

## Sybase® IQ Powers Data Aggregators to Deliver Value-Added Information Services



### SOLUTION OVERVIEW

In a recent independent benchmark, awarded the Guinness World Record for the largest data warehouse, Sybase IQ compressed one petabyte of raw data by 85 percent resulting in a “green” data environment with 90 percent less CO<sub>2</sub> emissions, and using 91% less energy, over the data warehouse’s life time.

[www.sybase.com/  
detail?id=1054011](http://www.sybase.com/detail?id=1054011)

### WHEN INFORMATION IS THE BUSINESS

More and more companies across numerous industries are beginning to rely on the data and analytic services provided by third-parties to drive some of the most critical decisions about their businesses—investments, customer relationships, product and service offerings and mix, competitive strategies, risk management, and more. These service providers, known as Data Aggregators, are companies whose business is to collect and sell information. These firms generate revenue by selling news, data, analysis, and reports (online or print) derived from the data they hold for other businesses that need it.

Data aggregators are information brokers, collecting industry- or society-wide data to provide value-added services to customers and subscribers, often over the Web. The services provided by data aggregators go by many names: database marketing services, financial and credit information services, audience measurement services, market research providers, national statistical agencies, online shopping comparison sites, and many others. They must provide differentiated, high-value data and analytics services to be competitive, and must deliver on customer expectations: faster turnaround on their analytics, more accurate predictive models, more detailed analytics to discover trends, and greater flexibility with “what if” queries.

### DATA AGGREGATOR INFRASTRUCTURES ARE UNDER PRESSURE

Data aggregators face the challenge of converting vast stores of data into a product or service, since their operations depend on reselling data and/or related analytics. Some aggregators manage the most complex and demanding data integration and analytics challenges imaginable: capturing millions of data points at the most detailed level from hundreds or thousands of locations every day, only to turn that data around for customer use in very short order. Others face more modest requirements. Typically, all will face one or more of the following challenges: large data sets, large numbers of concurrent users, complex queries, and ad hoc queries.

- **Large data sets.** Data aggregators are managing much more data than they did even one or two years ago because the analysis and information reporting services they provide require more and more data to be available online. The more data that is accessible for analysis, the more accurate predictive models for their customers are likely to be. For example, performing analysis of trends sufficient to make meaningful predictions requires many years’ worth of historical data, and in fact many data aggregators are now maintaining data warehouses that incorporate tens or even hundreds of terabytes of raw data.
- **Large numbers of concurrent users.** Because many of the services data aggregators provide are made available via on-demand access over the internet, the numbers of concurrent users that are accessing their portals can become very large. In order to ensure customer satisfaction, data aggregators must be able to provide their customers with same level of speed and quality of service no matter how much traffic their web-based applications receive, with no degradation in performance.

“SYBASE IQ IS THE FASTEST IN THE INDUSTRY. WE LOAD ALMOST 10 BILLION ROWS A MONTH, A TERABYTE A QUARTER. WITH SYBASE IQ IT JUST FLIES.”

CRAIG SILVER, SENIOR DATABASE ARCHITECT, NIELSEN MEDIA RESEARCH

- **Complex queries.** Data aggregators run complex queries that compare and contrast different data sets in order to spot trends, correlate cause and effect, or recognize patterns. “How has this new product affected sales of earlier products?” or “which customers are more likely to respond to this new campaign?” represent complex queries. This analysis also considers a wide range of predicates—selection criteria such as those based on sex, age, income, location, or other factors. The growing amount of data mentioned above also makes these complex queries more challenging because of the sheer volume of data that needs to be analyzed to identify root causes of events, data relationships, and the likely outcomes of potential business decisions.
- **Ad hoc queries.** Data aggregator customers require the ability to explore vast amounts of business data with an unconstrained “what if?” approach. Due to this fact, ad hoc (vs. planned or scheduled) queries comprise a significant amount of the query workload, can be very large, and can potentially be run at any time by 100s or 1,000s of customers. Since the timing and volume of these ad hoc queries is unpredictable, the “right” capacity needed for them is elusive, raising the twin risks that resources will fall short of demand or will be under-utilized much of the time—with disastrous consequences for the data aggregator in either case.

These massive infrastructure demands that data aggregators have consume significant computing resources and place added pressures upon the systems they use for their data and analytics services. This can result in significant efforts to tune and optimize the data warehouse, or to add significant hardware and storage capacity to handle the additional workloads. These pressures can result in two undesirable outcomes related to data aggregator infrastructures: slower performance and higher costs.

- **Slower performance.** The sheer volume and complexity of queries required by data aggregator customers can have a dramatic effect on system performance. Slow response times have a direct impact on revenues, productivity, and customer satisfaction—it is clearly unacceptable for customers to wait hours for results.
- **Higher costs.** It takes more system resources to handle more data and higher workloads. Traditional enterprise data warehouses or OLTP systems consume large amounts of CPU cycles to read every byte of every row of a large database and deliver the query result. In order to keep performance at target levels, more hardware must be added to the system. This can also require more DBA time to tune queries, adding indexes and summary tables to ensure acceptable response times and meet customer expectations.

Many data aggregators are realizing that traditional transactional-oriented database systems cannot deliver the performance required to overcome these daunting analytical challenges. Using traditional database infrastructures often ends up in compromised results—needing to limit the number of users, complexity of queries, or amount of available data. Hence, these issues are driving the need for a faster, more efficient way for data aggregators to manage and store business data.

#### **ANALYTICS SERVERS PROVIDE THE PERFORMANCE DATA AGGREGATORS REQUIRE**

Many data aggregators are realizing the benefits of using an optimized analytics server that is uniquely designed to meet their requirements for collecting large volumes of data from different sources, and providing very fast processing of concurrent customer queries. Such architectures enable complex analysis for large numbers of users and the business imperative to increase the amount of information and the number of users.

#### **SYBASE IQ—THE COMPETITIVE WEAPON FOR DATA AGGREGATORS**

Sybase IQ, the world’s leading column-based analytics server, empowers data aggregators to not only overcome their data management challenges, but to thrive in the face of the most demanding technology and business requirements. Sybase IQ’s column-based architecture and innovative features deliver ultra-high performance to support large numbers of ad hoc and complex queries. Sybase IQ’s data compression algorithms cost-effectively support the very largest databases. In-database analytics capability takes that performance to a new level. In addition, Sybase IQ’s multi-node architecture helps data aggregators to simultaneously support a large clientele of concurrent users while allowing continuous data feeds into their data warehouses.

Unlike traditional databases that were designed to support business transactions, Sybase IQ was architected for analytics. Transactional databases require complex, space-consuming indexing and summary tables to perform query-intensive workloads well. These indexes and summary tables actually explode database sizes, often requiring 5 or 10 times more storage in the analytical system than in the original operational system.

Traditional databases are also more complex to implement for data aggregator environments. They require more time to load and refresh, due to labor-intensive steps of creating backups, tables, and indexes. It also takes a lot of tuning to maintain query performance with a traditional database—diagnosing, testing, and tuning queries over and over again.

Sybase IQ is the competitive weapon for data aggregators—with extremely high performance for analyzing massive amounts of data, without requiring special tuning or expensive hardware and storage resources.

**“THE SPEED OF SYBASE IQ IMPROVES OUR ABILITY TO MINE THE DATA AND PRODUCE RESULTS FOR OUR CUSTOMERS MUCH MORE QUICKLY. THAT HELPS THEM MARKET MORE EFFECTIVELY AND GENERATE MORE BUSINESS.”**

**RIC ELERT, VICE PRESIDENT OF ENGINEERING, COMSCORE**

### THE ADVANTAGES OF SYBASE IQ FOR DATA AGGREGATORS

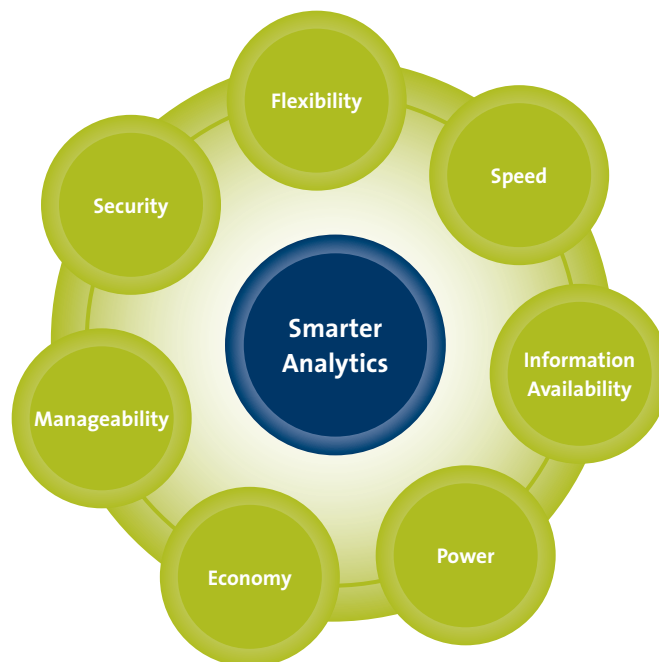
For data aggregators looking for more accurate and cost-effective analytics solutions that deliver winning business results in a volatile and competitive world climate, Sybase IQ is the world’s leading column-based analytics server. Sybase IQ delivers the smartest approach to enable enterprises to turn raw data into actionable information—to increase revenue, customers and margins.

Sybase IQ is the only analytics server that delivers dramatically faster, more accurate analytics—to all your customers, from all your information, on your terms. Sybase IQ’s column-based core architecture and innovative features provide flexibility and scalability in multiple dimensions, enabling unsurpassed query performance for the most complex analytics tasks, for the rapidly expanding BI user base, on the largest datasets—all within an environment that is both affordable, eco-friendly and manageable at the implementation stage and throughout the information lifecycle.

All of this is evidenced by award-winning results: more than 1,500 customers with more than 3,000 projects worldwide, top industry analyst, media and pundits support, and record-breaking benchmark results.

Specific capabilities of Sybase IQ that are critical for data aggregators include:

- Supports as many queries as possible running in parallel
- Column-based encryption as well as database-level encryption
- In-database analytics capability for greater speed and security
- Multiplex architecture enables readers to be allocated to separate subscribers
- Add new services and service levels, including web-based self-service
- Faster time-to-revenue for new products, services
- Add new analysis options and types of inquiries



## SYBASE IQ: A CLEAR WINNER

Sybase IQ infuses organizations with fast, flexible access to information. With Sybase IQ, you can analyze business performance, detect fraud, gauge marketing results, manage customer relationships, and ensure financial controls—in ways never before possible. The economics of Sybase IQ put a faster, efficient, cost-effective advanced analytics system well within reach.

For more information, contact us today at [infobi@sybase.com](mailto:infobi@sybase.com) or visit [www.sybase.com/iq](http://www.sybase.com/iq).

<b>Speed</b>	<ul style="list-style-type: none"><li>• Architected for queries, not transactions</li><li>• Query parallelism per node</li><li>• Query speed 10X – 100X faster than traditional row-based</li><li>• In-database analytics capability</li></ul>
<b>Databases</b>	<ul style="list-style-type: none"><li>• Patented indexing technology</li></ul>
<b>Economy</b>	<ul style="list-style-type: none"><li>• Supports all standard connectivity protocols</li><li>• Supports all standard hardware platforms/disk technologies</li><li>• Patented Data Compression (reduces storage needs up to 90%)</li><li>• Configurable tablespaces</li></ul>
<b>Manageability</b>	<ul style="list-style-type: none"><li>• One-touch multiplex grid configuration</li><li>• Index advisor utility guides index selection for top performance</li><li>• XML &amp; GUI-based query plan generator for diagnostics</li><li>• Versatile GUI-based admin console for server management</li></ul>
<b>Power</b>	<ul style="list-style-type: none"><li>• One-touch multiplex grid configuration</li><li>• Query parallelism per node</li><li>• High performance from any schema</li><li>• Supports hundreds or thousands of users &amp; petabytes of data</li><li>• In-database analytics capability</li><li>• Foundations for analytical applications</li></ul>
<b>Flexibility</b>	<ul style="list-style-type: none"><li>• Independent compute power and storage capacity scalability</li><li>• On-the-fly schema changes without query blockage or downtime</li><li>• Handles structured and unstructured data</li><li>• Works with industry-leading BI ISV products</li></ul>
<b>Information Availability</b>	<ul style="list-style-type: none"><li>• Multi-node loading</li><li>• Enhanced large table loads</li><li>• Direct loads from clients</li><li>• No impact to query performance while loading</li><li>• Supports rules-based information lifecycle management</li></ul>
<b>Security</b>	<ul style="list-style-type: none"><li>• Password authentication/encryption for data-in-flight</li><li>• Database and column encryption for data-at-rest</li><li>• Common Criteria certified at EAL-3</li><li>• FIPS (on Unix), Kerberos, and IPv6 support</li></ul>