Mobile Agent based Communication Platform for Heterogeneous Distributed Database

Nikhil Gajjam¹, Dr. Mrs. S. S. Apte²
Walchand Institute of Technology, Solapur

Abstract
With tremendous growth of data in database, distributed database is the current requirement with additional feature of heterogeneity. Today's database is not stored on single site, rather they are stored on multiple sites with extra features like transparency, performance. Such databases are termed as distributed database. Heterogeneous Distributed Database is combination of fragments of database with different schema belonging to different vendor technology namely Oracle, MySQL and MSSQL stored on different sites.

Aim of this work is to build a platform which enables a user to communicate with all the databases with a single query. We use a mobile agent which acts as a communicating agent between all the databases. Horizontal and vertical fragmentations are the two techniques for partitioning the tables of the database. In our work we mainly focus on horizontal fragmentation method.

Keywords: Heterogeneous Distributed Database, Mobile Agent, Aglet.

1. INTRODUCTION
Database is collection of data in specified format. There are 2 types of database on the basis of schema.

a. Homogeneous Database
   - When database consisting of single schema then it is homogeneous.

b. Heterogeneous Database
   - When database is of multiple schemas then it is heterogeneous.

The three important components of a Heterogeneous Database System as pointed out in are autonomy, heterogeneity and distribution.

There are 2 types of database on the basis of locality.

Centralized Database
- When database is stored at single site/single machine then it is centralized.

Distributed Database
- A distributed database is a set of fragments distributed over a set of computers. A distributed database management system (distributed DBMS handles data distribution and provides feature of transparency.

Use of horizontal and vertical partitioning makes database distributed. Horizontal partitioning divides the table row wise and vertical partitioning divides table column wise. We use union operation for merging the result in horizontal partitioning and join operation in vertical partitioning.

Distributed Database is again divided into Homogeneous Distributed Database and Heterogeneous Distributed Database.

Homogeneous Distributed Database
- All sites consisting of same database management system software.

Heterogeneous Distributed Database
- Fragments of table is stored on multiple sites in different schemas is termed as heterogeneous Distributed Database.

Fig 1: Heterogeneous Distributed Database

Mobile Agent
- Mobile agent is nothing but the object whose position is not fixed. It is free to travel among the hosts in the network, Created in one execution environment, it can transport its state and code with it to another execution environment in the network, where it resumes execution [11].
Mobile agent is an entity that has the ability to migrate from one machine to another in a network and resume from that site. They can:

- reduce network traffic
- overcome network latency
- make better use of resources
- adapt dynamically
- be robust and fault tolerant

**Aglet as a Mobile Agent**

It is a Java-based autonomous software agent. An Aglet is a Java based mobile agent platform and library for building mobile agents based applications. An Aglet has the ability to move autonomously and spontaneously from one host to another carrying some data with it. It can be programmed to execute at a remote host and show dissimilar behaviors at dissimilar hosts [3].

## 2. RELATED WORK

S. Papastavrou, G. Samaras, and E. Pitoura [12] worked on Mobile Agents for WWW Distributed Database Access. They concentrated on the platform of Distributed Database but not on Heterogeneous Distributed Database.

Danny B. Lange and Mitsuru Oshima [3] presented Mobile Agents with Java: The Aglet API and gave a brief introduction regarding mobile agents. They also gave Aglets' event-based programming model with functionality aglet listeners.

JOHN MILES SMITH, PHILIP A. BERNSTEIN, UMESHWAR DAYAL, NATHAN GOODMAN, TERRY LANDERS, KEN W. T. LIN, and EUGENE WONG [4] proposed a multibase system for integrating heterogeneous distributed database. They also proposed an architecture for integrating database consisting of mapping language, query graph decomposition.


YIN-FU HUANG AND JYH-HER CHEN [6] worked on the concept of Fragment Allocation in Distributed Database Design. They address issues on allocation of fragments like

1. How a global relation should be fragmented,
2. How many copies of a fragment should be replicated
3. How fragments should be allocated to the sites of the communication network
4. What the necessary information for fragmentation and allocation is.


## 3. PROPOSED WORK

We concentrate on 3 sites (site1, site2, site3) for storing 3 different databases. We can use horizontal partitioning technique for making database distributed and use different databases like MySQL, MSSQL, and ORACLE for making database heterogeneous. These 3 databases are stored at 3 different sites.
Components
1: Client
User enters the single query irrespective of the number of databases. All the databases, query processors are transparent to the server.

2: Server
Server takes the query as input, initialize aglets. Server sends this query & aglet to the server side aglet host present at 3 different sites.

3: Aglet host
Aglet host from both the sides does all the operations on the aglet like receive aglet, transfer aglet and handling other functions.

4: Query builder
Query builder tokenize the input query and forms new standard query.

Methodology
1) Phase I
When client enters any query to the server, that query may contain table from any database. Server receives that query and tokenize. Query Builder residing at server side will form new standard queries. Aglet Host at server side will take queries from query builder. Aglet Host forms 3 new aglets and stores respective query in that aglet and transfers that aglet to all respective sites.

2) Phase II
Aglet Host present at 3 sites will receives the aglet and retrieves query from that aglet. The query will be executed on respective databases and result is stored in aglet. Aglet host then send that aglet again to the server. Now server will retrieve data from all the aglets. It will merge the result and output is given back to the client.
Fig 6: Phase II

Advantages
1) Irrespective of which table is stored on which database, client can execute the query.
2) Client doesn’t need to know about syntax of all the databases. Using only one query language he can execute query of any of 3 databases.

EXPECTED RESULTS
Following results will be expected for 4 standard queries.

SELECT
Select query should gives result including join if it contains more than one table and should work normally if it contains only one table.
E.g.: select roll FROM student
It should retrieve column from table student which may stored on any database at any site.
Select college_name from staff, student
It should retrieve column from table staff and student which may stored on any database at any site and perform the join at server.

INSERT
Insert query should insert multiple records in the table. It should also insert multiple records in single table from single query.
E.g.: insert into student values (10,’abc’, 55.55)
It should insert record in the table student which may store on any database at any site.
Insert into student values (10,’abc’, 55.55), (20,’pqr’, 70.70)
It should insert 2 records in table student

DELETE
Delete query should delete records from the table according to the condition specified in the query irrespective of table present at any site in any database.
E.g.: delete from student where roll=5
It should delete the record of roll number=5.

UPDATE
Update query should modify records in the table on the basis of condition specified in the table irrespective of table present at any site in any database.
E.g.: Update student set name=’xyz’ where roll=5
It should modify record in the table student where roll number is 5.

4. Conclusion and Future Work
Goal of our proposed work is “single query from a client can retrieve joined data from the table on the local database and the table on the remote database with different databases”.
This work will be helpful in areas where companies with different databases decide to merge. It is not feasible for companies to transfer data of one database into another database. Our proposed work will provide uniform schema for them.
We also conclude that heterogeneity which is obstacle on modern database can be removed by concept of aglet and mapping of query.
Still the research is going on in this area with some issues like query optimization, security, concurrency control etc.
REFERENCES


