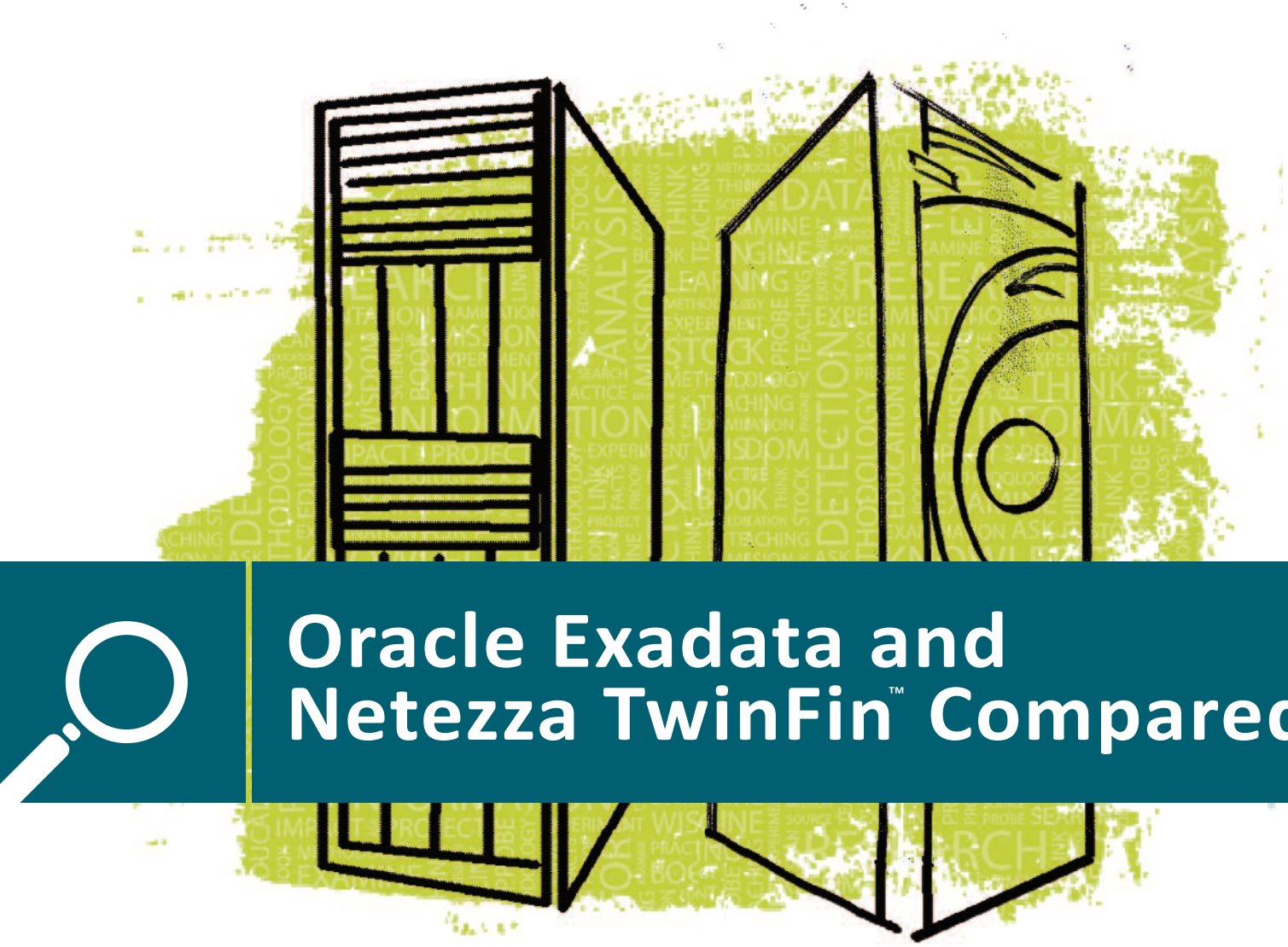
NETEZZA®
an IBM® Company



Oracle Exadata and Netezza TwinFin™ Compared.

Less than one year since announcing the release of Netezza's TwinFin, over a hundred customers have adopted the appliance including:

- **Blue Cross Blue Shield of Massachusetts**
- **BlueKai**
- **Catalina Marketing**
- **Con-way Freight**
- **DataLogix**
- **Epsilon, an Alliance Data Company**
- **interCLICK**
- **IntercontinentalExchange**
- **Japan Medical Data Center**
- **Kelley Blue Book**
- **Marshfield Clinic and Marshfield Clinic Research Foundation**
- **MediaMath**
- **MetroPCS**
- **MicroAd**
- **MyLife.com**
- **NYSE Euronext**
- **Pacific Northwest National Laboratory**
- **Premier, Inc.**
- **The Nielsen Company**
- **Wind Telecom**

Table of Contents

1 Introduction

Netezza focuses on technology designed to query and analyze big data. The company's innovative data warehouse appliances are disrupting the market. Wishing to exploit data at lower costs of operation and ownership, many of our customers have moved their data warehouses from Oracle. Oracle has now brought Exadata to market; a machine which apparently does everything TwinFin does, and also processes online transactions. Exadata's strength is simplifying and removing cost from Oracle customers' online transaction processing (OLTP) systems. It consolidates multiple systems, removes duplication and replaces complexity with a single scalable high performance OLTP environment. However, these benefits are compromised when organizations attempt to consolidate their analytics databases into Exadata. Netezza has a track record of delivering value to Oracle customers. By removing the analytics workload from Oracle OLTP databases, today over 200 organizations realize the best of both worlds.

This examination of Exadata and TwinFin as data warehouse platforms is written from an unashamedly Netezza viewpoint, however to ensure credibility we have taken advice from Philip Howard, Research Director of Bloor Research and Curt Monash, President, Monash Research.

To innovate requires us to think and do things differently, solving a problem using new approaches. Netezza focuses exclusively on customers' needs and wants for data warehousing. TwinFin delivers excellent **performance** for our customers' warehouse queries. TwinFin offers customers **simplicity**; anyone with basic knowledge of SQL and Linux has the skills needed to perform the few administrative tasks required to maintain consistent service levels through dynamically changing workloads. TwinFin's performance with simplicity reduces their costs of owning and running their data warehouses. More important, our customers create new business value by deploying **analytic applications** which previously they considered beyond their reach.

> ***"Netezza was part of the inspiration for Exadata. Teradata was part of the inspiration for Exadata,"*** acknowledged Larry Ellison on 27th January 2010. ***"We'd like to thank them for forcing our hand and forcing us to go into the hardware business."***

"Netezza was part of the inspiration for Exadata. Teradata was part of the inspiration for Exadata," acknowledged Larry Ellison on January 27, 2010. "We'd like to thank them for forcing our hand and forcing us to go into the hardware business."[1] While delivered with Larry Ellison's customary pizzazz, there is a serious point to his comment: only the best catch Oracle's attention. Exadata represents a strategic direction for Oracle; adapting their OLTP database management system, partnering it with a massively parallel storage system from Sun. Oracle launched Exadata V2 with the promise of extreme performance for processing both online transactions and analytic queries. That Oracle excels at OLTP is a given. But data warehousing and analytics make very different demands of their software and hardware than OLTP. Exadata's data warehousing credentials demand scrutiny, particularly with respect to simplicity and value.

This white paper opens by reviewing differences between processing online transactions and processing queries and analyses in a data warehouse. It then discusses Exadata and TwinFin from perspectives of their query performance, simplicity of operation and value.

All we ask of readers is that they do as our customers and partners have done: put aside notions of how a database management system should work, be open to new ways of thinking and be prepared to do less, not more, to achieve a better result.

...be prepared to do less, not more, to achieve a better result.

One caveat: Netezza has no direct access to an Exadata machine. We are fortunate in the detailed feedback we receive from many organizations that have evaluated both technologies and selected TwinFin. Given Oracle's size and their focus on Exadata, publicly available information on Exadata is surprisingly scarce. The use cases quoted by Oracle provide little input to the discussion, which in itself is of concern to several industry followers, e.g. Information Week.[2] Information shared in this paper is made available in the spirit of openness. Any inaccuracies result from our mistakes, not an intent to mislead.

[1] *See http://oracle.com.edgesuite.net/ivt/4000/8104/9238/12652/lobby_external_flash_clean_480x360/default.htm*
[2] *See http://www.informationweek.com/news/business_intelligence/warehouses/showArticle.jhtml?articleID=225702836&cid=RSSfeed_IWK_News*

2 Online Transaction Processing (OLTP) and Data Warehousing

OLTP systems execute many short transactions. Each transaction's scope is small, limited to one or a small number of records and is so predictable that often times data is cached. Although OLTP systems process large volumes of database queries, their focus is writing (UPDATE, INSERT and DELETE) to a current data set. These systems are typically specific to a business process or function, for example managing the current balance of a checking account. Their data is commonly structured in third normal form (3NF). Transaction types of OLTP systems are stable and their data requirements are well-understood, so secondary data structures such as indices can usefully locate records on disk, prior to their transfer to memory for processing.

In comparison, data warehouse systems are characterized by predominantly heavy database read (SELECT) operations against a current and historical data set. Whereas an OLTP operation accesses a small number of records, a data warehouse query might scan a table of a billion rows and join its records with those from multiple other tables. Furthermore, it is impossible to predict what queries business users will ask of data in their warehouse, reducing the value of caching and indexing strategies. Choices for structuring data in the warehouse range from 3NF to dimensional models such as star and snowflake schemas.

Data within each system feeding a typical warehouse is structured to reflect the needs of a specific business process. Before data is loaded to the warehouse it is cleansed, de-duplicated and integrated.

This paper divides data warehouses as either first or second generation. While this classification may not stand the deepest scrutiny, it reflects how many of our customers talk about their evolutionary path to generating greater and greater value from their data.

First-generation data warehouses are typically loaded overnight. They provide information to their business via a stable body of slowly evolving SQL-based reports and dashboards. As these simple warehouses somewhat resemble OLTP systems – their workload and data requirements are understood and stable – organizations often adopt the same database management products they use for OLTP.

With the product comes the practice: database administrators analyze each report's data requirements and build indices to accelerate data retrieval. Creep of OLTP's technology and techniques appears a success, until data volumes in the warehouse outstrip those commonly managed in transactional systems.

In this century, corporations and public sector agencies accept growth rates for data of 30-50% per year as normal. Technologies and practices successful in the world of OLTP prove less and less applicable to data warehousing; the index as aid to data retrieval is a case in point. As the database system processes jobs to load data, it is also busy updating its multiple indices. With large data volumes this becomes a very slow process, causing load jobs to overrun their allotted processing window. Despite working long hours, the technical team misses service levels negotiated with the business. Productivity suffers as business units wait for reports and data to become available.

Technologies and practices successful in the world of OLTP prove less applicable to data warehousing...

Organizations are redefining how they need and want to exploit their data; this paper refers to this development as the second-generation data warehouse. These new warehouses, managing massive data sets with ease, serve as the corporate memory. When interrogated, they recall events recorded years previously; these distant memories increase the accuracy of predictive analytic applications. Constant trickle feeds are replacing overnight batch loads, reducing latency between an event's recording and its analysis within the context of millions of other events. Beyond the simple SQL used to populate reports and dashboards, the warehouse processes linear regressions, Naïve Bayes and other mathematical algorithms of advanced analytics. Noticing a sudden spike in sales of a high-margin product at just five stores drives a retailer to understand what happened and why. This knowledge informs strategies to promote similar sales activity at all 150 store locations. The computing system underpinning the warehouse must be capable of managing these sudden surges in demand without disrupting regular reports and dashboards. The business users are demanding the freedom to exploit their data at the time and in the manner of their choosing. Their appetite for immediacy leaves no place for technologies whose performance depends upon the tuning work of administrators.

3 Query Performance

To analyze these terabyte-to-petabyte volumes requires a computer capable of spreading data over many compute nodes, each of which analyzes a smaller set of data in parallel with all its other peer nodes in the grid. Warehouses capable of sustaining fast, consistent query performance under constantly changing workloads enable a new class of applications, creating valuable insights and spurring the adoption of data-driven decision-making.

Query Performance with Oracle Exadata

In acquiring Sun, Oracle has come to the conclusion Netezza reached a decade earlier: data warehouse systems achieve highest efficiency when all parts, software and hardware, are optimized to their goal. Exadata is created from two sub-systems connected by a fast network: a smart storage system communicating via InfiniBand with an Oracle Database 11g V2 with Real Application Clusters (RAC). A single rack system includes a storage tier of 14 storage servers, called Exadata cells, in a massively parallel processing (MPP) grid, paired with the Oracle RAC database running as a shared disk cluster of eight symmetric multi-processing nodes.

In acquiring Sun, Oracle has come to the conclusion Netezza reached a decade earlier: data warehouse systems achieve highest efficiency when all parts, software and hardware, are optimized to their goal.

Oracle labels Exadata's storage tier as smart because it processes SQL projection, restriction and simple join filtering,[3] before putting the resulting data set on the network for downstream processing by Oracle RAC. This technique is called smart scan. However, whether data is processed in Exadata's MPP tier is conditional on a number of factors, including the type of table used to store data, how recently data arrived in the warehouse and other factors discussed below.

[3] *A Technical Overview of the Sun Oracle Exadata Storage Server and Database Machine – An Oracle white paper, October 2009.*

Data managed within Oracle structures called index-organized tables (IOTs) are not processed by smart scan. On page 3-20 of Oracle® Database Concepts 11g Release 2 (11.2), dated February 2010, the documentation recommends "Index-organized tables are useful when related pieces of data must be stored together or data must be physically stored in a specific order. This type of table is often used for information retrieval, spatial, and OLAP applications". Page 19-53 of Oracle® Database Administrator's Guide 11g Release 2 (11.2), dated February 2010, states "Index-organized tables are suitable for modeling application-specific index structures. For example, content-based information retrieval applications containing text, image and audio data require inverted indexes that can be effectively modeled using index-organized tables". Netezza customers run these second generation data warehouse applications in TwinFin. For example, the Levenshtein edit distance algorithm is invaluable in applications ranging from cleansing of name and address data for marketing campaigns to complex analysis of text undertaken by national security agencies as they scrutinize names of persons of interest listed in port of entry data. Organizations following Oracle's advice and managing text, image, audio and spatial data in IOTs, will find their applications perform sub-optimally as Exadata is unable to bring the power of massively parallel processing to their data.

First generation warehouses, commonly running as OLTP databases on SMP servers, simply lacked processing power to run all the queries a business required. The operational data store (ODS) represents one structure adopted by architects to work around this short-coming. Regularly updated by multiple operational systems, the ODS integrates data and provides both transactional and analytical services to applications. Second generation warehouses running MPP architecture offer organizations the opportunity to consolidate operational data stores into the warehouse. Exadata is unable to bring the power of smart scan to blocks containing a single active transaction (INSERT, UPDATE, DELETE) and an Exadata block can contain many records. This architectural quirk of Oracle creates complexity for customers considering Exadata as an ODS.

While Oracle provides 511 functions in the latest release of their database, only 319 of these can be processed in Exadata's MPP storage tier. Smart scan's conditionality extends to predicates of a function – for example Exadata will not exploit datetime functions using predicate months_between(d,sysdate) = 0, or months_between(d,current_date) = 0, or months_between(d,to_date('01-01-2010','DD-MM-YYYY')) = 0[4].

[4] *http://antognini.ch/2010/08/exadata-storage-server-and-the-query-optimizer-%E2%80%93-part-4/*

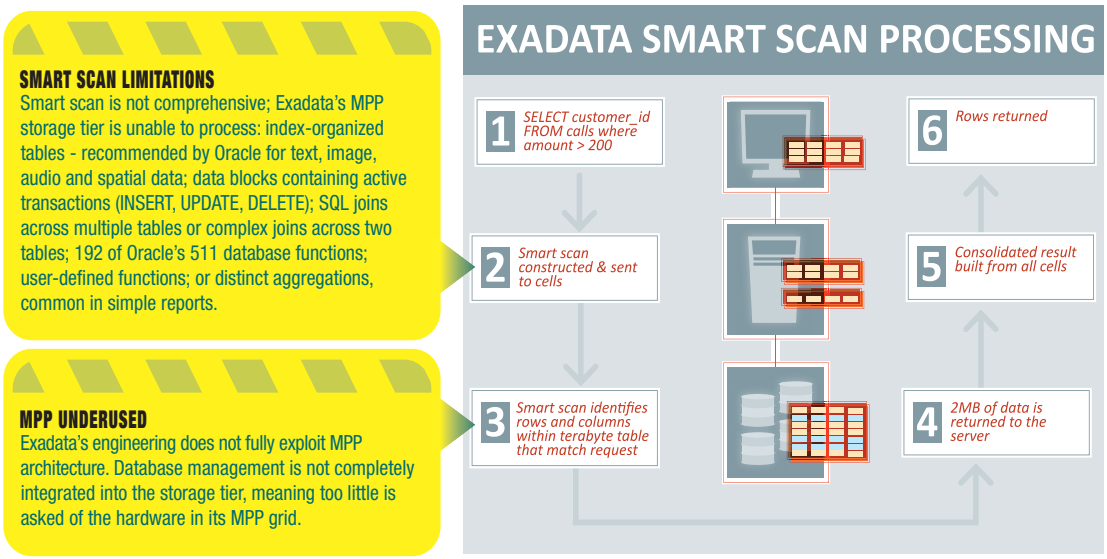
Exadata's storage tier provides Bloom filters to implement simple joins between one large and one smaller table, anything more complex cannot be processed in MPP. Analytical queries commonly require joins more complex than those supported by Exadata. Consider the straightforward case of an international retailer needing insight to the dollar value of sales made in stores located in the UK. This simple SQL query requires a join across three tables – sales, currency and stores.

```
select sum(sales_value * exchange_rate) us_dollar_sales
from sales, currency, stores
where sales.day = currency.day
and stores.country = 'UK'
and currency.country = 'USA'.
```

Exadata is unable to process this three table join in its MPP tier and instead must inefficiently move all the data required by the calculation across the network to Oracle RAC.

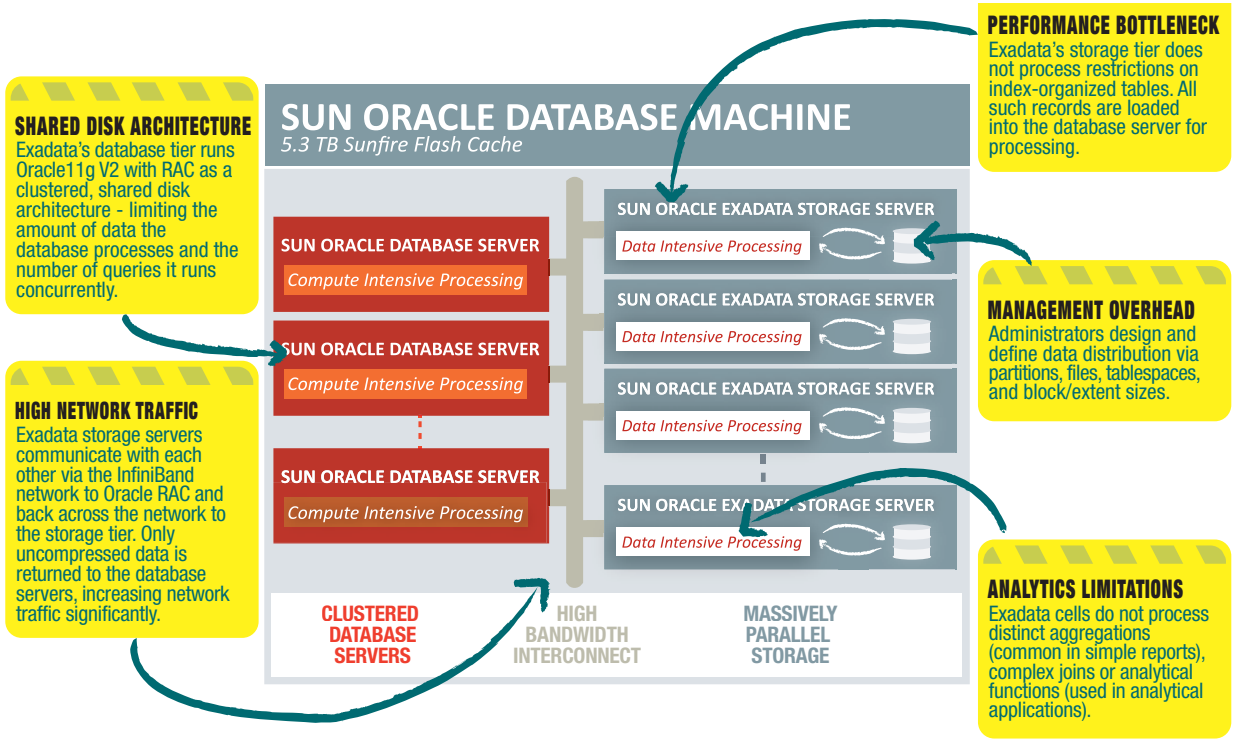
User-defined functions (UDFs) fulfill a central role in the second generation warehouse. UDFs extend query functionality beyond SQL and enable embedded analytic processing. Exadata is unable to process UDFs in MPP but must push responsibility for their processing to Oracle RAC, an OLTP database neither designed to nor capable of analyzing big data sets.

Exadata's storage tier cannot process distinct aggregations, common even in simple reports, and heavily used in analytic queries.



Exadata storage servers cannot communicate with one another; instead all communication is forced via the InfiniBand network to Oracle RAC and then back across the network to the storage tier. This architecture is beneficial to online transaction processing where each transaction, with a scope of one or few records, can be satisfied by moving a small data set from storage to the database. Analysis of very large data sets forces Exadata to move massive volumes of data between its storage and database tiers, adversely effecting query performance.

Oracle positions its use of 40 Gb/sec switch InfiniBand as an advantage over TwinFin; in reality Exadata needs this expensive network because of the system's imbalance and inefficiency. Exadata storage servers do too little work, so more data than necessary is put on the network to be moved downstream for processing by Oracle RAC, which is asked to do too much work.



Every disk in Exadata's storage tier is shared by all nodes in the grid running Oracle RAC. This communal storage creates the risk of a page being read by one node while it is being updated by another. To manage this, Oracle forces coordination between nodes. Each node checks the disk activity of its peers to prevent conflict. Oracle technicians refer to this activity as block pinging. Compute cycles consumed as each node checks disk activity of its peers, or lost as one node idly waits for another to complete an operation, are wasted. While useful for OLTP, in a database specifically designed for data warehousing, these cycles would not be wasted but employed processing queries, mining data and running analyses.

Consider the case of an organization interested in analyzing equity trade data to identify correlations in a sector of a stock market. Their algorithm calculates Spearman's rank correlation coefficient (Spearman's rho), measuring statistical dependence between two variables by assessing how well the relationship between them can be described. This analysis creates valuable insight in to whether specific equities influence behavior of other equities in the same market sector within a window of one to ten minutes. The data set is large: a long history of market data and daily trades, to be analyzed against their historic record as the market moves. The analysis requires a Cartesian join between all the equities in the sector while at the same time calculating a Volume-Weighted Average Price and a Return From Previous Close value for the equity under investigation. The results pass to Spearman's rank correlation coefficient function to calculate the Population Covariance and the standard deviation of every equity combination for the time period. After several hours testing the customer was still waiting for a result from Oracle 10g, whose poor performance excluded it as a technology in a business case with a maximum window of 10 minutes.

In Exadata Oracle 11gR2 is paired with an MPP storage system and a fast network. However smart scan is unable to process Cartesian joins, the first step of in this analytic process, and brings zero value to this calculation. Exadata will put all the records on the network and send them across to Oracle RAC. The question becomes: how much more efficient is Oracle RAC than Oracle 10g? In processing the algorithms Oracle must create and manage temporary data sets and write these out of memory for storage. Exadata's flash cache may play some role here, but the size of the data sets and the complexity of the algorithms will force database processes to write to disk. This flow from Oracle RAC is back across a network still clogged with coming from the MPP storage tier data, queued and unprocessed waiting for attention from a fully-consumed Oracle RAC. Exadata's network connection between the servers and the shared-disk is a bottleneck. Oracle RAC, designed as an OLTP database, is a bottleneck.

For all but simple queries Exadata must move large sets of data from its storage tier to its database tier, raising questions on its suitability as a platform for a modern data warehouse.

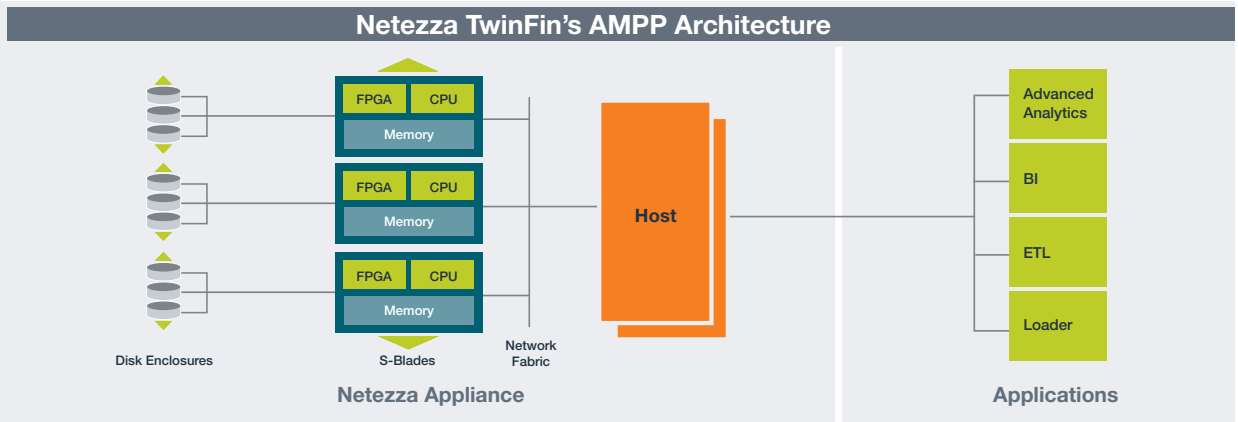
For data warehousing, Exadata's MPP storage tier is limited primarily to performing predicate level filtering (row-level restrictions based on the "where" clause of the SQL statement and column-level projections based on the "select" clause) and providing the CPU processing power for decompression of data stored on disk as it is scanned. And, processing in Exadata's MPP tier is conditional, forcing administrators to delve into the machine's complexity. All remaining complex join, aggregation and filtering logic, the heavy lifting of warehousing and data analysis, then executes on the results from all Exadata nodes in the clustered SMP head nodes running Oracle 11g RAC. Exadata presents interesting opportunities for CIOs looking to consolidate multiple OLTP systems to a single platform. For all but simple queries Exadata must move large sets of data from its storage tier to its database tier raising questions on its suitability as a platform for a modern data warehouse.

Query Performance with Netezza TwinFin

TwinFin is designed from the ground up as a data warehousing platform. Netezza employs an Asymmetric Massively Parallel Processing (AMPP) architecture. A Symmetrical Multiprocessing host[5] fronts a grid of Massively Parallel Processing nodes. TwinFin exploits this MPP grid to process the heavy lifting of warehousing and analyzing data.

A node in TwinFin's grid is called an S-Blade (Snippet-Blade), an independent server containing multi-core central processing units (CPUs). Each CPU is teamed with a multi-engine Field Programmable Gate Array (FPGA) and gigabytes of random access memory. Because the CPUs have their own memory, they remain focused exclusively on data analysis and are never distracted to track activity of other nodes, as occurs with block pinging in Oracle RAC.

[5] *TwinFin has two SMP hosts for redundancy but only one is active at any one time.*



An FPGA is a semiconductor chip equipped with a large number of internal gates programmable to implement almost any logical function, and particularly effective at managing streaming processing tasks. Outside of Netezza, FPGAs are used in such applications as digital signal processing, medical imaging and speech recognition. Netezza's engineers have built software machines within our appliances' FPGAs to accelerate processing of data before it reaches the CPU. Within each Exadata rack Oracle dedicates 14 eight-way storage servers to accomplish less than Netezza achieves with 48 FPGAs embedded within our blade servers. Each FPGA – just a 1"x1" square of silicon – achieves its work with enormous efficiency, drawing little power and generating little heat.

Netezza didn't take an old system with known shortcomings and balance it with a new smarter storage tier; TwinFin is designed as an optimized platform for data warehousing.

Inter-nodal communication across Netezza's MPP grid occurs on a network fabric running a customized IP-based protocol fully utilizing total cross-sectional bandwidth and eliminating congestion even under sustained, bursty network traffic. The network is optimized to scale to more than a thousand nodes, while allowing each node to initiate large data transfers to every other node simultaneously. These transfers bring enormous efficiency to the processing tasks typical of data warehousing and advanced analytics. Just as SQL statements benefit from processing within TwinFin's MPP architecture, so too do computationally complex algorithms at the heart of advanced analytics. Previous generations of technology physically separate application processing from database processing, introducing inefficiencies and constraints as large data sets are shuffled out of the warehouse to the analytic processing platforms and back again. Netezza brings the heavy computation of advanced analytics into its MPP grid, running the algorithms in each CPU physically close to the data, making data movement redundant and boosting performance. The algorithms benefit from running on the many nodes of Netezza's MPP grid, freed from constraints imposed on less-scalable clustered systems.

TwinFin demonstrates how a true MPP architecture excels in calculating Spearman's rank correlation coefficient - a real workload on a real dataset. The customer loads a massive volume of trading data into TwinFin and constantly trickle feeds data from live markets into the warehouse. TwinFin executes every step of the query in parallel utilizing all its hardware and software resources. Netezza's intelligent storage selects only the rows needed for that market sector and projecting only the columns needed for assessment. The join result is directly streamed to the code implementing the statistical analysis which TwinFin downloads to every processor in its MPP grid, running the complex calculations in parallel. Results from each node in the MPP grid are returned via the network to the host for final assembly and rendering back to the requesting application. TwinFin completes the analysis in a few minutes, well-within the organization's 10 minute window, and then runs it regularly for as long as the market is open and as long as the organizations wishes to assess behavior of different equities in dynamic markets. Beyond the financial services industry Spearman's rank correlation coefficient is valuable to statisticians working in econometrics, quality control, bioinformatics and marketing.

Netezza didn't take an old system with known shortcomings and balance it with a new smarter storage tier; TwinFin is designed as an optimized platform for data warehousing. TwinFin unconditionally runs all queries in its massively parallel grid, delivers performance generously, and makes life easy for programmers, administrators and users.

4 Simplicity of Operation

A customer of Netezza's from the financial services industry used the Lean approach to analyze resource expenditure required to manage their Oracle data warehouse. They learned in building and maintaining indices, aggregates, materialized views and data marts that

more than 90% of their IT team's work was either required waste or non-value added processing.

Simplicity of Operation with Oracle Exadata

Oracle relies on trained, experienced, and skilled database administrators designing databases to deliver performance to queries. One design technique is the physical partitioning of data, meaning queries can be restricted to partitions known to contain relevant data. Partitions are a heritage of Oracle's OLTP roots. In the second generation warehouse administrators cannot know what queries and analyses will be run tomorrow or next week. Exadata includes Oracle's Automatic Storage Management which automates striping of partitions across all available disks. However an administration team must still create partitions, configure and manage disk groups for shared storage across instances, choose and implement either 2-way mirroring or 3-way mirroring, and configure Allocation Unit sizes. Additionally, Exadata configuration requires administrators create and manage tablespaces, index spaces, temp spaces, logs and extents. As observed by industry analyst Curt Monash, "Even better might be a system that doesn't lean heavily on complex partitioning to achieve good performance." [6]

[6] *Curt Monash at http://www.dbms2.com/2009/09/21/notes-on-the-oracle-database-11g-release-2-white-paper/*

As described in the previous section, whether Exadata conditionally processes data in its MPP storage tier; this complication requires administrators, even those experienced with previous versions of Oracle, attend a series of education classes. Oracle recommends "Make sure your DBAs have up to date (11gR2) knowledge. Besides Exadata knowledge[7]" Gaining up to date knowledge requires each administrator takes a minimum of four courses totaling 18 days. Administrators new to Oracle RAC need study for another 5 days to learn intricacies of its administration.

A customer of Netezza's from the financial services industry used the Lean[8] approach to analyze resource expenditure required to manage their Oracle data warehouse. They learned in building and maintaining indices, aggregates, materialized views and data marts that more than 90% of their IT team's work was either required waste or non-value added processing. The cost of this waste translates to unnecessary hardware and software license costs, terabytes of wasted storage, elongated development and data load cycles, long periods of data unavailability, stale data, poorly performing loads and queries and excessive administrative costs.

Exadata does little to simplify managing an Oracle data warehouse. Administrators must manage multiple server layers, each with operating system images, firmware, file systems and software to be maintained. Oracle suggests that DBAs should expect to spend 26% less time managing 11g[9], the database version in Exadata, than they spend on older 10g deployments. If this is confirmed in practice and Exadata reduces by a quarter the time customers waste in valueless administration, Oracle has taken a step in the right direction. Netezza's appliances are designed not to waste any of the customers' time. "The DBA team only backs up the environment and manages the high level security model for the appliance and that is it. They don't need to do anything else (for example, the concept of indexing is foreign to them when dealing with Netezza)." [10]

7 *http://www.oracle.com/global/hu/events/20100126/exadata_migration_budapest_robert_pastijn.pdf*

8 *With roots in manufacturing, "Lean" is a practice using tools and techniques of Six Sigma to analyze wasteful expenditure of resources, and target activities adding no value to the product or service for elimination.*

9 *http://www.dbms2.com/2009/09/21/notes-on-the-oracle-database-11g-release-2-white-paper/*

10 *Customer using Oracle for OLTP and Netezza for data warehousing quoted from Linked-In Exadata Vs Netezza forum at http://www.linkedin.com/groupAnswers?viewQuestionAndAnswers=&gid=2602952&discussionID=11385070&sik=1275353329699&trk=ug_qa_q&goback=.ana_2602952_1275353329699_3_*

Not only do business users demand that their queries complete quickly, they also expect consistent performance; a report that completed in five seconds yesterday and three minutes today will likely create a ticket requiring response from IT helpdesk staff. Warehouses are inevitably subject to the demands of varied, dynamic workloads. Data arriving from OLTP systems via batch jobs or trickle feeds are loaded, administrative tasks such as backup and restore and grooming run in the background out of view of the business and dashboards are constantly updating. At the same time, computational intensive applications – such as those predicting which claims or trades might be fraudulent or irregular – create sudden, heavy load on the warehouse infrastructure. Delivering consistent performance to the business makes two requirements of the warehouse: consistent query performance and effective workload management which simplifies allocation of available computing power to all the jobs requiring service, usually based on priorities agreed with the business.

To manage workloads in Oracle requires administrators to learn multiple tuning parameters, many with a high degree of dependence on others. The situation is more complicated in Exadata where certain parameters must be set to the same value for every processor in their grid. Rather than a top-down system of control, administrators must work bottom-up - experimentally changing individual parameter settings and observing the effect in a production environment. Oracle' philosophy of work load management succeeds for online transaction processing, where the load each application places on databases and servers can be profiled in a development environment, effects of parameter settings noted and understood, and incremental tweaks made. Once a stable system is achieved the parameter settings can be put into production.

There is no "normal load" in second generation warehouses; business users will unleash computationally-demanding analyses on massive data sets as events occur in the outside world. Analysts will not wait for a quiet time when other loads subside. The dynamic nature of the warehouse exposes the fragility of Oracle's tuning approach to work load management. Achieving and maintaining consistent performance for large communities of users, with different application and data requirements, through rising and falling loads, is a complex task. Even a team of highly trained and experienced Oracle experience administrators will struggle to maintain a stable warehouse capable of delivering consistent performance to regular report and load jobs while leaving headroom for unexpected, but important, surges in analyst activity.

"The way we did a proof of concept with them [Netezza] was, they shipped us a box, we put it into our data center and plugged into our network. Within 24 hours, we were up and running. I'm not exaggerating, it was that easy."

Oracle RAC is a complex technology and its tuning parameters arcane. Exadata's complexity is confirmed by the number of training days Oracle recommends even for administrators experienced in previous versions of their database.

Simplicity of Operation with Netezza TwinFin

Netezza's customers willingly confirm on public record that our appliances are simple to install and use. "The way we did a proof of concept with them [Netezza] was, they shipped us a box, we put it into our data center and plugged into our network," he said. "Within 24 hours, we were up and running. I'm not exaggerating, it was that easy."[11] This commentary is from Joseph Essas, vice president of technology at eHarmony, Inc., a company already using Oracle's database and RAC software.

Reducing the time to get productive is a good start; Netezza's philosophy is to bring simplicity to all phases of data warehousing. The first task facing a customer is loading their data. TwinFin automates data distribution. Experience from proof-of-concept projects is that customers load their data to Netezza using automatic distribution, run their queries and compare results to their highly tuned Oracle environments. For all but the simplest queries, automatic distribution is good enough for TwinFin to outperform Oracle. Customers may later analyze all their queries to identify those that can be accelerated by redistributing data on different keys. TwinFin makes this task simple.

[11] *http://www.itworld.com/software/61575/eharmony-finds-data-warehouse-match-netezza*

There's something to be said for a simple approach

- **NO cluster interconnect (GES & GCS) monitoring/tuning**
- **NO RAC-specific knowledge/tuning (DBAs with RAC experience are less of a commodity)**
- **NO dbspace/tablespace sizing and configuration**
- **NO redo/physical log sizing and configuration**
- **NO journaling/logical log sizing and configuration**
- **NO page/block sizing and configuration for tables**
- **NO extent sizing and configuration for tables**
- **NO temp space allocation and monitoring**
- **NO integration of OS kernel recommendations**
- **NO maintenance of OS recommended patch levels**
- **NO JAD sessions to configure host/network/storage**
- **NO query (e.g. first_rows) and optimizer (e.g. optimizer_index_cost_adj) hints**
- **NO statspack (statistics, cache hit, wait event monitoring)**
- **NO memory tuning (SGA, block buffers, etc.)**
- **NO index planning/creation/maintenance**
- **Simple partitioning strategies: HASH or ROUND ROBIN**

All queries submitted to TwinFin are automatically processed in its massively parallel grid with no involvement of database administrators. Queries and analyses enter TwinFin through the host machine where the optimizer, the compiler and the scheduler decompose them into many different pieces or snippets, and distribute these instructions to the MPP grid of processing nodes, or S-Blades, all of which then process their workload simultaneously against their locally-managed slice of data.

A Snippet arriving at each of TwinFin's S-Blades initiates reading of compressed data from disk into memory. The FPGA then reads the data from memory buffers and utilizing its Compress Engine decompresses it, instantly transforming each block from disk into the equivalent of 4-8 data blocks within the FPGA. Netezza's engineering accelerates the slowest component in any data warehouse – the disk. Next, within the FPGA data streams into the Project Engine which filters out columns based on parameters specified in the SELECT clause of the SQL query being processed. Only records fulfilling the SELECT clause are passed further downstream to the Restrict Engine where rows not needed to process the query are blocked from passing through gates, based on restrictions specified in the WHERE clause. The Visibility Engine maintains ACID (Atomicity, Consistency, Isolation and Durability) compliance at streaming speeds. All this work, the constant pruning of unneeded columns and rows, is achieved in an energy efficient FPGA measuring just one square inch. If TwinFin doesn't need to move data, it doesn't.

The FPGA's pre-processing complete, it streams just the resulting trimmed down set of records back into S-Blade memory where the CPU performs higher-level database operations such as sorts, joins and aggregations, doing this in parallel with all other CPUs within the MPP grid. The CPU may also apply complex algorithms embedded in the Snippet code for advanced analytics processing. The CPU finally assembles all the intermediate results from the entire data stream and produces a result for the Snippet, sent over the network fabric to other S-Blades or the host, as directed by the Snippet code. When data required by a JOIN is not collocated on a node, TwinFin's inter-nodal network fabric efficiently and simply re-distributes late in the processing cycle after the database has completed restrictions and projections. Some highly complex algorithms require communication among nodes to compute their answer. TwinFin exploits a message passing interface to communicate interim results and to produce the final result.

And, as the original compressed data blocks are still in memory, they can be automatically reused in later queries requiring similar data via TwinFin's table cache – an automated mechanism requiring no DBA training or involvement.

Just three months after moving to Netezza, a customer relates that his team delivered more analytical applications than they could in the previous three years with Oracle.

Because TwinFin applies full parallelism to all tasks, its workload management system plays a critical role in controlling how much of the appliance's computing resources are made available to each and every job. In Netezza's appliance architecture, one software component controls all system resources: processors; disks; memory; network. This elegance is the foundation of TwinFin's Workload Management System. TwinFin's Workload Management System makes it simple for administrators to allocate computational resources to users and groups based on priorities agreed with the business and maintain consistent response times for multiple communities.

TwinFin eliminates the wasted work of database tuning. Equipped to make their own intelligent decisions, Netezza's appliances require no tuning and little system administration. The few administrative tasks necessary to maintain consistent performance through dynamic, changing workloads are within easy reach of anyone with experience of Linux and SQL. All that is required of the administrator is to allocate TwinFin's resources to groups within the user community and hand control to the Workload Management System. Freed from constant cycles of database administration, technical staff engages with the business to investigate new, value-creating ways of exploiting data. Just three months after moving to Netezza, a customer relates that his team delivered more analytical applications than they could in the previous three years with Oracle. Processing analytical applications close to where data is managed, exploiting the same MPP platform as used for processing SQL, represents a real opportunity for organizations to dramatically increase the value they derive from data.

5 Value

Value with Exadata

With enormous volumes of data and complex, often computationally-intensive workloads, the hardware and software needed to run a data warehouse represent significant expenditure in their own right. The experience of many organizations using OLTP databases to build their first generation warehouses is that costs of training and employing specialist technical staff not only outstrip asset costs but recur year-after-year. These costs burden data warehouse projects and detract enormously from the value they create for an organization. As the waste analysis conducted by the financial services customer of both Netezza and Oracle highlights, using Oracle for data warehousing is labor intensive. Netezza customers attest that low level, technically- demanding administration tasks are simply unnecessary; in this light it is indefensible that operating an Oracle database demands administrators spend the majority of their time on care and feeding of the underlying technology, while Netezza customers spend that time creating value by exploiting their data.

Exadata's new storage tier adds another layer of complexity for administrators to tune and manage. Because Exadata is very new, and so few data warehouses using the technology are in production, projections on its cost of ownership are premature. However, customers should expect that achieving consistently high performance from Exadata will incur substantial costs in training, database design and administration.

While adding a new storage tier conditionally reduces the disk throughput bottleneck to Oracle's database, Exadata's engineering is more adaption to massively parallel processing than full exploitation of the architecture. Oracle's failure to integrate data management fully into Exadata's storage tier means too little is asked of hardware in its MPP grid. This inflates the cost of acquiring Exadata; customers pay for hardware that will never be fully exploited by its software. These costs build over the lifetime of the warehouse. Customers pay for under-utilized space in their data centers which would return greater value if used to house a more efficient computer system.

Evaluating the Systems		Netezza	Oracle
Item		TwinFin 12	Exadata v2 (SAS)
Performance & Architecture	MPP	· True MPP · Optimized for Data Warehousing & Analytics	· Hybrid – parallel storage nodes & SMP clustered head node · Optimized for Transaction Processing (OLTP)
	Hardware Architecture	· Full processing S-Blades (1 CPU core + 1 FPGA core / 1 disk drive) · SMP host node used primarily for user/applications interface · Independent blade-to-blade redistribution	· Intelligent storage (1 CPU core / 1.5 disk drives) · SMP Cluster nodes running Oracle 11g RAC · InfiniBand (Exadata nodes to SMP cluster) · Head node engagement in all data redistributions
	Data Streaming	· FPGA performance assist on S-Blade – decompression, predicate filtering, row-level security enforcement · >95% of work done on S-Blades	· Exadata nodes primarily used for decompression and predicate filtering. · Most DW & Analytics work done in SMP head node
	In-DB Analytics	· Fully engaged MPP platform for analytics · User-defined functions, aggregates and tables · Language support: C/C++, Java, Python, R, Fortran · Paradigm support: SQL, Matrix, Grid, Hadoop · Built-in set of >50 key analytics (fully parallelized) · Open source: support for GNU Scientific & CRAN libraries · Integrated Development Env.: Eclipse & R GUI w/ wizards	· Analytics processing limited to head node cluster only · User-defined functions and aggregates · Language support: C/C++, Java · Paradigm support: SQL, Matrix (minor) · Basic analytics functions
	Scale	· Linear performance and data size scalability · Full-featured, enterprise-class work-load management & other features	· Non-linear performance & data size scaling – performance and i/o bottleneck at/to head node cluster
Simplicity	Appliance System Mgmt & Integration	· No tuning, no indexing, no partitions · Balanced system developed to deliver best price-performance	· Heavily tuned performance dependency · Performance depends on physical database design skills, including indices and partitions.

To compensate for the poor performance of Oracle RAC as an analytic database, Exadata contains a lot of hardware. The machine is expensive to acquire and expensive to run in the data center. Powering and cooling a single rack of Exadata consumes nearly 60 per cent more power than a single TwinFin rack.

While costs destroy value, a fundamental question is whether Exadata helps customers to create value. First-generation data warehouses play an important role in keeping an organization informed of the recent past, yet data unleashes greater potential through advanced analytics and other capabilities of second-generation warehouses discussed earlier in the paper. Oracle RAC teamed with traditional storage hasn't proved a success in this role to date. Exadata's storage tier is unable to process complex joins, distinct aggregations and analytical functions. It is difficult to envisage how two technologies, individually ill-equipped to analyze deeply very large data sets with high performance, will achieve this feat when connected by a fast network and housed in the same rack.

Value with Netezza TwinFin

Netezza's engineers integrate data management and analysis deep within massively-parallel, shared-nothing grids. One result of this innovation is simplicity for our customers, which translates directly to dramatically lower costs of owning and operating data warehouses than is possible with OLTP-oriented database products, such as Oracle's.

Demands on data warehouses have moved beyond processing simple SQL; to fully exploit data requires the warehouse be capable of running spatial analyses, predictive models, investigative graphs and other analytic applications. To illustrate, an organization wanting to assess risk created by catastrophes such as wildfire and flooding on a large portfolio of insured properties built an application using Oracle Spatial and Oracle 10g to identify insured assets in hazard-prone areas. Analyzing more than 4 million sites requires spatial joins and processing of computationally intensive spatial algorithms. Assessing wildfire risk requires 43 billion computations on 10,000 polygons, while assessing flood risk requires 30 billion computations on 7,000 polygons. Oracle took more than one hour to complete each assessment. Winds suddenly change direction and intensity; floods surge. The dynamic behavior of natural hazards means risk profiles can change enormously within one hour. The company required greater agility, to be free to analyze the situation whenever they thought it necessary for their business. Unlike Oracle 10g or Exadata, Netezza runs the spatial joins and computationally intensive algorithms of spatial analysis in its massively parallel processing grid. Netezza completes the wildfire risk assessment in 1.6 seconds and the flood risk assessment in 4.4 seconds, giving the company the agility to derive full value from their data.

TwinFin is designed from the ground up for massively parallel processing both SQL and the applications of advanced analytics. Netezza frees organizations from proprietary languages. Partners and customers can port existing applications to TwinFin or choose to develop new analytic applications in the language of their choice, including C++, Java, Python, R and Fortran. Organizations using C can take advantage of more than 1,000 analytic functions available as free software from the GNU Scientific Library.[12] Organizations using R can also make use of the more than 2,000 packages publicly available in the Comprehensive R Archive Network (CRAN).[13] Additionally Netezza's partners and customers can choose to work with MapReduce / Hadoop as, for example, a highly scalable ingestion mechanism to preprocess enormous data sets generated by public facing web applications and web logs before they are loaded into TwinFin for analysis.

[12] *www.gnu.org/software/gsl/*

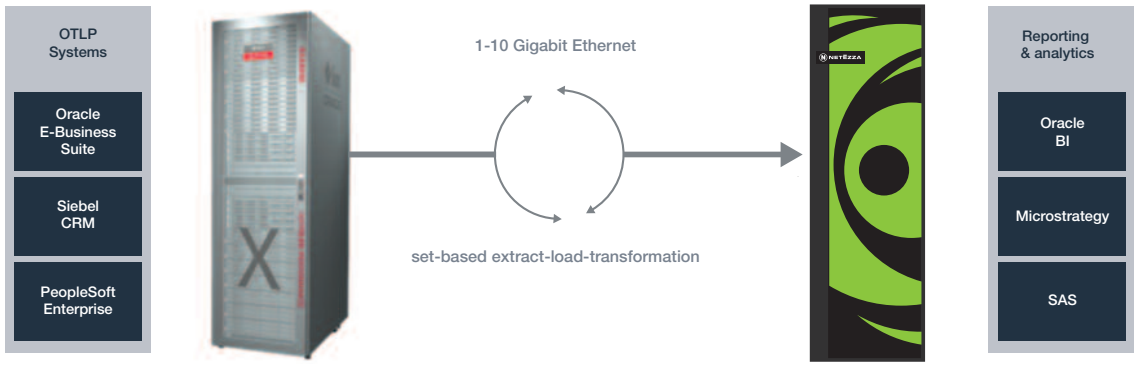
[13] *http://cran.r-project.org/*

6 Teaming Exadata with TwinFin

Section 2 discusses the differences between online transaction processing and data warehousing; OLTP's optimizations, data caching strategies and performance tuning have little relevancy to data warehousing. With its database designed for OLTP, Exadata offers CIOs the opportunity to consolidate multiple disparate transactional systems, each with their own hardware and Oracle databases, into a single manageable environment.

This consolidation likely presents opportunities to reduce costs and improve performance of extract-transform-load (ETL) stages used to feed the warehouse from source systems. As they add more sources to their first generation warehouses, some organizations unwittingly create a complex web of ETL. Many Oracle customers build their own ETL applications using Oracle's database with PL/SQL and other syntaxes and interfaces proprietary to that database system. While PL/SQL's row-based processing may have been fast enough to transform relatively low data volumes typical of early first generation warehouses, the big data sets normal in second generation warehouses require the higher performance available in set-based processing. Oracle's GoldenGate software and other products from specialist ETL vendors offer this capability. Although beyond the scope of this paper, Netezza Migrator™ and Netezza Data Virtualizer™ can also contribute to the success of ETL re-engineering projects and accelerating query performance.

Exadata, running multiple Oracle applications, passes data to TwinFin for Analysis



7 Conclusion

Netezza has emerged as the principal alternative to Oracle for data warehousing. Moving data warehouses and marts from Oracle to Netezza creates new opportunity, not risk. A majority of Netezza customers have already walked this path, many of them by partnering with system integration companies with strong track records for successful migrations.

Exadata is an evolution of Oracle's OLTP platform. Oracle's database management system is designed for OLTP where data volumes are relatively modest compared to data warehouses. The database activity of an OLTP system can be assessed before it is put into production; administrators have the time to design, test and optimize each transaction's data retrieval. Data warehouses must immediately process whatever query the business needs to ask of their data; technologies requiring administrator mediation are ill-suited to the task. Conscripting this technology into a role other than transaction processing places enormous stress on people and processes harnessed to manage and operate a data warehouse.

> ***"This [Netezza] is the first database product with a long term product roadmap that aligns perfectly with our own roadmap. We call this our on-demand database."***
>
> *– Steve Hirsch, Chief Data Officer at NYSE Euronext*

Oracle advises customers that Exadata is architecturally similar to Netezza but better because TwinFin doesn't support every data type or SQL standard, and that it doesn't support data mining or high concurrency. Netezza's customers disagree: "This is the first database product with a long term product roadmap that aligns perfectly with our own roadmap. We call this our on-demand database,"[14] said Steve Hirsch, Chief Data Officer at NYSE Euronext.

[14] *www.netezza.com/customers/nyse-euronext-video.aspx*

Given their different workload characteristics, few customers attempt to run OLTP and data warehouse systems on the same infrastructure; doing so demands constant tuning and optimizing. Technicians are placed in a difficult situation: either accept compromised performance for both OLTP and data warehousing, or ceaselessly reconfigure the database in a vain attempt to satisfy conflicting demands of the different workloads. Organizations will continue to run their OLTP and warehouse systems on different platforms, each specifically configured to the needs of their workloads. Organizations planning to use Oracle Exadata for OLTP[15] can get the best of both worlds by pairing it with Netezza TwinFin for data warehousing.

The only data warehouse that really matters is your data warehouse – your applications running on your data in your data center. An on-site proof-of-concept (PoC) creates the opportunity for an IT department to thoroughly investigate a technology, learning how they can use TwinFin to help their business peers extract greater value from data. Making the most of this opportunity requires the PoC to be managed with the same discipline afforded other projects. Curt Monash offers sage advice in his blog "Best practices for analytic DBMS POCs,"[16] including involving an independent consultant to steer the project to a successful outcome. For organizations wanting to understand how their warehouse performs on TwinFin, at no cost and with no risk, Netezza offers its Test Drive.
To book one, go to www.netezza.com/testdrive.

Netezza TwinFin: to use it is to enjoy it.

Learn More

Netezza is so sure you'll like the Netezza TwinFin solution, we invite you to try it at your site with your data – at no charge. Taking the TestDrive is easy.
Learn more here: www.netezza.com/testdrive

[15] See Curt Monash's DBMS2 Blog http://www.dbms2.com/2010/01/22/oracle-database-hardware strategy/ for his discussion of the role of Exadata-like technology as a platform for consolidating an enterprise's many Oracle databases rather than running a few heavy database management tasks.

[16] http://www.dbms2.com/2010/06/14/best-practices-analytic-database-poc/" \l "more-2297

Pass it Along

Please share this eBook with your friends and colleagues.

- Send it
- Post a link: *www.netezza.com/eBooks/Exadata-TwinFin-Compared.pdf*

Give Us Feedback

What did you think about the ideas and arguments in this eBook? Let us know what you liked, disliked or might want to discuss further.

- Chat with the Netezza Community: *www.enzeecommunity.com/groups/twinfin-talk*
- Tweet about the paper: *#twinfin*

Contact Us

Visit our blog: *www.enzeecommunity.com/blogs/nzblog*
Visit the Netezza website: *www.netezza.com*
Visit the Netezza Community website: *www.enzeecommmunity.com*
Send us your comments: www.netezza.com/company/contact_form.aspx

About the Author

Phil Francisco, Vice President, Product Management & Product Marketing, Netezza

Phil Francisco brings over 20 years of experience in technology development and global technology marketing. As Vice President of Product Management and Product Marketing at Netezza, he fosters new business and product strategies, directs the product portfolio and drives product marketing programs. Prior to Netezza, Francisco was the Vice President of Marketing at PhotonEx, a leading developer of 40 Gb/s optical transport systems for core telecommunications network providers. Before PhotonEx Francisco served as Vice President of Product Marketing for Lucent Technologies' Optical Networking Group, where he worked with some of the world's largest telecommunications carriers in planning and implementing optical network solutions. Mr. Francisco holds a patent in advanced optical network architectures. He received B.S. in Electrical Engineering and B.S. in Computer Science degrees magna cum laude, from the Moore School of Electrical Engineering at the University of Pennsylvania. He earned his Master's degree in Electrical Engineering from Stanford University and completed the Advanced Management Program at the Fuqua School of Business at Duke University. Read Phil's blog: www.enzeecommunity.com/blogs/nzblog.

About Netezza

Netezza – Pioneer, Leader, Proven

With the invention of data warehouse appliances, Netezza revolutionized and simplified analytics for companies drowning in data and struggling to find the processing speed and power to analyze and understand what it all meant.

Today, Netezza is the data warehouse appliance leader, combining storage, processing, database and analytics into a single system that delivers 10-100x the performance, at one third the cost of other approaches. We are once again resetting the bar on price/performance. Think of Netezza as a Ferrari, with the price and efficiency of an economy car.

With hundreds of customers including Nationwide, Neiman Marcus, Orange UK, The Sherwin-Williams Company, Virgin Media and more, and offices worldwide, Netezza (NYSE: NZ) is a proven solution to the rising costs and complexity of data warehousing and analytics. Our impressive ecosystem of global partners, including Ab Initio, Business Objects, Cognos, EMC, IBM, Informatica, Microsoft, MicroStrategy, SAS and others and an expansive list of system integrator, reseller and developer partners worldwide means our customers can rest assured that Netezza will live comfortably within any existing infrastructure they have in place.

We introduced the world's first data warehouse appliance and challenged the status quo. We created a new market segment and set the agenda for an entire industry. And now, driven by our customers, we are developing solutions to solve bigger, more complex enterprise-wide challenges including global federation, data integration, legacy system integration, business continuity, advanced analytics and compliance. With our strong financial base, profitable business model, focus on growth and extremely enthusiastic “enzee” community, you can be sure that Netezza will always be a strong partner for analytics, scaling with you as your data and demands grow.

NETEZZA® an IBM® Company

Netezza Corporation
26 Forest Street
Marlborough, MA 01752

+1 508 382 8200 TEL
+1 508 382 8300 FAX

www.netezza.com

About Netezza Corporation:
Netezza, an IBM Company, is the global leader in data warehouse and analytic appliances that dramatically simplify high-performance analytics across an extended enterprise. Netezza’s technology enables organizations to process enormous amounts of captured data at exceptional speed, providing a significant competitive and operational advantage in today’s data-intensive industries including digital media, energy, financial services, government, health and life sciences, retail and telecommunications. Netezza is headquartered in Marlborough, Massachusetts and has offices in North America, Europe and the Asia Pacific region. **For more information about Netezza, please visit www.netezza.com.**