

Choosing a Database for Your Data Warehouse

Choosing a Database for Your Data Warehouse

This document helps administrators evaluate various database options for a large, campus-wide data warehouses. There are three databases evaluated here: MySQL, Oracle and SQL-Server and evaluate them based on licensing and maintenance costs, functionality, and security.

Licensing and Maintenance Costs

Oracle is without a doubt the most powerful database available. Without a doubt it is also the most expensive database available. Licensing fees alone can easily be over \$100,000. In addition, Oracle can be very complicated to set up and administer, so a full-time database administrator is necessary, and expensive. This can cost another \$100,000 or more per year.

MySQL is on the other end of the cost spectrum. Total licensing and support costs are only a few thousand dollars. The total cost of ownership is much higher though, because it is still considered fairly difficult to maintain, integration can be difficult and it can be difficult to find database administrators skilled in MySQL. Notably, the user interface has been improving in recent versions.

SQL-Server falls between MySQL and Oracle in terms of price. It is by far the easiest database to maintain. The tools are very user-friendly and the help files are thorough. It is much easier and cheaper to find a skilled database administrator for SQL-Server than for Oracle or MySQL. Maintenance can also

be easier due to SQL-Server's tight integration with Windows. Trusted connections (discussed below) can be used to avoid creating and maintaining another set of usernames and passwords.

Functionality

In the area of functionality SQL-Server and Oracle are far more mature in terms of their data warehouse support than MySQL. MySQL lacks several features necessary for building a data warehouse that come standard with Oracle and SQL-Server:

1. a strong packages for handling the extract, transform, and load (ETL) phase
2. server support for OLAP
3. client support for OLAP
4. reporting services
5. data mining

To get this functionality in MySQL, several tools would have to be pulled together from several small vendors to get this functionality. Getting all of these tools to work with each other while keeping up with new versions and security patches can be time-consuming and aggravating. With Oracle and SQL-Server, these tools are integrated with the database and come from the same vendor, making maintenance much easier.

Developing a front-end is also much easier in Oracle and SQL-Server due to ODBC drivers for these databases included with Windows. Although MySQL has an ODBC driver, this is not supported by Microsoft. In the past, such features have failed to work with subsequent Microsoft releases – this could completely break a data warehouse.

Security

SQL-Server has become a victim of its own user-friendly interface. Because it is so user-friendly, administrators often neglect to tune and secure it on a regular basis. If it is secured properly, however, it can be much more secure than MySQL and even Oracle in some respects. Because of this, there are many examples of organizations using SQL-Server for sensitive information. For example, the next

version of Datatel's Colleague will be in SQL-Server. A separate document describes how to properly secure SQL-Server to have enterprise-level security.

For schools that are using Windows, SQL-Server can be the most secure database choice in many respects. For example, SQL-Server allows trusted connections. This means a user must be logged into the Windows network in order to access the database. In contrast to Oracle and MySQL, the user's username and password are never sent across the network, eliminating the risk of someone intercepting it in transit. It also eliminates the security risk caused by people accessing the database through scripts or third-party tools that cache the user's name and password. Furthermore, because the logons are integrated with Windows, the network administrator can control password expiration, complexity using the same methods currently used to control passwords for Windows users.

In case there are any breaches of security, they are very easy to track in SQL-Server and Oracle due to their implementation of C2 Level Auditing – approved by the Department of Defense. For example, it's possible to automatically send an email when someone is trying to access information that they should not be accessing, such as Social Security Numbers. MySQL's auditing features are still rudimentary.

Conclusion

For schools that already have Oracle databases and an Oracle database administrator, running Estudios Enterprise on Oracle is probably the best option. Although MySQL is great for general-purpose database tasks, it lacks the tools necessary for an effective and secure data warehouse. If development continues at the rate that it has, MySQL will certainly have these features, but it will take a few years. SQL-Server is probably the most secure and cost-effective choice for most schools, especially those that already use Microsoft products and even have some SQL-Server databases running.