BLOCKCHAIN – IT'S PRACTICAL USE FOR NATIONAL DATA CENTRES

A Survey of the pactical use case of Blockchain Technology in National Data Centres

Mulima Chibuye
Department of Computer Science
University of Zambia
Lusaka, Zambia
Mulima.Chibuye@cs.unza.zm

Jackson Phiri
Department of Computer Science
University of Zambia
Lusaka, Zambia
Jackson.Phiri@cs.unza.zm

Abstract— E-government has been an essential part of the targets set forth by all governments that are signatory to the United Nations 2030 sustainable development goals which am to create effective, accountable and inclusive governments. In order to achieve e-government in the most efficient manner, more and more governments around the world are embracing cloud storage and cloud applications as the de facto standard for data storage, archiving, analysis and manipulation. The amounts of money that governments will spend are huge and the infrastructure requirements far exceed the capabilities of most government institutions. Models such as centralized management of cloud infrastructure have been adopted but this defeats the purpose of cloud computing where resources and management of those resources are decentralized. In this paper we present a proposal for a standard model for governments to utilize their existing or planned cloud infrastructure and leverage private and public cloud infrastructure in order to achieve true cloud decentralization. We propose that Blockchain technologies provide the most optimal and quickest methods towards e-government. We look into the problem of analyzing big data in national data centres and methods of optimization based on already available technology.

Keywords—internet of things; blockchain; big data applications; cloud computing;

I. INTRODUCTION

National data centres are a basic requirement in e-governments across the world. The traditional models of running government worldwide have remained unchanged despite the advancements in data science and storage technologies. In the 1990's Government Process Reengineering which relied on Business Process Reengineering was proposed as a method to lead to leaner and more efficient governments[1]. This naturally led to what was termed the information based government which also evolved to the current e-government being implemented in most counties across the world.

In order to introduce efficiencies into the running of government institutions and avoid the duplication, unnecessary storage of data and global sustainable development goals as set forth by the United Nations, governments around the world have embarked on e-government projects. A typical example is the government of Zambia which has embarked on the e-government project which is a $440 million[2] project that aims to connect all government offices with 12,000km[3] of fiber using a FTTB(fiber to the building) model. These individual units will then have access to a common data centre where all data will be stored. The case of the Zambia national data centre is used as a generic model for most national data centres worldwide.

A. Cloud Computing

With the spread of broadband internet across the globe, more emphasis has been placed on how to effectively and efficiently utilize and allocate all the available computing, processing and storage capacity available globally. A reduction in latency and increased throughput has made it possible for remote computers to be managed and applications run as though they were physically local. Resource sharing has enabled concepts such as Software as a Service, Platform as a Service and Database as a Service where the end user does not need to worry about the technical requirements of systems, software and the technical specifications of the database but just focuses on service usage[4]. Cloud computing has made it possible for emerging companies to rollout their services at a much faster pace as the cost of datacentres is slowly becoming a non-factor with cloud based solutions such as Data Centre as a Service (DCaaS).

A very good summarization of what cloud computing is has been coined by the United States National Institute of Standards and Technology which states that cloud computing is a model that enables the convenient, ubiquitous network access of on demand and shared resources with rapid deployment, decoupled management and limited service provider interaction.

B. Case Studies – National Data Centres

1) The Zambia National Data Centre

The Zambia National Data Centre (NDC) is an initiative by the government of Zambia to deploy cloud infrastructure to be used for the following services;

- Infrastructure-as-a-Service
- Storage-as-Service - StaaS
- Backup-as-a-Service
- Disaster-Recovery-as-a-Service DRaaS
- Software-as-a-Service

The main purpose of the Zambia National Data Centre which is managed by Center of Excellence for e-Government and ICT, a wing of the Ministry of Communications and Transport is to consolidate government ICT resources hence introduce efficiencies in the acquisition, management and maintenance of those resources. Figure 1 below is a summary from the Ministry of Finance and National Planning page on the goals of the project.
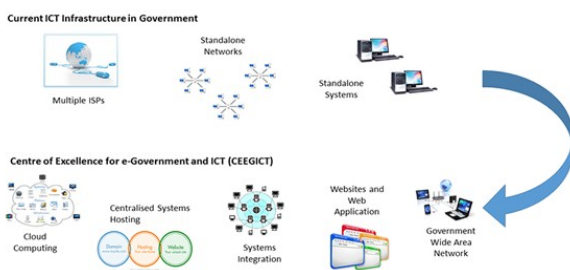

Fig. 1. *An example of the implementation of the Zambia National Data Centre*

From a redundancy and resilience point of view, we can already note that centralized systems hosting presents significant future challenges for this project. It is quite possible that the government will build a tier IV data centre that provides active redundant components as concerns power and cooling but this does not protect from localized geographic force majeure.

The current implementation of the Zambia National data is an example of a centralized approach towards national data centre implementation. We also looked at the case of a more advanced data centre implementation in the United States prior to the terrorist attacks of September 11, 2001 under the Department of Homeland Security. The United States Department of Homeland Security Data Centres were picked as a case study because they were the first data centres in the United States to undergo optimization projects.

2) The United States Department of Homeland Security National Data Centre

Prior to 2001, the United States government's department of homeland security had several data centres spread out in various areas of the country. According to a report on data centre consolidation by the department of homeland security, part of the reason for the success of the mission set forth by the department was "integrating the incumbent workforce,

processes, and systems into a unified organization"[5]. In the case of systems, these were related to the various data centres spread out throughout the country. The data centres were linked by 7 individual wide area networks. It was noted that despite the advances in e-government, there remained one fundamental problem which led to inefficiencies in collaboration amongst government stakeholders following the 2001 terrorist attacks an according the report these are highlighted below;

- Limited corporation
- Limited communication
- Limited data sharing
- Limited protection

It was therefore decided by the United State Government that a data centre consolidation project be undertaken in order to address the inefficiencies highlighted above. These consolidations would however maintain the basic requirements of distributed cloud architecture and bring about the following benefits;

- Significant operational cost reduction of more than $200m per year
- Standardizing IT resources
- Streamlining maintenance
- Expediting response times in case of emergencies

The United States Government has not ended its optimization campaign and through two more iterations, several government organizations are now implementing the Federal Data Center Consolidation Initiative[5].

C. The Shared Economy

By its definition, the shared economy aims to make optimal use of resources. Since the premise on cloud computing is also optimizing resources, we infer that a natural evolution of cloud technology is towards the shared economy[6]. With concepts such as the internet of things, it becomes possible to implement machine to machine (M2M) communication in a trustless manner because computer code will never lie and will always follow the programmed algorithm unless it is programmed as artificial intelligence neural networks with autonomy. Classic implementations of the shared economy include Airbnb[7] where hospitality is abstracted as a "thing" on the cloud and the rent a bicycle concept in China where each bicycle is a "things" and supports machine to machine communication[8]. End beneficiaries use a mobile app to get permission to the use of either an apartment in the case of Airbnb or a bicycle in case of the e-bike in China.

As the government of Zambia begins to implement cloud technology, we can argue that the implementation will naturally progress towards an optimal and cost effective model leading to the shared economy model. We build a model

towards the shared economy beginning with the implementation of private clouds, hybrid clouds and finally true shared decentralized clouds with end to end encryption for all the data.

II. BLOCKCHAIN

A. Definition

A Blockchain is an immutable public ledger of all transactions that have ever happened in a system. Blockchain relies on private/public key cryptography in order to achieve the immutability of stored records. While past records may not be changed, future records can link to past records with transactions thereby making a recent changed copy of a past record. The beauty about Blockchain is that it leaves an audit trail. It is possible to erase all past records in a Blockchain using a consensus algorithm whereby all participating systems within a Blockchain agree to change a record in the past and replay all future transactions after the change. Fortunately, there will be an audit trail of a past record having been altered.

In this paper, we make the argument that governments make use of publicly available Blockchain in order to provide data redundancy for the national data centres and once the technology advances further, governments may entirely rely on Blockchain technology to store and manage their data.

While there may be a concern as to how secure a public Blockchain would be in keeping government data secure, there exists advanced encryption standards such as the BB84 protocol that can be applied. This protocol makes it mathematically very hard to break the encryption code. An argument is put forth that information stored digitally and securely is safer than information stored in an office and on a file that even a secretary can have access to[9]. Information stored on file has the following risks;

- Third party access
- Files accidently left in an unsecure location
- Ease of duplication
- Natural disasters
  - Fire
  - Floods
- Theft
- Trust

With Blockchain Technology, only the user with access to the private key may have access to the files. Also, it is possible to implement multi key signatures for approval hierarchies where more than one user is required to approve a transaction.

B. Benefits of Blockchain to Governments

It has already been noted as to the broad benefits of Blockchain technology but here we look at the specific benefits to the quest for government's world over to implement national data centres.

Blockchain technology allows for the storage of data on every participating party's system hence offering unparalleled redundancy. They also offer the possibility of the implementation of smart contracts which are defined as peer to peer arguments enforced by code on a Blockchain network. These smart contracts can do many things such as copy and manipulate data that is stored on a Blockchain. The network can be sponsored by untrusted parties because only the user with a corresponding private key has access to data that was encrypted with a public key. It is even possible for a public/private key combination corresponding to a certain data set to be changed every time a transaction occurs on a Blockchain network. There are various possibilities of implementation of a Blockchain network and its characteristics are far superior to the traditional databases as used in cloud computing today.

Because of the fundamental characteristics of Blockchain technology, it is suggested that this be a standard implementation model for the e-government requirements of data storage facilities. Eventually, it might not be necessary for a government to maintain multiple data centres because Blockchain technology is true infrastructure as a service, storage as a service and data manipulation as a service.

C. Big O of Data Storage, Access and Manipulation

The big O notation is the standard method of analyzing the performance or complexity of an algorithm. Given that the ideal national data centre is required to host all the government data requirements from each of its institutions in a decentralized manner; it can be argued that the storage and analysis of data becomes a complex activity. E-government projects are primarily concerned with big data and according to the United Nations 2016 E-government survey report, the 2030 sustainable development goals have at their coal an open and inclusive government with big data available for interpretation by citizens and the world in general[10]. Storing and analyzing big data brings inherent storage and processing complexity. In big O notation, we refer to the storage and analysis complexity as space and time complexity respectively.

Because the nature of a data centre is to have high availability and data redundancy, there will be multiple instances of big data stored on servers which increases the data access and storage algorithm complexity. Fortunately, there are schemes such as Hadoop and Map Reduce that can be employed in the data centre. Hadoop is the open source implementation of Map Reduce.

With decentralized storage, the order of magnitude of computational complexity becomes exponential unless parallel search algorithms are deployed. It has been shown that multiple sources of data and parallel search algorithms reduce the space and time complexity of big data hence providing a more efficient big O rating. Because Blockchain are distributed

ledgers, we infer that we will have better time and space complexity if used as a national database.

1) Geographic Redundancy

While there does not seem to be any plans yet to enable geographic redundancy for the Zambia National Data Centre, this is the natural upgrade from the current tier VI setup.

Phase one of the national data centre will be located in Lusaka only. This presents an obvious problem as concerns availability and data redundancy. With no geographic redundancy, data loss and corruption become a big issue.

III. METHODS

Using a mixed methods research approach, we explore the proliferation of data centres in developing countries and rely on past literature in order to define a conceptual model for the optimized evolution of national data centres. We then build a model and glide path for the progression of national data centres towards a true decentralized cloud platform benefitting from a shared economic model using Blockchain technology and discuss each of the steps with pros and cons. We look at practical implementations from other governments and corporations.

For the protection of data, we recommend further research into the best possible encryption standard to follow. Since we are proposing a potentially future proof true decentralized national data centre, we employ the use of more advanced data encryption standards to outlast progress in computers that make use of Shor's algorithm and Grovers's algorithm to crack encryption codes. More work will be required in the future in order to improve cryptography post quantum computers.

A. Data Centres in Developing and Developed Countries

1) Honduras Land Titles on Blockchain

Honduras is a mostly Spanish speaking republic in Central America with a population of 9.1 million[11]. The case of Honduras is particularly interesting as this country opted to use Blockchain technology exclusively for the storage of land title[12], unfortunately, the project stalled but it is our view that we will continue to see such initiatives as the drive towards resource optimization, efficiency and the shared economy is pushed world over. The Honduras government partnered with a company called Factom[13] and their Factom Blockchain[14] which is in its most basic term a distributed database or ledger with Blockchain technology as the underlying foundation. Factom builds application on top of their Blockchain and supports large transactional volumes.

2) Swedish Land Registry using Blockchain Technology

In early 2017 The Swedish Mapping, Cadastre and Land Registration Authority embarked on a conceptual project to record land title information and real estate transactions using Blockchain technology for the following reasons[15];

- Faster and more transparent transactions
- Elimination of the possibility of selling a property more than once
- Greater security for users of the system due to the possibility of independent verification of transactions

B. Glide Path to True Cloud Architecture

Traditionally, governments embark on e-government projects with the following stages;

1. Set up national communications network
2. Set up data centres

The two stages highlighted above have a significant capital and operational expenditure. We make the argument that both (1) and (2) above have been set up or are in the process of being set up by private corporations in several countries around the world. Governments can use regulation to leverage existing IT infrastructure thus reducing their capital expenditure and relying more on operational expenditure in order to achieve the goals of e-government.

The characteristics of true cloud architecture are highlighted below;

- On demand self-services [16]
- Broad network access
- Resource pooling
- Fault tolerance
- Rapid elasticity
- Measured service
- Multi Tenacity

We construct a matrix and explore through literature the best model to follow.

Selecting thegBes gDa ag Cei regModel			
Ctharac eristc	Researcthgquestoi	Me thodolony	S atstcalgaialysis
On demand self-services	Which data centre model has the quickest deplyment	literature review	mean
Broad network access	Which data centre model provides the best	literature review	mean
Resource pooling	Which data centre model provides the best model for	literature review	mean
Fault tolerance	Which data centre model provides the most optmal	literature review	mean
Rapid elastcity	Which data centre model is	literature review	mean
Measured service	Which data centre model	literature review	mean
Mult Tenacity	Which data centre model can accommodate several	literature review	mean

Fig. 2. *How we analysed the required and generally accepted characteristics of true cloud computing*

The questions in figure 1 above were answered by looking at 3 main contributors to the field of cloud computing and what their conclusions were on the best model that fit the required characteristics of cloud computing. The contributors that were purposively selected include Massachusetts Institute of Technology (MIT) due to their research in methods of improving time complexity of messages in data centre environments[17], Gartner research[18] [19], due to their track record in consistently providing factual data and research into technological trends and the United States Department of Homeland Security because they were selected as a case study for this proposal.

Research question	Gartner research	US Dept. Homeland Security	MIT
Which data centre model has the quickest deployment time?	Hybrid clouds are favoured.	Outsourced	-
Which data centre model provides the best networking model?	Software defined networking is preferred.	Cloud	Distributed storage, centralized management
Which data centre model provides the best model for resource pooling?	Cloud computing.	Hybrid cloud	-
Which data centre model provides the most optimal time and space complexity?	-	-	Distributed storage, centralized management
Which data centre model is easily scalable?	Open platforms	-	-
Which data centre model has the least cost?	Cloud based	Hybrid cloud	-
Which data centre model can accommodate several users easily?	Cloud based	-	-

Fig. 3. *Synthesis matrix used in selecting the best data centre deployment model*

C. *Blockchain Proposal and Synthesis Matrix*

It can be noted that none of the contributors selected in the synthesis matrix actually mention Blockchain as an option for a data centre model. We make the argument that Blockchain data centre models being built on top of cloud infrastructure is therefore a superset of what is termed cloud based. Blockchain based models can encompass private clouds, public clouds and hybrid clouds. They can also offer the major requirements in data centres such as the platform, the security and the networking, all provided as a service. Interestingly, the model suggested by MIT in decentralized storage but centralized management does not restrict the use of Blockchain technology as applications can be built with flexibility in the mode of networking because we can view Blockchain solutions as software defined networks. We therefore propose that since cloud based approaches are preferred in most data centre leverage matrix, a Blockchain based approach provides the most optimal model. We propose that any government wishing to implement a national data centre immediately consider the Blockchain option for the following reasons;

1. Blockchain provides true Platform as a Service, Database as a Service and Infrastructure as a Service.

This is further highlighted in the choice by the analyzed contributors to mostly prefer a cloud based approach towards key characteristics of data centre requirements.
2. The infrastructure costs of Blockchain implementations for data centres are much lower.
3. The time to deployment is the quickest as governments may leverage already existing Blockchain implementations.
4. Citizens and private corporations within a country may be incentivized to contribute free resources to the countries Blockchain hence easily providing freely available scalability and additional computing resources.

We propose that any government's data centre must adhere to the above generally accepted characteristics of true cloud systems. Our glide- path therefore proposes a 4 stage process towards the evolution of government's data centres into Blockchain architectures. We have already looked at the case of the United States which has advanced data centre implementations but was forced to embark onto optimization projects. We have also analyzed the case of Zambia where the data centres are currently being deployed. From the study of the US case, it will save governments progressive amount of money if cloud based approaches are used towards data centre implementation and optimization. Our four, stage approach is as follows;

1. Establish a data centre or use an existing data centre
2. Backup data on a third party network and outsource selected services
3. Use a hybrid cloud model with Blockchain integration
4. Outsource the data centre and progress towards a true cloud model and shared economic model relying on Blockchain technology

Most governments are currently in the first phase while the United States Government is in the second phase. With the current projects being implemented by various departments within the US government, we can clearly see that they are heading towards the third stage of the model. We argue that further progressing towards the fourth stage or immediately implementing a national data centre using the fourth stage approach will reduce overhead and introduce further savings making the governments IT infrastructure more agile, robust, redundant and easily scalable with minimal operational overhead.

IV. CONCLUSION AND RECOMMENDATIONS

From the study of the two cases of Zambia and the United States presented in this paper, it can be seen that cloud computing and the shared economy are the apparent goals of e-government but the implementation approach lacks

standardization therefore introducing unnecessary costs associated with the implementation and further optimization of these projects. The largest costs associated with e-government projects are the infrastructure costs which this paper argues could be reduced to fractions of what they are today or completely eliminated if existing Blockchain are leveraged. With advancements made in cryptography, cloud computing, data access and storage techniques, governments world over can evolve seamlessly towards true cloud governments by taking advantage of the models implemented in current and future Blockchain technologies. The inherent nature of governments is to keep information in an unaltered state for reference by future generations. The internet and its latest application, Blockchain technologies provide the best solution for the continuation of government.

V. REFERENCES

[1] Z. Li and F. Yang, "The E-government Information Model Based on GPR," *Procedia Comput. Sci.*, vol. 91, pp. 193–200, Jan. 2016.

[2] "Data centre launch set – Zambia Daily Mail." [Online]. Available: https://www.daily-mail.co.zm/data-centre-launch-set/. [Accessed: 03-Aug-2017].

[3] "Tuesday, 20th June, 2017 | National Assembly of Zambia - E-Government." [Online]. Available: http://www.parliament.gov.zm/node/6419. [Accessed: 03-Aug-2017].

[4] S. Sharma, "Evolution of as-a-Service Era in Cloud." Center for Survey Statistics and Methodology, Iowa State University, Ames, Iowa, USA.

[5] United States Department of Homeland Security, "Federal Data Center Consolidation Initiative." Oct-2011.

[6] J. Sun, J. Yan, and K. Z. K. Zhang, "Blockchain-based sharing services: What blockchain technology can contribute to smart cities," *Financ. Innov.*, vol. 2, no. 1, Dec. 2016.

[7] L. Zekanovic-Korona and J. Grzunov, "Evaluation of shared digital economy adoption: Case of Airbnb," 2014, pp. 1574–1579.

[8] K. Flüchter and F. Wortmann, "Implementing the Connected e-Bike: Challenges and Requirements of an IoT Application for Urban Transportation," in *Proceedings of the First International Conference on IoT in Urban Space*, ICST, Brussels, Belgium, Belgium, 2014, pp. 7–12.

[9] G. P. Johnston and D. V. Bowen, "The benefits of electronic records management systems: a general review of published and some unpublished cases," *Rec. Manag. J.*, vol. 15, no. 3, pp. 131–140, Dec. 2005.

[10] "UN E-Government Survey 2016." [Online]. Available: https://publicadministration.un.org/egovkb/en-us/reports/un-e-government-survey-2016. [Accessed: 03-Aug-2017].

[11] "Honduras | Data." [Online]. Available: http://data.worldbank.org/country/honduras. [Accessed: 03-Aug-2017].

[12] Heimler, Marissa, "Intellectual Property Rights: A proxy to the continuous economic dependence of South American economies on the United States," University of Colorado, Boulder, 2017.

[13] "FAQ - Factom." [Online]. Available: https://www.factom.com/about/faqs. [Accessed: 03-Aug-2017].

[14] S. Singh and N. Singh, "Blockchain: Future of financial and cyber security," in *2016 2nd International Conference on Contemporary Computing and Informatics (IC3I)*, 2016, pp. 463–467.

[15] Lantmäteriet, Landshypotek Bank, SBAB, Kairos Future, Telia company, and ChromaWay and, "The Land Registry in the blockchain - testbed." Mar-2017.

[16] C. Gong, J. Liu, Q. Zhang, H. Chen, and Z. Gong, "The Characteristics of Cloud Computing," in *2010 39th International Conference on Parallel Processing Workshops*, 2010, pp. 275–279.

[17] "Fastpass: A Centralized 'Zero-Queue' Datacenter Network." [Online]. Available: http://fastpass.mit.edu/. [Accessed: 23-Aug-2017].

[18] "Gartner Says the Future of the Data Center Is Software-Defined." [Online]. Available: http://www.gartner.com/newsroom/id/3136417. [Accessed: 23-Aug-2017].

[19] "The Hybrid Cloud, Cloud Services, & Your Cloud Strategy - Gartner." [Online]. Available: http://www.gartner.com/technology/topics/cloud-computing.jsp. [Accessed: 23-Aug-2017].