Virtual Database Technology for Distributed Database in

Ubiquitous Computing Environment

Submitted by

Shunqi Zheng

Junior

Computer Science

Gildart Haase School of Computer Sciences and Engineering

A **project report** paper submitted in partial fulfillment of the requirements for the course - **CSCI 3331 / 7781 Advanced Database Systems**.

Under the supervision

of

Dr. Avimanyou Vatsa

Assistant Professor, Department of Computer Science



(Fall 2020)

**Virtual Database Technology for Distributed Database**

**in Ubiquitous Computing Environment**

**Introduction**

Nowadys, there are many popular databases, such as Centralised database, Distributed database, Personal database, End-user database, Commercial database, NoSQL database, Operational database,Relational database, Cloud database, Object-oriented database, Graph database.

These massive types of databases are full of our data world, which is called “ Ubiquitous databases hereinafdatater”. If we spent time on database selection and data collection, that would be a waste of time doing this unnecessary work, but what really data analyst should do is to take time analyzing data and extract the potential and useful information that could be used for business value.

So in order to achieve this goal, the primary objective is therefore to develop a virtualization technique in order that the data analyst or

other users can use all ubiquitous databases as if they were recognized as a single database, therefore it can help to reduce the user’s workload. Although it is a better way to use multi-databases in the meantime, it still has its drawback that need to be done such as when such a system encounters a trouble, some incompatibility might occur between the virtual database and the associated real databases. And also, it should be fixed with the recovery that the feature of virtual database need to have.

**Related Research**

There are many related researchs that have done before by other reearchers, for example, for data sources: a way to disseminate information to all users (in a mobile computing environment) without fail, a way to integrate data, and a way to wrap data from various sources (RDBs, Web, and Excel files)

And for XML data structure schema: a XML-to-relational mapping framework and system, a mechanism (for modifying and querying database contents), the cost-based XML storage (mapping engine and space), the integration of XML, the common data model and schema based on XML.

In addition, the researcheres of this paper have done a few things before: examined XML schema advantages, studied recovery techniques. Therefore, this paper itself is to solve such remaining problems: Propose a means to recover the associated databasesallowing user to examine the virtual environment, and ensure the integrity between virtual and real database.

**The Database Schema Virtualization**

Regarding to the common schema structure, a common schema provides the user with informationabout the structure of the virtual database, and is also usedfor checking query statements and constraints. Users can only see a single common schema of all distributed databases.

Next for the common schema creation, this is almost the same as we do in the normal database sql query, such as a database containing the following information is created and an XML schema is then constructed using PostgreSQL Similarly, to verify the basic constraints, in a column, the?primary key constraint is assigned to"Employee ID, Affiliation ID", the unique constraint is assigned to"Affiliation", the external key constraint is assigned to"A ffiliation ID", the check constraint is assigned to "salary", and the"NotNull"constraint is assigned to the"pame". Next regarding to the query conversion, if we use the RDB and XMLDB at the same time, differences will exist in the query language used for each database, e.g., SQL for the RDB and XQuery for the XMLDB.

**The Recovery of Virtualized database**

A common schema should be prepared (XML && RDB), the common-schema() statement in the query is a newly. And in the case of when recovering from trouble at each DBMS, an incompatibility might arise between the virtual database and the real database depending on the timing with which the trouble occurs Defined function for the virtual database, and is used to indicate the root of the common schema. An environment should be deployed and a log and back-up data needed.

So Overall, their technology is going to combine these above together into virtual DBMS as the recovery feature to overcome the trouble happening in the database system.

**Conclusion**

In this research paper, they comes up with: A method of virtualization of ubiquitous databases and a unified fashion way of how XML describes this database schema. And in the future, they are still continued working on: Examine procedure in system trouble recovery and the confirmation of compatibility of the log, and also I learned many knowledge from this paper, such as what location the virtualization should be in our history from SQL, to NoSQL, to distributed systems, to Virtualization, to blockchain technology nowadays, it has massive learning knowledge while reading this paper, and it has similar features and functions that the dataware house has. They all try to combine different types of database sources in order to help data related analysts focusing more on data analysis and extract potential information.

**Reference**

