Fall 2015

Data Framework Management System

Firasat Ali Mohammed  
Governors State University

Ahmad Munir Rizwi Syed  
Governors State University

Follow this and additional works at: http://opus.govst.edu/capstones

Part of the Databases and Information Systems Commons

Recommended Citation
http://opus.govst.edu/capstones/169

For more information about the academic degree, extended learning, and certificate programs of Governors State University, go to  
http://www.govst.edu/Academics/Degree_Programs_and_Certifications/

Visit the Governors State Computer Science Department
This Project Summary is brought to you for free and open access by the Student Capstone Projects at OPUS Open Portal to University Scholarship. It has been accepted for inclusion in All Capstone Projects by an authorized administrator of OPUS Open Portal to University Scholarship. For more information, please contact opus@govst.edu.
# Table of Contents

1. **Project Description** .......................................................................................................................... 3
   1.1 Project Abstract .................................................................................................................................. 3
   1.2 Competitive Information ..................................................................................................................... 3
   1.3 Relationship to Other Applications/Projects .................................................................................... 3
   1.4 Assumptions and Dependencies ......................................................................................................... 3
   1.5 Future Enhancements ......................................................................................................................... 4
   1.6 Definitions and Acronyms .................................................................................................................. 5

2. **Technical Description** ....................................................................................................................... 6
   2.1 Project/Application Architecture ......................................................................................................... 6
   2.2 Project/Application Information flows ............................................................................................... 6
   2.3 Interactions with other Projects (if Any) ............................................................................................ 6
   2.4 Interactions with other Applications ................................................................................................. 8
   2.5 Capabilities ....................................................................................................................................... 9
   2.6 Risk Assessment and Management.................................................................................................... 10

3. **Project Requirements** ....................................................................................................................... 10
   3.1 Identification of Requirements .......................................................................................................... 10
   3.2 Operations, Administration, Maintenance and Provisioning (OAM&P) ............................................ 11
   3.3 Security and Fraud Prevention .......................................................................................................... 12
   3.4 Release and Transition Plan .............................................................................................................. 12

4. **Project Design Description** ............................................................................................................... 12

5. **Project Internal/external Interface Impacts and Specification** .......................................................... 12

6. **Project Design Units Impacts** ........................................................................................................... 12
   6.1 Functional Area/Design Unit A ......................................................................................................... 14
      6.1.1 Functional Overview .................................................................................................................. 15
      6.1.2 Impacts ..................................................................................................................................... 15
      6.1.3 Requirements ............................................................................................................................ 15
   6.2 Functional Area/Design Unit B ......................................................................................................... 15
      6.2.1 Functional Overview .................................................................................................................. 16
      6.2.2 Impacts ..................................................................................................................................... 16
      6.2.3 Requirements ............................................................................................................................ 16

7. **Open Issues** ...................................................................................................................................... 17

8. **Acknowledgements** .......................................................................................................................... 18

9. **References** ......................................................................................................................................... 18

10. **Appendices** ...................................................................................................................................... 19
1 Project Description

The aim of this project is to design a system for managing multiple data source through networks. The focus is on applying best practices using the ITIL framework to achieve reliability. The IT Infrastructure Library (ITIL) Framework, managed by AXELOS Limited, defines a commonly used, best-practice approach to IT Service Management (ITSM).

1.1 Project Abstract

The goal of this project is to design a system for managing multiple data source through networks. The end user is concerned about the computations that depend on information from several data sources. The access will be from a portal. The system should use grid computing standards to fetch information from the different information sources, consolidates them, presented them as required.

1.2 Competitive Information

Qlik Sense and Snowflake are the major web based cloud systems which are close competitors.

1.3 Relationship to Other Applications/Projects

The application interacts with AWS. Amazon Web Services own and maintain the network-connected hardware required for these application services. Cloud Computing provides a simple way to access servers, storage, databases and a broad set of application services over the Internet.
### 1.4 Assumptions and Dependencies

<table>
<thead>
<tr>
<th>General Architecture principles based on past experiences and agile models best practices &amp; methodologies will be used in designing the solution.</th>
</tr>
</thead>
<tbody>
<tr>
<td>The basic TCP/IP (HTTP) protocol will be the only one used to access the application</td>
</tr>
<tr>
<td>The web browser will be the primary client used by employees and public users</td>
</tr>
<tr>
<td>This system supposed to be a web based application which can be accessed 24/7.</td>
</tr>
<tr>
<td>Network connection should be available to use the application.</td>
</tr>
</tbody>
</table>

### 1.5 Future Enhancements

Future customization will include features that will simplify data storage process. It is very important for the code to be customizable due to the distinct needs of different organization. The application will be integrated with the cloud.
### Definitions and Acronyms

<table>
<thead>
<tr>
<th><strong>Acronym</strong></th>
<th><strong>Definition</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DFMS</strong></td>
<td>Data Framework Management System</td>
</tr>
<tr>
<td><strong>DB</strong></td>
<td>Data Base</td>
</tr>
<tr>
<td><strong>ITIL</strong></td>
<td>The IT Infrastructure Library (ITIL) Framework</td>
</tr>
<tr>
<td><strong>PHP</strong></td>
<td>PHP Hypertext Preprocessor</td>
</tr>
<tr>
<td><strong>AWS</strong></td>
<td>Amazon Web Services</td>
</tr>
<tr>
<td><strong>HTML</strong></td>
<td>Hyper Text Markup Language</td>
</tr>
<tr>
<td><strong>FR</strong></td>
<td>Functional Requirement</td>
</tr>
<tr>
<td><strong>SOW</strong></td>
<td>Statement of Work</td>
</tr>
<tr>
<td><strong>CS</strong></td>
<td>Computer Science</td>
</tr>
<tr>
<td><strong>WAMP</strong></td>
<td>Windows, Apache, MySQL, PHP</td>
</tr>
</tbody>
</table>
2.0 Technical Description

AWS: Cloud Computing provides a simple way to access servers, storage, databases and a broad set of application services over the Internet. Cloud Computing providers such as Amazon Web Services own and maintain the network-connected hardware required for these application services, while you provision and use what you need via a web application.

MySQL Workbench 6.2

MySQL Workbench 6.2 is comprehensive, integrated database system that enables organizations to reliably manage mission-critical information and confidently run today’s increasingly complex business applications. MySQL Workbench 6.2 allows companies to gain greater insight from their business information and achieve faster results for a competitive advantage.

MySQL Workbench functionality:

- **SQL Development**: Enables you to create and manage connections to database servers while help in storing data sources. Along with enabling you to configure connection parameters, MySQL Workbench provides the capability to execute SQL queries on the database connections using the built-in SQL Editor.

- **Data Modeling (Design)**: Enables you to create models of your database schema graphically, reverse and forward engineer between a schema and a live database, and edit all aspects of your database using the comprehensive Table Editor. The Table Editor provides easy-to-use facilities for editing Tables, Columns, Indexes, Triggers, Partitioning, Options, Inserts and Privileges, Routines and Views.

- **Server Administration**: Enables you to administer MySQL server instances by administering users, performing backup and recovery, inspecting audit data, viewing database health, and monitoring the MySQL server performance.

- **Data Migration**: Allows you to migrate from Microsoft SQL Server, Microsoft Access, Sybase ASE, SQLite, SQL Anywhere, PostreSQL, and other RDBMS tables,
objects and data to MySQL. Migration also supports migrating from earlier versions of MySQL to the latest releases.

- **MySQL Enterprise Support**: Support for Enterprise products such as MySQL Enterprise Backup and MySQL Audit.

### 2.1 Project/Application Information flows

![Diagram of Project/Application Information flows]

### 2.2 Interactions with other Projects (if Any)

The data frame work management systems interacts with Amazon Web Services (AWS). AWS is a collection of remote computing services, also called web services, that make up a cloud platform offered by Amazon.com. These services operate from 11 geographical
regions across the world. The most central and well-known of these services arguably include Amazon Elastic Compute Cloud, also known as "EC2", and Amazon Simple Storage Service, also known as "S3". Amazon markets AWS as a service to provide large computing capacity more quickly and more cheaply than a client company building an actual physical server farm.

2.3 Interactions with other Applications

The DFM S application interacts with Amazon Web services. Below diagram explains how the interactions between the end user and Amazon RDS are implemented.
2.4 Capabilities

The DFMS application has the capability to store the data source.

The end user can use the information to merge or analyze the information.

The application has the ability to store the data sources in various formats.

The storing feature of the application is made after careful review of the existing
application available in the market.

3 Project Requirements

3.1 Identification of Requirements:

<GSU-001-0.1 data source 000101>

The application should have the ability to store data sources.

<GSU-001-0.1 data source format /000102>

The application should be compatible with various data source formats.

<GSU-001-0.1 adminstrator/000103>

The webapplication should have the secure login funtionality.

<GSU-001-0.1 registration /000104>

The user (guest) should be able to register if not a registered user.

<GSU-001-0.1 view files /000105>

Previously uploaded information visible to end user.

<GSU-001-0.1 feedback/000106>

The system should be able to capture the user feedback

3.1 Operations, Administration, Maintenance and Provisioning (OAM&P)

i) The ability to build and maintain positive relationships with team-members and stakeholders.
ii) Mandatory implementation of the client or supervisor feedback.

iii) The desire and ability to change one's approach to project management and/or course of action in response to business needs.

vii) The ability to get things done efficiently and effectively

3.2 Security and Fraud Prevention

The application coding has been designed to prevent code injection attacks. When the application is interacting with cloud, then the data exchange will happen over HTTPS.

3.3 Release and Transition Plan

Release Process

This statement of work (SOW) shall commence on August 21st, 2015 and shall continue until 29th Nov, 2015.

Below are the timelines and Deliverable details.

- Interim Project I has been given on Nov 03, 2015
- Final Project II been given on Nov 29, 2015
### 3.4 Transition Schedule

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Member1</strong></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td><strong>Member2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Member1</strong></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td><strong>Member2</strong></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### 4 Project Design Description

The purpose of this system is to support the end user or data scientist in storing and analyzing the data sources in multiple formats. Since DFMS is an online system, it can be easily accessed from web-browser with internet access, thus removing any constraints of time or place. The system also sends relevant notifications and information to respective users through emails when they contact for any help. The system will have a user friendly interface which will make it easier for all kinds of participants.

Using Data Frame work management system, you can get your data stored and sorted quickly. In addition to that as the application has been integrated with AWS, it provides you with a comprehensive ecosystem of technology and consulting partners through our AWS Partner Network and AWS Marketplace. Our partners have created innovative big
data analytics solutions or have implemented successful big data platforms for other customers in the AWS Cloud.

5 Project Internal/external Interface Impacts and Specification

i) Structure:

Structure is an internal Factor that impacts operation of Data Framework Management System.

ii) External Communications:

The impact web application has while interacting with cloud services such as AWS have to be closely monitored to achieve performance.

6 Project Design Units Impacts

The Techniques we will be using in our project will be focused on Agile Model driven development, the iterative nature of agile development means features are delivered incrementally, enabling some benefits to be realized early as the product continues to develop. A key principle of agile development is that testing is integrated throughout the lifecycle, enabling regular inspection of the working product as it develops. This allows the product owner to make adjustments if necessary and gives the product team early sight of any quality issues. Agile development principles encourage active ‘user’ involvement throughout the product’s development and a very cooperative collaborative approach. This provides excellent visibility for key stakeholders, both of the project’s progress and of the product itself, which in turn helps to ensure that expectations are effectively managed.
6.1 Functional Area/Design Unit A

The objective of the project is to develop a data framework management system within allocated time and specified quality. We combine 2 data sources to generate user desired output.
### Datasource 2

<table>
<thead>
<tr>
<th>education</th>
<th>age</th>
<th>job</th>
<th>marital</th>
<th>default</th>
<th>housing</th>
<th>loan</th>
</tr>
</thead>
<tbody>
<tr>
<td>high school</td>
<td>39.0</td>
<td>services</td>
<td>single</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>high school</td>
<td>25.0</td>
<td>services</td>
<td>married</td>
<td>no</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>high school</td>
<td>36.0</td>
<td>admin.</td>
<td>married</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>high school</td>
<td>29.0</td>
<td>admin.</td>
<td>single</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>high school</td>
<td>39.0</td>
<td>technician</td>
<td>divorced</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
</tbody>
</table>

**Output**
6.1.2 Functional Overview

With this system, users can do the following major functions:

I. Store data source.
II. Download stored data source.
III. View data source

6.1.3 Impacts

Impacts of the system are:

I. User should be able to access the system over the network.
II. Participants can be a normal user, researcher or data scientist.
III. Participants must have the valid email for the authentication purpose.

6.1.3 Requirements

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>DFMS Requirement</td>
</tr>
<tr>
<td>2</td>
<td>The user needs to provide the correct user ID and password</td>
</tr>
<tr>
<td>3</td>
<td>The user needs to provide the correct username</td>
</tr>
<tr>
<td>4</td>
<td>The application should have the capability to capture the news in the Information technology world</td>
</tr>
<tr>
<td>5</td>
<td>The application should have the capability to capture the different datasources such as CSV and Txt etc</td>
</tr>
<tr>
<td>6</td>
<td>The application should have the capability to check the different data sources which have been already been uploaded</td>
</tr>
<tr>
<td>7</td>
<td>The application should have the capability to let the end user provide the feedback or contact the management</td>
</tr>
<tr>
<td>8</td>
<td>The application should have the interface to select the datasource type which is being uploaded</td>
</tr>
</tbody>
</table>
6.2 Functional Area/Design Unit B

Impacts

I. The AWS service needs at least two data sources for the functionality to work.

II. The user should be able to gain access to AWS console by submitting valid credentials.

7 Open Issues:

i) Third party – Relating to issues with Amazon Web services or MySQL. The team encountered connectivity issues with MySQL and AWS as they are different project which the application is interacting.

ii) Technical – This segment deals with the technical aspects of the project. It is important for the team to decide the compatible cloud service.

iii) Business process – Relating to the project's design. Here the focus is towards how the design helps in achieving a better business model

iv) Change management – This segment deals with the end user change requests. The team needs to discuss if the requested change is valid and then implement it.
8 Acknowledgements:

As a team we would like to thank our supervisor for providing feedbacks and necessary information to help the team complete the project.

All the team members of group 4 express special gratitude and thanks to the computer science department and industry persons for giving such attention and time.

9 References

- https://d0.awsstatic.com/whitepapers/aws_cloud_adoption_framework.pdf
**Description**: Using the above screen where user has to enter login credentials to get inside the system. If the user has provided as correct user credentials then he will access the DFMS dash board to upload the data sources and view files, RSS feed to browser the latest information technology news.
**Description**: The above screen shot is the dashboard of the DFMS application. It offers the interface using which the user can upload different files and then can also view the Feed news. In addition to that the user can view the files which have been already uploaded them. The user can also provide the feedback using the contact us tab available.
**Description:** The DFMS application uses WampServer for hosting purpose. WampServer is a Windows web development environment.

Description: This interface lets the user to select the data source type such as CSV, Txt etc which the end user selects before uploading the file into the database which in our case local host.

**View Files**

- PDF
- CSV
- Excel

EbookDB.pdf
**Page Description:** Using the above webpage the user can view the previously uploaded data sources.

![Data Management System](image)

**RSS Feed**

- [http://feedproxy.google.com/~r/technradarallnews/~3/fIK4cIws02E/story01.htm](http://feedproxy.google.com/~r/technradarallnews/~3/fIK4cIws02E/story01.htm)

**Description:** Using the above displayed webpage the end user will be able directly visit other website WebPages.