

## **The Design of Data Hub System for Integration of Group In the Cloud Environment**

Hyung-Seok Kim\*, Jong-Yong Lee\*\*\*, Kye-Dong Jung\*\*\*

*\*Department of Information System Kwangwoon University Graduate School of Information Contents, 20  
Kwangwoon-ro, Nowon-gu, Seoul 139-701, Korea*

*\*\*\*Department of Electronic Engineering, Kwangwoon University, 20 Kwangwoon-ro, Nowon-gu, Seoul  
139-701, korea*

E-mail : tyuio7779@kw.ac.kr, jyonglee@kw.ac.kr, gdchung@kw.ac.kr

### **Abstract**

*For a recent most companies to make efficient business management, we are using a groupware service of integrated information system of the entire group. Groupware service integrates the cooperation in excellent synergy and duplicates has been business functions of the business, through the improvement of multi-purpose business processing capacity, there is an advantage of reduced operating costs. However, if the parent company and subsidiary, or to handle common tasks such information agency, which may cause differences in the format of the data passed in the case of a need to provide a document. Therefore, in this paper, in order to solve the heterogeneity problem in data between groups, the data system of the hub base of the cloud is provided. The proposed system is intended to improve the groupware environment including the interoperability of integrated standardized environmental data sharing service.*

**Keywords:** DBaaS, Integrating Groupware, Cloud System

## **1. Introduction**

Business integration in each group of affiliates, aim a system according to an integrated environment for the holding and management data that public of consistency. Therefore Groupware services, the population as a company, it has been used for the construction of integrated environment, its market size, tend to have increased daily in accordance with the development of the IT technology. Current groupware service shows a steadily improved performance. Many developers are doing more active research and development of this[1]. Moreover, cloud computing technology to a new paradigm has been exploited for the development of a stable integrated groupware systems in various parts. Representative examples, cloud storage (Cloud Storage) service is used in the role that primarily stores the customer information such as a portal site[2][3]. However, recently, there is a tendency to be able to use even for the purpose of sharing the service, and is variously applied in such populations comprising companies that are connected by the Internet[2][4]. However, groupware, a company connected to the Internet, and comprises each other use the same DBMS(DataBase Management System), and to share the data with each other, the possibility that the heterogeneity between

the stored data type occurs there is a problem[5]. To solve these problems, in this paper, there is provided a data hub system for a cloud environment based group integration. That, for data exchange between different companies, it is an object to construct a business environment interoperability and data integration service. In this paper, through a system of DBaaS(DataBase as a Service) based Data Hub, to solve the compatibility and heterogeneity issues between the data that is stored.

## **2. Related works**

### **2.1 Integrating Groupware**

Groupware solution for increasing the efficiency of the joint is a group connected to the same computer as the enterprise, are used for the purposes for institutional collaboration. This system, modules that provide the basic functionality, information sharing function module, decision support module, information storage module, flow management function module is divided into database connection module[6][7]. Within the companies that are using groupware, and efficiency of cooperation operations and data inconsistencies or duplicate that there is a problem to be between each sector using a business environment that is integrated utilizing the cloud, also reduce costs and increase efficiency and productivity of the parts of the business, it is possible to influence with much help[8][12].

### **2.2 DBaaS(DataBase as a Service)**

The purpose of cloud computing leverages Internet technology, is to provide a more scalable IT-related technologies and capabilities. This is based on the "as a service" IaaS (Infrastructure as a Service), PaaS (Platform as a Service), such as SaaS (Software as a Service), as each of the form, is expressed in a variety of services. Among them, DBaaS mainly when it is intended to develop a system that requires the essential collaborative environments, such as corporate, acts that will be transmitted on On Demand system IT resources and databases through the Internet[10][11]. Therefore DBaaS is characterized that they provide an environment that can be integrated efficiently database. And is possible efficient data sharing environment integrated into data compatibility[12]. Therefore, in the groupware environment that is integrated on the basis of DBaaS, the overall service of the database, it will provide a solution between enterprises and companies[8].

### **2.2 Query Management**

Query Management is the ability to generate and transmit the SQL statements global query that does not include the FROM clause that points to the table. Why does not contain a field for global query represented, it is because it is possible to transmit by using the stored in the Local Schema and Meta Schema <standard field name requirements[13]. In response to a request of the user who requests, <condition n> is to be able to extract the search information that is specified as the respective conditions. Also, it is possible to request the data of the DB corresponding to the Enterprise, it is possible to convert a local schema global query is transmitted[13][14].

## **3. DBaaS-based Data Hub System**

System environment of cloud-based Data Hub proposed in this paper is as shown in FIG. 1. The proposed system, specific groups that are connected to the Internet network using the groupware service, if necessary specific groupware DB that stores, sharable on demand via the Data Hub It will provide an environment. Most in DBaaS solution that provides advanced data integration environment is capable of effective data

management, to solve the interoperability issues of different types of data for each DB. In addition, in an integrated environment for sharing and collaboration of data, it is efficient, shows the excellent and in terms of business processing, superior performance than any other such systems.

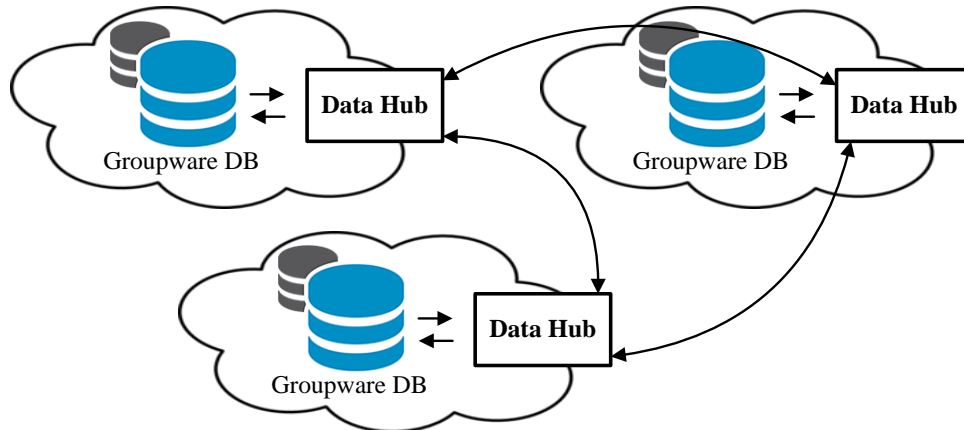


Figure 1. Data Hub System

### 3.1 DBaaS-based Data Hub Architecture

The overall structure of the proposed cloud-based Data Hub system, it was designed as shown in Figure 2.

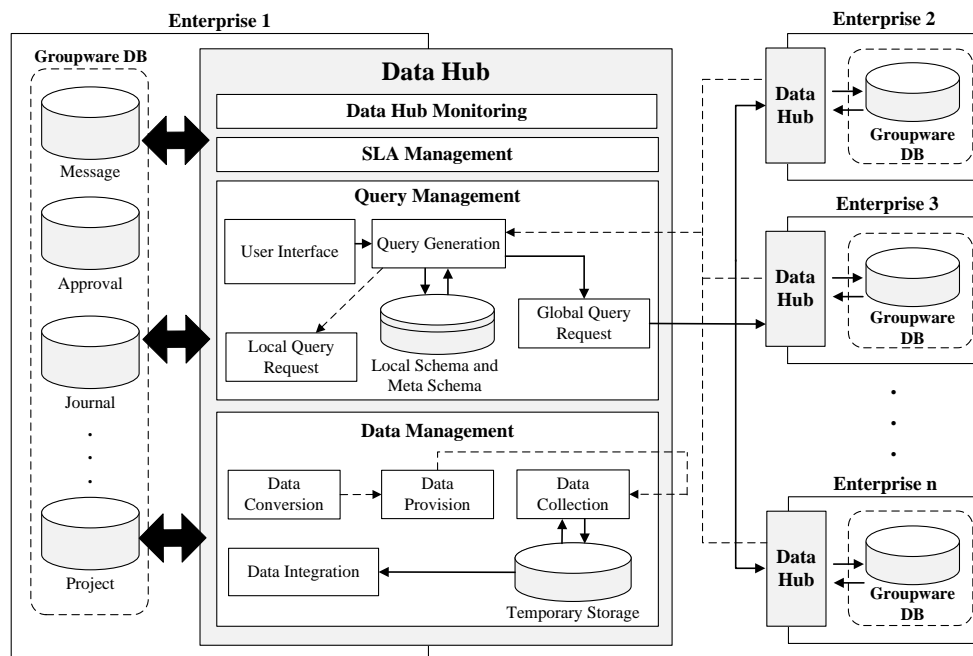


Figure 2. DBaaS-based Data Hub Architecture

Each function of DBaaS-based Data Hub system depicted in Figure 2 is as follows.

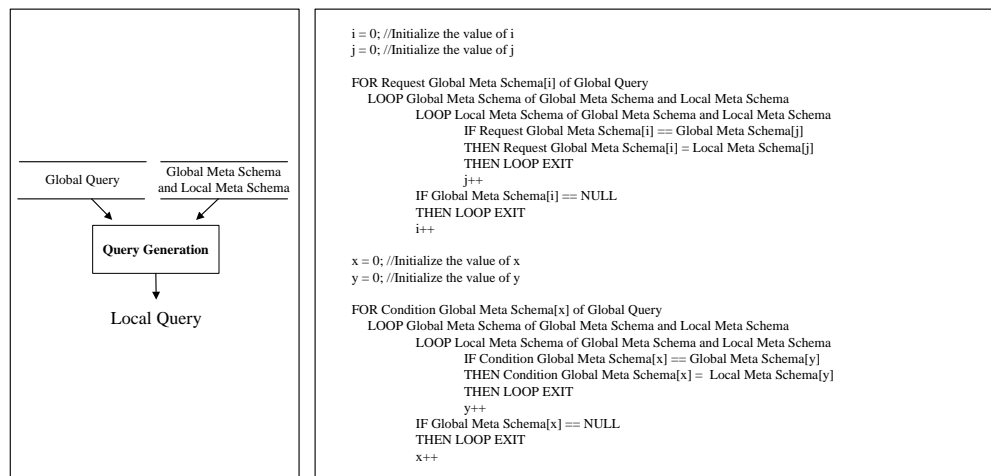
- Data Hub Monitoring: It is an area to monitor the overall queries and data input and output of the Data Hub system.
- SLA Management: It will manage the definition between data providers and service request to use the cloud. It will guarantee a certain level of service, such as performance and availability, and operating system.
- User Interface: The user is an interface for requesting data to the cloud.

- Query Generation: Using the storage local schema data standard schema and data and Enterprise1 have been mapped including the requirements that are transmitted via the interface, to produce a global query that can request data. Moreover, the global query which is transmitted, by reusing the above information, it is converted to local queries that can request data from the DB.
- Local Schema and Meta Schema: Local schema data in the standard schema data and Enterprise1 is stored. Standard schema is a standard defined based on the schema information of the DB which is connected locally.
- Global Query Request: This it will forward the global query that has been generated for each of the Enterprise Data Hub connected locally.
- Local Query Request: This transfers the local query that has been converted into groupware DB of Enterprise1.
- Data Conversion: This converts the data imported from the DB by the local query to the document-oriented data.
- Data Provision: It supplies the data of the converted document directed to the Data Hub system query is requested.
- Data Collection: It collects data sent from the Data Hub system each Enterprise.
- Data Integration: It is to integrate data collected document-oriented, managing duplicate data processed in version method.
- Temporary Storage: Data collection documents directed to temporarily store, and is used in place to integrate data.

### 3.2 Query Generation Algorithms

Utilizing the generated Global Query and “Global Meta Schema and Local Meta Schema” store, the algorithm for generating the Local Query with “Query Generation” is represented in Table 2.

**Table 2. Query Generation Algorithms**



Global Query Table 2 includes a defined user request “Request Global Meta Schema” and conditions, Condition includes a “Condition Global Meta Schema” and requirements on demand. “Global Meta Schema and Local Meta Schema” is, a Global Meta Schema and Local Meta Schema of each Enterprise by mapping, save in the table. Results of the value Local Query is actually to be able to retrieve data from the DB, in the detailed Condition including the coupling with the FROM clause that can know the exact position of the

Local Meta Schema and data defined by the user is constructed.

### 3.3 DBaaS Hub System Workflow

The order of the system and overall processing for the integrated operating environment for companies and businesses need is as follows in FIG.

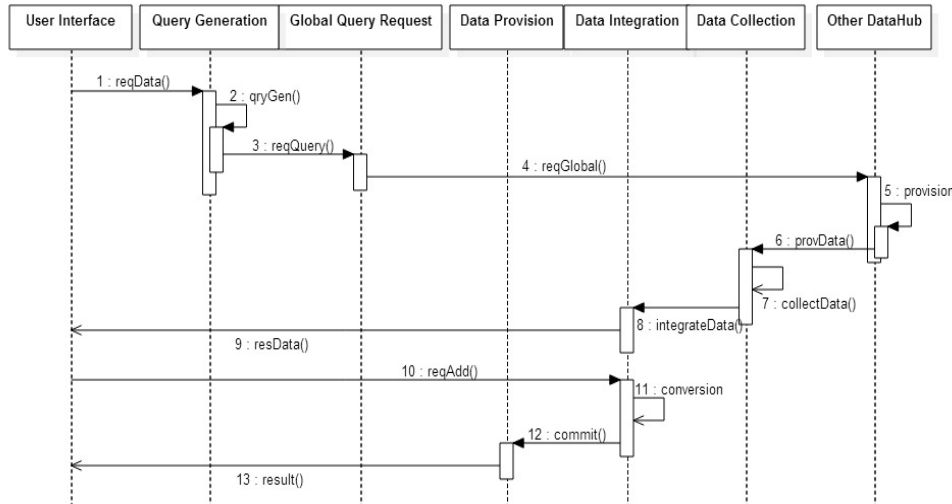


Figure 3. DBaaS System Workflow

Each statement represented in Figure 3, has the following functions.

1. reqData(): Data of Enterprise1 users approval information DB other Enterprise n - document form, approval document name, approval history, date, and approval person to request through the interface of the Data Hub.
2. qryGen(): To generate a global query from the rest of the Enterprise n to the cloud to the data information of the purpose can be obtained. Global query generated is sent to the Global Query Request using reqQuery().
3. reqGlobal(): It runs a global query that is generated. Enterprise n having received the global query via respective Query Generation, generates a local query to be able to get the data information from the groupware DB.
4. provision: "Data Conversion&Provision" of course, collects the requested data from the groupware DB, and converted into document-oriented data, and supplies the converted data to the user's Data Hub.
5. collectData(): It collects document-oriented data supplied from the Data Hub connected.
6. integrateData(): The integrated data collection document-oriented processing of the duplicated data, and manages it by cumulative updates by using the version method.
7. resData(): To provide a document-oriented data that has been integrated into the user via the interface.
8. reqAdd(): To be able to store the data of the provided document-oriented to the user's groupware DB, to request the requirements of through the interface of the Data Hub.
9. conversion: It wants to convert the data to a local format to match the groupware DB type in the "Data Integration".
10. commit(): Use the "Data Provision", and stores the data in the local format to the user groupware DB.
11. result(): The values of the results stored and provided via the interface.

## 4. Application Example and Comparative Analysis

### 4.1 Application Example

Table 3, the table is composed of Local Meta Schema of Global Meta Schema and Enterprise1 stored in the Data Hub system. This can be used to configure the mapping information between the database. Local Meta Schema is defined differently depending on the groupware DB in each Enterprise.

**Table 3. Standard Meta Data stored in the Enterprise1 data Hub System**

Global Meta Schema	Local Meta Schema
approval_id	Approval.a_id
approval_document_form	Approval.a_doument_form
approval_drafter	Approval.a_drafter
approval_drafter_position	Position.p_name
approval_drafter_department	Department.d_name
approval_subject	Approval.a_subject
approval_contents	Approval.a_contents
approval_status	Approval.a_approval_status
approval_public	Approval.a_public
approval_create_date	Approval.a_create_date
approval_modify_date	Approval.a_modify_date
approval_file	Approval.a_file

Table 3 assumes, Table 4, there is shown the format of the data of the document-oriented and process of executing the query process are stored based on user requirements.

**Table 4. Query transformation and document-oriented format**

<b>Global Query</b>	<pre>SELECT APPROVAL_ID, APPROVAL_DOCUMENT_FORM, APPROVAL_DRAFTER, APPROVAL_DRAFTER_POSITION, APPROVAL_DRAFTER_DEPARTMENT, APPROVAL_SUBJECT, APPROVAL_STATUS, APPROVAL_CREATE_DATE WHERE APPROVAL_PUBLIC = 'public' AND APPROVAL_CREATE_DATE &gt;= '2014-01-01'</pre>
<b>Local query</b>	<pre>SELECT app.a_id, app.a_document_form, app.a_drafter, pos.p_name, dept.d_name, app.a_subject, app.a_approval_status, a_create_date FROM approval app, employee emp, position pos, department dept WHERE app.a_public = 'public' AND to_char(app.a_create_date,'YYYY-MM-DD') &gt;= '2014-01-01' AND app.e_id = emp.e_id AND emp.p_id = pos.p_id AND emp.d_id = dept.d_id ;</pre>
<b>Document-Oriented File</b>	<pre>{ APPROVAL_ID: " 204404", APPROVAL_DOCUMENT_FORM: " Application " , APPROVAL_DRAFTER: " Hong Gil-Dong " , APPROVAL_DRAFTER_POSITION: " Staff " , APPROVAL_DRAFTER_DEPARTMENT: " Marketing Dept " , APPROVAL_SUBJECT: " Proposal project " , APPROVAL_STATUS: " Complete payment " , APPROVAL_CREATE_DATE: " 2015-02-05" }</pre>

Table 4 of the execution procedure is as follows.

- Phase 1: FROM and is a representation of the query in the absence Global Query.
- Phase 2: It is converted to local query to generate a FROM clause, to represent the shape of the expanded query.

- Phase 3: Collects data locally query, it shows the values of the results are converted into data in the file-based document-oriented.

#### 4.2 Comparison of Related System

Table 5 shows the results of the comparative analysis of DBaaS system with Oracle DBaaS in the cloud environment. Overcoming and heterogeneity that both the distributed data to the system by using a standardized during consolidation, it is possible to construct a unified work environment to minimize the effects of local application[14][15].

**Table 5. Comparison of Related System**

	Oracle DBaaS	Data Hub System
<b>Integration</b>	By using the schema level and the application level, which is defined in the schema type of master data, to integrate the data.	By using the schema level and application level, which is defined in the DBaaS-based Data Hub system, to integrate the data.
<b>Data Scalability</b>	Data standardization considering the integration and centralized system to collect information.	Data standardization considering the integration and centralized system to collect information using Meta Data.
<b>Query</b>	SQL	Global query based on the Meta Data stored in the standard schema format.
<b>Type</b>	Static	Dynamic
<b>Result</b>	Table Format	Document-oriented File Formats

In the process of integrating data, Oracle DBaaS is a method of centralized control of data from the central[15]. In the Data Hub system utilizes the meta data to create a global query, connect directly to a local system, using the method of extracting data. The heterogeneity of the schema using the mapping using the global meta-schema and the local metadata schema, with minimal changes in the system during the expansion of the local system, and to be able to seamlessly integrate.

Data Hub system, via a newly stylized processing technology, and Oracle DBaaS offers a different differentiated solutions. This indicates the desired performance in the cloud environment for the purpose of integration. For integration groupware goal that is discrimination of the company, from Oracle DBaaS, author that the performance of the Data Hub system proposed in this paper has a high advantage point it is expected.

## 5. Conclusion

A group of companies, for the purpose of sharing and effective data transfer, if it is that the business monitoring of a number of companies, issues and economic costs of the problem occurs in the development environment to. Even using the same groupware services, because the schema format table that manages the database are different, the only coupling occurs even interoperability. Therefore, in this paper, we have proposed a cloud-based Data Hub system for the business environment that is enhanced by integrating the

data. Data Hub system, through the work to be purified by extracting data in real time, consistent and can generate some information, the quality of information can be greatly improved. Also, by integrating databases associated service is used, it is possible to require various databases to match the groupware service. In the future, not being dependent on the device operations are performed in an enterprise, so as not limited by the movement, are focused only on the potential for development based on the cooperation with mobile cloud groupware service.

## References

- [1] Hyun-Sun Ryu, et al., "Empirical Study on Factors Affecting the Adoption of Software as a Services(SaaS) for Integrating Group of Companies," Proceedings of the Korea Multimedia Society Conference, pp. 652-655, 2010.
- [2] Jeong-Su Park, Yu-Mi Bae, and Sung-Jae Jung, "Technical analysis of Cloud Storage for Cloud Computing," *Journal of the Korea Institute of Information and Communication Engineering*, 17.5, pp. 1129-1137, 2013.
- [3] Wu, Jiyi, et al., "Cloud storage as the infrastructure of cloud computing," Intelligent Computing and Cognitive Informatics (ICICCI), 2010 International Conference on. IEEE, 2010.
- [4] Abu-Libdeh, Hussam, Lonnie Princehouse, and Hakim Weatherspoon, "RACS: a case for cloud storage diversity," Proceedings of the 1st ACM symposium on Cloud computing, ACM, 2010.
- [5] Ellis, Clarence A., Simon J. Gibbs, and Gail Rein, "Groupware: some issues and experiences," *Communications of the ACM*, 34.1, pp. 39-58, 1991.
- [6] Coleman, David, *Groupware technology and applications: an overview of groupware*, Groupware. Prentice Hall International (UK) Ltd., 1995.
- [7] Gutwin, Carl, and Saul Greenberg, "Workspace awareness for groupware," Conference Companion on Human Factors in Computing Systems, ACM, 1996.
- [8] Zohreh Sanaei, Saeid Abolfazli, Abdullah Gani, Rajkumar Buyya, "Heterogeneity in Mobile Cloud Computing: Taxonomy and Open Challenges," *IEEE Communications Surveys & Tutorials*, 16, 1, pp. 369-392, 2014.
- [9] Subashini, S., Kavitha, V., "A survey on security issues in service delivery models of cloud computing," *Journal of Network and Computer Applications*, 34, 1, pp. 1-11. 2011.
- [10] Lehner, Wolfgang, and K-U. Sattler, "Database as a service (DBaaS)," Data Engineering (ICDE), 2010 IEEE 26th International Conference on. IEEE, 2010.
- [11] Dillon, Tharam, Chen Wu, and Elizabeth Chang, "Cloud computing: issues and challenges," Advanced Information Networking and Applications (AINA), 2010 24th IEEE International Conference on. Ieee, 2010.
- [12] Kiefer, Tim, and Wolfgang Lehner, "Private table database virtualization for dbaas," Utility and Cloud Computing (UCC), 2011 Fourth IEEE International Conference on. IEEE, 2011.
- [13] Kye-Dong Jung, et al., "The Study of DBaaS Hub System for Integration of Database In the Cloud Environment," *Journal of Digital Convergence*, 12.9, pp. 201-207, 2014.
- [14] Jung, Kyedong, et al., "A Study on Abstract Query for Data Integration of DBaaS in Cloud Environment," International Conference on Convergence Technology, Vol. 4, No. 1, 2014.
- [15] Nayak, Ameya, Anil Poriya, and Dikshay Poojary, "Type of NOSQL Databases and its Comparison with Relational Databases," *International Journal of Applied Information Systems*, 5.4, 2013.