**ASSIGNMENT – 1**

**Aim:** Study of Open Source Databases: MySQL/mongodb/Hbase.

**Theory:**

A brief description of Open Source Database :

**Open Source Software:**

Open-source software (OSS) is [computer software](http://en.wikipedia.org/wiki/Computer_software) with its [source code](http://en.wikipedia.org/wiki/Source_code) made available and licensed with a [license](http://en.wikipedia.org/wiki/Open-source_license) in which the [copyright](http://en.wikipedia.org/wiki/Copyright) holder provides the rights to study, change and distribute the software to anyone and for any purpose.

A database is a base for data. An Open Source database is a base for data that includes Free and Open Source Software. Open source [software](http://it.toolbox.com/wiki/index.php/Software) is software that makes the source code available to anyone. The user is allowed to implement, share and further develop the database software to suit various needs.

## Popular Open source Databases

**1)   MYSQL**

Mysql is the world’s most widely used open-source [relational database management system](http://en.wikipedia.org/wiki/Relational_database_management_system) (RDBMS). The [SQL](http://en.wikipedia.org/wiki/SQL) phrase stands for [Structured Query Language](http://en.wikipedia.org/wiki/Structured_Query_Language). Mysql is a popular choice of database for use in web applications, and is a central component of the widely used LAMP open source web application software stack.

**Features of mysql:**

* Because of its unique storage engine architecture mysql performance is very high.
* Supports large number of embedded applications which makes mysql very flexible.
* Use of Triggers, Stored procedures and views which allows the developer to give a higher productivity.
* Allows transactions to be rolled back, commit and crash recovery
* Embedded database library
* Full-text indexing and searching
* Updatable views
* Cursors
* Triggers
* Cross-platform support

**2) MONGODB**

Mongodb is a cross-platform document-oriented database. Classified as a [nosql](http://en.wikipedia.org/wiki/NoSQL) database, mongodb eschews the traditional table-based [relational database](http://en.wikipedia.org/wiki/Relational_database) structure in favor of [JSON](http://en.wikipedia.org/wiki/JSON)-like documents with dynamic [schemas](http://en.wikipedia.org/wiki/Database_schema) (mongodb calls the format [BSON](http://en.wikipedia.org/wiki/BSON)), making the integration of data in certain types of applications easier and faster. Released under a combination of the [GNU Affero General Public License](http://en.wikipedia.org/wiki/GNU_Affero_General_Public_License) and the [Apache License](http://en.wikipedia.org/wiki/Apache_License), mongodb is [free and open-source software](http://en.wikipedia.org/wiki/Free_and_open_source_software).

**Some of the main features include:**

* Ad hoc queries
* Indexing
* Replication
* Load balancing
* File storage
* Aggregation
* Server-side javascript execution
* Capped collection

**Limitations of mongodb:**

* 4 MB/16 MB document size limitation depending on version
* Read/write lock is currently global level
* No joins across collections
* No transaction support
* No referential integrity support

**3)   Hbase**

**WHAT IS HBASE ?**

It is a distributed column oriented database. Which is built on top of HDFS.

An open source project, which is horizontally scalable.

Designed to provide quick random access to huge amount of structured data. It is part of hadoop ecosystem, giving real time read-write access to data in HDFS

**FEATURES**

Linearly scalable.

Automatic failure support

Provides consistent read and writes.

Integrates with Hadoop, as a source and a destination.

Has easy java API for client.

Provides data replication across clusters.

**Conclusion:** Awareness of Different Open Source Databases.