UNIVERSITY OF NAIROBI
SCHOOL OF COMPUTING AND INFORMATICS

UTILIZATION OF CLOUD DATABASE ON CRIMINAL BACKGROUND CHECK

MICHAEL NDONYE MUTISYA
P53/79214/2015

SUPERVISOR
DR. ANDREW KAHONGE

NOVEMBER 2017

A project report submitted to the School of Computing and Informatics in partial fulfillment of the requirements for the award of the degree of Master of Science in Distributed Computing Technology of the University of Nairobi

Declaration

This research project is my original work and has not been presented for a degree or any other award in any other educational institution.

Signature: _____   Date: _____

Michael Mutisya

P53/79214/2015

Approval:

This research project has been submitted for examination with my approval as the university supervisor.

Signature: _____   Date: _____

Dr. Andrew Kahonge

School of Computing and Informatics

University of Nairobi

Dedication

Specially dedicated to Almighty God, Happy my daughter, beloved Brothers Sammy and Juspar

Acknowledgement

I wish to acknowledge various people who have thrown in immensely to the accomplishment of this research project and my academic journey. First and foremost, I thank my supervisor Dr. Andre Kahonge, who has been on the forefront in stimulating new ideas and providing guidance in this research. Secondly, I give special gratitude to Prof Kimathi Kigatiira, Dr. Lucy Kibe and Charles Karuga who always encouraged me to believe in myself and overcome all hurdles in academic realm. Lastly but not the least, my beloved family for the moral support they have accorded me.

Abstract

The rate of Criminal background check has recently gone high in Kenya. Career candidates have been subjected to pre-screening they are engaged by employers. Employers have been slack in hiring or retaining employees who have been having defective history and who subsequently engage in violence.

Pre-screening has also been conducted whenever a person has committed a crime. This gives the CID detectives a substantial criminal history of the person who has been implicated in a crime.

Criminal history records consist of identifiable descriptions and notations, detention and indictments or other formal criminal charges and any disposition arising there from, including acquittal, sentencing, correctional supervision or release

In Kenya, the responsibility of issuance of police clearance certificate and investigation of complex charges is bestowed on the Criminal Investigation Department of the Kenya Police.

Emergence cloud computing service models has made computing resources available to the consumers the internet connections. This study focuses mainly on application of Cloud Database model in enhancing progression in screening industry. This model provides for access of computing resources when needed and allows network entrance to a common pool of networks, servers, storage space, applications, and services, that can be swiftly offered and unrestricted, with little clients endeavor or interaction by provider of the service.

Utilization of cloud database will trigger for migration of application by the CID hence minimizing the cost of infrastructure and technical expertise. These applications will be accessed from every CID office countrywide on demand.

Table of Contents

List of Figures

List of tables

DEFINITION OF TERMS

DBaaS- Database as a Service

NGO's- Non Governmental Organizations

NIST - National Institute of Standards and Technology

SLA- Service-level agreements

CID- Criminal Investigation Department

VMs- Virtual machines

CJIS- Criminal Justice Information Services

NGI- Next Generation Identification

APFIS - Automated Palms and Fingerprints Integrated System

Definition of Terms

Database   A collection of related information that is organized so
                                    for ease retrieval, access, management and updation.

Database as a Service (DBaaS)   an approach of cloud that enables the storage and
                                    administration of ordered data.

Cloud database   This is a content database that has an ability to scale up or
                                    down and which typically runs on a cloud computing
                                    platform.

Cloud computing   This is a model of computing in which computational
                                    resources are provisioned as services and availed to the
                                    clients over the Internet

Virtualization   Virtualization is the creation of a near version of
                                    something. Eg operating system

CHAPTER ONE

INTRODUCTION

1.1 Background of the study

There has been great progression in screening industry mostly the use of criminal background checks. Career candidates have been undergoing pre-screening before they are offered employment. Employers have been slack in hiring or retaining employees who have been having defective history and who subsequently engage in violence. Pre-screening has also been extended to tenants, vendors and volunteers. At several circumstances citizens might be asked to produce a criminal background check. Nowadays, companies, government sectors along with NGO's demand Police clearance certificate during vetting of prospective employees or when applying for a visa. Companies who wish to place tenders for the supply of goods are also requested to produce the certificate from the director of companies. Criminal background records provide substantial information for an individual who is under investigation by the Criminal Investigation Department.

Criminal history record is information collected by criminal justice agencies on individual consisting of identifiable descriptions and notations, detention and indictments or other formal criminal charges and any disposition arising there from, including acquittal, sentencing, correctional supervision or release (Jacobs et al., 2007). Criminal records are rapidly becoming negative curriculum vitae (Jacobs, 2005) used to determine eligibility for occupational licenses, social welfare benefits employment and housing (Pager, 2003). In Kenya, the certificate is basically a look up for persons criminal records stored by the CID and is usually known as Police clearance certificate. The certificate serves as prove that the holder has or has not been associated with any criminal act, civil or administrative criminal offenses in the preceding 6 months and the verdict of the case. According to Allen (2002), the criminal background check ensures that reports on a much broader range of crimes and data on the circumstances of the crime are included. Even though this check serves a good purpose, it is limited in sense that it covers previous activities from the time of application backwards. Brody (2010) argues that one commonly used assessment tool which is a standard practice among companies for screening potential employees is background checks; however they do not always unearth the entire truth because there are several practical limitations associated with these searches

A set of fingerprints is generally mandatory to obtain the certificate. The applicants entire set of rolled and plain fingerprints are appended to printed forms by the staff of Kenya Police ready for search. Emergence of new technologies has spearheaded adoption of better systems that are able to solve real world problems. The solution to the most challenges facing the criminal background check is cloud database.Wyld (2010) notes that Cloud computing is an emerging concept. He further states that the basic idea of cloud computing is that computing will become location- and device-independent- meaning that it increasingly will not matter where information is housed nor where computation or processing is taking place. This enables computing tasks and information to be available anytime, anywhere from any device- so long as there is access to the internet.

Cloud Computing paradigm of computing in which computational resources are availed as services to the users over the internet (Buyya, 2011).

The National Institute of Standards and Technology (NIST, 2011) defines cloud computing as "model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction"

In this model computing resources are delivered through the internet. As Buyya et al., (2011) stipulates, "Cloud is a parallel and distributed computing system consisting of a collection of inter-connected and virtualized computers that are dynamically provisioned and presented as one or more unified computing resources based on service-level agreements (SLA) established through negotiation between the service provider and consumers."

In this paradigm the use of computational resources is determined by the clients' demand and has transparent scalability where these resources are allocated in a changing and precise manner when they are strictly on demand, devoid of the client understanding the supporting infrastructure or the resources origin

Cloud database provide considerable prospective gain to enforcers of law and government agencies. The tangible benefits of cloud computing in criminal records check are: - Cost savings, dynamic allocation of new and additional resources when necessary, offsite storage and disaster recovery.

Recent growth of data gathering by law enforcement authorities through, for instance, the use of sensor devices and CCTV cameras, only compel to put more emphasis on the requirement for

clearly put together policies regarding cloud-based data storage. The data capacity has been mounting in the database. It is therefore foreseeable that most of that data will be stocked up and processed by the use of cloud services.

To meet the dynamic operational needs, while maintaining the security of systems and data, law enforcement agencies can contemplate on the use of cloud computing services to administer criminal background check.

1.2 problem statement

Criminal background check has become a major concern for employers, organizations or associations. Human beings are not prone to crimes and therefore it is necessary to know of any other formal criminal charges and any disposition arising there from, including acquittal, sentencing, correctional supervision or release before engaging someone in employment, association or even when carrying out criminal investigations. Criminal background check has been ever increasing. The current system allows applicants of Police Clearance Certificate to open e-Citizen account through which they can apply for the certificate. This system also allows the applicants to pay processing fee online via credit and debit cards or through mobile payment platforms like Mpesa, Airtel money etc. The applicant downloads and generates a printout of duplicates of their invoice, a duplicate of C24 under two faces on an A4 size paper and then take the C24 together with the invoices with an original National Identity Card for those above 18 years or Original Birth Certificate for minors, to the nearest Huduma Center or Directorate of Criminal Investigations Department headquarters for the purpose of finger print processing. If the finger scanning is done at Huduma centers, the filled forms are transported to the Directorate of criminal investigation for search. Foreign citizens are required to submit original copies of their passport and a documentary proof of being in the country for six months or more. After fingerprint processing, the officers carry out background check and then send the Police Clearance Certificate to the applicant for download. This takes an average of two weeks. Criminal background check by Directorate of Criminal Investigations Department of Kenya is rocked by several challenges as far as processing and issuing of police clearance certificate is concerned.

With the increased number of police clearance certificate applications, the current system is handling increased workload thus taking much time to process these certificates. The government agency has centralized system of processing police clearance certificate and this

poises threats of possible breakdown due to prolonged service and heavy workload. Currently the system receives about 1000 applications per day with the number bound to increase because of increasing demand from employers and government agencies. There is therefore a need for migrating to another platform that will address this issue at a low cost and at efficient way. Cloud database paradigm has proved to address and offer a relief to such challenges at a low cost. Migrating current application database to the cloud will boost efficiency and flexibility in a way that can support a wide range of on-demand requirements and enable citizens to obtain Police Clearance Certificate from their nearest CID office countrywide.

Of recent past, the applications of police clearance certificate in CID offices in various offices countrywide had to be transported to CID headquarters after collection. With long distances, this system was very costly and time consuming, with the applicants waiting for an average of one to two weeks before they receive the police clearance certificate. Despite some improvements on the system where applicants can apply the Police clearance certificate online through e-citizen, they are required to present themselves personally at the Directorate of criminal investigations for fingerprint scanning. Consequently, waiting time has not changed but still remains an average of one to two weeks as the system still remains centralized. After fingerprinting, the applicants are nowadays able to download the Police clearance certificate online.

When a crime has taken place from any part of the country and some suspects are held responsible, access of the suspect's background information for the investigations have to be referred to the CID headquarters. This makes the process a bit slow and this may cause tampering with the crime scene.

Another issue is that there is little academic reference or research that has been done on the application of technology on criminal background check in Kenya.

1.3 Research Objectives

Main Objective

- The overall objective of this study is to develop a prototype that will provide solution to effective Criminal background check in the background screening industry by utilization of cloud database.

Specific Objectives

1) To design and develop a prototype for Criminal Background Check under cloud database environment with utilization of database nodes.

2) Review the methods used for verification of police clearance certificate in current systems

3) To investigate and evaluate Biometric Recognition technology used in Policing and Criminal Background Check

1.4 Research questions

This research study is aimed at answering the following questions: -

1. What method may be used to develop and offer Criminal Background Check through cloud database?

2. Which methods are currently used to verify police clearance certificate?

3. Which Biometric Recognition technology is used in Policing and Criminal Background Check?

1.5 Scope of the study

This study will major on database as a service covering handling of distributed queries and node splitting and reporting on criminal background information. The prototype will not cover other cloud storage services like iCloud, drobox. The study will be limited to the Criminal Investigation Department since they are the ones involved in handling of criminal background records.

1.6 Justification

Criminal Investigation Department, being a Government agency, is guided by the current Kenya Vision2030 project which puts much emphasis on technology development by using Information Technology. This is aimed at making work easier and manageable as there is a tremendous increase in data volume. The system will enhance service delivery by the Criminal Investigation Department in issuance of police clearance certificates and boost investigations on individuals who may be involved in a crime. The access of individuals background information by the agency will not be limited to time and geographical site. With this in place, applications will not need to be transported to CID headquarters for processing but will be processed from the applicants nearest CID offices.

The authorized staff will be able to store and access encrypted background information from the police datacenters over the web. This will enable the investigating officers to acquire information from both datacenters and produce federated information about a person seeking to be issued with police clearance certificate and the persons under investigation. Backup will not be a necessity at every local CID office as there will be a repository of background information where is can be accessed and shared by all CID offices countrywide. This repository keeps on growing as more content is uploaded (Ouf et al.,2011). In event of clients' computer crash there will be no data loss as everything will be stored in the repository. (Pocatilu et al., 2010).

1.7 Assumptions of the study

i) The CID officers and detectives will be having knowledge on the basic operation of a computer system
ii) Every CID offices county wide will be having internet connection
iii) Criminal records are for public consumption and not private data
iv) If there is a miss of criminal record in the system and the applicant had not applied before revokes a manual search.
v) All the new criminal records entered into the system will be using biometrics as finger scanning technique.

1.8 Limitations of the study

i) A cloud database phenomenon is relatively new in Kenya and literature baseline around the area of cloud database is still limited.
ii) The questionnaire to be administered maybe somewhat technical for the Non-Technical staff and this might have affect the accuracy of the data to be collected in one way or another.
iii) The small sample size may relatively make the data statistically unreliable

CHAPTER TWO

LITERATURE REVIEW

2.0 Introduction

This chapter reviews literature of related perceptions by various researchers. This chapter opens with the preface of the key concepts of cloud computing paradigm and also discusses database as a service (DBaaS) model. Further, I will also discuss the utilization of Cloud database in Criminal Background Check.

2.1 Cloud computing Concepts

"Cloud Computing is a computation paradigm in which the resources of an IT system are offered as services, available to the users through net connections, frequently the Internet" (Buyya, 2011).

The National Institute of Standards and Technology (NIST, 2011) define cloud computing as "model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction". This technology is budding as a new computing paradigm where computational power is provisioned dynamically through the internet. This technology has seen many organizations and enterprises migrate their applications in order to cut down cost as the technology is grounded on pay-as-you-go model where services are provided on demand (Mell and Grance., 2009). Services are provisioned in a similar manner to commodities like water, electricity, telephony etc. In this model, clients concern is not the origin or the source of the computing resources and how they are delivered (Buyya et al., 2009). The main interest is acquisition of quality service from the service provider and this is achieved through Service Level Agreements(SLAs), According to this publication businesses and users are able to access applications from anywhere in the world on demand.

Many computing service providers are deploying data centers for hosting cloud computing applications in various locations around the world to deliver computing services. This is meant to ensure reliability and continuous provision of services in case of site failures. These providers include Amazon, Google, Yahoo, Microsoft and IBM. Content providers have to ensure they are flexible in-service delivery because service requirements vary from one consumer to another.

This allows content consumers to balance infrastructure to cut down on investments. Cloud computing allow for dynamic scaling up and down of computing resources as demands fluctuate (Ercan, 2010).

Recent advances in microprocessor power and software has led to cloud computing where hardware of a commodity can run applications within Virtualized environment. Buyya et al., (2009) asserts that Providers can expose applications running within VMs, or provide access to VMs themselves as a service (e.g. Amazon Elastic Compute Cloud) thereby allowing consumers to install their own applications. This raises new challenges with the allocation of resources within consumers competing resource demands.

According to Boss et al., (2007), a cloud is a collection of virtualized computing resources that accommodate a variety of diverse workloads and assigns them through the provisioning of physical or virtual machines.

The merits of cloud computing have been much discussed. According to a report by IDC (2011), cloud-based IT spending is going to rise from its 2011 level of less than 2% of IT spending to nearly 20% by 2015, which means about 10% of all information could be stored within a cloud database by this time

2.2 Database As a service (DBaaS)

Cloud database is frequently utilized as a service although it is less talked about compared to the above-mentioned services. According to Curino et al. (2011), this technology is presented as a new transactional "database-as-a-service" (DBaaS) called Relational Cloud.

Cloud computing is becoming widely accepted technology in the current times, and the database application has also migrated to cloud computing environment now. According to Waleed (2013) most enterprises will employ usage of cloud database to store and access huge data. With similar remarks, Gartner (2011) stipulates that cloud computing service is among top 10 strategic technologies for 2012 with the potential for creating a significant impact on business enterprises in the foreseeable years. DBaaS promises to migrate most of the operational problem of provisioning, configuration, scaling, performance tuning, backup, privacy, and access control from the database users to the service operator, offering lower overall costs to users (Curino et al., 2011)

Many corporations and e-commerce companies have already adopted Database as a service and they are currently reaping full benefits from this service.

Clients access the database on demand through the internet from the database service provider hence cloud database is designed for virtualized computer environment. Database application can also be accessed by people with mobile phones that are powered by 3G or 4G services (Pizzete and Cabot 2012). Cloud computing enables implementation of cloud database where the cloud computing service provider allows consumers utilize the hardware and software resources.

Numerous companies have begun migrating their applications to cloud computing where they get entrance to their data from cloud database. In a recently conducted survey, 36 percent of the companies are running their applications through cloud services (Mimecast Survey, 2011). Cloud computing is rising very fast in the IT industry globally. This trend clearly signals that soon and very soon, corporations will begin to be dependent on cloud applications.

The Vendors should ensure that database is accessible round the clock so that the service consumers can get the data as and when they need it. The database should also be managed with ease and at the same time should cut down on cost (Curino et al., 2011). In Cloud, information recovery is extremely efficient after a disaster in the database. In event of clients' computer crash there is no data loss as all data is deposited in the cloud database. (Pocatilu et al., 2010). Backup is not a necessity as users can create a repository of information where is can be accessed and shared by all authorized parties.

2.3 Utilization of cloud database in developed countries

Microsoft has recruited a number of law enforcement agencies to store criminal records, fingerprints, and video in its data centers.

Every time an armed robbery takes place in Los Angeles, the surveillance cameras captures and records the break-in and that video footage is send to the cloud database.

Microsoft in 2016 reported that a number of agencies that enforce the law including the San Bernardino County Sheriff's Department, and the Charlotte-Mecklenburg Police Department in North Carolina are good examples of law enforcement authorities that are employing usage of the company's cloud data centers to accumulate law enforcement record.

According to the report, law enforcement agencies are making use of Microsoft Azure to store surveillance video, fingerprint data, criminal record and even the footage from cameras implanted on police bodies.

Once the video gets onto the Azure platform it is aggregated and examined and makes database proactive to display information. The Oakland police search from the cloud database for specific

keywords and retrieve the audio or video of interest. Criminal records are critical and therefore Law enforcement agencies can only trust and store sensitive data with vendors that observe the FBI's Criminal Justice Information Services (CJIS) Security Policy. Microsoft (2015) argues that it complies with the security policies, which includes stringent security standards, and regular audits that are required of cloud-service provider's.

Microsoft is not the only cloud computing provider that has been on the forefront in sensitizing government agencies to use cloud storage but Amazon Web Services have also developed a special private cloud infrastructure for the CIA. This project is believed to have cost Amazon more than $600 million.

The FBI has taken a momentous milestone by making it clear that cloud services are tolerable, while also clarifying that conventional commercial cloud computing service models is not implementable in the CJIS environment. Under the new policy, criminal justice information may be processed using cloud services, but the service provider must ensure that many of its metadata analytical functions are disabled.

The FBI policy supports philosophy newly developed by the International Association of Chiefs of Police (Ponemon, 2013) and fittingly places the equilibrium between data security, convenience and efficiency. The strategy recognizes the prospective payback and security test offered by the technology, while offering the agencies of law enforcement with resources and equipment that can aid them in deployment of a CJIS compliant cloud.

2.4 Related work on law enforcement using cloud

Cloud computing has been used by law enforcing agencies in crime reporting and storage. According to a report from the International Association of Chiefs of Police there is need for a set of steering principles for cloud computing by enforcers of law.

Developments in technologies of cloud computing can deliver a number of benefits for both government agencies and law enforcers, including saving of cost and speedy provisioning of critical resources. One of the principles formulated was meant to compel the provider to ensure that law enforcement data maintained by the providers is compatible and transferable to other systems and can integrate with other operating systems with no breach on security and data integrity. The provider of the cloud service has an obligation to enforce all needed technical, administrative, physical and procedural measures to guard the confidentiality of law enforcement data. These steps include physical protection actions, permissions to access data, cyber security measures, and geographical location limitations among other key requirements.

A survey of 272 officials, most of them chief executives of police or sheriff departments (Ponemon et al., 2013) reveals that 46 percent of law enforcement agencies are not considering utilizing cloud computing. More than half of respondents are making use of or considering to make use of cloud computing with sixteen percent of the responding agencies already using the Cloud, and 38 percent are considering or planning to use it within the next 2 years

2.4.1 Applications that use Cloud
The survey further shows that among the agencies that have adopted cloud computing, email is the most popular application with 17 percent followed by cloud storage with 15 percent of respondents. 11 percent confirmed to be accessing criminal justice information system while crime reporting, analysis, and mapping had weight of 10 percent.

2.4.2 Plans or Consideration to implement Cloud in the Next 2 Years
However, within the next 2 years, respondents are planning or considering using Cloud for criminal justice information system access (51 percent); data storage (50 percent); RMS, crime reporting, analysis, and mapping (47 percent); and email (46 percent). This shows that cloud is highly preferred for information access and data storage.

2.4.3 Factors leading to cloud usage
According to the survey, when asked why they were going to use cloud, it turned out that 61 percent of the respondents cited the reason as saving on money; acquisition of software and hardware is not a necessity (52 percent); dynamic provisioning (39 percent); replacing legacy applications (34 percent); new features (33 percent); easier operation for end-users (19 percent); enhanced security (16 percent); improved technical support (15 percent); payment per usage (5 percent); and political mandate (5 percent).
The survey further revels that agencies that utilize cloud computing tend to be above average. In a sample size of 216 sworn officers per agency, departments utilizing the cloud comprised an average of 335 officers.

2.5 Utilization of cloud database in developing countries

The developing counties are yet to fully embrace cloud computing technology. This places cloud computing at an infant stage in the developing world (Nir, 2010). According to a Garner survey (2009) carried out on large enterprises, nearly half the respondents in rising markets either had not heard of cloud computing model or didn't have an idea of what it meant. The call center industry in South Africa has been a most increasing area for the cloud (Nir, 2010).

Theoretically, it is achievable for developing economies to be at par with the West as the cloud equalizes them in IT infrastructure, data centers and applications. The use of cloud cuts down the cost of infrastructure and makes playing field for business enterprises to be at the same level (Nir, 2010).

2.6 Utilization of cloud database in Criminal Background Check

With the changing trends in technology, emergence of cloud and its utilization law enforcement agencies can pursue smarter policing having one huge potential benefit on the horizon. Cloud computing can minimize up-front investment and ongoing costs for IT systems and applications, makes sense in this era of fiscal austerity. With some law enforcement agencies utilizing cloud database to securely store criminal data, it is possible for criminal Investigation Department to catch up with the FBI and CIA as the cloud allows them to have access to the same IT infrastructure, data centers and applications. This will enable the department to efficiently serve all citizens from every corner of the country. Like other government agencies, police hope they can save some money and get rid of legacy hardware and software by using the cloud. With cloud database storage of criminal records on the cloud will help in disaster recovery and backup, crime analysis and records management. The cloud opens up sophisticated technology tools and services to smaller agencies that don't have the funds to purchase an entire application on their own. Cloud database will give a leg up to the CID department as it would follow that the small police department also could move quickly as far as using the latest digital productivity tools without big up-front costs. What is holding back many police departments whether large or small are security concern. The cloud providers therefore should meet stringent security standards for data sharing on the network.

2.7 Biometric Recognition technology in Policing

Human are basically identified based on common methods such as credentials like documents and personal identification number (PIN). Due to increasing demand in screening industry, these methods have been unable to cope with stringent security check in applications such as National IDS and cards (Jain et al., 2012). As technology advances, new methods of identification are emerging especially that uses human features commonly known as biometrics. Biometrics refers to metrics used to verify the identity of individual based on unique human characteristics. Some

of the techniques in place today are eye recognition, fingerprint scan, voice recognition and facial recognition.

Efforts have been made by the Federal Bureau of investigation to develop a new biometrics database since 2008, a program called Next Generation Identification (NGI). This database is the largest in the world housing 125 million criminal and civilian fingerprints and 24 million mug shots. This system will be able to analyze the tattoos, scars to the voice and eyes of the suspect. (FBI, 2015). According to EPIC (2017), NGI system will store biometrics of millions of individuals who are neither criminals nor suspects without being aware of whether their biometrics are captured by CCTV surveillance cameras. According to this publication, the number of surveillance cameras in the United States is estimated to be 30 Million. These cameras will be integrated with NGI system for both law enforcement and non-law enforcement. The benefits of retaining civil fingerprints together with criminal data provides continuous background check in order to allow employers, interested parties and other authorized individuals to monitor criminal conduct of potential employers and trusted individuals (FBI 2015).

The Directorate of criminal Investigation uses APFIS to process fingerprint scans. This is an integrated database that has two other databases i.e. Manual database that contains files of the people who have been convicted and Integrated/Automated database that store soft copy of all criminal records.

This system stores soft copy of fingerprints taken from Police station once a person has been arrested.

When an applicant delivers the downloaded form from the e-Citizen, fingerprints are taken on the hard copy and the application is send to APFIS. In APFIS, before search start, they compare fingerprints on the hard copy taken from the applicant with the fingerprints of the thump prints taken when the applicant applied for the National ID. Mostly the Right and the Left thump fingerprints are used for comparison

This process of comparing fingerprints is done manually using magnifying lenses. After the process, they confirm that the document has been certified before taken to the next phase. Here, they fill the bio data from the application form into the system

After feeding the data they run APFIS scanner that scans and send the fingerprint to the database for comparison with the criminal database that stores all criminal records. This compares the two to see if they match for every individual. If the fingerprints match that's a hit. Meaning the

applicant has criminal records in the database. This record is then printed on the certificate and then uploaded on the e-Citizen for downloading by the applicant. This process of Biometric verification is limited to fingerprints scans. Elsewhere biometric technologies are being used at local level by police departments for public oversight. According a publication by Caplan et al. (2015), the center for department of investigation and the electronic Frontier foundation has shown how US military and intelligence agencies have used facial recognition software in Iraq and Afghanistan to identify suspected criminals in real time by capturing photos and sending them to criminal database for search. Use of biometrics plays a major role in laws enforcement and fighting of terrorist threats.

2.8 Related Work on Verifications

Dictionary.com defines verification as "process of research, examination, etc., required to prove or establish authenticity or validity; Evidence that establishes or confirms the accuracy or truth of something". Research on verification of credentials via the cloud has been carried out by several researchers. Musee (2015), for instance, researched on academic verification system using cloud environment. According to this research, most organizations and institutions don't have instant verification technique but purely relies of traditional paper verification technique. He further asserts that most people especially employers are facing it tough when verifying academic credentials presented to them as they do not have an instant solution for verifying the authenticity and validity of these credentials. Lack of verification of credentials can be costly for an organization. According to estimates, employers lose approximately $600 billion annually from resume fraud (Musee, 2015). This study further shows that employers must have a mechanism to independently check the background of prospective employees and that the employees shun hiring those with defective criminal background. In a separate study by Muthoni (2015), selection of candidates from a list that is not verified leads to fake certificates passing on hence recruitment of employees who may not have qualified for the job. This demoralizes the candidates who have qualified and also employers who believe in hiring the candidates who have met some specific threshold. Verification of credentials can improve on efficiency and handle issues on authenticity of these credentials. Criminal data can be shared between the investigating agency and the employers in order to ensure they don't recruit employees with defective history and the same time spending minimal effort and resources.

Verification through the cloud using SMS gateway can instantly give details on whether a police certificate is authentic or not.

Police clearance certificate is not only required when one is seeking employment but also required when one is vying for elective posts in the government of Kenya. The constitution of Kenya demands aspirants for various elective positions must be vetted to ascertain that only those with integrity and upright morally assume public offices. The electoral body (IEBC) is tasked to carry out vetting of all aspirants and certificate of good conduct is mandatory for all elective positions from the presidency to the member of county assembly. Most recently some of the aspirants have leveled blame against the Criminal Directorate for frustrating their efforts to vie for various positions by delaying to release the police clearance certificate on time thus jeopardizing their eligibility for failure to beat the set deadline. For verification of academic credentials, academic institutions, academic certification bodies, government and employers are provided with interface to query its authenticity of these credentials (Muthoni,2015). Screening of credentials gives a true picture of an individual. According to Otieno (2013), when hiring someone, the employer should be sure to screen their papers thoroughly, lest they earn themselves a fake friend, wearing a fake smile to impress the employers.

Much of verification is done manually. Employers usually look at the hologram, watermarks, seals and other security features that either qualify or disqualify a certificate. Human resource managers utilize background information by contacting referees to have an insight on the integrity of a potential employee (Mukami, 2014), but this process becomes difficult when the potential employee colludes with the referee (Muthoni,2015).

2.9 Structure of Cloud Database

When classifying DBMSs, there is on one side traditional relational databases (RDBMSs) that provide low latency and high throughput transaction processing at a single node, but lack features such as elasticity and scale-out. Some recently proposed cloud enabled relational databases target scale-out but do not consider lightweight elastic scaling as a first priority. An example of such is Cloud SQL Server (Bernstein et al., 2011a) and Relational Cloud (Curino et al., 2011). On the other side, key-value stores, such as Bigtable (Chang et al. 2006) PNUTS (Cooper et al. 2008), and Dynamo (DeCandia et al. 2007), support scalability, elasticity, and are highly available, but only support simple functionality such as key-based lookups and single-row transactions.

The cloud database uses data centers distributed at various localities to store data. This aspect draws a disparity between the cloud database structure and the normal database management system hence leading to a complex structure of the cloud database. Cloud database has multiple

nodes that services queries meant for data centers that are dispersed in diverse geographical locations as well as enterprise data centers. For comprehensive access of the cloud database services this joining is obligatory. Waleed (2013) demonstrates the structure of cloud database through Business Intelligence (BI) application used for storing huge customers data by the enterprise. The customers usually access the cloud database via the internet which acts as conduit between the data center, cloud databases and the client querying the database.

Cloud database is not composed of a single node but has different nodes are in place (Curino et al., 2013). Due to this reason, peer-to-peer interactions are favored. The reason behind implementation of peer-to-peer communication is that, a sole node can service any kind of request generated by the application user. The goal of having different node is to lower the number of cross-node distributed transactions, which leads to overhead both because of the additional work performed on each node and because of the increase in the time consumed during holding of locks at the back-end. With this complexity, an easy remedy for this sort of node system would be to have map pointing the stored data for every node in the cloud database. This data pointer helps in replicating and makes access of data easy for every issued query.

When a user executes a query through their computer, the head node decides the type of the query, determines the most suitable node to handle the query and then forwards the query to that particular node. The assigned node services the query and returns an answer to the client. For instance, when node 1 receives the query, it checks the data map and determines the node that holds the data e.g. Node 8 and then forwards the query to that node for servicing. Node 8 retrieves the data and responds to the user. The below figure demonstrates the basic architecture of the cloud database by Waleed (2013)
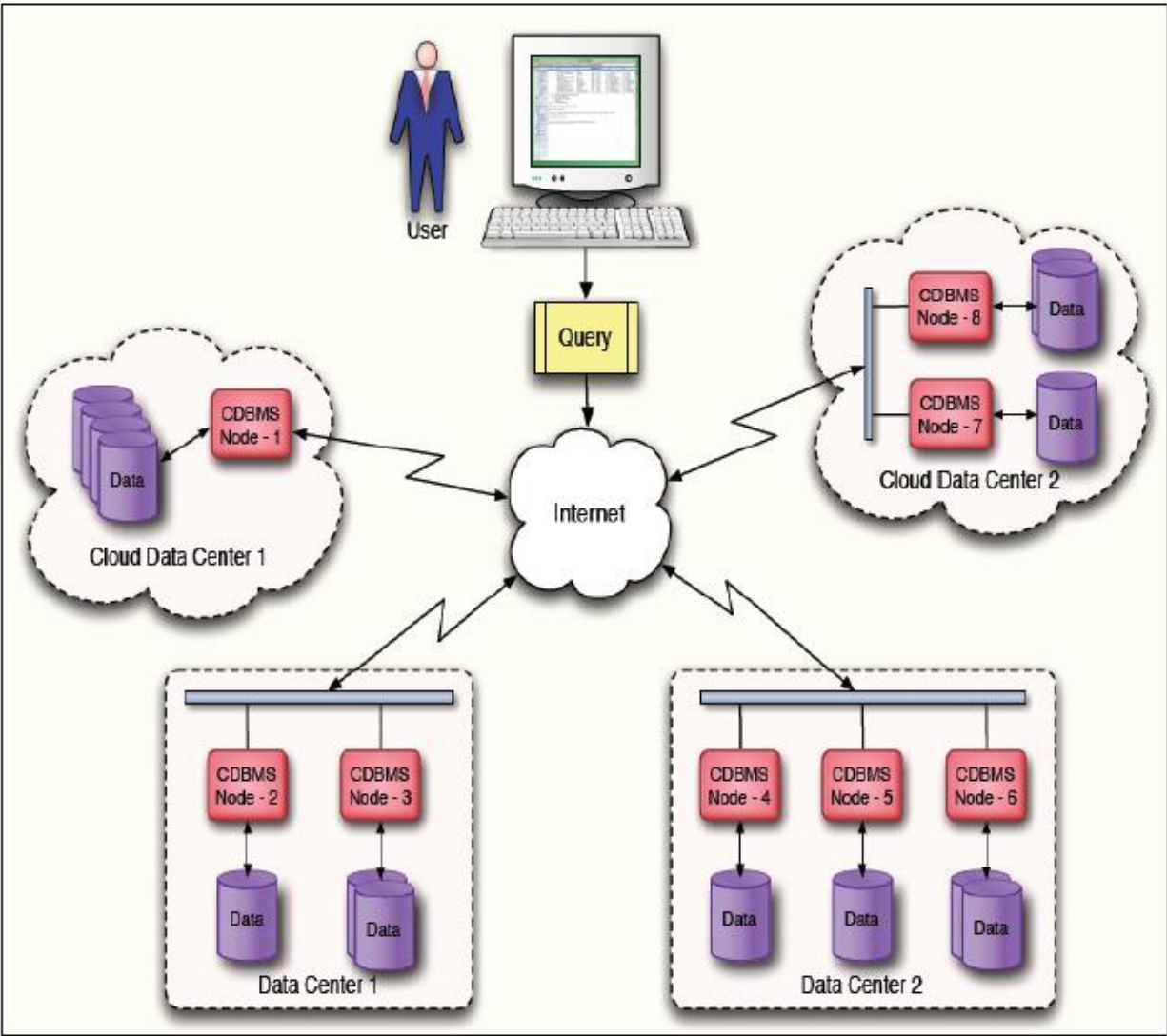
Figure 1:Structure of Cloud database

Source: Waleed et al. (2013)

2.10 Conceptual Model

The use of database nodes will help improve service delivery in access of criminal information. The system will enable CID detectives and staff in different CID offices located countrywide to search for data in police databases concerning a particular individual under investigation or applicants of police certificate. Due to increased rate of certificate forgery, employees and other interested parties will be able to verify the authenticity and validity of a certificate by sending the ID number of the certificate holder and get a confirmation SMS on whether the certificate is genuine or not.
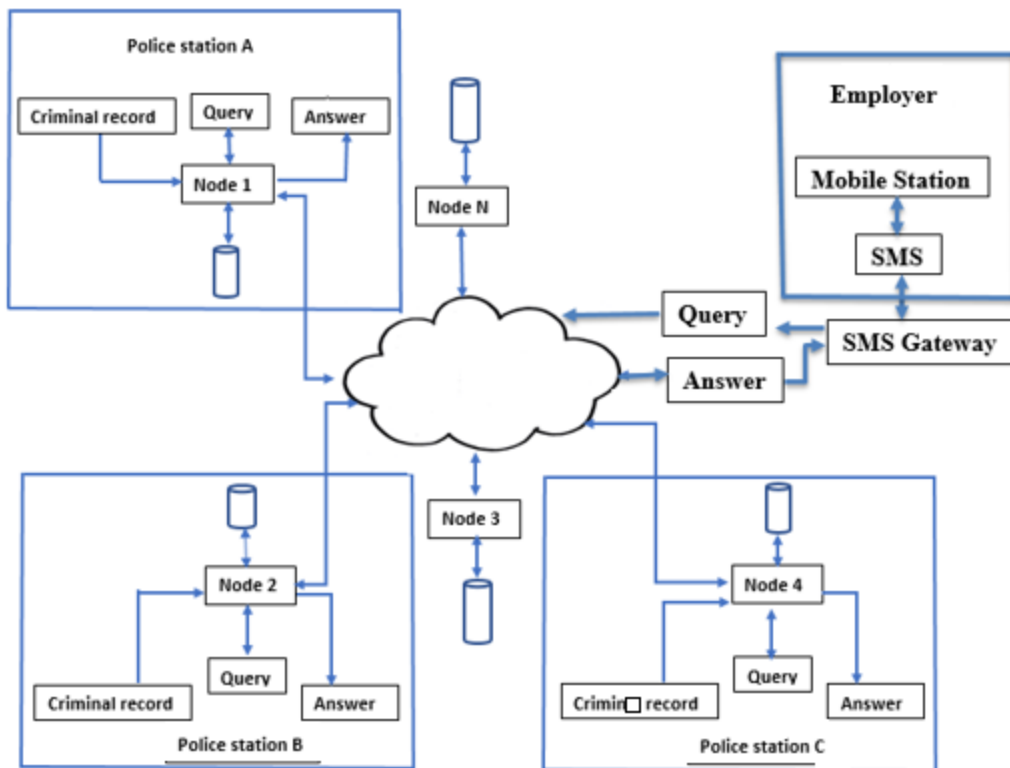
Figure 2: Conceptual Model

Source: Researcher 2017

Multiple Nodes

Due to increasing workload or multiple access of the cloud database, it becomes complicated for the cloud to handle growing number of queries, however multiple nodes in a data center handles shares the workload hence ensuring timely response

Distributed Queries

Different sort of queries can be involved to fetch data from the cloud database. Curino et al., (2011) asserts that when an enterprise is holding a large repository of data different types of queries can be involved to fetch data from the database making use of different nodes. The Distributed queries ensure the most efficient way to deal with the cloud database.

The distributed query can be viewed as the amalgamation numerous queries and each query will be directed to every distributed node for the access of record. According to (Bloor, 2011), the number of results are multiplied as a result of different queries then joined at the end to generate federated answer. A distributed query may further be broken down into sub-queries and then each sub-query will be assigned to a specific node. The answers retrieved from the node are

returned in a distributed form (Waleed et al, 2013). The answers from the nodes are then merged and returned to the requesting user. This can only work perfectly if all databases attached do not contain similar information. As for replicated database, it would make more sense if a particular query is processed by a specific node since all nodes contain the same version of data. This study all nodes contain replicated criminal record hence any request is serviced by any node. The system sends query randomly to any node for processing and then returns an answer to the user.

CHAPTER THREE

RESEARCH METHODOLOGY

3.0 Introduction

This chapter outlines study area, research design and methodology used in the study, which is further subdivided into; target population, data collection procedure, research instruments, validity and reliability of research instruments, and data analysis and presentation, assumptions of the model and ethical consideration.

3.1 Research design

This study used descriptive research approach. This approach is used when the researcher wants to provide answers to the questions of who, what, when, where, and how associated with a particular research problem (Bickman et. al, 1998).The researcher used Mixed methods research. This is a research design with philosophical assumptions as well as inquiry methods that guides the path and process that is involved in the collection and analysis of data. This research design also includes the combination of qualitative and quantitative data in a single study or chain of studies. Its central principle is that the use of combined quantitative and qualitative techniques provides a researcher with better understanding of research problems. (Creswell, et al. 2007).Questionnaires were used to collect quantitative data while interviews were used to collect qualitative data. After collecting relevant data, the researcher developed a prototype to demonstrate how distributed database nodes can make criminal background check process effective.

3.2 Target Population

This is the aggregate set of elements of analysis of data Kothari(2008). This study categorized target population into two: Officers at the Directorate of criminal investigation, a unit of Kenya Police responsible for investigating complex cases, applicants of police clearance certificate together with selected employers. The employers were examined to answer whether they have been verifying the authenticity of police clearance certificate presented to them by applicants. Currently there are 200 officers who are involved in criminal investigations and issuance of these certificates at the CID headquarters.

3.3 Population Sample

Population is the entire group of individuals, events or objects having a common observable characteristic (Mugenda and Mugenda, 2003). It includes a list or collection of the whole population of individuals or events or documents that the researcher wants to include in the study from which he chooses a sample (Oates, 2005)

The level of anticipated precision and the population size are the crucial factors to consider when determining the sample size in order to fulfill the requirements of efficiency, representativeness, reliability and flexibility (Kothari, 2004).During the research, the study adopted stratified sampling technique since it will be targeting specific group of CID officers who deal with the issuance of Police certificate. If a sample is to be drawn from a population that is not comprised of a homogeneous group, a sample representative sample can be obtained through stratified sampling technique (Kothari, 2004).

Qualitative research methods use a smaller sample size as compared to quantitative research methods because they are aimed at gathering an in-depth understanding of phenomenon which focus is on how many and why a particular process or situation (Dworkin,2012). He further stipulates that the intention of interview work is not to draw generalizations to a larger population of interest to the researcher neither is it based on testing hypothesis but is inductive in its process.

There has been debates on the sample size for qualitative research. According to Mason(2010), majority of scholars concur with the fact that saturation is the most vital factor of consideration when determining the sample size of a qualitative research. Saturation is the point at which data collection process no longer yields new data or data which is relevant to the research. According to Dworkin(2012), most researchers tend to evade the thorny question of the number of interviews that are sufficient for a qualitative research with many giving answers of "it depends". Baker & Edwards(2012) maintains the view in their review paper that published some of the responses from the pioneering qualitative researchers like Harry Wolcott who for instance termed the question of "how many" as perennial and that answer to this question depends on factors that can be controlled by the researcher while others are beyond their control. These factors include available resources, the time available to conduct the research, the importance of the question to the research and the number of respondents enough to satisfy committee members for a dissertation. Other researchers like Doucet postulates that the researcher must be conscious of the type of facts that will make their mentors, peers and readers to be contented before they

arrive on the sample size. According to Dworkin (2012), anywhere between 5-50 participants as viable sample size for qualitative research. This helps build and maintain more close relationship and interaction between the researcher and respondents. It also facilitates open and frank exchange of information. Based on publication by Dworkin (2012), 5 participants from the CID would be sufficient enough to gather requirements that can help develop a prototype that can improve on the performance and increase efficiency in issuing of police clearance certificate and boost criminal justice.

In a related study, Muthoni (2015), asserted that the total population of job seekers and employers is too large thus treated it as 'unknown' for the purpose of research and used 200 as a representative population for the unknown. Other parameters to consider are confidence level of 95% and confidence interval of 5%.

To acquire the sample size, this study will use several approaches and formulas.

$$n = \frac{z^2 \times P(1 - P)}{e^2}$$

where n=desired sample size

z-score is the confidence level, for this study 95% thus z-score is 1.96

P is the expected true proportion which will be 0.5 for this study

margin of error (e^2) will be 5%.

Thus, necessary sample size

$$n = \frac{1.96^2 \times 0.5(1 - 0.5)}{0.5^2}$$

$$= 384.16$$

=385 persons

Based on Mugenda&Mugenda (1999), the formula below can be employed to compute sample size where

nf= desired sample size when the population is less than 10,000,

n= desired sample when the population is more than 10,000,

n= estimate of the population size (this study will set it at 200).

$$nf = \frac{n}{1 + \dfrac{n}{N}}$$

Therefore

$$Nf = \frac{385}{1 + \dfrac{385}{200}}$$

$$= 131.6$$

The sample size chosen(nf)=132 persons

The researcher therefore used sample size of 107 job seekers/applicants and total of 30 employers in different sector to gather data.

3.4 Collection and analysis of Data

The researcher put into consideration several options of data collection methods by examining performance of tools in terms of efficiency and effectiveness in collecting the required data with minimal biasness, low cost and less duration of data collection. The researcher used various methods and instruments of data collection including questionnaires to collect quantitative data from police certificate applicants, selected employers and structured interviews to collect qualitative data from CID officers. In order to understand the activities involved in processing police certificate the researcher interviewed the officers at the CID headquarters. This was of great importance because their input would contribute immensely in describing the aspects and the requirements of the system. Interviews were also vital in acquiring insights on the current system of criminal background check and the entire process involved. The choice of interview by the researcher was based on the ability by respondents to provide immediate feedback and also allow the interviewee to give in-depth information that cannot be gathered using questionnaire. This also gave the researcher an opportunity to investigate challenges facing effective and timely issuance of police certificate in an in-depth way as well as discovering how CID officers think and feel about cloud database and their opinion on the same. The researcher created online questionnaire using Google forms. This helped in collecting data at first hand as well as guarding on data anonymity. The choice of online Google forms was informed by its ease of access and has no usage restrictions and also its capabilities to generate data in excel sheet for

further analysis. One set of online questionnaires was shared to selected employers. This was intended to provide an insight as to whether they verify the authenticity of police clearance certificate and the process they use to do so before they recruit prospective employees. On the other hand, the job seekers and applicants of the police clearance certificate were also issued with questionnaires in order to gather information about their satisfaction on the current process of application and issuance of Police certificate. This design is accurate because it involves depiction of current status of activities in a cautiously intended manner (Babbie, 2002). This method of data collection is also free form bias of the respondent (Kothari, 2004).

Qualitative data was used to inform the study and help the researcher in collecting requirements of prototype development. On the other hand, quantitative data was analyzed using Google forms data analysis utility and charts available in Microsoft Excel to discover relationships that are of interest for the research then drew general conclusion. Quantitative data was analyzed using descriptive statistics. As Kiriinya (2014) stipulates, descriptive statistics involves transforming raw data into tables, charts, with frequency distribution and percentages to allow the researcher to explore and derive values and patterns in it.

3.6 Limitations in Research methodology

Google forms was an accurate tool for data collection however it also comes with its limitations. In as much as the researcher wanted to preserve anonymity, he had to ensure that he restricts the respondents from filling up the form more than one time. This called for restriction on the Google forms settings so that the respondents can fill the form only once. This restriction was enforced by ensuring one response per email address however this also poised a challenge in controlling the number of responses from one respondent especially when the respondent has multiple email accounts.

Another limitation that the researcher encountered was laxity by police officers to respond to the interview questions, due to bureaucracy involved in the department.

3.7. Extreme Programming

This is a software development technique that is designed to improve the quality of software that is produced and also improve its ability to properly adapt to the changing requirements of the customer or client.

According to Beck, K. (2000), extreme programming consists of the following steps.

24

The first step was to specify all the features that are desirable in the system then development of the first prototype with minimal functionality followed

The next step is known as iteration planning. Under this phase, the most valuable part of the system known as stories were picked and implemented over a few weeks. Stories use cases that briefly capture functional requirements. Developer typically wrote stories on index cards to describe each system task. Stories should be testable therefore after implementation of the system segment, that segment was released and put into operation.

If new stories that are more valuable than the stories that are in the current release, the new stories can be substituted to the existing stories and released.

After each story was released, an acceptance test was conducted to ensure that the code is satisfactory. At the end of the first iteration, we had the seed, which is a system that is recognizable as the final system, although not everything will work in that system.

The system was integrated continuously as each feature was implemented. Tests were written as the code was written, and the design of the system continuously evolved. This involved the addition of flexibility when necessary and the elimination of complexity where it did not add value to the system. This ensured that changes are delivered at reasonable cost.

The diagram below summarizes the process of extreme programming.
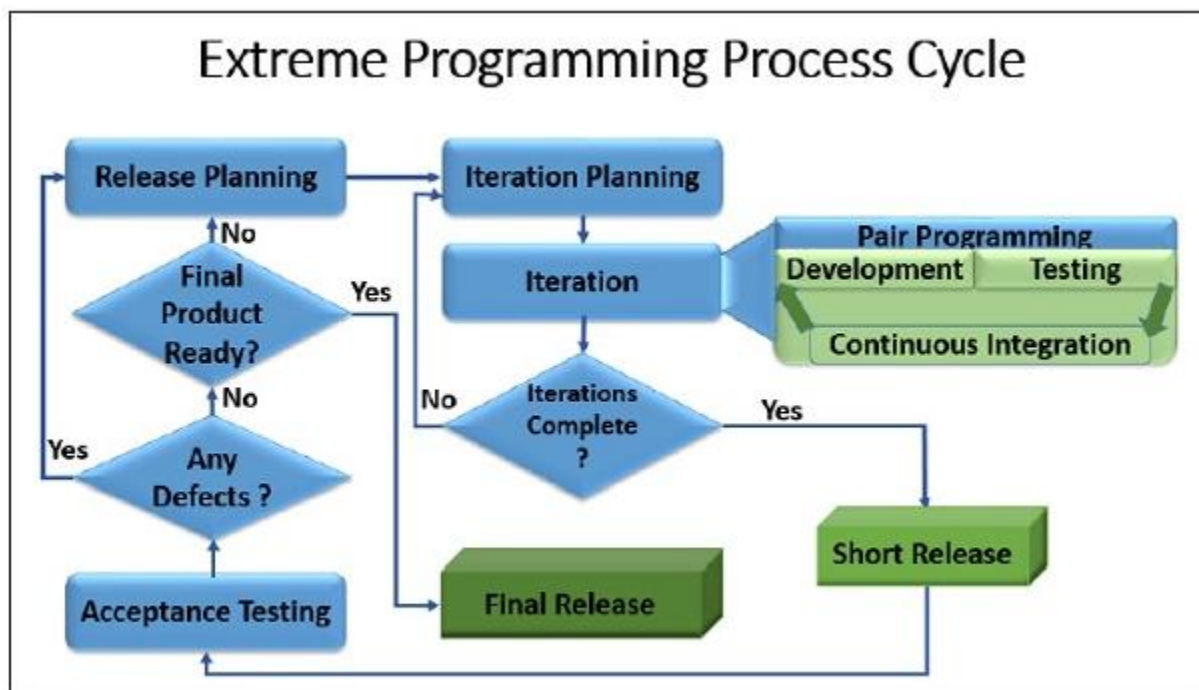
3.7.1 Reasons for Extreme Programming

First and foremost it enables the programmers to focus on the code rather than doing paper work that might not be very necessary, and also eliminates having too many meetings with the client. This increases the time that it takes for a system to be developed up to production level.

Is also ensures that a system is created with minimal defects, as opposed to the traditional methods. You will be able to change requirements quickly at minimal cost, thus making it more effective than traditional methods of software development.

You will be able to have small releases of the software that is being built. The releases are also tested at the end of iteration. That means that the software will be visible to the stakeholders and the client at the end of each iteration. They will be able to have a feel of the software even if the development process might not be complete.

3.8 Ethical consideration

The study acknowledges the importance of ethical issues in a research study and therefore there will be no need to observe the ethical issues of integrity, consistency and honesty while obtaining data from Criminal Investigation Department. Accuracy and consistency of information was observed while analyzing the data to eliminate the problem of missing data. The researcher also will ensure that any references made will be attributed to the actual author by indicating the name of the author, and the year his/her work was published.

CHAPTER FOUR: ANALYSIS, DESIGN AND IMPLEMENTATION

This chapter covers on system's functional and non-functional descriptions, the prototype design, the system use cases and contextual diagrams and finally the system implementation.

4.1 Functional Requirements

These are the specific statements of services that prototype should provide and how it should respond to particular inputs.

4.1.1 The administrator

The admin panel can be accessed using the link localhost/good_conduct/admin.php

The administrator will be able to:

- Key in new police officers

- Create new nodes

- Key in new police stations and counties to which they belong.

- Delete all the above data.4.1.2 Police Officer

The police officer will be able to perform the following functions:

- Input a new criminal record.

- Input a new record of police clearance certificate applicant

- Input a new record of a person with clean record incase manual search has generated clean record.

- View all the above data

- Search for a record

- Generate a new certificate of good conduct

4.1.2 The employer

- The employer will be able to send prospective employees ID number to a certain dedicated telephone number via SMS.

- View criminal background record for the certificate holder via text.

4.2 Non-functional Requirements

These are requirements that specify the characteristics and constraints of the system to enable fast and seamlessly search of the criminal background records by the application

4.2.1 The Biometric System

When the police officer logs in and gains access to the police officer portal, they fill a form with the details of the applicant of police clearance certificate and uploads fingerprint scans. The system then compares the binary value of the uploaded fingerprint with those stored in the database (both the tables of the criminal record and that of the one containing clean records).If a match of the finger print or the ID number is found of either the clean record or the criminal record, a police clearance certificate is generated. If no match is found the system notifies the police officer that no record of the applicant is found and a manual search is triggered.

It is expected that once the records are found during the manual search, the policemen will key in the records in the database so that the next time the applicant needs a police clearance certificate, there will be no need of another manual search being conducted.

4.2.2 SMS verification system

When the employer sends prospective employees ID number to a certain dedicated telephone number via SMS, the system compares the ID number of the applicant to see if there is any match to the records contained in the database (both the criminal records table and the clean records table).

If a criminal record is found, the record is sent to the employer via text. If the user is found in the clean records table, a message is sent to the user informing the user has no criminal record.

If there is no match in either of the tables, a message is sent to the user informing him/her that currently there are no records available in the database.

4.3 Architectural Design of the prototype

The prototype will run under cloud architecture- Database as A Service (DBaaS). The system is web based and the client can access the application via laptops, desktops and mobile device to make SMS requests to the server side which used PHP Script and MYSQL database.
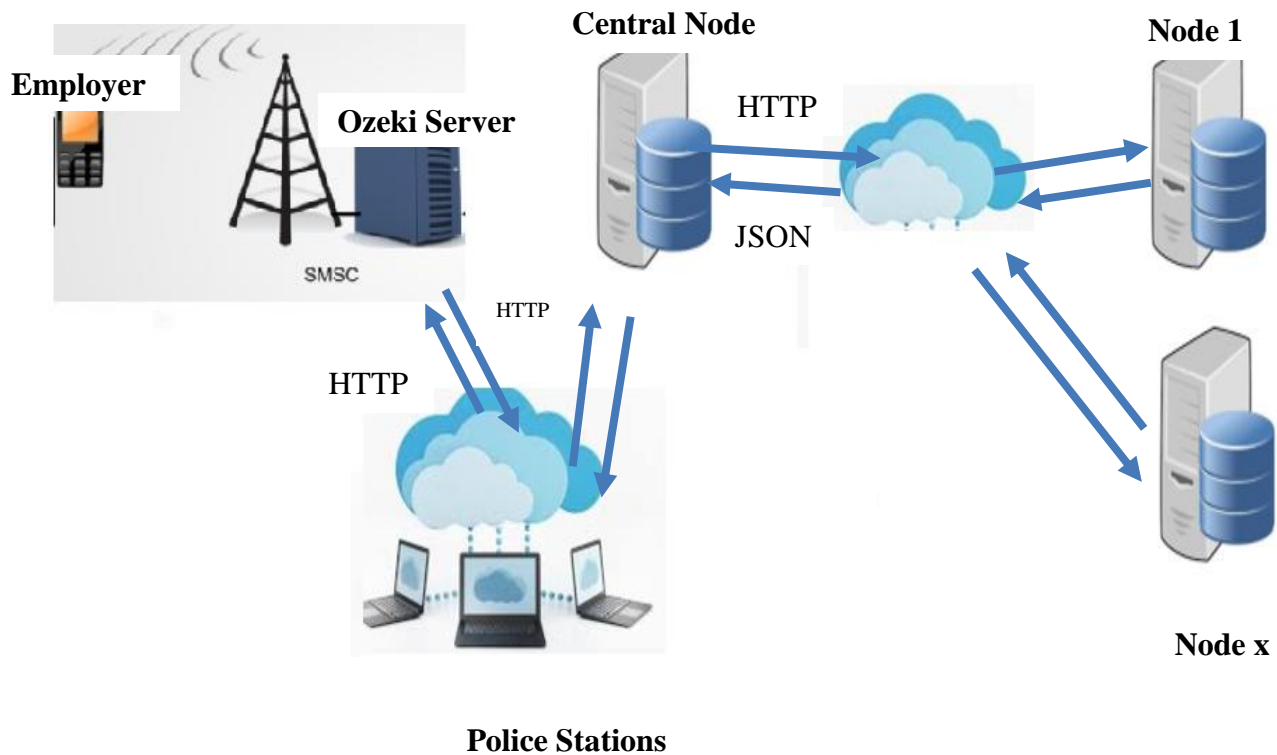
Figure 3: Prototype architecture diagram

Source: Author 2017.

4.3.1 Clients infrastructure

Below is a detailed explanation of the access infrastructure

1. A PC or laptop with internet connection for every CID office in every police station countrywide. This will be used by police officers to make applications for the police certificate for the applicants and upload the fingerprints scans

2. A mobile phone for sending and receiving SMS. This will enable the employers and recruiting bodies to verify the authenticity of police clearance certificate.

4.3.2 Database nodes and servers

The database nodes including the central node will be hosted in distributed Apache servers. There will be more as many central nodes as possible as well as other nodes. The mobile station will be able to interact with the central nodes hosting the web application through SMS gateway. This will enable the SMS server to send, receive and redirect messages accordingly.

4.4 System Design

This is the process in which different elements of a system are defined. These elements may include the components, the interfaces, the architecture and the modules of that system.

The various components that were designed include the System Flowchart, the Class Diagram, The Sequence Diagram, the Use Case Model, the Context Diagram, the Database Design, the Entity Relationship Model and the Interface Design.
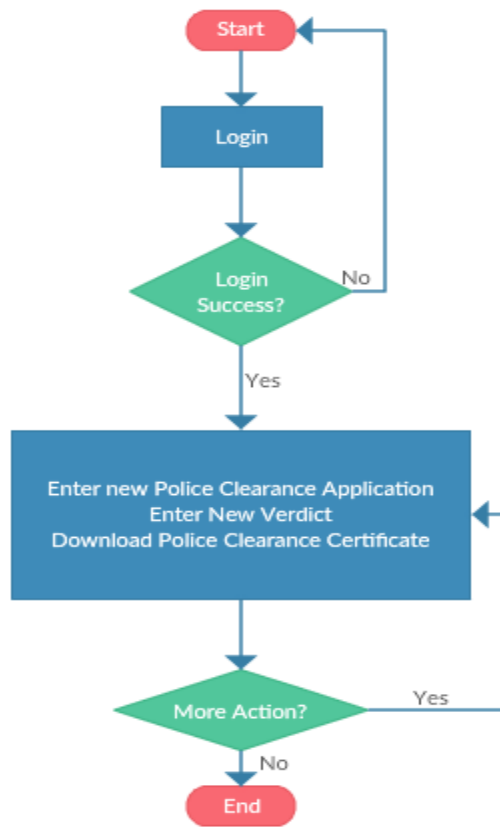
4.4.1 System Flowchart

This is a diagram that represents the flow of data and the workflow in the system using various symbols to designate specific actions.

The two diagrams below illustrate the system flowcharts of the administrator module and the police module.
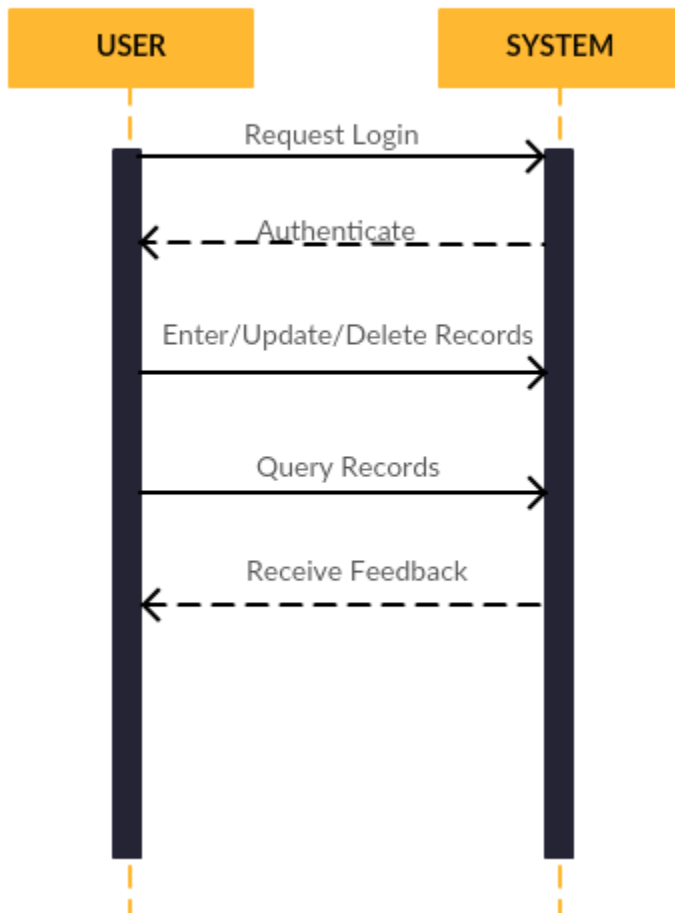
The Administrator Module
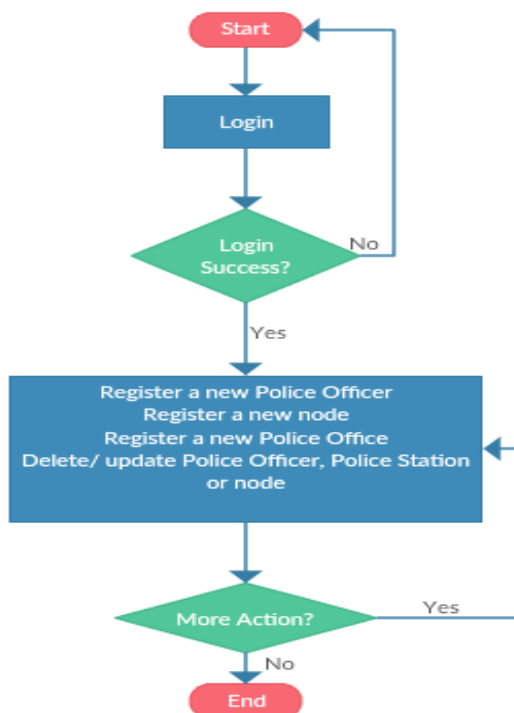
The Police Module

4.4.2 Sequence Diagram

This is a diagram that is used to show how different objects in a system operate with one another as well as the order in which they operate.

Below is the sequence diagram.

4.4.3 Use Case Model

This is a diagram that is used to display how different users of a system interact with it to solve a task. It is used to define the objectives of the users while using the system. It also shows the interactions between users and the system as well as how the system behaves to satisfy the set goals.

The diagram below illustrates the Use Case Model of the system.
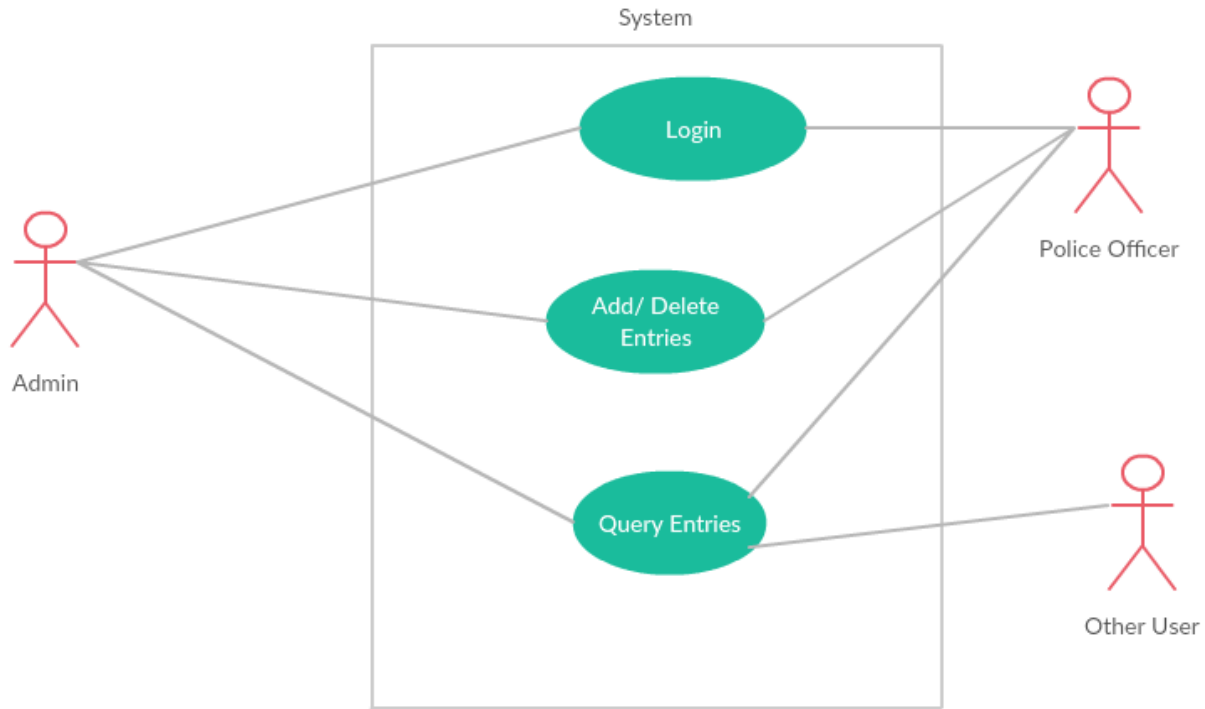
4.4.4 Context Diagram

This is a diagram that is used to display the high-level view of a system in that it defines the boundaries between the system or a segment of the system with its environment. It also shows the different entities that interact with the system.

The diagram below is the context diagram for this system.
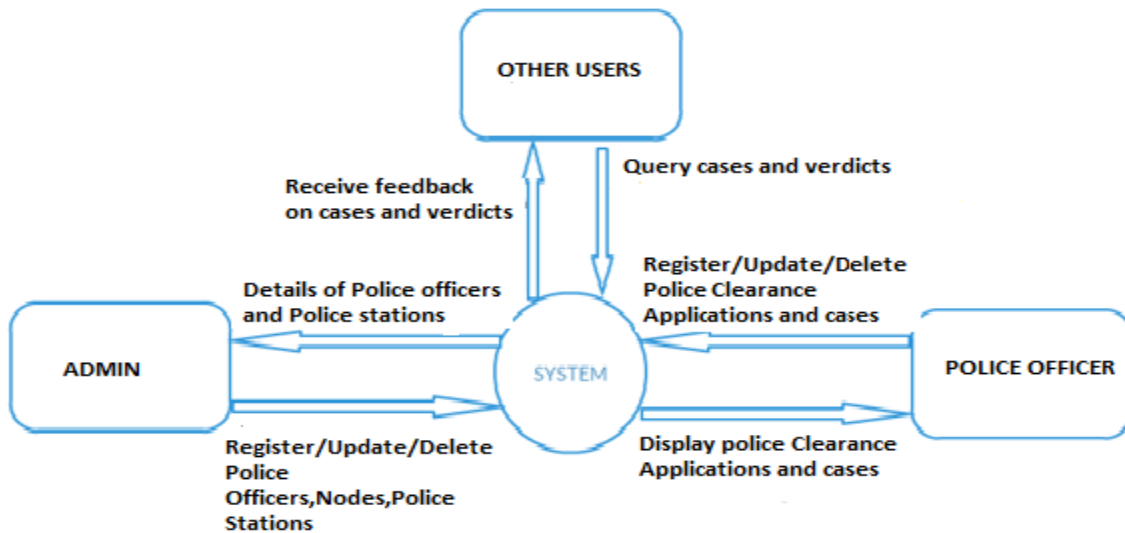
4.4.5 Database Design

A database is a grouping of related data that is structured in a way that the data can be easily managed, accessed, and updated. Below is the database design for the central node and the other nodes.

Below is the database design of this system. The first part is the design of the central node and the second one is the design of the other nodes. The design of the other nodes is identical.

Central Node
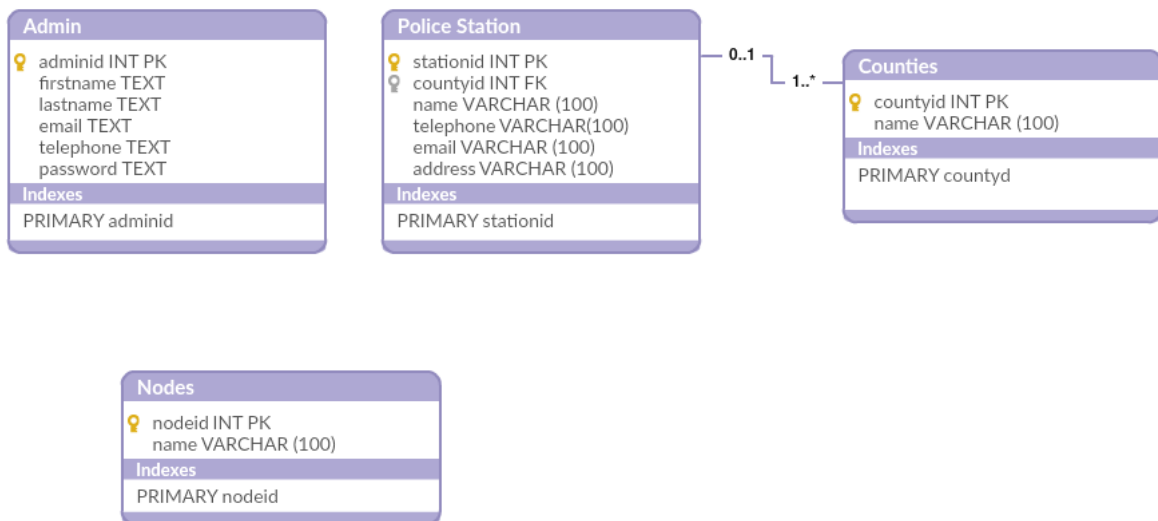
Other Nodes
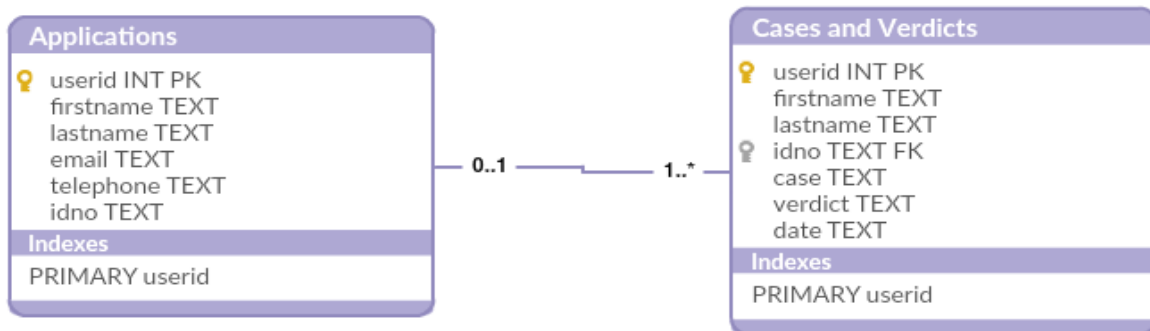
During the search from criminal database, the central node first searches from its database to see if the criminal record is available. If the record misses it searches the record from available nodes. The below loop is meant to search for a random node. If a record is found in the node, the loop would terminate and the central node would get the data it requires from that node.

```php
<?php
        $feedback="false";
        $x=0;
        $page="verdicts.php";
        $page2="registered.php";
        $page3="cleared.php";
        while (($feedback=="false") && ($x < 10))
        {
          $random="SELECT * FROM nodes ORDER BY RAND() LIMIT 1";
          $random2=mysql_query($random);
          while ($random3=mysql_fetch_array($random2))
          {
            $link=$random3['url'];
            echo $link;

                }
        ?>
```

4.5 Implementation

4.5.1 Hardware Resources
1) Modem for SMS gateway
2) Phones
3) Laptop or desktop with the following hardware specifications:
   - A Pentium 4 processor and above
   - A 1 GB RAM and above
   - At least 20 GB Hard disk and above.

4.5.2 Implementation Technologies
Several software applications and packages were used.

1. Netbeans IDE - This is a development framework used in developing mobile, web and desktop applications.

2. Bootstrap – Bootstrap is an open-source front-end application development framework built using JavaScript, HTML5 and CSS3. It is used in constructing websites and web applications.

3. PHP (Hypertext Preprocessor) Is web programming language for server-side and is designed for web development but can also be used as a general-purpose programming language.

4. **JavaScript-** It is a lightweight web scripting language used for making interactive web pages. It can be embedded in HTML and interpreted using a web browser.

5. **Ajax (Asynchronous JavaScript and XML)** is an application that allows for creation of speedy and dynamic web pages. In AJAX, web pages are updated asynchronously through exchange of small amounts of data with the server in the rear scenes.

9. **Ozeki SMS Gateway -** The Ozeki Message Server is a gateway of integrating SMS in various applications and routing messages to email addresses, mobile message centers and web pages.

10. **Apache -** This is an open source Web Server.

11. MySQL Database - this a popular database with web developers used in various web applications, it is an open-source software.

CHAPTER FIVE: SYSTEM IMPLEMENTATION, TESTING AND DISCUSSION

This chapter presents results and the discussion obtained from the system implementation and the research study. The chapter also includes the main screenshots to illustrate operation and proposed data formats.

5.1. System Implementation.

5.1.1 Interface Design.

These are mockup images that represent the flow of the different screens or interfaces. Is used to display how the system will be navigated from the first screen to the last.
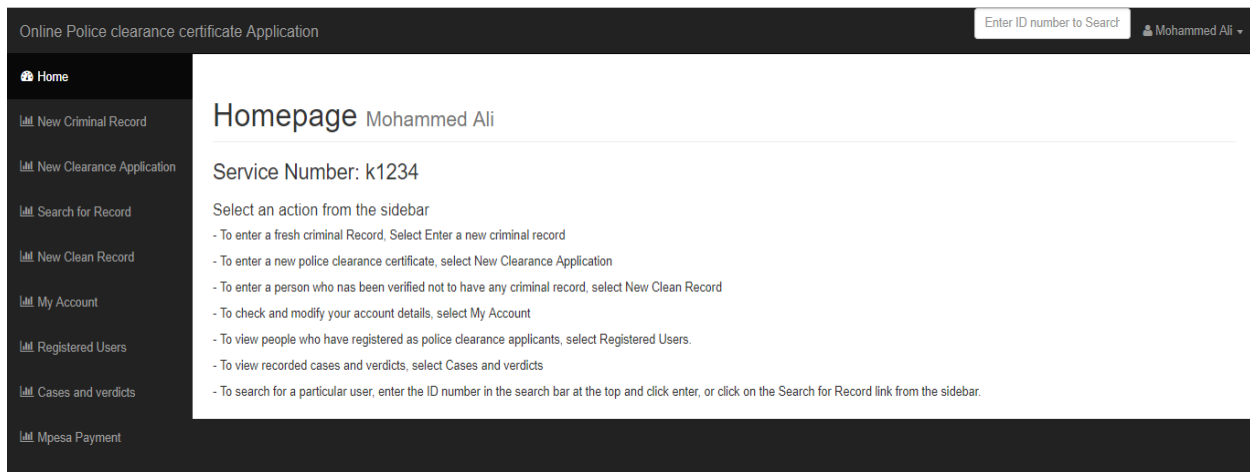
Below are the user interface diagrams.

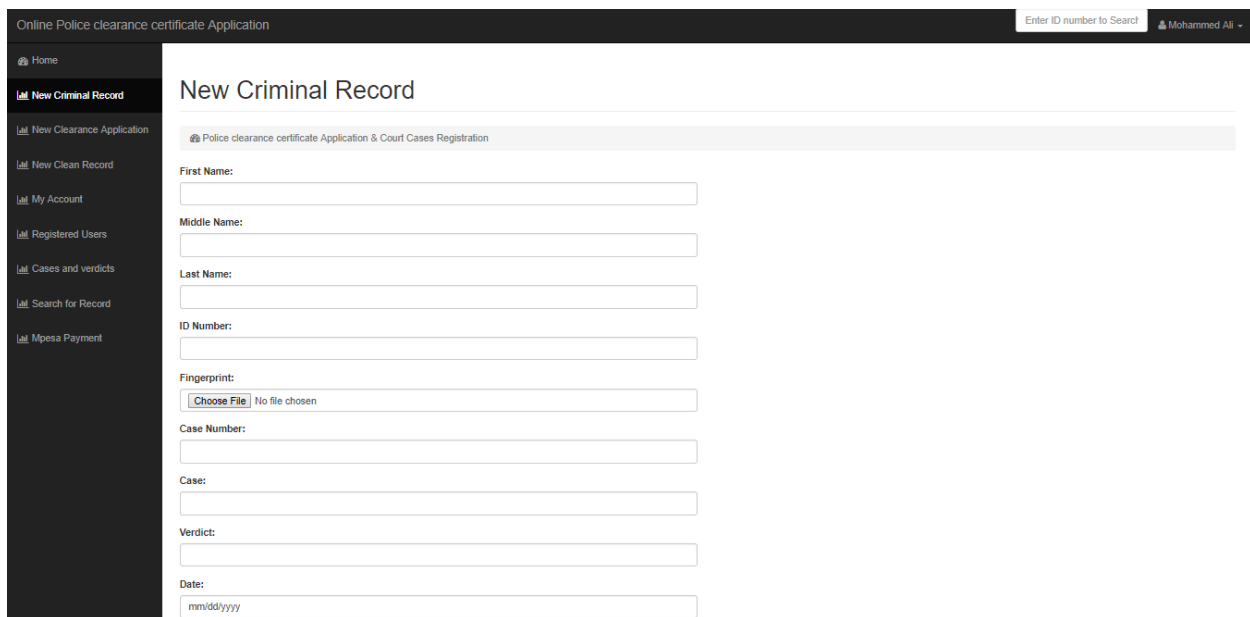*Figure 4:***Login page**
Source: Author 2017.

System dashboard

Figure 5: Homepage

Source: Author 2017.

Interface for entering criminal record

Figure 6: Record entry page

Source: Author 2017.

5.2 SYSTEM TESTING AND EVALUATION

5.2.1 Validation Testing

Each and every data entry field was validated to ensure that the correct data type was entered. This was done through PHP in the server side and JavaScript in the client side. When wrong input was inserted, the user was alerted and given a chance to correct the wrong input. No wrong input was allowed to be transmitted successfully.

The screenshot below illustrates the reaction of the system when wrong input was inserted.

5.2.2 Unit Testing

The various modules were tested independently to ensure that they met the desired standards. The result of the unit testing revealed that all the modules met the required standards.

5.2.3 Data Entry Testing

Test data was entered into the system to check if it was able to capture data and to store it into a database in the correct format.

To be able to draw an input, several input data is required. The first-hand data entry of criminal record is entered by police after a case has been determined and closed in courts of law. The ID number and the fingerprint scans are entered into the systems and the central node replicates the information to all available nodes. If any of the information about the criminal is not entered an error will be generated to ensure comprehensive information is submitted.

The prototype input screenshot is shown in the figure below.
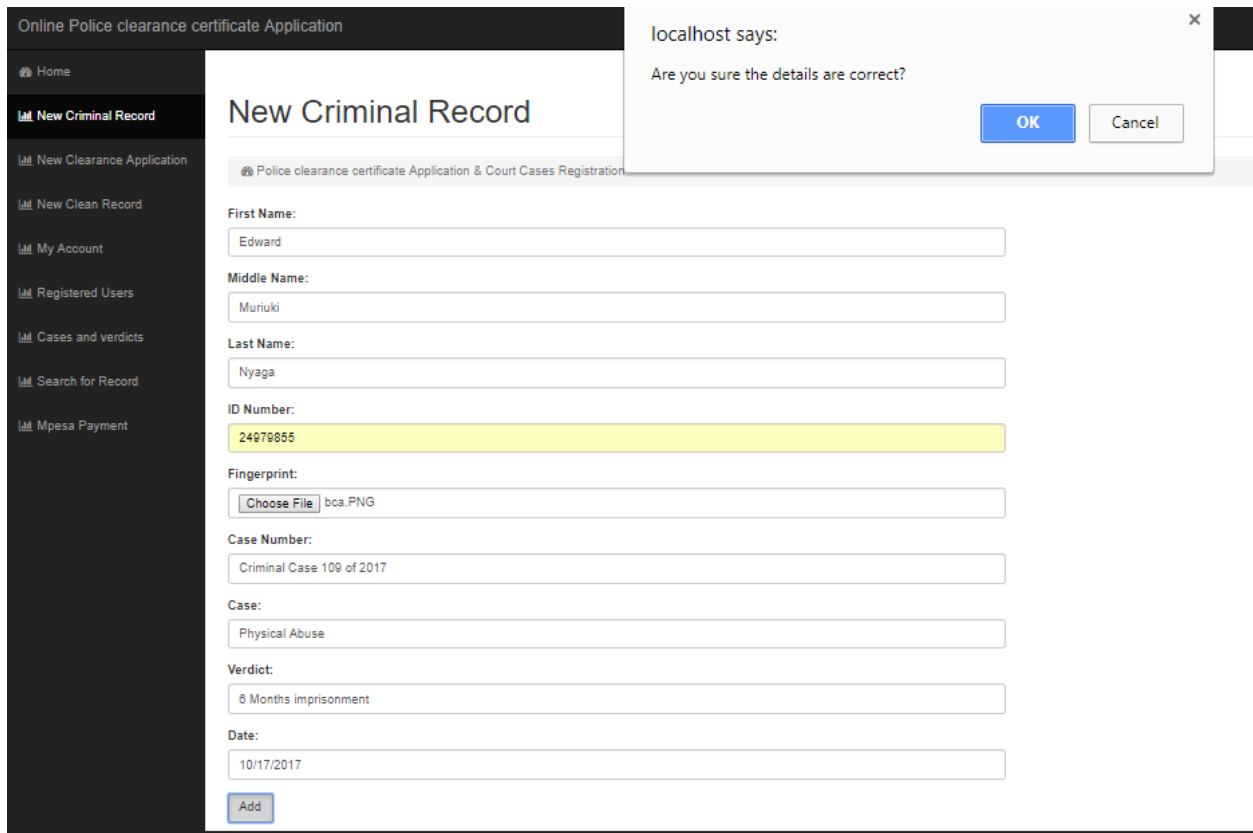
Figure 7: Data entry test
Source: Author 2017.

5.2.4 System Acceptance

System acceptance testing was conducted to test the basic functionalities of the system and how user requirements of the system and user requirements were directly achieved. The main objective of this test was to examine the success of the prototype usage in the actual environment and identify issues that would arise for further refinement

Table 1: The summary of Police ability to perform actions on the integrated system

	Feedback
Success	5
Failed Test	0
%Success	100%

From the sample 100% of the users appreciated the system was a success and that this provided an efficient way to process Police clearance certificate without strain regardless of locality from

where the applicant is making the application. The evaluation of the overall system showed that with the system in place, workload will be distributed hence improve the delivery of service in the application of Police clearance certificate.
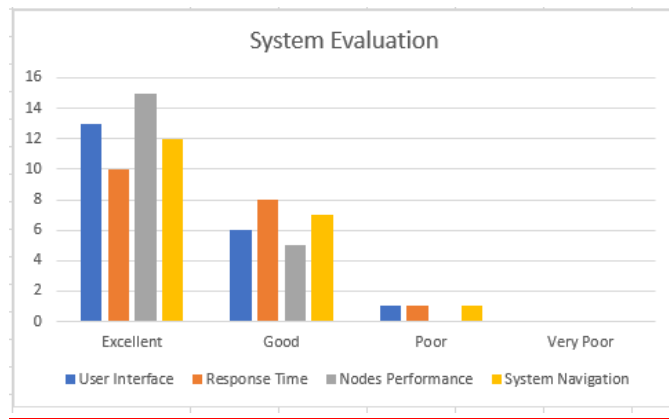
5.3 System Evaluation

System evaluation was conducted to gather user views on the performance of the prototype and their perception towards utilization of cloud database. The key parameters of this test were on:

- User Interface
- Response time
- Nodes performance
- System navigation

The evaluation was carried out with the selected users of the system. A sample of 20 respondents was used to evaluate the prototype as well as rating its performance. Among the respondents were 10 Police officers and ten professional developers. During evaluation, the researcher issued questionnaires to collect data in order to understand the views and perception towards the performance of the prototype. The collected data was analyzed using Microsoft excel in order to generate tables and graphs

Table 2: System Evaluation

|	Excellent	Good	Poor	Very Poor
User Interface	13	6	1	0
Response Time	10	8	1	0
Nodes Performance	15	5	0	0
System Navigation	12	7	1	0
Overall rating %	62.5	32.5	3.75	0

Figure 8: The summary of system evaluation

Source: Author 2017

The greatest number of system users, represented by 62.5% supposed that the performance of the system was excellent, 32.5% believed that the system was good to some extend while 3.75% thought that the system was poor. The overall sample rating on the evaluation of the system showed that the system would be of great help when operational zed in processing and issuing of Police clearance certificate.

5.4 Output Data
Sample Police Clearance Certificate

When data has been input, a search is carried out in the database and the criminal record of the applicant is retrieved and the softcopy of the output is generated then emailed or downloaded as pdf document for printing. The certificate generally contains the printout of any criminal record if there is any previous record, or if there is no existence of any criminal record, manual search is carried out and the record entered into the system if any, else a clean record is entered for that applicant and a certificate with no criminal record is generated.
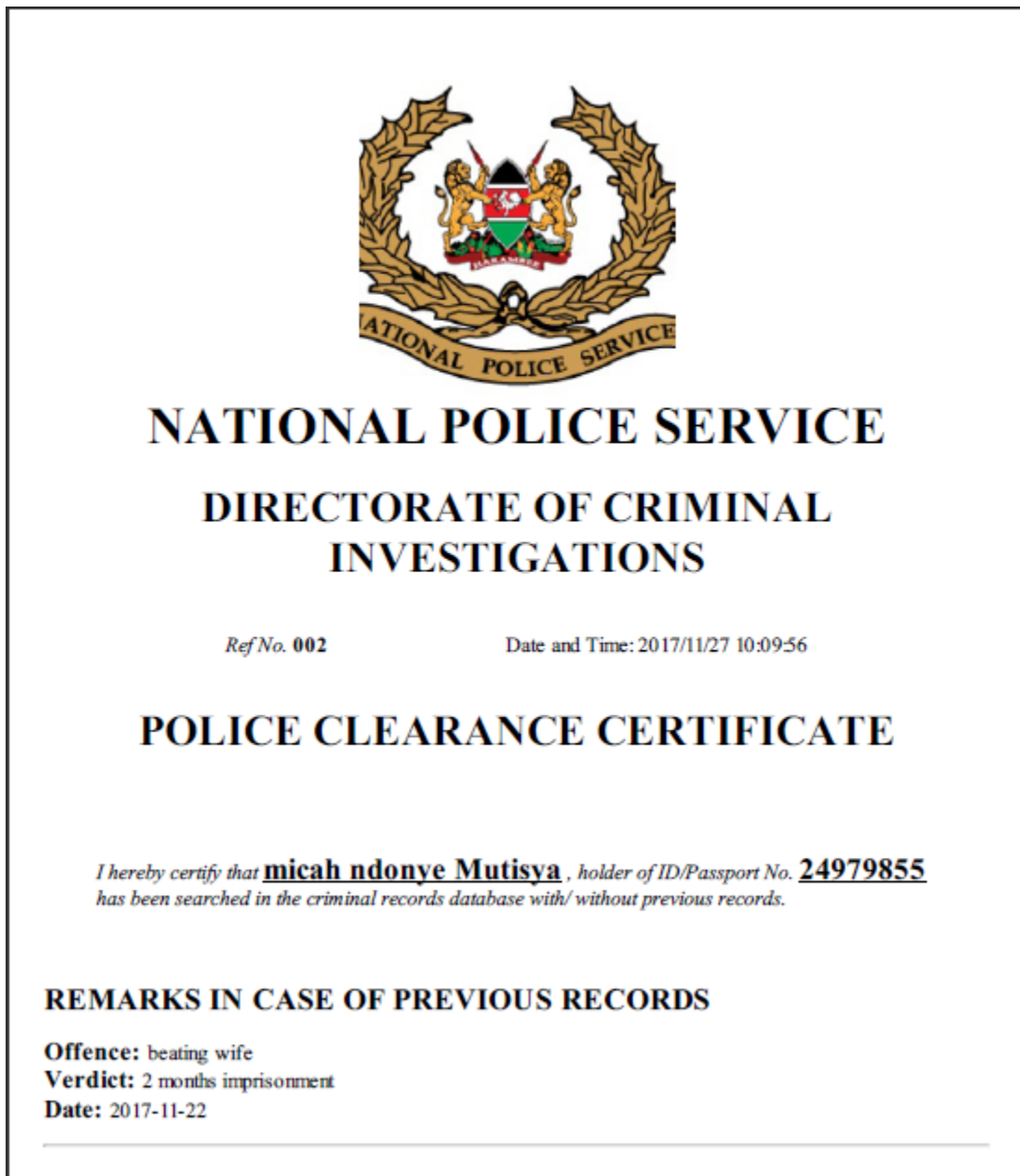
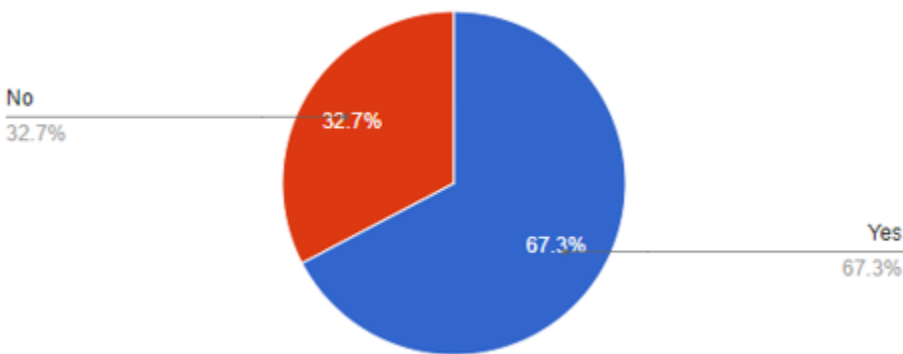*Figure 9: Sample Police Clearance Certificate*

Source: Author 2017.

5.5 Data analysis and findings

This is the process by which data that has been collected is inspected, cleansed, transformed as well as modeled with the aim of useful information being discovered, conclusions being suggested with the aim of supporting decision-making processes.

5.5.1 Analysis of response from the applicants of police clearance certificate

The researcher manged to interview 107 applicants and holders of Police Clearance Certificate out of a target sampe size of 132 respondents as specified, the following were the responses from those whom the researcher manged to get feedback from.

As to whether they had previously pplied for a Police Clearance Certificate before, out of the 104 responses we got from the respondents 67.3% said that they had appied the Police Clearance Certificate before while the remaining 32.7% said that they had never applied for the certificate before.

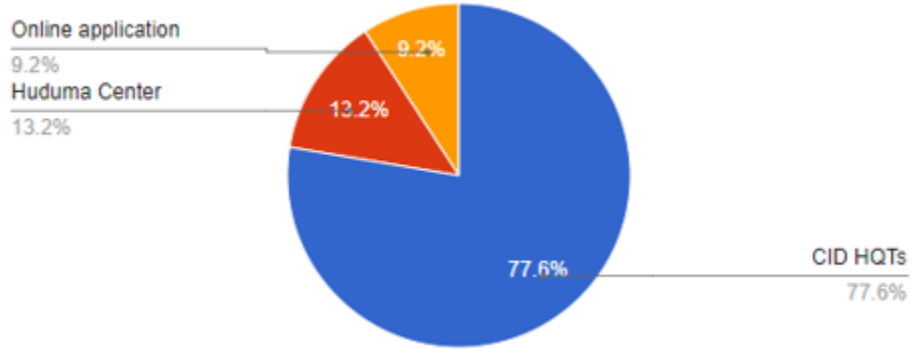*Figure 10: **Application of Police clearance certificate***
Source: Author data 2017. Derived from Google forms

From this study it was evident that the majority of respondents 67.3% had experience had already applied for this certificate at least once.

Having applied for the certificate, the researcher was then interested to find out their County of origin and where they had made the application from. Out of the 107 responses received, 77.6% had travelled to made the application from CID headquarters, 13.2% applied for the certificate from the Huduma Centers located in counties headquaters countrywide while the remaining 9.2% used online application as the mode through which they applied for the certificate.

3. County of origin	Top results ▼
Machakos County	29
Nairobi County	12
Makueni County	10
Kiambu County	6
Meru County	5
Nakuru County	5
Kisii County	4
Embu County	3
Nyeri County	3
Uasin Gishu County	3
Bungoma County	2
Homa Bay County	2
Kajiado County	2
Kwale County	2

Bungoma County	2
Homa Bay County	2
Kajiado County	2
Kwale County	2
Nyamira County	2

Figure 11: **Most preferred place of Application**

Source: Author data 2017. Derived from Google forms

The researcher also asked the respondents whether they had been asked to produce Police Clearance Certificates before and under what circumstances were they asked to produce the certificate.
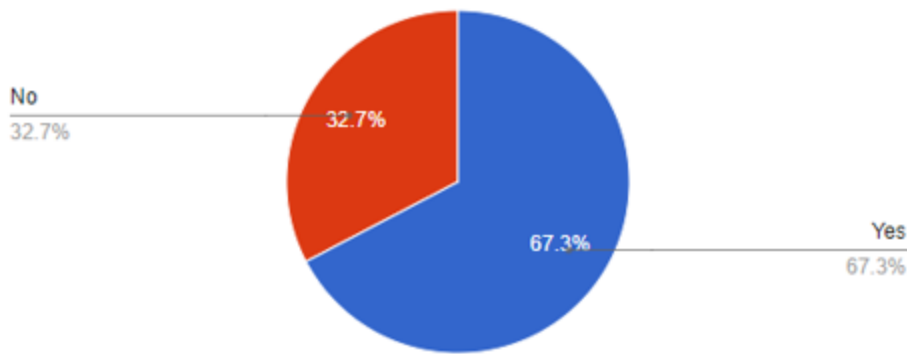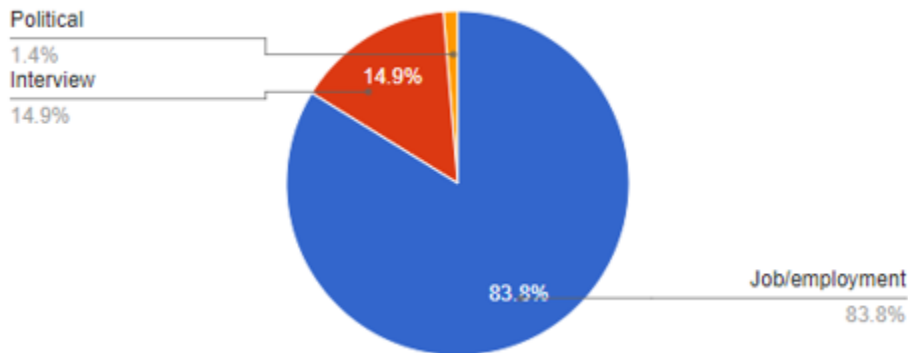
Figure 12: **Response on Application of Police Clearance Certificate**
Source: Author data 2017. Derived from Google forms

As shown above, 67.3% of the respondents had been asked to produce Police Clearance Certificates before while 32.7% had never ben asked to produce the certificate.

The respondents who said yes to the previous questions were asked under which circumstances they were asked to produce Police Clearance Certificates.


Figure 13: Reason **behind Application of Police clearance certificate**
Source: Author data 2017. Derived from Google forms

Out of 74 responses received, 83.8% said that they were asked to produce the certificate as a requirement for job or employment offer, 14.9% responded that they produced the same after they were invited for an interview. However, one respondent represented by 1.4 percent said that it was because of political requirement inorder to vie for an elective post.

When asked how long it took them to receive Police Clearance certificates, their responces were as below.
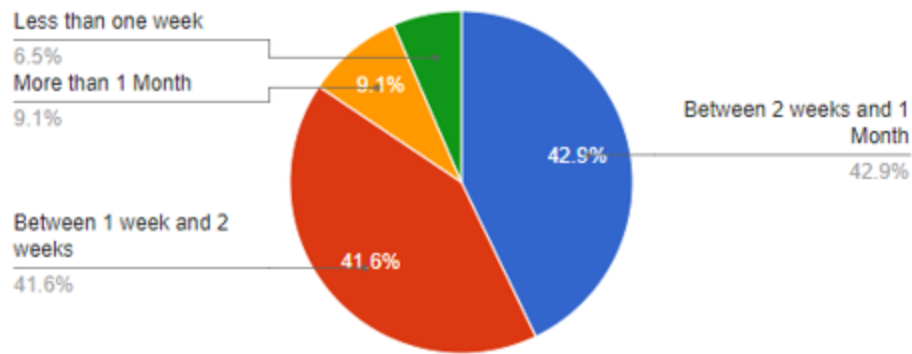
Figure 14: Waiting time for police clearance certificate

Source: Author data 2017. Derived from Google forms

Out of 77 respondents who reacted to this question, the majority represented by 42.9% said that it took them between two weeks and one month. 41.6% of the respondents said that it took them between one and two weeks whereas 9.1% said that it took them less than one week and the remaining 6.4% said that it took them more than one month. After analysis, the reseacher discovered that those who took more than one month were from far counties.

As to whether the respondents were satisfied with the process of acquiring Police Clearance Certificates. Their response was as below.
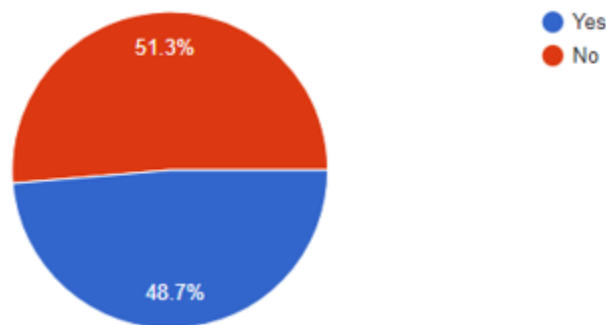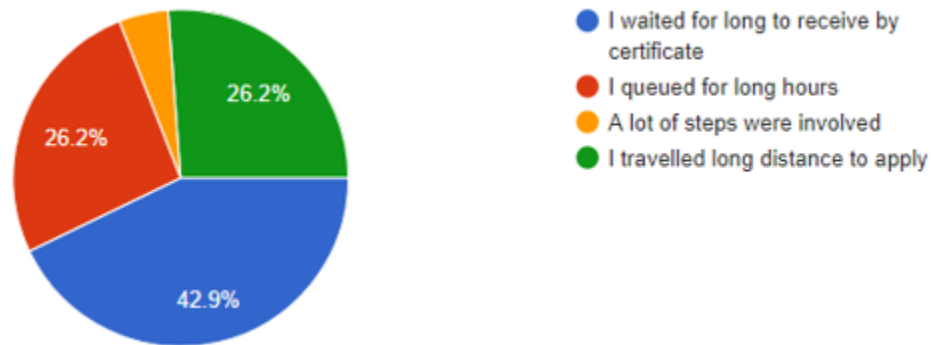

Figure 15: Level of satisfaction with the certificate application process

Source: Author data 2017. Derived from Google forms

Out of 76 responses on this question, 51.3% were not satisfied with the process, where the remaining 48.7% were satisfied with the process it took to apply or acquire the police clearance

certificate. With the majority not being satisfied by the process,the researcher then wanted to know the reason behind their disatisfaction on the process. Theirrespondends were as below.


Figure 16: Reasons behind dissatisfaction of application process
Source: Author data 2017. Derived from Google forms

There were 42 respondents to this question. Majority of the respondents(42.9%) who were not satisfied with the process claimed that they waited too long to get the certificate. Another group of respondends representing 26.2% argued that they queued for long hours. Another 26.2% said that they travelled long distance with the least percantage 4.7% of the respondents saying that there were a lot of steps involved in the process of acquiring the Police clearance certificate.

The researcher tried to find out from the respondents which improvements they would advocate for, in the process of issuing Police Clearance Certificates inorder to have effective service delivery.
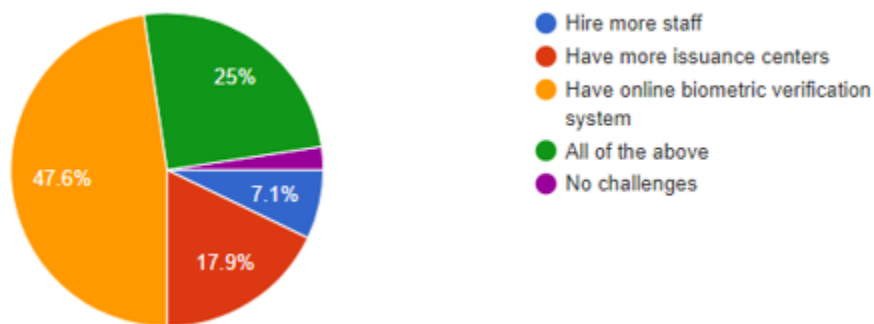

Figure 17:Proposed solutions by applicants
Source: Author data 2017. Derived from Google forms

Many applicants seems to understand the importance of online biometric system on enhancing the process of aquiring Police clearance certificate. Of the 84 responses received, as shown

above, the greatest percentage represented by 47.6% advocates for online biometric verification system as this would eliminate the need to travel to the CID headquaters or Huduma centers for finger scanning. With the system in place, applicants will be visiting nearest police station to have finger prints taken from there and have their police clearance certificate processed from there. 25% of the respondents believed that with all the options in place, then the process will be made easy.17.9% said that they would prefer having more issuance centers. 7.1% of the respondents said that they would prefer if more staff were hired, while the remaining respondents were satifsied with the process and advocated for no changes on the current system.

The researcher also inquired if an SMS verification system would help verify criminal records easily.
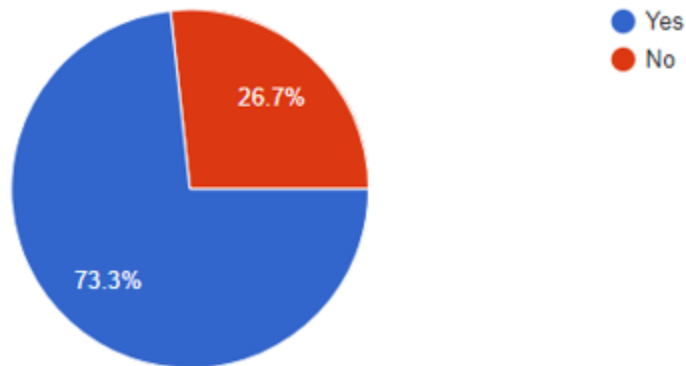
Figure 18: **Verification of criminal records by SMS**
Source: Author data 2017. Derived from Google forms

Out of the 101 respondents, 73.3% thought that SMS verification system would help verify criminal records easily while the remaining 26.7% of the respondents had a different view on sms verification with most believing that once it comes in place then it would breach on privacy.

On the issue of fingerprint scanning, the reseacher asked if the respondents had ever undergone fingerprint scanning by any organization for computer-matched biometric comparison.
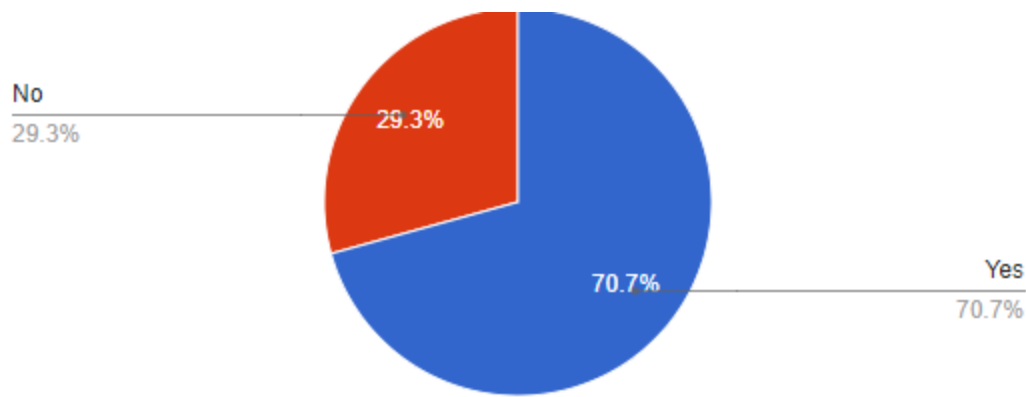
Figure 19: Computer-matched biometric comparison
Source: Author data 2017. Derived from Google forms

Out of the 99 responses received, 70.7% of the respondents had undergo fingerprint scanning mostly based on a computer reading of the line patterrns on the fingers, while the remaining 29.3% had not undergone any fingerprint scanning process. This shows that the use of biometric verification is becoming a common practice with many institutions.

5.5.2 Employer's and police clearance certificate verification

After interviewing the applicants, the researcher went ahead to interview some 30 employers. The researcher wanted to know the nature of their work and whether they demanded Police Clearance Certificates when recruiting employees.
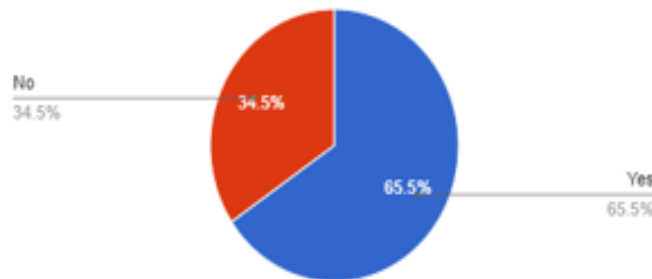
Figure 20: Proposed solutions by applicants
Source: Author data 2017. Derived from Google forms

Out of 29 responses, 65.5% have been demanding Police clearance certificate while 34.5% have not been asking for the same.

When asked whether they had ever verified the authenticity and validity of Police Clearance Certificates, their responses were as follows.
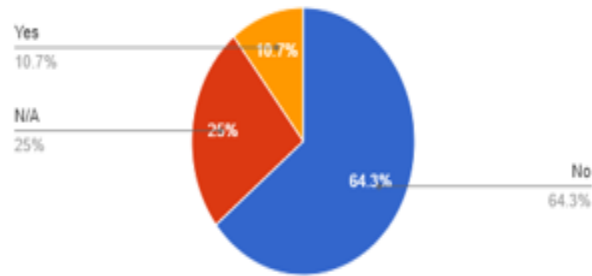

Figure 21: Verification of Police clearance certificate
Source: Author data 2017. Derived from Google forms

There were 28 responses on this matter. A majority (64.3%) of the respondents did not verify the authenticity of these certificates but accepted them based on mere trust. 10.7% Of those who verified the authenticity were from NGO's, Embassies and Parastatals. This question did not apply to 25% of the employees as they did not demand job applicants to present the certificate.

For employers who verified the authenticity and validity of Police Clearance Certificates Out of the 26 responses received, The greatest percentage of 38.5 did not use available channels of verification of Police clearance certificate but based the authenticity on mere trust. 26.3% made applications and physically presented the certificates at CID headquarters for verification of the certificates, while 35.2% were able to verify the authenticity by sending SMS to a designated number.
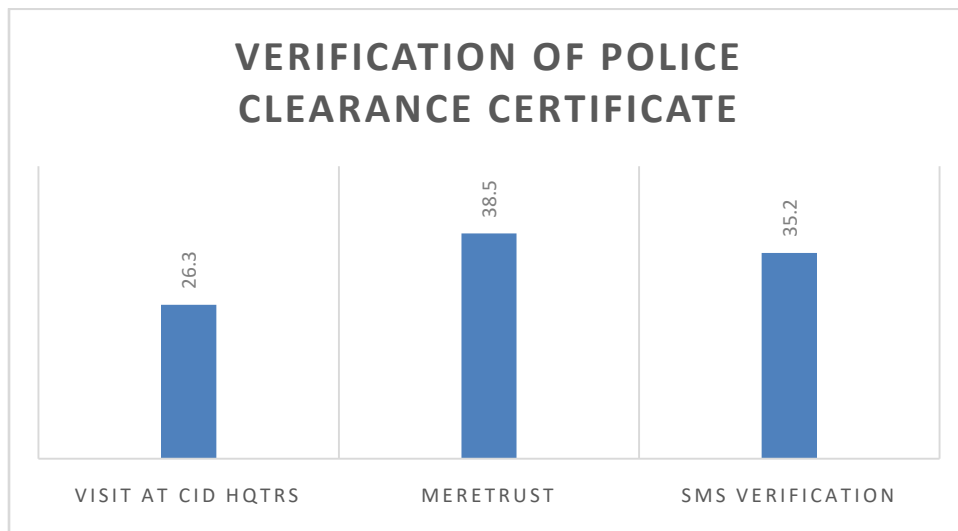
Figure 22: Verification method used in Police Clearance Certificate
Source: Author data 2017. Derived from Google forms

CHAPTER SIX: CONCLUSION AND RECOMMENDATIONS

6.1 Achievements

The first objective of this study was to develop a prototype that will provide solution to effective Criminal background check in the background screening industry by utilization of cloud database. We managed to analyze, design, develop and test the prototype. Several database nodes were designed and deployed in different data centers and tested to see if they could handle shared workload.

The system had three modules, the administrator module, the police officer module and the other user's module. The administrator was able to login, and to view, register, update and delete nodes, police officers, counties and police stations.

The police were able to register new clearance certificate applications, enter new cases and verdicts, query applications and cases and to generate Police Clearance Certificates in PDF format.

The other users were able to view cases and verdicts via an SMS based application that was able to query from a database.

The second objective was to review the methods used for verification of police clearance certificate in current systems. After collecting data from selected employees, it emerged that most do not verify the authenticity of Police Clearance Certificate despite availability of various means through which they can confirm the same. The available verification methods used by the current system are making application and physically presenting the certificates at CID headquarters for certification as true copies. The other commonly method of verification is use of Short Message Service where the interested party sends a code to a designated number then they are prompted to key in details of the certificate holder to confirm if the certificate is not a forgery.

The third objective was to evaluate and investigate how Biometric Recognition technology has been used for Policing and Criminal Background Check. Research of related work on Biometrics technology and Policing was achieved. Several biometric technologies have been discussed in literature review. Some of the techniques currently used are eye recognition which is based on the pattern of fleck in the iris of the eye, fingerprint scan, which is based on a computer reading of the line patterns on the fingers, voice recognition which is based on the frequency patterns in person's voice and facial recognition, which is based on the shape and contours of persons face. The most common used Biometrics technology is finger print scans. Currently the Directorate of

Criminal Investigation holds all fingerprint scans from everyone who ever applied for National Identification Card. Usually anyone who has attained age of majority i.e. 18 years and above is entitled for a National ID.

6.2 Finding

According to the evaluation and test carried out, it was settled that this study has proved that utilization of cloud database in criminal background check is vital. The results suggested that using the prototype poised great advantage in processing and issuance of Police clearance certificate from every corner of the country where applicants can visit any nearest police station and leave the station with the certificate on their hands.

6.3 Recommendation

Based on the research experience during the study, we would like to recommend that more research be carried out in criminal background check using Cloud database and a lot of literature review be done for related work.

After successfully testing and implementing the prototype, in order to come to full realization of the study, we highly recommend adoption of the prototype by the Directorate of Criminal Investigation in order to solve the issues facing the current process of background check and issuance of Police Clearance Certificate. Further study should also focus on integration of the prototype wit Court system where criminal records can be easily entered and accessed. This will enable easy sharing of criminal records between the two Justice systems.

6.4 Challenges and limitations

Installing datacenter in a cloud computing environment comes with high cost. Despite the challenge we managed to host the prototype in one data center and web server package then used local servers to host other database nodes. Cloud technology is still an aspect that is virgin and many enterprises, businesses and government agencies in Kenya have not fully adopted it.

A lot of bureaucracy was also involved in the process of acquiring information from respondents. Some of them were not willing to provide required data but after lengthy talk they agreed to offer the required information.

Another challenge we encountered was inability to acquire cheap fingerprint scans that could be used in the development of the prototype. Other development tools used were trial versions thus could not run throughout the entire prototype development and testing phase.

6.5 Conclusion

The findings from the project showed that the current system and processes takes too long to process criminal background information and ensure timely issuance of Police Clearance Certificate. From my research findings, I can state with confidence that Cloud database and its application in criminal background check will yield positive results by far extend. This study also revealed that with the full development and implementation of the system, all the issues associated with centralized access and check of criminal background record will come to halt. The integration of fingerprint scanners into the system will eliminate the need for many applicants to travel long distance.

REFERENCES

Al Shehri, W., 2013. Cloud Database Database as a Service. *International Journal of Database Management Systems*, *5*(2), p.1.

Allen (2002), "Meaning of Crimes", Crime Management

Babbie, E. (2002). Survey research methods (2nd ed.). Belmont: Wodsworth.

Baker, S. E.,&Edwards, R. (2012). How many qualitative interviews is enough? National Center for Research Methods. Available at: http.

Creswell, J.W. and Clark, V.L.P., 2007. Designing and conducting mixed methods research.

Baker, S.E. and Edwards, R., 2012. National Centre for Research Methods review paper: How many qualitative interviews is enough? Expert voices and early career reflections on sampling and cases in qualitative research. *National Centre for Research Methods, Southampton, UK. Available at: http://eprints. ncrm. ac. uk/2273/4/how_many_interviews. pdf (accessed 4 October 2013).*

Beck, K., 2000. *Extreme programming explained: embrace change*. addison-wesley professional.

Bernard, H. R. (2002). Research methods in anthropology: Qualitative and quantitative approaches (3rd ed.). Walnut Creek, CA: Alta Mira Press.

Bernstein, P. A., Cseri, I., Et Al. 2011a. Adapting Microsoft SQL Server for cloud computing. In Proceedings of the 27th International Conference on Data Engineering. IEEE, 1255–1263.

Bickman, L., Rog, D.J. and Hedrick, T.E., 1998. Applied research design: A practical approach. *Handbook of applied social research methods*, *19*, p.19.

Brody, R. G. 2010. "Beyond the basic background check: Hiring the 'right' employees," Management Research Review 33(3): 210–223.

Buyya, R., Broberg, J., Goscinsky, A.: Cloud Computing: Principles and Paradigms. John Wiley and Sons (2011)

Byrne, J. A. 1999. Chainsaw: The Notorious Career of Al Dunlap in the Era of Profit-at-Any-Prince. New York: HarperBusiness.

C. Curino, E. Jones, Y. Zhang, and S. Madden. Schism: A Workload-Driven Approach to Database Replication and Partitioning.In VLDB, 2010.

Chandran, E. (2004). Research Methods. Nairobi: Starbright services limited.

Chang, F., Dean, J., Ghemawat, S., Hsieh, W. C., Wallach, D. A., Burrows, M., Chandra, T., Fikes,A., And Gruber, R. E. 2006. Bigtable: A distributed storage system for structured data.In Proceedings of the 7th Symposium on Operating Systems Design and Implementation. USENIX Association, Berkeley, CA, 205–218.

Cooper, B. F., Ramakrishnan, R., Srivastava, U., Silberstein, A., Bohannon, P., Jacobsen, H.-A.,Puz, N., Weaver, D., And Yerneni, R. 2008. Pnuts: Yahoo!'S hosted data serving platform.Proc. VLDB Endow. 1, 2, 1277–1288.

Curino, C., Jones, E., Popa, R., Malviya, N., Wu, E., Madden, S., Balakrishnan, H., And Zeldovich, N. 2011. Relational Cloud: A database service for the cloud. In Proceedings of the 5th Biennial Conference on Innovative Data Systems Research. 235–240

Decandia, G., Hastorun, D., Jampani, M., Kakulapati, G., Lakshman, A., Pilchin, A.,Sivasubramanian, S., Vosshall, P., And Vogels, W. 2007. Dynamo: Amazon's highly available key-value store. In Proceedings of the 21st ACM Symposium on Operating System Principles. ACM, New York, NY, 205–220.

Dworkin, S.L., 2012. Sample size policy for qualitative studies using in-depth interviews.

Duncan, A., Creese, S. and Goldsmith, M., 2012, June. Insider attacks in cloud computing. In *Trust, Security and Privacy in Computing and Communications (TrustCom), 2012 IEEE 11th International Conference on* (pp. 857-862). IEEE.

Ercan, T.: Effective use of cloud computing in educational institutions. Procedia – Social and Behavioral Sciences 2(2), 938–942 (2010)

"EPIC - EPIC v. FBI - Next Generation Identification." Accessed July 14, 2017. http://epic.org/foia/fbi/ngi/.

"FBI Using Facial Recognition despite Privacy Concerns." Accessed July 14, 2017. http://www.valleynewslive.com/home/headlines/FBI-using-facial-recognition-despite-privacy-concerns-311738091.html.

Etikan, I., Musa, S.A. and Alkassim, R.S., 2016. Comparison of convenience sampling and purposive sampling. *American Journal of Theoretical and Applied Statistics*, *5*(1), pp.1-4.

G. Boss, P. Malladi, D. Quan, L. Legregni,H. Hall. (2007). "Cloud Computing" www.ibm.com/developerworks/websphere/zones/hipods/

Gartner. (2011). *Gartner Identifies the Top 10 Strategic Technologies for 2012*. Retrieved fromhttp://www.gartner.com/it/page.jsp?id=1826214

Gilleo, W.M., 2014. *Examining continued use of software as a service in organizations* (Doctoral dissertation, Capella University).

Golafshani, N., 2003. Understanding reliability and validity in qualitative research. *The qualitative report*, *8*(4), pp.597-606.

http://enterprise.microsoft.com/en-us/industries/government/start-your-microsoft-cloud-for-government-trial-today/

http://www.theatlantic.com/technology/archive/2014/07/the-details-about-the-cias-deal-with-amazon/374632/

International Assoc of Chiefs of Police and United States of America, 2013. Guiding Principles on Cloud Computing in Law Enforcement.

International Data Corporation (IDC), "The 2011 digital universe: Extracting value from chaos," Online, June 2011, http://www.emc.com/collateral/demos/microsites/emc-digital-universe-2011/index.htm.

Jacobs, J. and Crepet, T., 2007. Expanding Scope, Use, and Availability of Criminal Records, The. NYUJ Legis. & Pub. Pol'y, 11, p.177.

Jacobs, J.B., 2005. Mass incarceration and the proliferation of criminal records. U. St. Thomas LJ, 3, p.387.

Journal of Network and Computer Applications, 34(1), 1-11. http://dx.doi.org/10.1016/j.jnca.2010.07.006

Judith, M. M. (2011). Cloud services: Mitigate risks, maintain availability. Retrieved from http://www.ibm.com/developerworks/cloud/library/cl-cloudservicerisks/cl-cloudservicerisks-pdf.pdf

Karuna (2011), What are the merits and demerits of purposive sampling method as used instatistics?, viewed on 17th July 2014

http://www.preservearticles.com/201107189302/whatare-the-merits-and-demerits-of-purposive-sampling-method-as-used-in-statistics.html

Kiriinya, D., 2014. *Factors Influencing Cloud Computing Adoption in Organisations: A Case of Cooperative Insurance Company of Kenya*(Doctoral dissertation, United States International University-Africa).

Kothari, C. R.(2000) *Research Methodology: Methods and Techniques*. New Delhi: Wiley

Kothari, CK (2004), *Research Methodology: Methods and Techniques (Second revised Edition)* New Age International (P) Limited, Publishers.

Kshetri, Nir (2010) "Cloud Computing in Developing Economies", IEEE *Computer*, October 43(10), pp. 47-55.

Loudhouse Research: Cloud Barometer Survey 2010. www.mimecast.com/barometerresearch2010.

Mason, M. (2010). Sample size and saturation in PhD studies using qualitative interviews. Forum: Qualitative Social Research, 11(3)[Article No. 8].

Mell,P.,&Grance, T. The NIST Definition of Cloud Computing. Gaithersburg: NIST.2011.

Minqi, Z., Rong, Z., Wei, X., Weining, Q., &Aoying, Z. (2010). Security and Privacy in Cloud Computing: A Survey. *Semantics Knowledge and Grid (SKG)*. 2010 Sixth International Conference on (pp.105-112).

Mugenda, M., O., &Mugenda, G., A. (2003). Research Methods. Nairobi: Acts Press.

Mugenda, O.M., 1999. *Research methods: Quantitative and qualitative approaches*. African Centre for Technology Studies.

Musee, N.M., 2015. *An academic certification verification system based on cloud computing environment* (Doctoral dissertation, University of Nairobi).

MUTHONI, J., 2015. E-VERIFICATION–A CASE OF ACADEMIC TESTIMONIALS.

Mukami, T. (2014, January 11). *Got FAKE Documents? How HR Managers Will Get You*. Retrieved July 7, 2014, from http://www.careerpointkenya.co.ke/2014/01/11/fake-documents-how-hr-managers-will-get-you/

"Next Generation Identification (NGI) ? Retention and Searching of Noncriminal Justice Fingerprint Submissions." FBI.Accessed July 14, 2017. https://www.fbi.gov/foia/privacy-impact-assessments/next-generation-identification-ngi-retentionand-searching-of-noncriminal-justice-fingerprint-submissions.

Oates, B.J., 2005. *Researching information systems and computing*. Sage.

Otieno, D. (2013, September 2013). *Kenya, land of fake goods, fake leaders, fake smiles.* Retrieved June 26, 2014, from http://www.standardmedia.co.ke/?articleID=2000092585&story_title=kenya-land-of-fake-goods-fake-leaders-fake-smiles

Ouf, S., Nasr, M.: Business intelligence in the cloud. In: IEEE 3rd International Conference on Communication Software and Networks (ICCSN 2011), pp. 650–655 (2011)

P. Mell and T. Grance, "Draft nist working definition of cloud computing - v15," *21. Aug 2009,* 2009.

Pager, D., 2003. The Mark of a Criminal Record1. American journal of sociology, 108(5), pp.937-975.

Pocatilu, P., Alecu, F., Vetrici, M.: Measuring the efficiency of cloud computing for e-learning systems. W. Trans. on Comp. 9, 42–51 (2010)

Ponemon, L., Gould, J., Roberts, D.J., International Assoc of Chiefs of Police, United States of America, SafeGov, United States of America, Ponemon Institute and United States of America, 2013. Leveraging the Cloud for Law Enforcement.

Powell, R. R. (1997). Basic Research Methods for Librarians. 68.

R. Buyya, et al., Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility, Future Generation Computer Systems (2009), doi:10.1016/j.future.2008.12.001

Robyn, U., Ifeoma, A., Alex, R., &Danah, B., (2015). Biometric Technologies in Policing

Somani, U., Lakhani, K., &Mundra, M. (2010). Implementing digital signature with RSA encryption algorithm to enhance the Data Security of cloud in Cloud Computing. *Parallel Distributed and Grid Computing(PDGC)*, 2010 1st International Conference on (pp. 211-216).

Subashini, S., &Kavitha, V. (2011). A survey on security issues in service delivery models of cloud computing.

Tsai, H., Siebenhaar, M., Miede, A., Huang, Y., & Steinmetz, R. (2012). Threat as a Service?: Virtualization's Impact on Cloud Security. *IT Professional, 14*(1), 32-37. http://dx.doi.org/10.1109/MITP.2011.117

Vaquero, L. M., Rodero-Merino, L., & Moran, D. (2011). Locking the sky: a survey on IaaS cloud security.Computing, 91, 93-118. http://dx.doi.org/10.1007/s00607-010-0140x

Venkatesh, V., Morris, M.G., Davis, G.B. and Davis, F.D., 2003. User acceptance of information technology: Toward a unified view. *MIS quarterly*, pp.425-478.

Voorsluys, W., Broberg, J. and Buyya, R., 2011. Introduction to cloud computing. Cloud computing: Principles and paradigms, pp.1-44.

Williams, L., 2010. Agile software development methodologies and practices.*Advances in Computers*, *80*, pp.1-44.

APPENDICIES

Appendix A:Police Interview Questionnaire

Dear Sir/Madam

You are invited to participate in a research study investigating the implementation of a prototype for criminal background check using a cloud database. The interview questions are designed to collect data to help in the implementation of the prototype

The Information you avail will be used for academic research purpose only and will be held with high confidentiality.

Thank you
1. What process takes place when doing background check for individuals?

2. Which Method of Biometric Recognition technology do you use for fingerprint

   matching?

3. What challenges do you encounter when matching fingerprints?

4. Do you encounter occasional server failure due to increased workloads?

5. Do you have comprehensive disaster recovery plan for criminal records in case of server

   failure?

6. Do you know what cloud database means?

7. How long does it take to process and issue police clearance certificate?

8. Will having distributed database nodes improve access of criminal record?

9. Using the same distributed database nodes, will this ease your job workload?

10. Will Processing the police certificate at police station level improve process delivery?

11. If a prototype is developed for criminal background check, will it be of help to you?

12. Will having biometric system for fingerprints improve on the process?

Employers Questionnaire Form

Dear Sir/Madam

You are invited to participate in a research study investigating the implementation of a prototype for criminal background check using a cloud database. The questions are designed to collect data to help in the implementation of the prototype

The Information you avail will be used for academic research purpose only and will be held with high confidentiality.

Thank you

1. Gender
   - ☐ Male
   - ☐ Female
2. Level of Education.
   - ☐ PhD
   - ☐ Masters
   - ☐ Bachelors
   - ☐ Diploma
   - ☐ Certificate
   - ☐ N/A
3. What is the nature of your organization?
   - ☐ Government/Parastatal
   - ☐ Business
   - ☐ Educational
   - ☐ Other _____

4. Do you demand a Police Clearance Certificate when offering job to someone?
   - ☐ Yes
   - ☐ No
5. Have you ever verified the authenticity and validity of the Police Clearance Certificates?
   - ☐ Yes
   - ☐ No
6. If yes, what method of authentication do you use?

- ☐ Physically going to the CID Office
- ☐ Sending SMS/ Short Code/ USSD
- ☐ Mere Trust
- ☐ Other_____

Appendix B: Sample Source Code

Code for Central node

```php
<?php

//start the session
session_start();
ob_Start();
//include the database
require'db.inc.php';
//get the variables and urlencode them
$fname=urlencode($_POST['fname']);
$mname=urlencode($_POST['mname']);
$lname=urlencode($_POST['lname']);
$email=urlencode($_POST['email']);
$telephone=urlencode($_POST['telephone']);
$idno=urlencode($_POST['idno']);
$pic_data = base64_encode(file_get_contents($_FILES['image']['tmp_name']));
//$pic_data = addslashes(file_get_contents($_FILES['image']['tmp_name']));
$pic_size=$_FILES['image']['size'];
$pic_data= urlencode($pic_data);
$pic_size= urlencode($pic_size);

//get all the nodes
$thenodes="SELECT * FROM nodes";
$thenodes2=mysql_query($thenodes);
//for each node, send the data
while ($thenodes3=mysql_fetch_array($thenodes2))
{
    $base_url=$thenodes3['url'];
    $page="newapplication.php?";

$other="fname=$fname&mname=$mname&lname=$lname&email=$email&telephone=
$telephone&idno=$idno&pic_data=$pic_data&pic_size=$pic_size";
    $url="$base_url$page$other";
    echo $url;
    $response = file_get_contents($url);
    echo $response;
    //if the record is found in any node
 if ($response="found")
   {
   $_SESSION[response]=$response;}}
```