INSIGHT

SAP Business Intelligence Accelerator: A High-Performance Analytic Engine for SAP NetWeaver Business Intelligence

Dan Vesset

IDC OPINION

The business analytics market continues to grow at a healthy pace as more organizations embrace fact-based decision making. The two primary drivers of the demand for business analytics solutions are increasing data volumes and end-user populations. These drivers place new performance requirements on existing analytic platforms. To address some of these requirements, SAP has introduced SAP BI accelerator functionality with SAP NetWeaver — deployed as an appliance combining software and hardware to increase the performance characteristics of SAP NetWeaver BI deployments. With this introduction, SAP has the opportunity to:

☑ Expand the use of its analytic server technology

☑ Reach out to its existing customer base of SAP NetWeaver BI users to help them improve the query processing and data compression performance characteristics of their business analytics solutions

IN THIS INSIGHT

This IDC Insight evaluates the potential impact on the business analytics (BA) market of the SAP NetWeaver Business Intelligence (BI) accelerator appliance. Announced first in cooperation with HP and Intel at the SAP Sapphire Conference in Boston in July 2005, SAP announced a second configuration of the BI accelerator with IBM and Intel in September 2005. This Insight will discuss the role BI accelerator can play in the market in promoting the use of SAP NetWeaver as a platform for business analytics solutions.

SITUATION OVERVIEW

Market Definition

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. This software market includes both application development tools and packaged analytic applications.

The worldwide business analytics market stood at $15 billion in 2004 and is expected to grow at a compound annual growth rate of 12–15% through 2009. The market continues to enjoy robust growth driven by the competitive need of organizations to make better — faster and more insightful — decisions and by ongoing regulatory compliance requirements.

Primary Market Trends

As business analytics gain more mainstream acceptance, two primary trends are driving the technical requirements of software platforms to support decision making. These trends include:

☑ Growth in the number of end users:

❑ Internal access to business intelligence data is being expanded beyond the analyst/power user community to include all knowledge workers.

❑ Externally, IDC research shows that already 44% of large companies provide business intelligence reports to external users, such as suppliers, customer, partners, and other stakeholders.

☑ Growth in data volumes:

❑ Thirty percent of companies with $500 million or more in revenue expect their data warehouses to growth at least 100% over the next three years.

❑ Another recent IDC survey showed that 30% of organizations use intraday or continuous data feeds for their business analytics solutions, suggesting that refresh rates for data warehouses and marts are increasing.

☑ Shorter decision cycles requiring more flexibility:

❑ Decision-making flexibility comes from the ability to perform continuous "what-if" analytics and ongoing scenario planning. It is not enough to just deliver information faster to decision makers in the form of reports, dashboards, or alerts. Just as important is support for evaluating decision alternatives.

❑ To support these decision-making tasks, IT infrastructure must be able to support rapid prototyping and more flexible ad hoc querying. Traditional data warehousing and OLAP platforms with their aggregations and materialized views have been too rigid to provide such support.

Market Shortcomings and Challenges

Only 15% of managers agree with the statement that most reports developed in their organizations deliver the right data to the right people at the right time (source: *IDC Business Intelligence Survey*, May 2005).

At the same time, business analytics deployments are becoming more operational. For example, 40% of organizations indicate that their business intelligence solutions can be down for more than a few hours without causing significant disruption to ongoing operations. This is a relatively new phenomenon on the heels of a long-time practice where BI was viewed as a standalone and often "nice-to-have" system that didn't have direct operational impact.

To address these requirements and shortcomings, organizations are increasingly beginning to look for alternatives to traditional business analytics platforms that improve the scalability of such platforms. One such solution is SAP NetWeaver with its recently introduced BI accelerator.

SAP Business Intelligence Accelerator

SAP BI accelerator presents itself like an appliance because it combines in one package software and hardware. It's important to note that organizations looking to deploy the SAP BI accelerator will still need to deploy an instance of the SAP NetWeaver with its own associated software and hardware. 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.

In essence, 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 BI accelerator highly scalable is the use of SAP's TREX search technology in conjunction with blade server architecture provided by its hardware partners.
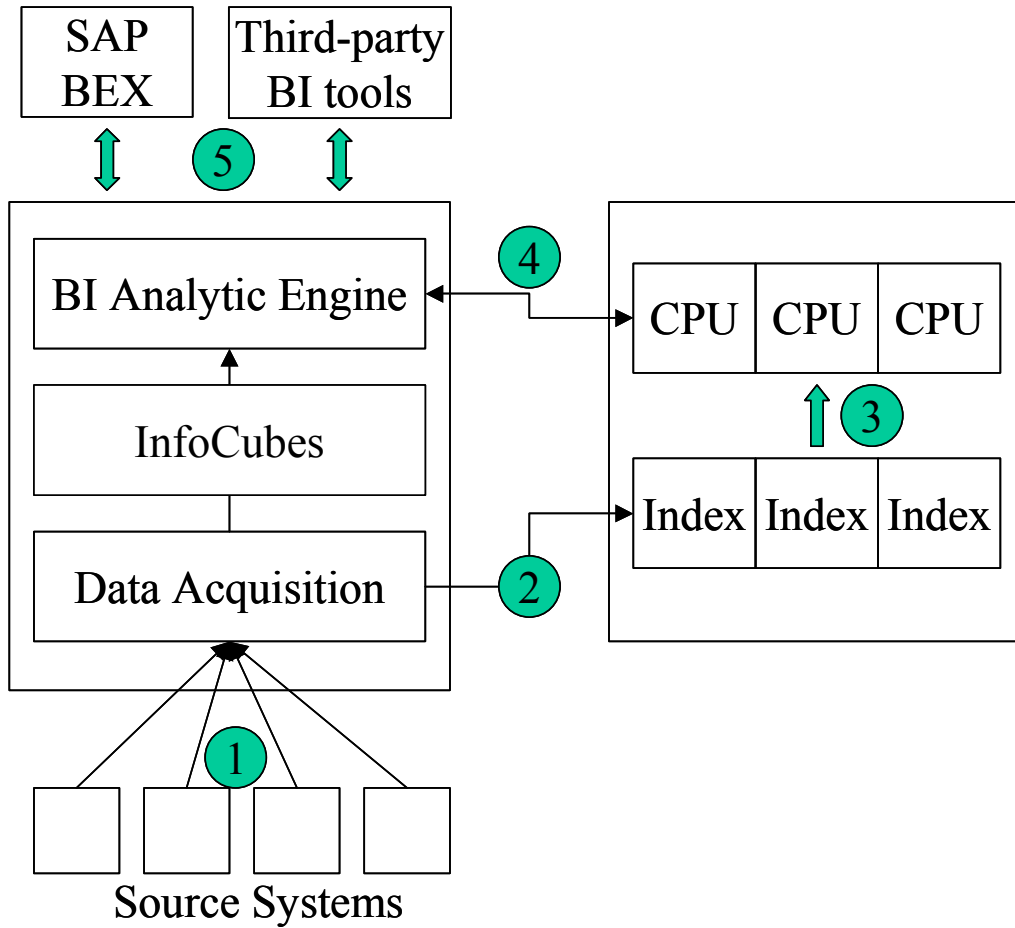
SAP BI Accelerator Architecture

A high-level BI Accelerator architectural diagram is shown on the right side of Figure 1.

SAP NetWeaver BW and SAP BI Accelerator Architecture

SAP NetWeaver BI   SAP BI accelerator

```
  ┌─────────┐  ┌──────────────┐
  │  SAP    │  │ Third-party  │
  │  BEX    │  │  BI tools    │
  └─────────┘  └──────────────┘
       ↕       (5)      ↕
  ┌──────────────────────┐        ┌──────────────────────────┐
  │  ┌────────────────┐  │  (4)   │  ┌─────┬─────┬─────┐      │
  │  │ BI Analytic    │◄─┼────────┼─►│ CPU │ CPU │ CPU │      │
  │  │    Engine      │  │        │  └─────┴─────┴─────┘      │
  │  └────────────────┘  │        │           ↑ (3)           │
  │  ┌────────────────┐  │        │  ┌───────┬───────┬───────┐│
  │  │   InfoCubes    │  │        │  │ Index │ Index │ Index ││
  │  └────────────────┘  │        │  └───────┴───────┴───────┘│
  │  ┌────────────────┐  │  (2)   │                           │
  │  │ Data           │──┼────────┼─►                         │
  │  │ Acquisition    │  │        └──────────────────────────┘
  │  └────────────────┘  │
  └──────────────────────┘
      (1) │ │ │ │
  ┌──┐ ┌──┐ ┌──┐ ┌──┐
  └──┘ └──┘ └──┘ └──┘
```

Source Systems

Source: IDC, 2005

SAP BI accelerator includes indexes that are vertically inverted reproductions of all the data included in InfoCubes (i.e., fact and dimension tables as well as master data). Note that there is no relational or other database management systems in BI accelerator. There is only a file system, and indexes are essentially held as flat files.

The second primary component of SAP BI accelerator is the engine that processes the queries in memory. The software is running on an expandable rack of blade servers. The operating system used for BI Accelerator is 64-bit Linux, so in addition to having no database license cost, there is also no OS license cost.

SAP BI Accelerator at Work

1. Data is loaded from source systems into an SAP InfoCube.

2. An index is built for this InfoCube and stored inside the BI accelerator appliance. These are search engine indexes built using SAP's TREX search technology. They are stored in a file system (not a database system) using vertical decomposition (a column-based approach as opposed to the row-based approach that requires more read time). This results in highly compressed data sets that further contribute to fast processing speeds.

3. BI accelerator indexes are loaded into memory where the query is processed. In memory, joins and aggregations are done at run time. Loading of indexes into memory happens automatically at first query request, or it can be set for preloading whenever new data is loaded.

4. At run time, query requests are sent to the analytic engine, which reroutes the query to the BI accelerator.

5. Query results are returned to the end-user application.

Note: Steps 1 through 3 above typically are performed offline, (e.g., during less critical times). Steps 4 and 5 below are executed at actual query time.

SAP BI Accelerator Query Processing Steps

To describe how queries are processed by SAP BI accelerator, it is first worthwhile to describe how they are processed within the traditional SAP NetWeaver BI architecture. The left side of Figure 1 depicts this SAP NetWeaver BI architecture. In this case, the steps are:

1. Query is launched from SAP Business Explorer (BEX) or a third-party BI tool.

2. Query evaluates whether there is a precalculated data set (usually calculated during off-hours). If one exists, the query retrieves data from that data set.

3. If a precalculated template does not exist, the query checks the OLAP Cache for the necessary data. The OLAP Cache doesn't benefit the first person launching the query, but will benefit all subsequent requests for that same query.

4. If the required data does not exist in the OLAP Cache, then the query looks for aggregate tables or materialized views that may exist. These preaggregated views are not as fast as processing the query against precalculated data sets or OLAP Cache, but they are still faster than going against the final layer, the InfoProvider.

5. The final option to execute the query is to run it against the InfoProvider, in this case the SAP InfoCube. This results in the slowest processing times as compared to the other three choices listed above.

Executing the same query with SAP BI accelerator results in a somewhat different set of steps:

1. The first three steps remain the same (i.e., the query first checks if there is a precalculated data set and then checks the OLAP Cache). These two sources still result in the best processing performance.

2. However, if neither of these options exists, the query checks if the BI accelerator option is available. In this case, the slowest options of using aggregates or the InfoCube itself are eliminated. Instead, the BI accelerator processing as described above kicks in.

Note that early tests by SAP suggest that the query-processing speed with BI accelerator is approximately the same as that with the OLAP Cache. However, the BI accelerator also provides additional flexibility and scalability. This suggests that even the planning and maintenance associated with OLAP Cache could disappear in deployments with the BI accelerator appliance. Although SAP itself does not make this recommendation at this time, if the processing times are indeed going to be proved to be the same, it seems that using the OLAP Cache loses its value.

Deployment

SAP BI accelerator is optionally deployed in conjunction with SAP NetWeaver BI. In other words, it is not a required component of the overall SAP NetWeaver BI solution. Rather, it is a scalable analytic processor that can be deployed where appropriate, depending on the specific end-user concurrency and data volumes within each organization.

In that case, the BI accelerator appliance is shipped directly from one of SAP's preferred hardware vendors (currently these include IBM or HP), with the BI accelerator software preloaded. There is a requirement that the software has been licensed directly from SAP (as part of SAP NetWeaver).

Competition

SAP BI accelerator appliance includes innovative use of existing, albeit not new technology. For example, search vendors such as FAST Search & Transfer, ClearForest, and Endeca, among others, are increasingly putting emphasis on using their technology in a business intelligence context.

On the appliance side of the market, companies such as Netezza, Datallegro, and Calpont have introduced true data warehousing appliances that combine software and hardware components and are specialized for analytic processes.

The column-based or vertical decomposition approach used by SAP for data compression within BI accelerator is also not new, it has been employed by Sand Technologies and Sybase for some time and more recently by Clareos.

Other high-performance analytic servers come from companies such as Greenplum and Kognitio that bill their offerings as virtual appliances, which means that their software components are similar to those of Netezza and Datallegro, but are not prepackaged with the hardware.

Finally, two companies that take yet another approach to data management and have provided evidence of high-performance analytic engines and high data-compression rates are QlikTech and ClearPace.

Further detail on these vendors will be available in an upcoming study in October 2005.

Benefits and Shortcomings of SAP BI Accelerator

Benefits

☑ Faster query processing and response time

☑ Faster load times, as aggregate change runs due to master data changes are handled by the BI accelerator rather than on top of InfoCubes

☑ Lower maintenance costs:

- ❑ BI accelerator eliminates the need to create relational aggregates.

- ❑ BI accelerator may eliminate the need to deal with an OLAP Cache.

- ❑ BI accelerator may decrease the need for logical partitioning on the NetWeaver BI side. However, there are other benefits beyond improving query processing speeds to having logical partitions.

- ❑ BI Accelerator results in less planning and tuning on the part of DBAs.

☑ Attractive packaging as an appliance that is preconfigured for analytic processing using SAP software and partners' hardware, which allows nonintrusive implementation

☑ High potential scalability — as demands grow, system scales up by adding blades

Shortcomings

☑ Currently the data source for BI accelerator can only be an SAP InfoCube. It does not work with other SAP data sources such as ODS.

☑ There's currently a one-to-one relationship between an instance of SAP NetWeaver BI and a BI accelerator. Sharing of multiple SAP NetWeaver BI instances with a single BI accelerator is not yet supported.

◻ There is currently no failover mechanism for BI accelerator (i.e., if the system goes down, the indexes will have to be rebuilt from beginning). However, this is unlikely to be a very time-consuming task due to the search paradigm involved.

◻ SAP BI accelerator is still in fairly early stages of delivery and will require further proof points and more precise published metrics on the actual performance improvements. Having said this, as of today, SAP claims that already over 50 million query tests have been performed over the last six months, including several pilot projects with actual customers.

FUTURE OUTLOOK

Although specific technologies employed by SAP in BI accelerator, such as search indexes, column-based (or vertically inverted) data structures, and blade server architecture, are not new, SAP has brought them all together into a cohesive solution that has the potential to make a significant impact on the business analytics market. Large companies, whether in IT or other industries, are rarely "bleeding edge" innovators when it comes to new technology. However, they are unquestionably innovators in the deployment and processes improvements around such technologies.

With BI accelerator, SAP has the opportunity to extend the scalability of its NetWeaver BI offering that is already deployed broadly, especially within its vast user base. This step in turn will increase the adoption of and satisfaction level with business analytics solutions and set yet another roadblock in the way of SAP's competitors in attracting business analytics market share away from the company.

In this scenario, in the highly competitive market for business analytics, it is the end users who will benefit from SAP raising the bar with its new BI accelerator solution.

Copyright Notice

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