

Magic Quadrant for Data Warehouse Database Management Systems, 2006

Gartner RAS Core Research Note G00138797, Donald Feinberg, Mark A. Beyer, 12 September 2006, R1985 09152007

The data warehouse DBMS market is heating up, with fierce competition from newcomers using open-source software and commodity hardware to increase the pricing pressures on this market. New offerings are also entering from proven vendors, adding further confusion to the landscape.

WHAT YOU NEED TO KNOW

The data warehouse (DW) database management system (DBMS) market continues to show intense competition among the leaders. During the past 18 months, some of the newer entrants have gained small to moderate market share. As IBM, Oracle and Teradata continue to battle for larger-size DWs with increased marketing and new functionality, they are faced with new competition from the appliance vendors, such as DATAlegro and Netezza, and a competitive DBMS from Microsoft. DWs are serving in an increasingly mixed workload capacity in which deep mining analysts, exploring business analysts running less-structured but equally complex queries and fast-running tactical queries all compete for CPU, memory and disk access with differing service-level expectations. Ignore marketing claims and base your decisions on customer references and proof of concepts to ensure that claims made by vendors will hold true in a real-life environment – more specifically, your own environment. Although this is a mature market with the full attention of large vendors seeking to make inroads with scale and innovation, smaller entrants often deliver a more focused, innovative solution.

Market Overview

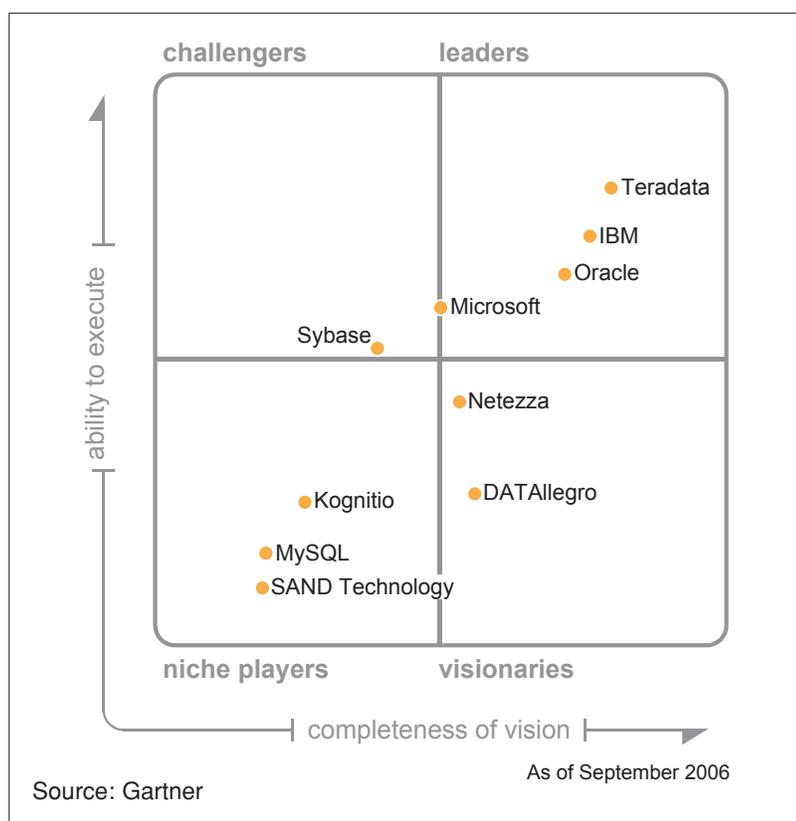
The DW DBMS market has started to focus on business intelligence (BI), with new vendors entering the sector and mature vendors offering new solutions. The lessons learned from Teradata and, more recently, Netezza about the advantages of pre-configured “appliance” solutions are catching the interest

of mature players, such as IBM and Oracle. Even Sun Microsystems is joining in with the announcement of a low-priced DW appliance using the Greenplum massively parallel processing DBMS in the fall of 2006. In addition, the size of the database is becoming less important. In the past, buyers believed that the vendor with the largest database was the leader. Today, DWs of 5TB to 20TB are common. The emergence of issues based on a mix of the more traditional DW workload is created by three distinct workloads:

- Continuous (near-real-time) data loading – similar to an online transaction processing (OLTP) workload
- Large numbers of standard reports ranging in the thousands per day, requiring Structured Query Language (SQL) tuning and index creation

MAGIC QUADRANT

Figure 1. Magic Quadrant for Data Warehouse Database Management Systems, 2006



- An increasing number of true ad hoc query users with a random, unpredictable use of the data

A fourth workload – an increasing level of analytics and BI-oriented functionality in OLTP applications – is creating a new level of use of the DW. The four workload types are creating issues for vendors, more than the actual size of the DW, even manifesting in database sizes less than 1TB. In addition to service-level expectations, the size and duration of “useful” data for each community often differs significantly, forcing every aspect of the DW environment – from input/output (I/O) channel balancing through disk management and into memory and processor allocation – to become involved. During the next three years, mixed workload performance will become the single most important performance issue in data warehousing.

As a direct effect of the mixed workload, with continuous loading and the increase in automated transactions from the functional analytics in OLTP, the transactional DBMSs have an edge that challenges the DW DBMSs (such as Teradata) to stay ahead. Another effect of this growing complexity and size is that organizations are reporting differing estimates of the resources required to support the enterprise DW and accompanying warehouse-dependent data marts. Some of this variation can be attributed to the mixing of resources required to support the physical warehouse (such as storage management, database reorganization and resource balancing) vs. the resources needed to support logical modeling, business process understanding, and optimization and performance enhancements. The results of efforts by DBMS vendors to automate the physical DBMS support functions are appearing as the number of database administrators (DBAs) required for these functions steadily decreases. However, all DWs require optimization techniques. In

some cases, combinations of logical designs create physical approaches in the database, such as automated summary-type tables and different types of physical views.

In other cases, optimization requires the design and deployment of physical tables bulk-loaded from extraction, transformation and loading (ETL) processes in microbatches throughout the day, or even intra-hour. Database systems formerly classified as “general purpose” (such as DB2, Oracle and SQL Server), which once required manually maintained solutions, can be run with fewer administrative and DBA staff thanks to greater automation. At the same time, specialized DW DBMSs (such as Netezza and Teradata) are beginning to require more of these types of solutions as well. As the size of the source system extracted data grows beyond the 5TB to 10TB level, the use of industry standard best practices in DW design becomes extremely important. The need for additional people and storage (for structures used to tune the DW design) should not be seen as a negative, but as part of the process of increasing the value of information. Multiple application types are accessing the DW, and application optimization should not be considered a cost of the DW. For example, high-level summary tables should be considered an application development burden, as opposed to low-level summary tables that can be used by different applications.

Market Definition/Description

The DW DBMS market consists of the DBMS products that deliver the software infrastructure to support a DW of any size. We define a small DW as less than 5TB; a medium DW as 5TB to 20TB; and a large DW as greater than 20TB. For purposes of measuring the size of a DW, we specifically define data as source system extracted data, excluding all

The Magic Quadrant is copyrighted September by Gartner, Inc. and is reused with permission. The Magic Quadrant is a graphical representation of a marketplace at and for a specific time period. It depicts Gartner’s analysis of how certain vendors measure against criteria for that marketplace, as defined by Gartner. Gartner does not endorse any vendor, product or service depicted in the Magic Quadrant, and does not advise technology users to select only those vendors placed in the “Leaders” quadrant. The Magic Quadrant is intended solely as a research tool, and is not meant to be a specific guide to action. Gartner disclaims all warranties, express or implied, with respect to this research, including any warranties of merchantability or fitness for a particular purpose.

© 2006 Gartner, Inc. and/or its Affiliates. All rights reserved. Reproduction and distribution of this publication in any form without prior written permission is forbidden. The information contained herein has been obtained from sources believed to be reliable. Gartner disclaims all warranties as to the accuracy, completeness or adequacy of such information. Gartner shall have no liability for errors, omissions or inadequacies in the information contained herein or for interpretations thereof. The reader assumes sole responsibility for the selection of these materials to achieve its intended results. The opinions expressed herein are subject to change without notice.

DW design-specific structures (such as indexes, cubes, stars and summary tables). Source system extracted data is the actual row/byte count of data that is reconfigured and loaded into the warehouse, without regard to optimization techniques implemented to meet the requirements of your service-level agreement.

A data mart is defined as an application-specific analytic repository of any size, normally with a specific, smaller group of users. An application can be a traditional set of applications (for example, SAP) or a workload-specific application (such as business analysts using a subset of the data for data mining purposes). Because of the wider acceptance and increasing variety of DW-driven applications, workload variations (as described earlier) have become a primary concern. This market includes only those products that are an enterprise-scale, relational DBMS (RDBMS).

A DBMS is a complete software system that supports and manages a logical database(s) in storage. A DBMS is not a set of datasets managed effectively by the operating system. Enterprise RDBMS systems are those systems that, in addition to supporting the relational data model, also support data availability to independent front-end application software, and include mechanisms to isolate work-load requirements and control various parameters of end-user access within a single instance of the data. It is important to note that a DBMS cannot “be” a DW. It is the platform on which a DW (solution/data architecture) is deployed. This market is specific to RDBMS systems that are used to support a production DW.

Inclusion and Exclusion Criteria

Vendors in this market have DBMS software that has been available in production for at least one year. They have generated revenue from at least one production customer and the customer(s) in place has deployed enterprise-scale DWs that integrate data from at least two operational source systems for more than one end-user community (such as separate business lines or differing levels of analytics). Support for these products must be available from a vendor – community-supported open-source software (OSS) products are not included.

Products that include unique file management systems embedded in the front-end tools, or that exclusively support an integrated front-end tool, do not qualify for this market. Systems or products that support an integrated front-end tool, but can also open their DBMS to competing applications, are included if access is achieved via open-access technology, as opposed to custom-built application programming interfaces. Vendors participating in this market must demonstrate their ability to deliver the necessary infrastructure and services to support these objectives. New entrants into this market will only be considered after their product meets these basic criteria, and they are commonly considered by Gartner clients as an option for supporting an enterprise DW.

Added

In this version of the DW DBMS Magic Quadrant, we added several new vendors. SAND Technology, from Canada, has reemerged after several years of silence with its DBMS offering for the DW market, offering highly compressed storage with online and near-online access. A new appliance has emerged, from DATAlegro, which makes use of commodity Intel hardware, Linux and the Ingres OSS DBMS. The longtime U.K. vendor, Whitecross, recently merged with a BI company to form Kognitio. It offers a DW DBMS called WX2. Finally, we have added MySQL, with its OSS DBMS MySQL Version 5.0, because it has reference clients with DWs of various sizes.

Evaluation Criteria

Ability to Execute

Ability to execute is primarily concerned with the ability and maturity of the product and the organization. The product must be able to scale from small to large mixed workloads with little effect on performance. In addition, this scaling should not exponentially increase the resources necessary to manage the product in the DW. Another important criterion is the portability of the product and its ability to run and scale in different operating environments, giving the customer a range of options. The skills of the support and sales organization are also important to the vendor’s ability to deliver the product. These characteristics are critical to the level of satisfaction and success the customer has attained with the

product. Management’s ability to be responsive to market changes and, therefore, the ability of the company to survive through market ups and downs is critical to the long-term survival of the product. Although all criteria are important here, the customer and product criteria are weighted more heavily.

Marketing and sales efforts, although critical to a vendor’s success, are less important, because most organizations are adopting a “show me” attitude and insisting on proof of concept to demonstrate the ability of the product to support the DW. This has elevated the importance of credible customer references with a wide range of workload parameters, such as batch loading to real-time loading, batch reporting to high-end power user ad hoc queries and active data warehousing. Another source of input for the ability to execute criteria is the Gartner Server Evaluation Model, which analyzes the DW workload on a variety of hardware platforms.

Specific Criteria:

- **Product and service** includes the technical attributes of the DBMS. We include scalability, manageability, security, high availability, support of mixed workloads and data loading. These attributes are measured across a variety of database sizes and workloads.
- **Overall viability** includes the corporate aspects of ability to execute, such as personnel, financial stability, R&D investment, and merger and acquisition activity.
- Under **sales execution and pricing**, we examine the price and different pricing models of the DBMS, the ability of the sales force to manage accounts and if the sales team is compensated appropriately in line with the corporate marketing initiatives.
- **Market responsiveness and track record** covers the issue of references (for example, how many, what sizes, what configurations and workload mix). Also included is the ability of the vendor to adapt to market changes and its history of being flexible to market dynamics.
- **Market execution** explores how well the vendor understands and builds its products in response to customers’ needs, in addition to targeting offerings to these needs and to the needs of the market in

general. This criterion includes the completeness of the vendor’s offering as well.

- Customer support and professional services are evaluated as part of the **customer experience** criterion, together with input from customer references as described earlier. Also included is the track record for proof of concepts and customer perceptions of the product, as well as aspects of customer loyalty to a given vendor. This demonstrates customer tolerance of vendor practices and may indicate satisfaction.
- **Operations** cover the alignment of the company’s operations, as well as whether and how they enhance the ability of the company to deliver.

Table 1. Ability to Execute Evaluation Criteria

Evaluation Criteria	Weighting
Product/Service	high
Overall Viability (Business Unit, Financial, Strategy, Organization)	high
Sales Execution/Pricing	standard
Market Responsiveness and Track Record	high
Marketing Execution	standard
Customer Experience	standard
Operations	standard

Source: Gartner

Completeness of Vision

Completeness of vision encompasses the ability of the vendor to understand the functionality necessary to support the DW workload design, the product strategy designed to meet market requirements, and the ability to understand overall market trends and influence or lead the market when necessary. A vendor’s vision is enhanced by its willingness to extend its influence throughout the market by working with independent, third-party application software vendors who deliver DW-driven solutions (such as BI). A successful vendor will be able not only to understand the competitive landscape of DW, but also to shape the future of this field.

A vision leadership role is necessary for the long-term viability of the product and the company. This criterion also includes the company’s ability to communicate this vision to its field organization and,

therefore, to clients and prospects. This must include a solid understanding on the part of the vendor of configuration management relating to the performance of the DBMS. The vendor must be able to develop new ideas for the management of a DW in a mixed workload environment. This leads to many vendors creating a packaged solution to ensure the correct minimum configuration to attain the necessary performance. Although important, we don't believe that the vertical focus in this space is part of the success of the product itself in a DW. It does support the offering and is important for the delivery of the application on the DW.

The geographic reach of the vendor is increasingly important. Data warehousing and BI have become the No. 1 technical issue for CIOs worldwide. A vendor's success increasingly depends on its ability to market and support its DW in a geographically disperse area, using subsidiaries or distributors. This also includes the ability to support clients throughout the world, around the clock, in many languages.

Specific Criteria:

- **Market understanding** covers the ability of the vendor to shape the market and show leadership in the DW DBMS market. We also examine the core competencies of the vendor in the DW DBMS market.
- **Marketing strategy** refers to the vendor's marketing messages and its ability to choose appropriate target markets and third-party software vendor partnerships to enhance the marketability of the product. For example, does the vendor encourage and support independent software vendors in its effort to support the DBMS in native mode?
- An important criterion for vision is the **sales strategy**. This encompasses all the channels and partnerships developed to assist with sales. Further, the ability of sales and marketing management to communicate the message to the field is evaluated within this criterion.
- **Offering (product) strategy** covers the areas of portability and packaging of the products. Vendors must demonstrate a strategy that enables customers to choose what they need to build a complete DW solution.

- The **business model** covers how the vendor's model of a target market combines with product offerings and pricing, and whether it has the ability to produce profits with this model based on the packaging and offerings.
- Specifically for the DW DBMS market, we do not believe that **vertical industry strategy** is a major focus, but it does affect the ability of the vendor to understand its clients. Specific models for the DW, however, belong in a discussion of applications.
- **Innovation** is a major criterion for evaluating the vision of DW DBMS vendors in developing new functionality, spending in R&D, pushing the market in new directions and pushing the envelope in the market to new heights. This also includes the vendor's ability to innovate and develop new functionality in the DBMS, specifically for the DW.
- The worldwide reach of the organization and its **geographic strategy** is evaluated considering its ability to leverage the resources in geographic regions, as well as subsidiaries and partners in other geographies.

Table 2. Completeness of Vision Evaluation Criteria

Evaluation Criteria	Weighting
Market Understanding	high
Marketing Strategy	standard
Sales Strategy	standard
Offering (Product) Strategy	high
Business Model	high
Vertical/Industry Strategy	low
Innovation	high
Geographic Strategy	standard

Source: Gartner

Leaders

The Leaders quadrant for DW DBMS contains those vendors that demonstrate the greatest degree of support for DWs of all sizes (small to large), with large numbers of concurrent users and a high degree of mixed data warehousing workloads. These vendors lead the market in data warehousing by consistently demonstrating customer satisfaction, strong support and professional services, as well as longevity in the DW DBMS market, with strong hardware alliances. Leaders also represent vendors

with the lowest risk for successful DW implementation. Vendors in this quadrant have demonstrated consistent, extensive and durable execution in delivering DW databases at the enterprise level. Additionally, the maturity of this market demands that leaders maintain a strong vision regarding the key points emerging during the past year: mixed workload management for end-user service-level satisfaction and data volume management.

Challengers

This quadrant typically represents vendors with strong offerings for the client base. They have the market presence in the DW DBMS space but have not yet shown or proved their vision or leadership in the market. Challengers generally have a highly capable execution model. Ease of implementation, clarity of message and end-client engagement all contribute to making these vendors successful. Challengers show a wide variety of data warehousing implementations across different sizes of DWs with mixed workloads. Organizations often demonstrate concern regarding vendors' ability to deliver at the enterprise level in cases where growing data volume or high end-user counts are involved. This includes offerings with a weaker marketing message, but products exhibit the potential to move into the Leader's quadrant by demonstrating strong, new client acceptance.

Visionaries

Visionaries are those vendor's that represent a forward-thinking approach to managing the hardware, software and end-user aspects of the DW. Visionaries frequently suffer from a lack of global or even strong regional presence. They also exhibit a smaller market share. New entrants with exceptional technology may appear in this quadrant very early after their general availability release, but more typically, unique or exceptional technology will emerge in this quadrant after several quarters of general availability. The Visionaries quadrant is often populated by new entrants that have new architectures and functionality that is yet unproven in the market. The requirement for production customers and general availability of at least one year indicates they must be more than a startup with a good idea. Vendors must demonstrate customers

in production proving the value of the new functionality and architecture. Frequently, visionaries will drive the leaders toward new concepts and engineering enhancements.

Niche Players

A niche player has low market share or low market appeal. Frequently, a niche player provides an exceptional DW DBMS product, but it is isolated to a specific end-user community or vertical industry. Although the solution itself may be without limitations, the market adoption is limited. This quadrant contains vendors in two categories: smaller vendors with DW DBMS products that lack the customer base or smaller vendors with a DW DBMS that lacks the functionality of leaders. They offer smaller, specialized solutions that are used for specific DW applications depending on the needs of the client. This quadrant also includes new DW DBMS products that lack general customer acceptance or proven functionality to move beyond niche status. This is the starting point for many new entrants.

Vendor Comments

DATAIlegro

DATAIlegro is the newest vendor in the DW DBMS market with its commodity appliance. DATAIlegro balances hardware, software and data volumes, but it does so with commodity processors, disc drives and motherboard/CPU combinations. DATAIlegro also makes use of Linux and the Ingres OSS DBMS. Although it has made modifications to Ingres for DW, these modifications have been cycled back into the public OSS version.

DATAIlegro has additional intellectual property wrapped around the standard software for workload balancing, query parallelization and multidimensional partitioning of the data. The result is a DW appliance priced even lower than Netezza. DATAIlegro has been in business for about three years. It has been delivering its product for less than one year, with approximately three production customers and a growing number of clients in testing phases (resulting in its lower position on the Magic Quadrant). DATAIlegro claims to handle volumes of 40TB to 60TB easily. However, with few production customers, this remains isolated to its internal

testing, customer proof of concepts and the expectations resulting from those tests.

DATALlegro has one technical advantage in that, being based on a full RDBMS with indexing (unlike Netezza), it can more easily pursue a DW architecture that combines operational systems with analytics in a mixed workload. Execution will improve over time, but this is highly dependent on the company's early customer references and ongoing proof of concepts, reportedly accomplished in days. From a vision perspective, the DATALlegro's inclusion of the indexing, load balancing management and software/hardware balancing is the strategy that is being followed by IBM and Teradata. The company is expected to handle mixed workloads between analytics and operational analysis support well. Its deep mining capability is yet to be proven, but is expected to perform as required.

IBM

IBM is rated as a single entrant in the Magic Quadrant, but it has two solutions. The IBM Data Warehouse Edition (DWE) is a software-only solution. IBM's DW appliance solution, the Balanced Configuration Unit (BCU), is a combined server with AIX or Linux, storage, IBM's DWE DBMS and service solution. DWE includes the DB2 DBMS, including database features such as materialized query tables and query parallelization, which are not specifically focused on data warehousing, but help to enhance DW performance. Additionally, the DWE includes additional workload management software, data transformations in the DW, integration of SAS/SPSS supporting data mining and data visualization capabilities, OLAP support for DW modeling with bridges to common BI tools, and logical and physical data partitioning.

The BCU concept emerged from IBM's professional services and client experience as they worked together to optimize DB2 for warehousing. The Unix BCU is DB2 DWE on IBM System p hardware, with AIX pre-configured with memory and storage. The Linux BCU (IBM's newest addition) uses IBM System x hardware and SUSE Linux or, recently announced, Red Hat Linux. By taking control of the hardware configuration, the database is optimized for I/O and disk management. Customers report the Unix BCU

achieves strong results; the Linux BCU is recently released with few references.

Over the years, IBM has had good success with DB2 as a DW (especially on z/OS and AIX) and with large warehouses using the DWE (now in its third release) on System p with AIX. IBM still needs additional proof of the manageability of the mixed workload with the BCU to become an effective competitor with proven vendors (such as Teradata). Although IBM's DB2 on z/OS and AIX are priced competitively with Teradata, they will need to show (primarily using the Linux BCU) that they can compete on price and performance with DATALlegro and Netezza.

If you are evaluating DW DBMSs, then consider the BCU as a shortlist solution in your DW DBMS selection process. IBM is a solid solution when the mixed workload for operational analytics and deep mining are combined on data volumes under 20TB of source system extracted data. Similar to most vendors in warehouses larger than 20TB, customers report the use of additional resources for tuning and management (as described earlier in the Market Definition section).

Kognitio

Kognitio appears for the first time under this name. It is the merging of two companies, Kognitio (a consultancy in the BI space) and Whitecross (a 25-year-old company in the data warehousing appliance market that moved to a DBMS software solution about 10 years ago). The result is what they call a BI DBMS – WX2 – a revised version of the original Whitecross DBMS (x86 based).

Kognitio is based in the U.K. and has amassed a number of large clients, mostly in Europe. Compared with the large DW DBMS vendors, its customer base is small (under 100) and most are using a solution managed by and many times outsourced to Kognitio.

Kognitio's DBMS runs on Linux and blade technology for a highly parallel, scalable solution to many terabytes of data. Its DBMS is a fully functional RDBMS with SQL interfaces through standard Open Database Connectivity (ODBC) drivers. It supports all the BI tools through this interface. The company's

solutions require a high degree of consultancy services for the management of the DW. As a small company, relatively local to the European market, it remains in the Niche Player quadrant. Its solution is effective for warehouse applications that require the searching of very large tables of detail data for analytical uses where users prefer to outsource the management of their DW. Many customers are end-user departments with a data mart solution. Although Kognitio also offers a DBMS software solution, we believe that it is a possible choice for organizations that prefer to outsource their DW solution.

Microsoft

Until recently, Microsoft SQL Server was relegated to departmental-sized applications. During the past several years, using SQL Server 2000, organizations have grown DWs to several terabytes in sizes, although the average-sized SQL Server DW is more in the 100s of gigabytes. With the release in November 2005 of SQL Server 2005, Microsoft is moving into the mainstream DW DBMS market. Since its release, Microsoft has seen double-digit revenue growth and leads the Windows RDBMS revenue market with 46.8% of the market and 16.6% growth from 2004 through 2005. The company has many references greater than 1TB of source-system extracted data. It also has added functionality, such as indexed views, the partitioning of data and new management capabilities specifically to support the DW. With this new functionality in SQL Server 2005, Microsoft is competing with other DBMS vendors for small to medium DWs.

Most organizations are reluctant to move quickly to large DWs with SQL Server until references increase to demonstrate performance in larger DWs with mixed workloads. SQL Server also suffers from the lack of solid server clustering solutions for scalability and high availability where IBM and Oracle have well-proven technology. Finally, the restriction of a Windows-only solution has a negative connotation. Many potential users desire the flexibility of multiple operating system support (portability). However, the ability of Microsoft to deliver DBMS solutions for the enterprise is no longer in question. With the addition of integration, reporting and analytical services, and the extensions in Office 2007, Microsoft will have an offering for the enterprise DW workload. Microsoft is

rated between the Challenger and Leader quadrants, because it needs to prove itself as a DW solution. To move into the Leader's quadrant during the next few years, Microsoft will need to demonstrate the scalability and performance of the leaders in a mixed workload DW.

MySQL

MySQL has rapidly matured as a DBMS engine using the OSS model. It is being used for various types of applications, even mission-critical applications in some technically advanced organizations. Some would argue that MySQL is not an OSS product, because it is owned by the company with the same name. We believe this is a strength, because it is responsible for the code, supplying customer support on a subscription basis while using the OSS license model.

With the new version 5.0 released in the fall of 2005, the necessary functionality for OLTP applications has been added, but it still lacks much of the functionality required to be considered for DW applications (such as materialized views and range partitioning). Most of the installations using MySQL are implementing smaller (under 1TB) DWs with low numbers of users. However, during the next few years, we believe this DBMS will add additional DW functionality and become more widely used and proven in larger, more-critical DWs. A general lack of support by tool vendors exists, as well as a lack of internal management support for the DW, including the workload management software so important to a DWs success. MySQL is a possible choice for simple, small DWs to begin experimenting with OSS in IT organizations that are willing to accept the added risk of a new DBMS and the lack of a skill base in implementing, using and tuning MySQL for data warehousing.

Netezza

Netezza was the first of the new appliance vendors (with products in production for about three years) to appear as a niche player with potential. It is a small, growing company with about 75 customers, primarily in North America and Europe. Netezza balances hardware, software and data volumes, and does so with commodity processors and disc drives using a

proprietary motherboard. The DBMS is proprietary using the Linux Operating System. The result is an appliance with a pricing strategy that continues to have a disruptive effect on the DW DBMS market by driving other vendors to lower price points.

Netezza serves in organizations as an enterprise DW and as an existing DW performance accelerator, frequently winning proof of concepts against many of the other DBMS vendors. Netezza supports DWs of the 5TB to 20TB size (recently it released a version for 50TB and 100TB). Netezza references report clear simplicity of installation and management of the DW and data mart environment. Organizations normally report having less than 20 power users on a system, and Netezza recently found a stronger role with deep analytic-style data mining. As the number of power users grows – therefore increasing the mixed workload – there will be performance issues.

Frequently, an established warehouse suffers from performance issues because of the mixed workload issue of statistical analysts and data miners running long, resource-intensive queries and consuming warehouse resources. In these cases, the introduction of a low-cost “accelerator,” such as Netezza, provides a data mart as a mining-dedicated platform with significant cost reduction by removing the need for expensive upgrades to the established DW solution. Netezza pursues more than one market strategy in that it simultaneously seeks to serve as the enterprise DW and to accelerate established DW implementations on competitive platforms at an advantageous price point – giving customers an alternative to increasing commitment to the established vendor or to staffing to support it.

Netezza also provides good support for the casual DW user who usually accesses reports. The company does not currently have data indexing (which is planned in the next software release). Although this is not necessary in a data mining environment, it does have an effect on the mixed workload capabilities. Netezza has the potential to become a powerful data mining adjunct. With additional improvements in workload balancing and indexing, it could move higher in the execution scale. Netezza can be considered as an accelerator or data mart for a dedicated user group, but significant stress testing needs to be considered when using it as an

enterprise DW solution. Expect to see competition against some of the less-expensive solutions, such as DATAlegro, and proven vendors, such as IBM.

Oracle

Oracle is the most widely deployed DBMS as demonstrated by multiple studies implying that Oracle market share for DW deployments is predominant. These studies show organizations already have Oracle on many platforms, and DWs are frequently pursued on existing platforms with established DBMSs. Oracle warehouses operating in the 20TB-and-greater range pursue a variety of optimization strategies that frequently include materialized views, ETL direct-loaded summary and pre-joined tables, and the use of star schemas.

Warehouses frequently occupy a larger storage environment (even after accounting for time variance and data versions) when compared with the source data extracted to load the warehouse. Much of the larger size in Oracle warehouses is attributed to additional storage to support chosen optimization strategies. Oracle warehouses exhibit an increased level of manual administration and data architecture/design needs, with sizes ranging from 1.5TB and 5TB of source-extracted data. DBA count for Oracle systems is seeing a reduction, but conflicting reports come from customers indicating that warehouses with more than 10TB of source data report large DBA commitments. The key to a large Oracle warehouse is creating a physical optimization layer in line with traditional best practices. Such optimization layers are deployed by building data marts centered on a star schema approach (in the same database instance or otherwise) or using materialized views. Many of the larger DWs are focused on less-complex workloads, such as static reports.

Oracle stand-out features include excellent security at the database level. Hardware and storage optimization with remote access concentrator (RAC) – released with 9i – and automated storage management (ASM; released with Oracle Database 10g) are not specifically focused on DW performance, but enhance DW delivery. Oracle offers specific DW capabilities that include data architecture optimization with materialized views, Oracle

Warehouse Builder for ETL and well-proven best practices from the field.

A wide user base from which to draw expertise exists, and Oracle supports all types of DW deployments. Oracle is also highly effective in operational analytics due to its transactional roots and analytics extensions to SQL. The company is a leader, but some issues include the management of mixed workload and high-volume scalability. Oracle is focused on DW and continues to spend heavily on R&D, specifically for the DW – a trend we expect to continue. Oracle continues to struggle with issues of account management, licensing and pricing issues, and potential alienation of partners because of competitive issues.

SAND Technology

SAND Technology is a small vendor with a vertical or column-vectored DBMS. It has a small customer base (about 50 customers) with good references for the SAND/DNA Analytics DBMS. It has the same issues as Sybase IQ, with result sets that return large numbers of columns. For this reason, it positions SAND/DNA Analytics as an analytical DW engine. As with all column-vectored DBMSs, the company has a strong compression ratio, making this a good solution for archival DWs where the data is not often accessed but needs to be online and not archived offline. SAND is fully SQL compliant and has a good ODBC/Java Database Connectivity (JDBC) interface, enabling it to work with a complete line of BI tools.

As with smaller vendors, SAND has limited third-party alliances to tune tools for specific use with its engine. The company recently announced the SAND/DNA Access product, which offers significantly greater compression and full SQL search capabilities against read-only data that can be versioned for change tracking. This product has provided SAND with one important third-party alliance – SAP. It is the only DW DBMS to have such an alliance. SAND is now certified as a mirror online archive for less-used data in the SAP Business Information Warehouse (BW). This enables the user to move data (InfoCubes, ODS objects and persistent storage area data) to a SAND highly compressed database with transparent access to the data. Therefore, everyday

queries against current SAP InfoCubes will be operating on a smaller amount of data, increasing the performance of the SAP BW and only accessing the older data when necessary.

SAND is focused on SAP – a strategy that will afford it some good growth. With the help of SAP, SAND can be considered a near-line archive in situations where the enterprise DW and, especially, the SAP BW have grown to a large size, or for data marts where typical access requires operating on only a few columns in a large schema with many attributes.

Sybase

Sybase is a vertical or column-vectored DBMS. In a column-vectored database, the data is stored by column instead of row – essentially creating an index on every column with the added advantage of highly compressing the data with little effect on performance – with a three-to-five-times compression ratio. Queries against a column-vectored system are very fast when there are few columns. As the number of columns increases, the database must reassemble data rows, exhibiting slower performance compared with a large query in a row-based system.

Sybase does commit annual budgets in R&D to IQ, and it continues to improve the product. However, Sybase suffers from market identity issues relative to its many products – from ASE through IQ to Mobile Computing solutions – and how the market perceives its commitment to the IQ product. Organizations have discussed some concern over Sybase having two different DBMSs – one for OLTP (ASE) and one for the DW (IQ). Do not view this any differently than using another specialized DW DBMS tuned for data warehousing (such as Netezza or Teradata). Sybase IQ is a good solution for enterprise DWs focused on reporting and business analysis, and for optimized analytical data marts. Proof of this has been the growth of IQ, gaining new clients for Sybase as IQ is used as an analytical data mart in the overall DW architecture.

Data mart does not imply small DW sizes, and several Sybase references prove this with large data mart sizes (one with more than 1 billion rows) in the 10TBs and higher of source system extracted data.

The principle of column vectoring has direct applicability to data mart optimization. IQ is especially good when performing analytics on a small number of columns and large number of rows. Organizations have reported internal benchmarks where IQ outperformed other DBMSs by factors of 100 to one in these situations. Because Sybase has some of the best mobile solutions, intermittently connected BI solutions will also benefit from the Sybase IQ DBMS.

Teradata

Teradata has concentrated exclusively on data warehousing solutions for more than 25 years. It ranks third in Unix revenue, with growth of 6.8% from 2004 to 2005 (compared with overall -1.0% for the Unix DBMS market). Teradata has always been based on the principle that data volume, data storage, I/O channels, memory and CPU processing had to be balanced against the workload expectations inherent in analytic processing. This is the concept now pursued by IBM with its BCU technology, Netezza, DATAlegro and, more recently, Oracle (with its validated configurations).

Another strength of Teradata is its complete solution approach. It rounds out the physical warehouse with strong consulting services and the widest variety of vertical solutions. It also has one of the strongest customer reference bases in the industry. Teradata is reported by numerous organizations as handling 40TB of data routinely, with many organizations reporting much higher volumes (some reaching more than 100TB). Teradata uses hardware units (nodes) and virtual work units (AMPs) to break its environment into task engines that enable the implementation of a highly parallel query capability. In addition to the parallelism, Teradata includes workload management features, such as the Priority Schedule Manager, to prioritize queries by user group, individual or even query class in assigning which job will receive resource priority.

Historically, Teradata was the most expensive solution relative to initial cost, maintaining the position that users would recover these higher initial

costs during the life cycle of the system as lower administration and maintenance costs played out. However, Teradata has acquiesced to the pricing pressure brought on by Netezza, along with the greater capabilities of IBM and Oracle in the DW. The company has published pricing and offers clients discounted pricing when competing with other DBMS vendors. Teradata has also released a version of its solution using Microsoft Windows (although it has few users) and, in April 2006, released a Linux version – which has a couple of initial customers. This Linux version is expected to be the basis of offerings that will be at lower price points, competing more effectively with the other appliances and DBMSs.

Teradata's remaining issues include the use of only two types of storage devices in its pre-built configuration. This does not sit well with the storage administrators, because proprietary interconnect between storage and processors, and some data modeling CASE tools, do not interoperate with the Teradata platform – sometimes complicating implementation and system documentation. Teradata remains a strong candidate for all size DWs, but it generally achieves its best cost-effectiveness above 10TB. In the past year, Teradata has introduced pricing changes that enable it to remain price competitive in smaller warehouses.

Acronym Key and Glossary Terms

ASM	automated storage management
BCU	Balanced Configuration Unit
BI	business intelligence
BW	Business Information Warehouse
DBA	database administrator
DBMS	database management system
DW	data warehouse
DWE	Data Warehouse Edition
ETL	extraction, transformation and loading
I/O	input/output
JDBC	Java Database Connectivity
ODBC	Open Database Connectivity
OLTP	online transaction processing
OSS	open-source software
RAC	real application clusters
RDBMS	relational database management system
SQL	Structured Query Language

Evaluation Criteria Definitions

Ability to Execute

Product/Service: Core goods and services offered by the vendor that compete in/serve the defined market. This includes current product/service capabilities, quality, feature sets, skills, etc., whether offered natively or through OEM agreements/partnerships as defined in the market definition and detailed in the subcriteria.

Overall Viability (Business Unit, Financial, Strategy, Organization): Viability includes an assessment of the overall organization's financial health, the financial and practical success of the business unit, and the likelihood of the individual business unit to continue investing in the product, to continue offering the product and to advance the state of the art within the organization's portfolio of products.

Sales Execution/Pricing: The vendor's capabilities in all pre-sales activities and the structure that supports them. This includes deal management, pricing and negotiation, pre-sales support and the overall effectiveness of the sales channel.

Market Responsiveness and Track Record: Ability to respond, change direction, be flexible and achieve competitive success as opportunities develop, competitors act, customer needs evolve and market dynamics change. This criterion also considers the vendor's history of responsiveness.

Marketing Execution: The clarity, quality, creativity and efficacy of programs designed to deliver the organization's message in order to influence the market, promote the brand and business, increase awareness of the products, and establish a positive identification with the product/brand and organization in the minds of buyers. This "mind share" can be driven by a combination of publicity, promotional, thought leadership, word-of-mouth and sales activities.

Customer Experience: Relationships, products and services/programs that enable clients to be successful with the products evaluated. Specifically, this includes the ways customers receive technical support or account support. This can also include ancillary tools, customer support programs (and the quality thereof), availability of user groups, service-level agreements, etc.

Operations: The ability of the organization to meet its goals and commitments. Factors include the quality of the organizational structure including skills, experiences, programs, systems and other vehicles that enable the organization to operate effectively and efficiently on an ongoing basis.

Completeness of Vision

Market Understanding: Ability of the vendor to understand buyers' wants and needs and to translate those into products and services. Vendors that show the highest degree of vision listen and understand buyers' wants and needs, and can shape or enhance those with their added vision.

Marketing Strategy: A clear, differentiated set of messages consistently communicated throughout the organization and externalized through the Web site, advertising, customer programs and positioning statements.

Sales Strategy: The strategy for selling product that uses the appropriate network of direct and indirect sales, marketing, service and communication affiliates that extend the scope and depth of market reach, skills, expertise, technologies, services and the customer base.

Offering (Product) Strategy: The vendor's approach to product development and delivery that emphasizes differentiation, functionality, methodology and feature set as they map to current and future requirements.

Business Model: The soundness and logic of the vendor's underlying business proposition.

Vertical/Industry Strategy: The vendor's strategy to direct resources, skills and offerings to meet the specific needs of individual market segments, including verticals.

Innovation: Direct, related, complementary and synergistic layouts of resources, expertise or capital for investment, consolidation, defensive or pre-emptive purposes.

Geographic Strategy: The vendor's strategy to direct resources, skills and offerings to meet the specific needs of geographies outside the "home" or native geography, either directly or through partners, channels and subsidiaries as appropriate for that geography and market.