

SQL Server 2005 vs. Oracle 10g

A Neutral Comparison

Ahmed Bahaa- Jan 2007



Table of Contents

1-Introduction	3
2-Information sources for this paper	3
3- Information source Background	4
3.1 Performance Tuning Corporation Background	4
3.2 Computer World Magazine	4
3.3 Intelligent Enterprises Magazine	4
4-VLDB Enterprise Features	4
4.1 Oracle RAC Marketing Materials Claims	4
4.2 RAC in a More Practical Look vs. SQL Server 2005	5
4.2.1 Availability	5
4.2.2 Reliability	7
4.2.3 Availability/Reliability Comparison Conclusion.....	8
4.2.4 Manageability.....	8
4.2.6 Scalability and Performance	12
4.2.5 Security	14
4.2.7 Cost	17
5- Business Intelligence	18
5.1 Market shares and Rankings.....	18
5.2 Why SQL Server has the biggest Market Share?	19
5.3 SQL Server 2005 BI Solution Advantages	19
5.3.1 Data Warehouse and BI Phases coverage.....	19
5.3.2 Cross integration out of the box	20
5.3.3 Office integration.....	20
5.3.4 Broad Range of Data Mining Algorithms.....	21
5.4 BI Conclusion	21
6. References and Case studies	21
6.1 Education and Research References	21
6.1.1 International References.....	21
6.2 Applications Related to the Education Domain.....	22
6.2.1 International References.....	22
6.2.2 Local References	22
6.3 Other References	23
6.4 Other Local References	23
6.5 SQL Server 2005 Fast Facts	23

1-Introduction

The purpose of the following white paper is to provide a feature comparison between the two leading Relational Database Management Systems (RDBMS) products on the market; MSSQL 2005 and Oracle 10g (up to R2). We will compare the VLDB, OLAP, and development related features. This paper will unveil many of the marketing materials myths that have been used as the core evaluation base at many previous comparison papers. To be more subjective and neutral, we will try to use prestigious corporate reports in this comparison. This will help readers of this paper to be confident in the results and conclusion that will be gotten out of this paper.

2-Information sources for this paper

The sources of information used whilst compiling the data that forms the basis of this document include:

- A Report Generated by Performance Tinning Corporation and titled as: Solutions for Highly Scalable Database Applications An analysis of architectures and technologies
- Computer World Magazine
- The CVE website at Mitre :
<http://cve.mitre.org/cgi-bin/cvekey.cgi?keyword=Oracle+Database>
<http://cve.mitre.org/cgi-bin/cvekey.cgi?keyword=Microsoft+SQL+Server>
- Report Generated by David Litchfield and could be downloaded from NGSSoftware Insight Security Research (NISR) website (<http://www.ngssoftware.com>)
- [www.TPC.org](http://www.tpc.org)
- Some Oracle Technical Materials (e.g. **How RAC Works**).
- Some Microsoft Technical Materials.
- Other local as well as international references.
- Oracle Book: *Oracle 10g and Real Application Clusters*.
- <http://www.linxcel.co.uk>
- Intelligent Enterprises IT Business Magazine
- ACM Digital Library

3- Information source Background

3.1 Performance Tuning Corporation Background

Before getting into more details into this point, it is important to highlight who this company is. Performance Tuning Corporation is one of the largest corporates all over the world, who are interested in tuning the performance of the IT system for the major IT systems products. Its customers list contains; **Oracle, HP, Dell, US Navy, US Marines, US Air Force, Verizon, Microsoft, and more...**

3.2 Computer World Magazine

Computer world magazine is one of the most prestigious IT Magazines all over the world. It is more than 35 years old, and its editorial board consists of many very prestigious, very famous IT consultants.

3.3 Intelligent Enterprises Magazine

It is one of the top two magazines all over the world talking in BI and data warehouses. Its editorial board includes Ralph Kimball (the father of Data Marts), and Barry Devlin the builder of the oldest data warehouse known all over the world.

4-VLDB Enterprise Features

The most important features that any decision maker is highly interested to look for when taking a decision of buying certain VLDB (**Very Large DataBase**) are; availability, reliability, scalability, security, and performance. For VLDBs any weakness found in other features may affect the day-to-day operation of the system administrators, DB specialist, developers, or even the application for a while but, the weaknesses that could be discovered in the major features listed above here could be fatal as whole for utilizing certain VLDB. In fact, availability, reliability, scalability, security, and performance are the major reasons for any decision maker to use VLDBMS (**Very Large DataBase Management System**) instead of SDBMS (Small DataBase Management System). Based on that, this paper will concentrate at the beginning on these features.

4.1 Oracle RAC Marketing Materials Claims

Oracle Real Application Clusters (RAC) has been presented by Oracle as the one solution that addresses all scalability, performance, high availability requirements, at a lower cost than any other solution. Oracle claims that RAC running on a cluster provides the highest level of capability

in terms of availability, scalability, and low-cost computing. According to Oracle claims, RAC is supposed to provide these benefits:

Availability: If a node in the cluster fails for any reason, Oracle RAC continues running on the remaining nodes. All the applications (users) connected to the failed node are transparently to be completed in 20 seconds.

Scalability: When more processing power is needed, new nodes can be easily added to the cluster, without having to modify the application or the database in any way. The load is redistributed so that it is balanced across all the nodes of the cluster. Oracle 10g R2 RAC supports up to 100 nodes in a cluster.

Cost Savings: RAC reduces hardware cost by running applications just as efficiently on clusters of small (< 4 CPUs), standardized, low-cost commodity hardware as on the more expensive SMP systems. For example a 16 node cluster of 4 CPUs each costs significantly less than an equivalent 64 CPU SMP machine. There are a number of reasons for the price differential, including the fact that smaller boxes benefit from the economies of scale.

4.2 RAC in a More Practical Look vs. SQL Server 2005

The following analysis is based on the Performance Tuning Corporation analytical comparison report:

There have been numerous heated *technical* debates about whether RAC really works as advertised. In depth technical analysis here maybe difficult to most of the readers to understand that is why the paper will try to unveil facts without getting into deep detailed technical matters.

4.2.1 Availability

Oracle RAC Availability

It is true that RAC offers a good solution for server failures. This solution could be implemented on the low cost commodity servers. This makes some cost reduction on the customer hardware budget (this doesn't mean that RAC as a clustering solution is a cost

saving solution, please refer to the part of the price comparison latter in this paper). The RAC architecture is depicted in figure 1 that is shown below:

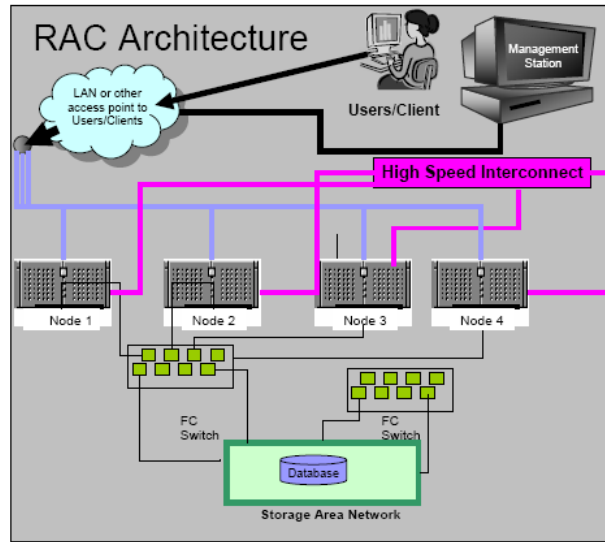


Figure1: A typical 4-node RAC configuration

Figure 1 shows a typical 4-node RAC configuration. The Oracle RAC Database Servers run on all the nodes in the cluster. The data resides in the shared storage. All the nodes in the cluster have equal access to all the tables in the database. There is no notion of data being “owned” by any particular node. As a result data does not have to be partitioned, though very often it is partitioned to increase performance. Applications just connect to the RAC cluster, not to a specific node in the cluster; RAC distributes the load evenly across all the nodes of the cluster.

SQL Server Availability

Database mirroring is a new SQL Server 2005 technology for increasing database availability. Database mirroring ships transaction log records directly from the primary server to a standby server, ensuring that at all times the standby is a mirror image of the primary database. Database Mirroring also quickly fails over to the standby server in the event of the primary server going down for any reason. You can code client applications to automatically redirect their connections so that in the event of a failover, they automatically connect to the standby database. While the precise time to failover depends on a number of factors, it is possible to failover in around 10 seconds (according to Microsoft materials, it could failover in duration between 3-5 seconds, but practical experiences says 10 seconds is more realistic number). Comparing the 10 seconds in SQL Server 2005 with the 20 seconds of Oracle RAC is something notable. The 10 seconds difference of waiting for high availability system may not be acceptable in many business domains users. SQL Server 2005 Database Mirroring does not require proprietary hardware and

is easy to set up and manage. Mirroring supports a feature that is called client redirect. Mirroring architecture is shown in the figure below.

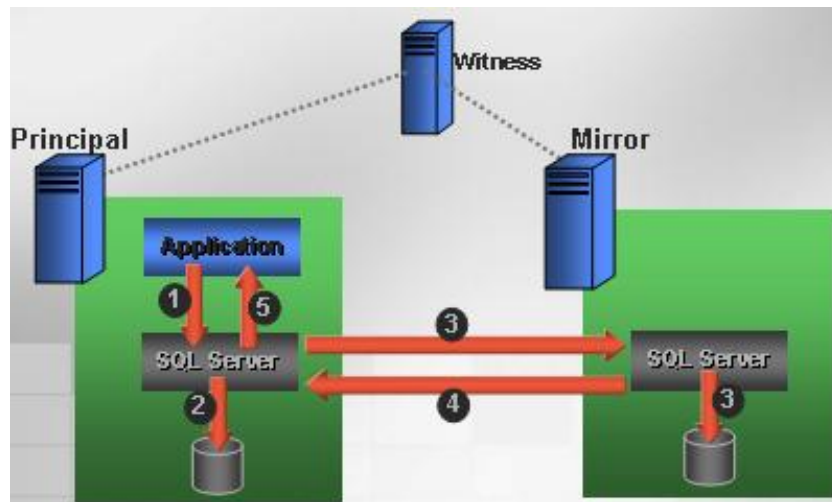


Figure 2: The SQL Server 2005 Mirroring Architecture

This feature enables system application to get automatically redirected to the mirrored server transparently if the session is dropped. This doesn't need any changes to application code. This happens transparently because of that MDAC (**M**icrosoft **D**ata **A**ccess **C**omponent) is aware of Principal and Mirror servers. Upon initial connect to the Principal server, MDAC caches Mirror name. When client attempts to reconnect, if Principal is available, it connects to it, if not; MDAC automatically redirects connection to Mirror. It is important to mention here that the witness server, that is optionally used to make automatic failover, could be SQL Server Express which is free software. Moreover, the licensing model of Microsoft doesn't charge for any standby server. Based on this, the standby server don't requires extra license. There is an important piece of information that should be highlighted here which is that, Oracle claims that one of the most important features in Oracle that it supports clustering while SQL Server doesn't. This is not true. Oracle has a clustering support because Oracle don't have its own operating system. This is not the case for Microsoft. SQL Server is a Microsoft Product that runs always on a Microsoft Operating System (Windows). Clustering in Windows is supported by using the Clustering service of the windows. SQL Server doesn't need to have its own separate extra layer for clustering instead, it integrated with the same vendor operating system to achieve this task without overwhelming the running sever with extra layers

4.2.2 Reliability

Oracle RAC Reliability

RAC by itself offers *no protection against disasters or storage failures*. RAC is based on shared-data architecture; therefore the storage is a single point of failure. If the storage fails for any

reason, the whole cluster fails. Oracle offers DataGuard as the solution to this problem. DataGuard involves having one or more duplicate databases called “Standby” database. “Standby” databases are kept in sync with the “primary” using log-shipping technology. It must be noted that Oracle will charge full price for each “Standby”, so the total cost of an Oracle solution has to be multiplied by the number of standby databases. Away of the costing aspects, installing the Oracle DataGuard solution over the RAC that is installed over the Oracle 10g DBMS itself complicate the management and administration massively. It is widely known for all oracle products that when it comes to integration, Oracle products always suffer from many problems and difficulties.

SQL Server 2005 Reliability

It is apparent from the description above for the SQL Server 2005 mirroring feature architecture that SQL Server 2005 doesn't work on a shared storage model. This solves the problem of single point of failure that could happen for the RAC solution without using the DataGuard solution. When the primary storage server failure happens the standby server with a complete another storage server failover and continue delivers system availability.

4.2.3 Availability/Reliability Comparison Conclusion

Both SQL Server 2005 and Oracle 10g can deliver a highly available solutions but, “SQL Server with Database Mirroring” can match the high availability capabilities of Oracle 10g RAC—at a substantially lower cost.

The above explanation shows that the Oracle RAC out of the box solutions is less reliable solution than the SQL Server out of the box solution. If DataGuard is added to the formula then, from technical point of view SQL Server 2005, and Oracle RAC delivers highly reliable solutions. It is important to highlight here that the Oracle DataGuard availability is based on a difficult to manage solution (this will be shown later.)

4.2.4 Manageability

Oracle Manageability

According to the Performance Tuning Corporation report, the Oracle RAC is very complex to get managed. The report stated that; *RAC is an extremely complex piece of*

technology. Unless the complexity is reduced by an order of magnitude, RAC falls far short of its promise as a viable technology for the vast majority of database applications. Probably the single biggest reason to avoid RAC might be its complexity. The complexity associated with RAC is documented succinctly in “Real-World Challenges for Oracle RAC Implementation” paper (The paper is available in the following URL:

<http://www.linxcel.co.uk/Whitepapers/RealWorld%20Challenges%20for%20Oracle%20RAC%20Implementation-Issue%201.pdf>). (i.e. linxcel is a famous Oracle partner that is specialized in delivering Oracle and Java Professional services in Europe.)

Here are the main problems in RAC manageability:

RAC requires Application and Schema Design Changes

Contrary to Oracle’s claims, according to many experiences has been that applications (and the associated database schemas) have to be specifically designed (or modified in the case of existing applications) in order to get them to perform on RAC. The nature and extent of the changes depends on a number of factors including the size of the clusters, the nature of the application (OLTP or DW), the speed of the cluster inter-connect and transaction volumes. Very often data has to be partitioned, especially tables that are hotspots.

RAC requires special storage solutions

Oracle RAC cannot run on a regular file system i.e. the file systems that usually ship with operating systems. Oracle RAC either works directly on **raw devices** or requires a clustered file system. This Includes:

- Oracle Cluster File System (OCFS)
- Oracle Automatic Storage Manager (ASM)

Due to the inherent complexity of using RAW storage devices, Oracle recommends OCFS or ASM, with more emphasis on ASM lately. While ASM has many interesting capabilities, it is very complex to manage and administer—it is a full-blown instance of Oracle requiring significant DBA time and skills to manage. Furthermore ASM is a proprietary file system; most of the backup & restore, diagnostics, monitoring and performance tuning tools that you are currently using will not work with ASM.

RAC is Composed of Many Components

Actually RAC into itself is composed of many several components:

- Global Cache Services (GCS)

- Cluster Read Service
- Automatic Workload Management
- Virtual Internet Protocol
- Oracle Cluster Registry

Patching RAC is difficult

Patches for RAC come in two flavors—those that can be applied one node at a time and those that need to be applied to the entire cluster at once. In the former case only the node to which the patch is being applied has to be down; the rest of cluster is functioning. In the latter case the entire RAC cluster has to be shutdown, thereby making the entire database unavailable. An analysis of Oracle's patches shows the majority of the patches belong to the latter category. In this case Oracle recommends a very complicated technique for "rolling upgrades" that involve two standby RAC clusters connected with Oracle Data Guard.

Tuning RAC is Complex

The complexity of RAC and the numerous moving parts involved in a RAC setup makes it hard to debug and tune. In addition to all the things that a DBA needs to know to tune an Oracle database, with RAC a DBA has to take numerous other factors into account including interconnect traffic, inter-connect latency, pinging of data blocks between nodes, disk I/O for each of the nodes, table hotspots etc. Here is just a partial list of some of the workarounds that Oracle has suggested to get around the performance issues with RAC.

- Assign transactions with similar data access characteristics to specific nodes, by partitioning users and applications.
- Create data objects with parameters that enable more efficient access when globally shared.
- Avoiding sequences as hotspots by creating node-specific staggered sequence ranges.
- Reduce the number of rows-per-block (RPB) in order to reduce page contention.
- Use as few indexes as possible to reduce intra-node pinging of index blocks.
- Pre-allocate space by turning on dynamic space management.
- Use reverse-key indexes to reduce index-page hotspots. This has the undesirable side-effect of eliminating the ability to use index-scans.
- Design indexes such that the clustering factor is as close to the number of used blocks as is possible.

The list is based on content from the book “Oracle 10g and Real Application Clusters”. These are very complex tuning recommendations, requiring a deep understanding of the inner-working of the Oracle database management system. The book summarizes the situation quite aptly when it says, ***“This may seem perplexing, since some of the suggestions are contradictory”!***

SQL Server 2005 Manageability

MSSQL 2005 administration is as simple as every Microsoft product administration with a nice GUI and only a few parameters required a manual change. The major Goals for the SQL Server 2005 manageability focus on some main objectives to achieve. These objectives are listed below:

- Integrated Management, Rich UI
- Scale down time/effort for common management tasks
- Server Transparency
- Everything to be scriptable

SQL Server has many very easy to use management tools to do whatever administration task you need. This includes;

SQL Server Management Studio

- This tool delivers integrated management environment for all SQL Server products including;
 - Relational DB
 - Analysis Services
 - Reporting Services
 - Notification Services
 - SQL Mobile

There are many tools that comes with MS SQL Server 2005 for manageability, Performance tuning, security analysis, and troubleshooting and diagnostics.

Manageability Conclusion

According to what is depicted above from much neutral technical comparison it could be discovered without any doubt that MS SQL Server is more manageable than Oracle enterprise solutions.

4.2.6 Scalability and Performance

Scalability is one of the most important enterprise features in any VLDB.

Oracle Scalability and Performance

It would appear that Oracle RAC has the edge here. Oracle 10g RAC R2 can support up to 100 nodes. Oracle does not specify whether there are any limitations on the number of CPUs per node. Therefore, **in theory**, Oracle RAC can scale beyond 64 CPUs. However, it must be noted that Oracle has not demonstrated publicly that Oracle RAC can scale beyond 64 CPUs. The **largest TPC-C benchmark with Oracle RAC has 64 CPUs**—a 16 by 4 cluster. The largest RAC installation widely known is a 64-CPU at Amazon, also a 16 by 4 cluster.

Oracle Claims that Oracle can scale out and always concentrate on scaling out rather than scaling-up. The main advantage of the scale-out solution of oracle is that it depends on low cost commodity servers. Oracle tries to promote this, claiming that this is a more cost effective solution (a more detailed look will concentrate on this).

SQL Server 2005 Scalability and Performance

Scaling Up With SQL Server 2005

SQL Server 2005 has proven that it can scale to 64 CPUs on a single SMP server. Interestingly, at some benchmarks, SQL Server 2005 has a better performance and price/performance than Oracle 10g RAC for 64 CPUs as can be seen from the table below.


Company	System	tpmC	Price/tpmC	System Availability	Database
	hp Integrity Superdome	1,231,433	4.82 US \$	06/05/06	Microsoft SQL Server 2005 Enterprise
	HP Integrity rx5670 Cluster 64P	1,184,893	5.52 US \$	04/30/04	Oracle Database 10g Enterprise Edition

Table 1 A TPC benchmarks for SQL Server 2005 and Oracle 10g

These two TPC-C benchmarks clearly demonstrate that on a system with 64 CPUs, SQL Server 2005 can outperform Oracle 10g RAC, and is more cost effective. It is important here to highlight that today the largest SMP server has 64 processors and can run over 99% of the world's real-world applications! According to Winter Corp's Annual Survey⁹ of the largest databases in production today, the largest OLTP and Data Warehousing database run on SMP servers, not clusters.

Scaling Out with SQL Server 2005

You might still prefer the scale-out architectures for the following reasons.

- Some companies have standardized on commodity servers (e.g. Plate Servers), refusing to deploy medium or high-end SMP (Symmetric Multi-Processing) servers. Typically these companies have standardized on vendors such as Dell who do not offer servers with more than 4 processors.
- Some customers have a highly distributed architecture that does not lend itself very well to large centralized SMP servers.
- Some customers are uncomfortable with the 64 processor limit of SMP servers, even if the servers meet their current requirements. They want to know that their database can scale beyond 64 CPUs if the need arises in some distant future.

SQL Server 2005 offers a choice of scale-out technologies. Unlike Oracle, which is single mindedly pushing RAC as the only viable scale-out architecture, Microsoft has taken a more nuanced approach, offering multiple scale-out technologies. Each of these technologies has its own advantages and disadvantages. You should pick the technology that best meets your business requirements. The following are four scale-out strategies with SQL Server 2005.

- Service Oriented Database Architecture (SODA)
- Shared Scalable Database (SSD)
- Peer-to-Peer (P2P) Replication
- Data Dependent Routing

Getting in more details regarding these strategies is out of the scope of this context.

Scalability Conclusion

SQL Server is as scalable as Oracle 10g R2. They are peers. At some cases SQL Server can outperform Oracle 10g. Moreover, it delivers scale-up as well as scale-out solution in less cost than Oracle 10g.

4.2.5 Security

This is one of the peaking points for SQL Server 2005 over Oracle 10g. For years Microsoft products have suffered from security issues but, nowadays this doesn't apply anymore. Microsoft has focused during the last couple of years in increasing the quality all its products from the security point of view. Bill Gates has launched and implemented an initiative to deliver a one week training for all Microsoft Campus developers (20,000 + developers). The training subject was (how to write a secure code). Moreover, Microsoft translates analyze every discovered security hole and learn lessons of it and transform this in a bad programming practices should be prohibited to be done by any Microsoft developer. The following are some very prestigious information sources that made a neutral comparison between SQL Server 2005, and Oracle 10g. As will be seen all of them depicts that SQL Server 2005 is more secured far beyond Oracle, and that one of the major advantages of SQL Server 2005 over Oracle 10g is Security.

4.2.5.1 David Lachfield Paper

David Litchfield who is one of the famous security specialists all over the world who used to criticize Microsoft for its poor security products has written a paper titled "**Which database is more secure? Oracle vs. Microsoft**"

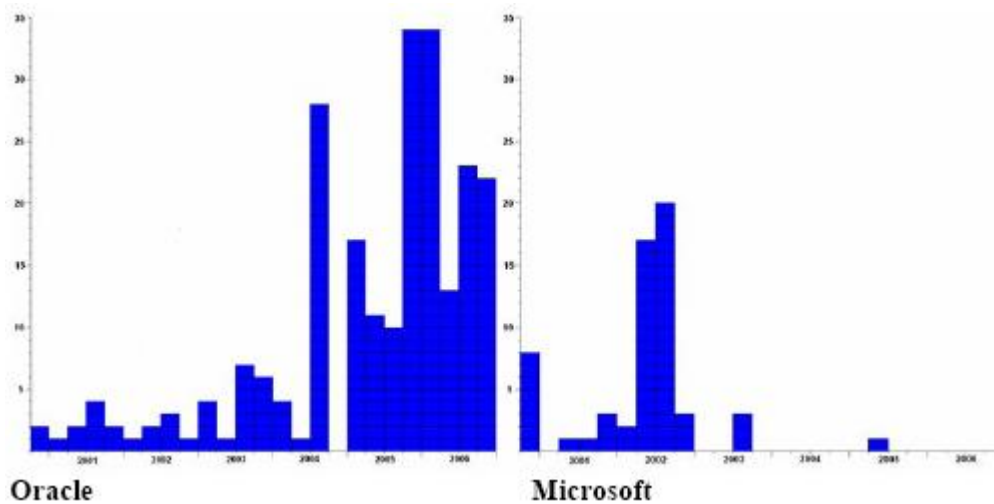


Figure3: The number of security flaws in the Oracle and SQL Server

According to David's paper, Figure 3 shows the number of security flaws in the Oracle and MS SQL Server database servers that have been discovered and fixed since December 2000 until November 2006. Each block

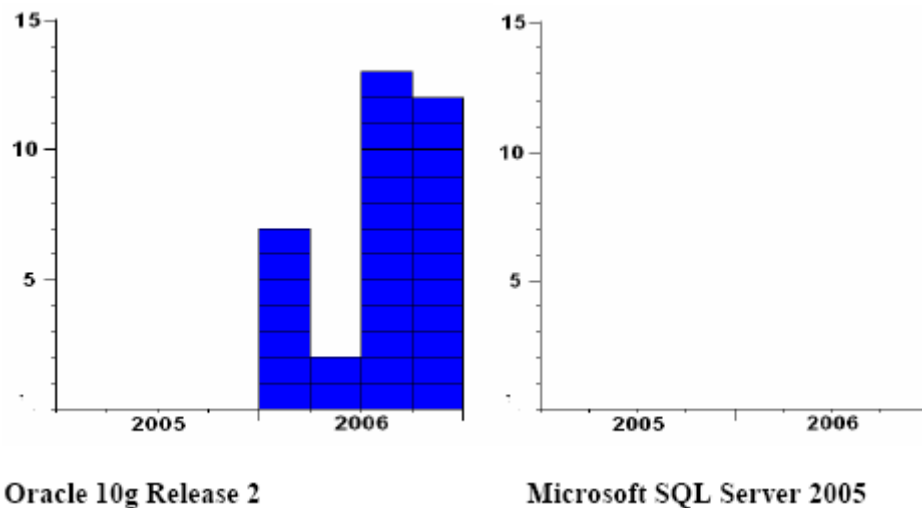


Figure 4 flaws that have been discovered by external security researchers in Oracle 10g R2 and SQL Server 2005

According to David's paper too, these two graphs indicate flaws that have been discovered by external security researchers in both vendors' flagship database products – namely Oracle 10g Release 2 and SQL Server 2005. No security flaws have been announced for SQL Server 2005. David has stated verbally that:

It is immediately apparent from these four graphs that **Microsoft SQL Server has a stronger security posture than the Oracle RDBMS.** Later at his paper he stated, SQL Server code is more secure than Oracle code. He expected that even if oracle have approached fixing these security bugs they may not be able to finish fixing them before end of 2007 or 2008!!

David said that he think the secret behind this gap between Oracle and SQL in the recent products is the Microsoft's Security Development Lifecycle which is not as same security oriented as Oracle – SDL. SDL is far and above the most important factor. A key benefit of employing SDL means that knowledge learnt after finding and fixing screw ups is not lost; instead it is ploughed back into to the cycle. This means rather than remaking the same mistakes elsewhere you can guarantee that new code, whilst not necessarily completely secure, is at least more secure than the old code.

He added that, Oracle keep making the same basic mistakes and that some of their security "fixes" indicate that they don't understand the problems they're trying to fix.

Note: The Whole paper is downloadable from the following URL:

<http://www.ngssoftware.com/research/papers/comparison.pdf>

4.2.5.2 The Computer World Business Magazine

On 19th October 2006 the 35+ years old very prestigious IT business magazine-ComputerWorld- has published an article titled (Let's give Microsoft SQL Server some credit) In which the author has declared that at the same time that SQL Server proofs his powerful security level, Oracle began to suffer massively from security vulnerabilities. He rooted this to the difference in the SDLC of Microsoft as well as Oracle. The article is available for public at the following web address: <http://www.computerworld.com/blogs/node/3778>

4.2.5.3 The Common Vulnerabilities and Exposures (CVEs) Website (<http://cve.mitre.org>)

CVE is A list of standardized names for vulnerabilities and other information security exposures - CVE aims to standardize the names for all publicly known vulnerabilities and security exposures The content of CVE is a result of a collaborative effort of the [CVE Editorial Board](#). The Editorial Board includes representatives from numerous security-related organizations such as security tool vendors, academic institutions, and government as well as other prominent security experts. [The MITRE Corporation](#) maintains CVE and moderates Editorial Board discussions.

At the time of writing this paper, the reported common CVEs for SQL Server 2005 are just four while they exceeds 40 for Oracle 10g. This could be validated through checking the following URLs;

Oracle

<http://cve.mitre.org/cgi-bin/cvekey.cgi?keyword=Oracle+Database>

SQL Server

<http://cve.mitre.org/cgi-bin/cvekey.cgi?keyword=Microsoft+SQL+Server>

Important Note:

It is important to mention here that these two URLs contain all known CVEs for both products across all their versions.

Security Conclusion

According to all reports, papers, and articles above, with a high level of confidence it is clear that SQL Server is more secured far beyond Oracle 10g. Some neutral parties like the famous security expert; David Lachfield, the prestigious IT Magazine; Computer World, and Enterprise Strategy Group considered Microsoft, with proper execution, years ahead of Oracle in producing secure, and reliable database solutions.

4.2.7 Cost

In spite of the fact that Oracle claims that its hardware requirements costs less, when it comes to calculate the overall budget that is needed for an overall oracle solution (on less expensive commodity servers) you will find it more expensive than the overall MS SQL Server 2005 solution. It is true that the single high-end SMP servers that MS SQL Server 2005 requires is more expensive than the commodity servers that oracle requires but, let us calculate the hole payment:

Oracle: \$40,000 (Oracle Price) + \$25,000 (RAC Price) + \$18,000 (the cheapest commodity server available) then, The total of Oracle solution on the Least expensive commodity server is 78,000\$ per processor

MS SQL Server 2005: \$25,000 (SQL Server Price) + \$50,000 (The most expensive known high-end server) then, the total of SQL Server solution on the most expensive high-end server is \$75,000 per processor.

Important note:

We are using the publicly available list prices for both Oracle 10g and SQL Server 2005. It is quite likely that academic as well as governmental accounts don't pays full price for either product. However that does not change the relative pricing differential between Oracle and SQL Server 2005.

Cost Conclusion

Oracle on the cheapest commodity server is more expensive than SQL Server solution. If the claim that the Oracle required hardware solution costs less considered true then, it should be phrased as it cost less for Oracle benefit not the customer benefit. Simply because, Oracle, not the customer, who gets the difference and more.

5- Business Intelligence

5.1 Market shares and Rankings

In spite of the fact that SQL server 2005 has added many very powerful features to the its business intelligence feature group. SQL Server is considered the best integrated BI solution out of the box. The highly prestigious Intelligent Enterprises business magazine within its annual 2002 review, has ranked the MS SQL Server 2000 BI solution as the best BI solution all over the world. Actually, this shows that SQL Server has an old history in excelling the BI area. According to a recent Paper published in Communications of ACM, (paper is available on the ACM digital library) SQL Server BI solution has been shown as the biggest market share for BI Market while Oracle has come in the Seventh place. The figure shown below shows the attitude of OLAP market share for both SQL Server and Oracle.

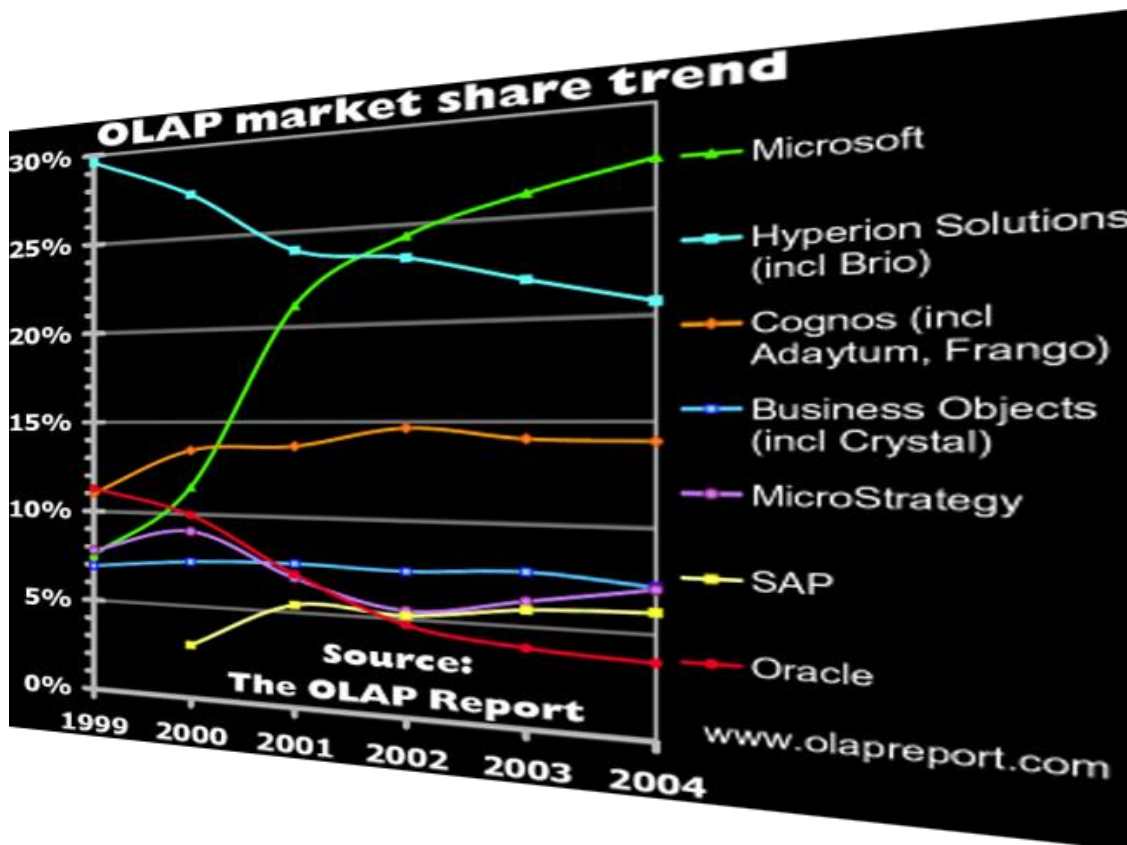


Figure 5: The market shares for different OLAP vendors till 2004

It is important here to mention that Oracle always claims that they have the biggest market share in BI industry. Actually, this is inaccurate information because they use the revenue market share measure which is misleading for customer. If Oracle has the revenue leading market share then this is true because their solutions are very expensive not because they are the most widely used by enterprises.

5.2 Why SQL Server has the biggest Market Share?

Actually, this question could be answered in a book but the paper will try to answer this very quickly and shortly through representing the major reasons for that.

5.3 SQL Server 2005 BI Solution Advantages

The following is a depiction for the main advantages for SQL Server BI solution as an overall solution not for depicting the detailed features. This may need more detailed technical context that is out of the scope of this paper.

5.3.1 Data Warehouse and BI Phases coverage

SQL server delivers tool for every phase within the data warehouse building life cycle. In addition these tools could work independently and away of any data warehouse. For enterprises that don't have real data warehouse but use OLAP for better decision support, Microsoft provide an easy to use solution with a broad range of tools. Figure 6 depicts those main tools for building a complete BI solution.



Figure 6: The SQL Server BI Tools

Figure 6 shows the set of services and tools that SQL Server provides out of the box. When the customer buys SQL Server 2005 he actually, buys all of these tools. For all other BI Tools vendors each of these tools is sold separately. Some very famous Data

Warehouse and BI tools Vendors (e.g. Ardent Solutions) has built their business on selling just one of these products. At regular scenarios, any data warehouse needs to begin by **Extraction, Transformation, and Loading (ETL)** process. This process takes a place on what is called the staging area that maybe a relational DB. The Microsoft Integrations Services is responsible of accomplishing the ETL task. The Loading is done from the Staging area to the actual data warehouse where all historical resides. Actually the SQL Server Relational Engine is available to take the responsibility of holding huge data warehouses. Later in this paper, some references will be given for this context.. At most cases people needs to navigate through the data in a multidimensional fusion for better understanding about their own business data. That is why a multi dimensional database management system is needed in order to build and manage the data cubes. Actually, this is what Microsoft Analysis Services does. After building the required cubes and mining models, it is time to expose this for users. Microsoft provides the Reporting Service in order to achieve this by the users themselves without the need of bothering, or depending on the IT team all the time. Reporting Services integrates with the Visual Studio.net to build very complicated and powerful reports. Many Management tools are available for developing and customizing your BI tools and managing your Meta Data as well.

5.3.2 Cross integration out of the box

All the previously mentioned tools and servers are natively integrated together. No all only these tools are integrated together but also a broad range of Microsoft Tools and Servers. This includes; Microsoft Office, Microsoft SharePoint Portal Server, Microsoft Active Directory, Microsoft Map Point (for geo-spatial data), and many other servers.

5.3.3 Office integration

Actually the Microsoft Office integrates with the SQL BI solution by a unique way. This includes;

- Integration with Excel through, excel charts, pivot tables, and more.
- Integration with the Windows SharePoint Service for building a BI dashboard.
- Integration with Microsoft Office Business Score Cards for building score cards as well as balanced score cards for better KPI monitoring.

5.3.4 Broad Range of Data Mining Algorithms

Additionally, On Microsoft Analysis Service you can build whatever Data Mining models using the existing available seven algorithms in order to mine into broad range of problematic data. In addition to that, third party ISVs can develop algorithms that snap in seamlessly to the Analysis Services data mining framework. Depending on the data and the goals, different algorithms are preferred, and each algorithm can be used for multiple problems. In fact, this feature of developing new data mining algorithm is a very powerful feature that is not available at most of the data mining tools (**Including Oracle tools**).

5.4 BI Conclusion

The SQL Server BI solution is really far beyond Oracle BI solution

6. References and Case studies

The Following are some hints about some notable references for SQL Serve that should be highlighted here in order to show up its capacity:

6.1 Education and Research References

6.1.1 International References

- **University of South Florida**

The University of South Florida chose the Microsoft platform and SQL Server to build a complete Web-based student information system that extracts information from an IBM mainframe and other database systems throughout the 36,000-student campus, increasing scalability while also lowering total cost of ownership and development costs of the previous deployed solution. The Old available system was previously built on Oracle.

- **The Henslow Research Group**

Until recently, it has been believed that British naturalist Charles Darwin developed his theory of evolution in isolation, basing it on observations he made during his five-year voyage on the

research ship Beagle from 1831 to 1836. However, an international research team based at Cambridge University in the United Kingdom, has used the analytical capabilities of Microsoft SQL Server 2005 to prove that one of the central tenets of Darwin's theory is actually based on work by his mentor, J. S. Henslow. The Henslow Group examined fragile paper-based data collected more than 160 years ago and used the business intelligence (BI) capabilities of SQL Server 2005 to prove the importance of Henslow's work.

6.1.2 Local References

- **Ministry of Education**

The Thanwaya Amma result web based exposure project is one of the hot references and examples that show how much SQL Server is scalable, reliable, and available. During the night of launching the Tnahwaya Amma result for education year, more than 4 million users has hit the website in one night. The Performance was outstanding. This could be validated easily through asking Ministry of Education people.

6.2 Applications Related to the Education Domain

6.2.1 International References

- **Barnes & Noble Giant Bookstores chain (Migrator from Oracle)**

Barnes & Noble is a giant bookstores chain with 40,000 employees, and 4.5 Billion gross in 2005. "We Are currently Processing anywhere between 1 million and 3 million sales transaction records daily, and it take no more than 10 minutes to get the records into the data warehouse. That rate was not possible before" Said Eliot Gardner, the Data Warehouse project manager in Barnes & Noble.

Barnes & and Noble has already a previously implemented OLTP on Oracle, but moved to build there BI solution on SQL Server.

6.2.2 Local References

- **IDSC and Helwan University**

The Information and Decision Support Cabinet for Prime Ministry in Egypt (IDSC) has their newest version of their Library Management system from scratch using SQL Server 2005, and Visual Studio.NET 2005. They faced some problems regarding huge amount of fields that should be added based on DIWI categorization Library standard. Using The available out of the box XML support feature, they designed the books table to have a an XML field that contain an xml document for saving the not commonly used DIWI categorization data. The

system is deployed now on Helwan University Central Library which is one of the Largest (if not being the largest) central library available in Egyptian universities (public and private) today.

Personal Contacts:

- Engineer Medhat Eltokhy, EALIS team leader, IDSC
- Engineer Medhat ElBakry, Applications Development CEO, IDSC.

6.3 Other References

- **Verizon Communications**

Verizon is one of the biggest two mobile as well as ground carrier in North America. They are operating a SQL Server database with a size of 50.7 Billion rows and it is expanding.

- **United States Department of Agriculture (USDA)**

United States Department of agriculture is responsible of about \$1.24 Trillion, or 12.3% of the overall American GDP. USDA is one of the most rigid agricultural ministries all over the world. They have very firm standards for everything. USDA is running a 25 Terabyte Data warehouse on and SQL Server platform.

6.4 Other Local References

- Engineer Taha Shindy, The Information Center Manager, MCIT. :

Information center Manager has built and currently host the Arab League website using Visual Studio.Net and SQL Server.

6.5 SQL Server 2005 Fast Facts

According to study held in 2006 by Winter's Corporation, about the database systems all over the world, the following findings has been produced:

- One of the top three largest OLTP databases in the world, by rows is running on SQL Server.
- Three of the largest OLTP databases in the world by volume are Running on SQL Server
- One of the Top Ten largest data warehouses in the world is running on SQL Server.
- 65 percent of all SQL Server entries are at 2 terabytes or larger