Stratecast *A Division of* FROST & SULLIVAN

Accelerating 'Time to Analytics' with Data Warehouse Appliances

Stratecast Perspectives & Insight for Executives (SPIE)
SPIE 2009 #34 — September 18, 2009

ACCELERATING 'TIME TO ANALYTICS' WITH DATA WAREHOUSE APPLIANCES

INTRODUCTION[1]

Business investment decisions are fraught with tradeoffs. Typical among these are cost, time, functional capability, ROI, longevity and a host of other drivers. Certainly this can be said about the choice between investing in information technology for day-to-day operations versus developing systems to facilitate more strategic pursuits such as planning, forecasting, market identification and product creation.

In support of these strategic endeavors, advanced data analytics[2] capabilities are gaining a foothold amongst Communication Service Providers (CSPs). Unlike simple reporting, analytics employs sophisticated statistical routines to mine the billions of transactions recorded each day. These routines search for meaningful patterns of activity that may provide visibility into business opportunities – functions not available in routine business applications. CSP executives are investing in hardware, software, education and professional services to help them with their analytics plans, as they seek to identify their most profitable offers; customers most at risk of churn; offers with a high likelihood of success; and opportunities for reduced operating expense, to name just a few.

[1] In preparing this report, Stratecast interacted with the following representatives of Netezza:

- Jim Baum, President & CEO
- Justin Lindsey, CTO
- Tim Young, Vice President, Corporate Marketing
- John Gillespie, Vice President, Telecom Industry Marketing

[2] Please note: the insights and opinions expressed in this assessment are those of Stratecast and have been developed through the Stratecast research and analysis process. These expressed insights and opinions do not necessarily reflect the views of the company executives interviewed.

Stratecast has published several reports on the use of advanced, predictive analytics for CSPs, including OSSCS 9-09 "CSPs Move Data Analytics Forward to Improve Competitive Odds" (November, 2008); SPIE 2009 #29 "IBM Drives Analytics into Mainstream with SPSS Acquisition" (August, 2009), and others. For more information on subscription options, contact your Account Executive, visit www.stratecast.com or telephone +1-877-463-7678.

But communications is an information intensive business, and the need to store and manage enormous volumes of data over long periods of time means maintenance and tuning of mission-critical production databases takes precedence over strategic analytical tasks. Moreover, database management systems used for daily business aren't necessarily well suited to the high-volume analysis and rapid feedback requirements of analytics, to say nothing of ad hoc query. To help analytics pay off, purpose-built data warehouse appliances now offer the kind of speed and volume handling needed for complex, strategic and fact-based decision support.

When one wishes to characterize the length of time needed to get a product from conceptual stages to full production, the term "Time to Market" is invoked. Likewise, "Time to Profitability" captures the earliest spending on a product to its first moment of solvency. In this SPIE, we coin the term 'Time to Analytics', with an eye toward today's data warehouse appliances and their applicability to rapidly expanding CSP analytic objectives.

THE ROAD TO THE DATA WAREHOUSE APPLIANCE

From Database to Warehouse

Databases[3] for commercial use were originally designed and established to assist businesses with storage of information for daily operational activities. Known as Transaction Processing (TP) and their interactive counterparts, Online Transaction Processing (OLTP), these systems excelled at organizing, storing and retrieving small amounts of consistently represented data, such as customer numbers, names, product numbers, dollar amounts, and the like.

Consider the example of a customer account. Information such as name, address, telephone number, credit limit and account number is stored in an account database. In that same database, transactions made by those customers may also be recorded. Every time a transaction occurs, this database is updated with those details.

In telecom, it is not uncommon for these databases to be tied to a single operational application such as CRM, Billing or Inventory. And with multiple product line 'silos' in place, a large incumbent telecom operator may easily have several hundred operational databases, particularly with older systems.

In contrast, the more recently developed *data warehouses* are databases optimized for business intelligence reporting, analysis, and high-volume, high-speed analytics. They often combine output from multiple operational databases. They also are driven by some

[3] The origins of relational technology, as opposed to earlier hierarchical data structures and management techniques, are attributable to Edgar F. Codd, Ph.D., whose seminal work in this field began in the early 1970s while Codd was employed by IBM. It was Oracle, however, that beat IBM to the market with a commercial database predicated on Dr. Codd's ideas. Source: Association for Computing Machinery.

form of a data model[4] – a combination card catalog/roadmap to help the user understand how the data is set up and, perhaps more importantly, connect similar information that may be arranged or recorded differently among two or more databases.

Data warehouses provide a purpose-built repository, optimized for analytical tasks. They are distinct from operational databases and do not compete for IT resources with production applications. As such, they are built for speed and volume.

As the adoption rate of data warehouses accelerated over the past 10-15 years, users embraced the advantages afforded by business intelligence and analytics applications. While simple reporting was left in the domain of the single application, data warehouses afforded cross-application and even enterprise-wide views of information.

Today most CSPs have some form of data warehouse for analytical activities such as customer churn analysis, product promotion or billing validation/revenue assurance checks against usage stores. Some CSP warehouses measure in terabytes with hundreds of users.

From Warehouse to Appliance

In an ironic twist, the more popular data warehouses became, the less productive they were, as users heaped on multiple production tasks while also issuing complex analytical queries. Industry standard hardware/software configurations designed for general purpose business computing began to bog down under the weight of increased use. Queries against 5 terabytes of data that took an hour to complete suddenly ran for half a day against a 30 terabyte warehouse. And adding more hardware did not provide a commensurate level of productivity improvement. This diminishing marginal return pushed up the per-terabyte cost to unreasonable levels, forcing industry executives back to the table to consider the next set of tradeoffs.

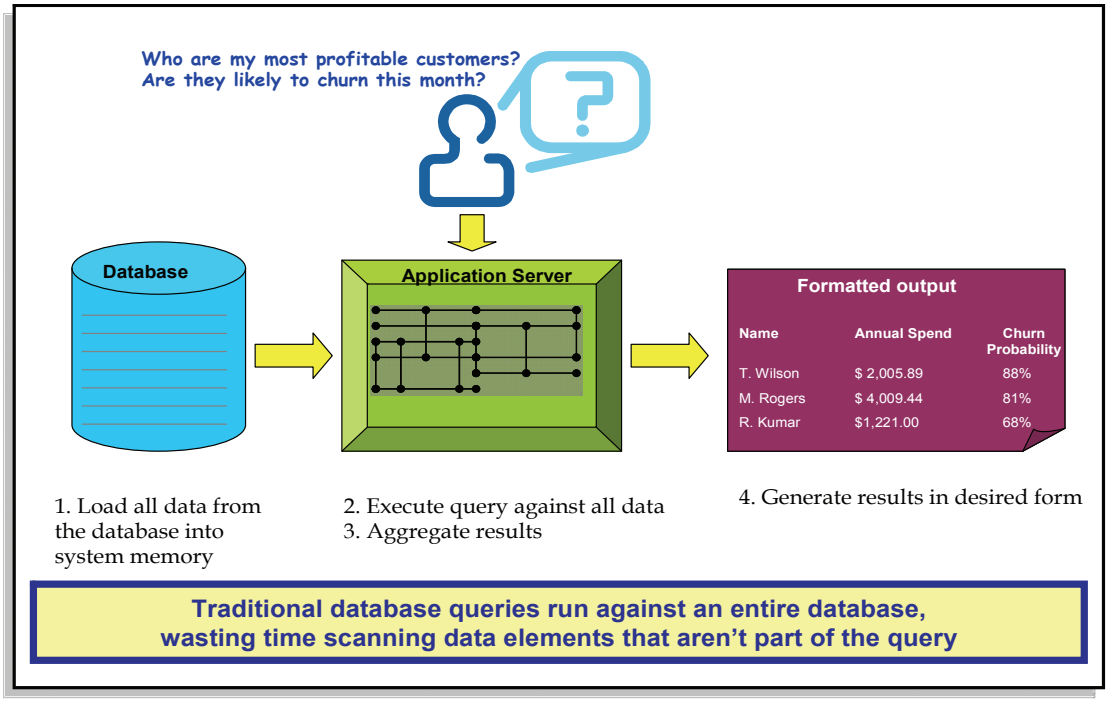
The logical evolutionary step was to fuse the data warehouse with purpose-built hardware that broke through the traditional database-server-storage configuration. As if making a return to the earliest days of IT where computers were designed for specific applications, data warehouse appliances emerged to address ever-growing database sizes along with ever more complex uses of the data.

[4] As used in this context, the term 'data model' is meant to encompass all related constructs used to connect disjointed stores of information, including schemas, metadata, and ontology. Technical attributes of data models are beyond the scope of this document. A more detailed discussion may be found in OSSCS 08-05, "Reducing OSS/BSS Integration Costs Through Common Information Modeling" (July, 2007).

UNDERSTANDING THE DATA WAREHOUSE APPLIANCE[5]

To achieve maximum productivity, data warehouse appliances re-configure both the hardware and software. Consider a traditional database query scenario, as seen in Figure 1.

Figure 1 – *A Traditional Approach to Analytical Queries*

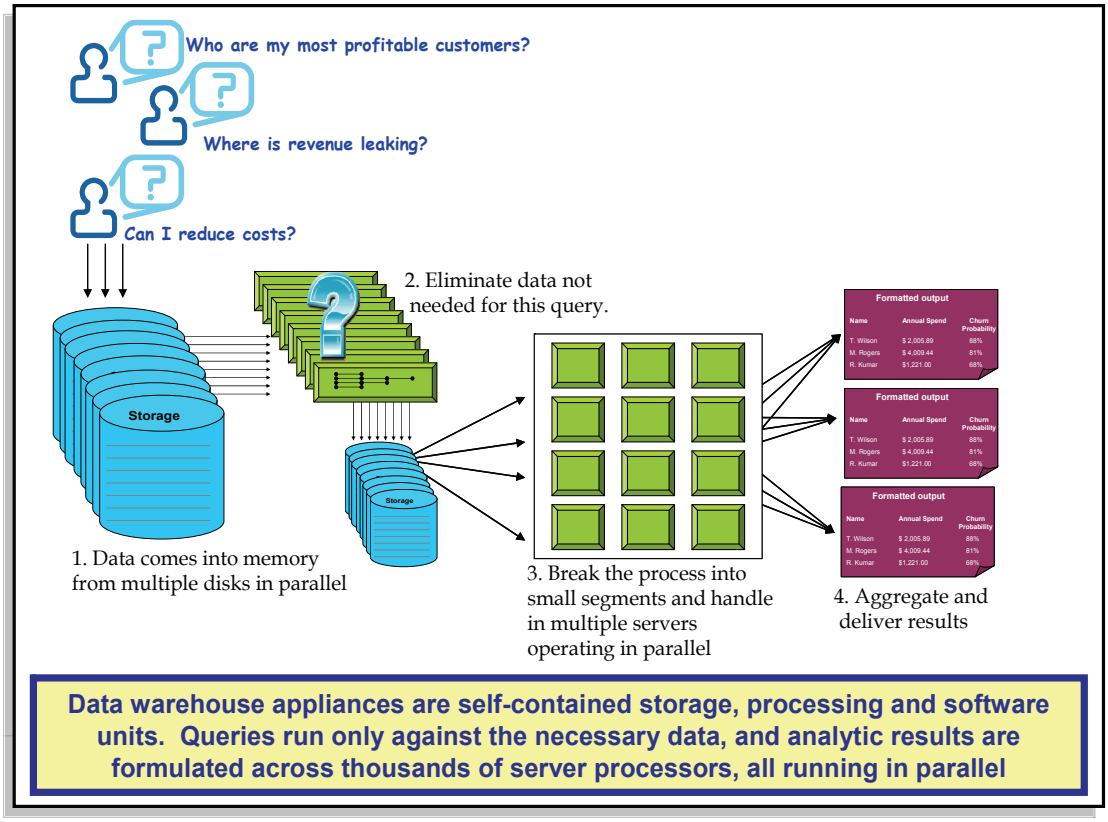


Source: Stratecast

Data warehouse appliances are much more effective in the way they execute queries. For example, Netezza's recently released TwinFin appliance determines which data is needed, segments the problem into manageable pieces, and then answers the query across multiple processors, all operating in parallel, as illustrated in Figure 2 below.

[5] An in-depth discussion of the hardware and software architectures of data appliances is beyond the scope of this document, as each vendor's design has its own nuances. This section provides information for the business reader only.

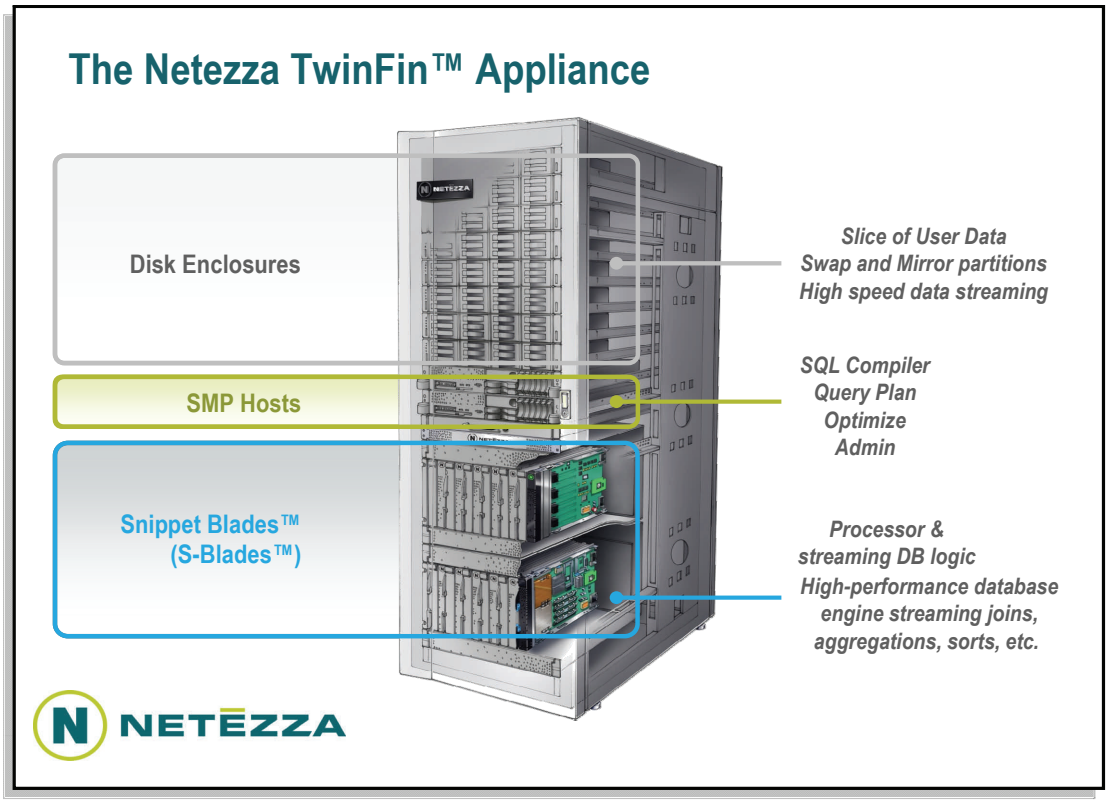
Figure 2 – *Data Warehouse Appliance Query Processing Example – Netezza TwinFin*



While not all appliances function in this exact fashion, it is true they are engineered for optimal data handling and analysis, consistent with business intelligence and analytics objectives.

Finally, data warehouse appliances are often provided as self-contained units of server, storage, and database, where data transfers and inter-component communications can occur at 'wire speed', that is, the shortest distance at the greatest speed allowable by the laws of physics. Figure 3 following provides an example of such a product.

Figure 3 – *Data Warehouse Appliances are Often Self-Contained Units*



Source: Netezza

WHY AN APPLIANCE?

Data warehouse appliances do more than address volume and speed requirements with more hardware. Were it that simple, additional conventional servers would suffice. Instead, by re-engineering both the hardware configuration and the software query handling sequence, appliances generally bring orders-of-magnitude improvement to analytic tasks. Some of the more noteworthy advantages include the following:

- ***Performance:*** Appliances are purpose-built for analytics and deliver significant improvements in the time required to deliver answers. With requirements for analytics moving toward more near-real-time, this can be a project differentiator. For example, Netezza executives explained to Stratecast that Orange UK noted an improvement in report generation for its senior management and competitive intelligence teams from eight weeks to one day, taking its output from marginally useful to immediate feedback and insight.
- ***Volume:*** Appliances have significant capacity by design. Current CSP case studies describe CDR volumes in the range of 200 Terabytes. Further, this volume of data serves multiple business objectives including customer behavioral analysis as well as revenue assurance.

- ***Granularity:*** Most traditional databases rely on some form of data summarization to achieve reasonable response times. While useful for higher level reporting, this is completely counter to analytical applications, whose value is predicated on the ability to ferret out and interpret relationships between the details of the data.
- ***Simplicity:*** Some appliances can be literally rolled in and set up in a couple of hours. Netezza executives told Stratecast they recommend that CSPs perform proof -of-concept tests in their own data centers with Netezza hardware. There are no special heating/cooling requirements.
- ***Cost:*** Traditional data warehouses have estimated total cost of ownership (TCO) of $100,000 to $200,000 per terabyte. Even at smaller sizes, costs of less than $100,000/terabyte are easily achieved, with some offering proof of TCO as low as $20,000/terabyte under certain conditions.
- ***Staffing:*** Unlike traditional databases, many appliances require little database tuning and maintenance by database administrators, saving on the cost of human capital.
- ***Reuse:*** Appliances generally support SQL[6] and will easily accept databases from current conventional database management systems (DBMS) and data warehouses. While this is perhaps sub-optimal, it does represent a fast means to results where high-volume analytical applications already exist.
- ***Non-disruptive:*** Appliances operate in parallel with daily business operations, providing a high-performance environment for multiple users to issue simultaneous queries.
- ***Standards-based:*** Some appliances offer rack-mounted blade server design based on widely available components using Intel chips, and applications run on standard operating systems such as Linux. There is little risk of component obsolescence.

Stratecast believes the data warehouse appliance is an essential foundation element of an analytics strategy. It mitigates (or eliminates) the time and expense tradeoffs normally encountered with multi-terabyte databases or data warehouses using conventional hardware.

Industry analytics and business intelligence software heavyweights such as SAS, SPSS, MicroStrategy, IBM Cognos, Oracle and SAP Business Objects have all embraced data warehouse appliances, having witnessed growing acceptance in multiple lines of business and government. Prominent systems integrators including Accenture, HP/EDS, Cap Gemini, TCS and Patni are also fully aboard as CSPs seek opportunities to compete on something other than pricing and bundling.

[6] Structured Query Language is the de facto standard for managing data in relational database management systems, based on Dr. Codd's original relational model.

Stratecast

The Last Word

The days of simple query and cross-tab reporting as the sole source of business application output are over. Today's CSP requires much more sophistication, including predictive modeling of marketing campaigns, forensic analysis of usage for fraud, and large-scale data mining for revenue enhancement and cost reduction opportunities.

To achieve this goal in a timely and cost-effective way, the general-purpose database gave way to the data warehouse more than 20 years ago. But with the practical limits of warehouse productivity being exceeded when deployed on general purpose computing hardware, costs began to rise as numbers of users and applications increased. The data warehouse was a victim of its own success.

In the last few years, the data warehouse appliance emerged as a solution to the tradeoffs of speed, volume, cost and user satisfaction. This special purpose hardware/ software combination provides the order-of-magnitude improvement needed to satisfy complex analytical queries in less time and across more data. Moreover, the declining cost to produce these results is fast approaching mid-sized company feasibility.

There is no doubt that general purpose computing capability will remain the dominant hardware and software choice in our data centers for many years to come, if for no other reasons than acquisition cost, availability of product, and access to human resources schooled in their use. But the technology evolution pendulum has swung back to the days of purpose-built products optimized for specific circumstances. The data warehouse appliance is evidence of this movement, and dozens of CSPs worldwide are now stepping forward to offer their testimony to the appliance's business value.

Whereas CSPs once talked in giga- and terabytes of storage, conversations about petabytes (millions of gigabytes) and exabytes (millions of terabytes) of storage are quickly becoming commonplace. Data warehouse appliances free IT departments from the hardware expansion conundrum and evolve analytics forward to carrier-grade.

Susan McNeice

Global Program Director – Communications Infrastructure and OSS/BSS Software

Stratecast (a Division of Frost & Sullivan)

smcneice@stratecast.com

CONTACT US

Silicon Valley
331 E. Evelyn Ave.
Suite 100 Mountain View, CA 94041
Tel 650.475.4500
Fax 650.475.1570

San Antonio
7550 West Interstate 10, Suite 400,
San Antonio, Texas 78229-5616
Tel 210.348.1000
Fax 210.348.1003

London
4, Grosvenor Gardens,
London SWIW ODH,UK
Tel 44(0)20 7730 3438
Fax 44(0)20 7730 3343

877.GoFrost
myfrost@frost.com
http://www.frost.com

Beijing
Bengaluru
Bogotá
Buenos Aires
Cape Town
Chennai
Delhi
Dubai
Frankfurt
Kolkata
Kuala Lumpur
London
Manhattan
Melbourne
Mexico City
Milan
Mumbai
Oxford
Palo Alto
Paris
Rockville Centre
San Antonio
São Paulo
Seoul
Shanghai
Singapore
Sydney
Tel Aviv
Tokyo
Toronto
Warsaw

ABOUT STRATECAST

Stratecast assists clients in achieving their strategic and growth objectives by providing critical, objective and accurate strategic insight on the global communications industry. As a division of Frost & Sullivan, Stratecast's strategic consulting and analysis services complement Frost & Sullivan's Market Engineering and Growth Partnership services. Stratecast's product line includes subscription-based recurring analysis programs focused on Business Communication Services (BCS), Consumer Communication Services (CCS), Communications Infrastructure and Convergence (CIC), OSS and BSS Global Competitive Strategies (OSSCS), and our weekly opinion editorial, Stratecast Perspectives and Insights for Executives (SPIE). Stratecast also produces research modules focused on a single research theme or technology area such as IMS and Service Delivery Platforms (IMS&SDP), Managed and Professional Services (M&PS), Mobility and Wireless (M&W), Multi-Channel Video Programming Distribution (MVPD), Network Infrastructure and OSS (NI&O), Secure Networking (SN) and Unified Communications (UC). Custom consulting engagements are available. Contact your Stratecast Account Executive for advice on the best collection of services for your growth needs.

ABOUT FROST & SULLIVAN

Frost & Sullivan, the Growth Partnership Company, partners with clients to accelerate their growth. The company's TEAM Research, Growth Consulting, and Growth Team Membership™ empower clients to create a growth-focused culture that generates, evaluates, and implements effective growth strategies. Frost & Sullivan employs over 45 years of experience in partnering with Global 1000 companies, emerging businesses, and the investment community from more than 30 offices on six continents. For more information about Frost & Sullivan's Growth Partnership Services, visit http://www.frost.com.