as a technology since 2006. They started it initially as a form of grid and utility computing through the provision of application services by application service providers (ASPs) and SaaS (Fang, Liu, Romdhani, & Pahl, 2015; Padilla, Milton, & Johnson, 2015). The concept of delivering computing resources through a global network is rooted in the 1960s (Bayramusta & Nasir, 2016). In 1969, Joseph Carl Robnett Licklider, the first director of the Information Processing Techniques Office (IPTO) at The Pentagon Advanced Research Projects Agency (ARPA), developed the idea of an *intergalactic computer network*. His innovation consequently enabled the establishment of ARPANET (Kaur & Kaur, 2014), whose initial purpose was to link computers at Pentagon-funded research institutions over telephone lines. However, the vision of organizing computing resources as a utility grid first came to fruition in the 1990s, as an effort to solve significant challenges in scientific computing.

Grid computing in practice involved interconnecting high-performance computing facilities across universities in regional, national, and pan-continent grids (Kaur & Rai, 2014; Sadiku, Musa, & Momoh, 2014). The scientists used these interconnected grids principally for transferring large amounts of data between nodes, executing computational tasks across administrative domains, and allocating resources shared across projects fairly (Sandholm & Lee, 2014). Grid computing evolved to utility computing (Koehler, 2014), which refers to efforts in the industry around the turn of the millennium to improve manageability and on-demand provisioning of these nodes and their compute clusters.

The service providers initially marketed utility computing on a cluster-by-cluster basis, and often situated on an organization's network. This distribution model made it very expensive to implement technological advancements to the platforms. The allocation of computational resources on these grid systems revolved around project membership enrollment, and scientists invested much effort in sophisticated security policy configuration and validation. The gap for IT managers was the need for a system that offered subscribers scalability, security, high operational capacity, and performance, with pay-as-you-use options (Padilla et al., 2015). These factors were the core motivations for organizations to subscribe to these systems instead of building their high-end computing infrastructures, sparing significant capital investments and time by migrating their workloads to these environments (Saha et al., 2016; Sandholm & Lee, 2014). These motivations remain true as of 2018, even though other factors such as agility and innovation have also become important reasons for organizations to choose cloud computing.

A contributing factor in the rapid rate of growth for cloud computing involves the relative rate of growth of data volumes and the array of Internet-connected devices and related applications. The rate of data accumulation is compelling organizations with data centers (DCs) to host large scale data-driven services and networked applications such as those with a heavy emphasis on media sharing and streaming (Hashem et al., 2015; Singh

& Rajdavinder, 2017). These types of applications and services need an efficient and low-cost platform, but their use can impose numerous types of high resource demands such as computer power, bandwidth, storage, and latency on the basic infrastructure.

The traditional DC architecture has limitations in its adaptability to support these data- or processor-intensive applications effectively, resulting in poor quality of service (QoS) support, deployment, management, and protection against security attacks (Boru, Kliazovich, Granelli, Bouvry, & Zomaya, 2015; Bruneo, 2014). Cloud computing, with its promises of scaling and elasticity, has become a preferred deployment model for such applications. By moving such applications and their related data services to cloud environments, organizations could avoid initial investment in expensive infrastructure setup and subsequent monthly maintenance costs. In cloud environments, organizations also can grow and shrink system capacity as rapidly as needed to meet any fluctuating business demands, without the need for or effort to directly acquire and set up the infrastructure required to do so.

In the early stages of the evolution of cloud computing, organizations struggled to understand the ever-evolving services offered by the cloud service providers such as Amazon Web Services, Microsoft Azure, and IBM (El Zant & Gagnaire, 2015; Hashem et al., 2015; Opara-Martins, Sahandi, & Tian, 2016; Padilla et al., 2015). Since 2010, these service providers have hastened the pace of service availability and distribution, but at the same time making it easier for subscribers to provision and use services. This ability to rapidly extend service capabilities to subscribers has been possible through the use of PaaS and CaaS offerings (Hintemann, 2015; Paschke, 2016; Peinl, Holzschuher, &

Pfitzer, 2016). The capability to offer subscribers a set of fully automated workflows, governed by cost, performance, and compliance specifications, has made adoption of cloud services easier for subscribers.

From the organizational perspective, the rapid growth of mobile devices and the increased use of computing resources worldwide have led to the growth of large data systems and the need for a different approach to resource provisioning. The initial architectural approach to addressing this problem was through DC consolidations (Sharkh, Shami, & Ouda, 2017). Large DCs are more cost-effective high quality of services (QoS) support because of consolidation of computing and storage clusters and are often on more reliable and efficient operating networks (Sandholm & Lee, 2014; Zhang, Yang, Shi, Wu, & Ran, 2015). Moreover, service providers often construct them in areas of cheap energy and labor costs (Saha et al., 2016; Zhang et al., 2015). This combination of favorable location and economics of scale help service providers to reduce their operational expenditures which they then pass on to subscribers as affordable pricing.

Any discussion of the growth of cloud computing is not complete without mention of virtualization technology. Malhotra, Agarwal, and Jaiswal (2014) described virtualization technology as the single most important element underlying cloud computing technology. Virtualization techniques such as Xen and VMW are how cloud providers run multiple virtual machines (VMs) on one physical infrastructure device (Mosa & Paton, 2016). Virtualization technologies enable users to access terabytes or even exabytes of storage and high processing power at high availability rates, and in a lower-cost, pay-as-you-use model (González-Martínez, Bote-Lorenzo, Gómez-Sánchez, & Cano-Parra, 2015). Virtualization technology also enables the storage of operating system and application code as images for execution on demand or replicated as other application instances, the benefit of which is the enablement of rapid provisioning, scalability, and fault tolerance. Virtualization enables cloud service providers to virtualize resources in an on-demand fashion and offer a high-processing-power environment, which is the most suitable and manageable framework for big data processing and management (Hummaida, Paton, & Sakellariou, 2016; Mosch, Groß, & Schill, 2014). Malhotra et al. (2014) pointed out that in addition to virtualization, other technologies such as multitenancy and shared infrastructure are also key to the success of cloud computing, but both of these technologies are possible only through virtualization.

Sociotechnical Systems Theory

I selected the STS approach as the conceptual framework for this study. The development of STS theory has been attributed to a group of researchers and consultants associated with the London Tavistock Clinic following the Second World War and was further enhanced to apply to the design of information systems by Trist and Bamforth (Carayon et al., 2015). STS theory states that the successful attainment of an organizational goal in STS entails a linear and nonlinear interaction of social and technical factors (Trist, 1981). The optimization of only one factor and not both may adversely influence the attainment of the goal (Carayon et al., 2015). A key tenet of STS theory is the decomposition of complex systems into two primary sub-systems: (a) the technology subsystem—equipment, machines, tools, technical environment, and related processes; and (b) the social subsystems—individuals, teams, management, and governance structures (Carayon et al., 2015; Cherns, 1987). The goal of the theory is the joint optimization of both subsystems to improve organizational effectiveness and the life of people in the workplace.

STS background and history. In the period proceeding post world war II, coal was the main source of power required to fuel the industrial reconstruction. The coal industry was underperforming through a combination of low morale, labor disputes, and high employee turnovers. The National Coal Board (UK) commissioned the Tavistock Institute to study the situation to determine what alternative approaches may help to improve the productivity of the mines (Trist, 1981). Until this period immediately post world war two, the predominant organizational model was a fusion of Weber's description of Bureaucracy with Taylor's concept of scientific management. The concept of STS theory evolved from these studies of the Tavistock Institute. The goal at conception was improving productivity through the diffusion of innovative work practices to enable the derivation of increased output from the application of scarce resources (capital expenditure) (Trist, 1981). During this post world war two period, coal was the main source of power required to fuel the industrial reconstruction, but the industry was underperforming through a combination of low morale, labor disputes, and high employee turnovers. These considerations drove the research that emerged from the Tavistock program and which led to the theoretical development of the core STS concepts. In addition to the development of the STS theory, the team also initiated other analytical field studies to develop methods to obtain the best match between the social and technical systems, how to improve the match, and finally, how to diffuse the improvements to the organizations.

The sociotechnical systems approach. Sociotechnical system approach has its roots in open systems theory developed by Von Bertalanffy in the 1930s and used by researchers in the study of complex systems. The premise of the open systems theory is that all open systems have similar underlying principles (Von Bertalanffy, 1972). This general systems theory views the organization as an entity made up of components or parts or subsystem. This consideration of the whole as a composition of parts is the same underlying theme for STS theory.

STS decomposes the organizational system into two primary subsystems (Carayon et al., 2015). These two primary subsystems are – the technology subsystem and the social sub-systems (Carayon et al., 2015; Cherns, 1987). The original formulation of the theory was to obtain a good fit between these two subsystems, but Trist and team later reformulated this as the joint optimization of the social and technical systems (Trist, 1981). Even though the two subsystems are independent of each other, they have a correlative relationship, with one depending on the other for the transformation of inputs to the outputs that define the work system (Trist, 1981). Because of this complex relationship between the two subsystems, the Tavistock research team postulated that the two subsystems must be jointly optimized and that attempts to optimize one alone and not the other will result in the suboptimization of the socio-technical whole (Trist, 1981).

Social subsystem. The role of human factors in sociotechnical systems is significant to the attainment of organizational goals. Keating, Fernandez, Jacobs, and Kauffmann (2001) defined the social subsystems as including attitudes and beliefs, contracts between employers and employees, reactions to work arrangements, and the relationships between individuals and among groups. They further stated that these social factors emerge through the people interacting with the internal and external work environment to attain organizational objectives. Kyriakidis, Kant, Amir, and Dang, (2017) stated that other human factors such as performance, decisions play a significant role in a vast range of operations in STS. They termed these as Performance Shaping Factors (PSF) and attempted to establish a generic framework that researchers can apply across organizations, industries, and sectors, to structure and manage these factors. They identified the categories of factors as personal, dynamic, organizational, task, team, systems, and environmental/climatic conditions (Kyriakidis et al., 2017). In this study, I will explore the roles some of these factors have in the supporting the site in managing their IT operational costs in cloud computing.

Technical subsystem. Another aspect of STS theory is the technical subsystem. Fox (1995) defined the technical subsystem is comprising of the structures, tools, and knowledge necessary to perform the work which produces products. An easier way to view the technical subsystem is as an aggregate of all the systems, machinery, equipment, inputs, and tools, and other non-human elements deployed by the organization and required by the human factors for use to meet the organizational objectives.

Design principles. With an objective to leverage the requirements of both the social and technical systems present in the workplace, Trist (1981) and team identified some key characteristics of STS design which was later refined by Cherns (1987) based on earlier work by Parsons. Parsons (1951), identified four conditions that STS systems must meet to be successful as (a) attainment of the goal of the organization, (b) ability to adapt to the work environment, (c) activity integration, and (d) continuity of occupation and roles in support of the systems. Both Cherns (1987) and Parsons (1951) recognized the dynamic nature of the relationship between the technical and social subsystems, and how changes in one can significantly affect the other and the subsequent ultimate system objective. I have included a listing of the key STS design principles as Appendix C in this document.

Applicability of STS to study. Researchers focused early applications of STS theory on the manufacturing and production industries such as coal, textiles, and petrochemical, as an attempt to humanize highly mechanized work environments (Baxter & Sommerville, 2011). Success in the use of the theory persisted to varying degrees through the 1980s and 1990, during which time such concepts as lean production techniques took hold (Baxter & Sommerville, 2011). With the advent of software engineering in the 21st-century researchers have once again turned to STS design, recognizing the participatory aspects of the theory as an approach to address some of the issues related to traditional software engineering methods (Maguire, 2014). STS has also gained some acceptance with cloud computing researchers (Gavin, Baxter, & Hester, 2014; Khan, Büyükşahin, & Freitag, 2015; Kolevski & Michael, 2015). One such study is by Jede and Teuteberg (2016) to understand the SocioTechnical impacts arising from the use of SaaS in companies, and another by Chen, Chuang, and Nakatani (2016) to explore the perceived business benefits of cloud computing.

Cloud Computing Adoption

Since the introduction of modern cloud computing services in 2006, service providers have continued to expand available service options at a rapid pace. The increasing adoption of cloud computing services by enterprises is driven in part by key benefits of cloud computing such as scalability, fault-tolerance, accessibility, high-performance, reliability, easy use monitoring, and management (Sandholm & Lee, 2014). Between 2009 and 2014, about 20% of published cloud computing research related to cloud computing adoption, with another 5% on benefits and challenges (Bayramusta & Nasir, 2016). The focus of research on adoption may be an indicator of the currency of the technology and a marketplace still assessing organizational and technology fit. For those IT managers who have embraced and adopted cloud computing, one key to rapid service acquisition and use is the ease with which they can subscribe to cloud computing services. Since 2015, anyone can provision cloud services from Amazon Web Services (AWS) or any of the top providers with a just a credit card and a few mouse clicks. To understand what motivates IT managers to adopt cloud computing technologies, I will review some of the key benefits in the following section.

Key benefits and drivers for adoption.

***Capital cost savings*.** The traditional approach to IT technology refresh or extension is through the acquisition, installation, and configuration of new equipment at an on-premise of colocation facility. With cloud computing, this capitalization activity is unnecessary, as the responsibility shifts from the organization to the cloud service provider reducing the total cost of ownership (TCO) of the IT infrastructure. TCO includes hardware expenses, software licenses, IT personnel, and infrastructure maintenance (Assunção, Calheiros, Bianchi, Netto, & Buyya, 2015; Church et al., 2017; Opara-Martins, Sahandi et al., 2016). Further, the absence or reduction of on-premise IT infrastructure also reduces an organization's operational costs associated with power, air conditioning, and building administration expenses (Assunção et al., 2015). The reduction or conversion of IT capital costs to operational costs has distinct tax implications for organizations.

Technology agility and innovation. Cloud computing delivers substantial business benefits, including lower capital costs and the opportunity for organizations to reduce IT infrastructure management activity (leaving it to the cloud service provider) and focusing rather on their key business activities and objectives (Assunção et al., 2015; Padilla et al., 2015). On-premise DCs typically require efficient assembly and configuring of computer hardware (racking and stacking), software patching, and other time-consuming IT management tasks. Cloud computing removes the need for several of these tasks so IT teams can spend time on achieving more critical business objectives (Saha et al., 2016). Subscribers can provision and implement services rapidly and de-provision when subscribers no longer need these services (Padilla et al., 2015). The ease of provisioning and use of these readily available services facilitates shorter product development cycles.

The ability to scale operations globally without the need to setup and manage IT facilities globally is another important adoption factor for IT managers. Almost all the major cloud service providers have offerings in such technological areas as media processing, machine learning, data analytics, high-performance computing, and artificial intelligence (Makoza, 2015). AWS, for instance, has facilities in 14 regions around the globe, each region hosting multiple facilities for redundancy and failover. Service providers often upgrade these global facilities to the latest generation of fast and efficient computing hardware (Kaur & Chana, 2015). Global DC networking offers several benefits over a single corporate DC, including reduced network inactivity for applications and more effective scaling mechanisms. IT managers looking for rapid application and service offerings in these areas can rapidly provision resources on these platforms for system development.

Scalability and elasticity. The cloud enables dynamic scale-in and scale-out as well as scale-up and scale-down for all core service offered through automated or manual provisioning and de-provisioning of resources as needed (Hummaida et al., 2016). System subscribers can quickly acquire these system resources in response to increases and decreases in application workloads (Coutinho et al., 2015). Both scalability and elasticity address system scalability, but the concept of elasticity addresses the ability of a system to scale in response to variations in system use and workloads (Coutinho et al., 2015). In a traditional DC, IT managers must size systems to address the maximum expected resource utilization, even if they need such capacity during select periods of time. This capability to scale on demand is another area where organizations can use cloud computing to reduce their IT capital and operational expenditures.

Reliability, flexibility, and performance. Cloud computing makes data backup, plausible disaster recovery, and business continuity easier and less expensive because of options to replicate data to multiple redundant sites on the cloud provider's network (Alkhalil et al., 2017; Padilla et al., 2015). Most service providers offer guaranteed 24-hour and 7-day per week service level agreements (SLAs) to help with troubleshooting and to resolve issues arising from the use of their services (Alkhalil et al., 2017; Ouedraogo, Mignon, Cholez, Furnell, & Dubois, 2015; Padilla et al., 2015). Beyond the SLAs, service providers can also often guarantee higher performance and Quality of Services across the global dispersion of services (He, 2016), even though He (2016) and Ghahramani, Zhou, and Hon (2017) admitted that some challenges still exist with fulfilling these guarantees.

Energy efficiency. Cloud service providers represent their facilities as energy efficient compared to the traditional DC. Researchers have shown that minimizing the power consumption of DCs is important from both the operational cost and environmental impact perspectives (Mastelic et al., 2015). Cloud service providers have been able to use energy more efficiently than traditional DCs (Kaur & Chana, 2015). They achieve this energy efficiency by use of highly energy efficient infrastructure and reduction in the IT infrastructure itself by multitenancy (Kaur & Chana, 2015). An Accenture study in 2010 found that small businesses saw about a 90% reduction in emissions while using cloud computing, and medium and large businesses can save about

30 to 60% and 60 to 90% respectively when they use cloud computing resources (Accenture Sustainability, 2010). Williams, Thomond, and Mackenzie (2014) also recognized this potential for cloud computing as a source for greenhouse gas abatement, but they conceded that computing the market penetration rate must be much higher than what it was in 2014 for this to be a significant source of abatement on a global basis.

Available web applications for high performance. Another area in which cloud computing has been beneficial for organizations is in the provision of interactive and user-friendly interfaces, which makes services easier to provision, and to manage cloud-based resources. Subscribers can easily access cloud-based resource management tools via web browsers, so there is no need to learn to use different system interfaces for different system components, as is the case for on-premise systems (Ashraf, Byholm, & Porres, 2016). Cloud computing platforms also enable component-based application development. In component-based application development, developers do not write code from scratch but use code snippets like web services and third-party software libraries by retrieving and assembling them dynamically. The use of these services reduces application development time and errors in software development and software quality assurance, further enhancing organizational business processes and innovation.

Reduced application development time is only one factor that has made cloud computing appealing to managers. The provision of cloud development platforms has lowered the complexity inherent in the development of such systems as mobile interactive applications such as location, environment, and context-aware (Avram, 2014). These applications can respond to real-time human and nonhuman inputs to generate outputs. Applications such as high-volume batch processing systems, complex business analytics, and artificial intelligent (AI) systems as well as systems that have either high data systems requirements or compute-intensive lend themselves to cloud computing platforms because of the nature of their computing resource requirements (Avram, 2014).

Beyond the benefits of cloud computing, organizations still need to make adoption decisions. The literature review showed an extensive listing of research studies with objectives to understand the drivers for technology adoption by IT managers and consumers (Awa & Ojiabo, 2016; Nguyen, Newby, & Macaulay, 2015; Tarhini, Arachchilage, & Abbasi, 2015). The common frameworks researchers have used in these studies are the technology acceptance model (TAM; Davis, Bagozzi, & Warshaw, 1989), diffusion of innovation (DOI; Baskerville et al., 2014; Rogers, 1995), and technology-organization-environment (TOE; Tornatzky, Fleischer, & Chakrabarti, 1990). TAM applies to consumer's adoption of technology, and DOI and TOE for organizational adoption (Oliveira et al., 2014). Lian, Yen, and Wang (2014) addressed the critical factors affecting an organization's decision to adopt cloud computing. Consistent with the STS theory, Lian et al. (2014) identified five human and technical factors as data security, perceived organizational, technical competence, costs, top management support, and complexity (Lian et al., 2014). They also qualified the most important dimension as technology, followed by nontechnology factors. Khanagha, Volberda, Sidhu, and Oshri (2013), on the other hand, identified management innovation as a critical factor in the adoption of cloud computing. Management innovation is the propensity for the management staff to introduce new to the firm structures, processes, and practices

(Khanagha et al., 2013). A review of these studies indicated the same general factors for adoption across business sectors and regions, factors that align with the two dimensions of STS theory – technology and social factors.

Barriers to cloud adoption. A review of the benefits of cloud computing may lead one to believe that key benefits such as the potential to transform business processes, reduce capital expenditures, lower operating costs, and accelerate innovation will compel organizations to rush to adopt cloud computing (Oliveira et al., 2014). However, even though there has been substantial growth in cloud computing business, and the prediction is for this to continue (IDC, 2017a), some identifiable barriers still exist for wide-scale adoption by organizations. The following is a discussion of some of these barriers.

Security, confidentiality, and trust. A review of the academic literature on barriers to adoption has indicated that the most critical adoption barrier to cloud adoption is security (Lian et al., 2014). The risk to data security exists at all stages in the interaction between the subscriber and the cloud service provider. Establishing data security during transfers, storage, and the processing of information are important for managers when determining final acceptance of cloud technology (Iqbal et al., 2016; Ouedraogo et al., 2015). The perception of a lack of transparency regarding the geographical location of data stored with the service providers (Iqbal et al., 2016) is another issue facing service providers. Other researchers identified other such factors as confidentiality, integrity, privacy, and accountability as barriers to cloud adoption (Iqbal et al., 2016; Opara-Martins et al., 2016). The security, trust, and privacy concerns have all contributed to what Ouedraogo et al. (2015) referred to as a lack of trust and with service providers. This lack of trust coupled with a lack of standards across service providers as well as an absence of clear global regulatory guidelines (Yeboah-Boateng & Essandoh, 2014) is another barrier affecting cloud computing adoption.

Control of client data. The risk of unauthorized access to data is also another concern of IT managers potentially affecting service adoption. The multiregional availability of services, the use of other regional locations for storage redundancy, and global content distribution also seem to present a noticeable impediment to the acceptance and market success of cloud services (Alkhalil et al., 2017). Potential clients of these services often think that they lose power over their data because they are not sure about the legal implications of locating their organizational IT data in locations outside of their operations (Ouedraogo et al., 2015). The uncertainty about true data ownership, the location of data on service provider network, legal ramifications of data stored at foreign locations, and a general absence of clear guidelines have all contributed to impact the potential growth in cloud computing services.

Cost savings. One of the principal reasons early adopters gave for their decision to migrate to cloud computing is cost savings. Cost savings is consistently high on the list of reasons why organizations migrate to cloud computing (Alkhalil et al., 2017; Hsu, Ray, & Li-Hsieh, 2014; Yeboah-Boateng & Essandoh, 2014). However, once organizations make this transition, service providers present them with a wide range of hosting options to choose from, each impacting cost in a different way. Tak et al. (2013) referred to this as *cloudy with a chance of cost savings* and stated that navigating one's way through these options requires an in-depth understanding of the cost implications of all the possible choices specific to a subscriber's instance. Rana (2014) added that the actual costs for the long-term use of outsourced computing infrastructure (as is the case in cloud computing) could be unfavorable for smaller businesses. This challenge is further compounded by higher bandwidth requirements to access services (Ouedraogo et al., 2015). The lack of clarity and direction on how to select the most cost-effective services to meet technical requirements is an ongoing challenge for organizations looking to adopt services.

In a study to determine decision-making factors for migration to cloud computing, Alkhalil et al. (2017) explored the drivers for an organization's decision to migrate to the cloud. They developed a research model dependent on a combination of the DOI and TOE framework and tested the model using exploratory and confirmative factor analysis simulations. They concluded that cost benefits, business agility, system performance, fault tolerance, and data redundancy (in this order) are the key reasons for organizational decision to migrate to cloud computing. These findings are consistent with a previous study by another study by Hsu et al. (2014). Other factors such as competitive pressure, the complexity of IT needs, technology readiness, and trading partner pressure can be key in managers decision to adopt the technology (Gutierrez, Boukrami, & Lumsden, 2015).

The academic literature has a well-defined list of cloud computing benefits, but some of these benefits are more important to enterprises and seem to feature on top of the lists. They include cost-savings, reduced downtime again, reliability, scalability, and reduced operational support (Church et al., 2017; Hsu et al., 2014; Opara-Martins et al., 2016). In spite of evidence in the literature review of ongoing challenges with cloud computing services, the consensus view is that the benefits readily outweigh the issues (Assunção et al., 2015; Padilla et al., 2015; Ward & Barker, 2014). As service providers work to resolve some of the common issues, expand services, and reduce prices, more organizations are transforming their IT operations to cloud computing.

Adoption models (service offerings). Cloud service providers traditionally offer services under three main service models. These models form the basis for the sustainability of the cloud systems, such as managing cloud services and evaluating its service performance (Cartlidge & Clamp, 2014; Khan, Kiani, & Soomro, 2014). The model architecture consists of three layers: Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service.

Infrastructure-as-a-service (IaaS). The IaaS offering is the closest service to the traditional DC service that most organizations implement. Banerjee, Gupta, and Gupta (2014) described the key characteristics of IaaS as computing resources (a) distributed as a service, (b) that support dynamic scaling, (c) that has a variable cost of a utility pricing model, and (d) that entails some degree of resources sharing. In this offering, subscribers can provision core IT resources such as computers, database systems and storage services and directly deploy `them in support of their applications (Padilla et al., 2015). AWS provides such IaaS services in the form of their Elastic Computing Service (EC2) and the S3 storage services. In general, service providers organize IaaS services in cluster-like structures to promote the use of virtualization technologies to create virtual machines (VMs) and subsequently *rented* to subscribers (Kumar & Charu, 2015). Cloud service providers are also responsible for the management and maintenance of the core infrastructure and the subscribers the virtual machines they acquire for their computing and storage needs.

IaaS services are a good fit for specific organizational computing needs. Their use is particularly suitable for organizations in instances where (a) volatile demand for infrastructure services, (b) lack of capital to invest in hardware or pressure to reduce such costs, and (c) need for rapid infrastructure scaling, (d) need for temporary IT infrastructure (Banerjee et al., 2014). The use of IaaS services may be unsuitable for subscribers where regulatory compliance limits data storage and processing locations, and where levels of performance may mandate on-premise or dedicated hosting services (Banerjee et al., 2014). An organization looking to subscribe to these services must access its needs and understand the options before adopting this model.

Platform-as-a-service (PaaS). Cloud service providers specifically target software developers with the PaaS category of cloud services. With PaaS services, providers present developers an option to use infrastructure and application stack services under the provider’s environment. The services include programming languages, libraries, and tools (Chun & Choi, 2014). The subscriber controls the software deployment and configuration, and the service provider the network, servers, storage, and the platform solution stack (Banerjee et al., 2014). The benefit to the subscriber is that they can focus on application related activities such as development, testing and release, and not on platform related setup and maintenance.

The use of PaaS is beneficial for organizations in situations where multiple development groups (including external parties) need to interact with each other as part of the development process. Banerjee et al. (2014) suggested that the use of agile software development may help the growth of PaaS services, because of the potential to ease the difficulties related to rapid development and rate of iteration. PaaS offering may not be suitable for organizations with a need for proprietary languages not offered by the service providers, or where there is a need to customize the underlying hardware and the development stack (Banerjee et al., 2014) as service providers do not permit such modifications.

Software-as-a-service (SaaS). Majority of prospective customers are familiar with this higher abstraction of service for subscribers. SaaS applications are distinguished from other service offerings because there is no need for the provisioning of hardware or platform services. Subscribers can provision the direct use of the application services, often on a user by user basis, and leave the management of the application stack and any related hardware to the service provider, and can assess the application anywhere (and on-demand). SaaS solutions are often available via a web interface and delivered in a one-to-many model, and the users or subscribing organizations do not have any responsibility for software updates, upgrades or patch application (Banerjee et al., 2014). Notable SaaS examples are SalesForce CRM, Google Apps, Outlook email and calendar applications.

Some types of applications and organizational needs lend themselves well to SaaS solutions. These include, (a) generic applications requiring little or no customizations to use (like email), applications used by organizations to interact with customers, applications used for web and mobile access, and applications with seasonal spikes

(Banerjee et al., 2014). SaaS offerings are traditionally not suitable for solutions requiring a high degree of customization or performance.

Cloud computing adoption frameworks. Beyond the service model, a review of publications in cloud computing indicated that researchers had suggested several frameworks for cloud computing adoption. The Cloud Computing Adoption Frameworks (CCAF) are guidelines for implementation to ensure that service deliveries can meet technical and organizational challenges (Chang, Kuo, & Ramachandran, 2016). Consistent with the concerns about security, the majority of the frameworks and related studies are in the area of security, including the studies by Chang et al. (2016), Chang and Ramachandran (2016), and Hussain, Fatima, Saeed, Raza, and Shahzad (2017). In another publication by Chang, Walters, and Wills (2015), they listed eleven of these frameworks and identified known limitations with each of them. Most of the limitations range from an absence of known field studies employing the frameworks to lack of sufficient details on how to apply them in organizational settings. None of the frameworks they reviewed met the technical, financial, user, and complexity requirements necessary for dealing with emerging issues and range of solutions organization need for cloud computing adoption (Chang et al., 2015). So even though these frameworks exist in principle, there is insufficient evidence in the literature, which they are anything more than academic.

The one CCAF that addressed cost concerns associated with the return on investment for cloud computing (ROICC) framework proposed by Skilton (2010). In this study, Skilton (2010) outlined the pricing and costing structures in cloud computing, funding approaches, ROI factors, key performance indicators, TCO, risk management, and the decision and evaluation choices available to organizations looking to adopt the technology. Chang et al. (2015) indicated that the limitation to this ROICC framework is the lack of any prescriptive approach to the calculation of ROI and cost the cost-benefit analysis. This weakness in the framework limits its application as an effective tool in an examination of cost factors in cloud computing adoption.

Cloud Computing Pricing and Operations Management

Cost components. The leading service providers like Amazon, Google, Salesforce, Oracle, SUN, and IBM continue to expand their computing infrastructures and platforms as the basis for providing excellent services for computing, data storage, database, and network applications (Ward & Barker, 2014). The service providers are also adding new features and capabilities to application services such as email, office applications, finance, video, audio, and data processing services (Dautov, Paraskakis, & Stannett, 2014). At the same time, these service providers are making it easier to access and use their services, often touting prices reduced prices for these offerings not easy for IT managers to ignore. The ease to with which organizations can acquire and use cloud computing services within the enterprise leads to a tendency to over-provision such resources, and sometimes to even under-utilize what they provision for organizational use. Over-provision and under-provision of cloud resources can both have adverse cost and performance consequences for the cloud service consumer (Sandholm & Lee, 2014). Over-provision, especially the type that leads to under-utilization of cloud resources is one key factor in cost management.

The easiest way to do a breakdown of cloud computing costs is relative to the traditional cloud computing service offerings – SaaS, PaaS, and IaaS. Mayo and Perng (2009) however, took a more practical approach and identified the five areas for the cost for cloud computing payback as qualifying the key cost areas hardware, software, automated provisioning, productivity improvements and systems administration. Of these five cost areas, the last three have a direct bearing on IT operational costs, which is the focus of this study. Software and hardware related costs, even though they are relevant from an enterprise viewpoint, fit more into the capital expenditure category. This view of cloud computing costs is specific to the determination of the payback period for cloud computing investments but is an effective way to categorize the costs for enterprises.

The most common types of articles that addressed cloud computing costs are related to the assessment of financial metrics like ROI (Maresova et al., 2017) and cost-benefit analysis studies (Mohammed, Aljumaili, & Salah, 2014). The study by Tak et al., (2013) was more specific to identify a taxonomy of costs involved in in-house application hosting compared to cloud computing. The taxonomy they developed categorized costs as indirect or direct, and the related expenses as quantifiable or less quantifiable (Tak et al., 2013). Direct, quantifiable costs included material, labor and utility expenses (applications), including charges for cloud computing use, whereas the quantifiable indirect costs included facilities, staff salaries, DC utility costs (Tak et al., 2013). The majority of the studies on cloud computing by researchers were focused more on resource utilization and optimization of direct computing resources (Khan et al., 2015; Manvi &

Shyam, 2014; Mosch et al., 2014; Reddy et al., 2017). These types of resource optimization studies outnumber other studies directly related to cost, and is an opportunity for additional studies in this area of cloud computing, especially on operational costs.

Surveys have been one way for researchers to understand the state of technology adoption and use in enterprises. Researchers completed a series of such surveys on cloud computing adoption between 2016 and 2017 to evaluate the use of cloud computing (Alkhalil et al., 2017; Hummaida et al., 2016; Opara-Martins et al., 2016; RightScale Research, 2018a). These studies focused on cloud customers and users, rather than cloud retailers. Notable among these surveys is the one by RightScale, which found among others, that (a) cloud users underestimate the cost of wasted provisioned, which respondents estimate around 40 – 45%, (b) optimizing cloud costs was the top initiative across all cloud users (53%) and particularly, among experienced cloud users (64%), and (c) lack of competent cloud technical resources/expertise continues to be a major concern for organizations (RightScale). What is noteworthy about these findings is that they are all related to the operational management of cloud environments. As more IT managers transition from cloud transformation activities to operational maturity in cloud environments, the concerns of managers are slowly shifting from cloud adoption and transformation challenges to operations management and is another indicator of the need for a better understanding of resource utilization and cost management in cloud computing.

Pricing models. Service provider pricing models, like all other vendor pricing models, have significant implications on the selection and use of cloud services for any organization using cloud resources. Irrespective of the service offering, all the vendors have various options available for selection and use by the service consumer. In general, the service providers classify their services based on the service types, like infrastructure, data storage, database, email, web services, data processing, data sharing, business process management, marketplace, billing, and accounting. What follows is a discussion of the common pricing models, with a specific emphasis on options by AWS, which the is the platform I selected to use for this study.

***Conventional pricing model*.** The most common pricing model used by cloud service providers is a pay-per-use model, where the user pays a fixed price for a unit of the resource per a unit of time. The computing resource may be CPU per hour (or an aggregate of CPU and GB of Memory for instance). In the case of storage, defined as GB of storage provisioned per month (Ward & Barker, 2014). Pay-per-use pricing models are simple widely used by service providers for products or services which mass production and universal delivery make price negotiation impracticable (Pakath, 2015). A comparable pricing model is a subscription model, where the customer (signs a contract for fixed charges (subscribe) per billing period, usually monthly or annually.

Dynamic pricing. Also called variable pricing, dynamic pricing is a pricing model in which the service provider fixes a target service price which is reflective of the ever-changing state of the supply and demand for the service. An example of dynamic price changes is what occurs in auctions or price negotiations (Hummaida et al., 2016). If the demand is high, pricing increases; if demand is low, pricing decreases. A reduction in price leads to tenants (consumers or clients) who cannot afford the higher pricing, so they withdraw from it, which decreases demand, and pricing falls again. In turn, this encourages some tenants to increase their use, allowing a dynamic flow of price changes with alterations in demand and supply. The main advantages to IaaS providers are the ability to make more profit on peak demand (a point in time in which the demand and provision of resources reach its maximum level) by increasing charges on premiums. A dynamic pricing mechanism is effective to move workloads off-peak hours by charging a premium for it and forcing consumers to consider using alternate periods for computation to reduce their costs and use the spare service provider capacity during such times.

Price setting. Price is an essential factor to determine profit for the service provider and essential cost factor for the service subscriber. For this reason, a good pricing strategy is necessary to set the right prices to attract users. In setting prices in computational markets, bid auctions are popular since they are efficient for setting prices that reflect the demand for the product or resource, with limited communication (Sandholm & Lee, 2014). Bidders typically place secret bids, and the highest bidder is the client who gets to the use of the resource. Conversely, if there is a single provider who is trading in computing resources, spot pricing (the current market price for purchasing or selling a resource for instant payment and delivery) is an efficient way of setting demand-based prices, for a guaranteed bid which would fulfill a single spot. Consumers can also purchase resources on a spot market (Sandholm & Lee, 2014). Spot pricing, however, is associated with the consumers being at high risk of either of paying higher for the same allocation or compelled to reduce the allocation to fit their budget. A standard way of reducing the consumers' pricing risk is to offer a reservation market price (Sandholm & Lee, 2014). A reservation market computes an expected spot demand for some time and adds an uncertainty premium to arrive at a reservation price to move the risk from the consumer of computing resources or the tenant or the provider. For any IT manager looking to transition to cloud computing and understanding of these basic pricing structures as well as the profiles of the applications, they plan to deploy in the environment are essential.

Application profiling. A discussion of resource utilization in any environment will not be complete without directly addressing the systems that require and use the resources. Most systems that consume computing resources perform optimally within a range of computing resources (specifications). These computing specifications are the basis for system provisioning, and since provisioning has a direct bearing on the cost of operation, an understanding of the application profile is necessary to avoid over/under-provisioning of resources for the application. Application profiling is an approach to qualify the use of computing resources by an application and its expected future requirements (Weingärtner, Bräscher, & Westphall, 2015). Application profiling is necessary for three key reasons (a) support application management to prevent service degradation, (b) optimize application resource utilization to meet SLAs, and (c) manage application cost impact (Weingärtner et al., 2015). Some of the core characteristics of application profiling include accuracy, application design, background workload, historical data, migration, network affinity and SLA (Weingärtner et al., 2015). Of these characteristics, historical data is the most prevalent in the academic models (Weingärtner et al., 2015). The main challenge for IT managers transforming to cloud computing is the lack of current profile data on applications. The absence of application profile information across an enterprise portfolio is a gap that makes the task of cost management in cloud computing a reactive instead of a proactive activity.

AWS pricing models. In this section, I address the current pricing models specific to AWS, the cloud computing environment I selected for use in this study. AWS has a myriad of pricing structure, but I will focus the discussion here on the core services such as compute, databases, storage, and networking services. The other IaaS vendors employ similar pricing models but may use different terminology to represent them.

AWS EC2 instance pricing. Amazon Web Services has pricing models for its flagship EC2 instances similar to the pricing schemes I discussed in this section. The EC2 instance pricing depends on the type of service required by subscribers. The services are available in three categories – On-Demand Instances, Reserved Instances, and Spot Instances (He, Wen, Huang, & Wu, 2014). Customers who purchase on-demand instances can pay off the rented computing environment they choose without any upfront payments or long-term commitments. The prices vary by the instant types (an aggregate of CPU, Memory, and OS specifications). Customers who choose reserved instances, on the other hand, make a relatively low (about 50% lower than a comparable on-demand option for the same instance), one-time payment for a long-term reservation of the instance (one to three years) (He et al., 2014). For applications with low variations in compute resource requirements and a stable operating profile, using a reserved pricing option can present subscribers with a lower cost option.

The final price category is for the spot instance. Spot instances require no upfront commitments, but the customers have to specify the maximum hourly price they are willing to pay for the instance types. AWS sets a spot price for the EC2 instances dynamically. If the customer's specified price is higher or equal to the spot price, the instance is activated and will continue to run until the spot price is higher than the customer set price, at which time the instance will be taken offline. This way, the customer does not ever pay the price higher than their specified price for any instance they use (He et al., 2014). This type of instance pricing is a good fit for applications with very high computing resource requirements which are also non-real-time.

AWS database services pricing. Amazon Web Services offer two types of database services. Subscribers can provision a compute environment and configure it to fit the traditional database services such as Oracle, SQL Server, or MySQL. Subscribers can also purchase Amazon Machine Images (AMI) from the AWS marketplace pre-configured with the install images to reduce the deployment effort. The pricing models for EC2 instances apply to this approach as well, and the standard pricing may include the necessary database licenses cost. Also, AWS offers a fully managed Relational Database Service (RDS), which is a fully managed version of the database services. AWS represents RDS as an easy to setup, cost-effective, easy to scale and operate, with AWS automatically performing data system patching, replication, backups, and other operational management activities (AWS, 2017d). The subscriber only has to focus on the data in the databases, and not on the normal operational maintenance activities. One database service which Amazon prices differently is Aurora. Amazon bases the Aurora pricing model on a combination of four factors—storage volume (GB), Input/ Output rates (number of requests per month), storage volumes (GB), and data transfers (GB) in and out of the instance (AWS, 2017a). For an IT manager not used to assessing costs at this component level, the various pricing options available with cloud service providers can create cost management complications.

AWS storage systems pricing**.** Amazon Web Services offer two main types of storage services, object storage they refer to as a Simple Storage Service (S3), and block-based service called Elastic Block Store (EBS). Subscribers can use S3 to store all types of objects, flat files, media files, and other such data files, and can access them through a web interface. The files are stored in buckets (folders) within S3, and users can access them using a web interface or API calls. EBS volumes, on the other hand, are used by AWS to provide persistent block storage for use with EC2 instances. AWS represent them as highly available and reliable storage volumes available to be attached to instances in the same region as the EBS volume (AWS, 2017c). The key difference between the two is that S3 storage cannot be mounted directly as drives for an EC2 instance, but APIs exist to allow such instances to access S3 buckets.

In the case of S3, Amazon categorizes the storage into three groups – standard, standard – infrequent access, and glacier storage. These categories reflect the SLA for data retrieval from the storage system. The final pricing for S3 subscribers is based on this storage category, as well on factors such as (a) a graduated threshold volume per month (GB/month), with the price per GB lower for higher volume tiers, and (b) number of upload requests per month (c) number of download requests from the bucket per month, and (d) number of lifecycle transition requests per month (AWS, 2017e). AWS pricing on EBS volumes, on the other hand, is based on the type and performance specifications for the volume. Four types of storage volumes are available, and AWS has standard unit costs for each of these volume types, expressed as unit cost per GB-month for the provisioned volume (AWS, 2017c). In some instances, subscribers can incur additional costs associated with additional provisioned IOPs per month.

Beyond the core AWS services, subscribers may incur additional charges related to other IT infrastructure components such as networking throughput and the use of networking load balancers. Maresova et al. (2017) stated that compute, database, and storage services are foundational for many of the services that higher cloud services stacks are based and form a solid foundation for use to assess general costs of cloud services. Some cloud service providers, including AWS, provide costing calculators to aid subscribers in calculating costs because determining monthly billing can be complicated for IT managers. But even with that, organizations still face challenges understanding all the computing needs and the best approaches to managing related costs on an ongoing basis.

AWS, like other IaaS service providers, have a virtual meter on all core resources, and a measure and access a cost for it as part of the infrastructure services they provide. As I stated earlier, for organizations that have historically owned and managed their IT infrastructure, transformation into an environment where the use of every resource has direct cost implication presents a challenge for IT operations managers. In this study, my objective is to explore the strategies IT managers use to manage operational costs in these cloud computing environments.

IT operations and cost management in cloud computing. IT operations managers are responsible for the smooth functioning of the IT infrastructure and operations environment for the applications and systems that support business processes and organizational activity. IT operations management is defined as the process of managing the provisioning, capacity, performance, and availability of an organization's IT infrastructure including on-premises data centers, private cloud deployments, and public cloud resources (IT Operations Management, 2018). The objectives of these operations activities for most organizations is to ensure performance and availability of digital services, optimize costs and risks to business needs, efficiently run technologies that underlie business processes, and respond to problems in the environment promptly (IT Operations Management, 2018). So operational costs management is the responsibilities of the operations team. In this study, I will explore how IT managers who effectively manage costs in cloud computing environments are doing so through the management of the technological and social factors (consistent with the STS theory) in such environments.

Transition

Since 2006, cloud computing has become the most dominant platform for computing in the world. The promise of on-demand computing, pay as you go price offering, capital cost savings, technological agility, and the possibility of limitless computing resources both elastic and scalable on demand makes it an attractive proposition and a viable platform for all IT professionals. Cloud computing adoption by organizations was slow at the onset, but the rate of adoption has grown considerably since 2010 (IDC, 2017a). An important motivator for cloud computing adoption by organizations is the cost savings. The perception of cloud computing a vehicle to rein in costs is consistently at the top of surveys of IT managers (RightScale, 2018a), but the realization of true cost savings, once organizations transform to cloud computing is not as obvious as it appears. The challenge to realize promised cost savings in cloud computing is complicated by the nature of the shift in the paradigm between traditionally in-house hosted IT infrastructure and cloud computing.

In this section of the study, I established the need for the study and completed a comprehensive literature review tracing the evolution of cloud computing, cloud computing adoption and the organizational motivation for doing so. To perform the study to explore effective strategies and processes used for cost management in cloud computing, I selected the STS theory as the framework. STS theory is a framework for researchers to use to study the optimization of complex systems by the optimization of its related technical and social systems (Trist, 1981). This view of looking at systems is consistent with other approaches such as TOE.

Because of the correlation between pricing and cost, I also performed a comprehensive review of the pricing models and options with some of the core technologies in the cloud computing environment. The actual computation of operations costs may be more difficult than the case for traditional DC environments and the representations I made of the pricing models in this section. In Section 2, I will also outline the research method and design, the population, sampling, data collection, data analysis, and validation techniques. I will complete the study in Section 3, in which I will document the outcome of the research with the findings and conclusions, including recommendations for further research and conclusions

Section 2: The Project

In Section 1, I addressed the foundation of the study, providing background and defining the purpose and problem statements for the study and leading to the overarching research question. I also identified a conceptual framework to use for the study and a statement on why I believe this study to be significant for professional practice and social change. I concluded the section with a review of the professional literature about the subject, offering syntheses and critical analysis of the related themes from previous publications and identifying gaps that may be bridged by this study.

In this section, I focus on my role as a researcher in the study and describe the ideal participants for the study, their eligibility criteria, and how I sought to gain access to them. I also provide a justification for the selection of the population and the sampling methods I used in the study, with due considerations for ethical research principles. Additionally, I address data collection strategies—instruments, techniques, and analysis—and conclude by this section by addressing how I sought to ensure reliability and validity in the study.

Purpose Statement

The purpose of this qualitative multiple case study was to explore strategies and processes that IT managers use for managing infrastructure operations costs in cloud computing environments. The target population was IT operations managers from two IT organizations in the Tampa, Florida area with a record of successfully managing IT infrastructure operational costs in a cloud computing environment. IT managers may benefit from the findings of this study because the strategies may help them understand the pathways for successfully managing their infrastructure costs through cloud computing. By managing and reducing costs, organizations can focus on their core business, improve innovation, and sustain competitive advantage (Oliveira et al., 2014). The economic viability of these organizations may lead to more job opportunities within the local communities where they operate and lead to positive social change. Cloud computing service providers use energy-efficient and environmentally friendly technologies for their facilities (Liu et al., 2015). Operational cost savings in cloud computing, which can have the long-term effect of increasing cloud computing adoption, also has the potential to reduce an organization's environmental impact and carbon footprint through reduced energy use.

Role of the Researcher

Several characteristics distinguish qualitative studies from those conducted with other methods. One such difference is the data collection methodology and the role of the researcher in that process (Turner et al., 2017). In their capacity as key instruments in data collection, Yilmaz (2013) recommended that researchers declare their relationship to the participants, the location, and/or the sites where data collection occurs. Furthermore, researchers must state any experiences they have relative to the topic to mitigate any potential issues with strategy and ethics, as well as any personal issues related to the research process (Yilmaz, 2013).

In this qualitative case study inquiry, I was the key instrument in the data collection process. I have worked as an IT consultant on projects for several organizations all over the United States. In the last 5 years, my work has focused principally on helping organizations migrate their on-premise IT infrastructure and applications to new DCs, private clouds, public clouds, or hybrid clouds. In this capacity, I have performed several in-person and over-the-phone interviews with potential candidates for employment directly for my company or on behalf of my clients. I have also performed other technical interviews (and other information-gathering sessions) as part of processes to gather business requirements for projects. Another normal part of the consulting work I do is reviewing customer data and reports to identify gaps or validate information from other data sources, so I have experience in reviewing multiple data sources for triangulation. Additionally, I am a certified AWS architect and a cofounder of an organization certified as an AWS partner.

In my role as the instrument for data collection, I was committed to the guidance, recommendations, and protocols of the Belmont Report. The Belmont Report, enacted into law in 1974, is an analytical framework designed to ensure the protection of the interests of participants in a research study. It serves as a guideline for the treatment of participants in a research study, with basic ethical principles such as respect, beneficence, and justice (Office of Human Research Protections [OHRP], 1979). To ensure that I treated participants in this study in a manner consistent with the Belmont Report, I made it clear to them that their role in the study was voluntary and ensured that there was a documented process for informed consent. I did not permit management staff to compel any individuals to participate against their wishes. There was no expectation of harm to participants from the data collection process, nor any direct benefits to participants in the study, so there was no risk of violating the principle of beneficence and justice in this study.

Closely related to issues involving the role of the researcher and ethical considerations in research is the effect of bias on the study. Bias is inherent in every research work and can influence any step in the process, the validity and reliability of the findings, and subsequent representation for practice (Noble & Smith, 2015). The concepts of validity and reliability, typically addressed statistically in quantitative methods, are also applicable to qualitative methods. Noble and Smith (2015) stated that researchers can infer validity from the integrity and application of appropriate methods in the study, and reliability from the consistent use of analytical processes.

To minimize bias, I outlined the limitations of this study and documented the extent of my exposure to the topic. I also avoided any prior personal or professional relationship to the case, site, and location selected for the study. I carefully selected the case and participants for the interviews to ensure the relevance of their knowledge and roles in the organizations. Before completion of the interview protocol, I performed a pilot activity with a colleague to ensure that the questions that I used for the interviews were correctly phrased, would be understood by the participants, and were open ended to elicit responses that reflected the exploratory nature of the study. I used a purposeful sampling technique to identify and select participants for the study. Palinkas et al. (2015) and Walker et al. (2015) both recommended this sampling technique in instances where the interest is in using individuals who are especially knowledgeable about the phenomenon and who are available, willing, and able to express their knowledge and opinions on the subject clearly. I planned the interviews to be conducted face to face, except for one participant who had a change of schedule at the last minute and could only take part in the interview by phone. During all of the interviews, I took notes in addition to recording the audio to ensure the accuracy of the data for analysis. I also encouraged the participants to share any documentation they had that might provide further elaboration or corroboration of responses from the interviews, as proposed by Yin (2017). Finally, in the data analysis phase, I was careful to ensure that themes and related outcomes were reflective of responses from the participants and not shaped by my personal views and prejudices on the topic.

Participants

Successful cost management entails a complex set of coordinated activities in the enterprise. Liu (2015) stated that modern enterprises often have a range of activities to be effective with cost management, and it is necessary to combine specific requirements of strategic cost control in all applicable departments operating in the company. To this effect, in this study exploring strategies used by organizations to manage IT infrastructure operating costs in cloud computing, individuals eligible to participate could come from any department in the organization—such as finance, accounting, IT systems operations, cloud architecture, and software development teams, including technologists from all departments. An important requirement for anyone from any of these departments to be eligible to participate in the study was that participants needed to have knowledge of (or be directly involved in) efforts to manage IT infrastructure costs for the cloud computing systems.

To gain access to the group, I solicited the assistance of the IT leadership team from the site after initial Institutional Review Board (IRB) approval to proceed with the study and provided the team with selection criteria for participants who would be the right fit for the study. I sought to obtain a description of potential participants' roles in IT infrastructure cost management and validated that they met the eligibility requirements. The selected individuals formed the population that I targeted for data collection. After identifying this subgroup, I requested that IT leadership send a prepared email to the group to introduce me and my study interests to the group, and to copy me on that message. The email directed potential participants to contact me directly if they had any questions or wanted to volunteer for the study. I was available to answer any questions that I received from them, including those related to the interview process, my ethical considerations for the study, and concepts of transcript review and member checking. In addition to helping to improve the validity of a study, member checking and transcript review can improve the overall trustworthiness of content in studies (Elo et al., 2014). Ensuring that participants understood my intent to use these methods and that I would accurately record information they provided during interviews (and not misinterpret data in the analysis phase) helped to promote a positive working relationship with them before the interviews. For those who responded to me with interest to participate, I sent a copy of the approved consent form with instructions to review and to respond if they still agreed to participate. After I received responses from those who consented, I contacted them directly to schedule the interviews onsite at a time of their convenience.

Research Method and Design

In this section, I discuss the research method and specific design under the method that I selected, providing a rationale for the decisions to use them for this study. Research methodology differs from research method in that methodology is a strategy or design that the researcher selects that directs the choice and use of specific methods to attain study objectives (Jamshed, 2014). The selection of methodology is driven in part by the nature of the research question, and the methodology, in turn, dictates the choice of methods and related designs available to the researcher for use to answer the question. When a researcher defines research methods and designs in a study, it gives both the researcher and evaluators alike a known and acceptable structure for a common understanding of the approach to answering the research question (Ioannidis et al., 2014). In the following sections, I discuss the rationale for my selection of the research method and design.

Research Method

I considered three possible research methods for use in this study—qualitative, quantitative, and mixed methods. Qualitative methods involve inductive thinking by nature, whereas quantitative methods are deductive. Creswell, Hanson, Plano, and Morales (2007) proposed a framework for identifying when to use a qualitative or quantitative methods in a study. Their framework was based on the researcher's worldview, the object of the study (who or what), the strategies to be used for the study, and the mechanism by which data will be collected and analyzed for the study.

Qualitative methods are applicable when the goal of the study is to explore a phenomenon (a single or collective case) with the objective of seeking perspectives and understanding the meaning that participants assign to their observations and experiences as it relates to the phenomenon under study (Pietkiewicz & Smith, 2014). An extension of this definition of qualitative research by Yilmaz (2013) is the view that qualitative research is a naturalistic approach to the study of a things (people, cases, phenomena, social situations, and processes), done in descriptive form, and in natural settings to understand the meaning that people associate with their experiences. In contrast, objectivist epistemology informs quantitative methods, for which researchers employ statistical methods to establish relationships between variables under study (Yilmaz, 2013). The emphasis in quantitative research is on the measurement and analysis of causal relationships between variables through hypothesis testing.

The third research method that I considered was mixed methods. A mixed method approach is a hybrid of the qualitative and quantitative methods and is used by researchers when there is a need for both qualitative and quantitative methods in the same inquiry (Percy et al., 2015). A fundamental element of this method that differentiates it from the other two methods is the sequential or parallel use (in the same inquiry) of qualitative and quantitative methods in data collection, analysis, and presentation (Turner et al., 2017). Because this study involved exploring strategies and processes used by organizations to manage IT operational costs in cloud computing from the perspective of IT operations managers, and there was no intent to perform any measurement and analysis of causal relationships or to test any hypotheses, quantitative and mixed methods were not applicable. I selected a qualitative method as the correct fit for this study.

Research Design

Even though the nature of qualitative research method involves exploring how people interpret their experiences, how they construct their world, or what meaning they attribute to their experiences, researchers using the qualitative research method of inquiry have the option of using different research designs under this method (Kahlke & Hon, 2014). Each design is a better fit for a specific type of inquiry. Lewis (2015) identified the four main research designs under the qualitative method as narrative, case study, phenomenological, and ethnographic.

Researchers using qualitative methods employ the use of narrative research design when the objective of the study is to detail an event or outline a series of sequential actions in relation to the phenomenon under study (Percy et al., 2015). Ethnographic research design, on the other hand, is applicable when the study is about a group sharing a common culture and the objective is to describe and explore behaviors, beliefs, language, and interaction of or within the group (Percy et al., 2015). Whereas researchers use narrative design to provide an account of a phenomenon, ethnography is about the understanding of the shared experience of a group with common attributes. Neither of these designs was applicable for use in my study because there was no intent to narrate or to explore belief systems of a group in the study.

Of the other two main qualitative research designs, researchers use the phenomenological design for studies with a goal to explore and understand the lived experiences of participants, focusing on attitudes, beliefs, opinions, and other internal cognitive dimensions (Percy et al., 2015). Finally, case study research is applicable when the object of the study is to explore a real-life situation in a specific setting (which also bounds the study), with the aim of understanding what is distinctive about that situation and setting (Hyett et al., 2014). Cronin (2014) extended this definition by qualifying the case study approach as systematic, conducted over a period, and characterized by in-depth data collection and analysis.

In summary, researchers using case study design can perform a more in-depth analysis of the phenomenon under study using multiple sources of data from a specific setting (the case) to facilitate a better understanding of the phenomenon (Yin, 2017). In this study, to explore strategies and processes used by organizations to manage IT operational costs in cloud computing, I did not intend to explore attitudes and beliefs of individuals, but rather to perform a detailed exploration of a real-life phenomenon in a specific setting. I selected a case study design as the one most applicable for use in this study.

In qualitative research methods, researchers consider data saturation to have an impact on the quality of a study and the validity of the content (Fusch & Ness, 2015). Researchers attain data saturation in a study when there is enough information to replicate the study (Fusch & Ness, 2015), and when additional information collected has no potential to add on to existing information about the phenomenon (Cronin, 2014). Data saturation is not usually easy to predetermine, but Fusch and Ness (2015) identified three areas that researchers must address to ensure data saturation: (a) using proven data collection techniques, (b) ensuring that the researcher's worldview does not influence data interpretation, and (c) data triangulation. In this study, I planned to interview a minimum of five IT managers within the organization to facilitate data saturation. I also reviewed transcripts from the interviews with the interviewees to ensure reliability and validity. I intended to collect other supporting documentation available such as cost management reports and resource utilization reports for review and to validate against data the interviewees provided to facilitate data triangulation, but none of the participants provided any supporting documentation.

Population and Sampling

Researchers recommend that a sampling technique selected for use in a study must be consistent with the study design method. Cleary, Horsfall, and Hayter (2014) identified some important participant selection principles as studying small number intensively, choosing participants purposefully, selection driven by the conceptual framework, and specification of rationale. In the case of a qualitative study method, it is necessary to achieve a depth of understanding of the phenomenon from the site and participants, as well as achieve data saturation (Roy, Zvonkovic, Goldberg, Sharp, & LaRossa, 2015). To identify and select the best available participants from the site that I select for this qualitative multiple case study, I used a purposeful sampling technique.

Purposeful sampling techniques are effective for use in selecting participants for cases where the depth of information (richness) is more important than breadth, and eliciting insights from individuals with special knowledge or experience with the phenomenon (Palinkas et al., 2015). Other commonly used strategies such as random or convenience sampling will not be appropriate for use in this study. Random sampling is a process researcher's use when selecting cases from a general population (a subset) to serve as a representation of the population at large, convenience sampling, as the name implies, is a selection criterion dictated by convenience factors such as proximity and availability (Robinson, 2014). In addition to experience relative to the subject of the study, Palinkas et al. (2015) also stated that selected individuals must be willing to participate and be able to articulate their knowledge on the subject coherently.

For the site for the study, I initially obtained permission from the IT leadership for the study. I submitted the permission to the IRB for their initial approval then worked with the IT leadership (Chief Technology Officer) to send email to the potential pool of participants from across multiple departments. Selecting individuals from multiple departments made it possible to collect data from varying sources and perspectives from the same site to improve rigor (Boddy & Boddy, 2016; Cronin, 2014). Yin (2017) recommended using about six sources of data for case studies, so I targeted5-6 individuals from different departments within the same organization to meet this requirement. I performed the interviews on site and based the interview schedule on the availability of the participants. I requested to use a room away from general human traffic to reduce potential distractions and ambient noise that could impact the clarity of audio recordings of the sessions. Providing an environment conducive to uninterrupted interview experience is consistent with the recommendations of Rimando et al. (2015) to conduct interviews at the convenience of the participants, and safe from interruptions. I shared the interview schedule and interview guide with each participant in advance of the actual sessions.

I planned to supplement the interviews on site with other data sources from the same organization to facilitate both data sources triangulation and saturation. The data sources I requested included IT operations management cost records by month, documented operating procedures, and IT assets provisioning and system use reports. Data triangulation is a technique researchers use to verify the accuracy of data they collect for analysis by checking it against other sources of the same data (Hussein, 2015). Data saturation, the point in data collection when (a) there is sufficient data collected for someone else to replicate the study (Fusch & Ness, 2015), and (b) collection of additional information does not add to any new themes or new information on known themes (Cronin, 2014). I extended interviews to additional participants until I achieved data saturation.

Ethical Research

Ethical considerations in academic research are important because the process of a research study involves steps such as study design, data collection design, use of participants for data collection, data recording, analysis, and reporting of findings (Baines, Taylor, & Vanclay, 2013). Navigating through the process entails having to make decisions regarding what available options to use to complete the steps, but the danger exists for lawful, yet unethical behavior in that decision-making process. College Institutional Review Boards (IRB) are one mechanism to ensure the protection, rights, and interests of research participants, but this is only one aspect of the ethical considerations for a researcher, and Walden University IRB has specific processes and procedures that students must adhere to as part of an approved research study.

One element of ethical considerations in research is the role of humans in the process. Baines et al. (2013) identified 18 ethical principles to guide researchers in this respect. The 18 principles are consistent with the provision of the Belmont Report and include respect for participants, informed consent, the full disclosure of funding sources, and avoiding harm to participants (Baines et al., 2013). Underlying several of these principles is the requirement for participants to provide consent for their role in the study through a consent form. Spike (2017) referred to this informed consent process as the single most important concept in assessing decision-making capacity of study participants. Researchers use it to inform potential participants of the study, the reason for their selection, what the researcher expects of them, what they should expect from their participation, and the voluntary nature of their role in the study (Kadam, 2017). I used the informed consent process in this study because it is a way to provide essential information to potential participants and enable them to make a rational and informed decision about participating in the study.

All potential participants could elect to withdraw from their names from consideration for the study before I notified them of their selection to participate. They could also withdraw their names from participation any other time after that. The consent form specified that participants can elect to terminate their participation all the way up to, and during the interview. Participants could communicate their intent to withdraw verbally, or via any form of written communications, including email or text messages.

The consent form included the IRB approval number 05-30-18-0482061 and the expiration date that I obtained from the University.

As part of the informed consent process, I set the expectation of complete confidentiality for the participants, the site, organization, or specific location. I extended this identity scrubbing activity to any other data I collected from the site. I only used the formal identities of individuals for purposes of scheduling interviews. I did not directly associate their identities with the information or data they provided (including during transcription of the recordings from the interviews). Upon completion of the interviews, I replaced any names with codes such as P1 and P2 and ensured that names did not appear in any subsequent write-up from the study. I will store all the data I collected from the interviews in an online password protected data storage vault for five years (as required by the university). There was no payment in cash or kind for participation in the study.

Data Collection Instruments

In this study, I was the primary instrument for data collection, and in this capacity, it was important to plan the interview and develop an appropriate interview guide to ensure that the participants were comfortable in their interaction with me. The preparation was important to be successful in collecting the participant's detailed accounts (Doody & Noonan, 2013). The first step in the planning and preparation process was determining the type of interview to conduct. There is a range of formats for conducting interviews including structured, semi-structured, and unstructured. The most common of this is a semistructured interview. In this type of interview, the interviewer uses predetermined questions as a basis for seeking clarifications from the participants but also has the flexibility to explore new paths that may emerge during the process (Doody & Noonan, 2013). I used semistructured interviews to collect data from the participants because it was the best fit for the type of exploratory data that I needed to collect for the study.

The selection of questions for the interview was also an important aspect of the data collection activity. In this qualitative case study, I used open-ended questions, clear, neutral, and based on knowledge and experience with the subject, as recommended by Doody and Noonan (2013) and Rowley (2014). I employed the use of an interview guide, consisting of a core central question and other associated questions related to the central question. The interview guide is a tool to ensure that the researcher achieves the optimum use of the interviewee's time during the process (Jamshed, 2014). A copy of the interview protocol I used is included in this document as Appendix A. Before the start of the interviews, I reviewed the process with the participants to make sure they were clear with what would happen and notified them that I planned to record the session for later transcription. Rimando et al. (2015) recommend this record and transcribe approach as a means to ensure that the data researchers capture is accurate and complete. I scheduled the interviews for 45 to 60 minutes and took handwritten notes during the process to supplement the audio recordings. In addition to the interviews, I also asked interviewees for access to any non-proprietary documents that may shed more light on their experiences and knowledge on the subject in the specific organizational setting.

The plan was to collect data initially from five participants from the site, but the main goal was to reach data saturation to enhance the reliability and validity of the data. For this reason, I continued the interview with a sixth participant, at which point I was confident of data saturation. Fusch and Ness (2015) recommended this approach to continue with interviews until data saturation. In addition to data saturation, and based on a recommendation by Birt, Scott, Cavers, Campbell, and Walter (2016), I also performed member checking with the participants as another way to enhance validity. To do this, I provided the transcripts to the participants for review as well as the output of the my initial analysis of their responses from the interviews. These actions were to ensure the accuracy of the data I collected and eliminate chances of misrepresentation of the data.

Data Collection Technique

One effective way for qualitative researchers to collect information for studies is through the interview process. Researchers can use interviews to gather descriptions of experiences from interviewees and their interpretations of those experiences (Anyan, 2013). Furthermore, Jamshed (2014) described interviews as the most common format for data collection in qualitative research and recommended in-depth semistructured interviews as most effective with individual participants in a one on one situation. One disadvantage of in-depth personal interviews compared to other data collection techniques (such as field interviews and surveys) is the relative time it takes to perform the interviews, collect, and transcribe the data (Carter, Bryant-Lukosius, DiCenso, Blythe, & Neville, 2014). One advantage of the in-depth interview is that it is an effective way to solicit additional information, generate new ideas with those participants more likely to discuss sensitive topics and stimulate discussion in that interview setting (Carter et al., 2014).

Before conducting the interviews, I worked with the administration at the site to ensure that we have a convenient, comfortable, and distraction-free location for the interviews. The interviews I scheduled began with a welcome statement, a summary of the objective of the interview, and validation that the participants were ready to start. I reminded them again before the start of the interviews about recording the session and the reason for doing so before commencing with the interview. I used a digital recording device to record the interviews. The digital recording device is a dedicated recording device unlike the cell phone, but with less likelihood of interruption during the interview by an incoming call, text message or an alert typical of multi-function cellular devices. The specific recording device I used for the interviews was Philips DVT2da710 Voice Tracer (with Built-in USB and Dragon Speech Recognition Software). I also informed the participants of their option to terminate the interview at any point in the process if they feel uncomfortable.

I used the interview guide with the questions, starting with some general questions that the participants can easily answer, and proceeding on to the main questions. Rimando et al. (2015), stated that most participants are eager to provide information but may need guidance on the level of details the interviewer is seeking. The questions I used were open-ended and designed to encourage the participants to provide in-depth information on their experiences with the subject. Also, I encouraged the participants to elaborate on their responses and share their insights as well as any documentation they have that can help shed light on the strategies and processes used at the site to manage IT operational costs in cloud computing. After I completed the interview with each participant, I thanked them for their time and explained to them that once I transcribe the recordings, I will send them a copy to review to ensure all responses were correctly documented. Member checking and transcript reviews are processes researchers use to ensure information collected during the sessions are accurate and reflective of the interviewee's intent and responses (Harvey, 2015). I used transcript review in this case to improve the reliability of the data I collected from the site.

Data Organization Technique

A commitment I made to the participants on the consent form was to ensure their complete anonymity. Lancaster (2017) identified anonymity as an important core principle of research ethics, and a mechanism by which the researcher can achieve privacy and confidentiality. To meet the anonymity requirements for the study, I replaced participants names with an alphanumeric code such as P1 and P2 and associated these naming codes to the data I collected from them. I tagged any other data provided by the participants with the same alphanumeric code assigned to the participants.

In addition to anonymity, there is also a need to effectively organize the data researchers collect for studies (Tsang, 2014). To organize the data, I initially saved all the transcripts from the interviews in an electronic vault I established specifically for the study and created folders to arrange the data in chronological order with the data collection activity, as recommended by Cronin (2014). The data included all the audio files from the interviews and the associated transcription files. This systematic multiple data sources handling approach is consistent with the recommendations of Houghton, Murphy, Shaw, & Casey (2015), who highlighted its importance to the quality of a study. Yost et al. (2014) suggested some light editing may be necessary to remove fillers such as *ums* and *ahs* and any ambient sounds, which I had to do to finalize the data for analysis. This collation of data facilitated easy uploading and segregation into the NVivo data analysis tool. Consistent with the requirements of the University, and as I specified on the consent form, I stored all raw data I collected for the study in an online password protected data storage vault and will preserve them there for five years and destroy them after that period.

Data Analysis

Because of the nature of case study research, researchers employ triangulation methods for data analysis to address the need for the data to converge, and to enable understanding of the whole case, recognizing that the data sources are not independent of one another (Choy, 2014). The four triangulation techniques researchers use for data analysis are – method triangulation, investigator triangulation, theory triangulation, and data source triangulation (Carter et al., 2014). Methodical triangulation involves the use of multiple research methods to collect data, and investigator triangulation the use of multiple people in the data gathering process (Wilson, 2014). Theory triangulation involves the use of different theories to for data analysis and interpretation, and data source triangulation the collection of data from different types of people to gain different insights into the case (Carter et al., 2014).

In this study, I used both method and data source triangulation for data analysis. I collected the data through semistructured interviews with multiple participants from the site. I also requested to review data from other sources such as process documentation and operational records. After validating all interview data with the participants in a follow-up review meeting, I organized and uploaded the data into NVivo (version 11) to commence the process of analyzing the data. I selected NVivo over other similar qualitative data analysis software like Atlas.ti because I have taken online tutorials in the use of the program, and I am more familiar with it than the others that serve the same purpose. Other researchers have successfully used NVivo in cloud computing studies including studies by Alharthi, Hussain et al (2017), Bauer and Bellamy (2017), and Hachicha and Mezghani (2018). It was a good fit for the study.

Organizing and systematic handling of data from multiple sources are important for qualitative case study researcher (Houghton et al., 2015). The source data is textual, and Ryan & Bernard (2003) stated that analyzing this type of data typically involves four steps – discovering themes, honing down the themes to a few main themes, developing a hierarchy of themes, and linking the themes into the conceptual framework for the study. Morse (2015) summarized this succinctly, identifying the key stages as comprehending, synthesizing, theorizing, and re-contextualizing. In this study, I used the simplified three-step approach suggested by Smith and Firth (2011). The steps consist of (a) familiarization with the data through reviews, developing initial categories and themes and assignment of data to the codes and categories, (b) summarizing, synthesizing, and refining the themes to form a complete view, and (c) developing associations within concepts and themes that reflect the original data, and aligns with the conceptual framework (Smith & Firth, 2011). During the data analysis phase, I continued review of the current literature for opportunities to correlate themes with new or updated findings from peer-reviewed publications.

Reliability and Validity

The application of concepts of reliability and validity to evaluate studies is an integral part of quantitative research methods, where researchers employ the use of statistical methods to determine the extent to which a study can be qualified as legitimate (Darawsheh, 2014). In this section, I address the concepts of reliability and validity and how I intend to mitigate them in this study.

Reliability

The reliability of a study is a measure of the consistency and rigor of the research design, research methods, and other processes used by the researcher for the study, accounting for biases (Noble & Smith, 2015). Golafshani (2003) suggested that a simple way to think about reliability is to associate it with dependability and recommended *audit trails* to enhance the dependability of a qualitative study. Audit trails refer to the preservation of all data collected, the related record of procedures used to collect them, and any notes related to the collection and analysis of the data. Birt et al. (2016) recommended member checking as an effective process to use to ensure the accuracy of the source data. Member checking is a review process with the participants to ensure the veracity of the data attributed to them (Kornbluh, 2015). In this study, I performed both transcript review and member checking by providing a copy of the interview transcripts to the participants for review and early detection of any discrepancies and following that up with another review of the interpretation of their responses with them.

Validity

The validity of a study is a measure of the extent to which the findings from the study reflect the data collected for the study (Noble & Smith, 2015). Threats to a study's credibility can happen at any stage of the study process but typically exist at three key stages – research design and data collection, data analysis, and data interpretation (Darawsheh, 2014). Unlike quantitative researchers, qualitative researchers do not employ the use of statistical methods to attest to the credibility of their studies, but there is still the need to establish minimal threats exist to both external and internal legitimacy for the study (Darawsheh, 2014). To accomplish this validity for qualitative researchers, Noble and Smith (2015) recommended the use of concepts such as truth value, consistency, confirmability, and applicability (transferability) as an equivalent set of concepts for evaluation of qualitative studies. Their recommendation is consistent with that of Walby, Marshall, and Rossman (2015), who also advocated for the replacement of validity and reliability with credibility, dependability, and transferability.

To realize these concepts of validity in practice and establish study credibility, various qualitative researchers have suggested different techniques for addressing the threats of legitimacy in a study (Darawsheh, 2014; Kornbluh, 2015; Leung, 2015). Most of these recommendations reflect three core concepts – data saturation (Fusch & Ness, 2015), data triangulation (Hussein, 2015), and member checking, (Kornbluh, 2015). Transferability of findings is another area that researchers have to address in a study.

Transferability refers to the consideration of the scope and extent to which any findings by the researcher in a study applies to other similar settings or contexts (Noble & Smith, 2015). Transferability has direct implications for recommendations for future research, and the potential application of the recommendations and findings in practice.

Researchers bring their own biases into studies, and these biases can influence all aspects of the studies (Noble & Smith, 2015). Confirmability is the accounting for the researcher's philosophical position, perspectives, and experiences; They likened the concept of confirmability to *truth value,* which is a general acceptance of the existence of multiple realities, and the relative effect of the researchers bias on the study. Closely associated with confirmability is the concept of consistency, the trustworthiness of methods, and data collection trail that can lead another researcher using the same data to reach a comparable finding. In various sections of this document, I have provided validation of the design method and methodology, and specified my role as a researcher in this study, outlining my experiences with the subject matter, and relation to site and participants in the study.

Researchers use data saturation technique to ensure the completeness of the data they collect. Researchers attain this in a study at the point where the addition of new data does not lead to the discovery of any new information or themes (Houghton et al., 2015). Data saturation reflects the sufficiency of data to support the findings and has a direct bearing on the reproducibility of the study and thus reliability. Achieving data saturation is a key requirement for validity and one good way to ensure the credibility, transferability, and confirmability of the findings (Yin, 2017). Data triangulation on the other hand, is a way to collaborate data the researcher collects from one source with data from other sources and in the process, verifying the accuracy of the data (Hussein, 2015). In this study, I continued the interview process (beyond the five participants that I originally targeted at the site) until I had indication that I had attained data saturation. I also solicited interviewees for any documentation they have that supported their responses from the interviews, and which may shed more light on the strategies and processes they used to manage IT operational costs, even though I did not receive any such documents. During the data analysis phase, I also reviewed emerging themes with the participants and finally followed that up with another review of the findings.

Transition and Summary

The purpose of this qualitative multiple case study is to explore strategies and processes IT managers use for managing infrastructure operations costs in cloud computing environments. In this section of the study, I outlined the research approach I used for the study and reviewed my role as a researcher, identified type of site I targeted for the study, and the participants I intend to solicit for the study. I also specified the research method and design, data collection and organization techniques, and approach to data analysis. I followed ethical research standards for the study and took the necessary steps to improve reliability and validity at all stages of the research process. In the next section, I present the findings from the study, and the implication for professional practice and social change. I also stated any recommendations for action and future research and completed the section with my reflection and conclusions of the study.

Section 3: Application to Professional Practice and Implications for Change

Introduction

The purpose of this qualitative multiple case study was to explore strategies and processes that IT managers use for managing infrastructure operations costs in cloud computing environments. In this section, I present my findings from the study and discuss applications of the findings to professional practice and implications for social change. I also provide recommendations for action and further research and close the section with reflections and a summary of the study.

The interviews I performed to collect the data occurred over a 2-week period and involved seven participants. The interviews took place in a quiet and comfortable location onsite at participants' organizations, where they were free to respond to seven semistructured questions (Appendix B) designed to elicit information on strategies and processes that IT managers use to manage costs in cloud computing. Based on the main research question and the analysis of the data I collected, I identified 12 emergent themes related to strategies and processes that managers use to manage cloud computing costs.

The conceptual framework that I used for the study was the STS framework for the study. The STS framework is used by researchers to analyze complex systems by viewing them through the decomposition of two primary subsystems. The two subsystems are (a) the technology subsystem, consisting of equipment, machines, tools, the technical environment, and related processes; and (b) the social subsystem, consisting of individuals, teams, management and governance structures (Carayon et al., 2015; Cherns, 1987). The key tenet of the theory is that it is only when managers optimize the effectiveness of both of these subsystems that they can achieve technological effectiveness and improve the lives of employees. Based on my reviews of the data I collected from the interviews, the participants' responses were consistent with strategies and processes related to the two subsystems. The findings presented in the next section are grouped into technical subsystem strategies, technical subsystem processes, social subsystem strategies, and social subsystem processes.

Presentation of the Findings

The overarching research question for the study involved what strategies and processes IT managers use to manage IT infrastructure operations costs in cloud computing. I identified 12 emergent themes from the analysis of the data from the interviews and categorized them into strategies and processes for each of the subsystems. Figure 1 is a summary of how the study's objective relates to the conceptual framework and how the themes, in turn, align with the framework. In the following sections, I present a description of each of the themes I identified from the study and how the findings relate to or deviate from the academic literature.

Technical Subsystem Strategies

Based on STS design, the technical subsystem in an organization consists of the technology subsystem—equipment, systems, machines, tools, technical environment, and related processes (Carayon et al., 2015; Cherns, 1987). Table 1 shows the themes that I identified for the technical subsystem from the data.

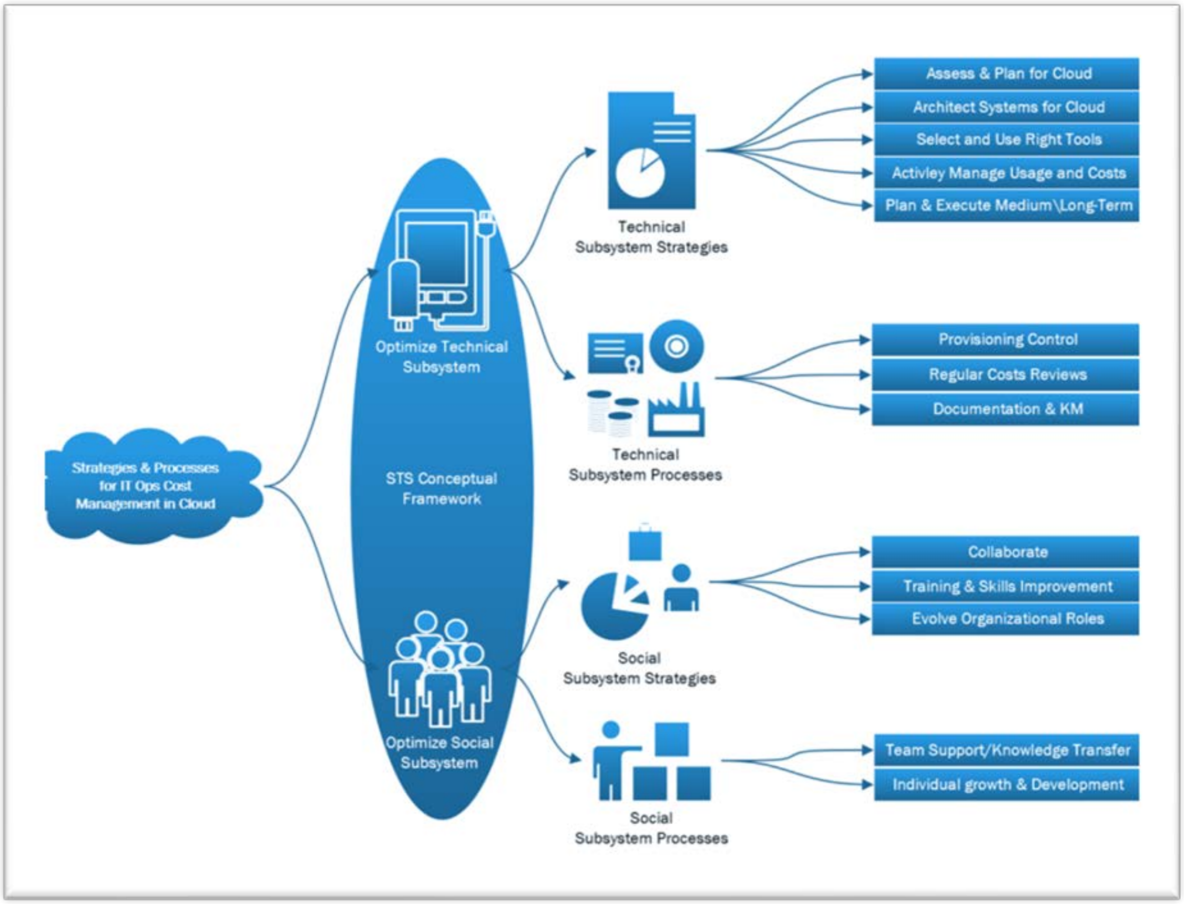

Figure 1. Summary of themes by STS subsystems.

Table 1

Technical Subsystem Strategy—Themes

Code	Number of participants	%
Assess and plan for cloud environment	4	75
Architect systems for cloud services	5	83
Select and use right tools	6	100
Actively manage resources and costs	6	100
Establish and execute medium- to long-term plans	6	100

The next section contains a discussion of the emergent themes from the study for the technical subsystem.

STS Technical Subsystem Strategy Theme 1: Assess and plan for cloud. Even though the study was on the strategies and processes for cost management for organizations with IT infrastructure in a cloud computing environment, it was evident from the data that efforts to manage costs start before the transformation of organizational assets to the cloud computing environment. Gholami, Daneshgar, Beydoun, and Rabhi (2017) noted that inadequate understanding of and preparedness for cloud computing migration are reasons for most of the failures of migrations to meet organizational goals. The participants identified the starting point as the assessment and planning for transformation to the cloud. Table 2 is a summary of the subthemes that I identified from the study for this strategy to assess, plan, and prepare for the cloud computing environment.

Table 2

Technical Subsystem Strategy Theme 1: Assess and Plan for Cloud Environment

Code	Number of participants	%
Understand the motivation for transformation	2	33
Understand current costs and baselines	3	50
Perform assessment to understand application profiles and environment	4	66

The IT manager's motivation for adopting cloud computing is important for setting the stage for the subsequent planning and transformation activities that enable effective cost management posttransformation. Yeboah-Boateng and Essandoh (2014) emphasized the importance of this when they stated that because of the complexity of the cloud environment, IT leaders must factor in planning and preparation as elements of the strategy to transform. Fahmideh and Beydoun (2018) added that the effort to transform to cloud may not be hazard-free and demands a proper understanding of requirements and risks involved prior to taking any actions. P2 put it succinctly, stating that as far as strategy goes for organization XYZ, "we always look at the purpose of what they are trying to do within the cloud," and P7 referred to this as getting an understanding of the overall business requirements for cloud adoption. This understanding of the overall business motivation and requirements for cloud adoption is a necessary framework to help set the stage for the subsequent change activities necessary to manage costs in a cloud environment. For most organizations, cloud computing is a paradigm shift from the way in which IT resources are acquired and used at their on-premise data centers, and this shift requires planning and preparation.

The preparation for cloud adoption goes beyond an understanding of the motivation for adoption. If cost savings is one of the primary factors, then it is necessary to have some basic understanding of current IT costs to use as a baseline. P7 stated that as part of the preparation to transform, "we work with the different groups to determine their budgets and application profiles." P5 added that as part of the preparation process, they identify what the specific budgets are, what they are currently spending at their data centers, and what baseline settings should be for tracking and reporting purposes. The baselines that the managers establish from this process form the basis for configurations of reports and subsequent tracking and reporting that constitute an integral part of ongoing cost management.

Another aspect of the preparation activity that emerged from the data is an understanding of application profiles. Understanding application profiles at a high level involves an understanding of application architecture, resource requirements, and performance specifications. Weingarten et al. (2015) defined application profiles as a way to qualify the use of computing resources by an application and its expected future requirements. Knowledge of application profiles is critical for the resource provisioning of the systems, ensuring that sufficient resources are allocated to meet performance requirements and that the resources can be optimized without degradation of the underlying business needs (Weingärtner et al., 2015). The need for application profiling was pointed out by P3, who stated that “from a strategy perspective, we always like to do that type of analysis first to make sure that when we bring it to the cloud, it actually is a solution that's going to meet the needs from a cost perspective.” P3 also emphasized that in general, “really looking at and understanding the analytics or the performance of the application is really important.” P5 identified the best approach to understanding these application profiles as performing an assessment using tools that they installed within the data centers and that facilitate a “very clear understanding of what your usage and capacity spend is going to be (monthly) before you even move to the cloud.” Even though Weingärtner et al. (2015) noted that this application profile information has historically been difficult for organizations to obtain, advances in tools for this purpose are making it easier for managers to do this, and based on the data analysis, should be an essential part of preparation to migrate to cloud computing environments.

STS Technical Subsystem Strategy Theme 2: Architect for cloud services. With an understanding of the business objectives and application profiles as well as a basic understanding of costs, the next logical theme that the participants identified involved the need to architect systems for cloud services. Table 3 shows a summary of the subthemes for this strategy and the number (and percentages) of participants who addressed this in their responses.

Table 3

Technical Subsystem Strategy Theme 2: Architect Systems for Cloud Services

Code	Number of participants	%
Align application profiles to best cloud services	3	50
Tag resources comprehensively	4	66
Use reserved instances (RI) and autoscale	5	83

Architecting systems for the cloud in this context implies identifying application components, understanding how those components work together, and mapping them to equivalent cloud services (Zhang, 2018). These components include types of servers, containers, storage, network access service models, and deployment models.

Cloud computing architecture consists of three layers—IaaS, PaaS, and SaaS—with each layer serving application requests and supporting upper layer functions (Weintraub & Cohen, 2015). Within each layer, cloud service providers present a myriad of services to fit a specific application architecture and business requirements. The choice of specific use of one service over another can have a significant effect on application performance as well as costs (Weintraub & Cohen, 2015). P3 stated that "architecting the cloud environment is an essential part of any strategy." P7 elaborated on this by stating that they perform an analysis of existing application and infrastructure and then "right size it to where there are some cost savings there" before starting the migration activity. In effect, this involves "translating what is in the current infrastructure, understanding the resource usage" before moving the assets to the cloud environment (P7). Cloud transformations may sometimes entail migrating systems from a server infrastructure to a PaaS environment instead of one-on-one mapping to IaaS equivalent. Managers must carefully consider all of these options must before migrating the workloads to the cloud.

Key objectives of architecting applications for the cloud services include understanding which cloud services will meet the operational and performance needs of the application and selecting the right set of cloud services (keeping in mind the cost implications of using one cloud service option over another). Evangelinou et al. (2018) recognized the challenges for customers seeking to know application behavior on cloud platforms before selecting services from a performance and cost perspective. The variety of service options, combined with consumption-based pricing models, makes it complicated to find the optimal deployment options. Evangelinou et al., proposed a combined methodology and a set of tools to help organizations whose staff are overwhelmed by the selection process. The methodology they proposed consisted of a performance assessment, a profiler that matches an application computing footprint with cloud services for optimal performance and costs, and a tool to help in identifying minimum deployment costs that reflects the application workload over time (Evangelinou et al., 2018). Application profile information from the application assessment activity is key to rightsizing cloud services for the applications as part of the provisioning process, and prior to migration.

Another aspect of architecting for cloud environments is the concept of resource tagging. Resource tagging is a way for service consumers to mark or add metadata to identify the location of their information assets (Pichan, Lazarescu, & Soh, 2015). P5's reference to this activity as mandatory was consistent with the representation that tagging of cloud resources is a critical element of any cloud governance initiative (RightScale, 2018c). Cloud service providers and third-party vendors provide tools that facilitate the monitoring and reporting of resource usage and aggregate them based on their tags. P3 stated that the tools they employ "give us the ability to tag and develop cost centers and develop different types of cost-bucketing and budgets based on a group or department or cost centers." I addressed the role of budgets and cost centers as an emergent theme in another section of the presentation of findings, but this activity is the basis for a comprehensive understanding of which resources are consumed by which applications, and by which teams, groups, or departments.

The tagging activity should be comprehensive and conducted in enough detail to enable the tracking of resource usage at all levels of the infrastructure. Resource tagging is also the basis for the allocation of service usage and associated costs to teams, groups, departments, and organizational units (P3). The best practice for tagging resources for cloud governance and cost management entails (a) establishing a global policy for doing it, (b) understanding what needs to be tracked and at what level of detail, (c) maintaining consistency with tag names, (d) tagging resources on launch or immediately thereafter, and (e) setting up automated "tag-checking" to ensure that no resources are left untagged (RightScale, 2018c). Tagging resources is a necessary part of any strategy to track and report resource usage across the IT infrastructure and to hold users accountable for their cloud usage.

One of the values of understanding application profiles in terms of the cost management process is that it gives cloud migration architects the source information they need to select the optimal cloud service(s) to meet the technical and performance requirements for the application. P7 described this process, stating, "we do an analysis on what they have, rightsize it to where there are some cost savings there, and then we start to migrate their data sets (their virtual machines, their applications or whatever) into this infrastructure." In addition to rightsizing resources to fit performance requirements, the two other mechanisms that the managers used to effectively manage costs were reserved instances (RI) and autoscaling. P2 identified the value in doing this, stating that they "buy reserved instances which, if you look at the chart, there was about a 45% savings if you take into account reserved instances that we purchased." P4 pointed this out as well, stating that the cost savings opportunities on "Amazon, in particular, leans towards cost control by either purchasing reserved instances, or using the spot market." Srirama and Ostovar (2018) agreed to this approach to cost management, recommended that autoscaling systems in the cloud address both *when* and *how* questions related to scaling, and presented a resource provisioning policy that can assist managers in finding the most cost-optimal setup of instances in the cloud to fulfill application performance requirements. Liu et al. (2018) also identified the value in using RIs as providing elastic and cost-effective infrastructure to support high-performance computing applications and proposed an instance reservation-based cloud service framework to help managers overcome cost-minimizing reservation decision issues. Researchers such as Kandpal and Patel (2018) and Goldsztajn, Ferragut, Paganini, and Jonckheere (2018) have proposed frameworks to aid IT managers in determining when (and how) to use RI to manage operational costs. In the end, the selection of which mechanism to use will depend on the application profiles and the performance specifications to meet business requirements, but to what extent it is feasible, RI and autoscaling must be an integral part of the strategy.

All the managers agreed that the need for good preparation before the transformation to cloud computing is very important. The challenge for managers adopting cloud computing is the pace at which transformations are occurring might imply shorter planning timeframes which may have implications for overall organizational effectiveness in realizing the full value of the technology (Garrison et al., 2015). Organizations that invest in these preparatory activities to understand their objectives, business requirements, costs and budgets, their applications profiles, and correctly architect their applications and services to capitalize on the correct service offerings, and tag all of their cloud resources will be better positioned to manage their applications performance and costs in cloud computing. For managers who have cloud-native environments, or who have already transformed to the cloud but are challenged with cost controls, it may be necessary to invest in these preparatory activities as a way to get back to basics and set the stage for managing their costs.

STS Technical Subsystem Strategy Theme 3: Select and use the right tools. The process to transform infrastructure to cost-effective systems is aided in part by tools provided by the cloud service providers or third-party vendors. The use of tools in support of cost savings was an underlying theme that consistently came up in all the interviews. Table 4 below shows the frequency and types of sub-themes from the data analysis.

Table 4

Technical Subsystem Strategy Theme 3: Select and Use the Right Tools

Code	Number of participants	%
Select right toolsets	6	100
Monitor performance and resource utilization	6	100
Automate reports and auto-orchestrate changes	3	50

Effective IT operations cost management is only possible with the monitoring of resource utilization across the whole IT infrastructure. Fatema, Emeakaroha, Healy, Morrison, & Lynn (2014), stated that the complicated nature of cloud environments can result in a multi-level and intricate layered structure and that the use of monitoring tools is the way to understand how to optimally manage them. Monitoring tools and techniques play an important role in facilitating the collection of information managers require to make informed decisions (Fatema et al.). P6 made this point when he stated that “without those tools, we won't be able to administer such an environment with such a small staff; without the oversight, there will just be too much that would be missed.” P6 went on to emphasize that the tools they use are “really our primary mechanisms for optimizing against cost.” The cloud service vendors provide some tools to use to manage the infrastructure, but they also make available a set of application interfaces to allow third-party vendors and developer to access and report on service performance and resource usage. P5 identified this capability for the right tools to have “extensible reporting” that integrates into the cloud environment as beneficial. The participants listed at least three tools they use at the site to manage costs.

Because of the importance of the tools, the participants also stressed the need to use best of breed tools to ensure a complete insight into cloud system usage. P4 made this point when he stated that “a sort of the best of breed of the... we're going to pull all your information and then give you a whole bunch of recommendations.” P2 also qualified the tools they use at the site as the “who of who” of tools. Fatema et al. (2014) identified some of the capability attributes that an ideal monitoring tool should possess to serve the objectives in the cloud operations as (a) accounting and billing, (b) SLA management, (c) resource and service provisioning, (d) capacity planning, (d) configuration planning, (e) security, and (f) privacy assurance. The tools the participants selected to use at the site reflected most of these capabilities, but they also highlighted the importance of product support and training from the vendors as important.

The requirements of the tools that were listed by the participants included the need to monitor and report on both application performance and resource utilization (P2), translate resource usage to costs (P7), provide pace and trend reporting (P2), and threshold alerting (P1; P7). These features are consistent with the findings of Fatema et al. (2014). In addition, the participants identified the need for the tools to incorporate best practices for cloud computing cost savings (P3; P4), as well provide proactive notification of cost reduction opportunities in the infrastructure (P2), and auto orchestration of cost-saving recommendations to reduce the time and effort it takes to capitalize on the opportunities for cost savings (P3). To obtain all these capabilities, the managers use a combination of tools to monitor, report, review, and make decisions on what actions to take to manage operational costs in the cloud environment.

The regular updates to cloud service options, the introduction of new services, and the variations in service pricing mean the tool vendors need to regularly update them to keep pace with the changes in cloud provider offerings. P2 pointed out the importance of this, stating that it is necessary to “make sure everything's up-to-date within the system.” P5 also mentioned the need for a collaboration and feedback mechanism with the tool vendors: “we have a very close relationship with the developers of those applications so they take our feedback very frankly and they will literally iterate (if we say this is missing, or we want this) and the next release we would often see that stuff added.” The close relationship, in this case, may be because of the strategic nature of the site to the tool vendor, but the ability to take feedback from valued customers to improve products promptly was important to the participants and a good attribute of the tools and vendors.

The best-of-breed tools also have built-in functionality to also analyze a client IT infrastructure and determine if managers have optimally provisioned application’s resources for performance and cost. The determination of optimal provisioning is made in part by incorporating in the tool, best practices for cost savings in the cloud computing environment. The incorporation of these best practices features makes it easier for managers to quickly understand issues with their resource allocations, and without much research, understand what corrective actions to take, and or where opportunities exist to fine-tune the resource provisioning to manage costs.

STS Technical Subsystem Strategy Theme 4: Actively manage resources and costs. The role of tools in the cost management strategy is to facilitate insight and visibility into resource utilization and cost in the IT environment, and this was one of the key themes that emerged from the interviews. To ensure visibility into resource consumption entails an understanding of the cost contributors by application, the monitoring, reporting, and regular review of the reports with stakeholders to correct any discrepancies that could adversely affect operations costs. Table 5 shows a summary of the sub-themes related to this strategy and the number of participants who addressed each in the interviews.

Table 5

Technical Subsystem Strategy Theme 4: Actively Manage Resource Usage and Costs

Code	Number of participants	%
Establish budgets and baselines	5	86
Create reports, dashboards, and alerts	3	50
Review performance and promptly address deviations (automate)	6	100

For cloud environments, in particular, Fatema et al. (2014) pointed out that deciding what to monitor is important because usage-based billing and the potential for elastic scaling are impossible to implement in the absence of the relevant metrics. P4 described their approach to this by stating that they “extract from the resource level stuff where you need to know what the costs are.” Extracting and organizing the data for reviews is aided by tagging applications resources to specific individuals or logical organizational units (groups, departments, or business units) to allow allocation of the related resource consumption to the right business units to hold them accountable for related costs.

The use of budgets is effective in organizations as a means to control costs. Bedford (2015) stated that organizations undergoing strategic changes have higher performance when they use budgets. P3 stated that they “work with the different groups to determine their budgets” as part of the strategic process for cost management. Managers can use the best of breed tools to do real-time tracking of resource costs against budgets so that stakeholders have early warning of potential cost overruns. P3 pointed this out when he stated that when they “change resources or add things, it will automatically change their projection so that way if they hit or if they get up into the yellow area where they potentially could exceed budget limitations, it will notify the manager.’ This alerting feature is very effective for cost controls, and as an example, P3 added that “if a development group logs into their dashboard within our system, they're going to see their spend on a graph, they're going to see what it is today, actual and then what the projection is for the remainder.” Having this feature available to managers “makes them more aware of where they're at throughout the month on what their spend is” (P3). The real-time display of cloud spend by team and business unit is an essential part of the process to make consumption information readily accessible to team members, putting the onus of control in the hands of the responsible managers.

Budgets are only one means of cost management. Another way to proactively manage costs is through provisioning controls. Managers use this tool feature to limit service provisioning capabilities of team members once their budget allocation reaches a predetermined threshold. P3 described this feature stating “we can also put in hard limits to say that once you get to a point where you're going to exceed your budget, it will no longer allow you to add any additional resources to your environment.” P5 added that what worked for them is to “limit the usage by department, limit who can do things, and be mindful of that as an organization.” Breitenbücher et al. (2014), introduced a policy-aware provisioning concept that enables defining non-functional system requirements (such as security requirements) on the execution of provisioning tasks using automated policies. The tools the managers use enable the application of provisioning constraints based on cost policies. When IT managers employ the right tools in a cloud environment, these types of controls can easily be instituted to manage operational costs.

The process to actively manage operations costs on an ongoing basis is not complete without the establishment of a forum for reviewing and analyzing the reports the managers collect from the tools. P2 specified that they had “weekly meetings, and now our bi-weekly meetings, we do go and look at the adjustments that need to be made, we will look back at the previous month to see how we were tracking from a week-to-week and month-to-month basis and then we'll compare that to our current trend.” P4, on the other hand stressed the value of these regular meetings stating that “even if you haven't gotten to fix the things that you need to fix or make the adjustments you need to make, just the fact that you're doing that regular reviews, that you have a cadence on that means that it stays up in the awareness and don't lose track of it". P6 noted that "we now get regular weekly updates on any idle or unused resources that allows us to watch for those kinds of isolated instances" that could impact costs. The key objective of this strategy is to establish a forum where managers are held accountable for any deviations in planned resource usage for their applications, as well as to review any requests for provisioning or allocation of new resources.

STS Technical Subsystem Strategy Theme 5: Establish and execute medium- to long-term plans. The effort by IT managers to control IT operational costs is not a one-time effort but an ongoing activity that requires long-term planning and that must integrate into organizational business objectives. Table 6 shows the subthemes for the strategy and the number of managers who addressed them during the interviews.

Table 6

Technology Subsystem Strategy Theme 5: Establish and Execute Medium to Long-Term Plans

Code	Number of participants	%
Formulate a long-term perspective	2	33
Regularly assess changes in infrastructure and cloud provider offerings	6	100
Prioritize activities to capitalize on opportunities	4	66

The managers identified three aspects of this strategy: (a) keeping a perspective of the whole IT infrastructure, (b) identifying opportunities to capitalize on technological changes that would improve system performance and/or reduce operational costs, and (c) creating, prioritizing, and executing projects to capitalize on the opportunities. P2 put it best when he stated that they used the cross-functional team of managers to make “a list on the board of all the issues we’re having (whether it be cost-related, difficulty operating things, or just keeping track of our infrastructure as a whole), and broke that down into the top 10 things we want to focus on for the year.” The focus, P2 added, was the “things that impacted our business the most.” One such architectural change implemented by the managers under this strategy was the introduction of containerization technology (kubernetes) to the environment. P6 stated that “we are moving toward putting everybody in the containerized environment.” The execution of this is an effort that takes careful planning and execution by the IT managers to ensure that business processes are not impacted by the platform change, and it may involve investments that may increase IT costs in the short term but will ultimately lead to long-term cost savings.

The incorporation of medium to long-term organizational objectives into the plan is important because the managers view effective cost management as not only about cost reduction, but also about managing application performance and operational continuity to meet business requirements. Infrastructure changes, even if they save money in the long term, need to be made with business needs in mind. P6 was clear on this, stating that making changes “becomes a very serious critical point for us because it’s not simply just saying okay where we are, we going to try new things; You can cause all sorts of issues by just being open-ended like that.” P6 added that one “must have that kind of dual perspective to be able to come up with the right solution, and currently when we look at it, we have to weigh utilization metrics (as an example) against potential risks, so some of our infrastructures are not, I would say not fully optimized." P2 made the same point, adding that "it's not always necessary to have the newest and biggest thing, even though it's good to understand how it will impact you." Having a long-term perspective on how to improve the IT infrastructure as whole to has to be done with due considerations to the organizational business objectives.

Technical Subsystem Processes

Based on my analysis of the data collected from the site, the managers identified three processes that pertained to the technical subsystem. Table 7 shows these themes and the number of participants who discussed them during the interviews. I discuss the processes and the subthemes associated with these processes in the next section.

Table 7

Technical Subsystem Processes—Themes

Code	Number of participants	%
Provisioning controls	5	83
Resource consumption and costs reviews	3	50
Documentation and knowledge management	3	50

STS Technical Subsystem Process Theme 1: Provisioning controls. One of the key contributors to cost overruns in cloud environments is server sprawl. P3 pointed this out when he stated that "the important thing, if a customer or a company is taking on cloud as its primary facilitation for IT services, is really looking and managing servers and service sprawl." The server sprawl problem is a situation in data centers in general, and cloud environments in particular, in which multiple underutilized servers or devices take up space and consume more resources than is justifiable from a workload or performance perspective (Shirvani & Ghojoghi, 2018). Researchers have come up with algorithms and frameworks to address this problem of managing or controlling server sprawl (Naeen, Zeinali, & Haghighat, 2018; Shirvani & Ghojoghi, 2018). The one effective approach to the sprawl problem in their environment involves the process of provisioning controls.

The managers use a provisioning control process to qualify and approve what new resources that are provisioned in the IT infrastructure, as well as examine if currently provisioned resources are being efficiently utilized. P2 described the provisioning control process as they use at the site as “if we do stand something up on a new instance for a client, there is an approval process, we have a form (basically an internal form) that just says -this is what we need, and this is how much will this cost; its's basically an estimation and review before you stand it up” process. P4 saw it as a “process to review whether consumption is warranted and what modifications are required to be made.” The managers also use the process to document cloud provisioning requirements for the application including what hours the instances will be used during the day, and any dates when they can be decommissioned from the environment. P3 gave an example of this saying “if developers work from 8 am to 6 PM, we say—okay, any development resources that you're going to spin up in here, we're going to automatically start at 5 o'clock, we're going to wait 45 minutes to make sure no other keystrokes or stuff works coming through, we're going to shut those down until tomorrow morning we're just not going to leave them up and running forever.” The provision control process when properly instituted by my managers can be an effective means to manage the introduction of new service items into the environment and the impact to the IT infrastructure and cost implications to the organization.

STS Technology Subsystem Process Theme 2: Resource consumption and cost reviews. To enable the effective control of the operational costs, it is necessary to establish a forum where the managers can review resource consumption and cost reports with all stakeholders. The managers pointed out that the reports are auto-generated and shared with the stakeholders for the preview (P3). Specifically, P3 stated that "we will pull in all of the historical billing data and it will actually show you that data over time, it will show you based on the performance footprints, the services that you're using and where your cost savings can be." To review these reports, P2 added that they have weekly meetings where "we do go and look at the adjustments that need to be made, we will look back at the previous month to see how we were tracking from a week-to-week, month-to-month, basis and then we'll compare that to our current trends." This level of visibility of information makes it possible to enroll all stakeholders in the process of understanding the cost impact of the systems they use, and any changes in performance requirements or the commissioning of new systems.

The managers who attend the meetings are from various functional teams consisting of managers from both IT and the business stakeholders. The use of a cross-functional team is a way for managers to ensure that costs information is visible to all stakeholders, and any decision made on provisioning, upgrading, or downgrading is a team decision (P2). P2 identified the three forums where meetings occur, corresponding to increasing aggregation of IT services. These forums are held weekly, bi-monthly and monthly. The weekly (and sometimes bi-weekly as needed) are to review resource consumption on a team level and determine any deviations from plans. The bi-monthly meetings are to do a review of the state of the infrastructure as a whole, and the monthly meetings to review the status of projects to introduce new services or upgrade to new and more efficient service provider options like the introduction of kubernetes to the infrastructure (P6). All these processes and controls are designed by the managers to ensure transparency of their service consumption and the cost implications, and also to minimize the risk of changes to the environment that may compromise IT services.

STS Technical Subsystem Process Theme 3: Documentation and knowledge management. When organizations migrate to cloud computing environments from an on-premise data center, managers are faced with a dynamic computing environment with a myriad of metered services. Cloud service providers also continue to update existing services and or introduce new ones which are often more efficient and have better pricing than the previous services. IT managers are constantly challenged to deal with the complexity of the cloud environment and services, and the new processes and procedures necessary to manage the dynamic environment (Fatema et al., 2014). The managers identified documentation as a process they use to collect and share information about their environments within the organization. Garousi et al. (2015) also identified technical software documentation as an important practice which when implemented, improves development and maintenance activities in an organization. They went on to state that making decisions on how much of such documentation to produce and the level of details was a major challenge for managers (Garousi et al., 2015). The degree to which this process has been implemented and working at the site varies with participant, with P4 pointing out that “the documentation pieces at this location are not particularly good for the DevOps side, but the developers actually have a better set – they’ve (at least one side of the house) made more of an effort to keep that sort of thing in line with requirements from the business side of things.” I concluded that the documentation process is still an evolving process at the site and has gaps, but is required to ensure people who need information about the state of the infrastructure and application requirements and readily obtain them.

Some issues exist with documentation processes in general. The challenge for managers is with how much documentation to produce on the one hand, and on the other hand dealing with the perception that documentation is an expensive activity, difficult to maintain, and not useful in all circumstances (Garousi et al., 2015). P7 confirmed one of these issues, pointing out that “I'm going through rewriting our operations guides for AWS and Azure because I just wrote them just a year ago, but they're already outdated.” The documentation process has not matured across all teams, but P4 brought up the relative importance of it, stating that “if you told me today that I had the resources to get something written and done, the one thing that I would do would be more documentation.” An important aspect of the process is documenting the basis for decisions made regarding modifications that the managers make to the IT infrastructure (P2; P4). In a study by Cozzolini and Berbegal-Mirabent (2017) on the importance of processes in innovation, they concluded that even the use of best of breed tools are often ineffective if not properly supported with an appropriate knowledge management process. The documentation of deviations from plans, opportunities that came up, and what corrective or proactive actions were taken to capitalize on them is useful information that forms part of the knowledge management system in the organization.

Social Subsystem Strategies

The social subsystem in an STS framework consists of the individuals, teams, management and governance structures in an organization. The framework posits that both the technical and social subsystems need to be optimized in order to derive the most organizational value from the technological adoption. Table 8 is a summary of the themes I identified for the social subsystem.

Table 8

Social Subsystem Strategies—Themes

Code	Number of participants	%
Collaboration	4	66
Employee training and job skills Improvement	6	100
Evolve organizational roles	4	66

I will discuss these themes and their related sub-themes in the next sections.

STS Social Subsystem Strategy Theme 1: Collaboration. Operational cost management is an enterprise-wide effort and an ongoing process. For managers to sustain this effort, a high degree of collaboration is necessary. Collaboration involves the coordination and engagement of all parties responsible for managing and accounting for cloud service usage to actively participate in the process to reduce cost. Driskell, Salas, and Driskell (2018) viewed collaboration from the perspective of team-work and defined it as the engagement of team members and the enactment of teamwork processes to achieve effective team performance. Table 9 shows the key groups that the managers actively collaborate with as part of their strategy.

Table 9

Social Subsystem Strategy Theme 1: Collaboration Groups

Code	Number of participants	%
Internal teams	4	66
Cloud service provider and partners	2	33
Community sources	2	33

To facilitate this collaboration or team-work, transparency into cloud spend data, and a genuine effort to engage managers with the intent to support them with managing their cloud consumption is necessary. P5 identified some of these teams, saying "we have an architectural board… we've got people who are from the operations side, we got people from the development side of the house to the straight up architecture side, and the business analysts' side." The collaborative effort and inclusion lead to quality input that P5 went on to describe as "getting a lot of input from the stakeholders" to rapidly make decisions. The support is in collaborating with stakeholders to perform application rightsizing activities or to capitalize on opportunities to reduce spending while maintaining application performance to meet business needs. In this respect, P3 noted that one of their objectives is to "really understand the impacts as far as growth and sprawl as opposed to cost and containment from that size is really important." The management of growth and server sprawl effectively translated to cost management for most organizations.

The need to collaborate is not only with and between internal teams. The managers indicated that it is important to extend the collaboration strategy to external partners and vendors alike. Among the external groups, the managers regularly collaborate with to keep pace with the rapid changes in the cloud environment are cloud service providers (P2; P3), tool partners and other vendors (P2; P5), consultants (P2), and community partners (P2; P5). With all these teams, groups, or individuals, the goal is the same – identify individuals with common interests, learn from one another, share experiences and techniques on improving efficient utilization of cloud services and operational costs.

The collaboration strategy starts early in the transformation process, when the managers engage responsible parties to understand (a) applications and other IT assets, (b) team structures for reporting purposes, and (c) current costs and budgets. It continues on with the development of dashboards and custom reports and alerts to help them understand their services usage. Finally, the active participation in regular meetings to discuss the fine-tuning of resource allocation, and or discuss and implement opportunities for service improvement or cost reduction.

STS Social Subsystem Strategy Theme 2: Staff training and skill development. With the understandable complexity of cloud computing environments, and the rapid growth in cloud computing. P7 articulated his direct experience with this, stating “as cloud changes (which as you know, just every week it seems like Microsoft and AWS are throwing something new out there at break-neck speed), it’s hard for me as someone who focuses on this 100 percent to keep up).” P2 also addressed the same point, stating that one has "to stay up to date with what Amazon and Azure are doing, what their tools are, even if it's not always necessary to have the newest and biggest things."   New and updated cloud services are often more efficient and have pricing incentives to motivate organizations to migrate to them.

In addition to the need to keep pace, IT managers face challenges with this growth even with the ability to source expertise to securely support their needs during and post-migration to cloud computing, particularly Small-Medium-Enterprises (SME) with limited resources (Foster et al., 2018). To keep pace with the changes in the industry and skills current with the needs of the cloud computing environment, some technical training and staff skills development is necessary. P6 also noted the importance of acquiring new skills, declaring that "the ability for my team members to be able to stay up to date on what's happening is going to determine success or failure so, it is that critical from that perspective." Table 10 shows the sub-themes representing the types of training that the managers referenced.

Table 10

Social Subsystem Strategy Theme 2: Types of Training and Areas for Skills Development

Code	Number of participants	%
General cloud computing	5	83
Cost management	4	66
Consulting services partnerships	2	33
Tools and technology partners	2	33

The role of training in facilitating organizational change cannot be understated. Sartori, Costantini, Ceschi, & Tommasi (2018) studied how concepts of training, skills development, and innovation, work together to affect organizations dealing with change. They defined training and development as educational opportunities to empower competencies of staff in the lifelong perspective of improving their performance, and competencies as personal characteristics that enable people to be effective in the changing contexts of both workplace and everyday life (Sartori et al., 2018). Another study by Sung and Choi (2014) indicated that organizational investments in internal training predicted interpersonal and organizational learning practices, which practices, in turn, increased the chances of performance with an adoption of innovation. Organizational investments in time, resources, and training spend is an effective strategy for success with operations cost management in cloud computing.

The skills that technical teams need to be successful in cloud environments can be different from those that exist in most organizations' pre-transformation, and training is the most effective means to bridge this gap. Khanye, Ophoff, and Johnston (2018) noted that managers should expect to change organizational roles citing for instance that, engineers may need to be more versatile and develop business analyst and other management skills. There is an overwhelming amount of content available from cloud service providers and partners, and some training options. Because of all the changes, P4 pointed out some degree of self-learning is necessary, and the managers are working to encourage team members to capitalize on opportunities to do so. P7 summarized this saying "to keep people trained on this, you have to encourage them to do at a high level, offer them direction, but they have to go out and do it on their own." All the participants agreed that some general cloud computing education for all team members is necessary, even for the non-technical team members. Managers can then supplement this general education with specific training that targets an individual's area of expertise, and ultimately the individual's other areas of interest.

For managers directly involved in cost management activities, some additional training in this area was is necessary. P3 understands the value of this, stating that "I think it's very important that there is a specific role that handles kind of like the nomenclature of the bills and how they allocate costs and can understand that correctly." To be more effective in his role as cost manager, P2 took a cost management course through a third party, and additional training related to the tools they use on site for monitoring, reporting, and managing cloud resources and spend. The tools partner training is also available online, and there is a subscription available for other managers to take that training as well. Other sources of skill improvement the managers listed included working with consulting services partners (P2), and community partners through technical events and other related seminars (P2). An organizational commitment to training and skills development is an important strategy and theme the managers identified, and critical for long-term success of the cloud computing adoption effort, and the development and maintenance of a successful operations management strategy.

Social Subsystem Strategy Theme 3: Evolve organizational roles/governance. Providing training and skills development for team members who are part of the IT operations and cost management team is one important way to improve organizational innovative performance, but the strategy to accomplish this must be backed up by an evolution of organizational roles. The roles must have the responsibility to maintain and ultimately integrate newly formed cloud management activities and processes into the organization. Table 11 shows the sub-themes the participants discussed under this theme.

Table 11

Social Subsystem Strategy Theme 3: Evolve Organizational Roles and Governance

Code	Number of participants	%
Cost manager role	4	66
Governance and executive staff support	2	33

Organization XYZ has an assigned technical analyst whose role is akin to that of a cost manager in other organizations. The cost manager role in the organization evolved from a technical analyst role to become the main champion of cost management activities in the organization, and P6 described the value of this stating that “I think that having that role associated with what's going on it's been beneficial for us.” In this role, the manager works with technical teams and business stakeholders to ensure clarity of responsibilities, and supports all the teams in the selection and use of economical cloud service options to meet their application performance needs, consistent with the nature of the role as defined by Mogouie, Arani, and, Shamsi (2015). Other desirable cost manager skills include financial analysis, project management, and basic technical background (RightScale, 2018b). The cost manager organizes and manages all cost management activities, and accounts for the cost performance to the CTO, and the skills are necessary to be successful in that role.

The effectiveness of the cost manager role is a function of the degree of executive sponsorship or senior management support, and the establishment of good governance within the organization. Cronemberger, Sayogo, and Gil-Garcia (2017), defined governance as the implementation of policies and standards to guide collaboration among members of an information sharing initiative, and found that there is a significant impact of executive involvement on the success of technology initiatives when mediated through that governance variable. This findings of Cronemberger et. al. (2015) is consistent with another study by Duryan and Smyth (2018), which study found that sponsorship of IS projects by senior level management gives the supporting roles ‘voice’ in the organization and increases the chance of successful integration of the initiative into the organization and the achievement of objectives.

The cost manager role is not the only one that the organization may have to evolve, the way IT managers view organizational roles with respect to vendors and internal business units may also need to evolve. Vithayathil (2018), studied the organizational impact on IT governance under cloud computing and concluded that successful IT departments under cloud computing will transform into new roles that address internal customer-facing issues and external cloud facing issues. In the external facing role, IT managers may need to manage activities such as service procurement, metering, billing, quality of service, service performance, reliability and compliance issues, and in the internal role; in the internal role acting as a consultant to understand business needs, and optimally matching the needs to cloud services (Vithayathil, 2018). Consistent with the evolution to these roles, organization XYZ has been effective deploying account managers in an external facing roles and business and technical analysts for internal customers (P2; P3; P5). This evolution of IT roles is necessary to align with the way IT managers acquire and manage services in a cloud environment, and are required for managers to continue to add value to their organizations after the transformation to cloud.

Social Subsystem Processes

STS Social Subsystem Process Theme 1: Team member support and professional growth.

One of the social subsystem strategies themes from the study was technical training and skills development. The skills IT operations management staff need to be successful in cloud computing are often different from those that have been successful in managing operation at the traditional data centers. This skills gap contributes to a fear of job loss, which Raza, Adenola, Nafarieh, and Robertson (2015) identified as one of the factors that initially slowed down the adoption of cloud computing. This concern for job security will persist if employees perceive themselves as lacking some of the skills to fully contribute to the pending technology change. Some of the DC infrastructure and platform management roles are assumed by the cloud service provider as a consequence of cloud adoption, and those roles may no longer be required in the organization. P7 pointed this out, saying “there are people in all organizations that are worried about what these Google and all these other people coming online are doing, is it going to make us obsolete?” Table 12 below shows the sub-themes I identified from the data.

Table 12

Social Subsystem Process Theme 1: Team Member Support and Professional Growth

Code	Number of participants	%
Cross-functional knowledge transfer	4	66
Manage team member growth and professional development	4	66

The managers identified technical training as one of the strategy areas, but there are other ways for staff members to acquire the skills they need for the job. Sung and Choi (2014) identified three types of learning practices as individual learning practices, interpersonal learning practices, and organizational learning practices, and went on to state that collective training that offers opportunities for communications among employees from same or different departments stimulate employees to share ideas and experiences. To facilitate this cross-functional team learning, P5 stated that “we've recently broken-down our siloed teams into more cross-functional resources so that you have your onboarding, also has someone from like a networking platform, an application, and an infrastructure perspective altogether.” P5 and also added that “knowledge sharing and collaboration has become even more important because the boundaries between IT disciplines is defusing.” In such an environment, Sung and Choi (2014) noted that employees were more likely to use the collaborative interactions to promote interpersonal learning in the form of active involvement in mutual learning, coaching, and cross-training of each other’s tasks and responsibilities. When employees feel that they have organizational support to develop new skills on the job, it has the potential to reduce their fear of job loss and contribute to job satisfaction and commitments to organizational outcomes.

Having all of the learning practices in place is useful, but the path to job growth and professional development should not be left to individual IT leaders to manage. Individual ownership for their professional development is necessary. P7 recognized the challenge with this, stating that “you have to encourage them to do at a high level, offer them direction, but they have to go out and do it on their own; I cannot sit in a classroom every day and train you, that's one of the biggest challenges we have.” In addition to guiding and encouraging employees through the process, organizations must also formalize the process for supporting employee development and growth to ensure that opportunities are made available to all teams and team members. P7 qualified this in general terms for the organization, saying “fundamentally as an organization, learning is a lifelong process, so we're constantly pushing our team to pursue professional development opportunities and extend that training for themselves.” This organizational perspective is important, and when employees understand this and recognize opportunities to develop new skills and professional growth within the context of the cloud transformation activities, organizations have a higher possibility of success.

Applications to Professional Practice

Cloud computing is now widely accepted as the IT platform of the future, with rapid adoption rates across all sizes of businesses and in all sectors of the economy (Garrison, Rebman, & Kim, 2018). RightScale (2018a), also found in a survey that the extent of cloud computing usage was very high at 96% of the respondents, an increase of

6% from 2017. The major advertised benefits of cloud adoption are cost reduction, improved capability, business agility, and enhanced scalability (Chen et al., 2016). The challenge for IT managers is that businesses using cloud seem to benefit more in enhanced scalability than in cost reduction and increased business capabilities (Chen et al., 2016). In the same survey by RightScale, they found that more than 30% of cloud resources provisioned were wasted, and that optimizing costs was the top initiative across all cloud users (RightScale), an indication that IT managers need a comprehensive approach to addressing the challenge of cost management in cloud environments. The majority of academic researchers in this area have focused on the mathematical algorithms to optimize resource provisioning and utilization for specific types of applications or workloads (Khan et al., 2015; Manvi & Shyam, 2014; Reddy et al., 2017). Other researchers who addressed costs in their studies focused on the realization of broader economic objectives such as ROI and NPV (Maresova et al., 2017; Mohammed, Aljumaili, & Salah, 2014). As valuable as these studies are, they do not directly address a broader set of strategies and processes that IT managers can use to meet this challenge of cost management.

This study, using the STS framework extends on the knowledge from the previous research work by exploring both the technical and social aspects of strategies and processes IT managers use to manage their operations costs. The findings can be used by managers to understand what planning activities are important to undertake before migration. Having a strategy for transformation to the cloud is particularly important. Goutas, Sutanto, and Aldarbesti (2015) noted that many IT managers had approached the cloud without a clear strategy in mind, which has resulted in numerous downfalls. The first part of the strategy is preparation and planning activity entails an understanding of the motivation for cloud adoption, assessing applications and the IT infrastructure as a whole to understand applications and systems profiles, then architecting the applications to leverage optimal cloud services to meet performance and cost requirements. The findings also highlight the need to comprehensively tag all cloud resources, to enable managers to use the right set of tools to track and report on all of cloud resource usage.

The findings also indicated that architecting systems for the cloud is more than rightsizing compute instances to meet performance requirements (Evangelinou et al., 2018), but must also include considerations for the use of reserved instances and autoscaling. These three techniques have the most potential for overall cloud cost savings, with some estimates of savings from use of reserved instances alone accounting for a potential 45% cost saving for the applications that can benefit from it (P2; RightScale, 2018b). Knowledge of application profiles and business requirements are useful in determining which of these techniques (or a combination thereof) are most effective for managers to use to reduce operational costs.

Most IT operations managers are familiar with operations monitoring for issues and alerts in the infrastructure. The findings of this study indicate that to be successful with cost management; managers need more extensive monitoring beyond what they did at the traditional data centers (monitoring issues with applications and IT infrastructure). A large amount of data points, the metered billing for resource usage, and the dynamic pricing associated with most cloud resources mean the normal monitoring activities need to be extended to include resource utilization at a level that commensurates with the billable components of the cloud service provider. The findings also indicated that effective cost management means that making teams (and departments, and or business units) accountable for their consumption and associated costs by a process that allocates their costs to them, gives them access to the reports, as well as ensure their active participation in regular reviews of the reports. Managers must promptly address any wastage they identify through this review process. With such an approach, all IT managers will have full responsibilities for managing their resource usage and cost.

In addition to presenting the key strategies for cost management, the findings also outlined key processes that have been effective when instituted to work in concert with the strategies to ensure organizational changes necessary with the adoption of cloud computing can be established and integrated within the IT organization. Some of these processes are also necessary to improve the medium to long-term potential to sustain the initial cost savings effort. Technical processes such as regular cost and resource consumption reviews, provisioning controls, and documentation and knowledge management processes are all important to institute in the organization to continue to leverage the most efficient cloud resources for system performance and costs.

The STS framework posits that to realize the full potential of technological innovation, both the technical and social subsystems must be optimized. Among the conditions that Parsons (1951) identified as necessary for technological innovation to be successful in an organization are activity integration, continuity of occupation, and consistent roles in support of the systems. While the majority of cloud computing researchers focused on the technical aspects of adoption and use of the systems, the findings of this study are consistent with these conditions, addressing both the technical and social subsystem strategies and processes.

The social subsystem strategies from the findings highlighted the value of collaboration, employee training, on the job skills improvement, and evolution of organizational roles. Managers cannot ignore the importance of the social subsystem aspects of cost management and must combine them with a process to provide team members support through training and knowledge sharing, and an organizational commitment to the professional growth and development of team members. With these practices, managers can ensure that team members can commit to and positively contribute to the effort to manage cloud costs, cloud adoption, and the value of the technology to the organization. The contribution of this study to professional practice extends to cloud service providers as well. Cloud service providers who understand the strategies and processes their subscribers need to be successful in cloud environments, can be more effective in supporting them with the tools and training they need. The success of current customers can only incent other clients to accelerate their adoption of the technology.

Implications for Social Change

Cloud computing services are now widely used in enterprises, and the rate of adoption is projected to reach $266B in 2021 (IDC, 2018b). The rapid adoption has come with challenges for IT managers, who report a fairly high degree of cloud wastage. But in spite of this estimated wastage, IDC also predicted that the worldwide public cloud services spending would reach $180B in 2018, an increase of 23.7% over 2017 (IDC, 2018b). The need for more cloud computing services combined with the what IT managers perceive to be an inefficient utilization of the resources is one of the reasons why cloud cost containment is the top initiative of IT managers for 2018 (RightScale, 2018b). This need to manage cloud costs is the principal reason why the findings from this study to address strategies and processes to manage IT operational costs may be of value to managers, and can have potential implications for positive social change.

Reducing cloud services expenditure has the potential to make a positive social impact for the organizations by making them more economically profitable, improving their competitive posture and viability as a business. Economic viability has the potential to support local economies through the provision of, and growth of employment opportunities for the communities. Potential organizational cost savings from the use of the findings from this study can also be re-invested in businesses to create new products or services, which in turn may lead to more jobs and improvement of stakeholders return on investment, all of which can have a positive social impact.

Some of the social subsystem strategies from this study, if implemented as part of the cost management effort may also potentially contribute to positive social change. Training and skills improvements, new and more relevant roles for team members, providing the necessary organizational support for team member growth and professional development all have the potential to improve job satisfaction. When employees get the training and support they need to continue to add value at work, extend their tenure with the organization, and increase their career advancement opportunities, they are more likely to be satisfied employees (Iliopoulos, Morrissey, Baryeh, & Polyzois, 2018). The combination of satisfied employees, longer-term job opportunities and advancement, and a stable work environment have the potential to contribute to a more stable local economy and economic growth.

Another area where the findings from this study have the potential for positive social change is in the area of reducing energy use. Balasooriya et al. (2016) found that cloud service providers engage in initiatives considered adaptive, energy efficient, and environmentally responsible in the way they efficiently use energy resources. When organizations are successful in managing their operations costs in cloud computing, they are more likely to migrate more workloads to the cloud and increase their cloud usage. The potential positive social impact of more cloud computing usage is an effective reduction in net energy use and carbon footprint which fosters environmental sustainability.

Recommendations for Action

The objective of this study was to explore strategies and processes IT managers use to manage IT infrastructure operations costs in cloud computing. The findings from the study provide a clear outline of the key strategies, specifying the importance of planning and preparation before transformation, architecting applications and systems to optimally fit cloud services, rightsizing systems, leveraging reserved instances, and (the elasticity of the cloud to) autoscale resources where applicable. Making cost data transparent to IT managers, and holding them accountable for the cloud spend were all important strategies, as were comprehensive monitoring of cloud services consumption and related costs to ensure prompt resolution of any deviations from targets. Regular reviews of new and often more efficient service options from cloud providers are also necessary to identify medium to long-term opportunities to improve service performance and still control costs. The use of tools is critical for monitoring and reporting, and as an aid to managers to quickly identify inefficiencies or underutilized resources in the infrastructure. The findings also indicated provisioning controls and documentation are important processes that support the overall technical strategies, and that IT managers must have a long-term perspective on cost management, and yet be ready to capitalize on newer and more efficient cloud technologies to further reduce costs.

The technical strategies and processes are important, but the findings indicated that it is necessary to combine them with social subsystem strategies and processes such as fostering internal and external collaboration, providing employees with training and skills improvement opportunities, and the evolving some organizational roles and governance. These changes are necessary to build the skills organization needs, as well as integrate any new processes into the organization. Because of the nature of cloud computing platform, managers may need to modify the traditional siloed IT organizational structure to a more cross-functional team dynamic to facilitate cross-functional support and knowledge transfer between the resources supporting the environment. The organization must also be prepared to support team members involved in the transformation with technical and professional development opportunities, and incent them to take advantage of them. If IT managers effectively practice the sociotechnical strategies and processes from the findings of this study, they stand a great chance to succeed in their post-transformation operational cost management effort and give their organizations a better chance of realizing one of their key objectives for cloud computing adoption.

Recommendations for Further Research

In this study, the focus was on the strategies and processes IT managers use to manage infrastructure operations costs in cloud computing. The findings from the study represent information I collected and analyzed from a single IT organization in the Tampa, Florida area in the United States. IT managers in the organization use AWS and MS Azure cloud services, so their responses to the interview questions reflect their shared experiences with these cloud services providers.

As recommendations for further studies, I will suggest additional studies using a similar single or multiple case study of other organizations outside of Tampa area of the US or even other regions of the world, with a successful record of IT operations costs management in cloud computing. Findings from additional studies outside of this geographic area if consistent with this study will help to lend credence to these findings and improve applicability. According to the RightScale 2018 report on the state of the cloud, organizations are using an average of over three clouds, and over half of them use hybrid cloud infrastructure (RightScale, 2018a). Since the site for this study used two such clouds—AWS and Azure, I will propose a study with an organization that uses a hybrid cloud, and another that uses four or more cloud services providers. Findings from such studies may potentially extend what we know about strategies and processes from this study, and or help managers understand what additional strategies and processes may be necessary for environments with more than two clouds. Other studies will extend the findings from this study will help IT managers in their effort to contain costs in different cloud environments and types.

Reflections

When I started this journey to complete this doctorate program, I expected that it would be interesting, and I will have the opportunity to learn new skills and improve my abilities along the way. Little did I realize what I was committing to and the level of perseverance which will be required to be successful with it, all the while balancing multiple businesses and family. For a while, it was hard to see how this it was ever going to end and what it would take to get there. But in hindsight, I am I made the decision, and for the opportunity to be in the program to the end. There was something new to learn in every class, from my colleagues and instructors, but in the process, I also learned a lot about myself and what motivates me to accomplish my dreams, something I did not expect when I enrolled in the program.

My research experience through this program has been an eye opener. It was interesting for me to research a topic that I have some passion for, and yet take steps not to bias the study with my own experiences. Through the research, I have been able to expand my understanding of IT operations cost management in cloud computing, and I hope that other IT managers may also find value for their organizations by applying the findings of this studies to their unique transformation situations. It is also my hope that other researchers may also be able to use this study as a basis to perform further studies on the subject and to continue to improve professional practice in this area that is so critical to cloud computing adoption and use in organizations.

Summary and Study Conclusions

The findings from this qualitative single case study indicated that to effectively manage IT infrastructure operations costs in cloud computing the key strategies IT managers need includes (a) assess and plan the migration, (b) architect applications and systems to optimally fit cloud services, (c) select and use the right tools to actively track and manage all cloud resources, and (e) to establish and execute a medium to long-term plan to capitalize on ongoing changes to their environments. The study also revealed that social strategies such as collaboration, employee training, and evolving organization roles are also necessary to support the human aspects of the innovation, and without which complete success will be difficult to achieve.

The findings indicated that managers must also be ready to support the strategies with new processes, some of which may be new to some organizations. The processes I identified from this study included regular reviews of cloud consumption and costs, provisioning controls, documentation and knowledge management, and management of team member growth and professional development. It is important to note that STS framework states that it is only when IT managers effectively implement both the technical and social strategies and processes that they will have the high likelihood of success from the application of technological innovation.

References

Abdelmaboud, A., Jawawi, D. N., Ghani, I., & Elsafi, A. (2015). A comparative evaluation of cloud migration optimization approaches: A systematic literature review. *Journal of Theoretical and Applied Information Technology*, *79*, 395-414. Retrieved from http://www.jatit.org/volumes.php

Accenture Consulting. (2010). Accenture sustainability: The environmental benefits of moving to the cloud. Accenture in collaboration with WSP Environment & Energy. Retrieved from https://aboutcloud.ch/en/studies/accenture%E2%80%8E-cloud-computing-and-sustainability-environmental-benefits-moving-cloud.

Alkhalil, A., Sahandi, R., & John, D. (2017). An exploration of the determinants for decision to migrate existing resources to cloud computing using an integrated TOE-DOI model. *Journal of Cloud Computing*, *6*(1), 1-10. doi:10.1186/s13677-016-0072-x

Amazon Web Services. (2017a). Amazon Aurora pricing. Retrieved December 21, 2017, from https://aws.amazon.com/rds/aurora/pricing/

Amazon Web Services. (2017b). Amazon EBS pricing. Retrieved December 21, 2017, from https://aws.amazon.com/ebs/pricing/

Amazon Web Services. (2017c). Amazon Elastic Block Store. Retrieved December 21, 2017, from https://aws.amazon.com/ebs/

Amazon Web Services. (2017d). Amazon RDS pricing. Retrieved December 21, 2017, from https://aws.amazon.com/rds/pricing/

Amazon Web Services. (2017e). Amazon S3 pricing. Retrieved December 21, 2017, from https://aws.amazon.com/s3/pricing/

Ananthakrishnan, R., Chard, K., Foster, I., & Tuecke, S. (2015). Globus platform-as-a-service for collaborative science applications. *Concurrency and Computation: Practice and Experience, 27*, 290-305. doi:10.1002/cpe.3262

Anyan, F. (2013). The influence of power shifts in data collection and analysis stages: A focus on qualitative research interview. *The Qualitative Report, 18*(2013), 1-9. doi:10.1177/1049732313502128

Ashraf, A., Byholm, B., & Porres, I. (2016). Prediction-based VM provisioning and admission control for multi-tier web applications. *Journal of Cloud Computing, 5*(15), 1-21. doi:10.1186/s13677-016-0065-9

Assunção, M. D., Calheiros, R. N., Bianchi, S., Netto, M. A., & Buyya, R. (2015). Big data computing and clouds: Trends and future directions. *Journal of Parallel and Distributed Computing, 79*(5), 3-15. doi:10.1016/j.jpdc.2014.08.003

Avram, M. G. (2014). Advantages and challenges of adopting cloud computing from an enterprise perspective. *Procedia Technology, 12*, 529-534. doi:10.1016/j.protcy.2013.12.525

Awa, H. O., & Ojiabo, O. U. (2016). A model of adoption determinants of ERP within TOE framework. *Information Technology & People, 29*(4), 901-930. doi:10.1108/ITP-03-2015-0068

Baines, J. T., Taylor, C. N., & Vanclay, F. (2013). Social impact assessment and ethical research principles: Ethical professional practice in impact assessment Part II.

Impact Assessment and Project Appraisal, *31*, 254-260. doi:10.1080/14615517.2013.850306

Balasooriya, P. N. L. N., Wibowo, S., & Wells, M. (2016). Green cloud computing and economics of the cloud: Moving towards sustainable future. *GSTF Journal on Computing*, *5*(1), 15-20. doi:10.5176/2251-3043_4.4.347

Banker, D., Mashruwala, R., & Tripathy, A. (2014). Does a differentiation strategy lead to more sustainable financial performance than a cost leadership strategy? *Management Decision, 52*, 872-896. doi:10.1108/MD-05-2013-0282

Baskerville, R., Bunker, D., Olaisen, J., Pries-Heje, J., Larsen, T. J., & Swanson, E. B. (2014). Diffusion and innovation theory: Past, present, and future contributions to academia and practice. *Creating Value for All Through IT*, *429*, 295-300. doi:10.1007/978-3-662-43459-8_18

Bauer, J., & Bellamy, A. (2017). Latent effects of cloud computing on IT capacity management structures. *International Journal of Computer and Communication Engineering*, *6*(2), 111-126. doi:10.17706/ijcce.2017.6.2.111-126

Baxter, G., & Sommerville, I. (2011). Socio-technical systems: From design methods to systems engineering. *Interacting With Computers*, *23*(1), 4-17. doi:10.1016/j.intcom.2010.07.003

Bayramusta, M., & Nasir, V. A. (2016). A fad or future of IT? A comprehensive literature review on the cloud computing research. *International Journal of Information Management*, *36*, 635-644. doi:10.1016/j.ijinfomgt.2016.04.006

Banerjee, S., Gupta, N., & Gupta, V. (2014). Implementation and management of framework for PAAS in cloud computing. *International Journal of Innovations & Advancement in Computer Science, 2*(2), 38-49. Retrieved from http://www.academicscience.co.in/publishedpaperlist.php?vol=IJIACS

Bedford, D. S. (2015). Management control systems across different modes of innovation: Implications for firm performance. *Management Accounting Research, 28*(2015), 12-30. doi:10.1016/j.mar.2015.04.003

Birt, L., Scott, S., Cavers, D., Campbell, C., & Walter, F. (2016). Member checking: A tool to enhance trustworthiness or merely a nod to validation. *Qualitative Health Research, 26,* 1802-1811. doi:10.1177/1049732316654870

Boddy, C. R., & Boddy, C. R. (2016). Sample size for qualitative research. *Qualitative Market Research: An International Journal, 19*, 426-432. doi:10.1108/QMR-06-2016-0053

Boru, D., Kliazovich, D., Granelli, F., Bouvry, P., & Zomaya, A. Y. (2015). Energy-efficient data replication in cloud computing datacenters. *Cluster computing, 18*, 385-402. doi:10.1007/s10586-014-0404-x

Botta, A., De Donato, W., Persico, V., & Pescapé, A. (2016). Integration of cloud computing and internet of things: A survey. *Future Generation Computer Systems*, *56*, 684-700. doi:10.1016/j.future.2015.09.021

Breitenbücher, U., Binz, T., Fehling, C., Kopp, O., Leymann, F., & Wieland, M. (2014). Policy-aware provisioning and management of cloud applications. *International*

Journal On Advances in Security, 7(1), 15-36. Retrieved from www.iariajournals.org/security/

Bruneo, D. (2014). A stochastic model to investigate data center performance and QoS in IaaS cloud computing systems. *IEEE Transactions on Parallel and Distributed Systems, 25*, 560-569. doi:10.1109/TPDS.2013.67

Carayon, P., Hancock, P., Leveson, N., Noy, I., Sznelwar, L., & Van Hootegem, G. (2015). Advancing a sociotechnical systems approach to workplace safety–developing the conceptual framework. *Ergonomics, 58*, 548-564. doi:10.1080/00140139.2015.1015623

Carter, N., Bryant-Lukosius, D., DiCenso, A., Blythe, J., & Neville, A. J. (2014). The use of triangulation in qualitative research. *In Oncology Nursing Forum, 41*, 545-547. doi:10.1188/14.ONF.545-547

Cartlidge, J., & Clamp, P. (2014). Correcting a financial brokerage model for cloud computing: closing the window of opportunity for commercialisation. *Journal of Cloud Computing, 3*(2), 1-20. doi:10.1186/2192-113X-3-2

Chang, V., & Ramachandran, M. (2016). Towards achieving data security with the cloud computing adoption framework. *IEEE Transactions on Services Computing, 9*(1), 138-151. doi:10.1109/TSC.2015.2491281

Chang, V., Walters, R. J., & Wills, G. B. (2015). Cloud computing and frameworks for organisational cloud adoption. *Delivery and Adoption of Cloud Computing Services in Contemporary Organizations, 1*, 1-28. doi:10.4018/978-1-4666-9466-8.ch030

Chang, V., Kuo, Y. H., & Ramachandran, M. (2016). Cloud computing adoption framework: A security framework for business clouds. *Future Generation Computer Systems, 57*(C), 24-41. doi:10.1016/j.future.2015.09.031

Chen, T., Chuang, T. T., & Nakatani, K. (2016). The perceived business benefit of cloud computing: An exploratory study. *Journal of International Technology and Information Management, 25*(4), 100-122. Retrieved from http://scholarworks.lib.csusb.edu/jitim

Cherns, A. (1987). Principles of sociotechnical design revisited. *Human Relations, 40*(3), 153-161. doi:10.1177/001872677602900806

Chou, D. C. (2015). Cloud computing: A value creation model. *Computer Standards & Interfaces, 38*(2*)*, 72-77. doi:10.1016/j.csi.2014.10.001

Choy, L. T. (2014). The strengths and weaknesses of research methodology: Comparison and complimentary between qualitative and quantitative approaches. *IOSR Journal of Humanities and Social Science, 19*(4), 99-104. Retrieved from http://www.iosrjournals.org/iosr-jhss.html

Church, P., Mueller, H., Ryan, C., Gogouvitis, S. V., Goscinski, A., & Tari, Z. (2017). Migration of a SCADA system to IaaS clouds–a case study. *Journal of Cloud Computing, 6*(11), 1-12. doi:10.1186/s13677-017-0080-5

Chun, S. H., & Choi, B. S. (2014). Service models and pricing schemes for cloud computing. *Cluster Computing, 17*, 529-535. doi:10.1007/s10586-013-0296-1

Cleary, M., Horsfall, J., & Hayter, M. (2014). Data collection and sampling in qualitative research: does size matter? *Journal of Advanced Nursing*, 70, 473-475. doi:10.1111/jan.12163

Coutinho, E. F., de Carvalho Sousa, F. R., Rego, P. A. L., Gomes, D. G., & de Souza, J. N. (2015). Elasticity in cloud computing: A survey. *Annals of Telecommunications*, *70*, 289-309. doi:10.1007/s12243-014-0450-7

Cozzolini, L., & Berbegal-Mirabent, J. (2017, September). The key lies in the process, not in the innovation. *In European Conference on Knowledge Management* (pp. 1241-1250). Academic Conferences International Limited.

Creswell, J. W., Hanson, W. E., Plano, V. L. C., & Morales, A. (2007). Qualitative research designs selection and implementation. *The Counseling Psychologist*, *35*(2), 236-264. doi:10.1177/0011000006287390

Cronemberger, F., Sayogo, D. S., & Gil-Garcia, J. R. (2017). Assessing the role of executive involvement and information needs as socio-technical determinants of governance in IIS success. *In Proceedings of the 50th Hawaii International Conference on System Sciences* (pp. 2923-2932). doi:10.24251/HICSS.2017.354

Cronin, C. (2014). Using case study research as a rigorous form of inquiry. *Nurse Researcher, 21*(5), 19-27. doi:10.7748/nr.21.5.19.e1240

Darawsheh, W. (2014). Reflexivity in research: Promoting rigour, reliability and validity in qualitative research. *International Journal of Therapy and Rehabilitation, 21*, 560-568. doi:10.12968/ijtr.2014.21.12.560

Dautov, R., Paraskakis, I., & Stannett, M. (2014). Towards a framework for monitoring cloud application platforms as sensor networks. *Cluster computing, 17*, 1203-1213. doi:10.1007/s10586-014-0389-5

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: a comparison of two theoretical models. *Management Science, 35*, 982-1003. doi:10.1287/mnsc.35.8.982

Dinh, H. T., Lee, C., Niyato, D., & Wang, P. (2013). A survey of mobile cloud computing: architecture, applications, and approaches. *Wireless Communications and Mobile Computing, 13*, 1587-1611. doi:10.1002/wcm.1203

Doody, O., & Noonan, M. (2013). Preparing and conducting interviews to collect data. *Nurse Researcher, 20*(5), 28-32. doi:10.7748/nr2013.05.20.5.28.e327

Driskell, J. E., Salas, E., & Driskell, T. (2018). Foundations of teamwork and collaboration. *American Psychologist*, *73*, 334. doi:10.1037/amp0000241

Duryan, M., & Smyth, H. (2018). Cultivating sustainable communities of practice within hierarchical bureaucracies: the crucial role of an executive sponsorship. *International Journal of Managing Projects in Business, 2018*(1), 1-23. doi:0.1108/IJMPB-03-2018-0040

El Zant, B., & Gagnaire, M. (2015). Towards a unified customer aware figure of merit for CSP selection. *Journal of Cloud Computing, 4*(24), 1-23. doi:10.1186/s13677-015-0049-1

Elo, S., Kääriäinen, M., Kanste, O., Pölkki, T., Utriainen, K., & Kyngäs, H. (2014). Qualitative content analysis. *SAGE Open, 2014*(4), 1-10. doi:10.1177/2158244014522633

Evangelinou, A., Ciavotta, M., Ardagna, D., Kopaneli, A., Kousiouris, G., & Varvarigou, T. (2018). Enterprise applications cloud rightsizing through a joint benchmarking and optimization approach. *Future Generation Computer Systems, 78*(2018), 102-114. doi:10.1016/j.future.2016.11.002

Fang, D., Liu, X., Romdhani, I., & Pahl, C. (2015). An approach to unified cloud service access, manipulation and dynamic orchestration via semantic cloud service operation specification framework. *Journal of Cloud Computing*, *4*, 1-20. doi:10.1186/s13677-015-0039-3

Fahmideh, M., & Beydoun, G. (2018). Reusing empirical knowledge during cloud computing adoption. *Journal of Systems and Software, 138*(2018), 124-157. doi:10.1016/j.jss.2017.12.011

Fatema, K., Emeakaroha, V. C., Healy, P. D., Morrison, J. P., & Lynn, T. (2014). A survey of cloud monitoring tools: Taxonomy, capabilities and objectives. *Journal of Parallel and Distributed Computing, 74*, 2918-2933. doi:10.1016/j.jpdc.2014.06.007

Foster, D., White, L., Adams, J., Erdil, D. C., Hyman, H., Kurkovsky, S., ... & Stott, L. (2018, July). Cloud computing: developing contemporary computer science curriculum for a cloud-first future. *In Proceedings of the 23rd Annual ACM*

Conference on Innovation and Technology in Computer Science Education (pp. 346-347). ACM. doi:10.1145/3197091.3205843

Fox, W. M. (1995). Sociotechnical system principles and guidelines: Past and present. *The Journal of Applied Behavioral Science, 31*(1), 91-105. doi:10.1177/0021886395311009

Fusch, P. I., & Ness, L. R. (2015). Are we there yet? Data saturation in qualitative research. *The Qualitative Report, 20,* 1408-1416. doi:10.1177/1468794107085301.

Ganapathy, M. (2016). Qualitative data analysis: Making it easy for nurse researcher. *International Journal of Nursing Education, 8*(2), 106-110. doi:10.5958/0974-9357.2016.00057.X

Gannon, D., Barga, R., & Sundaresan, N. (2017). Cloud-native applications. *IEEE Cloud Computing, 4(*5), 16-21. doi:10.1109/MCC

Garousi, G., Garousi-Yusifoğlu, V., Ruhe, G., Zhi, J., Moussavi, M., & Smith, B. (2015). Usage and usefulness of technical software documentation: An industrial case study. *Information and Software Technology, 57,* 664-682. doi:0.1016/j.infsof.2014.08.003

Garrison, G., Rebman Jr, C. M., & Kim, S. H. (2018). An identification of factors motivating individuals' use of cloud-based services. *Journal of Computer Information Systems, 58*(1), 19-29. doi:10.1080/08874417.2016.1180653

Garrison, G., Wakefield, R. L., & Kim, S. (2015). The effects of IT capabilities and delivery model on cloud computing success and firm performance for cloud supported processes and operations. *International Journal of Information Management, 35*, 377-393. doi:10.1016/j.ijinfomgt.2015.03.001

Gavin, J., Baxter, D., & Hester, A. (2014). Socio-technical systems theory as a diagnostic tool for examining underutilization of wiki technology. *The Learning Organization, 21*(1), 48-68. doi:10.1108/TLO-10-2012-0065

Ghahramani, M. H., Zhou, M., & Hon, C. T. (2017). Toward cloud computing QoS architecture: Analysis of cloud systems and cloud services. *IEEE/CAA Journal of Automatica Sinica, 4*(1), 6-18. doi:10.1109/JAS.2017.7510313

Gholami, M. F., Daneshgar, F., Beydoun, G., & Rabhi, F. (2017). Challenges in migrating legacy software systems to the cloud—an empirical study. *Information Systems, 67(*2017), 100-113. doi:10.1016/j.is.2017.03.008

Gibson, C. B. (2017). Elaboration, generalization, triangulation, and interpretation: On enhancing the value of mixed method research. *Organizational Research Methods, 20*(2), 193-223. doi: 10.1177/1094428116639133

Goldsztajn, D., Ferragut, A., Paganini, F., & Jonckheere, M. (2018). Controlling the number of active instances in a cloud environment. *ACM SIGMETRICS Performance Evaluation Review, 45*(2), 15-20. doi:10.1145/3199524.3199529

Golafshani, N. (2003). Understanding reliability and validity in qualitative research. *The Qualitative Report, 8*, 597-606. Retrieved from nsuworks.nova.edu/tqr/

González-Martínez, J. A., Bote-Lorenzo, M. L., Gómez-Sánchez, E., & Cano-Parra, R. (2015). Cloud computing and education: A state-of-the-art survey. *Computers & Education, 80*(2015), 132-151. doi:10.1016/j.compedu.2014.08.017

Goutas, L., Sutanto, J., & Aldarbesti, H. (2015). The building blocks of a cloud strategy: evidence from three SaaS providers. *Communications of the ACM, 59*(1), 90-97. doi:10.1145/2756545

Goyal, S. (2014). Public vs private vs hybrid vs community-cloud computing: a critical review. *International Journal of Computer Network and Information Security*, *6*(3), 20-29. doi:10.5815/ijcnis.2014.03.03

Gutierrez, A., Boukrami, E., & Lumsden, R. (2015). Technological, organisational and environmental factors influencing managers' decision to adopt cloud computing in the UK. *Journal of Enterprise Information Management, 28*, 788-807. doi:10.1108/JEIM-01-2015-0001

Hachicha, Z. S., & Mezghani, K. (2018). Understanding intentions to switch toward cloud computing at firms' level: A multiple case study in Tunisia. *Journal of Global Information Management*, *26*(1), 136-165. doi:10.4018/JGIM.2018010108

Harvey, L. (2015). Beyond member-checking: A dialogic approach to the research interview. *International Journal of Research & Method in Education, 38*(1), 23-38. doi:10.1177/1049732316654870

Hashem, I. A. T., Yaqoob, I., Anuar, N. B., Mokhtar, S., Gani, A., & Khan, S. U. (2015). The rise of "big data" on cloud computing: Review and open research issues. *Information Systems, 47*, 98-115. doi:10.1016/j.is.2014.07.006

He, H. (2016). A Resource reservation based framework for QoS-aware resource provision in cloud computing. *International Journal of Grid and Distributed Computing*, *9*(9), 193-204. doi:10.14257/ijgdc.2016.9.9.17

He, J., Wen, Y., Huang, J., & Wu, D. (2014). On the Cost–QoE tradeoff for cloud-based video streaming under Amazon EC2's pricing models. *IEEE Transactions on Circuits and Systems for Video Technology*, *24*, 669-680. doi:10.1109/TCSVT.2013.2283430

Hintemann, R. (2015). Consolidation, colocation, virtualization, and cloud computing: the impact of the changing structure of data centers on total electricity demand. *Advances in Intelligent Systems and Computing*, *310*, 125-136. doi:10.1007/978-3-319-09228-7_7

Houghton, C., Murphy, K., Shaw, D., & Casey, D. (2015). Qualitative case study data analysis: An Example from practice. *Nurse Researcher*, *22*(5), 8-12. doi:10.7748/nr.22.5.8.e1307

Hsu, P. F., Ray, S., & Li-Hsieh, Y. Y. (2014). Examining cloud computing adoption intention, pricing mechanism, and deployment model. *International Journal of Information Management*, *34*, 474-488. doi:10.1016/j.ijinfomgt.2014.04.006

Hummaida, A. R., Paton, N. W., & Sakellariou, R. (2016). Adaptation in cloud resource configuration: a survey. *Journal of Cloud Computing*, *5*(7), 1-16. doi:10.1186/s13677-016-0057-9

Hundu, J. T., & Mamman, E. S. (2017). Cloud computing and its benefits to libraries. *International Journal of Science and Applied Research*, 2(2), 18-32. Retrieved from http://www.ijsar.org.ng/index.php/ijsar/issue/archive

Hussain, S. A., Fatima, M., Saeed, A., Raza, I., & Shahzad, R. K. (2017). Multilevel classification of security concerns in cloud computing. *Applied Computing and Informatics, 13*(1), 57-65. doi:10.1016/j.aci.2016.03.001

Hussein, A. (2015). The use of triangulation in social sciences research: Can qualitative and quantitative methods be combined? *Journal of Comparative Social Work, 4*(1), 1-12. Retrieved from http://journal.uia.no/index.php/JCSW/issue/archive

Hyett, N., Kenny, A., & Dickson-Swift, V. (2014). Methodology or method? A critical review of qualitative case study reports. *International Journal of Qualitative Studies on Health and Well-being, 9*, 1-12. doi:10.3402/qhw.v9.23606.

Iliopoulos, E., Morrissey, N., Baryeh, K., & Polyzois, I. (2018). Correlation between workplace learning and job satisfaction of NHS healthcare professionals. *British Journal of Healthcare Management, 24(*5), 226-233. doi:10.12968/bjhc.2018.24.5.226

Iqbal, S., Mat Kiah, M. L., Dhaghighi, B., Hussain, M., Khan, S., Khan, M. K., & Raymond Choo, K. (2016). On cloud security attacks: A taxonomy and intrusion detection and prevention as a service. *Journal of Network & Computer Applications, 74*(C), 98-120. doi:10.1016/j.jnca.2016.08.016

IT Operations Management. (2017, December 21). Retrieved from http://www.bmc.com/it-solutions/it-operations-management.html

Ibrahim, A. B. (2015). Strategy types and small firms' performance an empirical investigation. *Journal of Small Business Strategy, 4(1),* 13-22. Retrieved from http://www.jsbs.org/

Ioannidis, J. P., Greenland, S., Hlatky, M. A., Khoury, M. J., Macleod, M. R., Moher, D., ... & Tibshirani, R. (2014). Increasing value and reducing waste in research design, conduct, and analysis. *The Lancet, 383*(9912), 166-175. doi:10.1016/S0140-6736(13)62227-8

International Data Corporation. (2017a). Worldwide public cloud services spending forecast to reach $122.5 billion in 2017, according to IDC. Retrieved February 20, 2017, from http://www.idc.com/getdoc.jsp?containerId=prUS42321417

International Data Corporation. (2017b, July 18). Worldwide public cloud services spending forecast to reach $266 billion in 2021, according to IDC. Retrieved September 9, 2018, from https://www.idc.com/getdoc.jsp?containerId=prUS42889917

Jamshed, S. (2014). Qualitative research method-interviewing and observation. *Journal of Basic and Clinical Pharmacy, 5*(4), 87-88. doi:10.4103/0976-0105.141942

Jede, A., & Teuteberg, F. (2016). Understanding Socio-Technical impacts arising from Software-as-a-Service usage in companies. *Business & Information Systems Engineering, 58*(3), 161-176. doi:10.1007/s12599-016-0429-1

Kadam, R. A. (2017). Informed consent process: A step further towards making it meaningful!. *Perspectives in Clinical Research, 8*(3), 107-112. doi:10.4103/picr.PICR_147_16

Kahlke, R. M., & Hon, B. (2014). Generic qualitative approaches: Pitfalls and benefits of methodological mixology. *International Journal of Qualitative Methods, 13(1),* 37-52. doi:10.1177/160940691401300119

Kandpal, M., & Patel, K. (2018). Survey of machine learning algorithms for dynamic resource pricing in cloud. *International Journal of Scientific Research in Science, Engineering and Technology*, *4(*2), 93-97. Retrieved from www.http://ijsrset.com/paper/?param=paper

Kaur, T., & Chana, I. (2015). Energy efficiency techniques in cloud computing: A survey and taxonomy. *ACM Computing Surveys, 48*(2), 22. doi:10.1145/2742488

Kaur, K., & Rai, A. K. (2014). A comparative analysis: Grid, cluster and cloud computing. *International Journal of Advanced Research in Computer and Communication Engineering, 3,* 5730-5734. Retrieved from http://www.ijarcsee.org/index.php/IJARCSEE/issue/archive

Kaur, R., & Kaur, A. (2014). A Review paper on evolution of cloud computing, its approaches and comparison with grid computing. *International Journal of Computer Science and Information Technologies*, *5,* 6060-6063. Retrieved from http://www.ijcsit.com

Keating, C. B., Fernandez, A. A., Jacobs, D. A., & Kauffmann, P. (2001). A methodology for analysis of complex sociotechnical processes. *Business Process Management Journal*, *7*(1), 33-50. doi:10.1108/14637150110383926

Khan, A. M., Büyükşahin, Ü. C., & Freitag, F. (2015). Incentive-based resource assignment and regulation for collaborative cloud services in community networks. *Journal of Computer and System Sciences*, *81*, 1479-1495. doi:10.1016/j.jcss.2014.12.023

Khan, Z., Kiani, S. L., & Soomro, K. (2014). A framework for cloud-based context-aware information services for citizens in smart cities. *Journal of Cloud Computing*, *3*(14), 1-17. doi:10.1186/s13677-014-0014-4

Khanagha, S., Volberda, H., Sidhu, J., & Oshri, I. (2013). Management innovation and adoption of emerging technologies: The case of cloud computing. European Management Review, 10(1), 51-67. doi:10.1111/emre.12004

Khanye, T., Ophoff, J., & Johnston, K. (2018). Issues in migrating legacy systems to the cloud. *In 2018 8th International Conference on Cloud Computing, Data Science & Engineering* (Confluence) (pp. 694-699). IEEE. doi:10.1109/CONFLUENCE.2018.8443029

Koehler, M. (2014). An adaptive framework for utility-based optimization of scientific applications in the cloud. *Journal of Cloud Computing*, *3*(1), 1-12. doi:10.1186/2192-113X-3-4

Kolevski, D., & Michael, K. (2015). Cloud computing data breaches a socio-technical review of literature. *In Green Computing and Internet of Things*, *2015*, 1486-1495. doi:10.1109/ICGCIoT.2015.7380702

Kornbluh, M. (2015). Combatting challenges to establishing trustworthiness in qualitative research. *Qualitative Research in Psychology*, *12*, 397-414. doi:10.1080/14780887.2015.1021941

Kumar, R., & Charu, S. (2015). An importance of using virtualization technology in cloud computing. *Global Journal of Computers & Technology, 1*(2) 56-60. Retrieved from https://globaljournals.org/papers/

Kyriakidis, M., Kant, V., Amir, S., & Dang, V. N. (2017). Understanding human performance in sociotechnical systems–Steps towards a generic framework. *Safety Science, 2017*, 1-12. doi:10.1016/j.ssci.2017.07.008

Lancaster, K. (2017). Confidentiality, anonymity, and power relations in elite interviewing: conducting qualitative policy research in a politicised domain. *International Journal of Social Research Methodology, 20*(1), 93-103. doi:10.1080/13645579.2015.1123555

Leung, L. (2015). Validity, reliability, and generalizability in qualitative research. *Journal of Family Medicine and Primary Care, 4,* 324-327. doi:10.4103/2249-4863.161306

Lewis, S. (2015). Qualitative inquiry and research design: Choosing among five approaches. *Health Promotion Practice*, *16*, 473-475. doi:10.1177/1524839915580941

Lian, J. W., Yen, D. C., & Wang, Y. T. (2014). An exploratory study to understand the critical factors affecting the decision to adopt cloud computing in Taiwan hospital. *International Journal of Information Management*, *34*(1), 28-36. doi:10.1016/j.ijinfomgt.2013.09.004

Liu, J., Mooney, H., Hull, V., Davis, S. J., Gaskell, J., Hertel, T., …Li, S. (2015). Systems integration for global sustainability. *Science, 347*, 1258832. doi:10.1126/science.1258832

"Liu, K., Peng, J., Yu, B., Liu, W., Huang, Z., and Pan, J. (2018). An instance reservation framework for cost effective services in geo-distributed data centers. *IEEE Transactions on Services Computing, 1*(1), 1-14. doi:10.1109/TSC.2018.2818121

Liu, Q. (2015). Research on strategic cost management in modern enterprises. *Management & Engineering, 21*, 206-210. Retrieved from http://www.seiofbluemountain.com/en/magazine

Lorido-Botran, T., Miguel-Alonso, J., & Lozano, J. A. (2014). A review of auto-scaling techniques for elastic applications in cloud environments. *Journal of Grid Computing*, *12*, 559-592. doi:10.1007/s10723-014-9314-7

Luftman, J., Derksen, B., Dwivedi, R., Santana, M., Zadeh, H. S., & Rigoni, E. (2015). Influential IT management trends: An international study. *Journal of Information Technology*, *30*, 293-305. doi:10.1057/jit.2015.18

Maguire, M. (2014). Socio-technical systems and interaction design– 21st-century relevance. *Applied Ergonomics, 45*(2), 162-170. doi:10.1016/j.apergo.2013.05.011

Makoza, F. (2015). Cloud computing adoption in Higher Education Institutions of Malawi: An exploratory study. International Journal of Computing & ICT Research, 9(2), 37-54. Retrieved from http://ijcir.mak.ac.ug/

Malhotra, L., Agarwal, D., & Jaiswal, A. (2014). Virtualization in cloud computing. *Journal of Information Technology Software Engineering, 4*(2), 1-3. doi:10.4172/2165-7866.1000136

Maresova, P., Sobeslav, V., & Krejcar, O. (2017). Cost–benefit analysis – evaluation model of cloud computing deployment for use in companies. *Applied Economics, 49*, 521-533, doi:10.1080/00036846.2016.1200188

Mastelic, T., Oleksiak, A., Claussen, H., Brandic, I., Pierson, J. M., & Vasilakos, A. V. (2015). Cloud computing: Survey on energy efficiency. *ACM Computing Surveys, 47*(2), 1-36. doi:10.1145/2656204

Manvi, S. S., & Shyam, G. K. (2014). Resource management for infrastructure as a service (IaaS) in cloud computing: A survey. *Journal of Network and Computer Applications, 41*, 424-440. doi:10.1016/j.jnca.2013.10.004

Mayo, R., & Perng, C. (2009). Cloud Computing Payback. IBM. Retrieved from ftp://ftp.support.lotus.com/software/uk/itsolutions/systems-management/cloud_computing_ROi.pdf

Mell, P., & Grance, T. (2014). The NIST definition of cloud computing. *National Institute of Standards and Technology (NIST).* Retrieved from http://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-145.pdf

Mogouie, K., Arani, M. G., & Shamsi, M. (2015). A novel approach for optimization auto-scaling in cloud computing environment. *International Journal of Modern Education and Computer Science, 2015*(11), 46-53. doi:10.5815/ijcnis.2015.11.05

Mohammed, M. A., Aljumaili, A. T. Y., & Salah, H. A. (2014). Investigation the role of cloud computing in the business value for optimal criteria. International Journal of Enhanced Research in Science Technology and Engineering, 3(6), 111-8. Retrieved from http://www.erpublications.com

Morse, J. M. (2015). Critical analysis of strategies for determining rigor in qualitative inquiry. *Qualitative Health Research, 25*, 1212-1222. doi:10.1177/1049732315588501

Mosa, A., & Paton, N. W. (2016). Optimizing virtual machine placement for energy and SLA in clouds using utility functions. *Journal of Cloud Computing*, *2016* (5), 1-17. doi:10.1186/s13677-016-0067-7

Mosch, M., Groß, S., & Schill, A. (2014). User-controlled resource management in federated clouds. *Journal of Cloud Computing*, *3*(10), 1-18. doi:10.1186/s13677-014-0010-8

Naeen, H. M., Zeinali, E., & Haghighat, A. T. (2018). A stochastic process-based server consolidation approach for dynamic workloads in cloud data centers. *The Journal of Supercomputing, 74*(1) 1-28. doi:10.1007/s11227-018-2431-5

Nguyen, T. H., Newby, M., & Macaulay, M. J. (2015). Information technology adoption in small business: Confirmation of a proposed framework. *Journal of Small Business Management*, *53*(1), 207-227. doi:10.1111/jsbm.12058

Noble, H., & Smith, J. (2015). Issues of validity and reliability in qualitative research. *Evidence-Based Nursing, 18*(2), 34-35. doi:10.1136/eb-2015-102054

Office for Human Protections (OHRP). (1979). The Belmont Report: Ethical principles and guidelines for the protection of human subjects of research. Retrieved from https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/index.html#xinform

Ouedraogo, M., Mignon, S., Cholez, H., Furnell, S., & Dubois, E. (2015). Security transparency: the next frontier for security research in the cloud. *Journal of Cloud Computing, 4*(12), 1-14. doi:10.1186/s13677-015-0037-5

Oliveira, T., Thomas, M., & Espadanal, M. (2014). Assessing the determinants of cloud computing adoption: An analysis of the manufacturing and services sectors. *Information & Management, 51*, 497-510. doi:10.1016/j.im.2014.03.006

Opara-Martins, J., Sahandi, R., & Tian, F. (2016). Critical analysis of vendor lock-in and its impact on cloud computing migration: a business perspective. *Journal of Cloud Computing, 5*(4), 1-18. doi:10.1186/s13677-016-0054-z

Padilla, R. S., Milton, S. K., & Johnson, L. W. (2015). Components of service value in business-to-business Cloud Computing. *Journal of Cloud Computing*, *4*(15), 1-20. doi:10.1186/s13677-015-0040-x

Pakath, R. 2015. Competing on the Cloud: A review and synthesis of potential benefits and possible pitfalls. *Journal of Organizational Computing and Electronic Commerce. 25*(1), 1-27. doi:10.1080/10919392.2015.990771

Palinkas, L. A., Horwitz, S. M., Green, C. A., Wisdom, J. P., Duan, N., & Hoagwood, K. (2015). Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. *Administration and Policy in Mental Health and Mental Health Services Research, 42,* 533-544. doi:10.1007/s10488-013-0528-y

Peinl, R., Holzschuher, F. & Pfitzer, F. J. (2016). Docker cluster management for the cloud - survey results and own solution. *Grid Computing, 14*, 265-282. doi:10.1007/s10723-016-9366-y

Paschke, A. (2016). Provalets: Component-Based mobile agents as microservices for rule-based data access, processing and analytics. *Business Information Systems Engineering, 58,* 329-340. doi:10.1007/s12599-016-0447-z

Parsons, T. (1951). The social system. The Free Press, Glencoe.

Percy, W. H., Kostere, K., & Kostere, S. (2015). Generic qualitative research in psychology. *The Qualitative Report, 20*(2), 76-85. Retrieved from http://nsuworks.nova.edu/tqr/vol20/iss2/7

Phaphoom, N., Wang, X., Samuel, S., Helmer, S., & Abrahamsson, P. (2015). A survey study on major technical barriers affecting the decision to adopt cloud services. *Journal of Systems and Software, 103*(2015), 167-181. doi:10.1016/j.jss.2015.02.002

Pichan, A., Lazarescu, M., & Soh, S. T. (2015). Cloud forensics: Technical challenges, solutions and comparative analysis. *Digital Investigation, 13*(2015), 38-57. doi:10.1016/j.diin.2015.03.002

Pietkiewicz, I., & Smith, J. A. (2014). A practical guide to using interpretative phenomenological analysis in qualitative research psychology. *Psychological Journal*, *20*, 7-14. doi:10.14691/CPPJ.20.1.7

Rana, O. (2014). The costs of cloud migration. *IEEE Cloud Computing, 1*(1), 62-65. doi:10.1109/MCC.2014.24

Rani, B. K., Rani, B. P., & Babu, A. V. (2015). Cloud computing and inter-clouds–types, topologies and research issues. *Procedia Computer Science, 50*(2015), 24-29. doi:10.1016/j.procs.2015.04.006

Raza, M. H., Adenola, A. F., Nafarieh, A., & Robertson, W. (2015). The slow adoption of cloud computing and IT workforce. *Procedia Computer Science, 52*, 1114-1119. doi:10.1016/j.procs.2015.05.128

Reddy, K. H. K., Mudali, G., & Roy, D. S. (2017). A novel coordinated resource provisioning approach for cooperative cloud market. *Journal of Cloud Computing, 6*(2017), 2-17. doi:10.1186/s13677-017-0078-z

RightScale. (2018a). State of the Cloud Report. (2018, March 23). Retrieved from https://assets.rightscale.com/uploads/pdfs/RightScale-2018-State-of-the-Cloud-Report.pdf

RightScale. (2018b): How to Set Up a Cloud Cost Optimization Process for Your Enterprise. (n.d.). Retrieved September 20, 2018, from https://www.rightscale.com/webinars/set-up-cloud-cost-optimization-process-for-enterprise

RightScale. (2018c). Tagging best practices for cloud governance and cost management overview by RightScale. (n.d.). Retrieved September 20, 2018 from http://assets.rightscale.com/uploads/pdfs/tagging-best-practices-for-cloud-governance-and-cost-management-overview-by-rightscale.pdf

Rimando, M., Brace, A. M., Namageyo-Funa, A., Parr, T. L., Sealy, D., Davis, T. L., Martinez, L. M., & Christiana, R. W. (2015). Data collection challenges and recommendations for early career researchers. *The Qualitative Report, 20*, 2025-2036. Retrieved from http://nsuworks.nova.edu/tqr

Robinson, O. C. (2014). Sampling in interview-based qualitative research: A theoretical and practical guide. *Qualitative Research in Psychology, 11*(1), 25-41. doi:10.1080/14780887.2013.801543

Rogers, E. M. (1995) *Diffusion of innovations*. Fourth Edition ed., New York, Free Press

Rowley, J. (2014). Designing and using research questionnaires. *Management Research Review, 37,* 308-330. doi:10.1108/MRR-02-2013-0027

Ryan, G. W., & Bernard, H. R. (2003). Techniques to identify themes. *Field methods, 15*(1), 85-109. doi: 10.1177/1525822X02239569

Roy, K., Zvonkovic, A., Goldberg, A., Sharp, E., & LaRossa, R. (2015). Sampling richness and qualitative integrity: Challenges for research with families. *Journal of Marriage and Family, 77*, 243-260. doi:10.1111/jomf.12147

Ruan, J., Chan, F. T., Zhu, F., Wang, X., & Yang, J. (2016). A visualization review of cloud computing algorithms in the last decade. *Sustainability, 8*(1), 1-16. doi:10.3390/su8101008

Sadiku, M. N., Musa, S. M., & Momoh, O. D. (2014). Cloud computing: opportunities and challenges. *IEEE Potentials, 33*(1), 34-36. doi:10.1109/MPOT.2013.2279684

Sartori, R., Costantini, A., Ceschi, A., & Tommasi, F. (2018). How Do You Manage Change in Organizations? Training, Development, Innovation, and Their

Relationships. *Frontiers in psychology*, *(2018)*9, 1-11 doi:10.3389/fpsyg.2018.00313

Saha, S., Sarkar, J., Dwivedi, A., Dwivedi, N., Narasimhamurthy, A. M., & Roy, R. (2016). A novel revenue optimization model to address the operation and maintenance cost of a data center. *Journal of Cloud Computing, 5*(1), 1-23. doi:10.1186/s13677-015-0050-8

Sandholm, T., & Lee, D. (2014). Notes on cloud computing principles. *Journal of Cloud Computing*, *3(*21), 1-10. doi:10.1109/CLOUD.2009.5071529

Seethamraju, R. (2015). Adoption of software as a service (SaaS) enterprise resource planning (ERP) systems in small and medium sized enterprises (SMEs). *Information Systems Frontiers, 17*, 475-492. doi:10.1007/s10796-014-9506-5

Sharkh, M. A., Shami, A., & Ouda, A. (2017). Optimal and suboptimal resource allocation techniques in cloud computing data centers. *Journal of Cloud Computing*, *6*(1), 1-17. doi:10.1186/s13677-017-0075-2

Shirvani, M. H., & Ghojoghi, A. (2018). Server consolidation schemes in cloud computing environment: a review. *European Journal of Engineering Research and Science*, 1(3), 18-24. doi:10.1016/0301-4215(96)00017-1

Singh, A., & Rajdavinder, S. B. (2017). A review on big data and cloud computing security. *International Journal of Engineering Science*, *7*, 11365-11369. Retrieved from http://ijesc.org/international-journal/downloads

Skilton, M. (2010). Building return on investment from cloud computing. Retrieved December 9, 2017, from https://www.vanharen.net/Player/eKnowledge/roi_on_cloud_computing.pdf

Smith, J., & Firth, J. (2011). Qualitative data analysis: the framework approach. *Nurse Researcher*, *18*(2), 52-62. doi: 10.7748/nr2011.01.18.2.52.c8284

Spike, J. P. (2017). Informed consent is the essence of capacity assessment. *The Journal of Law, Medicine & Ethics, 45*(1), 95-105. doi:10.1177/1073110517703103

Sultan, N. (2014). Making use of cloud computing for healthcare provision: Opportunities and challenges. *International Journal of Information Management*, *34*(2), 177-184. doi:10.1016/j.ijinfomgt.2013.12.011

Sung, S. Y., & Choi, J. N. (2014). Do organizations spend wisely on employees? Effects of training and development investments on learning and innovation in organizations. *Journal of organizational behavior*, *35*, 393-412. doi:10.1002/job.1897

Tak, B. C., Urgaonkar, B., & Sivasubramaniam, A. (2013). Cloudy with a chance of cost savings. *Parallel and Distributed Systems, IEEE Transactions on*, *24*, 1223-1233. doi:10.1109/TPDS.2012.307.

Tarhini, A., Arachchilage, N. A. G., & Abbasi, M. S. (2015). A critical review of theories and models of technology adoption and acceptance in information system research. *International Journal of Technology Diffusion, 6*(4), 58-77. doi:10.1007/978-3-319-11848-2_60

Tornatzky, L. G., Fleischer, M., & Chakrabarti, A. K. (1990). *Processes of Technological Innovation*. Lexington Books.

Trist, E. (1981). The evolution of socio-technical systems. *Occasional paper, 2*, 1-66.

Tsang, E. W. (2014). Case studies and generalization in information systems research: A critical realist perspective. *The Journal of Strategic Information Systems, 23*, 174-186. doi:10.1016/j.jsis.2013.09.002

Turner, S. F., Cardinal, L. B., & Burton, R. M. (2017). Research design for mixed methods. *Organizational Research Methods, 20*, 243-267. doi:10.1177/1094428115610808

Vithayathil, J. (2018). Will cloud computing make the I nformation T echnology (IT) d epartment obsolete? *Information Systems Journal, 28*(4), 634-649. doi:10.1111/isj.12151

Von Bertalanffy, L. (1972). *General System Theory: Foundations, Development, Applications: Rev. Ed*.

Walby, K., Marshall, C., Rossman, G. B. (2015). Designing Qualitative Research. *Canadian Journal of Sociology, 40*, 399-402. doi:10.17169/fqs-9.3.996

Walker, R. C., Morton, R. L., Tong, A., Marshall, M. R., Palmer, S., & Howard, K. (2015). Patient and caregiver preferences for home dialysis—the home first study: a protocol for qualitative interviews and discrete choice experiments. *BMJ Open, 5*(4), e007405. doi:10.1136/bmjopen-2014-007405

Ward, J. S., & Barker, A. (2014). Observing the clouds: a survey and taxonomy of cloud monitoring. *Journal of Cloud Computing, 3*(12), 1-30. doi:10.1186/s13677-014-

0024-2

Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS quarterly*, xiii-xxiii. doi:10.2307/4132319

Weingärtner, R., Bräscher, G. B., & Westphall, C. B. (2015). Cloud resource management: A survey on forecasting and profiling models. *Journal of Network and Computer Applications, 47*, 99-106. doi:10.1016/j.jnca.2014.09.018

Weintraub, E., & Cohen, Y. (2015). Cost Optimization of Cloud Computing Services in a Networked Environment. *International Journal of Advanced Computer Science and Applications, 6*(4), 148-157. Retrieved from http://thesai.org/Publications/Archives?code=IJACSA

Whaiduzzaman, M., Haque, M. N., Rejaul K. C., M., & Gani, A. (2014). A study on strategic provisioning of cloud computing services. *The Scientific World Journal, 2014*(1), 1-16. doi.org/10.1155/2014/894362

Williams, D. R., Thomond, P., & Mackenzie, I. (2014). The greenhouse gas abatement potential of enterprise cloud computing. *Environmental Modelling & Software, 56*(2014), 6-12. doi:10.1016/j.envsoft.2013.11.012

Wilson, V. (2014). Research methods: Triangulation. *Evidence Based Library and Information Practice, 9*(1), 74-75. doi:10.18438/B8WW3X

Yeboah-Boateng, E. O., & Essandoh, K. A. (2014). Factors influencing the adoption of cloud computing by small and medium enterprises in developing economies. *International Journal of Emerging Science and Engineering, 2*(4), 13-20. Retrieved from https://www.ijese.org/archive/

Yilmaz, K. (2013). Comparison of quantitative and qualitative research traditions: Epistemological, theoretical, and methodological differences. *European Journal of Education*, *48*, 311-325. doi:10.1111/ejed.12014

Yin, R. (2017). *Case study research: Design and methods*. Thousand Oaks, CA: Sage Publications, Inc.

Yost, J., Dobbins, M., Traynor, R., DeCorby, K., Workentine, S., & Greco, L. (2014). Tools to support evidence-informed public health decision making. *BMC Public Health*, *14*(1), 1-13. doi:10.1186/1471-2458-14-728

Zhang, R. 2018. The impacts of cloud computing architecture on cloud service performance. *Journal of Computer Information Systems*, *2018*(1), 1-9. doi:10.1080/08874417.2018.1429957

Zhang, S., Yang, J., Shi, Y., Wu, X., & Ran, Y. (2015). Dynamic energy storage control for reducing electricity cost in data centers. *Mathematical Problems in Engineering, 2015*, 1-13. doi:10.1155/2015/380926

Appendix A: Interview Protocol

A: Preinterview Protocol

1. Obtain IRB approval to conduct interviews
2. Work with CIO/CTO to send introductory email I create to potential pool of participants
3. Respond to employees who indicate interest to participate in study providing details on researcher/role, more information on project, and consent form.
4. Work directly with participants who return consent form to identify availability for interview (date/time)
5. Schedule interviews – quiet location onsite, away from distraction
6. Send separate email reminder to each participant 48 hours prior to their interview time
7. Complete other logistics to be on site and ready for interviews.

B: Day of Interview Protocol

8. Arrive early at location, ready to start with interviews.
9. At time of interview with a participant
10. Re-Introduce myself, and the objective of the study and meeting

a. Ask confirm with participant if they still want to participate
b. If no, thank participant for the consideration and terminate interview
c. If yes, continue to 5
d. Thank participant again for their participation, and review flow of activities
e. Remind them of recording, and why
f. Remind them of member checking transcripts review processes and why
g. Start recorder and commence with interview questions and data collection
h. At end of interview, inquire about other sources of data, and arrange with participant to send them to interviewer for review (Data Triangulation)
i. Summarize next steps and what to expect (2 Stage member checking)

Appendix B: Interview Questions

The overarching research question is what strategies and processes IT managers use for managing IT infrastructure operational costs in cloud computing.

1. What strategies have you used for managing IT infrastructure operational costs in cloud computing?
2. What processes did you find worked the best for developing and implementing IT infrastructure operational costs in cloud computing?
3. What tools did you find to be effective to use for managing IT operational costs?
4. What types of staff training, if any, did you find to be effective for managing IT operational costs in cloud computing?
5. What organizational changes, if any, did you find to be effective for managing IT infrastructure operational costs in cloud computing?
6. What ongoing adjustments to IT operational procedures have you found to be necessary for managing costs on a month-to-month or an ongoing basis?
7. What else would you like to add that we have not yet addressed?

Appendix C: STS Design Principles

The STS design principles according to Chern (1987) are:

(i) Design should be compatibility with organizational objectives—A capability of the design to accommodate self-modification through the creative abilities of the individual. People are extensions of the systems (Trist, 1981)

(ii) Employ minimal critical specifications—Some systems specifications are necessary, but their specification should not be too precise to mandate the specifics on the performance of the related tasks.

(iii) Sociotechnical criterion control—Variances in output must be controlled/reconciled as close to the point where they occur as possible. In other words, the integration of inspection/resolution into tasks in the workstream.

(iv) Capitalize on the multifunction nature of organism and mechanism – The design must support different ways to accomplish the same outcome using a different combination of elements to increase system adaptability and equifinality.

(v) Ensure boundary locations do not impede the desirable sharing of information, knowledge, experience, as well as the efficient transition of activities between groups or departments

(vi) Information systems should provide the right type and amount of feedback directly to the teams who need to act so that they can use the information to improve outputs

(vii) Systems of social support should reinforce the types of behaviors that the organization structure desires to promote. These systems include the system of selection, training, remuneration, and conflict resolution.

(viii) Design with human values in mind—The work design should focus on providing quality work, but that quality is subjective, so the characteristics of a good job (as defined by Thorsrud (1972) should guide work design.

(ix) Work design is always an incomplete process – The incompletion of design is a result of the reiterative nature of the process, requiring a redesign soon after one implementation completes.