

WHITE PAPER

W I N T E R C O R P O R A T I O N

SUN ENTERPRISE SYSTEMS

*A Scalable Platform for Very Large
Data Warehousing with Oracle*

Specialists in the World's Largest Databases



www.wintercorp.com

SPONSORED
RESEARCH
PROGRAM

SUN ENTERPRISE SYSTEMS

*A Scalable Platform for Very Large
Data Warehousing with Oracle*

JUNE 2005



WINTER CORPORATION
411 WAVERLEY OAKS ROAD, SUITE 327
WALTHAM, MA 02452
617-695-1800

Executive Summary

The need for proven large-scale architectures to support the requirements for Very Large Data Warehousing (VLDW) has never been greater. Data warehouse sizes continue to grow rapidly as businesses attempt to extract more value from existing operations. More data is being analyzed—with more complex questions being asked by more users—to provide insights into customer behavior, supply chain optimization, and other aspects of business operations. These forces continue to push database products across new frontiers of performance and scalability, along with the requirement for increased levels of system reliability, availability, and manageability.

For more than fifteen years, Oracle has been a database market leader and, in the past five years, has significantly increased its presence in the VLDW market. The foundation of Oracle's data warehousing success is a series of product enhancements that facilitate the management of large volumes of data accessible to a growing number of users. New features have been introduced in the past several major releases in the areas of parallelism, partitioning, clustering, transportable tablespaces, materialized views, compression, and analytic functions. The latest Oracle release, Oracle Database 10g, continues to build on this foundation.

Sun has been a key partner of Oracle throughout its years of database market success, providing a high-performance technology infrastructure for realizing the benefits of Oracle's VLDW capabilities. The Sun-Oracle alliance has included a close and effective engineering partnership for decades, which adds to the benefits of the Sun-Oracle VLDW solution along with the technology embodied in the Sun Fire server family, Sun StorEdge systems, and Solaris operating system. As further evidence of the Sun-Oracle focus on high-end data warehouse performance, an industry-leading TPC-H result was announced on January 31, 2005—3 Terabyte scale performance of 59,435 QphH, more than 30% higher than the previous leading result.

In this WINTER CORPORATION WHITE PAPER, we describe the features of the Oracle database and of Sun Fire servers that contribute to a solution that is able to satisfy the performance, scalability, availability, and manageability requirements of large-scale data warehouses. We also describe real customer data warehouses that have reaped benefits from the Sun-Oracle technology alliance and are successfully meeting their performance and reliability objectives.

Table of Contents

Executive Summary 3

1 Business Requirements for Very Large Data Warehousing 5

 1.1 Scalability and Performance 5

 Data Volume 5

 Large User Population 5

 Multiple and More Complex Applications 5

 1.2 Availability and Reliability 6

 1.3 Manageability 6

 1.4 Goals of the Paper 6

2 Oracle—A Database for Very Large Data Warehousing 7

3 Sun Enterprise Systems—An Infrastructure for Oracle VLDW 9

 3.1 Scalability and Performance 9

 3.1.1 Server Technology 9

 3.1.2 Solaris Operating System 10

 3.1.3 Scale-out via Sun Clustering Technology 11

 3.1.4 Enterprise Storage Technology 11

 3.1.5 TPC-H—A Significant Benchmark Proof Point 12

 3.2 Availability 12

 3.3 Manageability 13

 3.4 Future Directions 13

4 Customer Experiences 14

 4.1 Anonymous Customer 14

 4.2 Solectron Corporation 14

5 Conclusions 16

1 *Business Requirements for Very Large Data Warehousing*

The corporate appetite for information is expanding like never before. Users want to make fact-based decisions using the latest data. They want to move from debating what the facts are to analyzing what the facts mean. They also recognize that optimal solutions to business problems require insights from multiple disciplines and business areas. These forces are driving information technology (IT) organizations to integrate operational information gathered from across the corporation into a central information pool, the enterprise data warehouse (EDW). As the data warehouse expands to encompass more of an enterprise's information assets, the requirements on the data warehouse become more demanding and the ability of the hardware and software components of the warehouse to satisfy these requirements becomes more problematic.

These forces are not new, but their effects continually push database products across new frontiers of performance and scalability. As these data warehouses get larger, they also require increased levels of system reliability, availability, and manageability. IT executives seek to satisfy these requirements on a proven architecture that has already demonstrated its enterprise mettle—to minimize the risks associated with the project.

1.1 SCALABILITY AND PERFORMANCE

Data Volume

The increased integration of information inherent in the EDW drives the growth of the data warehouse along several dimensions. The most obvious area of growth is the size of the database. There is one constant in data warehousing today: volumes will grow. According to the 2003 Winter TopTen survey, the largest data warehouses are approaching 30 terabytes (TB) of user data and are expected to more than double in size over the next few years. In general, query processing on 50 TB of user data will consume more computer resources than on 5 TB or 500 GB. Nevertheless, analysts want their answers in roughly the same elapsed time, and database loads and extracts cannot exceed business requirements for timeliness of information.

Large User Population

As data from more areas of the business are incorporated into the EDW, the population of database users grows. In many organizations, thousands of internal users throughout the company access the EDW. Increasingly, companies are opening their databases to vendors, customers, suppliers, and other external business partners, further fueling growth of the user base. This growth of the user population drives up the volume of concurrent activity demanded from the system. Here too, however, users require the same response time despite the increased throughput demands placed on the system.

Multiple and More Complex Applications

The nature of the workload comprises a third dimension of EDW growth. More subject areas and the exploratory nature of data analysis increase the likelihood of new combinations of data. The breadth of novel data combinations generates queries that are more complex. In addition, competitive pressures on the business produce analytics of increasing sophistication, yielding query logic of growing complexity. Finally, the growing user base wants to pursue many different tasks through a diverse mix of applications, from business operations to data mining. The total workload mix is therefore increasingly diverse. More complex queries and a more diverse query mix both add to the system workload. This workload must be accommodated within ever-tightening time constraints for the delivery of results.

Meeting the business requirements driving the EDW, as ever more data becomes both accessible and critical to profitability growth, imposes serious demands on database technology. Not only do database products need to demonstrate broad-based efficiency and scalability today, but since the forces

A W I N T E R C O R P O R A T I O N W H I T E P A P E R

fueling database growth are accelerating, the underlying database architecture also needs to support continuous, rapid performance and scalability improvements in the future.

1.2 AVAILABILITY AND RELIABILITY

Along with a growth in data size and workload, the role that the EDW plays within an organization has become more visible and more critical to the success of the business. Through the 1990s, business intelligence systems proved their ability to help control costs and increase sales. If it was once acceptable for a data warehouse to be down for a day, that is the exception today. Many enterprise data warehouses are now considered mission-critical.

In addition, analytic results from business intelligence applications are frequently being fed back to operational systems to influence customer treatment, user behavior, inventory management, logistics, and other operational decisions. Increasingly this feedback mechanism needs to be dynamic and up-to-the-minute. These business operations are also increasingly global and need the latest analytic results round the clock.

While EDW systems are increasing in size and complexity, reliability and availability demands are also increasing rapidly. Keeping the data warehouse up and running is vital. Data and data integrity must survive inevitable disk failures. Reviving a system within minutes of a processor failure, without incurring major system degradation, is increasingly important. Organizations not only require a data warehouse that can grow with them as their data needs explode, but also require that their data warehouses have extensive safeguards to keep them running virtually flawlessly and to protect the value of their investment in the data itself.

1.3 MANAGEABILITY

As these systems grow in size and complexity and acquire ever more stringent reliability constraints, the challenge and costs of managing them—health checks, preventative maintenance, system upgrades, performance monitoring—should not grow nearly as fast as the growth rate of the system itself. The increasingly competitive business environment not only drives the effort to extract more value from a larger data volume, but also forces lower unit support costs. This climate will not permit the achievement of linear scalability in performance to imply a linear increase in administrative costs. As they grow, data warehouse systems must be easier and cheaper to manage than ever.

1.4 GOALS OF THE PAPER

In this WINTER CORP White Paper, we examine how well a Sun-Oracle solution meets these technical requirements. We explore the scalability and performance of the Sun technology architecture and the Oracle database architecture and their implications to the reliability, availability, and manageability of the data warehouse. We also describe representative examples of Sun-Oracle warehouses that demonstrate these advantages.

Sun has been a pioneer in high-end computing technology and Oracle has demonstrated leadership in solving the most demanding database problems. We believe this review of their architectures and their technology partnership reveals distinct advantages when employing Sun and Oracle for very large-scale enterprise data warehouses.

2 Oracle—A Database for Very Large Data Warehousing

Oracle has been a database market leader for more than fifteen years with particular strengths in transaction processing. In the past five years, Oracle has rapidly increased its presence in the large-scale data warehousing market. Over several releases, Oracle has developed and matured the necessary database infrastructure to support very large data warehouses. Today, as the results of the 2003 Winter Corporation TopTen Survey confirm, some of the largest data warehouses in existence at global businesses run on Oracle. In fact, the largest data warehouse in the latest survey is a 30 TB Oracle database at telecommunications giant, France Telecom, and the heaviest peak data warehouse workload recorded by the latest survey also runs Oracle, a Real Application Cluster configuration at Experian Marketing Services, hosted on a Sun Microsystems platform.

The foundation of Oracle's recent data warehousing success consists of a number of key building blocks introduced over the past several major releases of the product that enable successful management of very large volumes of user data and its efficient access by a growing population of users. These building blocks include parallelism, partitioning, indexing, real application clusters, transportable tablespaces, materialized views, compression, and analytic processing features.

- **Parallelism** – Oracle introduced query parallelism in the early 1990s and today offers comprehensive and mature parallel processing of all SQL operations. Parallelism improves response time on operations against large volumes of data, and is critical for analytic queries, decision support reporting, and database maintenance operations on very large Oracle data warehouses.
- **Partitioning** – Partitioning is a technique for breaking up very large tables and indexes into smaller pieces. This allows processing intensive operations, like table scans, to be easily divided among multiple parallel tasks, thus dramatically improving response time for these operations. It also provides an opportunity to quickly eliminate partitions that are not affected by particular operations. For example, new inserts into a time-partitioned table only affect the latest period; all other partitions can be ignored by these inserts. This not only improves performance by eliminating unnecessary work, but may also improve increase multi-user concurrency. Oracle also supports composite partitioning, using multiple partitioning strategies, such as range partition with hashed sub partitions, to maximize opportunities for exploiting parallelism. In Oracle, indexes can also be partitioned, limiting the need for index rebuilds following updates.
- **Indexing** – Oracle provides a rich set of indexing options including B-tree indexes, Bitmap indexes, function indexes, join indexes, and cluster indexes. B-tree indexes allow efficient, ordered access to data that is especially good at range and wild card searches. Bitmap indexes excel at resolving multi-key restriction predicates, often critical in business intelligence queries. Function indexes allow rapid searches against algorithmic derivations of column data; for example, searching for particular months using a date column. Join indexes allow queries to quickly retrieve information from multiple tables, often directly from the index itself. Finally, cluster indexes allow data from multiple tables to be stored together for efficient access. Together these indexing techniques provide complementary performance characteristics, and, when used in conjunction with partitioning, can be used in parallel to provide scalable access to large data warehouses.
- **Clustering** – Oracle's Real Application Clustering (RAC) supports multi-node access to a shared data warehouse. Parallel threads executing a single query can be distributed across the nodes of the cluster. Modified shared pages can be transferred from node to node across the cluster interconnect at high speeds without being landed to disk first. Read-only data pages can be cached in each node's buffer pool to enable shared access by multiple nodes without requiring continuous transfer between nodes.

A W I N T E R C O R P O R A T I O N W H I T E P A P E R

- **Transportable Tablespaces** – Oracle provides the capability of attaching and detaching populated tablespaces to and from existing tables. By synchronizing tablespace and partition boundaries, this provides high performance support for rolling window operations on data warehouses, a common data warehouse paradigm. In this scenario, new data can be inserted in an offline partition and rolled into an active table on a preset schedule. Likewise, the oldest table partition can be rolled out when its data becomes obsolete.
- **Materialized Views** – materialized views are used to compute and store commonly aggregated or joined data. A refresh of the materialized views is required when the underlying master table data changes, and the refresh method can be either immediate or deferred. The cost-based query optimizer can choose to use the materialized view to resolve a query instead of using the master tables when it provides improved performance.
- **Table Compression** – a dictionary compression algorithm eliminates data redundancy at the table or partition level. This can be used to improve storage efficiency, especially in read-only or read-mostly partitions of large data warehouse tables without limiting functionality. Oracle reports typical compression ratios ranging from 3:1 to 5:1
- **SQL99 Analytic Functions** – implementation of a broad range of statistical functions common to business intelligence applications harnesses the power of the parallel query engine for a larger share of the analytic workload. Large performance efficiencies can result from pushing more of the complex analysis into the query processing and closer to the data.

The latest Oracle release, Oracle Database 10g, continues to build on these capabilities with numerous performance improvements and ease of management features. These include improved table scan performance, partitioning memory efficiencies, and improved parallelization of correlated subqueries. Oracle 10g improves database management by simplifying the configuration of database instances, automating storage management via a brand new storage management facility, and adding a self-tuning SQL optimization capability. In addition, by enlarging the suite of supported analytic functions and supporting OLAP cubes virtually in relational tables, Oracle has significantly augmented its business intelligence capability.

These features allow Oracle to effectively exploit parallelism on very large data warehouses without sacrificing the ease of administration benefits of a shared data architecture. Oracle customers in ecommerce, retail, telecommunications, and other industries rely heavily on these features to support data warehouses containing terabytes—in some cases, tens of terabytes—of user data. Oracle's ongoing innovation in database and data management technology critical to VLDW applications, along with its strong marketplace leadership, has pushed it toward broad acceptance as a top tier VLDW technology provider.

3 Sun Enterprise Systems—An Infrastructure for Oracle VLDW

Throughout this period of rapid Oracle growth, Sun Microsystems servers have been a principal platform for deploying Oracle databases. There are more Oracle databases running on Sun and Solaris than on any other platform. Thousands of Sun-Oracle operational systems exist worldwide, and Sun has been leveraging this success to expand into data warehousing and business intelligence applications. In the 2003 Winter TopTen Survey, 46 Sun-Oracle systems were reported. Of these, 32 Sun-Oracle data warehouses were reported, 19 with validated user data volume greater than one terabyte. The largest was Claria Corporation—with a 12.1 TB data warehouse—who ranked seventh in the TopTen category, Database Size, DSS. The average data warehouse size for these validated Sun-Oracle customers was 3.5 TB.

Sun Microsystems offers a comprehensive enterprise system environment that offers an excellent technology infrastructure for realizing the benefits of Oracle's very large data warehousing capabilities. Sun invests heavily in innovative hardware and operating system features that complement and enhance the scalability and performance, availability, and manageability of the Oracle database engine, especially with features like partitioning and RAC. Sun and Oracle engineers work closely together to plan hardware and operating system features that extend Oracle capabilities and to enable Oracle to take maximum advantage of Solaris performance features like lightweight threads and memory placement optimization. These innovations and the cooperative engineering efforts that accompany and support them strengthen the competitive position of the Sun Enterprise platform as an option for businesses deploying critical Oracle Data Warehousing applications.

3.1 SCALABILITY AND PERFORMANCE

3.1.1 Server Technology

The Sun Fire server family scales up to 72 dual core processors containing up to 576 GB of physical memory and 72 hot swappable I/O slots that are capable of managing hundreds of terabytes of online storage. The Sun Fire server line is based on the 64-bit UltraSPARC chip technology and provides a balanced architecture well suited for the multi-user, I/O intensive workloads required by very large Oracle data warehouses.

A key to Sun's capability to support memory intensive data warehousing and business intelligence applications is its large processor count symmetric multiprocessor (SMP) architecture. This large-scale SMP architecture not only enables application of enormous compute power, but also supports efficient shared access to a very large common memory pool critical to caching large volumes of frequently accessed data. Sun's SMP architecture relies on a crossbar backplane to provide scalable, high bandwidth connectivity for up to 18 separate system boards, each containing four CPUs and up to 32 GB of RAM, with reasonable latency in the few hundreds of nanoseconds. While all major server vendors have since adopted a similar large SMP architecture, Sun was the first vendor to deliver this innovative SMP solution.

The ability to cache frequently accessed data improves database performance by servicing more data requests at memory speeds rather than disk or network speeds. In addition, large scale SMP systems provide the efficient data redistribution necessary to support the typical join and aggregate processing common to data warehouse and business intelligence applications.

Sun has also moved aggressively to develop a distinctive processor architecture. Sun has a history of processor innovation, starting with its adoption of 64-bit technology in the early 1990s. In its latest move, rather than focus on single process throughput, Sun has become an early proponent and adopter of placing multiple processors on a chip—what Sun calls Chip Multi-Threading (CMT). This innovation emphasizes multi-threaded performance, a key factor in data warehouse scalability. While all chip

manufacturers are pursuing variations on this technology, in fact, Sun has been the first manufacturer to market with large server exploitation of this “multi-core” technology.

Sun’s first generation CMT chip, the UltraSPARC IV, is a dual threaded processor that puts two UltraSPARC III cores on a single chip. Its largest configuration, the Sun Fire 25K, can contain up to 72 processors, or 144 processor cores. Early benchmarks have shown excellent scalability for Sun’s multi-core. On the SPECint test of processor performance, the UltraSparc IV turned in a peak performance rate 1.87 times that of the UltraSPARC III at the same clock speed. In a business application setting, on the Manugistics supply chain benchmark, the CMT chip achieved 1.92 times the throughput of its UltraSPARC III predecessor. These across-the-board results bode well for real world impact of CMT technology for Oracle VLDW applications.

Rapid growth in processor scale necessitates performance and scalability improvements in the I/O subsystem and operating system. In its latest generation of Sun Fire systems, Sun is releasing an upgraded hot swappable I/O interface. The Solaris operating system has also kept pace, with sophisticated multi-threading efficiencies that yield a performance boost and support for file system direct I/O that dramatically improves file system I/O performance.

3.1.2 Solaris Operating System

The Solaris operating system provides a mature 64-bit operating environment to support very large parallel Oracle warehouses that have prodigious I/O appetites and consume massive amounts of memory for database buffer pools. The Solaris and Oracle engineering teams collaborate closely to boost data warehousing performance and scalability. Two features stand out as significant examples of this engineering collaboration—multiple page size support and memory placement optimization (MPO). Multiple page size support, allowing operating system page sizes up to four megabytes, can dramatically reduce the system overhead of memory accesses in large memory applications like very large Oracle data warehouses. Oracle has been able to take advantage of this feature since Oracle9i.

Memory placement optimization allows the system to compensate for a significant scalability barrier encountered by applications exploiting very large SMP systems. Oracle’s shared disk architecture is ideally suited for SMP systems. In addition, SMP systems are often favored for large data warehouse applications because of the simplified system administration and database administration environment they offer. All large SMP systems, however, as noted above, scale up by distributing memory across a high bandwidth system interconnect. The Sun Fire SMP systems are no exception. This leads to non-uniform memory access times for retrieving data from memory locations further away from the active processor than from memory on the same processor board as the active processor. While the differential access time has decreased over the years, it still takes two to three times as long to access remote memory as local memory, and given the speeds of modern CPUs, this can result in a several thousand CPU cycle delay for a cache miss.

This inherent problem of SMP architectures results in decreasing throughput per processor as the number of processors in the SMP system rises. For large-scale data warehousing, this presents a barrier to scalability that is a critical problem to address. Solaris’s MPO feature directly attacks the memory latency problem by preferentially locating newly fetched data in memory local to the fetching processor. Oracle has been MPO aware since Oracle9i. When combined with processor affinity—a feature that schedules processes on the same locality group—and effective use of Oracle partitioning to maintain locality of reference at the database level, remote memory accesses in large-scale Oracle data warehouses can be dramatically curtailed.

Finally, in today’s 7x24 computing environment, planned downtime for system upgrades must be minimal, or even non-existent. To insure continued scalability as requirements grow and to permit recovery from component failures, VLDW systems need to provide a non-disruptive system upgrade

capability. Solaris provides a dynamic reconfiguration facility that permits system components to be added or changed while the system continues to run. It also provides a system interface allowing applications like Oracle to automatically respond to reconfiguration events including changes in the number of CPUs or the amount of physical memory in the system, under the direction of a Solaris service, the Reconfiguration Coordination Manager. Since Oracle9i, Oracle has exploited Solaris's DR facilities to support dynamic resizing of the System Global Area (SGA). This coordination between Oracle and Solaris enables system reconfiguration of very large Oracle data warehouses on Sun, e.g., adding system resources to meet emerging requirements while Oracle continues to operate.

CMT, Multiple Page Sizes, Memory Placement Option, and Dynamic Reconfiguration are some of the key server and operating system features that provide a high performance infrastructure for Oracle VLDW applications and dramatically boost their scale up potential. In Solaris 10, Sun has enhanced system performance further through additional kernel and network efficiencies that lower CPU cost and improve multi-user performance by reducing kernel and network latencies.

3.1.3 Scale-out via Sun Clustering Technology

As Oracle data warehouse size and workload expand beyond the capacity of a single large Sun SMP server like the SunFire 25K, multiple servers can be clustered via one of multiple interconnect technologies managed by Sun Cluster software and Oracle's Real Application Cluster (RAC) feature. The best cluster performance can be obtained by using the scalable coherent interface (SCI) interconnect that provides 200 MB per second bandwidth and exploits remote shared memory to insure low latency (4 μ s.). Sun Cluster environments support Oracle9i and Oracle 10g RAC configurations. Additional SCI links can be added for increased bandwidth.

Up to four Sun Oracle servers can be connected over SCI via redundant switched connections. This accommodates a theoretical scale-out solution harnessing the power of 288 processors (576 UltraSPARC IV cores). The maximum number of RAC nodes in a Sun Cluster environment is expected to grow over time.

3.1.4 Enterprise Storage Technology

To manage the enormous quantities of data concentrated in very large Oracle data warehouses, Sun markets the StorEdge 9900 Series of enterprise storage systems. The StorEdge 99xx SAN storage systems have a capacity of up to 1152 disk drives with a 128 GB data cache that can be linked to database servers by up to 64 2 GB per second redundant fibre channel connections. The StorEdge 99xx systems are capable of transferring data at 10.6 GB per second. Configured with 73 GB drives in a 4+1 RAID 5 setup, this provides approximately 60 TB of storage in a single unit. Multiple 99xx units can be configured, providing highly scalable capacity for very large Oracle data warehousing applications.

The StorEdge 9900 Series combines dense packaging with sophisticated storage management software that supports: storage virtualization for flexible allocation; graphical administration tools for simplified storage configuration; tuning and management; comprehensive fail-over support for availability; multi-device backup and restore support; remote online copy support for disaster recovery; and integrated security support. Storage can also be added to the system non-disruptively, and can be immediately put to use by Oracle Automated Storage Manager (ASM). The StorEdge 9900 Series integrates with Sun Cluster Manager and also supports popular third party tools like Veritas Volume Manager and Cluster Manager, and Oracle's ASM, released as part of Oracle Database 10g. A newly released product, SAM/FS bring information lifecycle management (ILM) capabilities to Sun's storage offerings that allows "cold" data to be maintained on lower cost media.

Sun storage systems are fully capable of handling customer's very large, mission-critical data warehouses.

3.1.5 TPC-H—A Significant Benchmark Proof Point

TPC-H is a decision support benchmark, consisting of a suite of business oriented ad-hoc queries and concurrent data modifications typical of systems that examine large volumes of data, execute queries with a high degree of complexity, and give answers to critical business questions. While no benchmark can predict the performance of a particular customer's application, TPC-H is the most frequently quoted metric to predict the relative performance of different system platforms and configurations for data warehousing applications. Platforms that consistently produce leading TPC-H benchmark results are likely to demonstrate similar performance advantages on similar application workloads. The TPC-H benchmark can be scaled by database size from 100 GB to 10 TB.

The TPC-H result announced on January 31, 2005 is a 3 TB scale benchmark on a 72-way 1.2 GHz UltraSPARC IV processor-based Sun Fire E25K running the Oracle 10g database. The Composite Query-per-Hour Performance (QphH) was 59,435@3TB—more than 31% greater than the previous leader, an HP Integrity Superdome. The Transaction Processing Council (TPC) also releases relative price-performance metrics for the TPC-H results. Significantly, the recent Sun result had an excellent price-performance of \$101/QphH, 7% less than the price-performance metric for the lower-performing HP system.

Compared to earlier Sun TPC-H results at the same 3 TB scale, these new UltraSPARC IV results indicate excellent scalability for the new CMT chips over the previous generation UltraSPARC III single-core chips running at the same clock speed. While the number of processors remained the same in both benchmark tests (72 processors) the query performance results (QphH) more than doubled, from 28,948 to 59,435, while the price-performance (\$/QphH@3000GB) of the latest results is 38% lower than the price performance of the earlier test. While differences in disk configuration and Oracle version between the two tests accounts for some of these differences, these results show significant scalability on data warehousing workloads for Sun's CMT architecture.

3.2 AVAILABILITY

Sun Fire servers provide full hardware redundancy and are designed to automatically detect, diagnose, isolate, and recover from hardware failures. In addition, Sun servers provide facilities to allow online system expansion and repair of all components, not just hot-swappable disk. With version 10, Solaris introduced fault and service managers that provide automatic online error detection and automate system recovery, by detecting failing components and removing them before failure, then restarting applications and services that were using the failed resource.

Sun Cluster Manager is integrated with Oracle RAC. Sun Cluster Manager software provides a heartbeat mechanism for all nodes participating in the cluster. In a high availability configuration, the Oracle RAC software supervises fail-over processing. The failing node is shutdown as gracefully as possible and any in-flight transactions initiated by users on that node are rolled back. Oracle RAC provides a Transparent Application Failover facility that automatically reconnects these users on surviving nodes in a balanced fashion, preserving their login sessions. In-progress queries are restarted automatically, and, except for an extended response time, the restart is transparent to the user.

To provide comprehensive disaster recovery solutions, Sun has also developed and validated a business continuity infrastructure that supports clustering and SAN data access across distances up to 200 kilometers. Area-wide clustering can be augmented by offsite media storage, electronic vaulting, and database replication to provide a flexible business continuity solution. Taken together, these capabilities allow enterprises to implement a solution to maintain data availability in the face of unplanned system outages. The customers we interviewed indicated their satisfaction with Sun's uptime record in their large data warehousing environments.

3.3 MANAGEABILITY

Solaris provides a broad suite of management tools to help system administrators effectively manage the complex hardware and software environments typical in large data warehouse systems. These include Solaris Volume Manager, Resource Manager, and Network Resource Manager. Solaris Volume Manager provides RAID services, logical volume management, and storage monitoring. Solaris Resource Manager provides monitoring and control of system resources, while the Network Resource Manager provides monitoring and control of network resources. These tools are all integrated components of Solaris available at no added charge. Sun also partners with 3rd party suppliers like Veritas, to provide an open management environment that fits into existing system management infrastructures.

3.4 FUTURE DIRECTIONS

Sun is continuing to develop innovations that directly influence Oracle VLDW performance and scalability. The next generations of CMT chips will feature expanded multi-threading, leading to eight cores per chip, each running four independent threads—a sixteen-fold hardware scalability improvement. Solaris 10, released in late 2004, has a number of performance features, including a rewritten TCP stack, improved thread management and more efficient system call libraries. Finally, Sun is working with Oracle to incorporate exploitation of software multi-threading, which offers significantly lower overhead than Oracle's traditional multi-processing approach. This is scheduled to ship in the coming months with Oracle 10g Release 2.

Over the last two decades, Sun has demonstrated technological leadership in hardware and operating systems. Solaris is still regarded by many as the gold standard of Unix operating systems—and Sun is maintaining a strong R&D commitment to continuous innovation for the foreseeable future. Significantly, Sun's greatest strength may be its large loyal customer base. Many of today's IT leaders were originally attracted to Sun in the 1990s because of its technology. The customers we spoke to noted that Sun was the only server player at the time of their selection willing and able to tackle their particular problems. Successful implementations, augmented by a strong, targeted service and support organization, cemented this bond and paved the way for expansion into new application areas, like business intelligence. Together, Sun's future innovation and the strength of past successes should allow Sun to expand its presence in the business intelligence market.

4 Customer Experiences

4.1 ANONYMOUS CUSTOMER

This Anonymous Customer processes millions of rows of data per day. Until 2002, the business ran on many, separate Windows servers (up to eight processors per), but, as its business became more sophisticated and the queries became more complex, it became increasingly difficult to satisfy business requirements. Business problems were no longer neatly segmented into functions—for example, business development and sales forecasting. Federated databases were utilized with some success, but the scalability limitations of a single server posed an obstacle to the development of more powerful applications and inhibited sophisticated ad hoc analysis.

The solution was a large-scale data warehouse consolidation with small data marts of aggregate data, such as sales reports. The Oracle database was selected over other scalable database platforms, such as DB2 and Teradata, partly because skilled Oracle users and administrators were readily available to the Anonymous Customer, and also because, at the time, the Anonymous Customer concluded that an Oracle solution was the most favorable cost alternative. Sun was chosen as the server platform because of hardware compatibility across its server product family and because the Anonymous Customer believed that Solaris was easier to tune, customize, and integrate with its storage system than other UNIX platforms for Oracle. The Anonymous Customer believes that its Sun E15000 database server has plenty of room to grow (currently 40 processors, soon expanding to 60 processors), and is very pleased with its performance on a very demanding multiprocessing workload (60-700 parallel threads of execution).

The net result of the data warehouse consolidation is that analytics that ran for 36-40 hours can now be completed in four hours. A great deal of time was spent prior to consolidation locating and moving data to meet an analytical requirement. The Anonymous Customer estimates that its IT operations cost has been reduced by 75%. Furthermore, unscheduled system downtime has been eliminated over the past three years. The operational improvements on its core application have allowed the Anonymous Customer to expand functionality, adding a second data warehouse, being built on a similar Sun-Oracle infrastructure. This is a testimony to customer satisfaction—doing it again the same way.

4.2 SOLECTRON CORPORATION

Solectron is a leading provider of electronics manufacturing and integrated supply chain services. With a presence in 50 countries, Solectron provides 7x24 access for customers worldwide—to a data warehouse that currently exceeds one Terabyte but is expected to grow to 3 TB within two years. The performance critical application at Solectron involves re-loading 600 Gigabytes into the data warehouse every day—within a short timeframe and without downtime.

The data warehouse requirements grew, following a series of acquisitions by Solectron over the past decade, merging multiple ERP systems and data from many different sources. In 2000, the architecture migrated from multiple commodity (Windows NT-based) servers to a “best of breed” approach with the database (DB2), the ETL processes (Informatica), and query server (Business Objects) on servers from multiple vendors running multiple operating systems. By 2002, the database and ETL servers were reaching their scalability limits and the query server required frequent rebooting. Solectron concluded that a single, scalable SMP was the best hardware alternative for the data warehouse. Oracle on a large SunFire server was selected because of the availability of resources and the knowledge base for Oracle and Solaris.

The new architecture was operational by the end of 2003. A key to the success of the solution has been Oracle transportable tablespaces, which allow rapid switchover to the latest data that had been held in an Oracle staging area. Whereas actual data movement of 200 GB across a fibre channel could take

A W I N T E R C O R P O R A T I O N W H I T E P A P E R

as many as eight hours, the export/import and cataloging within the same Sun 6320 SAN storage took only five minutes with transportable tablespaces. The resulting benefits for Solectron customers have been improved system availability and broader functionality for the business analysts. Solectron has strong praise for the Sun-Oracle technical support team that allowed the project to be completed on time and under budget.

5 *Conclusions*

The most recent WinterCorp TopTen survey results demonstrate Oracle's dramatic growth in the VLDW market. Winter Corporation anticipates a continued growth in Oracle's presence in this market based on the additional performance and system management features incorporated in Oracle Database 10g. Many of the largest Oracle data warehouses run on Sun's scalable SMP platform. Sun Microsystems has demonstrated a history of technology innovation, embodied in its high-end Sun Fire server family and Solaris operating system. Recent platform enhancements, such as chip multi-threading and the memory placement option, are evidence that this track record is continuing. The Sun-Oracle partnership benefits from their effective joint engineering efforts, resulting in a robust platform for supporting VLDW. A recently announced TPC-H benchmark result—59,435 QphH@3TB, #1 for this benchmark—provides additional evidence of the strength of this alliance and the scalability of the platform. We expect these developments to increase the enthusiasm from Sun's loyal customer base.

*A leading center of expertise in very large databases,
Winter Corporation provides services in
consulting, research, architecture and engineering.*

*We help users and vendors understand their opportunities;
select their database and data warehouse platforms;
define and measure the value of their strategies, architectures and products;
plan, architect and design their implementations;
and manage their scalability, performance and availability issues.*

Our focus is databases near, at and beyond the frontier of database scalability.



WINTER CORPORATION

411 WAVERLEY OAKS ROAD, SUITE 327
WALTHAM, MA 02452
617-695-1800

visit us at www.wintercorp.com