Migrating Databases from Sun Systems to Dell Servers

Running Red Hat Enterprise Linux AS 3

To take advantage of high-performing, industry-standard Intel® processor–based servers, many enterprises are migrating their mission-critical applications from proprietary UNIX® systems to cost-effective Dell™ servers. To demonstrate the ease and benefits of migrating from Sun Solaris–based servers, Dell engineers moved a large (100 GB) database—representing an online DVD store application and running on a leading enterprise database manager—from a four-processor Sun Fire V440 server running Solaris 9 to a four-processor Dell PowerEdge™ 6650 server running the Red Hat® Enterprise Linux® AS 3 operating system.

BY TODD MUIRHEAD AND DAVE JAFFE, PH.D.

Enterprises evaluating the move from proprietary UNIX/RISC-based servers often consider migrating their applications to high-performing, industry-standard Dell servers. This article examines a migration to a Dell PowerEdge server running Red Hat Enterprise Linux AS 3, which can yield significant benefits in performance and costs.[1]

To demonstrate the ease and potential benefits of a migration from Sun Solaris–based servers, Dell engineers moved a large (100 GB) database running on a leading enterprise database manager from a four-processor Sun Fire V440 server running Solaris 9 to a four-processor Dell PowerEdge 6650 server running Red Hat Enterprise Linux AS 3. Following the migration, the test application ran more than twice as fast on the PowerEdge 6650 as on the Sun Fire V440. Both servers were configured with four of the fastest processors available at the time the tests were conducted. In a second test, the PowerEdge 6650 with only two processors was still faster and 44 percent less expensive than the four-processor Sun Fire V440.

Configuring the hardware

Conducted in July 2004 by a team of Dell engineers, this study involved the migration of a large (100 GB) database from a four-processor system, the Sun Fire V440, to another

[1] For more information about the benefits of migrating from the Sun Solaris platform to Dell servers running Microsoft Windows Server 2003, see "Migrating Databases from Sun Systems to Dell Servers Running Microsoft Windows Server 2003" by Todd Muirhead; Dave Jaffe, Ph.D.; and Kerstin Ackerman in *Dell Power Solutions,* October 2004.

Enterprises can easily move from UNIX/RISC-based servers to industry-standard Intel processor–based servers running Linux and save significantly.

four-processor system, the Dell PowerEdge 6650. The initial migration used the fastest configuration of both systems available at the time of the tests: the Sun Fire used four Ultra SPARC IIIi processors at 1.28 GHz and with 1 MB of level 2 (L2) cache; the Dell PowerEdge 6650 used four Intel Xeon™ processors MP at 3.0 GHz and with 4 MB of level 3 (L3) cache. Both systems had 16 GB of memory. Visit *Dell Power Solutions* online at http://www.dell.com/magazines_extras for complete configuration and price information.

After successfully migrating the database and completing the initial test, the Dell team ran a second test on a less-expensive PowerEdge 6650 configuration with two Intel Xeon processors MP at 3.0 GHz and 4 MB of L3 cache. The list price of the database manager license, based on processor count, was included in the calculation of the total price. The OS cost was included in the server hardware price.

Red Hat Enterprise Linux AS 3 was installed on the Dell PowerEdge 6650 and Solaris 9 12/03 was installed on the Sun Fire V440. Because of the 32-bit architecture of the Dell servers, the test team had to add parameters to Red Hat Enterprise Linux to enable the database to use more than 4 GB of memory. The team created an in-memory ramfs file system in addition to enabling hugetable support.

A storage area network (SAN)–attached Dell/EMC CX600 Fibre Channel storage array was used for external storage by the Dell and Sun servers. The servers were connected to the SAN via QLogic host bus adapters (HBAs). Each server was assigned to an identical set of logical storage units (LUNs) that used the same number and type of disk drives. The Dell and Sun servers as well as the driver and firmware levels of the QLogic adapters were configured based on the EMC compatibility matrix.[2] Visit *Dell Power Solutions* online at http://www.dell.com/magazines_extras to see the storage components used in the test. Pricing of the configurations did not include the SAN hardware or software.

Setting up the database and performing the migration

A leading enterprise database server was installed and configured on both the Sun Fire V440 server and the Dell PowerEdge 6650 servers. Installation and configuration of this enterprise database server was completed according to the database provider's installation guide for the respective servers.

The 100 GB database was created on the Sun Fire V440 using scripts and then moved to the Dell PowerEdge 6650 server using migration tools included with the database server. Migration of the database included converting the data files to a Linux-compatible format, transferring the converted files across the network to the Linux-based PowerEdge server, and then connecting the transported data files into the running database instance on the PowerEdge server.

No data was lost during the migration, which took slightly more than two hours. The initial conversion of the tablespaces took 54 minutes, the ftp of the files across the network completed in 32 minutes, and the final step of moving the newly copied data into the database storage manager took 43 minutes.

The Dell team determined the initialization parameters for the database by enabling the auto-tuning features of the database server on the Sun V440 and letting the server run under load. The only parameter that was specified by the test team was a limit of 14 GB of RAM to be utilized by the database. The auto-tuned parameters were then applied to the database instance on the Linux server as closely as possible. Because of the differences in the memory models on the two platforms—64-bit versus 32-bit—the USE_INDIRECT_DATA_BUFFERS and DB_BLOCK_BUFFERS parameters were used on the Linux server, instead of the DB_CACHE_SIZE parameter, to specify the amount of memory. Visit *Dell Power Solutions* online at http://www.dell.com/magazines_extras to see the non-default database initialization parameters that were used in this study.

A total of 34 drives were used to set up the database tablespaces identically on each of the systems, with 30 drives for data and 4 drives for logs. Visit *Dell Power Solutions* online at http://www.dell.com/magazines_extras to see the tablespaces used in this study.

The application

The test described in this article used a different version of the online DVD store database application employed in previous Dell migration studies.[3] Called DVD Store 2, this application includes several features—the use of transactions, triggers, and referential integrity constraints—that are intended to model additional database management software features. In addition, functionality typical of some online e-commerce stores was added

The fastest Dell PowerEdge 6650 server, with four 3.0 GHz processors, handled 2.33 times more orders than the Sun Fire V440 system.

[2] The EMC compatibility matrix is available at http://www.emc.com/interoperability/index.jsp.

[3] For more information about the DVD Store application used in previous Dell migration testing, see "Migrating Databases from Sun Systems to Dell Servers Running Microsoft Windows Server 2003" by Todd Muirhead; Dave Jaffe, Ph.D.; and Kerstin Ackerman in *Dell Power Solutions*, October 2004.

to the DVD Store 2 application, including reporting previous purchases to the user and recommending titles enjoyed by others. However, this additional complexity significantly lowered order rates and changed the application profile, so results from the DVD Store 2 test scenario should not be compared to results from the original DVD Store program.

In a second test, the PowerEdge 6650 with only two processors was still faster and 44 percent less expensive than the four-processor Sun Fire V440.

The DVD Store 2 database—slightly larger than 100 GB and representing an online DVD store with 1 million DVD titles—was driven by a ProC-language program simulating users logging in to the online store; browsing for DVDs by title, author, and category; and then finally submitting an order. The driver program measured the number of orders per minute (opm) that the database could handle as well as the total response time as experienced by the simulated end users.

The database schema

The DVD Store 2 application used in the test scenario described in this article comprised six main tables. The Customers table was prepopulated with 200 million customers, with one logical partition containing 100 million U.S. customers and a second partition containing 100 million customers from the rest of the world. The Orders table was prepopulated with 10 million orders per month for 12 months; each month's orders were stored in a separate logical partition. The Orderlines table was prepopulated with an average of five items per order, also partitioned by month. The Products table contained 1 million DVD titles. Figure 1 shows the database schema for the online DVD store.

In addition, DVD Store 2 introduces the Reorder table. When the QUAN_IN_STOCK value for each product in the Products table falls below a specified value, a trigger is set off. This trigger calls code to write information about the product to the Reorder table. A separate process (not modeled in this study) can watch the Reorder table to initiate reordering of needed titles. Finally, the Categories table listed the 16 DVD categories.

The stored procedures

The DVD Store 2 database included two stored procedures for the login; three stored procedures for browsing by category, actor, and title; and a stored procedure for completing the purchase. A single order operation comprised a login, several browsing actions, and a purchase. Visit *Dell Power Solutions* online at http://www.dell.com/magazines_extras to see the Purchase stored procedure.

The Dell test team added several features to the stored procedures to model additional features used in some online e-commerce stores. During the Login procedure, for example, the customer's previous order (up to 10 titles) is reported, along with titles that other customers who purchased those titles have recommended. The Browse_by_category procedure returns those titles that are currently on sale in the specified category. And the Purchase stored procedure checks the QUAN_IN_STOCK field from the Products table to determine whether a title is available. This step is performed using a database transaction, so the QUAN_IN_STOCK data is not updated unless sufficient stock is available to fill the order. If insufficient quantity is available to fill the order, no new record is written to the Orders table.

The OLTP driver application

To model the online transaction processing (OLTP) order-entry workload, the Dell test team wrote a multithreaded program to drive a load against the database. Each thread of the OLTP driver application connected to the database and made a series of stored procedure calls that simulated customers logging in, browsing, and purchasing—without pausing to simulate customer think times or key times. By running the threads at full speed without simulation of user pauses, the database connections remained full and simulated what happens in a real-world multitiered application in which a few connections are pooled and shared among Web servers that may be handling thousands of simultaneous customers. In this way, the Dell team was able to achieve a realistic simulation without the need to model thousands of users.

Each thread of the OLTP driver application modeled a series of customers going through the entire sequence of logging in,

Table	Columns	Number of rows
Customers	CUSTOMERID, FIRSTNAME, LASTNAME, ADDRESS1, ADDRESS2, CITY, STATE, ZIP, COUNTRY, REGION, EMAIL, PHONE, CREDITCARD, CREDITCARDEXPIRATION, USERNAME, PASSWORD, AGE, INCOME, GENDER, PROD_ID_IDX, PROD_ID1, PROD_ID2…PROD_ID10	200 million
Orders	ORDERID, ORDERDATE, CUSTOMERID, NETAMOUNT, TAX, TOTALAMOUNT	120 million
Orderlines	ORDERLINEID, ORDERID, PROD_ID, QUANTITY, ORDERDATE	600 million
Products	PROD_ID, CATEGORY, TITLE, ACTOR, PRICE, QUAN_IN_STOCK, SPECIAL, COMMON_PROD_ID1, COMMON_RATING1, COMMON_PROD_ID2, COMMON_RATING2, COMMON_PROD_ID3, COMMON_RATING3, SALES	1 million
Reorder	PROD_ID, DATE_LOW, QUAN_LOW, DATE_REORDERED, QUAN_REORDERED, DATE_EXPECTED	Varied
Categories	CATEGORY, CATEGORYNAME	16

Figure 1. Database schema for online DVD store

System	Simultaneous database connections	Orders per minute (larger is better)	Average response time (seconds)	CPU utilization	Dell PowerEdge 6650 performance advantage over Sun Fire V440	Total hardware and software price	Price/performance ($/opm—lower is better)	Dell price/performance advantage
Four-processor Sun Fire V440 (processors: 1.28 GHz/1 MB of L2 cache)	7	8,156	0.050	89%	n/a	$193,739	$23.75	n/a
Four-processor Dell PowerEdge 6650 (processors: 3.0 GHz/4 MB of L3 cache)	9	18,970	0.028	89%	2.33×	$199,016	$10.49	126%
Two-processor Dell PowerEdge 6650 (processors: 3.0 GHz/4 MB of L3 cache)	5	10,865	0.027	89%	1.33×	$109,016	$10.03	137%

Figure 2. Sun-to-Dell migration results

browsing the catalog several ways, and finally purchasing the selected items. Each sequence completed by a customer counted as a single order. The driver measured order rates and the average response time to complete each order. Several tunable parameters were used to control the application; visit *Dell Power Solutions* online at http://www.dell.com/magazines_extras to see these parameters.

Examining the test results

The DVD Store 2 database running on the Sun Fire V440 was tested using the OLTP driver application described in the preceding section. The database was then moved to the Dell PowerEdge 6650 server, where it was tested using two different CPU configurations. The tests measured how many orders per minute each database server could handle while keeping CPU utilization under 90 percent—a typical system target to allow processing capacity for order spikes. CPU utilization was measured using the vmstat program under both Solaris and Red Hat Linux.

In each configuration tested, the Dell PowerEdge server offered substantially better price/performance at a fraction of the cost of the proprietary Sun server.

As seen in Figure 2, the fastest Dell PowerEdge 6650 server, with four 3.0 GHz processors, handled 2.33 times more orders than the Sun Fire V440 system. Although the PowerEdge server was priced $5,277 more than the Sun server (total server hardware and software price, including OS), it offered a price/performance of $10.49 per opm, whereas the Sun Fire V440 cost $23.75 per opm.[4] In other words, the Sun Fire V440 cost 126 percent more than the Power Edge 6650 per unit of work that it could handle. With two 3.0 GHz processors, the PowerEdge 6650 was 33 percent faster and 44 percent less expensive than the four-processor Sun Fire V440 configuration.

Reaping the benefits of industry-standard, Linux-based systems

Enterprises can easily move from UNIX/RISC-based servers to industry-standard, Intel processor–based servers running Linux and save significantly, as demonstrated by the results of this study. In each configuration tested, the Dell PowerEdge server offered substantially better price/performance at a fraction of the cost of the proprietary Sun server. The results described in this article prove that systems using industry-standard components, such as Intel processors and the Linux OS, can outperform proprietary systems while helping organizations meet the need for cost-efficiency in the data center.

Todd Muirhead is an engineering consultant on the Dell Technology Showcase team. He specializes in SANs and database systems. Todd has a B.A. in Computer Science from the University of North Texas and is Microsoft Certified Systems Engineer + Internet (MCSE+I) certified.

Dave Jaffe, Ph.D., is a senior consultant on the Dell Technology Showcase team who specializes in cross-platform solutions. Previously, he worked in the Dell Server Performance Lab, where he led the team responsible for Transaction Processing Performance Council (TPC) benchmarks. Before working at Dell, Dave spent 14 years at IBM in semiconductor processing, modeling, and testing, and in server and workstation performance. He has a Ph.D. in Chemistry from the University of California, San Diego, and a B.S. in Chemistry from Yale University.

FOR MORE INFORMATION

Dell and Red Hat:
http://www.dell.com/linux

[4] U.S. prices for the Dell PowerEdge 6650 servers and the Sun Fire V440 server are cited from the Dell and Sun online stores, respectively, as of September 27, 2004.