



How to evaluate data warehouse appliances

Business Intelligence and analytics remain top priorities for organizations seeking to increase revenues and cut costs, but these systems often require a data warehouse to feed them the correct, relevant information at the right time. Data warehouse appliances, a combination of preconfigured hardware and software, have entered the market to help ease implementation. Yet, not all data warehouse appliances are alike and just because the hardware and software come together doesn't mean it's an easy decision. This E-Book provides detailed expert advice on how to evaluate, select and prepare your organization for a data warehouse appliance.

Readers will:

- Get important background on the technology, with an overview of the data warehousing appliance market and a brief history of its evolution.
- Read step-by-step advice from a data warehouse expert about how to evaluate and compare data warehousing appliances

Sponsored By: **SYBASE®**



How to evaluate data warehouse appliances

Table of Contents:

[Data warehouse appliances go mainstream](#)

[How to select data warehouse appliances: 10 steps for evaluation and comparison](#)

[Resources from Sybase](#)

Data warehouse appliances go mainstream

By Jeff Kelly, SearchDataManagement.com News Editor

Thanks to the promise of quick deployments with little to no customization, data warehouse appliances are making their way into the mainstream data center, according to a recent Forrester Research report.

Companies are increasingly turning to preconfigured data warehouse appliances for specific data management and analytic functions, such as bulk data loading and multidimensional OLAP queries, the report said. Cheaper and less complex than traditional enterprise data warehouses, data warehouse appliances are being touted by vendors as no-fuss, no-muss, plug-and-play alternatives.

"What all these solutions do, at heart, is accelerate the deployment and processing of data warehousing applications, specifically complex queries against structured data sets maintained in star schemas within your analytic databases," said James Kobielus, the report's author and senior analyst with the Cambridge, Mass.-based research firm. "They're all basically OLAP [and data mart] accelerators."

This is good news for information and knowledge management workers, who are increasingly expected to quickly create and deploy data marts to address short-term business intelligence (BI) demands, Kobielus writes in the report. Data warehouse appliances can be deployed in a fraction of the time it takes to get a data mart or enterprise data warehouse off the ground, he said—and often with less customization.

Essentially prepackaged hardware and software, appliances are nothing new to IT. Firewall appliances and WAN optimization appliances, for example, have been around for years. The idea is that one convenient package helps companies with limited technical expertise quickly solve a particular problem.

Data warehouse appliances first hit the scene in 2002 with the emergence of data warehouse appliance vendor Netezza Corp., Kobielus said. The Framingham, Mass.-based vendor, and those that followed—like San Mateo, Calif.-based Greenplum Inc.—"stressed the whole quick deployment, quick time-to-value, preconfigured hardware and software bundle to process data warehouse workloads," he said.

Since then, many of the so-called mega-vendors have gotten into the data warehouse appliance game, most notably IBM, which re-launched its enterprise data warehouse portfolio as data warehouse appliances in March 2007. Oracle and Microsoft followed soon after with data warehouse appliances of their own, developed through partnerships with storage and hardware partners. As a result, Kobielus said, data warehouse appliances have come a long way in just a few short years.

"The recent embrace of the appliance go-to-market approach by the three top

"What all these solutions do, at heart, is accelerate the deployment and processing of data warehousing applications, specifically complex queries against structured data sets maintained in star schemas within your analytic databases."

James Kobielus, senior analyst, Forrester Research

enterprise DBMS vendors—IBM, Microsoft and Oracle—shows that appliances, however defined, are becoming the dominant approach for delivering [data warehouse] functionality to customers of all sizes," the report states.

Data warehouse appliances maturing; can't yet match enterprise data warehouse functionality

But most data warehouse appliances are not as robust as traditional enterprise data warehouses, which are still needed for complex BI and analytics functions. Enterprise data warehouses are scalable to manage increasing data volumes and can process multiple data domains simultaneously, Kobielus said. Data warehouse appliances, on the other hand, are much less complex, designed to work with relatively stable data volumes and usually just one type of data domain—such as customer, human resources or product data—to address tactical data mart requirements. Because they are less customizable, appliances are not flexible enough to support mixed-query workloads, he said.

Also, data warehouse appliances are not truly plug-and-play, Kobielus said. They are significantly faster to deploy than enterprise data warehouses, sometimes implemented in a week or two, but some customization is inevitably needed with appliances.

"When you plop them down into your data center and connect them to your existing ETL code and your existing BI application...more likely than not, you're going to need to make some modifications to those ETL scripts. You're going to have to rewrite your SQL queries inside your BI application to be able to work most efficiently," Kobielus said. "It's not a huge issue, but nonetheless, nothing's really plug-and-play."

Kobielus expects data warehouse appliances to continue maturing into enterprise-grade data warehouse platforms, but he doubts they will replace enterprise data warehouses altogether.

"Basically, one can easily come at this whole topic with the misconception that it's an either/or thing, that you've got an enterprise data warehouse or you've got a data warehouse appliance, and the two shall never meet," Kobielus said. "In fact, that's far from the truth."

Data warehouse appliance evaluation and buying advice

The report offered the following recommendations for companies considering data warehouse appliances:

- Consider BI and data warehouse applications' query workload needs and whether a data warehouse appliance can truly meet them.
- Use the same criteria to evaluate data warehouse appliances and traditional enterprise data warehouses.
- Deploy data warehouse appliances first in tactical roles, such as in function-specific data marts, then roll them out enterprise-wide.

Kobielus also warned companies considering data warehouse appliances not to be confused by the marketing terminology. Some vendors, like IBM and Teradata, have embraced the term "appliance" while others -- Kobielus declined to name them—have resisted, fearing that the term makes their products sound more lightweight or less robust than they truly are.

"[Either way] you'll know when an appliance is right for you when the appliance-based solution addresses all of your data warehousing requirements, irrespective of what they happen to go to market under in terms of banner or buzzword," Kobielus said.



SMARTER ANALYTICS.

TOMORROW'S ANSWERS GO FAR BEYOND YESTERDAY'S TECHNOLOGY.

Any user, any query, any time. Begin with those criteria, then add unlimited headroom for data and users, fast and accurate answers to complex queries, incremental scalability to grow and adapt, the freedom to leverage standard hardware and operating system, and the flexibility to choose your reporting and analytics tools. Next come faster, more accurate answers to complex queries, unbounded reporting, deep-dive data mining and predictive analytics, followed by insight-driven perspective into risks, opportunities, and rewards. Now you have information-based differentiation from high-performance business analytics proven in over 3,100 unique installations at 1,700+ companies.

It takes a smarter analytics platform to power the new business reality. It takes Sybase IQ.

www.sybase.com/bi

How to select data warehouse appliances: 10 steps for evaluation and comparison

By Mark Whitehorn, SearchDataManagement.com contributor

Data warehouse appliances have become popular in recent times, with the inevitable consequence that as more people jump aboard the bandwagon, the definition becomes wider and wider. For example, some of them can be used as mini data warehouses complete with limited data integration capabilities. However, the majority are far less adaptable and more function specific -- some highly so, capable of sucking data from only one precise data source (perhaps a particular finance application) and offering no integration functions at all. This diversity means that the most important question to answer when trying to select a data warehouse appliance is:

What exactly *is* your problem?

By determining what problems you hope the data warehouse appliance will solve, you have taken the biggest step toward reducing the number of products to consider. So if, for instance, your problem revolves around an existing system that holds data in a form resistant to analysis, you need look at only those appliances that work with your intransigent source.

Here are 10 "dos and don'ts" to help you narrow the field further.

1. Do perform analytical analysis.

Even if you've determined that you want the appliance to work with data from a particular source, before you talk to vendors, you should sit down with the potential users of the system and find out what kinds of analysis they wish to perform. This should be undertaken as a formal process, which creates a detailed specification of the required analytics.

2. Don't let the salespeople show you only the analyses they have in their canned demos.

A pre-built demonstration of analysis can border on the worthless -- who cares whether the data warehouse appliance can do what *they* want it to do? Whether it'll do what *you* want is the important issue. Now, this may seem a little hard-line, and there is nothing wrong with keeping an open mind. Indeed, the demo may show aspects of analysis that you or potential users of the system never dreamt were possible. But it is very important to keep focused on the problems you want it to solve rather than on the glitzy demo.

3. Do prepare a set of data before talking to vendors.

Create a set of your own data that you wish to analyze. Make details anonymous and remove any commercially sensitive material, if necessary.

4. Don't cleanse it.

Do not cleanse your test data set. True data warehouses have an extensive set of custom-built ETL (extract, transform and load) routines to cleanse data. Data warehouse appliances often lack these. Most data in source systems isn't clean, so it is instructive to ask how the data warehouse appliance is going to address this problem: Does it have any cleansing capabilities or will the dirty data just end up as inaccurate information that misleads your users? At this point, it is also worth testing to see whether the appliance can support multiple definitions of, say, "customer." Different users often wish to analyze on different definitions: new customers, important customers, established customers, and so on.

5. Do ask to see a full demonstration.

It is, after all, a data warehouse appliance, and it should be easy to set up, configure and get working. Ease of use is probably one of the main reasons why you're considering buying an appliance. This shouldn't be a problem to any vendor. If it is, you are entitled to ask whether they are offering a true data warehouse appliance.

6. Don't be too impressed when it works like greased lightning.

This may also sound a bit too hard-nosed, but it is a classic gambit for vendors to ask for your analytical query from hell and then to run it for you at the speed of light against your test data. If they can do it, then of course that is a good result. But it is also worth remembering that, until now, you have probably been running that query on a transactional system. The data in a transactional system is structured (as the name implies) to optimize transaction processing, not analytics. In addition, that transactional system is likely to be several years old and may have been bought when memory cost real money, rather than being given away free with your breakfast cereal, and when CPUs were slow. So of course your analytical query from hell is likely to be extremely fast when it is run on the data warehouse appliance. There'd be something wrong if it weren't.

A data warehouse appliance should be easy to set up, configure and get working. Ease of use is probably one of the main reasons why you're considering buying an appliance. This shouldn't be a problem to any vendor."

7. Do find out whether the users like the interface.

Perform some user testing to find out whether users like the look of the application and find it easy to use. It's an appliance, so it's supposed to be intuitive and straightforward to use.

8. Don't expect the data warehouse appliance to integrate.

A data warehouse appliance is unlikely to integrate well with other systems. This might not be a problem; indeed it is somewhat unfair to expect a data warehouse appliance to integrate with other BI systems -- its role in life is to address a specific need, not to become part of a far-reaching empire.

9. Do study and test the scalability.

Take an interest in scalability. Data sets generally grow, so you might consider taking your current largest data set and doubling it up for testing. Many data warehouse appliances are based on massively parallel processing (MPP) architecture, and this is particularly well suited to delivering scalability. Incidentally, you'll notice that this "do and don't" list does not focus on whether you should buy an MPP appliance or not -- the important issue here is functionality. Attractive as the nitty-gritty technicalities are, you should not be buying a data warehouse appliance because it's MPP but because it provides the functionality and scalability your users need.

10. Don't forget to negotiate.

Especially in a slow economy, there are many vendors out there and they want/need your business. For those data warehouse appliances that are built on commodity components, the main cost of the appliance is unlikely to be the hardware; it will be those somewhat more intangible elements, the software and the support. Intangibles are always up for negotiation, and in a recession you have more scope -- so negotiate hard.

Resources from Sybase

SYBASE®

[White paper: Sybase IQ Technology Overview – Bloor Research](#)

[Techcast: Sybase IQ 15: More Flexibility to Adapt and Grow](#)

[Webcast: Sybase IQ 15: Smarter Analytics](#)

About Sybase

Sybase is an industry leader in delivering enterprise and mobile software to manage, analyze and mobilize information. We are recognized globally as a performance leader, proven in the most data-intensive industries and across all major systems, networks and devices. Our information management, analytics and enterprise mobility solutions have powered the world's most mission-critical systems in financial services, telecommunications, manufacturing and government. With Sybase, enterprises can manage the high volumes and variety of data, analyze this vital information and mobilize it so people can conduct business, regardless of where they are located or what devices they are using. And since all Sybase solutions are built using open standards, organizations can leverage the IT investments they depend on today, along with the ones they'll need tomorrow.