Cloud Computing in Managing Big Data

Terry Mungai

***Abstract*—As the world becomes technologically advanced, so has the collection and accumulation of big data. With this comes the perceived vision that cloud computing will become inevitable in the near future. The purpose of this paper is to evaluate ways and means of how Cloud Computing in Managing Big Data can be advantageous in storing, analyzing and extracting useful information. The outcome is seen to favorably impact sustainability of enterprises as executives are afforded a wide array of choices for making predictive decisions in business models and improving customer satisfaction.**

***Index Terms*—Big Data; Cloud Computing; Data Centers; Enterprise Customers; Price and Performance; Security.**

I. OVERVIEW

During the last decade, big data has been growing exponentially leading to an urgent call for alternative storage. Cloud computing, due to its scalability, easy access, quick deployment as well as low costs in usage has allowed for users to embrace it over the traditional methods of data storage and analysis. More and more resources have been poured in to researching cloud storage and big data as governments and individuals realize its importance. The main concerns are applications scalability, security and how much value small enterprises can benefit from that [1]. There's a noted wealth of resources in cloud computing where users can optimize their needs based on data amount, analysis and dissemination. With these needs, it is of essence that cloud providers be in a position to provide quick deployment as well as efficiency. A review by both governments and enterprises to validate feasibility would be necessary [2].

II. RESEARCH METHODOLOGY

A. Review Stage

Using extensive course work notes, data was obtained to help with research of the paper. Also, with the help of databases supplied by University of Bridgeport, a wide variety of journals were obtained which included Library, Information Science & Technology, IEEE Xplore amongst many others. In this paper, information was collected from various information science academic journals, computer science academic journals as well as text books.

B. Research Process

The research paper is primarily focused on the exponential data collection by institutions as well as individuals and how best to manage, analyze and create value for users from this big data. With the help of databases provided by the University of Bridgeport, numerous journals have been collected and analyzed in an effort to bring to light how cloud computing can be of value to managing big data. It also emphasis on how best to utilize big data and examines its benefits. It reiterates on a definition of big data and offers various objectives of analyzing big data which essentially allows for better business decisions hence leading to successful business models [3].

III. CLOUD COMPUTING IN MANAGING BIG DATA

Fig. 1. Big Data Space Management

Cloud computing has become a fundamental need in our everyday lives [4]. Data is growing exponentially on daily bases and the only way for us to be able to handle the massive amount of data generated is by using cloud for storage[5]. When data is aggregated, then it becomes easier to maintain. Security is easily employed and wasteful use of resource minimized [2]. I agree with the fact that we as consumers should be able to have basic understanding on how the environment can be protected as cloud computing affects the environment directly [1]. Date integrity must be maintained at all costs in order to secure customers information [6]. I therefore support that cloud computing should be supported and encouraged by both the governments and consumers [7].
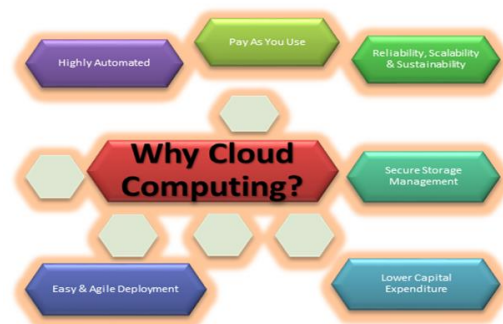
A. Data Centers

Fig. 2. Why Data Centers for Cloud Computing

Terry Mungai is with the Department of Computer Science and Technology, University of Bridgeport, Bridgeport, CT, USA.

Cloud data centers are on the rise as we witness big companies like Microsoft, IBM and google compete to launch data centers [8]. The main goal for cloud providers is to cater for the tremendous amount of data being populated daily by individuals and corporations [9]. I agree that data centers may have security issues which subsequently makes some people refuse to embrace the technology [10]. Also, I think that the massive amount of data could reach to bottleneck if data centers are not advanced in a timely manner [11]. I agree that there is a lot of promise with data centers and cloud computing [12]. Finally, it is fundamental to note that, with big data explosion, governments and enterprises across the globe will be willing to come to together in an effort to find better and efficient data storage methods. However, this does come with ramifications, one of which would be increased lawsuits to cloud providers [13].

B. Security in Building and Running Data Centers

Fig. 3. Building Security

I believe that fear plays a big hindrance when it comes to adopting cloud computing [14]. Over the course of time, there will be trust in cloud computing that will get instilled in people [15]. Technology is getting better [16]. Schools are now putting tremendous amount of time in technology investment therefore as students go on to graduate, they are equipped with proper tools of software development and engineering that can be applied in cloud computing security [17]. It would therefore be of importance for governments and enterprises to adopt to cloud computing as this is a better way to utilize the available resources [18]. Much as agility in adopting cloud computing is slow, I strongly feel that cloud computing will be the only option in managing and storing data as local servers and personal computers gets phased out [19].
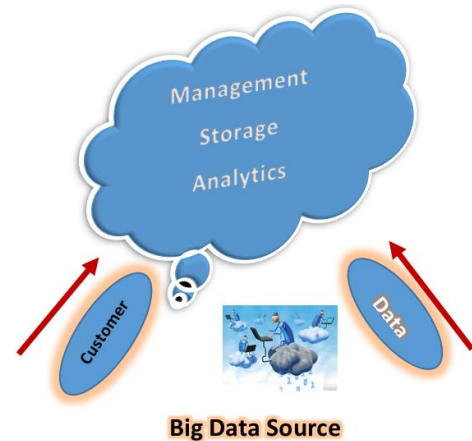
C. Enterprise Customers

Fig. 4. Sources of Big Data and its Management

The research that has been conducted shows a clear progress in cloud computing [20]. Intellectual resources have heavily been invested in cloud computing and especially in security [21]. The seamlessness in integrating data from enterprises like hospitals or government in to cloud is obviously very sophisticated and secure [22]. This is return is creating trust in companies like IBM, Google, Oracle and other cloud providers [23]. These companies are not only known to invest heavily in human intelligence but also hire the brightest of their engineers [24]. I strongly agree that a lot of emphasis should be put in schools to educate students on cloud computing and especially where security is concerned [25].
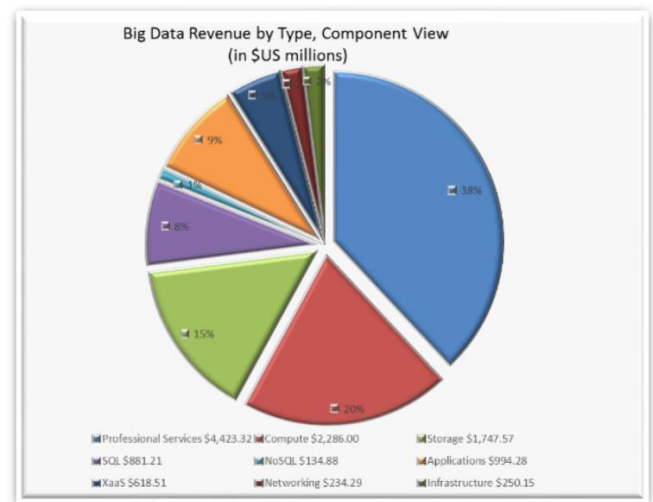
D. Price and Performance

Fig. 5. Big Data Revenue by Type

Cloud computing, even though relatively new in technology advancement should not be exploited by providers when it comes to pricing [26]. There should be a fair and reasonable scale in setting prices especially for small companies and individuals [27]. I strongly feel that the government and schools should put a big emphasis on educating students on cloud computing and especially security that pertain to it [28]. I agree that cloud computing is under-utilized and therefore, providers must come up with ways in which to maximize utility of the resources [29]. It is also important to note that relational IT capability is a

fundamental point in facilitating cloud success compared to technical and managerial IT capabilities [30]. I conclude that in order for cloud computing service to succeed, acceptable pricing mechanism both by users and the service provider should be maintained [31].

IV. DISCUSS MODEL

Even though still in its infancy, cloud computing is starting to gain momentum as governments, enterprises as well as individuals start to realize the urgency to store, manipulate and extract value from extremely large data sets. New computing paradigm also known as cloud computing has become a household name and as such, is being integrated or applied to enterprise architectures [32]. Cloud computing is breaking barriers in computing world and is seen to deliver favorable outcome to users. "The approach relies on a number of existing technologies e.g., the Internet, virtualization and grid computing. However, the provision of this service in a pay-as-you-go way through the popular medium of the Internet renders this computing service approach unique compared with currently available computing service modalities [33]." In the past few year, medical facilities have increasingly been adopting Electronic health records partly due to increased funding from the US federal government. Data collected offer the opportunity to identify possible improvements in health care delivery through use of data analysis as well as data mining and will also enable a wide variety of new applications that will be meaningful to patients and medical professionals. The US federal government in collaboration with research labs offer researchers access to health care data to assist in data analysis. It is important to note that HIPPA safe guards patients data at all means therefore monitoring access and usage of patients' data [34].

V. RESULTS AND FINDING

The impact of cloud computing for organizations around the world has led to the advocacy for a better understanding of how different geographic regions could assimilate their difference in technology in order to come to a better and stronger usage of information technology. Only through identification of these differences can countries come to better the structure of cloud computing and monitor data collection as well as usage [35]. It is true to say that both governments and enterprises have, for the past decade, put tremendous amount of efforts in understanding big data and essentially it's dissemination [36]. ICT (information and communications technology - or technologies) has been known for its growth mainly in part because of its low prices, adaptability to digital services, cloud computing as well as high customer margin. Therefore, a lot of companies geared towards social media have embraced it due to its innovative tendencies. Out of this, estimation on how much usage of ICT services by small enterprises has become hard to gauge [37].

VI. CONTRIBUTION AND NEW INTUITION

Day in day out, cloud computing is starting to become a household name. Both adults and children are endlessly generating data, especially from social media. How to manage this tremendous amount of data while making sense out of it brings to light the phenomenon of cloud computing. Cloud computing is seen to have penetrated deep in usage within our daily lives and this phenomenon spreads across the globe. Education is certainly not left behind in integration and usage. It is true to say that cloud computing can certainly bring an increased number of benefits to education institutions due to its cost effectiveness as well as students cravings to indulge in technology. Therefore, cloud computing in managing big data becomes an extremely hot topic as an area of study [38]. "Today, MapReduce and virtual cluster are sharp swords for this big data and cloud computing era. To combine these two emerging technologies, it brings feasible-scalability, easy-management, fast-deployment and high-efficiency with the system [39]." Learning in the cloud environment is by far beneficial in comparison to ordinary brick and mortar learning environments as it offers tremendous amount of learning materials as well as endless interactions between students. Therefore, by use of a value-sensitive approach, cloud computing in managing big data allows users to explore through use of data analysis and subsequently offer feedback on its importance. Through this approach, users are therefore able make decisions promptly [40].

VII. CONCLUSION

It is without a doubt that cloud computing is gaining momentum in our daily lives. Given the affordability of internet connectivity, majority of the people across the world are constantly engaging with social media. Both the government as well as enterprises collect data from individuals on each and every single transaction. This data can be stored as structured, semi-structured or raw data. This is where cloud computing comes in to play. An infrastructure of data storage therefore becomes of essence. The goal behind data storage should be scalability, quick deployment as well as reasonable pricing. Cloud computing has yet to see a breakthrough in the world of technology as businesses slowly integrate cloud usage into their business models. Cloud computing does come with its ups and downs and it is important to keep in mind that costs incurred during design and implementation could be expensive while security remains a number one concern. Data integrity and security should be a key focus especially in cloud computing pertaining to healthcare industry. Companies like IBM's project – Watson, currently focusing on patients' data should be vigilant on data security and integrity. Individuals want to feel save whenever they pay a visit to their doctors. There has to be a level of assurance that there will be no tampering or wrongfully disseminating of personal data stored in their databases. Cloud computing providers should not only emphasize on efficiency of data storage but should keep pricing fair in order to encourage usage of their services. One very important aspect of cloud computing and managing big data should be in human intelligence. Schools should be encouraged to offer programs related to big data and especially security of the data in cloud computing. Governments should also play a fair share in encouraging these programs as well as helping in funding them.

R<small>EFERENCES</small>

[1] H. M. Sabi, F.-M. E. Uzoka, K. Langmia, and F. N. Njeh, "Conceptualizing a model for adoption of cloud computing in education," *International Journal of Information Management,* vol. 36, pp. 183-191, 2016.

[2] S. Jones, "Cloud computing procurement and implementation: Lessons learnt from a United Kingdom case study," *International Journal of Information Management,* vol. 35, pp. 712-716, 2015.

[3] M. Vael, "Top tips to manage Big Data and reap the benefits," *Managing Information,* vol. 20, pp. 4-5, 2013.

[4] M. Bayramusta and V. A. Nasir, "A fad or future of IT?: A comprehensive literature review on the cloud computing research," *International Journal of Information Management,* vol. 36, pp. 635-644, 2016.

[5] V. J. Sosa-Sosa and E. M. Hernandez-Ramirez, "A File Storage Service on a Cloud Computing Environment for Digital Libraries," *Information Technology & Libraries,* vol. 31, pp. 34-45, 2012.

[6] H. Yueshun, Z. Wei, J. Nan, L. Lihua, and L. Xiaoyan, "The Research of Scale-free Sensor Network Topology Evolution Based on Cloud Computing," *Journal of Digital Information Management,* vol. 13, pp. 266-275, 2015.

[7] V. Chang, R. J. Walters, and G. B. Wills, "Organisational sustainability modelling—An emerging service and analytics model for evaluating Cloud Computing adoption with two case studies," *International Journal of Information Management,* vol. 36, pp. 167-179, 2016.

[8] A. Wolke, B. Tsend-Ayush, C. Pfeiffer, and M. Bichler, "More than bin packing: Dynamic resource allocation strategies in cloud data centers," *Information Systems,* vol. 52, pp. 83-95, 2015.

[9] Y. Ying, W. Cui-rong, and W. Cong, "Fair Virtual Bandwidth Allocation Model in Virtual Data Centers," *Journal of Digital Information Management,* vol. 11, pp. 179-183, 2013.

[10] F. Akhbar, V. Chang, Y. Yao, and V. Méndez Muñoz, "Outlook on moving of computing services towards the data sources," *International Journal of Information Management,* vol. 36, pp. 645-652, 2016.

[11] P. Laughton, "Data Centres Operating in the Knowledge Economy," *Journal of Information & Knowledge Management,* vol. 11, pp. 1250028-1-1250028-6, 2012.

[12] E. S. Dunn Jr, "The Idea of a National Data Center and the Issue of Personal Privacy," *American Statistician,* vol. 21, p. 21, 1967.

[13] J. Agee and S. Naper, "Off-site storage: an analysis," *Collection Building,* vol. 26, pp. 20-25, 2007.

[14] Y. K. Sinjilawi, M. Q. Al-Nabhan, and E. A. Abu-Shanab, "Addressing Security and Privacy Issues in Cloud Computing," *Journal of Emerging Technologies in Web Intelligence,* vol. 6, pp. 192-199, 2014.

[15] Y. Choi, S. Lee, J. Kim, Y. Kim, H. Pak, G. Moon, J. Ra, and Y.-G. Jung, "The method to secure scalability and high density in cloud data-center," *Information Systems,* vol. 48, pp. 274-278, 2015.

[16] K. J. Knapp, G. D. Denney, and M. E. Barner, "Key issues in data center security: An investigation of government audit reports," *Government Information Quarterly,* vol. 28, pp. 533-541, 2011.

[17] N. M. Yaghoubi, H. R. Jafari, and J. Shukuhy, "Identification and Ranking of the Risk Factors of Cloud Computing in State-Owned Organizations," *Journal of Information Processing & Management,* vol. 30, pp. 1-27, Spring2015 2015.

[18] M. Ramachandran, "Software security requirements management as an emerging cloud computing service," *International Journal of Information Management,* vol. 36, pp. 580-590, 2016.

[19] K. Irion, "Your digital home is no longer your castle: how cloud computing transforms the (legal) relationship between individuals and their personal records," *International Journal of Law & Information Technology,* vol. 23, pp. 348-371, Winter2015 2015.

[20] C. McLellan, "From solid stone temples into the heavens: the impact of cloud computing on the task of recordkeeping," *Archivi & Computer,* vol. 20, pp. 49-65, 2010.

[21] V. Chang, M. Ramachandran, Y. Yao, Y.-H. Kuo, and C.-S. Li, "A resiliency framework for an enterprise cloud," *International Journal of Information Management,* vol. 36, pp. 155-166, 2016.

[22] S. Kumar, G. D. Abowd, W. T. Abraham, M. al'Absi, J. G. Beck, C. Duen Horng, T. Condie, D. E. Conroy, E. Ertin, D. Estrin, D. Ganesan, L. Cho, B. Marlin, C. B. Marsh, S. A. Murphy, I. Nahum-Shani, K. Patrick, J. M. Rehg, M. Sharmin, and V. Shetty, "Center of excellence for mobile sensor data-to-knowledge (MD2K)," *Journal of the American Medical Informatics Association,* vol. 22, pp. 1137-1142, 2015.

[23] M. Yuvaraj, "Perception of cloud computing in developing countries," *Library Review,* vol. 65, pp. 33-51, 2016.

[24] N. M. Yaghoubi, J. Shukuhy, and H. R. Jafari, "Identifying and Ranking Key Factors Influencing the Adoption of Cloud Computing in Electronic Health," *Journal of Information Processing & Management,* vol. 30, pp. 1-25, 2015.

[25] S. Schneider and A. Sunyaev, "Determinant factors of cloud-sourcing decisions: reflecting on the IT outsourcing literature in the era of cloud computing," *Journal of Information Technology (Palgrave Macmillan),* vol. 31, pp. 1-31, 2016.

[26] K. PushpaLatha, R. S. Shaji, and J. P. Jayan, "A Cost Effective Load Balancing Scheme for Better Resource Utilization in Cloud Computing," *Journal of Emerging Technologies in Web Intelligence,* vol. 6, pp. 18-28, 2014.

[27] K. Ketler and J. Walstrom, "The outsourcing of information systems," ed, 1992, pp. 255-255.

[28] Z. Zheng-ping, P. Ren-fang, M. A. Jie, C. Zhe, and Z. Pang-yun, "Study on the Key Technology for Establishing a Cloud Platform-Oriented Digital Oilfield Based on High-performance Computing," *Journal of Digital Information Management,* vol. 12, pp. 429-439, 2014.

[29] N. Susila, S. Chandramathi, and R. Kishore, "A Fuzzy-based Firefly Algorithm for Dynamic Load Balancing in Cloud Computing Environment," *Journal of Emerging Technologies in Web Intelligence,* vol. 6, pp. 435-440, 2014.

[30] G. Garrison, R. L. Wakefield, and S. Kim, "The effects of IT capabilities and delivery model on cloud computing success and firm performance for cloud supported processes and operations," *International Journal of Information Management,* vol. 35, pp. 377-393, 2015.

[31] L. Chu-Fen, "Cloud Computing System Management Under Flat Rate Pricing," *Journal of Network & Systems Management,* vol. 19, pp. 305-318, 2011.

[32] H. Wang, W. He, and F.-K. Wang, "Enterprise cloud service architectures," *Information Technology & Management,* vol. 13, pp. 445-454, 2012.

[33] N. A. Sultan, "Reaching for the "cloud": How SMEs can manage," *International Journal of Information Management,* vol. 31, pp. 272-278, 2011.

[34] N. Regola, N. V. Chawla, and G. Eysenbach, "Storing and Using Health Data in a Virtual Private Cloud," *Journal of Medical Internet Research,* vol. 15, pp. 1-1, 2013.

[35] J. Luftman, H. S. Zadeh, B. Derksen, M. Santana, E. H. Rigoni, and Z. Huang, "Key information technology and management issues 2011-2012: an international study," *Journal of Information Technology (Palgrave Macmillan),* vol. 27, pp. 198-212, 2012.

[36] N. Jukić, A. Sharma, S. Nestorov, and B. Jukić, "Augmenting Data Warehouses with Big Data," *Information Systems Management,* vol. 32, pp. 200-209, 2015.

[37] F. M. Cesaroni and D. Consoli, "Are Small Businesses Really Able to Take Advantage of Social Media?," *Electronic Journal of Knowledge Management,* vol. 13, pp. 257-268, 2015.

[38] C. Changchit, "Cloud Computing: Should it be Integrated into the Curriculum?," *International Journal of Information & Communication Technology Education,* vol. 11, pp. 105-117, 2015.

[39] Y. Yang, L. Xiang, and S. Biaobiao, "Spanning Tree Method for Minimum Communication Costs In Grouped Virtual MapReduce Cluster," *Journal of Digital Information Management,* vol. 11, pp. 213-219, 2013.

[40] D.-H. Shin, "User value design for cloud courseware system," *Behaviour & Information Technology,* vol. 34, pp. 506-519, 2015.