A Survey on Digital Universe and Industry Migration to Optimum Cloud Storage

P. T. Mary Theresa Rani* and S. Prasanna

VELS University, Pallavaram, Chennai - 43, Tamil Nadu, India; pt_mary@yahoo.com
prasanna@velsuniv.org

Abstract

Objectives: Due to phenomenal growth of digital data and social media, industries need tool to migrate from traditional to agile approach with less investment to meet the requirements of customer. **Statistical Analysis:** In this paper, we have analyzed multiple research works and identified cloud storage is the most appropriate approach for digital universe. By understanding various techniques, we confirm that the common fear factors like data loss, personalization, security, reliability and availability have been completely addressed in cloud computing. Hence cloud storage can be suggested to industry for digital transformation **Findings:** There are 3 models in cloud computing and suitability of model based on need of industry requirement is not identified so far. We found open data source in hybrid cloud computing is the best arrangement for industry to fulfill the demand of various customers. Among the existing open databases Mongo DB is identified as faster and consistent. **Application/Improvements:** The proposed model suggests Mongo DB storage in Hybrid cloud computing to provide rich customer experience with less investment for industries. This need to be further analyzed with specific case studies in different types of organization using application tools.

Keywords: Digital Universe, Hybrid Cloud, Industry and Cloud, Mongo DB and Hybrid Cloud, Optimum Cloud Storage

1. Introduction

EMC Digital Universe analysis by IDC states that the universe is becoming digital and is growing by 40% every year. It declares that 44 trillion gigabytes will be the usage in 2020. The dependency on data is huge. It also stated that usage of data in cloud will be double up in the coming years.

Organizations are at the high risk of storing, managing and analyzing this huge data. Convergence technologies are playing major roles. With convergence technology support wired and unwired data built up Internet of Things (IoT) will be the main source for data growth. Figure 1 describes the digital universe.

Technologies in cloud, big data, analytics, mobility, social media and the Internet of Things (IoT) have become the order of the day. Companies are keeping track of these trends and gearing up to change their products and delivery to accommodate these areas where clients are looking for more innovations to enhance delivery. Industries that are traditionally customer-facing are succeeding as digital, simply as an extension of their business.

Retail companies have had to improve their abilities to drive customer reliability and operational costs. Companies are gearing up to change their model and services to accommodate digital areas where clients are looking for more revolutions to enhance delivery.

In this paper we will study the optimized accessing of data using cloud data management framework as well as open source languages. We will review the need of data storage in hybrid cloud model and supporting tools which will make data intensive computing more accessible to all users. The Table 1 defines the need and challenges to focus.

This literature review analyzes the following factors

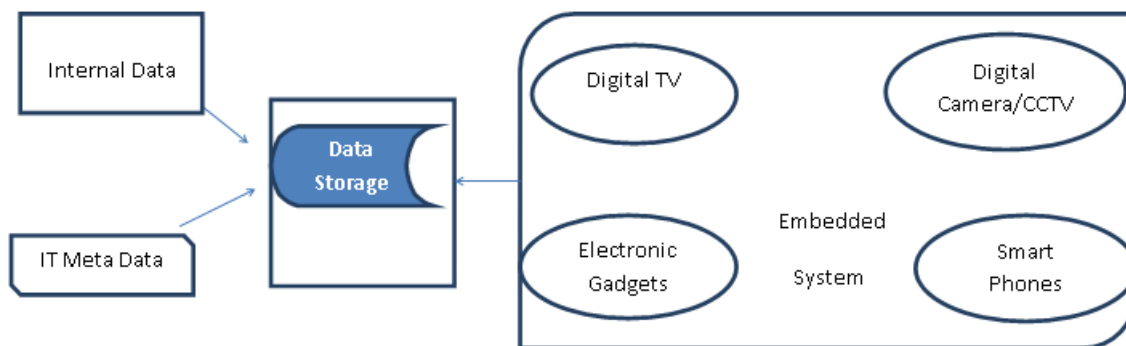- Understanding the growth of data and Cloud computing

Figure 1. Digital universe data storage.

Table 1. Digital Transformation

Challenges	Need
Massive Data collection from various sources. Data duplication across devices	Technology should be based on web and control data duplication
Analyzing data based on business requirements	Tools which support huge data analysis that works on internet
Industry Requirements	
Many small scale and large scale industries will grow in market and resource utilization will be high	Resources to be easily available based on demand
Cost effective	Models that supports various business units
Customer personalization and authentication to be maintained	Severe security mechanism
Enhancement	
Unstructured data models that supports all electronic gadgets in ecommerce	New type of databases with different mechanism to access unstructured data
Data to be scalable and readily available	Change in Locking mechanism of standard databases
De duplication of data	Mechanism which avoids duplication

- The different types of services in Cloud computing and data store
- The challenges in industries and how cloud computing models supports to overcome
- Efficiency of hybrid cloud
- Opportunities for Hybrid cloud with NoSQL model for any industry

2. Data Storage and Cloud Computing

In[1] mentioned that cloud storage is the common available data store in the internet which can be accessed by individual or industry. Cloud computing is the technology that supports storage on the internet. Cloud Computing is the delivery of services for computing a service rather than a product. This includes shared resources, software and infrastructure[2] Cloud computing technology is the most effective way of storing and accessing data. Cloud storage is the cost effective and secured data storage[3].

Small as well as large scale industries are moved into cloud due to its various services. Cloud services are classified as Software as a Service (SaaS), Platform as a Service (PaaS) and Infrastructure as a service (IaaS)[3,4]. SaaS supports on applications, PaaS focus on framework and IaaS is focusing on hardware.

2.1 Cloud Storage

SAAS services of cloud computing provides the application software over the internet[5]. SaaS utilization will grow to $30 billion in 2016. SaaS application can be loaded into multiple servers when required and can be administrated[6]. SaaS cloud storage is the effective way of data storage and maintenance.

In SaaS consumer investment for installation of software and maintenance is absolutely zero. So the end user need not spend on the cost pertaining to maintenance when the data is massive. The consumer or the application users can either consume data through SaaS services or built data store on PaaS. The types of cloud and services are explained in Table 2.

2.2 Business Requirements and Cloud Models

There are various cloud storage models and they are highly accountable[7]. We discussed about the models provided in cloud computing Table 2. The choice of cloud model is only depending on the organization and the investment plan. The deployment model in cloud services are public, private, hybrid and community[7].

There are many advantages of cloud storage. However there are fewer guidelines for large and small scale business units to decide the cloud models based on their requirements. The main two factors which the business units are forced to pay attention is security and cost effectiveness.

3. Industry Challenges

There are two main challenges that need to be analyzed while migrating to cloud.

- Data Security
- Investment
- Agile Model
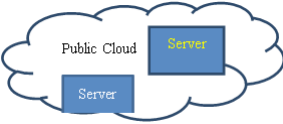
3.1 Security Threatens and Resolutions

There are many approaches to reduce security issues that occurred when data is stored in multiple cloud storage. Group auditing and batch auditing can be applied to verify the security of data[8,4]. Different types of security attacks and to prevent the same has been successfully analyzed and solutions discussed[3,4,8,9]. For database security suggested to avoid dynamic SQL code generation to avoid unauthorized users accessing database[4].

In[10] mentioned that currently the uses of public cloud are enormous. Loss of Information as well as lack of personalization is more. Encryption with hashing data compressing mechanism can be used for protection. So though there are security threats, the model we choose and the data store mechanism can easily address these challenges[9,11,12].

3.2 Cost Effectiveness

Infrastructure and maintenance cost are the major investment for any organization[13]. In Cloud computing these costs are shared by multiple users when the organization moves to public or hybrid cloud. So based on case to case, business units can share their data in private and public. In hybrid model the business unit

Table 2. Cloud models and industry preferences

Private Cloud [Server]	**Owned by Single user or one organization** **Investment cost is high** **Data is highly secured** **Any Data File system can be used**
Public Cloud [Server] [Server]	Owned by Any authorized users Resources provided by external providers Cost effective Data will be shared and accessed across public servers Any available data file system can be used Pay on demand
Hybrid Cloud [Private Cloud] [Public Cloud]	Blended technology Consumer / organization can use private or public based on need Commonly accessed resources can be shared from public cloud Secured and confidential data can be hosted in private cloud

can store confidential data in private cloud data and for expansion and flexibility of data storage; they can use external providers of public cloud[13-16.]

In[17] mentioned hybrid cloud as the best architecture that helps organization to access data in low cost. Hybrid cloud is preferred by organization, as the change in system will not affect the entire application. The next main concern, which is data reduplication, can be avoided using confidential encryption in hybrid cloud, thus helps in minimizing the storage cost. Case studies have been done on e-learning portal using hybrid cloud and promoted hybrid model as the best ROI (Return On Investment) model[15,18].

4. New Era in D

4.1 ATA Storage

We have discussed the growth in data and how cloud computing provides cost effective and secured solution. Among the different model, hybrid cloud will be more reliable and supports extensibility as it has features of both private and public clouds[15,19,20]. So far the common form of data storage is database. However any standard database provides limited support to convergent technologies due to its architecture.

The post by Danyl Bosom worth and Dave Chaffey, described that mobile data usage is at 51% compared to desktop (42%) and other connected devices (7%). Mobile Apps and social networking, video and mailing services accessing in smart phones becomes the default need[21].

4.2 NoSQL

The new change for the industry is data accumulation and data accessing bycustomers using connected devices. This business model needsto be reliable and scalable as industries can leverage global market.

Due to the ACID (Automacity , Consistency , Isolation and Durability) properly of databases like SQL, the availability of data stored in public cloud model across companies will be time consuming[22]. As an alternate NoSQL databases are introduced. NoSQL databases are schema less and comes with a key value pair accessing mechanism.

4.3 MongoDB

There are hundreds of NoSQL open source database is available in the market. Among them Mongo DB is faster, provides high output and consistent. Mongo DB uses document model and supports querying and sorting[23,24].

Mongo DB is supported by multiple programming interfaces. Using Mongo DB data can be divided into small units and shared across machines[25]. This has been used by Hadoop based analysis system to store the log in Monog DB which will be processed by parallel systems.

Many reviews analyzed the performance overview of multiple NoSQL databases and stated the capabilities of this schema less databases are not utilized properly and need to be focused more[26].

The below code represents simple document type model in Mongo DB

```
{ _id: Object Id (7df78ad8904c)
title: "MongoDB in Digital Universe",
description: "mobile data",
by: "Mary Peter",
url: "http://www.scholar.com",
tags: ["nokia", "samsung", "sony"],
likes: 105,
comments: [
    {
user:"Dan",
message: "Wanna Share",
dateCreated: new Date(2016,6,20,2,15),
like: 0 },] }
```

5. Opportunities

In[11] stated that the data storage needs to be dynamic and optimum. We analyzed that hybrid cloud computing has got enhanced security and leverage the opportunities of private and public clouds[27,28].

 For database accessing Mongo DB is suggested for the organizations as they can satisfy the desktop as well as mobile customers. When NoSQL databases deployed in the premises as well in the hybrid cloud, it's cost effective and highly reliable. Mongo DB in hybrid cloud is the area to be focused more that will help all scale industries to satisfy the need of digital data accessing.

Figure 2 describes the new model to be followed in the industry to implement the open data store Mongo DB in hybrid cloud computing model.
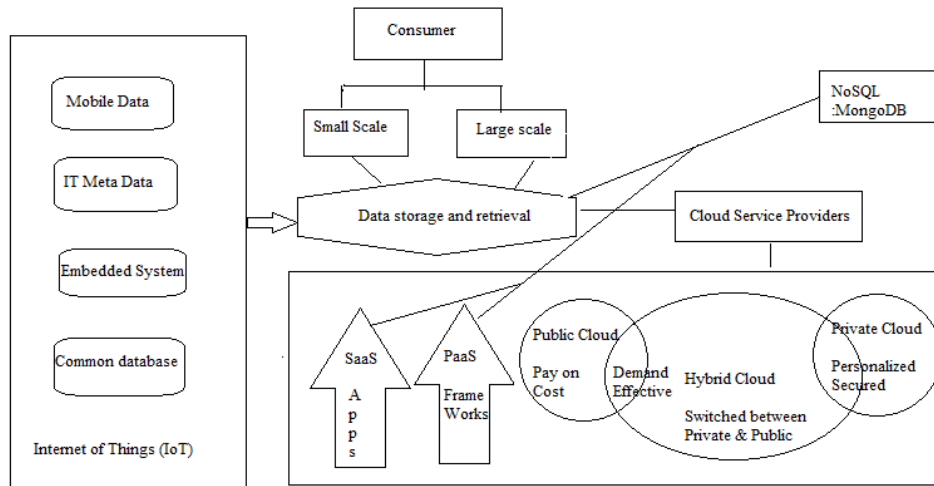
Figure 2. Data storage in cloud models.

6. Conclusion

In this paper we have discussed the growth of digital data in convergence technologies. We have suggested cloud computing models are the most suitable way for handling this growth. We have identified security and cost effective feature of cloud computing. Opportunities in e-business provide the industries a way to expand and reach their global customers[29]. Managing these data and performing scalable and optimized results to the customer only will retain the business. Since there are many cloud models depending on the need of organization we suggested hybrid cloud model can be suggested, as it supports both private and public. We have analyzed the mobile data which plays vital role for global customers and NoSQL structures which is the open source database which is scalable and faster.

We arrived the opportunity of Mango DB deployment in hybrid cloud computing. Both have been proved to be scalable and effective. The future study will be to complete case study framework that deploys Mango DB in hybrid cloud. The case study need to be performed with various types of industries based on their investment

7. References

1. Nitin P, Doiphode HP, Channe C. A survey on accountability of data usage in cloud computing. IJSETR. 2015; 4(3):2278–7798.
2. Nagaprasad S, Vinaya Babu A, Madhukar K, Marlene D, Verghese G, Mallaiah V, Sreelatha A. Reviewing some platforms in cloud computing. International Journal of Engineering and Technology. 2010; 2(5):348-53.
3. Brindha T, Shaji RS, Rajesh GP. A survey on the architectures of data security in cloud storage infrastructure. IJET. 2013; 5(2):1-7.
4. Charanya R, Aramudhan M, Mohan K, Nithya S. Levels of security issues in cloud computing. IJET. 2013; 5(2):975-1024.
5. Omotunde O, Ayokunle A, Okolie SO, IdowuS A, Ajayi OB. Resource allocation in cloud computing – An Expose. International Journal of Advanced Research in Computer Science and Software Engineering. 2015; 5(9):2277 128X.
6. Shingade R, Patil A, Suryawanshi S, Venkatesan M. Efficient resource management in cloud computing. IJET. 2015; 7(6):975-1024.
7. Kirubakaramoorthi R, Arivazhagan D, Helen D. Analysis of cloud computing technology. Indian Journal of Science and Technology. 2015 Sep; 8(21):1-3.
8. Suveetha K, Manju T. Ensuring confidentiality of cloud data using homomorphic encryption. Indian Journal of Science and Technology. 2016 Feb; 9(8):1-17.
9. Beulah S, Dhanaseelanc FR. Survey on security issues and existing solutions in cloud storage. Indian Journal of Science and Technology. 2016 Mar; 9(13):1-8.
10. Mutharasi N, Ajitha PME. Improved secure storage as service in cloud computing. IJET. 2015; 7(5):975-1024.
11. Ahmed ESA, Saeed RA. A survey of big data cloud computing security. IJCSSE. 2014; 3(1):2409-4285.
12. Raghavi J, Anand SK. A survey on cloud storage systems and encryption schemes. IJET. 2013; 5(2):1-5.
13. Dalapati P. A survey on cloud computing. IJCSE. 2013; 5(6):1-14.

14. Jambunathan B, Kalpana Y. Multi cloud deployment with containers. IJET. 2016; 8(1):2319-8613.
15. Sbihi B, Kadiri KEE, Aknin N. Towards a collaborative learning process based on the hybrid cloud computing and Web 2.0 tools. IJET. 2013; 5(3):1-31.
16. Snyder B, Ringenberg J, Green R, Devabhaktuni V, Alam M. Evaluation and design of highly reliable and highly utilized cloud computing systems. Journal of Cloud Computing: Advances, Systems and Applications Springer Open Journal. 2015; 4(11):12-24.
17. Rao VNT, Naveena K, David R, Narayana MS. A new computing environment using hybrid cloud. Journal of Information Sciences and Computing Technologies. 2009; 3(1):1-6.
18. Prakruthi KC. Efficient and secure data reduplication in cloud storage with hybrid cloud approach. International Journal of Computer Science and Information Technology Research. 2015; 3(2):544-8.
19. Kala DDM, Begum SK, Sasikala M. Managing the Storage space in hybrid cloud using reduplication Technique. National Conference on Research Advances in Communication, Computation, Electrical Science and Structures NCRACCESS; 2015. p. 1-5.
20. Saha1 S, Sarkar J, Dwivedi A, Dwivedi N, Anand M, Narasimhamurthy N, Roy R. A novel revenue optimization model to address the operation and maintenance cost of a data center. Journal of Cloud Computing: Advances, Systems and Applications. 2016; 5(1):1-15.
21. Kalpana V, Meena V. Study on data storage correctness methods in mobile cloud computing. Indian Journal of Science and Technology. 2015 Mar; 8(6):1-6.
22. Grolinger K, Wilson A, Higashino H, Tiwari A, Capretz MAM. Data management in cloud environments: NoSQL and NewSQL data stores. Journal of Cloud Computing: Advances, Systems and Applications. 2013; 2-22.
23. Lourenco JR. Choosing the right NoSQL Database for the Job: A Quality Attribute Evaluation. Springer; 2015. p. 2-18.
24. Ozarkar K, Rajani R. Optimization technique for efficient dynamic query forms with NoSQL. IJSR. 2014; 3(11):1-4.
25. Han SH, Kim M, Lee1 H. Design and Implementation of a MongoDB-based Log Processing System in Cloud Computing Environment. 2012. p. 1-4.
26. Abramova V, Bernardino J, Furtado P. Which NoSQL Database? A Performance Overview. 2014; 1(2):11-5.
27. Trust mechanisms for cloud computing. Available from: http://www.journalofcloudcomputing.com/content/
28. Ahirrao S, Ingle R. Scalable transactions in cloud data stores. Journal of Cloud Computing: Advances, Systems and Applications Springer Open Journal. 2015; 4-21.
29. Turner V, John F, Gantz G, Reinsel D, Minton S. The Digital Universe of Opportunities: Rich Data and the Increasing Value of the Internet of Things. White Paper. EMC Corporation; 2010.