Chapter 3

TOOLS AND TECHNOLOGIES USED IN THE RESEARCH

3.1 CHAPTER OVERVIEW

This chapter describes various tools and technologies used in the research to set up heterogeneous distributed database environment, data collection, analysis, implementation, result comparison and hypothesis testing.

3.2 ORACLE HETEROGENEOUS GATEWAY

To set up the heterogeneous distributed database environment, Oracle’s heterogeneous gateways are used [45][47][100][102]. These gateways work as a communication channel between Oracle DBMS and non-Oracle DBMS. Using these gateways, distributed data stored on the remote databases could be accessed and manipulated from the Oracle RDBMS. It provides location and database transparency in the distributed environment, which means there is no need to provide physical address of the machine where the remote database is stored. Using Oracle’s commands, user can access data from the remote databases. Oracle provides gateways to many non-oracle systems such as DB2, Sybase, Informix, MS SQL Server, etc. It also provides gateways for ODBC which is general solution that uses ODBC driver to access any ODBC compliant database system such as MS Access, PostgreSQL, MySQL, MS Excel, etc.

Oracle Heterogeneous Gateway is composed of two components: Heterogeneous Service (HS) and an agent. Heterogeneous Service is common for all the non-oracle systems. It provides the technology to connect with non-oracle systems, recognize SQL, process PL/SQL and mapping to access metadata. It also maintains the transactions coordination between oracle and non-oracle...
databases. The agent provides datatype conversion, mappings and interface between oracle and non-oracle databases. It interacts with the HS to provide transparent connectivity between oracle and remote databases. Figure 3.1 shows the architecture of Oracle heterogeneous gateway. To connect Oracle and non-oracle database, heterogeneous gateway and client libraries are required to be installed. The database link to connect with a non-Oracle database instance should be created in Oracle, ODBC driver of specific database to be installed and required parameter files need to be created. The detailed process is described in Chapter 4.

![Figure 3.1 Architecture of Oracle Heterogeneous Gateway](image)

The following sequence is followed to process the query from Oracle to non-Oracle database.

1. The query is sent by user application to Oracle through oracle net.
2. HS gateway converts query into the form which is understandable by non-oracle database.
3. Then it is sent through oracle net to the HS gateway.
4. Using username and password specified in the database link created on oracle for the non-Oracle database, the gateway logs in to non-Oracle database.
5. Gateway fetches data using query statement of non-Oracle database.
6. Gateway converts the fetched data into the format of oracle database.
7. Result is sent by the gateway to oracle database using oracle net.
8. Oracle sends result to the user application using oracle net.
The Oracle Heterogeneous Gateway for Oracle 11g is installed using the installation files downloaded from OTN (Oracle Technology Network) web site. OTN is a web site which provides many Oracle DBMS products for educational purpose for free.

### 3.3 ORACLE CLIENT LIBRARY

Oracle 11g client library should also be installed along with the gateway to enable the heterogeneous services.

### 3.4 ODBC DRIVERS

ODBC (Open Database Connectivity) is used to access data from database using the ODBS interface. ODBC driver for each database is different. It is not provided by Oracle. User has to install the driver for specific database. To connect Oracle with PostgreSQL, MySQL and MS Access DBMS, the respective drivers are installed.

### 3.5 PROGRAMMING LANGUAGE – JAVA

Java is a very powerful programming language which enables user to write and compile programs which could later be used on any platform without recompilation which contains JVM (Java Virtual Machine), irrespective of computer architecture. To develop the GUI (Graphical User Interface) of TIER, Java programming language has been used.

### 3.6 EDITOR – NETBEANS IDE

NetBeans is an open source Java-based Integrated Development Environment (IDE) which is used to develop desktop and web applications in Java and other languages. NetBeans IDE 8.0.2 has been used as an editor during this research to develop the tool TIER.

### 3.7 OPERATING SYSTEM – WINDOWS

Microsoft Windows 32-bit is used as an operating system. However, the tool which is developed could be executed on any version of the Windows Operating system.
3.8 C#.NET

C#.Net is used to implement Apriori Algorithm. After implementation of Apriori Algorithm in C#.Net, the executable file of the code is called on the TIER interface.

3.9 BENCHMARK FACTORY FOR DATABASES

There are number of benchmarks available for the database [101]. The well known database benchmarks are given by Transaction Performance Processing Council (TPC). These benchmarks are widely used by researchers, academicians and IT professionals. The freeware mode of Benchmark Factory 7.0 has been installed to evaluate the results obtained from experiments. It provides TPC-C and TPC-H benchmarks. In TPCC and TPCH, C and H are used as version numbers. TPC-C is a benchmark for online transaction processing (OLTP) and TPC-H is a benchmark for decision support. TPC-C benchmark has been used to load data into oracle and other RDBMS, while TPC-H benchmark is used to run scalability test to evaluate and compare performance of two transactions which are executed concurrently by 1 to 10 users. The performance of transactions before index recommendation and after index recommendation by TIER is compared using benchmark factory. The standard transactions (query) are generated by TPC-C benchmark itself.

3.10 RELATIONAL DATABASE MANAGEMENT SYSTEMS

The following RDBMSs are used to establish multiple heterogeneous distributed database environment.

3.10.1 Oracle

Oracle 11g is installed from OTN. This version includes functionalities of partitioning and materialized views. The SQL Plus Oracle client is also installed to access data from Oracle Server. This edition is available for free on OTN for educational use. It is installed on a stand-alone machine. The database instance and required users are created after installation to access the data.
3.10.2 MySQL

WAMP (Windows, Apache, MySQL and any one of PHP, Perl or Python) Server is installed to access MySQL 5.6.12 from it. WAMP Server could be installed only on windows. Using MySQL Console, the database is accessed and data could be managed within it. MySQL ODBC 5.3 Unicode Driver is also installed to connect MySQL and Oracle databases through oracle heterogeneous gateway.

3.10.3 PostgreSQL

PostgreSQL 9.2 is installed which is available for free. Data is managed using either PGAdmin or PSQL (SQL Shell) clients. PostgreSQL ODBC Driver (ANSI) is installed and used to connect PostgreSQL and Oracle databases through oracle heterogeneous gateway.

3.10.4 MS Access

Microsoft Access is used which is an RDBMS available with Microsoft Office 2007 package. Microsoft Access Driver is used to connect MS Access and Oracle databases through oracle heterogeneous gateway.

3.11 PHStat FOR STATISTICAL ANALYSIS

To do statistical analysis of data collected from experience survey, PHStat (for Microsoft Excel 2007) is used which is available for free.

3.12 CHAPTER SUMMARY

This chapter covered all the tools and technologies which are used in research. Using these, the tool TIER is developed, statistical analysis is done and performance evaluation of obtained results is done.