

## INDUSTRY DEVELOPMENTS AND MODELS

---

### Business Analytics Appliances Are Here to Stay

---

Dan Vesset

---

#### IDC OPINION

---

IDC believes that adoption of data warehouse (DW) and business intelligence (BI) appliances is not just a passing trend based on limited anecdotal evidence. Early quantitative surveys lend credence to strong interest in, awareness of, and adoption of DW appliances. For the foreseeable future, we expect DW and BI appliances to be deployed alongside traditional DW solutions at large organizations and in some cases as new primary DW solutions, especially at midsize organizations that may not have the resources for traditional DW solution acquisition, deployment, and maintenance. DW and BI appliances are increasingly making an impact in the market as more end users evaluate, adopt, and deploy these solutions in support of their business analytics needs. These appliances are fulfilling real, existing needs of a sizable market segment that is looking for:

- Lower total cost of ownership
  - Preconfigured solutions with lower maintenance requirements
  - Highly scalable business analytics platforms to address growing data volumes
-



## IN THIS STUDY

In this study, IDC examines the market forces, demand drivers, and suppliers that have converged to usher in a new cycle of innovative solutions in the business analytics market. The study focuses on IT vendors that supply appliances for business analytics and classifies them according to their primary offering and market position. In addition, the impact of this emerging group of vendors on established data warehousing and other business analytics companies is discussed.

## SITUATION OVERVIEW

### **Introduction: Tremors in the Data Warehousing Market**

In 2003, IDC published a case study about Netezza as a revolutionary data warehousing appliance (see *Netezza: A Business Analytics Revolution in the Making* (IDC #29323, May 2003). Some questioned whether this description was accurate given that data management appliances or purpose-built IT solutions for data warehousing were in existence or at least attempted before Netezza.

Dedicated database machines, as they were called, are indeed not new. Several attempts were made to develop them over 20 years ago. At that time, the impediments to successful commercialization were lack of low cost, standardized hardware, and algorithms for parallel processing. In 1983, there was a much cited paper on this subject by David DeWitt and Haran Boral titled *Database Machines: An Idea Whose Time Has Passed? — A Critique of the Future of Database Machines*. However, in 1995, referring to DeWitt's paper, Kjell Bratbergsengen of the Norwegian University of Science and Technology wrote in his *Parallel Database Machines*, "The paper [DeWitt's] 'proved' that it was not possible to get performance improvements beyond 4 nodes in a parallel system for doing algebra operations. Luckily, both David himself and others, later — proved this to be wrong." It took another several years before the market began to see commercialized solutions as data warehouse appliances. Although not a new concept, are data warehouse appliances still revolutionary and therefore disruptive to the broader DW industry?

Merriam-Webster's dictionary defines "revolution" as being derived from the Latin *revolvere* or to revolve — the period made by the regular succession of a measure of time — but also as a sudden, radical, or complete change and as a fundamental change in the way of thinking about something. Based on the combination of these definitions, the emerging market for data warehouse and business intelligence appliances certainly qualifies as revolutionary.

Although current solutions reference research that was conducted years ago, the latest patents and marketing methods demonstrate that vendors in this market are also displaying renewed innovation. The latest technical innovation in software is coming from new methods of harnessing the power of parallel processing — in hardware from the ability to leverage commodity products (a result of design and manufacturing process innovations that have led to commoditization), in tightly combining software and hardware in new ways, and in marketing by pinpointing specific needs of the right niche for these products.

Although current solutions reference research that was conducted years ago, the latest patents and marketing methods demonstrate that vendors in this market are also displaying renewed innovation.

## Market Definitions

Before the latest market trends surrounding business analytics appliances are presented, it is important to lay out the definitions of technologies discussed in this study. At the highest level is the all-encompassing market of business analytics.

### ***Business Analytics Software***

The business analytics software market comprises tools and applications for tracking, storing, analyzing, modeling, and presenting data in support of automating decision making and reporting processes. The tools segment of the business analytics market is further composed of data warehouse generation, data warehouse management, and business intelligence (query, reporting, analysis, and advanced analytics) tools as well as spatial information management. The applications segment in turn is composed of customer, financial, and business performance management as well as supply chain and other operational analytic applications.

The business analytics software market comprises tools and applications for tracking, storing, analyzing, modeling, and presenting data in support of automating decision making and reporting processes.

Business analytics software is deployed on a variety of hardware platforms, combining server and storage technologies from different IT vendors.

### ***New Appliance Solutions***

This study focuses on solutions that fall into the data warehouse management segment of the broader business analytics market. Table 1 lists the vendors highlighted in this study in the first column and identifies three vendor focus areas in the second column:

- ☒ **"Complete stack" data warehouse appliances.** These solutions combine software and hardware. There is no uniformity about what is included in software or hardware. For example, two of the vendors include OS and RDBMS software and server and storage hardware. Another vendor doesn't require any database software in one version of its appliance.
- ☒ **Virtual data warehouse appliances.** These solutions provide only software but are deployed on clusters of commodity hardware. These vendors are not prepackaging the hardware with their software; instead, they are leaving this to their clients. Nevertheless, the software component of these vendors are very similar to those marketed by complete stack DW appliance vendors.
- ☒ **Business intelligence (BI) appliances.** These solutions combine software and hardware. Unlike the first two categories, BI appliances are focused not on the data warehouse platform but on the front-end BI software packaged with data integration and management components.

Regardless of the solution, the value proposition of all appliances is based on features and functions described in this study.

**TABLE 1****Data Warehouse Appliance Vendors**

Vendor	Type	Year Founded	First Appliance Product Launch	Product Name	Patents/Innovation
Netezza	Appliance	2000	2002	Netezza Performance Server (NPS)	Intelligent query streaming, combining massively parallel processing (MPP) and symmetric multiprocessing (SMP) techniques
Datallegro	Appliance	2003	2Q05	P- and C-series Datallegro Appliance	I/O optimization, use of MPP and a proprietary encryption technology
Calpont	Appliance	2000	1H05	Calpont Database Appliance	Methods for processing, accessing, and storing data based on network routing technology
Greenplum	Virtual appliance	2003	2Q05	Bizgres MPP	MPP used on clusters of commodity servers
Kognitio	Virtual appliance	2005	1992 (first implementation of WhiteCross database)	WX2	MPP with proprietary, high-performance WX2 database for BI
PARACCEL	Virtual appliance	2006	Expected end of 2006	PARACCEL Accelerator	Grid warehouse architecture with column-based memory mirror
SAP	Virtual appliance	1972	4Q05	NetWeaver BI Accelerator	BI data processing utilizing SAP's search technology
Celequest	BI appliance	2002	2Q06	LAVA	First BI appliance based on BAM and dashboarding software
IBM	Appliance		2005	IBM Data Warehousing Balanced Configuration Unit (BCU)	Combines IBM's full-featured DB2 Data Warehouse Edition software with hardware and add-on services

Note: See the Complete Stack Data Warehouse Appliances and Virtual Data Warehouse Appliances sections for further descriptions of these vendors.

Source: IDC, 2006

## **Appliances Value Proposition and Selection Criteria**

There is no single criteria by which end users are selecting DW/BI appliances. It is not simply the amount of data, processing speed, solution cost, or ability to integrate the appliance into the existing business analytics architecture in isolation. Instead, end users are basing their decisions on a combination of all these variables. Although the three types of solutions differ, their value propositions are remarkably similar. These characteristics are not shared universally, but many of them are present in all the technologies discussed in this study. They include high performance and scalability as well as simplification and usability, which are described in the sections that follow.

### ***High Performance and Scalability***

Performance and scalability metrics must always be evaluated relative to several variables, such as price, data volume, number of users, and specific hardware components. Discussions with end-use organizations point to the fact that DW/BI appliances provide strong performance characteristics at a given price point. In most cases, the price/performance characteristics of these appliances when taken in isolation from other IT purchase considerations are more attractive than those from established database software vendors and associated separate hardware costs. High performance is based on functions such as:

- ☒ Parallel processing
- ☒ Fast data loading
- ☒ Deployment of innovative I/O management techniques
- ☒ Purpose-built solutions for analytics (not for mixed use data processing [i.e., analytics and transaction processing])

### ***Simplification and Usability***

In researching end-user reactions to DW appliances, IDC found the following characteristics as most attractive:

- ☒ **Minimal maintenance requirements by database administrators.** In some cases where RDBMSs are used, maintenance of indexes is not required, resulting not only in fewer tuning requirements but also in lower data storage requirements because the data overhead due to indexes is not present.
- ☒ **Rapid deployment.** A decreased level of integration with existing enterprise information management architectures has made appliances easier and faster to deploy than traditional DW and BI tools. Also, with complete stack DW appliances, the job of integrating software with hardware has already been made. Partnerships among DW appliances and BI software vendors have proven that minimal if any schema changes are required to existing BI tools to "point" them at the DW appliances as the source system.

DW/BI appliances provide strong performance characteristics at a given price point.

- ☒ **Rapid prototyping (e.g., disposable marts).** Data warehouse appliances are sometimes used to pull data in and run ad hoc queries and then "flush" out the data — a use pattern very different from traditional data warehouses.
- ☒ **Open source technologies.** Data warehouse appliance vendors specifically utilize both Linux for OS and open source RDBMSs as their embedded databases, contributing to lower costs for the solutions. In some cases, these open source technologies are modified or enhanced resulting in proprietary solutions. However, the use of these technologies nevertheless enables appliance vendors to decrease product costs. Their goal is not to create "open" modifiable technology as is the case with truly open source projects.
- ☒ **Single source for customer problem resolution.** This issue is raised by all end users and is something that established DW vendors are not fully addressing. The finger-pointing that often occurs among software, server, and storage vendors whose technology is used in a DW environment is something that the complete stack DW appliance vendors are able to eliminate.

The finger-pointing that often occurs among software, server, and storage vendors whose technology is used in a DW environment is something that the complete stack DW appliance vendors are able to eliminate.

Besides the similarities in value propositions among the appliance vendors, it is not surprising that even the corporate vision statements and marketing messages from most of these appliance vendors have similar tones. Consider the following:

- ☒ "The Power To Question Everything" and "Answers at the Speed of Thought" (Netezza)
- ☒ "Data at the Speed of Business" (Datallegro)
- ☒ "Redefining the Performance Game" (Calpont)
- ☒ "PostgreSQL for High-Performance Data Warehousing" (Greenplum)
- ☒ "Competitive Advantage from Data" (Kognitio)

We will not discuss the creative merits of these statements, but their overall thrust is clear.

---

## **Market Drivers: Why Now?**

All except one of the vendors have launched their appliance products since 2000 to address the market needs and challenges discussed in the sections that follow.

### ***Growth in Data Volumes***

Growth in data volumes available for analysis in support of decision making is due to:

- ☒ Standardization on enterprise application in the 1990s
- ☒ Growth in ecommerce and resultant availability of large volumes of click-stream data
- ☒ Emerging deployments of RFID-based solutions

- ☒ Compliance strategies with expansive data retention policies
- ☒ Increasing use of external data sources including GIS, demographics, weather, and so forth for various analytic needs

Recent IDC research supporting this trend includes the following findings:

- ☒ 23% of organizations expect their DW to at least double in size over the next three years. That's up from 18% in 2004.
- ☒ 45% of the respondents from a recent survey said they expect to have to upgrade their data warehousing software to benefit from RFID. 46% of the respondents expect their data warehouses to grow at least 25% over the initial two years when RFID initiatives will be launched in their organizations.
- ☒ About 12% of organizations indicate that their data warehouse is loaded using continuous feeds rather than batch loads. When asked about expectations for the next 12 months, that ratio increases to 20%.
- ☒ End users continued to demand access to more data. 23% of organizations say that BI solutions could better meet their organizations' main objectives if there were more and better information available to end users.

23% of organizations expect their DW to at least double in size over the next three years. That's up from 18% in 2004.

### ***Growth in Business Analytics Users***

In addition to more data, there is a trend of increasing populations of internal employees and external stakeholders who have access to business intelligence tools. With the increase in user populations, the DW and BI offerings must be able to extend support to both traditional analysts and power users that represent about 15–20% of a typical large organization and all other knowledge workers. In a 2005 survey:

- ☒ 70% of organizations indicated that they are planning to increase the number of internal users of BI tools over the next 12 months.
- ☒ Organizations indicated that they will be providing direct BI tool access to suppliers (33%), to customers (40%), and to government agencies (20%).

The number 1 reason (41% of respondents to a 2004 IDC survey conducted with *InformationWeek*) for purchasing any new applications software is the need to improve access to information. Currently, however, only 15% of organizations state that they are confident that the majority of reports/dashboards developed by their organizations deliver relevant data to the right person at the right time.

### ***Ad Hoc Query and Analysis***

There is a need for more ad hoc access to data by different end-user groups. In more detail:

- ☒ According to an IDC survey conducted with *InfoWorld*, the two BI software segments that are expected to attract most end-user investments in the short term are data mining (44% of respondents) and ad hoc query (42% of respondents).

The two BI software segments that are expected to attract most end-user investments in the short term are data mining and ad hoc query.

- ☒ The top impediments to broader adoption of BI include, among others, more ad hoc query and analysis functionality and a wider selection of end-user interfaces for different user groups.
- ☒ Supporting potentially complex ad hoc queries seems to be more of a challenge than dealing with the user concurrency issue:
  - ☐ In terms of specific IT challenges, twice as many organizations stated that data scalability is currently a bigger challenge (20% of respondents) than user scalability (10% of respondents).

### ***IT Resource Constraints***

Resource constraints of IT departments in addressing the ever-growing demand of end users for BI is another issue facing the market. In more detail:

- ☒ According to survey conducted by IDC and *InfoWorld*, the biggest challenges facing organizations in BI/DW projects are lack of budget (50% of respondents) and time required to implement the project (40%). These two challenges are followed by business rules analysis/gathering user requirements (30%) and education and training of end users (29%).
- ☒ 41% of organizations in a 2005 IDC survey conducted with *InformationWeek* on data integration said that one of the three biggest challenges their company faces in information integration projects is lack of internal expertise.
- ☒ With respect to data warehousing or other business analytics projects, the two primary issues facing IT departments are ongoing maintenance costs and deployment speed of both initial implementation and subsequent upgrades and enhancements. In particular:
  - ☐ Faster implementation of requested changes by end users is the second-highest factor influencing wider BI adoption.
  - ☐ Maintenance overhead is an especially tough issue and something that the new appliance vendors have actively addressed. A typical maintenance task such as index maintenance (planning, rebuilding), storage management (disk stripping, partitioning, defragmentation), and tuning (analysis of logs, cache sizing) adds a significant amount of overhead to any project. Solutions that can decrease or eliminate typical data warehousing maintenance overhead processes have the opportunity to address a real market need.
  - ☐ The side benefit of addressing the maintenance issue is the ability to faster address end-user requirements for new functionality, new data sources, and new calculations.

Solutions that can decrease or eliminate typical data warehousing maintenance overhead processes have the opportunity to address a real market need.

## **Competitive Forces**

### **Strategic Importance**

DW and BI are increasingly becoming more critical to organizations. Recent survey findings show that when asked about a wide range of IT investment priorities, organizations regard their BI investments and projects as increasing in priority. For example:

- For the short term (12 months), 56% indicated BI as a critical or high priority.
- For the intermediate term (3–5 years), 65% indicated BI as a critical or high priority.
- For the short and intermediate terms, only 16% indicated BI was a low priority or no priority.

### **Operational Business Intelligence**

BI and DW efforts are becoming more operational, meaning that they are no longer nice-to-have solutions that are disconnected from core operational applications. For example:

- 43% of organizations state that to avoid severely impacting business operations, their BI solutions can't be down more than a few hours. For 26%, this time frame can't exceed one day, while only 8% claim that the BI systems being down will not have a severe impact on business operations (source: IDC's May 2005 *BI Survey*).

### **Technology Legacy**

- IT departments' familiarity with RDBMS persists. Databases were originally developed for OLTP and even if they were to be used for reporting, the expectation was that only a few users would be running queries for reports. Over time, the performance of RDBMSs has improved, but the underlying fundamentals haven't changed. Although traditional RDBMS vendors have databases that are deployed as multiuse (transaction and analytic processing) solutions, the adoption of such multiuse solutions is limited.
- In the mid-1990s, RDBMS vendors took their first steps to incorporate OLAP and data mining engines into their databases. These now include Microsoft SQL Server Analysis Services, Oracle 10g OLAP and Data Mining, Teradata Data Miner, and IBM Cubeviews. However, traditional OLAP engines still can't handle very large data volumes for true ad hoc analysis at attractive price points.

In response to these trends, organizations are increasingly open to exploring alternative DW and BI solutions. The survey results discussed in the Adoption of DW/BI Appliances section give first indications of the real demand for and awareness of DW appliances.

Organizations are increasingly open to exploring alternative DW and BI solutions. Survey results give first indications of the real demand for and awareness of DW appliances.

## Adoption of Data Warehouse/Business Intelligence Appliances

Adoption of DW and BI appliances is not just a passing trend based on limited anecdotal evidence. Early quantitative surveys lend credence to strong interest in, awareness of, and adoption of DW appliances. 2005 survey results are discussed in the sections that follow.

### ***Interest and Awareness***

- ☒ Survey data suggests healthy interest in and awareness of DW appliances (in fact surprisingly strong indicators for such an immature market with so few small players). In more detail, 32% of organizations indicated that they would be interested in evaluating a data warehouse appliance (source: IDC and *DM Review*, July/August 2005, n = 315, United States only; respondents were provided with the definition of "full stack" appliances). No discernable differences existed among companies of various sizes, industries, or titles of respondents.
- ☒ 10% of IT purchasing decision makers indicated that they are aware of Netezza (source: IDC and *InfoWorld's* March 2005 *BI Survey*, n = 286, United States only; other appliance vendors were not identified by name).
- ☒ 1% indicated that they would be likely to purchase Netezza in the next 12 months. Based on the current customer acquisition run rate, Netezza has already captured 1% of the DW management market (when only software for the overall market is counted; see *Worldwide Data Warehousing Tools 2005–2009 Forecast*, IDC #34316, November 2005).
- ☒ Correlating the results of organizations' interest in DW appliances with their overall data warehouse project stage indicates that those who are currently in planning stages of their DW projects are more willing to evaluate DW appliances; however, 43% of those willing to evaluate are already at the production, maintenance, upgrade stages (see Table 2).

32% of organizations indicated that they would be interested in evaluating a data warehouse appliance.

Those who are currently in planning stages of their DW projects are more willing to evaluate DW appliances.

### ***Expected Benefits***

Expected benefits from DW appliances are shown in Figure 1. Note that this survey did not ask about whether the lower up-front cost of DW appliances is an attractive variable. However, other research suggests that while an important consideration, the lower up-front cost of DW appliances is not the primary consideration for purchasers (source: IDC interviews with end users). Usually, initial cost savings are partly due to lower software licensing costs and hardware costs. However, lower TCO from a maintenance and upgrade perspective seems to be more important.

**TABLE 2**

## End Users' Interest in Data Warehouse Appliances (% of Respondents)

Q. *Would you be interested in evaluating and purchasing a data warehouse appliance from one IT vendor that combines server, storage, and software into a single preconfigured bundle?*

Q. *At which stage is your organization in the data warehousing process?*

Stage of DW Project	Interested in Evaluating a DW Appliance	Not interested in Evaluating a DW Appliance
Initial planning	22.8	11.2
Prototype or pilot	7.9	6.5
Design	3.0	5.6
Construction	5.9	4.7
Implementation	7.9	7.5
Production/maintenance	38.6	36.9
Upgrade	4.0	12.1
No plans at this time	9.9	15.4
n =	101	214

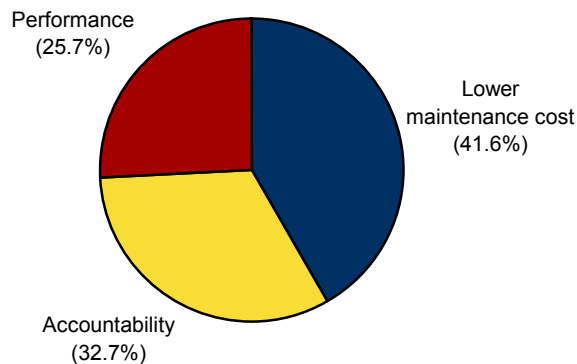
Note: Data reflects only U.S. respondents.

Source: IDC and *DM Review*, July/August 2005

**FIGURE 1**

## Expected Benefits of Data Warehouse Appliances

Q. *What benefits do you expect such a data warehousing appliance to provide?*



**n = 183**

Note: Data reflects only U.S. respondents.

Source: IDC and *DM Review*, July/August 2005

---

## **Current Data Warehouse and Business Intelligence Appliance Offerings**

As mentioned previously, DW and BI appliances come in three primary flavors: complete stack data warehouse appliances, virtual data warehouse appliances, and business intelligence appliances (see the Market Definitions section of this study).

### ***Complete Stack Data Warehouse Appliances***

#### **Netezza**

Founded by market veteran Jit Saxena (founder and former CEO of Applix), Netezza is the pioneer and current leader in the market for DW appliances. With 65 customers and 195 employees, it is enjoying first-mover advantage and has been successful partly through direct sales and partly through partnerships with most major BI and ETL software vendors, systems integrators, and industry-specific service providers that in turn process and analyze large amounts of data on behalf of their clients. Since its inception in 2000, the company has amassed an impressive lineup of customers in retail, ebusiness, telecommunications, services, healthcare, and media verticals. Netezza offers solutions for customer segmentation, revenue leakage management, and fraud prevention, among others.

The market's awareness of Netezza is rising. IDC's May 2005 *BI Survey* showed that 10% of respondents were aware of Netezza as a DW vendor (the highest awareness percentage, 84%, went to Oracle). Netezza's primary innovation, besides its patented software, was in putting together the first appliance based on open source OS and DBMS software and combining it with off-the-shelf hardware. Netezza's innovation was also in combining SMP and MPP techniques into a technology it calls asymmetric massively parallel processing (AMPP) — and in its Intelligent Query Streaming technology, in which processing is done in parallel at the disk level. The Netezza product portfolio includes a range of appliances called Netezza Performance Server (NPS) that are priced depending on their data and processing capacity. The entry price for a 1.5TB NPS 5000 appliance is about \$200,000.

#### **Datallegro**

Founded in 2003, Datallegro provides a DW appliance that combines both software and hardware. Its first commercial product launch was in March 2005, and since then the company has gained a number of marquee clients in the retail, telecommunications, and services verticals. The company has also established partnerships with Business Objects, Informatica, and Microstrategy. Datallegro's core innovation has come from its methods of I/O optimization as well as use of MPP.

Datallegro appliances are available in two types of configurations: high performance (P-series) and high capacity (C-series). A typical base price for a P-series DW appliance is \$250,000. Datallegro also has encryption technology, which can be used to protect inactive data from theft and address certain compliance requirements. Given the company's pipeline, IDC expects Datallegro to be a fast climber in the DW appliances market.

## **Calpont**

Calpont was founded in 2000 and started focusing on the data warehousing market in 2004. The company currently expects to complete the commercial development of its DW appliance later in 2006 and to put this product into general release in mid-2007. The Calpont DW appliance is based on a massively threaded database processor (with a thread count that is significantly higher than its original 64 threads), the use of a large amount of memory (over 256GB), and the appliance's efficient memory structures for fast storage and access of data in memory and on disk.

Calpont's database processor enables its engine to continuously process workloads at high throughput rates. Using this processor and its efficient memory and storage, Calpont's solution focuses on eliminating the I/O bottleneck that occurs in traditional RDBMS-based data warehouses. Calpont's massively threaded DW appliance processes ANSI SQL functions and uses Linux as the OS. The first version of the product was completely memory based. The soon-to-be-released version of the product will have an added embedded database. Initially, the Calpont appliance is recommended in configurations ranging in size from half a terabyte to over 20TB per single box.

Sitting between the categories of complete stack DW appliances and BI appliances is IBM.

## **IBM**

IBM has introduced two versions of its IBM Data Warehousing Balanced Configuration Unit (BCU) to address its clients' demand for preconfigured data warehouse appliances. The first BCU released in June 2005 was for AIX operating system, and the second BCU, released in June 2006, is for Linux operating system. Both appliances combined IBM's storage, server, and DB2 Data Warehouse Edition software into a single preconfigured unit.

The most fundamental difference between IBM's BCU appliances and those from other full stack DW appliance vendors, is that it includes not only software for data warehousing, but also for business intelligence, including the database-embedded data mining, OLAP, and other related tools. Some of the other differences include the use of SMP (as opposed by MPP by other appliance vendors) and read-write functionality as well as the ability to purchase several related IBM services.

However, several characteristics of IBM BCUs are similar to those of other DW appliances, including single point customer service group and preconfigured hardware and software components. Nevertheless, because of all these add-ons, IBM's DW appliances come at a higher price tag than those of its smaller competitors. Currently, IBM is using its BCU products to address the needs of its higher end customers — a user group that has somewhat different requirements from the typical use cases of purchasers of other DW appliances.

## ***Complete Stack Business Intelligence Appliances***

### **Celequest**

Founded in 2002 by industry veteran Diaz Nesamoney (cofounder and former COO of Informatica), Celequest released in April 2006 the first BI appliance, named LAVA. The difference between this solution and that of the DW appliance vendors is that Celequest incorporates its core operational BI (data integration) software, analytics, end-user dashboards, and tools for collaboration. However, as with DW appliance vendors' solutions, Celequest's appliance incorporates open source database and operating system software as well as its proprietary data monitoring and integration software. The LAVA appliance includes preintegrated and pretested software and hardware. The entry-level price point for the appliance, which the company claims can support several thousand users, is \$125,000. There is no per-user cost to the appliance. The solution is sold with a 20-day installation service agreement.

LAVA is fueled by Celequest's patented streaming data flow engine, which provides continuous integration of operational and historical data, time-based calculations, business rules processing, and event-driven alerting. Its Performance Dashboard technology displays information in an intuitive, personalized format so that any user can immediately see what is happening in the key areas of business operations.

## ***Virtual Data Warehouse Appliances***

### **Greenplum**

Greenplum was formed in 2003 as a merger of Metapa (with its analytic infrastructure product) and Didera (with a supercomputing platform). Both companies trace their beginnings to 2000. The company offers two primary products: a high-performance database server named BizgresMPP (a commercial product based on PostgreSQL) and a free open source database named Bizgres. BizgresMPP is free for development and evaluation but requires a paid software license for production use.

BizgresMPP is a high-performance data warehousing system built using large clusters of open source software and low-cost, commodity hardware (note that Greenplum itself does not sell the software). The product is optimized for multiterabyte data warehousing environments.

The open source Bizgres product is part of the Bizgres.org project, whose mission according to Greenplum "is to make PostgreSQL the most robust open source database for business intelligence." Bizgres is designed for small to medium-sized BI applications and is best suited to support departmental data marts ranging from 10GB to 300GB. Greenplum provides support for both BizgresMPP and Bizgres.

Greenplum counts among its partners other open source vendors, such as JasperSoft for queries and reporting, and Kinetic Networks, producers of open source ETL software.

## **Kognitio**

Kognitio was founded in 2005 as a merger between two U.K. companies — WhiteCross (founded in 1988) and PMPL (founded in 1991).

Kognitio continues to develop WX2, a scalable analytical database that is designed specifically to enable an ad hoc analysis of complex, large volume granular (atomic-level) data. It is a software-only solution that uses high-speed massively parallel processing technology to create a fast data mart/warehouse platform.

WX2 runs on low-cost industry-standard hardware, does not use indices or data partitions, and can be scaled to handle hundreds of terabytes of data. WX2 was initially created and developed by Whitecross Systems in 1988, and former versions included WX/DES, a proprietary hardware-based solution. In 2005, Whitecross Systems launched the new version of its analytical database, WX2, which is designed to run on nonproprietary industry-standard blade servers from HP, IBM, and Dell under Linux and is offered as either a licensed software solution or a managed service.

In August 2005, Whitecross Systems merged with Kognitio, and Kognitio now owns the IP to WX2. Kognitio continues to develop WX2 at its Technology and Data Centre in Bracknell, Berkshire, United Kingdom.

## **SAP NetWeaver BI Accelerator**

In July 2005, SAP introduced its NetWeaver BI Accelerator as an appliance that combines software and hardware in one package. To create the BI Accelerator appliance, SAP has partnered with Intel, which provides the processors, and HP and IBM, which provide their respective server and storage technologies. The BI Accelerator is a highly scalable analytic server that processes queries initiated by users of SAP NetWeaver BI. Its uniqueness and the features that make the BI Accelerator highly scalable are the use of SAP's TREX search technology in conjunction with blade server architecture provided by its hardware partners. It is important to note that organizations looking to deploy the SAP BI Accelerator will still need to deploy an instance of the SAP NetWeaver BI (formerly SAP BW and related software) with its own associated hardware (i.e., the Accelerator only works by accelerating an existing SAP BI software implementation). (For more information on SAP NetWeaver BI Accelerator see *SAP Business Intelligence Accelerator: A High-Performance Analytic Engine for SAP NetWeaver Business Intelligence*, IDC #34234, October 2005.)

## **PARACCEL**

Formed in February 2006 by, among others, Barry Zane, PARACCEL acquired the IP of XPrime. With an infusion of angel investments, the company has been in stealth mode since February 2006, focused on R&D based on the former XPrime technology. PARACCEL expects to have beta trials of its appliance in the third quarter of 2006 and a general release at the end of 2007. PARACCEL's modularly stacked appliance combines several techniques to achieve high performance in its DW appliances. These include a column-based database, grid-based architecture, and commodity hardware. The appliances interface with BI tools through SQL.

While the company's appliance accelerates query performance of any existing data warehouses based on RDBMSs, the initial market entry is expected to focus on the Microsoft SQL Server market. The PARACCEL appliance uses a technique called logical mirroring, where it mirrors data from the original DW inside PARACCEL's own appliance. The appliances will be sold in increments of a blade (not a rack) and are expected to find an initial client base among existing Microsoft SQL Server-based large data warehouses.

## FUTURE OUTLOOK

As the market for DW and BI appliances develops, existing vendors as well as end users should evaluate the three market scenarios discussed in the sections that follow.

---

### Scenario 1

Success of the vendors highlighted in this study will attract new competitors and innovators into the market. The data warehouse management market that has for many years been dominated by four to five leading database vendors will experience disruptive change. Although we can only be sure in hindsight, it seems that the market is developing by the book — literally! The inability of large competitors to react to the emerging DW appliances market quickly is playing itself out based on the scenario laid out in Clayton Christensen's book *The Innovator's Dilemma*. In other words, DW appliance vendors are being given the opportunity to grow organically and take increasing share of a certain segment of the market. Larger, established vendors (incumbents) will often say that they don't see the new DW appliance vendors in competitive deals. That was one of the points made by Christensen. End users whose needs have not been fulfilled by these incumbents don't even consider them for their projects in which they evaluate DW appliances. Although IBM and SAP are two large IT vendors listed in this report, their respective appliances address primarily the needs of their installed base.

Should this scenario of continued organic growth of DW appliance specialists play itself out over the next few years, IDC estimates that the current DW appliances market, which early estimates put at about \$50 million to \$75 million in worldwide revenue in 2005, will have a CAGR of 70% over the next five years to reach just over half a billion dollars. Forecast assumptions that were used in this market evaluation include:

- ☒ Three to five new entrants will emerge over the next three years (two seem to be already under development, with PARACCEL being in stealth mode [see the PARACCEL section of this document]).
- ☒ No other major DBMS or systems vendor will enter the market over the forecast period. (Note: Eliminating this assumption would create a scenario of a higher growth rate due to the sales execution power of such market entrants. Theoretically, primary candidates could include HP, IBM, Unisys, Sybase, and EMC; Oracle and Microsoft would be less likely candidates.)

DW appliances market, which early estimates put at about \$50 million to \$75 million in worldwide revenue in 2005, will have a CAGR of 70% over the next five years.

- ☒ Forecast assumptions take into account a known number of customers at current vendors and initial growth rates of these vendors (vendor-specific data under NDA).
- ☒ SAP BI Accelerator is assumed to be adopted by 3% of SAP's BW clients over the next three years. (This may seem a low adoption percentage assumption, but it is in line with the information on SAP's current beta runs and future product release schedules.)
- ☒ Over the next five years, a challenge to RDBMS-based DW appliances that are used for ad hoc query and analysis will come from content access (search technology) vendors that provide many of the same value propositions of a high-performance, pretuned, and low-maintenance DW appliance.

The forecasted growth rate of the DW appliances market compares to about a 10% growth rate over the same time period for traditional BI and traditional DW tools. Given current market intelligence, most of the vendors in this study have seen revenue growth rates in excess of 100% over the past three years. Although these growth rates are coming off of very low base year revenue levels, the potential for extremely strong growth continues. While none of the vendors are individually a major threat to any of the leading database vendors, as a group or emerging market segment, they are clearly beginning to have an impact — a classic example of the phenomenon of *The Innovator's Dilemma*.

---

## Scenario 2

Large data warehousing and other systems vendors will act rapidly to take advantage of the innovation provided by the new vendors. Several acquisitions will occur in the short term (one to three years), and the newly acquired technology will become part of the broader portfolio of data management products of the larger vendors. This scenario requires the larger vendors to look beyond their own short-term interests and acknowledge that the new entrants are addressing legitimate market needs not being met by larger vendors' solutions.

Who are the likely acquirers of the new technology? The fastest-growing data warehousing company, Microsoft, is the least likely acquirer in the short term. For example, the DW appliance vendors are all using open source software either for the OS or the database or both. This product strategy is part of the equation that has enabled such vendors to keep their prices down. Also, Microsoft already has the lowest average price point for DW solutions among its peers. It is therefore difficult to see Microsoft acquiring such solutions in the short term.

Oracle is also an unlikely acquirer but for different reasons. As the current leader of the DW market, Oracle has the broadest (in terms of DW size) customer base of all its major competitors. It might therefore not see the value of solutions targeting a market segment where Oracle already has a strong presence. IBM, on the other hand (like Teradata), is skewed toward larger DW implementations, and solutions for targeting the segment in the sub-10TB range may be attractive to IBM. In addition, IBM's investment in open source initiatives predisposes it well toward embracing such technologies. However, there are many other potential acquirers such as enterprise

applications vendors (e.g., SAP, Lawson), large BI vendors (e.g., Business Objects, Cognos, SAS, Hyperion), or server and storage vendors (e.g., EMC, HP, Hitachi, Fujitsu).

Regardless of who the specific players in the proposed M&A activity will be, this scenario will mean that customers will continue to enjoy the benefits of the new technology while being supported by the large services organizations of the leading IT providers. On the other hand, it will mean that the independent market for DW appliances will not survive on its own.

---

### **Scenario 3**

The third scenario requires a more long-term perspective. It is legitimate to ask the question whether in the long term, traditional DW software will be relegated to just storing historical data for compliance and long-term trending analysis purposes, while all operational BI needs will be serviced by other technologies. In general, few organizations store several years of history for active/operational decision making. The data does exist, but it is usually not utilized for operational BI. As the trend of bringing BI to the masses progresses and ad hoc analysis requirements of power users increase while IT resources can't keep pace with such requirements, new operational BI tools and techniques will become much more widely accepted. This scenario will be affected by the emerging push of content access or enterprise search vendors into the business analytics market.

These scenarios don't represent a zero sum game. For example, to address the points in scenario 3, Teradata has over the past several years advanced its concept of active data warehousing by opening its core database to inputs from real-time data feeds and incorporating processing techniques to combine historical and real-time data. In the final analysis, while IDC believes that over the intermediate term (three to five years) additional large vendors will step into the market either as acquirers or OEM partners of appliance software, the independent vendors will not disappear. Instead, an equilibrium of a few large vendors dominating the market with several smaller niche players will persist.

Although over the intermediate term, additional large vendors will step into the market either as acquirers or OEM partners of appliance software, the independent vendors will not disappear.

---

## **ESSENTIAL GUIDANCE**

### **Impact on Competition**

In the past, developers and DBAs have had to compromise between the up-front costs of creating complex indices and aggregations and doing "quick and dirty" loads and dealing with the consequences of poor query performance afterward.

At the most fundamental level, the reason for the attractive price/performance level of the solutions addressed in this study is the fact that they are purpose built and dedicated to analytics rather than being general-purpose RDBMSs. However, even this point needs to be qualified because purposed built solutions are not new. Teradata is one of the leading DW vendors whose business and technology is dedicated to data warehousing. So was Red Brick Software, which is now part of IBM.

At the most fundamental level, the reason for the attractive price/performance level of the solutions addressed in this study is the fact that they are purpose built and dedicated to analytics.

In fact, on initial view, the vendor most threatened by DW appliance vendors is Teradata. However, this view is, at least for now, misleading. The company continues to enjoy success (in fact accelerated revenue trends) in the multiterabyte enterprise DW (EDW) deployments. Teradata remains the preferred vendor for many organizations when it comes to the deployment of the largest EDWs for customers that require not just software and hardware but also a full menu of professional services — something that the smaller vendors are not able to supply at this stage in the game. As IDC research has shown in the past, larger DWs are becoming larger at higher growth rates than midsize and smaller DWs. In general, it looks like the rising tide is lifting all boats. At a time when appliance vendors are experiencing rapid growth rates, Teradata has turned in double-digit software revenue growth in the past two years.

Similar trends exist with respect to BI software. Some of the most sophisticated users of BI and DW software continue to invest in specialty products they integrate on an as-needed basis using either internal IT staff or external systems integrators. This continued trend for specialty tools mirrors that found in other industries. Although software tool consolidation is a factor, market research shows that it is much less of a factor (especially in BI) than large vendors' messaging would let us believe it to be. The market is always divided into the specialist and the generalists, and business analytics is no exception.

The market is always divided into the specialist and the generalists, and business analytics is no exception.

The question is what large database vendors can do, and when will they do it? Christensen's concept of *The Innovator's Dilemma* suggests that they won't do anything for some time. The trend has been identified and highlighted at an extremely early stage in large part due to savvy marketing on the part of the emerging vendors.

An evaluation of the above scenarios shows that vendors should reexamine their existing business analytics platform products and partnerships. The new appliances market is presenting some attractive opportunities that are finding a receptive audience from end users.

End users need to also evaluate their business analytics solutions and weigh the benefits provided by DW and BI appliances against the risks of working with small vendors.

---

## Conclusion

IDC believes that adoption of DW and BI appliances is not just a passing trend based on limited anecdotal evidence. Early quantitative surveys lend credence to strong interest in, awareness of, and adoption of DW appliances. For the foreseeable future, we expect DW and BI appliances to be deployed alongside traditional DW solutions at large organizations and in some cases as new primary DW solutions, especially at midsize organizations that may not have the resources for traditional DW solution acquisition, deployment, and maintenance.

The goal of this study is not to suggest that traditional data warehousing solutions based on separate database software (primarily RDBMS) and server and storage hardware will disappear. Too much has been invested in existing data warehouses and the business logic of applications that sit on RDBMSs. Today, the emerging

appliance vendors don't yet have the resources to compete with the breadth of offerings from larger vendors, which, besides professional services, include DW-specific services such as full breadth of failover, disaster recovery, backup, and remote database replication services, among others. However, appliance vendors can continue to take away share in their chosen specialty niches addressing the business needs outlined in this study.

Prepackaged DW appliances are going to make a difference in the market as more end users evaluate, adapt, and deploy these solutions. DW appliances are fulfilling real, existing needs of a specific segment of the market instead of being a novelty (i.e., an offering for a few adventurous IT hacks). The current trends show that:

- The initial wave of DW appliance purchasers are skewed toward larger companies that have experience with DW, have been disappointed by existing solutions, and are willing to experiment with new technologies. However, as market research shows (refer back to Table 2), new prospects are also willing to consider DW appliances, but perhaps for different reasons (further research would need to be done to fine-tune the purchasing criteria of different market segments).
- Complete stack DW appliances seem to have more of the important characteristics desired by end users than do virtual DW appliances. Pretuning, preconfiguration, single-vendor responsibility, and fast implementation times are the highest valued features of these DW appliances.

Complete stack DW appliances seem to have more of the important characteristics desired by end users than do virtual DW appliances.

---

## **The Sweet Spot for Data Warehouse Appliances**

The current sweet spot for DW appliances involves business initiatives that require:

- Ad hoc querying of large data volumes
- Focus on power users
- Fast data load speeds, which have been mentioned as an attractive feature of full stack DW appliances
- Single-subject domain focus (at least initially):
  - Initial, narrow domain focus is not directly correlated to the amount of data, which does not seem to be the deciding factor. Similar to the debate about what is a DW versus a data mart, DW appliances can come in a variety of sizes.
  - It's already clear that the single-subject domain focus is only an initial market entry strategy for appliances products. For example, Netezza has several customers that are using its appliance as a multisubject data warehouse.
- Sample applications that have been deployed on DW appliances, including:
  - Patient care records analysis

- eCommerce, click-stream analytics
- Telecom CDR analysis
- Inventory analysis
- Loan transactions analysis
- Marketing campaign analysis

## LEARN MORE

---

### Related Research

- Worldwide Business Intelligence Tools 2006–2010 Forecast* (IDC #200973, April 2006)
- Worldwide Data Warehousing Tools 2005–2009 Forecast* (IDC #34316, November 2005)
- SAP Business Intelligence Accelerator: A High-Performance Analytic Engine for SAP NetWeaver Business Intelligence* (IDC #34234, October 2005)
- Boral, H. and D.J. DeWitt. *Database Machines: An Idea Whose Time Has Passed? — A Critique of the Future of Database Machines*, Proceedings of the 1983 Workshop on Database Machines (Springer-Verlag), (1983), 166–187
- Bratbergsengen, K. *Parallel Database Machines*, Rivista di Informatica, Vol. XXV, n. 4 (ottobre–dicembre 1995)
- For an excellent review of the history of the decision support IT market, see D.J. Power's article at [www.dssresources.com/history/dsshistory.html](http://www.dssresources.com/history/dsshistory.html).

---

### Copyright Notice

This IDC research document was published as part of an IDC continuous intelligence service, providing written research, analyst interactions, telebriefings, and conferences. Visit [www.idc.com](http://www.idc.com) to learn more about IDC subscription and consulting services. To view a list of IDC offices worldwide, visit [www.idc.com/offices](http://www.idc.com/offices). Please contact the IDC Hotline at 800.343.4952, ext. 7988 (or +1.508.988.7988) or [sales@idc.com](mailto:sales@idc.com) for information on applying the price of this document toward the purchase of an IDC service or for information on additional copies or Web rights.

Copyright 2006 IDC. Reproduction is forbidden unless authorized. All rights reserved.

---

**Published Under Services:** Analytics and Data Warehousing Software; Business Intelligence and Data Warehousing Strategies