

Enterprise Data Architecture of Data as a Service in a Cloud: The Case of Ethiopian Banks

Meskerem Hassen

HiLCoE, Software Engineering Programme, Ethiopia
meskihass@gmail.com

Nassir Dino

HiLCoE, Ethiopia
ndinob@yahoo.com

Abstract

Cloud computing services are on the rise across the business landscape. The promise of low-cost, data availability, solution scalability and operational and infrastructure efficiency have attracted big and small business interests alike. Banks are looking for the means to efficiently store and mine their data for their success, while protecting privacy, integrity, and sharing without a compromise to security. On premise Software as a Service cloud resembling technologies have been successfully used to support the existing banking business processes, but to date no further cloud based services are known. For this reason, this paper aims to assess the existing data architectures employed in the Ethiopian banking industry, evaluate challenges that prevent the Ethiopian banking industry from employing Data as a Service, review existing data architectures for Data as a Service on the cloud in the Ethiopian banking industry, and also propose a data architecture model that shows acquiesces for Data as a Service.

We first examined the current data architecture and management trends of three commercial banks by interviewing selected staff using a purposeful sampling technique. The interview result shows that banks are practicing Data as a Service for delivering their business to users using file transfer protocol, but lack of proper data architecture and data management in place is the major factor preventing them from using Data as a Service in the cloud. The case study also shows that enterprise data is not properly managed and hence data sharing, availability, security, and integrity are among the basic challenges that banks are facing today. This paper proposes an architecture that helps holistic view of data within the enterprise which help manage data within an enterprise.

Keywords: Cloud; Banking Industry; Data Architecture; Data as a Service

1. Introduction

Cloud computing is accepted as one of the factors impacting IT today. Enterprises are ready to migrate their data and applications to a cloud with key concerns including latency, security, and control over critical data and applications. Major enterprises are opting to migrate to the cloud cautiously in progressive manner that aligns with their IT and business governance, risk mitigation and compliance guidelines [1].

The rapid adoption of cloud-based applications, platforms, and infrastructure has resulted in more fragmented data scattered across an enterprise both inside and outside of the firewall, increasing the demand for robust cloud data management practices

[2]. To make the most of a company's return on investments in the cloud, data within these systems should be accurate, complete, and up-to-date. It must be available in time when users need it. Cloud data must be fully synchronized with data in on-premise applications and databases, backup as well as with data in other cloud-based systems [3].

Data becomes the great resource [4]; enterprises are looking for the means to efficiently store and mine their data for their success, while protecting privacy, integrity, and sharing without a compromise to security [5].

The concept of Data as a Service (DaaS) basically advocates the view that, with the emergence of Service-Oriented Architecture (SOA), which includes

standardized processes for accessing data "where it lives", the actual platform on which the data resides doesn't matter. With DaaS, any business process can access data wherever it resides. DaaS began with the notion that data quality could happen in a centralized place, cleansing and enriching data and offering it to different systems, applications or users, irrespective of where they are in the organization or on the network. For an enterprise to benefit from DaaS, it requires proper data architecture in place.

Enterprise Data Architecture (EDA) refers to a collection of master blueprints designed to align information assets (data, device, or other components) that support information-related activities with business strategy. EDA is used to guide integration, quality and data delivery [6]. EDA is part of the overall enterprise architecture, which has several integrated aspects, including hardware, applications, business processes, technology choices, networks, infrastructure and data. EDA plays a vital role in developing and implementing data strategy and also guides developments across systems, such as common reporting, Enterprise Application Integration (EAI), and data warehousing initiatives [7].

The growing desire to seek competitive advantage from the use of data and the challenge of managing an increasingly complex and heterogeneous data landscape have created the right conditions for DaaS to emerge. The debate over what format DaaS should take, who is positioned to provide it, and how it will evolve into a component of the "as-a-service" category is only just beginning. The ongoing consumerization of IT – an overarching trend affecting every aspect of IT use in the enterprise – is shifting user expectations. It represents the enablement of regular, non-expert users to effectively take control of often highly complex and traditionally inaccessible IT tools. DaaS is no exception to this trend and is a critical underlying capability that allows business users to focus on business outcomes rather than attempting to master IT [7].

This paper considers DaaS cloud computing model and data architecture model types with the intent of

data management considerations that are significant for the Ethiopian banking sector opting to move into the cloud computing environment. To do so, the paper proposes an enterprise data architectural model that guides the beneficiary cloud services implementation and acquisition.

2. Related Work

2.1 Enterprise Data Management (EDM)

Enterprise Data Management (EDM) helps Communication Service Providers (CSPs) address the challenges caused by convergence of technologies and frequent mergers and acquisitions. It provides a single view of the truth, unique reference data and a unified data quality framework to integrate, validate and migrate data. In addition, it enables continuous monitoring of the quality of data and establishes standards across the enterprise data lifecycle [8].

EDM helps cloud service providers manage heterogeneous data sources, validate the quality of data, devise a common data model by integrating information, build analytical and presentation layers, and manage end-to-end metadata in the analytical and presentation layers. EDM also guides and governs the data architecture, while managing data assets.

EDM ensures consistency of information with a 'single version of truth' by providing reference data requirements on an integrated data platform. It supports operations and enhances decision making capabilities by helping CSPs migrate from disparate data silos to an integrated, enterprise wide data environment.

2.2 Data Management Maturity (DMM) Model

For organizations that seek to evaluate and improve their data management practices, CMMI has developed the Data Management Maturity (DMM) model to bridge the perspective gap between business and IT.

It provides a common language and framework depicting what progress looks like in all of the fundamental disciplines of data management and offers a graduated path to improvement which is easily

tailored to an organization's business strategies, strengths and priorities [9].

2.3 Reference Architecture for Centralized RDM

Many investment banks have implemented multi-year enterprise data management programs aimed at integration of reference data including legal entity data, instrument data, pricing, and corporate actions data. A scalable, flexible platform based solution capable of providing cost and operational efficiencies that reduces manual intervention, speeds up the onboarding process, simplifies risk analysis and increases automation necessary to meet a firm's business needs [8]. A strategic approach would be to automate reference data management across the value chain by defining a centralized model.

2.4 Banking in the Cloud

With the rise of existing and new non-traditional competition, the banking sector faces a changing business landscape. Satisfying customer demands has become more complex as customers demand more convenience and control over their banking services. At the same time, regulators are ushering in a new era of government oversight. Banks currently face challenges in a number of key areas. Capital inadequacy that depresses profit margins, emboldened customers who expect rapidly evolving new services and offerings, and fierce competition for customers have spawned industry consolidation and the entrance of nontraditional firms. Globally, 90% of financial services clients' surveyed in [10] believe they need to transform from the status quo for future profitability.

a. Microsoft Industry Reference Architecture for Banking

Microsoft is a leading provider of Data and Business Intelligence platforms providing a robust, end-to-end data analytics and collaboration platform. The platform provides Master Data Management (MDM), Data Quality Services (DQS) and pre-defined BI Semantic Metadata (BISM) which overlay BI capabilities delivered via pre-tuned data warehouse configurations, near real-time analytics delivered through High-Performance technical Computing

(HPC) and Complex Event Processing (CEP). The platform is pre-integrated into the scorecard and collaboration platform (SharePoint) that helps data consumers discover, understand, share and collaborate in relation to specific data.

b. Oracle

Oracle Industry Reference Model for Banking is a comprehensive repository of pre-engineered generic processes for banking and capital markets. This pre-analyzed repository is enriched with best practices assimilated from experience in offering specialized IT solutions to more than 840 financial institutions in over 130 countries. The unique value of Oracle Industry Reference Model for Banking lies in its structure, depth, and comprehensive coverage. Its technology-independent framework provides the basis for subsequent translation of abstract process models into executable workflows. Workflows that provide end-to-end automation of banking transactions, on a single operating platform, serve as a foundation for offering insights into future performance [11].

2.5 Banking Applications on the Cloud

The T24 (Temenos 24 hours) on Windows Azure offering is based on Software as a Service (SaaS) model. This model allows financial institutions of all sizes and locations to quickly take full advantage of the rich functionality of T24, without having to manage and invest upfront capital in a complex on-premise deployment. Because T24 is offered as a pre-configured model bank, the solution can not only be rolled out quickly, but it requires very little customization [12].

Temenos is already in the process of taking financial institutions from a traditional hosted environment onto the Windows Azure platform, to help them lower costs, enhance operations, and provide the flexibility to rapidly build and bring to market new products to support mission-critical business opportunities.

3. The Proposed Architectural Model

Enterprise data architecture model, which is business-driven, should describe a structured

interrelationship between the technical and procedural solutions to support the long term needs of a business. If the architecture is to be successful, then it must provide a rational framework within which decisions can be made upon the selection of data management solutions. The decision criteria should be derived from a thorough understanding of the business requirements, including the need for cost reduction, modularity, scalability, ease of component reuse, operability, interoperability both internally and

externally, and integration with the enterprise IT architecture and its legacy systems.

EDA for DaaS cloud architecture model is partly adopted from banking reference architecture models and their existing problem, gap and the remaining work that must be addressed and the way to solve the problem on data management for DaaS cloud for the banking industry. Our proposed architecture model is shown in Figure 1. It is composed of three layers each representing the theoretical foundations.

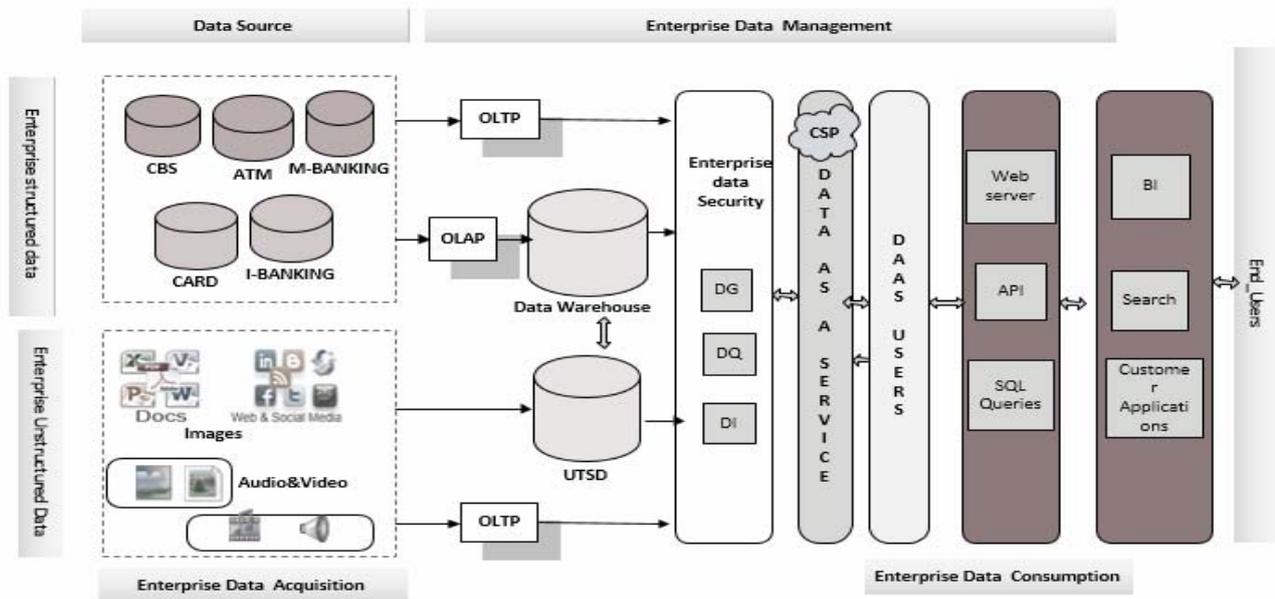


Figure 1: The Proposed Enterprise Data Architecture Model for DaaS Cloud

3.1 Requirements Gatherings

Based on the literature review, we analyzed the bank data management requirements through analysis of how data is accessed, stored, searched, and archived with respect to DaaS cloud services. In line with this, bank data management issues and drivers are identified. Moreover, the main data characteristics and management requirements of the banking industry are analyzed. The details are presented as follows.

a. Data Collection

Literature is reviewed in the area of data architecture and cloud computing in general and cloud computing for enterprise data in particular. The primary data is obtained from interviews.

b. Samples and Population

In order to discover what the perspective on enterprise data architecture for DaaS cloud is from the point of view of financial institutions, an interview has been developed. The main targets of this approach are divided into three viewpoints.

- Business view: Account and Finance Department, Human Resource Department and Credits Relations Department.
- Information and Technology view: Database, Application and System Administrators, CTOs and CIOs.
- CSP/Consultants view: Bank external consultants.

Two individuals from each department and one external consultant per bank were selected based on

experience. The selected banks access their core application from central data center which can be taken as having a look and feel of a private cloud.

c. Interview Results

The interview is conducted to assess existing DaaS and data architecture, trends, challenges and availability of such architectural models.

i. Enterprise Data Sharing and Storage

The interview result shows that most banks share their enterprise unstructured data with others using FTP. The FTP server client will be installed on each user machine. Folders will be created for each department and branch. There will be one or two users to access the FTP folder. The user with credential will upload reports, memos and other MS Word/Excel generated files into his/her folder. A branch or department can only view/access a folder owned by the department not others. Each department or branch will process its daily tasks using MS Word, MS Access, etc. Regarding storage, they use local disk or external disks to store their data. They sometimes use FTP to store their files in addition to exchanging files. The storage structure in general is both central (FTP) and distributed (local disk/client machine).

ii. Data as a Service

Interviewees indicate that they are already practicing DaaS. They provide data to branch and department users through FTP. The users will then access the data and use different tools to convert it into information. Most branches and departments use MS Office applications and others use MS reporting tool to exchange reports within the bank or third parties. But all claim that they don't have enterprise wide data and DaaS architecture to make data available to all applications and systems. They add that they don't have data architecture to support all data to be delivered as service.

iii. Enterprise Data Architecture Model

The interviewees responded that they don't have a data architecture model, but say that they have designed and documented data center architecture. They also mentioned that there is lack of a clear

design as to how all application systems, storage, data and databases relate with each other.

iv. The Challenges while Providing DaaS

Most interviewees say that they have encountered the following challenges: data availability, data security, data indexing, data change control/data audit, data classification, data backup and recovery and data storage space.

v. Enterprise Data Access

In order to access data from an FTP server, the user must have a username and password. A department can only access its own shared folder but not that of others. User credentials will be shared by users amongst the department or give responsibility to one or two staffs based on their position. There is no mechanism to track change made to the data on shared environment. The users who have access to shared folder are accountable or assumed to have made all the change if needed.

All access to banking applications like CBS, CRM, ERP, ATM, etc. is through users, whether they are administrative users, application accounts, or regular users. To access banking database through applications, users will first have to be registered using AD (Active Directory) or LDAP (Lightweight Directory Access Protocol). After being authenticated they will log into their applications using application credentials (username and password).

3.2 Validation

Enterprise Data Sharing System (EDSS) validates part of the proposed architecture model. The system serves as a prototype to demonstrate the feasibility of enterprise unstructured data, specifically the data sharing part, for progression towards DaaS. In its development process, first an appropriate model for demonstration purpose was selected, and major system components were designed and implemented. Finally, the system is hosted on selected bank data center and tested with selected bank users. The EDSS implements only the data sharing and access portion of the proposed architecture model element, i.e., unstructured data. It does not implement data

generated from social media sites, audio and video files and also text processing.

i. Features of the Prototype

The EDSS provides the following basic services: data upload, download, grant and revoke permissions, view, different search options, audit, etc. by storing it into a database.

ii. System Development Architecture

Model-View-Controller Architecture is followed for developing the web application. There are three layers in the overall system architecture. These are: Model Layer (Data Access Layer), View Layer (Presentation Layer), and Controller Layer (Business Layer). The Model represents the application data. The View renders a presentation of the model data and The Controller handles and route requests made by the client [13].

5. Conclusion and Future Work

This paper offered an architectural model for DaaS cloud adoption that covers both the enterprise architecture and business perspective which intern applied to the Case Study Bank. Based on the usage experience, this research shows the need for having common interface portal for DaaS by demonstrating an end to-end data architecture model for DaaS cloud adoption.

This paper has three key contributions. Firstly, various methodological supports in relation to cloud and generic business strategy have been elaborated to harmonize The Open Group Architecture Framework Architecture Development Method. The suggested architectural model has emphasized the importance of having DaaS, a proper data architecture, data management within an enterprise and also identifies the cloud-enabled architecture vision.

Secondly, we have revealed the gaps on enterprise data and data management trends, the difficulties before enterprises move to DaaS not only from technical concerns but also from the inter-organization management aspects. Lastly, the research demonstrated the use of the proposed EDA model for DaaS cloud services and approaches by applying it to

a real case of cloud implementation in the Analyzed Case Study Bank.

Further research and analysis is required to improve the suggested architecture model and make it more comprehensive and viable for DaaS cloud implementation. There are several opportunities to extend the current research.

The paper focuses more on phase C (data architecture) of The Open Group Architecture Framework Architecture Development Method. Hence, exploring the following parts of the method will help enrich the architecture model.

- Data security for DaaS on public cloud,
- Data management policy for DaaS cloud, and
- Implementation of DaaS interface that integrates different banks to share data using a common API.

References

- [1] R. Harms and M. Yamartino, "The Economics of The Cloud," November, 2010.
- [2] P. Kumar and H. S. Arri, "Data Location in Cloud Computing," International Journal for Science and Emerging Technologies with Latest Trends, Jalandhar, Punjab, February 5, 2013.
- [3] Z. Mahmood, "Data Location and Security Issues in Cloud Computing," Int. Conference on Emerging Intelligent Data Web Technology, pp. 49–54, Sep. 2011.
- [4] K. Hwang and D. Li, "Trusted Cloud Computing with Secure Resources and Data Coloring," IEEE Internet Computing, Vol. 14, No. 5, pp. 14–22, Sep. 2010.
- [5] C. M. Lawler, "A Cloud Framework for Integrated Security, Governance, Compliance, Content and Service Management", Gartner Symposium ITXPO, October 25, 2011.
- [6] EWSolutions, "Enterprise Data Architecture" white paper, 1999, retrieved from <http://also.co.uk/docs/speed.pdf>, Last accessed on June 10, 2011.
- [7] Raghunath Rajamony and Jamuna Ravi, "Enterprise Data Architecture Opportunities, Challenges and Expectations", Beye Network UK.

- [8] G. A. Lewis and P. Place, "An Enterprise Data Architecture Guide", October, 2001.
- [9] C. M. Model and C. M. Modeling, "Data Management Maturity Model Introduction University of Ottawa," 2014.
- [10] Srinivasa Kumar Yerchuru, Ganesh Padmanabhan and Ramakoteswara Rao T, "Cloud Solutions for Centralized Reference Data Management", No.1 Dec, 2011.
- [10] "Microsoft Industry Reference Architecture for Banking (MIRA-B)," May, 2012.
- [11] O. E. Architecture, "Reference Library of Architecture Best Practices IT Strategies from Oracle (ITSO)", Oracle enterprise architecture white paper, December 2014.
- [12] Temenos, "Banking in the Cloud Windows Azure", Temenos white paper, 2011.
- [13] C. S. Foundation, "CakePHP Cookbook Documentation", 2016.