

## CLOUD BASED BANKING DATA ANALYSIS SYSTEM

Akshat Zala<sup>1</sup>, Anand Ashok<sup>2</sup>, Siddharth Namboodiri<sup>3</sup>, Vishnu Nair<sup>4</sup>, Samit Shivadekar<sup>5</sup>  
<sup>1,2,3,4</sup>BE (IT), <sup>5</sup>Asst. Prof. (IT)

Department of Information Technology, SIES Graduate School of Technology, Mumbai, Maharashtra

**Abstract:** *The Cloud Based Bank application presented here will be used by the Admin, Back office or Branch manager. With help of this application he/she can view various data in graphical form for easy assessment. The data is stored in cloud, so that it can be accessed from anywhere, through this application through internet. The data stored in cloud will be big data on which the various methods will be called. Using this we can apply data warehousing on various databases or big data. The manager will be able to see various data in graphical format. This application will take out all IT Management function from Bank team and decrease manager's work load. Only the authenticated user can login to the application and use it. This application designed is for use by the bank members and not the customers.*

**Keywords:** *Cloud computing, Hadoop, Big Data, Banking Analysis.*

### I. INTRODUCTION

Cloud computing today encompasses every vertical in the Market across sectors. Organizations are adopting Innovative cloud apps to support their everyday business Operations. To drive growth and innovation in banking, it is increasingly necessary to dramatically leapfrog the Competition using IT and business model transformation. The dramatic changes taking place in banking require new Ways to maximize profitability and returns. Cloud Technology offers secure deployment options that can help Banks develop new customer experiences, enable effective Collaboration and improve speed to market—all while Increasing IT efficiency. Banks that take advantage of Cloud computing are better positioned to respond to Economic uncertainties, interconnected global financial systems and demanding customers.

### II. OVERVIEW

Banks may have various reasons for migrating to the cloud, but the main reason as under- A pivotal stumbling block for huge investments in new technologies has always been the capital expenditure needed for advance infrastructure. With cloud computing, various financial institutions only have to budget for operational expenses towards the services they use. They can avoid captive. IT Infrastructure capital investment. This will also help them avoid deployment of expert IT Team to manage the infrastructure. It would also help them pay for exclusive licences of Basic software as well as security applications. This is also followed by heavy expenses towards maintenance of the same. All these tasks are now being managed by Cloud service provider. So the management bandwidth is freed from IT Segment and they

can focus on growth of business by working on outcome of data and associated analytics. This makes it effortless and more cost effective to test new applications on the cloud versus prevailing conventional. It also allows them to grow the application suite as they get the economic benefit of basic applications.

### A. MODELS

Cloud service models offer financial organization the option to move from a capital-intensive way to a more malleable business model that minimize operational wage. The key to achievement lies in choosing the right cloud services model to meet business needs. In this section we review various models for cloud computing services, functions and deployment

Cloud Service Models

- Business Process-as-a-Service (BPaaS): The cloud is used for standard business processes such as billing, payroll, or human resources. BPaaS combines all the other service models with process expertise.
- Software-as-a-Service (SaaS): A cloud service provider manage the business software and related data and users access the services and data via their web browser.
- Platform-as-a-Service (PaaS): A cloud service provider offers a complete platform for application, interface, and database development, storage, and testing. This allows businesses to streamline the development, maintenance and support of custom applications, lowering IT costs and minimizing the need for hardware, software, and hosting environments.
- Infrastructure-as-a-Service (IaaS): This cloud model allows businesses to buy those resources as a fully outsourced service rather than purchasing servers.

### B. ARCHITECTURE



Fig (a)

A. *User Interface:* User has provided with the graphical user interface in which he has given provision of running analytics over single-node cluster or multi-node cluster. User can distribute data over multi node cluster and results of analytics over that data stored on cloud are provided to the

user in the form of pie charts and bar charts. So user can check the analyzed results by viewing graphs implemented based on analysis made on the data from cloud. User can only be the Branch Manager. Through the analyzed results they will come to know how to increase the profit for the next financial year which will in turn increase the sales in the future.

**B. Business Logic:** It is the middle tier which executes users request by performing operations on the data stored in data store (Cloud). In this system, user invokes to distribute log file on nodes in a Hadoop cluster and run MapReduce in a distributed way. So the logic for these users request is provided in this tier. So it distributes log files by dividing them in equal sized blocks over multiple nodes in a Hadoop cluster, performs Map and Reduce operations over blocks of log file to generate results depending on the key attributes in the log file. Also it does pig query execution over this results to separate out results according to key attributes to generate pie charts and bar charts for them. So when any change occurs we need to change business logic accordingly.

**C. Data Store:** This layer provides data availability. All the bank data including customer information are stored on Cloud. Log file is the required data to process in this system. This log file is stored in HDFS by distributing it over a Hadoop cluster. HDFS works well for text files so it is a good fit to store log files in the simple text format.

### C. ADVANTAGES OF CLOUD

- **COST SAVINGS:** Business sharpness is determined by the cost an organization incurs. There are a few self-service based, and perceptually cost effective public cloud computing solutions. Low-cost price plans advertised by public cloud vendors have inspired IT departments to gain an insight into costs, resource allocation models and the variety of cloud models, including public, private and hybrid. Billing is a non-core process for banks, and outsourcing it to a less expensive mediator allows them to route their capital into core technology-based functions.
- **SCALABILITY:** If well designed, cloud solutions empower banks to meet customer demands and scale quickly, dynamic provisioning of computing resources, will save business users and IT experts from engineering the systems for peak loads. Banks can tackle the challenges of security and data privacy by devising a hybrid cloud where precise data can reside on a private cloud and computing power can be available on a public cloud. These private and public clouds can be integrated in a virtual private network to forge a single scalable hybrid cloud.
- **TIME TO MARKET:** With cloud computing, time to market can be curtailed from months to weeks or days, depending on the size of a bank. A self-service based, on-demand and real-time monitored cloud helps by:
  - Phasing out procurement delays for computing hardware and software
  - Accelerating computing power for when current applications need to deal with peak loads
  - Eradicating the capital and time investment for procuring hardware for proof of concept work.
- **DATA VIRTUALIZATION:** Data virtualization is the assimilation of data from multiple and diverse sources across the enterprise or external sources for the on-demand consumption by a wide range of applications in a virtualized manner. Many mandates in context with the regulations and performance of banks require a data virtualization strategy. This strategy can be used to provide a single source of reference data, such as security master data. Also, risk and analytics calculations rely on many different types and sources of data, including relational and semi-structured XML. Accordingly, accessing that data from a single virtual source would drive scores of data consolidation within banks.
- **Mobility:** Many of today's corporate world techno savvy workers want to access risk and analytics reports while they are on the move. They see the benefits of accessing the internet on their smart phones and iPad's, instantly even in remote locations. Likewise, they want similar interfaces for banking services-specific applications and since a cloud facilitates users to access systems and infrastructure using a web browser or customized clients regardless of location and time, advancement of such interfaces has started taking shape.
- **With Cloud computing,** unnecessary capital expenditures as well as large upfront costs of the infrastructure can be avoided as banks and financial industries can focus on all of the important businesses and projects. The Cloud computing system does not require banks and financial industries to purchase budget shortening hardware.
- **Improved Manageability:** With the help of Cloud computing, banks and large financial industries can make rapid adjustments to their resources for all the unpredictable and fluctuating business requests as well as quickly have their applications uploaded online in no time due to the vastly improved management of Cloud computing that does not require much maintenance.
- **Resiliency:** The systems present in the Cloud computing are extremely useful for banks and financial industries due to the ability of creating a wide enough enterprise availability that is greatly helpful for the continuity of a business venture which is an important part for banks and financial industries.
- **Scalability:** Large acquisitions and mergers are very frequent in the banks and the financial industries which is why easy integrations as well as good scalability are important. The Cloud computing,

works as a very affordable concept that can scale the IT operations according to the needs of the company.

- Accessibility: With Cloud computing the limitations of client server environments are not applied as accessing data and applications from any other computer is achievable anytime, everywhere.
- Security: The Cloud computing system provides a very high level of data protection, especially for sensitive data that includes customer information. The data is kept in a centralized data storage that can only be accessed through strict authentication methods. The security of data is of top priority for Cloud computing, which is why it is beneficial for banks and the financial industry.

#### D. Issues in Banking on cloud

- Security-The confidentiality and security of commercial and personal data and mission-critical applications is preminent. Banks cannot allow the danger of a security breach. Despite economic strain for business to cut down charges and fervent assurances from cloud computing technology providers, security remains a top barrier to cloud technology acceptance. Ultimately, for cloud computing to gain full acceptance within the banking services sector, cloud services must be harmlessly integrated into existing security platforms and processes.
- Regulatory and compliance: Customers are basically responsible for the security and integrity of their own data, even when it is govern by a service provider. Conventional service providers are subjected to external audits and security certifications. Cloud computing providers who ignores to undergo this evaluation are signalling that customers can only use them for the most superficial activities .Many banking mangers require that financial data for banking consumers stay in their native country. Certain compliance arrangements require that data not to be mixed with other data such as on shared servers or databases. As a result banks must have a fair understanding of where their data is stored in the cloud. Security issues which cloud clients should advert are.
  - Privileged user access: There dwell sensitive data that is processed outside the organization inherent risk of security of data because outsourced services bypass the physical and logical IT controls.
  - Regulatory compliance: Customers are responsible for the security of their data. Traditional service providers are subjected to external audits and security certifications.
  - Data location: When users use the cloud, they have no knowledge about the hosted data. Distributed data storage is a main

reason of cloud providers that can cause lack of control and that is risky for customers.

- Data segregation: As cloud is typically in a shared environment in that data can be shared. So there is the danger for data loss. Is encryption available at all phases, and were these encryption patterns designed and tested by experienced professionals.
- Recovery: It is very essential to recover the data when some problem occurs and creates failure.so the main question arises here is that can cloud provider restore data completely or not ,this issue can cause a stalemate in security.

### III. EXPERIMENTAL RESULTS

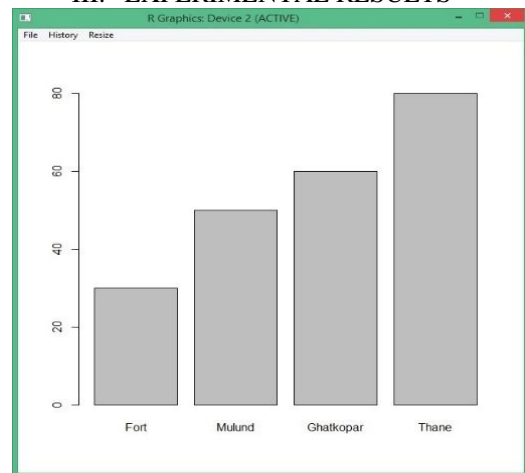


Fig (b) Bar graph displaying assets of certain branches in crores

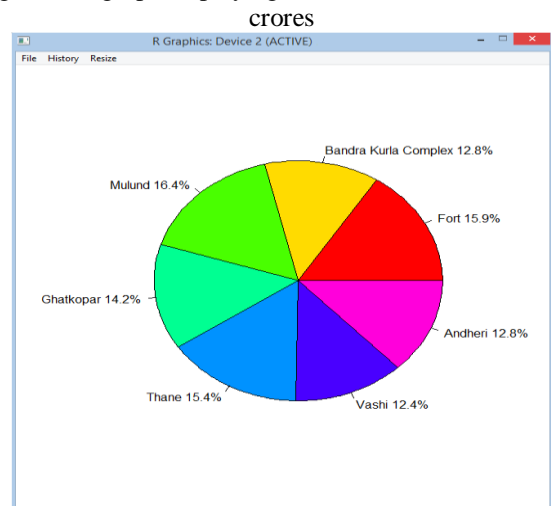


Fig (c) Pie chart displaying distribution of assets in percentage

### REFERENCES

- [1] R. Buyya, J. Broberg, A. Goscinski, "Cloud Computing: Principles and Pradigms", Wiley, 2011.J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68-73.
- [2] John M. Chambers and Trevor J. Hastie, "Statistical

- Models” in S. Chapman & Hall, New York, 1992.
- [3] John A. Rice, “Mathematical Statistics and Data Analysis”, Duxbury Press 1995.
  - [4] Sahil Patani, Sumesh Kadam, Prateek V. Jain. Cloud “Computing in the Banking sector: A survey. International Journal of Advanced Research in Computer and Communication Engineering”, 2014, pp. 1-3.
  - [5] Larry Pace, “Beginning R: An Introduction to Statistical Programming”, 2012.
  - [6] Robert J Knell, “A beginner’s guide to data visualization and statistical analysis and programming in R”, 2014.
  - [7] M. Friendly, “A Brief History of Data Visualization,” Handbook of Computational Statistics: Data Visualization, vol-2, 2008.