

April 4, 2008

Appliance Power: Crunching Data Warehousing Workloads Faster And Cheaper Than Ever

by James Kobielus

for Information & Knowledge Management Professionals



April 4, 2008

Appliance Power: Crunching Data Warehousing Workloads Faster And Cheaper Than Ever

by **James Kobielus**

with Connie Moore, Boris Evelson, Rob Karel, and Norman Nicolson

EXECUTIVE SUMMARY

Appliances are taking up permanent residence in the heart of the enterprise data center — the data warehouse (DW). DW appliances — in all their bewildering proliferation — are moving into the mainstream. The reason? They are preconfigured, modular devices that support quick deployment for DW killer applications — most notably, accelerating online analytical processing (OLAP) queries against large, multidimensional data sets. DW appliances prepackage and pre-optimize the processing, storage, and software components for fast OLAP and fast data loading. Information managers should now factor DW appliances into data center deployments. But don't forget the basics, and use the same core criteria to evaluate DW appliances as for equivalent DW software, including price-performance, functionality, flexibility, scalability, manageability, integration, and extensibility.

TABLE OF CONTENTS

- 2 **Appliances Are Pervading All Niches of Enterprise Infrastructure**
- 2 **DW Appliances Are Coming To The Fore . . .**
- 4 **Pure Plays And Established Vendors Are Jostling For DW Appliance Dominance**
- 8 **Competitive Fray Strains The Appliance Definition To The Breaking Point**
- 10 **Still, Appliances Are Beginning To Deliver On Their Promise**

RECOMMENDATIONS

- 12 **Focus On DW, Data Mart, And Analytic Price-Performance Requirements**
- #### WHAT IT MEANS
- 14 **I&KM Pros Will Continue To Grapple With The DW Appliance Value Proposition**

NOTES & RESOURCES

Forrester interviewed 10 vendors: Business Objects (an SAP company), DATAlegro, Dataupia, Greenplum, IBM, Microsoft, Netezza, Oracle, ParAccel, SAP, and Teradata. Forrester also interviewed three user companies in the telecommunications and retailing industries.

Related Research Documents

["Data, Data Everywhere!"](#)
June 23, 2007

["IBM Leads In Enterprise ETL With A Comprehensive Information Management Suite"](#)
May 2, 2007

["The Forrester Wave™: Enterprise ETL, Q2 2007"](#)
May 2, 2007

APPLIANCES ARE PERVADING ALL NICHEs OF ENTERPRISE INFRASTRUCTURE

Enterprise software vendors have devised many ingenious new approaches for crafting solutions. Over the past few years, vendors have ventured well beyond their traditional focus on licensed software packages. Many have begun to offer solutions that incorporate such diverse approaches as open source software, service-oriented architecture (SOA), and software-as-a-service (SaaS).

Some software vendors have even taken a bold step into the world of hardware. A growing number are offering “appliances,” which integrate software with processors, storage, and other hardware to deliver function-specific, performance-optimized solutions for quick deployment. An appliance essentially allows a vendor to pre-equip a “shrink-wrapped” infrastructure component to fit a particular functional role and support a specific capacity, throughput, and performance profile. Appliances hold the promise to buyers of being easier to procure, and potentially cheaper to deploy, as they require only one vendor to deal with and fewer resources for upfront integration and optimization due to pre-configuration by the vendor.

Appliances — in all their bewildering proliferation — are here to stay, and they are moving into the mainstream of enterprise computing and networking. To alleviate performance bottlenecks at various points in their distributed infrastructures, enterprises are deploying appliances to process growing transaction volumes at the application server, intranet perimeter, network backbone, and enterprise service bus.

DW APPLIANCES ARE COMING TO THE FORE . . .

Appliances have even begun to take up permanent residence at the heart of the enterprise data center: in the data warehouse. Increasingly, DW vendors focus on integrating, packaging, and pricing their products as preconfigured, modular appliances for quick deployment.¹ DW appliances consist of processing, database, storage, interconnection, and software components that have been prepackaged, preconfigured, and pre-optimized to accelerate core DW functions such as multidimensional OLAP queries and bulk data loading. Appliance vendors vary widely in how they implement these features within the context of their commercial solution, but the general overview of the typical appliance is relatively constant (see Figure 1).

Information managers should begin now to factor appliance-based solutions into data center deployments. DW appliances appeal to information managers because:

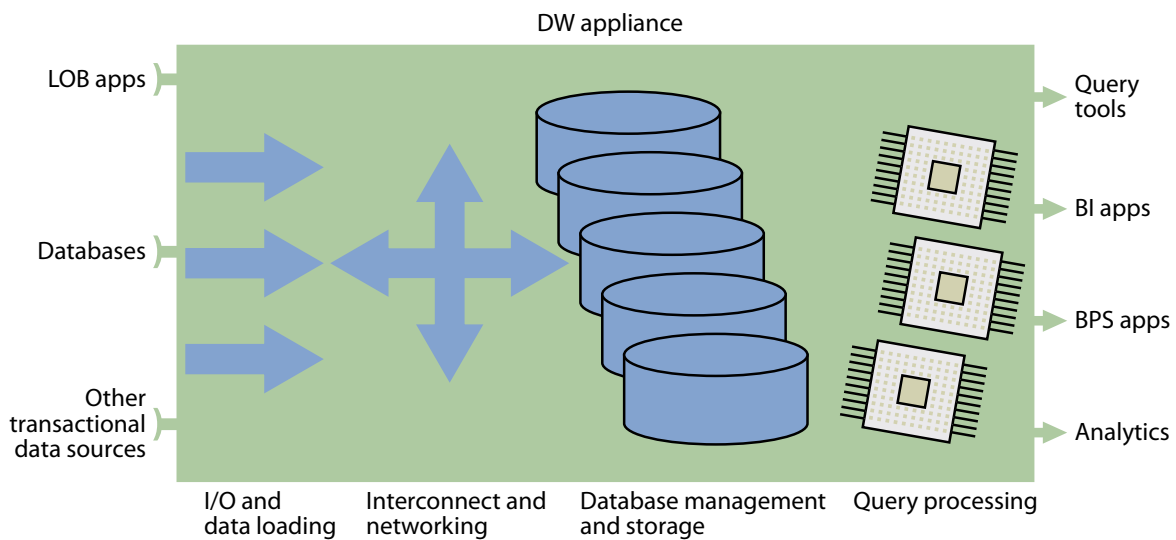
- **Speed and simplicity are a powerful value proposition.** I&KM professionals must often deploy data marts quickly to address short-fuse BI and analytics requirements. What’s more, organizations often need to do so without the complex configuration and optimization tasks associated with full-blown enterprise DWs. Appliances address these imperatives by promising time-to-value that’s closer to the “plug and play” end of the deployment spectrum.

- **Business analytics thrives on cheap horsepower.** Enterprises run their businesses on analytics, which means that the amount of processing and volume of data being handled by DWs continues to skyrocket. Consequently, DWs are an expanding component of the enterprise computing budget, a trend that causes organizations to search for cost-effective alternatives. Appliances generally offer high-performance analytics processing at a fraction of the cost of running equivalent workloads on traditional enterprise DWs.
- **Appliances are rapidly scaling and maturing into enterprise-grade DW platforms.** Though they still primarily address tactical data mart requirements, DW appliances from established pure plays such as DATAlegro, Dataupia, Greenplum, and Netezza are rapidly surmounting the scalability curve. Increasingly, enterprises are offloading more and more of their analytics processing to robust DW appliances, paving the way for these solutions to someday address high-end enterprise DW (EDW) requirements.

... But They Haven't Dislodged Traditional DW Solutions From The High-End Market

DW appliances are the upstart approach, so pure-play vendors of these solutions recognize that the top tier of the EDW market may still be slightly beyond their grasp. To understand the pros and cons of DW appliances versus traditional EDW solutions, focus on how DW appliances are currently being used in enterprises and when it is appropriate to use a DW appliance (see Figure 2).

Figure 1 Functional Overview Of A Typical DW Appliance



45641

Source: Forrester Research, Inc.

PURE PLAYS AND ESTABLISHED VENDORS ARE JOSTLING FOR DW APPLIANCE DOMINANCE

Pure-play DW appliance vendors have gained considerable mindshare over the past several years, but this growing market segment is really up for grabs. Enterprises are deploying DW appliances from several distinct groups of solution providers (see Figure 3).

Teradata Pioneered The Concept Of Purpose-Built EDW Hardware/Software Solutions

Teradata is widely acknowledged as the pioneering vendor in providing appliance-like DW hardware/software solutions, having provided such solutions since the mid-1980s. Teradata continues to provide purpose-built, integrated solutions for DWs of all sizes, though its primary emphasis is on EDWs for large enterprises. To date, though, Teradata has not aggressively positioned its EDW offerings as DW appliances.

IBM's Dynamic Warehousing Initiative Validated DW Appliance Concept In The Enterprise

Startups such as DATAlegro, Greenplum, and Netezza have popularized the DW appliance concept over the past several years. However, in terms of legitimizing DW appliances among large enterprises, the market experienced an important turning point in March 2007. That was when an established EDW vendor, IBM, announced its Dynamic Warehousing initiative. What Big Blue did was re-architect and reposition its entire EDW product family as DW appliances, ranging in functionality and price-performance from enterprise-grade to midmarket-focused. Under the Dynamic Warehousing initiative:

- **IBM launched the “Balanced Warehouse.”** Forrester believes that IBM provides the most comprehensive, scalable enterprise DW appliance family on the market. IBM's Balanced Warehouse solutions address diverse DW price points and requirements, ranging from high-end enterprise DWs down to smaller, function-limited DWs and departmental data marts. Balanced Warehouses build on and extend the Balanced Configuration Unit (BCU) DW appliance solutions that IBM had introduced almost two years prior.
- **IBM later introduced a blade-based appliance.** In October 2007, IBM moved its DW appliance strategy forward with the announcement of a blade-based appliance that incorporates the core data integration (DI) and data quality (DQ) functionality of its IBM Information Server suite.² As a result, over the past several months, other leading database vendors — most notably, Oracle and Microsoft — repositioned their strategies more aggressively toward DW appliances.

Figure 2 EDW Still Has The Upper Hand, But DW Appliances Are Often Cheaper, And Just As Fast

	Traditional custom-built EDW	Appliance-based DW deployment
What it is	User acquires DW server hardware, software, storage, and interconnect from one or more vendors, then custom integrates, configures, optimizes all components for OLAP, ETL, and other DW functions.	User acquires pre-integrated, preconfigured, pre-optimized DW server hardware, software, storage, and interconnect from one vendor.
When to use	Enterprise data warehouse managing multiple analytic applications, domains, and entities	Subject-specific, departmental, and/or tactical data marts for focused BI/analytics requirements, though may be deployed in hub-and-spoke implementation with central DW feeding multiple data marts
Pros	<ul style="list-style-type: none"> • Often (not always) more scalable (in database size, high-volume data loading, greater usage concurrency) • Acquire DW, BI, ETL, DQ, and other solution features and components from single source • Building future data marts from an established EDW may be quicker and cheaper than separately deploying disparate data marts. • More unified administration and support than going with multivendor DW deployment 	<ul style="list-style-type: none"> • Less expensive to acquire, deploy, and administer • Less complex to integrate, configure, optimize, and administer • Often just as fast as EDWs in processing multidimensional OLAP queries and loading data in bulk
Cons	<ul style="list-style-type: none"> • May be more expensive to acquire, deploy, and administer, though there are cases where, on a per-user or per-terabyte basis, across a large organization and a large set of workloads, a large EDW system is more cost-effective than a DW appliance • May be more complex to integrate, configure, optimize, and administer, though managing one EDW with multiple applications may be easier than trying to manage an equal number of disparate data marts 	<ul style="list-style-type: none"> • May be (but not always) less scalable than EDWs that users custom-build from server, storage, interconnect, and software provided by one or more vendors • Must often (but not always) acquire DW, BI, ETL, DQ, and other solution components from multiple sources • If using appliance from pure play, less unified administration and support than going with single-vendor enterprise DW deployment

45641

Source: Forrester Research, Inc.

Figure 3 Startups Driving DW Appliance Market; Established DW Vendors Responding Aggressively

Appliance market segment	Principal vendors*	Their differentiation
Startup DW appliance pure plays	DATAlegro, Dataupia, Greenplum, Netezza, ParAccel	These vendors offer solutions for online analytical processing (OLAP) mixed-query workload optimization and fast extract transform load (ETL).
Established DW vendors	IBM, Microsoft, Oracle, Teradata	Teradata has long offered high-end, appliance-type, integrated EDW solutions, but it keeps its distance from marketing its offerings under the banner of "appliances." Other long-term DBMS/DW vendors offer preconfigured hardware/software bundles. IBM includes its own server and storage offerings in its DW appliances, while Oracle and Microsoft provide their DBMS and ETL offerings as DW appliances through hardware partners.
Established vendors of servers, storage, and other data center hardware	Bull, Dell, EMC, HP, Sun	HP, keeping its distance from the appliance label, has rolled out appliance-like hardware/software DW solutions under the Neoview brand. Bull, Dell, EMC, and Sun partner with pure-play DW appliance vendors.

*List is representative, not exhaustive

45641

Source: Forrester Research, Inc.

Oracle Has Followed Suit . . .

In September 2007, Oracle announced the Oracle Optimized Warehouse Initiative. Under this effort, Oracle is accelerating customer EDW deployments by offering a choice of prebuilt, pre-optimized solutions that combine the performance, reliability, and scalability of Oracle Database with hardware and storage from partners. Each partner-provided Oracle Optimized Warehouse is an integrated, prebuilt, prevalidated solution consisting of database, storage, and servers that have been validated and tested to deliver optimal performance. Each is sold as a single product under a single support contact. Under this initiative, several Oracle hardware partners — EMC/Dell, IBM, and Sun — have introduced their respective Oracle Optimized Warehouses.

. . . As Has Microsoft . . .

Microsoft has also rolled out a partner-focused DW appliance product initiative. In May 2007, Microsoft announced that it had teamed with HP — and separately in September of that year, with Dell — to offer DW appliances that are powered by SQL Server 2005. Microsoft's DW appliance strategy is similar to Oracle's, but has, to date, resulted in fewer strategic partnerships and has not been fully integrated into Microsoft's go-to-market strategy for SQL Server.

... And The Established Vendors Of Servers, Storage, And Other Data Center Hardware ...

Many real-world DW deployments include servers, storage, and other hardware components from vendors such as Bull, Dell, EMC, HP, IBM, and Sun. So it makes strategic sense for these vendors to explore opportunities in the DW appliance segment — either with their own branded hardware/software solutions (à la HP and IBM), or as the hardware substrate underlying partner appliances. For example:

- **HP has achieved considerable mindshare, if not market presence.** Over the past year and half, HP has focused on its DW appliance-like Neoview solution (the vendor studiously keeps its marketing distance from the “appliance” label), while also, as noted above, partnering with Oracle and Microsoft on the latter’s DW appliance initiatives.
- **Sun has also partnered widely in the DW appliance arena.** Its partnerships include Greenplum, Oracle, and ParAccel (the latter a columnar database vendor that provides a hardware appliance-based version through Sun). Having recently acquired its own database product through the MySQL AB acquisition, it’s not clear whether or when Sun plans to develop a DW appliance product in-house.

... And — No Surprise — So Have The BI Vendors

Appliances have also gained a key foothold among BI vendors, often handling the same core application — OLAP acceleration — as solutions from Greenplum, Netezza, and the like, a phenomenon that blurs the practical distinction between DW and BI appliances. So far, though, the BI vendors’ own OLAP-acceleration (and report-acceleration, dashboard-acceleration, and business-activity-monitoring-acceleration) appliances have not achieved the degree of user adoption enjoyed by solutions from DW appliance pure plays. In part, that may be due simply to the fact that these BI vendors (i.e., Business Objects, IBM Cognos, SAP) have come into the appliance arena slightly later than the DW pure plays. It may also be due to the fact that the BI vendors still largely position themselves as purveyors of packaged software solutions (first and foremost), albeit with increasing productization focus on SaaS and appliances.

In the BI arena, SAP got the ball rolling in May 2006 when it announced general availability of its BI Accelerator (BIA) solution. SAP BIA is a hardware appliance technology that accelerates database queries when deployed in blade servers in conjunction with SAP NetWeaver’s Business Intelligence (BI)/Business Warehouse (BW) product. BIA’s hardware/software technology — co-developed by SAP and Intel — offers significant OLAP performance improvements on ad hoc queries. It accelerates data-intensive queries more cost-effectively than purely software-based OLAP architectures. It does so, in part, by using an in-memory column-oriented DBMS index.

Also, BI market leader Business Objects (recently acquired by SAP) announced in mid-2007 that it too had put appliances at the core of its go-to-market strategy. To develop BI appliances for

various customer segments, incorporating (potentially) its full BI solution stack, Business Objects has partnered with a wide range of complementary vendors, ranging from large established server/storage providers to pure-play DW appliance startups. Just as important, SAP BIA is a core technology for evolution of Business Objects' own BI solution stack.

Another BI leader, Cognos, bought a BI/BAM-acceleration appliance vendor, Celequest, in late 2006, and has itself recently been acquired by IBM. Already, IBM has promised to include Cognos' BI solutions as an option in the IBM InfoSphere Warehouse (the new branding of its DW appliances) and to increase integration and optimization of Cognos' offerings within InfoSphere Warehouse.

COMPETITIVE FRAY STRAINS THE APPLIANCE DEFINITION TO THE BREAKING POINT

The DW appliance wars are upon us. This can be seen in vendors' eagerness to slap the appliance label on a growing range of hardware-integrated solutions, most of which are far more complex and costly than true appliances — though, ostensibly, less so than the software-centric solutions they hope to supplant. Each vendor claims that its appliances are true plug-and-play solutions, though few customers are so naive as to imagine that a complex IT solution can be as easy to install and setup as, say, a toaster oven. Indeed, all DW appliance vendors provide at least some basic on-site deployment, configuration, optimization, and troubleshooting services to help customers get started with all the plugging and playing.

DW Appliances: The Concept Is Growing Bigger Than A Breadbox, Softer Than The Bread

In addition, vendors are still debugging their definitions of what exactly a DW appliance is and what value it offers vis-à-vis traditional DW software-centric solutions. Indeed, vendors and industry observers now proliferate alternate definitions of what constitutes a true appliance. As a result, appliances span this wide spectrum:

- **Simple DW appliances.** Some DW appliances may be regarded as “black boxes” that are designed and optimized for a single function or transaction type. Simple appliances, often packaged as blades or standalone assemblies, allow little if any modification or customization by the customer. For example, Dataupia provides function-specific, black-box DW appliances, though the vendor also offers massively parallel processing (MPP) functionality, linear scalability, and support for mixed workload environments.
- **Complex DW appliances.** Most DW appliances fall into this category, including those from leading pure plays (e.g., DATAlegro, Greenplum, and Netezza) and from DBMS/DW vendors that have begun to emphasize appliance-style packaging (e.g., IBM, Oracle). This sort of appliance is a complex assemblage of processing, input/output, storage, and other components integrated across one or more racks in an enterprise data center. Often, a complex appliance consists of one or more modular blades, which may or may not be able to stand alone. To address growing enterprise workloads, complex DW appliances are increasingly relying on

clustering, grid, MPP, symmetric multiprocessing (SMP), federated deployment support, column-oriented storage, and other approaches for scaling up and out.

- **Virtual software appliances.** Some vendors are now stretching the concept of appliances to the breaking point. One of the core features of appliances — physical tangibility — is starting to fall by the wayside. For example, Business Objects, an SAP company, partners with VMware on the nouveau notion of the “virtual appliance” — a self-contained software package that can be deployed rapidly to diverse operating and hardware platforms through virtualization technologies. Likewise, Ingres rolled out a BI software appliance through a partnership with open source BI software vendor JasperSoft. It’s not clear how these “virtual software appliances” differ from existing development paradigms, such as Java, that also promise, where BI/DW solutions are concerned, the ability to “write once, run anywhere.”
- **Prepackaged professional services.** To further soften the concept of a DW appliance, more and more vendors incorporate prepackaged professional services into their solutions. IBM, in particular, stresses global services as a core feature of its appliance offerings. With this approach, the appliance is a broad solution package that includes services to help customers install, set up, configure, optimize, and maintain the prepackaged hardware/software assembly. This heralds a curious, though understandable, trend in a market segment that touts its ability to deliver DW solutions that are “preconfigured” and “pre-optimized.”

Summing up these trends, we see DW appliances evolving along three dimensions: features, software-centricity, and services (see Figure 4).

Figure 4 DW Appliances Are Evolving Along Three Dimensions

Dimensions of appliance evolution	What it means
More feature-rich	DW appliances are expanding to address the full spectrum of enterprise requirements. Offerings range from single-function “black boxes” to multifunction, multirack, grid-enabled assemblages of hardware, software, and storage.
More software-intensive	DW appliances are becoming more “software-centric” to support a wider range of deployment scenarios. Solutions range from traditional appliances (tangible hardware solutions) to a new generation of self-contained software packages designed for quick deployment over diverse operating systems, hardware platforms, and virtualization layers.
More services-reliant	DW appliance-based solutions are getting bundled with a wider range of prepackaged professional services. Offerings range from self-contained turnkey systems to solution packages that include cradle-to-grave vendor-provided professional services for installation, configuration optimization, and management.

45641

Source: Forrester Research, Inc.

STILL, APPLIANCES ARE BEGINNING TO DELIVER ON THEIR PROMISE

Appliances are no marketing gimmick, though it can be difficult to distinguish between traditional DW offerings and those that choose to sail under this trendy new banner. Indeed, the recent embrace of the appliance go-to-market approach by the three top enterprise DBMS vendors — IBM, Microsoft, and Oracle — shows that appliances, however defined, are becoming the dominant approach for delivering DW functionality to customers of all sizes.

In real-world deployments, appliances are beginning to deliver on their promise of being quicker to implement, as well as easier and less resource-intensive to support, than traditional DW solutions. Based on the findings from our case study interviews, the companies that have deployed Netezza, DATAlegro, and Greenplum see appliances as an attractive alternative to traditional DW solutions. Information and knowledge management (I&KM) professionals turn to appliances not just to accelerate time-to-value on short-fuse DW deployments, but also to deliver an order-of-magnitude greater performance on OLAP and BI transactions at a fraction of the cost of traditional solutions.

The three companies we interviewed are relying on appliances to address various high-volume, high-capacity analytical requirements. These companies use DW appliances in lieu of traditional DW deployment, relying on them to offload specific analytical workloads from production DW environments.

- **Appliances cost-effectively offload high-volume query processing from the EDW.** A US-based software firm has deployed DW appliances to support analytical processing of hundreds of millions of call detail records (CDRs) daily for long-distance and wireless telecommunications carriers. Based in Fairfax, Va., the company found that traditional relational database management systems (DBMSes) were more suited to online transactional processing (OLTP) than to the high-volume, mixed-query OLAP required by its CDR auditing, antifraud, and predictive analytics service offerings. It also evaluated traditional DW offerings but found that they were considerably more expensive than appliances — as much as three to four times as costly — on a per-terabyte basis when managing large data sets. Nevertheless, the firm — which runs both Netezza's and DATAlegro's appliances in its data center — says traditional DW offerings are more scalable in support of highly concurrent usage.
- **Appliances provide a modular stepping stone to EDW-scale analytic workloads.** A Finland-based hosted analytical processor of retailers' point-of-sale (POS) data benchmarked commercial DW appliances against traditional DW solutions (using the services of an independent consultant). One of its core evaluation criteria was the ability of the chosen solution to cost-effectively process high-volume, mixed-query OLAP workloads. It chose Greenplum (on Sun hardware) based on its ability to handle analytic workloads at a fraction of the cost — one-eighth, to be specific — of a traditional DW solution. Another factor in Greenplum/Sun's favor was, per the benchmarks, faster loading of an equivalent amount of

data (20 minutes versus 3 hours). Yet another point was the ability to integrate the appliance rapidly to its existing BI (MicroStrategy), ETL (Informatica), and operational DBMS (Oracle) infrastructures. Other DW solutions that the company benchmarked include offerings from Teradata, HP, and Sun. Having proven itself in a production environment with dozens of internal and external users, the Greenplum deployment will be migrated over time to take over the entire analytic workload, scaling up to manage several years of POS data, from a traditional DW solution that will be phased out.

- **Appliances play well with customer investments in BI and analytic applications.** A British telecommunications company deployed Netezza's DW appliance two years ago to handle analytical processing on a growing database of CDRs, migrating away from a traditional DW deployment based on an enterprise RDBMS on general-purpose hardware. Upon testing the Netezza platform, the company saw dramatic speed increases across nine query types, with one complex query that formerly took more than 13 hours to return results clocking in at less than 6 minutes. Since the initial rollout, which took four weeks to complete, the carrier has continued to add capacity to its Netezza deployment while launching new BI and analytic applications — in Business Objects and SAS environments — that tap into the data running on the DW appliance. One migration detail — which took a few days to execute — was the need to rewrite some SQL queries to convert them to ANSI standard SQL from Oracle specific syntax to enable them to run against Netezza's database.

There Are Still The Inevitable Sticking Points

So, yes, DW appliances are providing powerful, modular building blocks for a new generation of mission-critical business analytics. But as the companies we spoke with bear out, commercial DW appliances vary widely in maturity, and vendors' promise of quick "plug and play" deployment must be countered with skepticism. Here are some issues to watch for:

- **Most DW appliance installations involve some on-site support by vendors.** Users often need to work through various deployment details with their solutions. Just because a vendor claims its appliance is "preconfigured" does not mean that it will be ready for production deployment right out of the box. Deploying a DW appliance may also involve some rewriting of existing BI applications, revision of ETL scripts, migration and transformation of large data sets, porting of in-database analytic applications, and other tasks necessary to begin processing of OLAP workloads at maximum efficiency. To the extent that the DW appliance uses parallel or grid-based scaling, developers may also need to rewrite and recompile their BI and analytic applications to take advantage of these performance-enhancing features.
- **Technical debugging, application development, and data preparation may take time.** Expect to spend up to several weeks or months to complete these tasks, thereby delaying the process of putting the DW appliance into full production. These necessary chores may also increase the total life-cycle cost associated with ripping and replacing an established DW solution with an appliance-based alternative.

- **Hands-on testing is needed to decide which approach and vendors support your requirements.** These requirements might include processing complex OLAP workloads, concurrent DW usage, fast bulk loading, workload optimization, and analytic-database scalability.³ Enterprises considering the appliance approach should perform extensive benchmark testing of alternative solutions, both from appliance-oriented solution providers and from established DW vendors.

Clearly, DW appliances are far more complex and costly an investment than the average toaster. For database administrators and other information and knowledge management professionals, the consequences of making the wrong choice in a DW solution — and implementing it sub-optimally — are far more dire than burnt toast.

RECOMMENDATIONS

FOCUS ON DW, DATA MART, AND ANALYTIC PRICE-PERFORMANCE REQUIREMENTS

For information and knowledge managers, it's critically important to focus on your core requirements and not be distracted by vendors' self-serving definitions of what constitutes a "true appliance." Use the following guidelines to sort through the field of appliance-based DW alternatives and to deploy these offerings effectively:

- **Consider the specific query workloads of your BI and DW applications.** Comparing diverse DW appliance solutions is difficult if you do not consider the types of query execution that you need them to execute. For example, brute-strength table scans are the forte of DW appliances that incorporate relational databases, such as solutions from Greenplum and Netezza. However, these table scans are sub-optimal if your BI application is doing mostly single-row lookups. In a similar vein, DW appliances that incorporate columnar databases — from vendors such as ParAccel — support a high degree of compression that is tuned to each column/attribute's specific data type. But columnar databases lose their efficiency advantage on queries that involve SELECT * operations across many columns.
- **Rethink your enterprise data management architectures.** Consider offloading as much CPU- and storage-intensive functionality to appliances as possible, as commercial appliance offerings become available. As borne out by case studies, cost-effective scaling of your DW, OLAP, and BI environments depends on accelerating complex queries and bulk data loading through pre-optimized solutions that were designed for these functions. But also consider the life cycle costs of duplicating data to multiple appliance-based data marts and having to manage these systems in a hub-and-spoke configuration.
- **Use the same core criteria to evaluate DW appliances as traditional DW approaches.** In other words, use those criteria that involve customer integration of database, servers, storage, and other components. Evaluate alternative DW deployment approaches, including rival

appliances, based on price-performance, functionality, flexibility, scalability, manageability, integration, and extensibility.

- **Determine the extent to which commercial appliances integrate with your environment.** Learn whether the appliance will easily snap on to your existing BI, DI, DQ, and DBMS environments or whether, to operate at full efficiency, it will require extensive modifications to existing enterprise software. For example, ask if you will need to make modifications to your SQL queries, analytic applications, and ETL scripts to integrate completely with a DW appliance.
- **Consider deploying DW appliances initially in tactical roles.** Find functions — such as standalone data marts for department- or function-specific data marts — in which these solutions can address your core OLAP and ETL requirements and integrate into your existing DBMS and BI environments. Proving out DW appliances in specific, quick-payback deployments will help you gain support for more broad-based enterprise deployments going forward.
- **Consider using DW appliances enterprisewide after trying them in tactical deployments.** While leveraging the appliances' native symmetric multiprocessing, grid, clustering, load balancing, and failover features, implement them in distributed and federated configurations and scale them to handle growing BI, DI, and DQ workloads.⁴
- **Demand that DW appliance vendors back up their claims.** Vendors should provide independent performance and availability benchmark results, explanations of proposed savings, detailed case studies, and evaluation models that you can run against your own workloads.
- **Determine if a DW appliance vendor has the R&D resources and partnerships to survive.** Continuing hardware price-performance improvements translate into ever-faster DW appliances. Established vendors such as HP, IBM, and Teradata will continue to scale their DW appliances by leveraging their unparalleled R&D resources and by making steady improvements to the server, storage, and interconnect components within their appliances.
- **If you're a user of traditional high-end DW products, continue to evaluate DW appliances.** Determine whether any of these offerings have the maturity and cost-effectiveness that might merit a migration toward this new prepackaged approach to building scalable DWs. Ask DW appliance vendors for reference customers, preferably those that have similar requirements to yours. Then follow up by contacting these customers to determine whether their experience with the vendor's DW appliance corroborates the vendor's grand claims.

WHAT IT MEANS

I&KM PROS WILL CONTINUE TO GRAPPLE WITH THE DW APPLIANCE VALUE PROPOSITION

Given the immaturity of the DW appliance market, the burden of proof is still on vendors to show that these pre-integrated offerings can scale in capacity and broaden in scope to support users' most demanding ETL, OLAP, BI, and other requirements. Even in those scenarios where DW appliances have a clear price-performance advantage — such as offloading high-volume query workloads from EDWs — the upstart platforms are primarily suitable for these function-limited deployments.

No DW appliance pure play can offer a commercial platform as scalable — in database capacity, mixed-workload support, concurrent usage, and fast data loading — or as flexible in deployment, optimization, and administration features as DW market leader Teradata, which has long had an appliance-like value proposition though it has never chosen to set sail under that marketing banner. Just as important, no DBMS vendor that has chosen to embrace the appliance approach — such as IBM, Oracle, and Microsoft — has demonstrated, when you factor in all hardware, software, database, migration, and professional services costs, that its appliance solution is any more cost-effective than traditional “bespoke” DW deployments. Indeed, the life cycle cost per-user, per-terabyte, or per-application of a feature-complete BI/DW deployment with an appliance platform at its heart may not necessarily be any lower than for an equivalent, custom-configured EDW deployment.

So, in that sense, the ROI for these nouveau solutions in enterprisewide deployments is still as soft and squishy as the industry's myriad, overstuffed notions of what a DW appliance is or is not.

SUPPLEMENTAL MATERIAL

Companies Interviewed For This Document

Business Objects, an SAP company	Netezza
DATAAllegro	Oracle
Dataupia	ParAccel
Greenplum	SAP
IBM	
Microsoft	

ENDNOTES

¹ Quick deployment is essential in many DW deployments. Leaders of DW initiatives have a limited amount of time to establish the planning processes, select the solution platform, and make the deployment decisions that will chart the course of the initiative's long-term success. See the February 13, 2007, "[The 10 Mandatory Steps For The First 90 Days Of A Data Warehousing Project](#)" report.

² Specific components of IBM Information Server Blade include WebSphere DataStage and WebSphere QualityStage. See the December 28, 2006, "[Information Server Takes IBM One Step Closer To Delivering Information-As-A-Service](#)" report.

Forrester evaluated leading enterprise extract, transform, and load (ETL) vendors across 68 criteria and found that IBM and Informatica maintain leadership positions in enterprise ETL thanks to their ability to scale and perform batch and operational data integration (DI) in complex environments, as well as to maintain a consistent focus on providing robust data management capabilities. Business Objects and Oracle have also emerged as Leaders with significant usability and scalability improvements, but they are still primarily used in data warehousing environments and have not been widely adopted for operational DI needs. SAS is a Strong Performer, but it remains most attractive as an integrated piece of a SAS BI platform. Ab Initio offers a highly scalable and configurable data processing platform, but its secretive corporate culture limits prospective customers' visibility into its strategy. iWay Software, Microsoft, Pervasive, and Sunopsis (acquired by Oracle) round out the Strong Performers best suited for a more targeted subset of DI professionals. As a new player in enterprise ETL, Sybase has some of the raw materials needed to develop a competitive solution, but it must integrate the tools it has acquired into a DI suite with a clear differentiation strategy to gain traction in this crowded market. See the May 2, 2007, "[The Forrester Wave™: Enterprise ETL, Q2 2007](#)" report.

³ Configuring DW environments to support highly scalable OLAP and BI applications against very large databases demands careful consideration of many technical deployment options, including attention to parallelization, partitioning, caching, indexing, compression, and materialized views. See the June 23, 2007, "[Data, Data Everywhere!](#)" report.

⁴ Distributed and federated DW deployment topologies, incorporating conformed dimensions, are a key approach for converging BI silos. See the February 15, 2007, "[Death to the BI Data Silo? Not So Fast](#)" report.

DW workloads may be distributed across several linked platforms to address the distinct loading, performance, and availability requirements of different classes of users for BI — power users, casual users, and business users, for example — with each class exhibiting unique query workloads. See the December 27, 2004, "[The Emergence Of Process-Centric BI](#)" report.

FORRESTER®

Making Leaders Successful Every Day

Headquarters

Forrester Research, Inc.
400 Technology Square
Cambridge, MA 02139 USA
Tel: +1 617.613.6000
Fax: +1 617.613.5000
Email: forrester@forrester.com
Nasdaq symbol: FORR
www.forrester.com

Research and Sales Offices

Australia	Israel
Brazil	Japan
Canada	Korea
Denmark	The Netherlands
France	Switzerland
Germany	United Kingdom
Hong Kong	United States
India	

*For a complete list of worldwide locations,
visit www.forrester.com/about.*

For information on hard-copy or electronic reprints, please contact the Client Resource Center at +1 866.367.7378, +1 617.617.5730, or resourcecenter@forrester.com. We offer quantity discounts and special pricing for academic and nonprofit institutions.

Forrester Research, Inc. (Nasdaq: FORR) is an independent technology and market research company that provides pragmatic and forward-thinking advice to global leaders in business and technology. For more than 24 years, Forrester has been making leaders successful every day through its proprietary research, consulting, events, and peer-to-peer executive programs. For more information, visit www.forrester.com.