white paper

Sybase® IQ, A Competitive Analysis

Brought to you by Think88
September 2010
Introduction

With the release of the pioneering Sybase IQ more than 10 years ago, Sybase originated the column-store analytic and data warehouse marketplace. In production with more than 1800 customers supporting more than 3,200 projects and with numerous instances of enormous data volumes under management (including holding the Guinness record for its 1-petabyte data warehouse), Sybase IQ has proven to be a well-regarded, highly successful solution. Sybase IQ’s ability to consistently deliver high performance and scalability has been key differentiators with customers such as the TransUnion, AdOn Networks, and BNP Paribas Securities.

In the past few years, a number of enabling technologies and customer-driven information processing trends have led to a significant expansion in the number of suppliers serving the analytic and data warehousing marketplace. First, in tandem with the open source movement, commodity hardware and operating systems have greatly expanded the information management options for organizations of all sizes. Virtualization is also playing a significant role in helping to derive added value from investments in technology. The net result of all these innovations has been a dramatic increase in data processing capabilities and storage volume capacity at a constant or even falling price.

As quickly as processing power and storage capacity grow, enterprises in industries as diverse as government, financial

Performance Considerations

Earlier generations of analytic and data warehousing solutions were commonly deployed in support of a relatively small number of non-mission-critical applications, such as business intelligence and reporting. While the responsiveness of these implementations was certainly important, day-to-day organizational operations were generally unaffected by the amount of time it took to extract intelligence from these systems.

The situation is radically different today: usage of these solutions has notably expanded beyond the original business intelligence and reporting missions to buttress a much broader range of operational activities. Furthermore, the information provided by modern analytic and data warehousing implementations is needed to provide guidance on running the enterprise’s operations in real-time. Thus, degraded responsiveness is a direct contributor to disappointing operational results.

Combining these realities with the expectations of today’s users for instantaneous access to information, and it’s easy to see that it’s vital to ask the right questions when procuring this type of technology. To help you manage this significant responsibility, we present a series of performance-related guidelines in two major categories: mixed workload support and scalability.
services, telecommunications, and media have kept pace by generating and managing massive amounts of new data. Created by internal personnel and applications as well as customers and partners, all of this additional knowledge is leading to decidedly larger quantities of data under management — often orders-of-magnitude greater than in the recent past. To make matters more challenging for IT leadership, user requirements and expectations are growing even faster than all of the new sources of information. The speed at which business is conducted continues to accelerate, with all of the commensurate competitive pressures. Finally, to help them do their jobs, entirely new classes of users want access to all of this knowledge, wherever and whenever they want it, including in the office, from home, and on the road.

Naturally, this expanding and dynamic market opportunity has attracted many new entrants, from established database vendors to newer, niche providers. Each of these suppliers is attempting to solve the particular challenges of analytics and data warehousing through their own set of techniques. To help the reader gain further insight into this ever-changing landscape, this paper will explore a series of high-level best practices for analytics and data warehousing based on Sybase’s extensive experience with successful implementations. We’ll also contrast these recommendations with observations of less effective strategies.

This paper is intended for both IT and operational management. In particular, we’ll provide suggestions in three essential categories: performance, cost control, and ease-of-deployment. Each dimension will be further broken down into a number of important sub-topics. We’ll also summarize our observations in a collection of tables.

Mixed workload support

Deploying an analytics and data warehousing solution entails a significant amount of money, time, personnel, and equipment. If established properly, these solutions have the potential to provide valuable intelligence for many different uses, which can deliver dramatic advantages for the entire organization. However, many enterprises make the mistake of deploying a niche, point solution that is missing key capabilities. The results are often underwhelming, and lead to users and management believing that the investment wasn’t worth the effort. To counteract this unfortunate outcome, it’s critical that the selected solution is capable of efficiently addressing a broad range of workloads, including:

- Fixed, pre-determined queries
- Dynamic, ad-hoc access
- Analytics and business intelligence
- Reporting
- Reporting

To ensure that your selected solution will be able to cope with the mixed workloads found in today’s analytic and data warehousing applications, ask the following questions:

- **Was the product specifically designed for data warehousing and analytics?** In contrast with other packages that are simply repurposed online transaction processing (OLTP) offerings, Sybase IQ was expressly constructed to address the unique challenges found in this domain, which are very different than the requirements for a successful OLTP solution.
Is it capable of incremental scaling? Sybase IQ manages its resources carefully, making it possible to gracefully support a variety of workloads while coaxing as much performance as possible from your hardware resources. We’ll discuss the full topic of scalability in detail a little later, but for now it’s important to note that Sybase IQ makes it easy to scale up or out for scenarios where the workloads exceed capacity.

Have the data integration and loading processes been optimized? Sybase IQ was designed to work with a broad range of information, including diverse data sources and unstructured data. It’s able to consume data from active, operational systems via source data transfer and very large table /multi-table loads. The data loading process may also be launched from a client.

Since enterprises are increasingly mandating that data warehousing platforms provide answers to ad-hoc queries, it’s essential that these solutions be flexible enough to return results for these types of inquiries as quickly as possible. Unfortunately, a number of other entrants in this marketplace have experienced challenges handling the dynamic nature of ad-hoc queries. These technologies include IBM DB2, Oracle 11g and Exadata, Microsoft SQL Server, Teradata, Vertica, and Greenplum. A number of product design and architectural factors contribute to these limitations, including:

- Products that are targeted at simple data models and elementary decision support inquiries
  - Indexes must be carefully designed to deliver acceptable performance
  - More complex queries must be tuned, which is not possible in an environment featuring numerous ad-hoc requests
- A focus on returning results on cache-based data
  - Disk-based information access may be dramatically slower
- Difficulty supporting multiple concurrent users

<table>
<thead>
<tr>
<th></th>
<th>IBM DB2</th>
<th>Oracle 11g</th>
<th>Oracle Exadata</th>
<th>Microsoft SQL Server</th>
<th>Teradata</th>
<th>Vertica</th>
<th>Greenplum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complex queries must be tuned</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Requires careful index design</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Few sites with large numbers of concurrent users</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Limited to simple schemas</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Table 1: Causes of ad-hoc query issues

An analytic/data warehousing solution is only as good as the data that it contains. The data loading process needs to be as efficient as possible. Otherwise, relevant information may be missing (or arrive too late), and decisions may be made based on stale or incorrect data. Both IBM DB2 and Greenplum impose intricate indexing requirements, which degrade responsiveness for this essential task.
Scalability

When first introduced, data warehousing solutions were commonly deployed as departmental data marts. This had the unfortunate side effect of exacerbating the already-present phenomenon of detached islands of information, inaccessible to large parts of the business. The situation is quite different now: many mid and large-sized enterprises are centralizing their analytic and data warehousing platforms, viewing them as strategic differentiators, and making them available to broad swaths of users working in different departments.

Now that these technologies are being deployed for the entire organization, it’s essential that they be able to scale to support massive amounts of data and hundreds-to-thousands of concurrent users. This added scalability also needs to be affordable; budgets for additional hardware and software continue to be constrained. Scalability-oriented product evaluation questions include:

- **Has the solution been validated by independent benchmarks?** Sybase IQ has excelled in the TPC-H benchmarks. In some of the recently published numbers in the 100 GB, 1TB, and 3TB categories with popular SMP configurations, Sybase IQ has consistently figured among the top vendors in performance measures.

- **Does the platform attempt to squeeze as much performance from its hardware?** Sybase IQ caters to real-world environments and uses its parallel facilities (via Multiplex) to support multiple, concurrent queries rather than focusing on a fixed number of pre-determined queries. This means that performance degradation is negligible as additional concurrent users are added.

- **Can scalability be attained and managed by administrators via an easy-to-understand user interface?** Sybase IQ provides unmatched resources for managing scale with intuitive, graphical grid control icons for simple right-click operations and extensive diagnostics-gathering tools for quick turnaround in situation analysis and timely resolution.

As a side benefit, the shared-disk Multiplex approach followed by Sybase IQ allows for storage and compute resources to be scaled independently. Finally, in contrast with other products that impose a heavy operating system burden, it requires only one operating system process per database.

Other technology providers have placed varying amounts of emphasis on scalability, often leading to offerings that aren’t able to support the amount of data and process-intensive operations required for an enterprise-caliber deployment. These products include Oracle 11g and Exadata, IBM DB2, and ParAccel. There are a number of differing causes for this diminished scalability among these packages, including:

- An operating system process-intensive architecture
- A focus on comparatively small, in-memory databases
- Unbalanced resource allocation due to the partitioning architecture
- A requirement to purchase expensive additional modules to achieve greater scalability
Table 2: Examples of reduced enterprise capabilities

<table>
<thead>
<tr>
<th></th>
<th>IBM DB2</th>
<th>Oracle 11g</th>
<th>Oracle Exadata</th>
<th>ParAccel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operating system process-intensive architecture</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aimed at relatively small, in-memory databases</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Resource allocation issues because of partitioning</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Architecture requires add-on purchases</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Small number of concurrent queries consume all bandwidth</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Table 3: Sources of limited concurrency

<table>
<thead>
<tr>
<th></th>
<th>Netezza</th>
<th>Vertica</th>
<th>Greenplum</th>
<th>ParAccel</th>
<th>Microsoft Madison</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lack of TPC-H benchmarks (Which ones that would determine the Xs)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Architected for relatively simple topologies</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Designed for a small number of concurrent users</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Additional concurrency requires significantly higher outlays</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
Cost Control Considerations

As we illustrated earlier, enterprises are continually discovering new and innovative ways to gain competitive advantages from the extensive amounts of new information generated during the course of daily operations. User expectations continue to rise as well: they presume that they will have instant access to this intelligence no matter where they might be located.

Sadly, despite the momentous rise in both data volumes as well as user demands, IT continues to be asked to do more with less. Budgets for new hardware and software acquisitions are pinched, training is increasingly seen as a luxury, and lengthy professional services commitments are becoming a thing of the past.

Thus, faced with these two conflicting realities, it’s critical that acquirers of analytic and data warehousing platforms pay careful attention to keeping expenditures in check. The nature of these classes of solution makes this responsibility especially important: for unsuspecting buyers, it’s easy to get trapped into a never-ending, costly upgrade cycle as platform usage and data volumes inflate.

To help you manage this important responsibility, this section supplies some guidance in two major cost containment categories: total cost of ownership and administration.

Total Cost of Ownership

When procuring an analytic/data warehousing platform, it’s wise to remember that there are many components to the total cost of ownership (TCO) of these selected application. These include:

- Software licensing
- Installation
- Energy
- Hardware
- Professional services
- Administration

In fact, controlling administrative costs is such an important responsibility that the next section of the paper is dedicated to this topic.

As you conduct your evaluation of these types of solution, be sure to take the following considerations into account:

- **Does the product lock you in to proprietary hardware or operating systems?** You’re free to deploy Sybase IQ onto commodity hardware running your choice of operating systems, including Windows, UNIX, and Linux. Being able to scale compute and storage resources separately offers additional cost control benefits.

- **Are you at risk of base license cost increases as your data expands?** Sybase IQ’s price is fixed: you’re free to add as much information to your database as needed.

- **Is the package designed for rapid deployment?** Sybase IQ supplies standards-based, easy-to-learn toolsets for developers and administrators. Its consulting-lite approach means that you get up-and-running quickly. In keeping with this philosophy, it’s available both as a traditional software package as well as a comprehensive, pre-configured analytic appliance.
Does it devote sufficient efforts to compress data? The unique column-based storage architecture of Sybase IQ enables it to compress data efficiently on the fly; as each column is made up of a number of records of the same data type and size, compression can be very efficient and rapid. In 2008, Sybase IQ powered the world’s largest data warehouse, compressing one petabyte of data into 160 terabytes of storage. In the real world, Sybase IQ’s industry leading compression capabilities have been proven time and again — some notable examples are Statistics Canada (up to 80% compression), HealthTrans, USA (up to 70% compression), and Korean Health Insurance Review Agency (up to 80% compression).

Technologies such as Oracle 11g and Exadata, IBM DB2, Teradata, and Greenplum commonly demand significant investments from customers in software licensing and professional services charges. Reasons for these expenditures are varied among these providers, and include:

- A vendor business model focused on the biggest deployments with the biggest budgets
  - Modular licensing mandates many add-ons to purchase, such as for partitioning

- Product complexity
  - Even simple customizations necessitate consulting

<table>
<thead>
<tr>
<th></th>
<th>Oracle 11g</th>
<th>Oracle Exadata</th>
<th>IBM DB2</th>
<th>Teradata</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aimed at largest projects with biggest budgets</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Consulting-intensive approach</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Additional functionality mandates licensing more modules and tools</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Table 4: Origins of increased licensing and professional services costs

Vertica, and ParAccel are examples of two vendors that charge customers by the amount of data under management. By steadily increasing costs, this naturally serves as a disincentive for enterprises hoping to take full advantage of the power of an analytic and data warehousing solution.

Microsoft SQL Server and Madison, Teradata, and Netezza all require the deployment of proprietary operating systems and/or hardware platforms. This makes it difficult for buyers to profit from the proven cost-control advantages of open source software and commodity hardware. Meanwhile, Oracle (both 11g and Exadata) and IBM DB2 generally require high-end hardware and tuning for adequate performance. This leads to additional administration expenditures as well as increased outlays for energy.

As we mentioned earlier, it’s essential that your data warehousing platform expend sufficient efforts to compress information: while storage is cheaper than ever, it’s still not free. Oracle 11g and Exadata, IBM DB2, Microsoft SQL Server, Teradata, and Greenplum do not feature the same amount of compression as found in Sybase IQ. This shortfall occurs for a number of reasons, including:

- A traditional (i.e. non column-based) storage architecture
- Indexing and partitioning requirements can dramatically expand storage consumption
- A business model that dictates charging for more data
- Add-on services that have their own storage requirements, which can balloon to as much as six times the original data
Oracle 11g  |  Oracle Exadata  |  IBM DB2  |  Teradata  |  Greenplum
--- | --- | --- | --- | ---
Lacks support for end-to-end column storage architecture  | ✓  | ✓  | ✓  | ✓  | ✓
Indexes, aggregates, and partitions consume considerably more storage space  | ✓  | ✓  | ✓  | ✓  | ✓
Business model that charges by amount of raw data  |  | ✓  | ✓  | ✓
Add-on services significantly add to storage layer overhead and complexity  | ✓  | ✓  | ✓

Table 5: Factors inhibiting effective data compression

Administration

When calculating the total cost of ownership, many organizations forget to factor in ongoing administration expenditures. To make matters even more confusing, these outlays often vary widely among different solution providers. To compare apples-to-apples, ask the following questions when evaluating a package:

- **Is the solution in question designed for minimal administration?** Sybase IQ’s unified environment means less effort to configure, tune, monitor, and manage your environment.
- **Can your existing DBA staff come up to speed quickly?** For administrators familiar with traditional RDBMS technologies, Sybase IQ is simple to maintain.
- **Will your database administrators need to learn a specialized, non-standard language to administer your environment?** Since Sybase IQ uses standard SQL, training requirements are markedly diminished in contrast with proprietary offerings.
- **Does it require extensive manual configuration and ongoing re-tuning?** As a product purposely tailored for analytics and data warehousing, Sybase IQ is specifically tuned for these types of usages.

Packages such as Oracle 11g and Exadata, IBM DB2, Teradata, Vertica, Greenplum, and ParAccel all impose significantly higher administrative burdens such as:

- Partitioning design and management
- Logical and physical data modeling
- Query tuning
- Resource balancing
The reasons for these extra burdens vary across products, and include:

- Platform complexity
- Proprietary administrative languages
- A smaller selection of management tools
- Manual schema modification
- Diminished visibility into underlying storage

<table>
<thead>
<tr>
<th></th>
<th>Oracle 11g</th>
<th>Oracle Exadata</th>
<th>IBM DB2</th>
<th>Teradata</th>
<th>Vertica</th>
<th>ParAccel</th>
<th>Greenplum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Platform complexity due to indexing, storage, and data distribution</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Immature administrative tools</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Manual management necessary</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Lack of visibility into underlying storage</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
</tbody>
</table>

Table 6: Reasons for extra administrative workload

Netezza, Teradata, and Vertica each require database administrators to learn proprietary languages and acquire additional, non widely-available skills, while Oracle 11g and Exadata each entail adjusting and maintaining numerous manual settings.

Ease-of-deployment Considerations

Enterprises can afford to take more chances when employing analytic and data warehousing products in relatively small, departmental data marts. The appetite for risk is markedly diminished, however, when making the investment in an enterprise-class implementation. Deploying this class of solution is a big decision that involves a significant commitment, and is often undertaken as a response to competitive pressures.

Since few organizations have the time, budget, or personnel to experiment with choices of this magnitude, it’s imperative to ask the right questions and validate that the recommended product will be the optimal selection. In this section, we offer advice on evaluating the completeness of the analytic and data warehousing product, as well as the adopted vendor’s flexibility.
Completeness of platform

Driven by the explosion in new information, attempts to gain competitive advantage, and the desire for better, real-time decision support, analytic and data warehousing solutions are now viewed as mission-critical enterprise assets.

When deciding on a product for this type of undertaking, IT organizations need to balance its feature set with cost control and speed-of-deployment determinants: a lightweight, niche solution won’t have enough capabilities to support all of the users and workloads, while an expensive, heavyweight package will take longer to deploy and consume far more resources than most enterprises can afford. To help fine-tune this determination, here are a few questions to ask:

- **Is the product compatible with leading database products, SQL standards, and data integration tools?** Although designed as a specialized analytic and data warehousing platform, Sybase IQ interoperates with the data management toolsets in place within mainline IT organizations, thus providing stability and compatibility.

- **Has it been available and in use for a respectable amount of time?** Sybase IQ has been in production at thousands of sites for more than 10 years. It has a predictable, mature product release cycle, with a significant updates every 18 months and patches as needed. The market has validated it as well, with more than 1,800 customers and a 96% customer satisfaction rating.

- **Does it offer enterprise-grade capabilities?** Important capabilities such as stored procedures, OLAP functions, and extended data types provide a solid foundation for constructing the most comprehensive applications. Five core and five specialized indexes furnish application designers and database administrators with ample choices for supporting diverse platform usages.

- **Is security a core aspect of product capabilities?** Selected by some of the world’s most security-focused financial institutions, Sybase IQ offers multiple encryption capabilities and protocols, data encryption at the database and column levels.

Oracle 11g and IBM DB2 are general purpose products that haven’t been specifically designed for analytics. This translates into a number of unfortunate side effects, including:

- OLAP extensions are limited in functional depth
- Ad-hoc query and analysis capabilities are not as powerful as in specialized analytic products
- Fewer industry-specific analytics offerings

Oracle 11g and Exadata, IBM DB2, and Teradata all have chosen to decouple key functionality from their core product offerings. This is usually a result of a conscious architectural decision to deliver a modular product, along with a business model based on selling multiple software components. In fact, the data warehousing and analytic solution is often an amalgamation of multiple products.

Microsoft SQL Server, Netezza, Vertica, and ParAccel have each placed diminished focus on enterprise-grade capabilities. Depending on each product’s feature set, these gaps include:

- Auditing
- Archiving
- Referential integrity
- High-availability/disaster recovery
- Automated backup management
- Rich SQL extensions
- A procedural language for complex analytics
- Administrative tools
Table 7: Examples of reduced enterprise capabilities

<table>
<thead>
<tr>
<th></th>
<th>Netezza</th>
<th>Vertica</th>
<th>ParAccel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Missing archiving solution</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>No referential integrity</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diminished high availability/disaster recovery capabilities</td>
<td>✓</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Omitted SQL extensions/SQL compatibility issues</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Minimal source-to-target replication solution</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Mature administrative tools</td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

With more timely and sensitive information stored in data warehouses, it’s essential that this data be safeguarded. Oracle 11g, Microsoft SQL Server, IBM DB2, Netezza, and Vertica are each hampered by their own, unique gaps in security coverage, including:

- Requiring additional modules and options at extra cost
- Mandating frequent security patches
- Needing third party security add-ons
- No column-level security
- Lack of encryption

Table 8: Security gaps

<table>
<thead>
<tr>
<th></th>
<th>Oracle 11g</th>
<th>Microsoft SQL Server</th>
<th>IBM DB2</th>
<th>Netezza</th>
<th>Vertica</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequent security patches necessary</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Additional modules and/or third-party security products necessary</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Minimal/missing encryption</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>No column-level security</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Significant amount of common vulnerabilities and exposures</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No FIPS certified encryption algorithms</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Vendor flexibility

Single-vendor IT organizations are extremely rare: most rely on technologies produced by multiple suppliers. With this fact in mind, it’s easy to observe that analytic and data warehousing offerings that feature proprietary technologies and closed systems will not meet the needs of the enterprise. In addition, some IT teams seek to establish their own hardware and software infrastructure, while others prefer to install specialized, pre-configured appliances to handle major computing tasks.

To determine if your software supplier will offer sufficient adaptability, ask questions such as these:

- **Does the platform offer support for heterogeneous environments?** Sybase IQ can be populated from a variety of data sources, including structured and unstructured data from Sybase and non-Sybase originators. It’s also available on multiple operating systems, including Windows, UNIX, and Linux.

- **Is there a solid ecosystem in place for the product?** With more than 10 years of success and greater than 1,880 customers, there are numerous trained personnel familiar with Sybase IQ as well as a strong partner and professional services network.

- **Can you choose a software or appliance deployment model?** Sybase Analytic Appliance is configured and tuned out of the box, providing extreme analytics capabilities for overburdened enterprise DW systems. Combining category-leading technologies, including Sybase IQ, PowerDesigner, IBM Power Systems and MicroStrategy business intelligence technology, the Analytic Appliance gives all the benefits of a custom-built enterprise data warehouse, but is easy, fast and affordable.

- **Is the platform open or proprietary?** With Sybase IQ, you’re free to select the development, analytic, business intelligence, and administrative tools of your choice.

Oracle (11g and Exadata) and Microsoft (SQL Server and Madison) have followed a more restricted, closed architecture philosophy, translating to the following types of side-effects:

- Oracle’s limited, proprietary APIs
- Requiring the use of vendor-specific front-end tools, which are often charged separately
- Reduced operating system choices in the case of Microsoft’s products

While targeted, niche offerings such as Teradata, Vertica, and Greenplum do a good job of serving as a point solution, this tactic usually implies:

- A small number of qualified service partners
- Scarcity of trained personnel increases costs
- A relatively small number of sites
- Reduced emphasis on building an ecosystem

Microsoft SQL Server and Microsoft Madison, and ParAccel (which is primarily a query accelerator for SQL Server databases) have all seen most of their success in smaller warehouses, for diverse reasons such as:

- Not architected for large implementations
- Designed for simple data models
- Structured for low concurrency
- Not meant for high availability
Finally, Microsoft Madison, ParAccel, Vertica, and Greenplum are each deployed in a relatively limited number of environments, making it difficult to assess their suitability for the mass analytic and data warehousing marketplace.

**Conclusion**

IT organizations will increasingly rely on robust, far-reaching analytic and data warehousing solutions to cope with ever-expanding amounts of information under management. These applications will be a key technique for satisfying highly sophisticated, demanding end-users and keeping up with competitors.

The growth of this category has led to numerous new suppliers providing offerings designed to address these needs. Unsurprisingly, each supplier brings a unique perspective to this landscape, which has led to major product differences among these firms. Given these variations, it’s critical that enterprises seeking to deploy an analytic and data warehousing package apply an objective, vendor-neutral set of considerations when evaluating these offerings.

To help you come up with evaluation criteria tailored to your specific needs, table 9 shows the relative strength of each of the vendors discussed in the areas evaluated in the preceding sections.

<table>
<thead>
<tr>
<th></th>
<th>IBM DB2</th>
<th>Greenplum</th>
<th>Microsoft Madison</th>
<th>Microsoft SQL Server</th>
<th>Netezza</th>
<th>Oracle 11g</th>
<th>Oracle Exadata</th>
<th>ParAccel</th>
<th>Sybase IQ</th>
<th>Teradata</th>
<th>Vertica</th>
</tr>
</thead>
<tbody>
<tr>
<td>Powerful performance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simple and flexible scalability</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affordable total cost of ownership</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ease of deployment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Solution Flexibility</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
About the Author

Robert D. Schneider is a Principal at Think88 Ventures LLC. Based in Silicon Valley, and drawing from a deep pool of global talent and expertise in a wide variety of disciplines, Think88 delivers high-value solutions quickly and cost-effectively in the following practice areas:

- High performance data warehousing design, development, and optimization
- Consulting and training based on modern Service Oriented Architecture (SOA) and Cloud Computing technologies
- Technical marketing content creation, including competitive analysis, case studies, and white papers

He has provided database optimization, distributed computing, and other technical expertise to a wide variety of enterprises in the financial, technology, and government sectors. Clients have included JP Morgan Chase & Co, VISA, HP, S.W.I.F.T., and the governments of the United States, Brazil, Malaysia, Mexico, and Australia. Robert has written 6 books and numerous articles on database technology and other complex topics such as Cloud Computing, and Service Oriented Architecture (SOA). He is a frequent organizer and presenter at technology industry events, worldwide.