

Gartner Identifies Nine Key Data Warehousing Trends for the CIO in 2011 and 2012

Key Data Warehousing Issues to be Discussed at Gartner Business Intelligence Summits, February 22-23 in Sydney, and May 2-4, Los Angeles

STAMFORD, Conn., February 9, 2011— CIOs need to familiarize themselves with nine key trends in data warehousing and how they will impact the cost-benefit balance of technology deployed to deliver business analytics value during 2011 and 2012, according to Gartner, Inc.

Gartner analysts said the data warehouse is set to remain a key component of the IT infrastructure and believe that, as the demand for business intelligence (BI) and the wider category of business analytics increases, optimization, flexible designs and alternative strategies will become more important.

"The data warehouse remains one of the largest — if not the largest — information repository in the enterprise," said Mark Beyer, research vice president at Gartner. "Only by being aware of the key market trends and how emerging technology solutions will blend with proven practices can the CIO avoid budget waste through 'misdirection' by the data warehouse management and delivery team."

Gartner has identified nine key trends in the data warehousing market for 2011 through 2012:

Optimization and Performance

Advanced functionality for hardware management of input/output (I/O), disk storage and CPU/memory balancing are now included almost as a matter of course in data-warehouse-capable platforms. Some new entrants are focusing on optimization as a differentiator and nearly every data warehouse vendor is now addressing the issue of optimizing storage for the warehouse via compression and usage-based data placement strategies. Vendors are also expending great effort differentiating their products on performance claims and technology, in ways that are not necessarily significant to the use case.

Data Warehouse Appliances

Although there are many reasons why organizations consider buying an appliance, the main reason is simplicity. The vendor builds and certifies the configuration, balancing hardware, software and services for a predictable performance. The appliance is delivered complete and installs rapidly. If there are any problems, a single call to the appliance vendor is the first course of action. There is a secondary effect as well, in that appliances can speed delivery by avoiding time-consuming hardware balancing.

The Intensive POC

During 2010, most organizations heeded Gartner's advice to perform a proof of concept (POC) with a "shortlist" of vendors during the selection phase of the data warehouse database management system (DBMS). Gartner recommends that POCs use as much real source-system extracted data (SSED) from the operational systems as possible, while performing the POC with as many users as possible, creating a data warehouse workload that approaches that of the environment to be used in production.

Data Warehouse Mixed Workloads

There are six workloads that are delivered by the data warehouse platform: bulk/batch load, basic reporting, basic online analytical processing (OLAP), real-time/continuous load, data mining and operational BI. Warehouses delivering all six workloads need to be assessed for predictability of mixed workload performance as failing to plan for mixed workloads will lead to increased administration costs over time, as volume and additional workloads are added, potentially leading to major sustainability issues.

The Resurgence of Data Marts

A data mart is defined as an application-specific analytic repository of any size, normally with a specific, smaller group of users than a data warehouse. Data marts can be used to optimize the data warehouse by offloading part of the workload to the data mart, returning greater performance to the warehousing environment.

Column-Store DBMSs

Column-store DBMSs generally exhibit faster query response than traditional, row-based systems and can serve as excellent data mart platforms, and even as a main data warehouse platform. Gartner foresees several vendors changing the pricing model for the software from a more traditional per-user or per-core model to a price based on the volume of data loaded into the database.

In-Memory DBMSs

In-memory DMBS (IMDBMS) technologies exhibit extremely fast query response and data commit times and introduce a higher probability that analytics and transactional systems can share the same database. Analytic data models, master data approaches and data services within a middle tier will begin to emerge as the dominant approach, forcing more traditional row-based vendors to adapt to column approaches and in-memory simultaneously. BI solutions will emerge sooner rather than later, and these will leverage IMDBMSs with superior-performing products and will quickly become acquisition targets for megavendors.

Data Warehouse as a Service and Cloud

In 2011, data warehouse as a service comes in two "flavors" — software as a service (SaaS) and outsourced data warehouses. Data warehouse in the cloud is primarily an infrastructure design option as a data model must still be developed, an integration strategy must be deployed and BI user access must be enabled and managed. Private clouds are an emerging infrastructure design choice for some organizations in supporting their data warehouse and analytics.

Using an Open-Source DBMS to Deploy the Data Warehouse

Open-source DBMSs are still being used in both experimental and more formalized approaches. At this point, open-source warehouses are rare and usually smaller than traditional ones and also generally require a more manual level of support. However, some solutions are optimized specifically for data warehousing.

Additional information is available in the Gartner report "Data Warehousing Trends for the CIO, 2011-2012." The report is available on Gartner's website at <http://www.gartner.com/resId=1530314>.

About Gartner Business Intelligence Summits

The Gartner Business Intelligence Summits provide the insights, frameworks and best practices to implement BI, analytics and performance management initiatives that drive better decision-making and business performance across the organization. Through a unique mix of analyst and peer interaction, it helps make the case, drive adoption, create and sustain successful BI strategies.

The Gartner Business Intelligence & Information Management Summit in Sydney takes place February 22-23 at the Sydney Convention & Exhibition Centre. Additional information is available at www.gartner.com/ap/bi. Members of the media can register for the event by contacting Susan Moore at susan.moore@gartner.com.

The Gartner Business Intelligence Summit in Los Angeles is being held May 2-4 at the Hyatt Regency Century Plaza. Additional information is available at www.gartner.com/us/bi. Members of the media can register for the event by contacting Christy Pettey at christy.pettey@gartner.com.

Additional information from the events will be shared on Twitter at http://twitter.com/Gartner_inc using #GartnerBI.