Impact of Dell FlexMem Bridge on Oracle Database Performance

A Dell Technical White Paper

Dell Database Solutions Engineering

Balamurugan B

January 2011

January 2011

Contents

Tables

Figures

Introduction

As data center size and complexity grow, the demand for resources such as faster CPUs, larger memory capacity, and storage capacity continues to increase. This requires careful resource planning techniques to ensure that requirements are met.

At the same time, the nature of database workloads also plays a major role in determining memory and processor requirements. For example, Online Transaction Processing (OLTP) workloads exhibit very different resource usage patterns than Online Analytical Processing (OLAP) workloads. Many of the major database applications consume large amounts of memory, and adding additional memory is only possible by activating additional processor sockets on the same server or by adding an additional server. This in turn leads to higher software licensing costs or the need for additional data center space.

Dell developed FlexMem Bridge technology to mitigate the memory scalability challenges of small 2U servers without adding extra data center space requirements. This whitepaper explores the flexibility, ease of use, and performance impacts that FlexMem Bridge technology has on Oracle database performance.

FlexMem Bridge technology also introduces a number of business benefits. Apart from scalability and performance benefits, FlexMem Bridge technology also enables:

- **Cost savings**: Adds flexiblity to inepensively scale up infrastructure when the need arises.
- **Efficiency**: Besides memory scalability, the FlexMem Bridge 2P configuration consumes less power than the standard 4P configuration.
- **Ease of implementation**: Straightforward memory scalability by means of FlexMem Bridges inserted in the CPU slots.
- **Performance**: More system memory helps to achieve good database response time as well as increases in user load capacity.

Overview of FlexMem Bridge Technology

FlexMem Bridge is an innovative technology introduced in selected 11th generation (11G) server platforms such as Dell™ PowerEdge™ R810 and M910 servers. These servers use Intel® Nehalem–based EX processors (7500 and 6500 series) that support DDR3 memory technology. FlexMem Bridge is designed to provide customers extended memory scalability while populating only a subset of processor sockets.

Traditionally, on four-socket servers, the memory subsystems of CPU3 and CPU4 are accessible only when those sockets are populated. In other words, if only CPU1 and CPU2 sockets are populated, the memory subsystems of CPU3 and CPU4 are inaccessible. Adding memory to the existing server requires the addition of two more processors, which adds to the database licensing cost.

FlexMem Bridge is designed to resolve these challenges using the following architecture:

- **Extended Scalable Memory Interconnects (SMI):** SMIs from CPU1 and CPU2 are extended to the memory subsystems of CPU3 and CPU4 to ensure that the memory subsystems of CPU3 and CPU4 are accessible even when only CPU1 and CPU2 are populated.
- **Quick Path Interconnect (QPI) Inter-Processor Communication:** FlexMem Bridge reroutes and terminates the second QPI link to provide better inter-processor bandwidth. This provides increased performance for a 2P configuration as it allows two full-bandwidth QPI links between CPU1 and CPU2 instead of a single link.

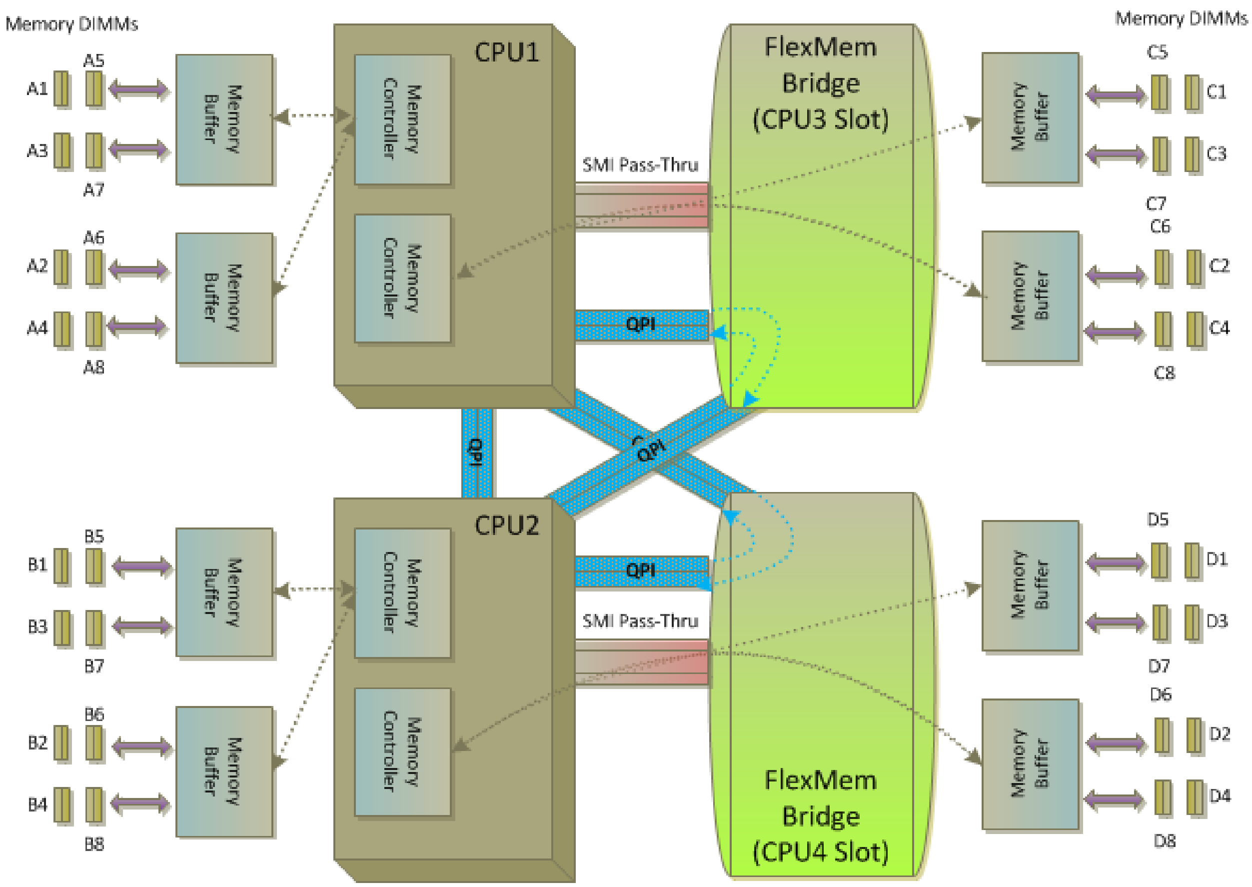
In a four-CPU configuration, the R810 and M910 servers use only one memory controller per CPU. This single controller connects to two memory buffers via SMI links. Each memory buffer in turn connects to four DDR3 DIMMs with a total of 32 DIMMs accessible. In a two-CPU configuration, this would normally mean that only four memory buffers are connected; therefore, a total of only 16 DIMMs are accessible. To overcome this limitation in the case of single- and two-CPU configurations, the R810 and M910 use a pass-through called the FlexMem Bridge. The FlexMem Bridge is installed in the CPU2 Slot in a single-socket configuration, and CPU3 and CPU4 in a two-socket configuration. Consequently, with the FlexMem Bridge installed, CPU1 can access an additional 8 DIMMs of CPU2 in a single-socket configuration, while in a two-socket configuration an additional 16 DIMMs of CPU3 and CPU4 can be accessed by CPU1 and CPU2 .

The FlexMem Bridge provides two pass-through links for SMI, and one pass-through link for QPI. The pass-through SMI links connect the two installed CPUs to additional Intel 7500 Scalable Memory Buffers; therefore, the two CPUs will have the following memory attached:

- CPU1 will have access to DIMMs [A1:A8], plus DIMMs [C1:C8] (those normally associated with CPU3)
- CPU2 will have access to DIMMs [B1:B8], plus DIMMs [D1:D8] (those normally associated with CPU4)

Figure 1 depicts the FlexMem Bridge layout with two processors populated.

Figure 1. **FlexMem Bridge Architecture**



Benefits of Implementing Dell FlexMem Bridge for Databases

Customers can leverage FlexMem Bridge technology to improve the following Oracle Database performance factors:

Response Time

Oracle Database performance is directly impacted by factors such as processor, memory, disk, and network configuration. Scaling up the memory available to the database allows more data buffering for fast retrieval without the overhead of disk access, which directly influences database response time.

User Scalability

Memory is one of the critical components in terms of user scalability. Using FlexMem Bridge, memory capacity can be internally scaled without adding new processors or servers, which in turn helps the customer to support more user access.

Dell FlexMem Bridge technology adapts the platform for workloads where there is more demand for memory per processor. This technology allows easy internal scalability of both memory and processors, and addresses customers' concerns such as memory-bottleneck issues and rack-density problems by allowing them to maintain the same software license cost scheme and use 32 DIMMs with two processors in a slim 2U form factor. If the customer runs out of CPU resources in the future, the FlexMem Bridges can be replaced with CPUs (for models with the Intel Xeon 7500 CPU only), and instantly double processor performance – all without taking up additional rack space.

In summary, the FlexMem Bridge offers the following benefits for Oracle database configurations:

- More memory scalability with the same licensing scheme
- No demand for extra data center space if memory capacity is scaled
- Additional processors needed only when more processing power is required

Test Methodology

The experiments were designed specifically to analyze the improvement in database performance by scaling up available database memory. Separate databases were configured for analyzing the OLTP and OLAP workload performance benefits. Dell solution engineers used Quest Software® Benchmark Factory® TPC-C and TPC-H workload to explore the benefits of FlexMem Bridge technology with Oracle OLTP and OLAP workloads on PowerEdge 11G R810 servers.

Test Configuration

Table 1 describes the complete software and hardware configuration that was used throughout testing, without and with FlexMem Technology.

Table 1. Test Configuration Details

Test Component	Details
Server	PowerEdge R810; BIOS : 1.2.0
CPU	Intel Xeon® CPU 7560 @2.27GHz
Number of Cores	8
Memory	16 * 2GB DDR2 1067GHz DIMMs (without FlexMem) 32 * 2GB DDR2 1067GHz DIMMs (with FlexMem)
Network Adapters	2 * Broadcom BCM57711 NetXtreme II 10GB NICs
Network Switch	PowerConnect™ 8024F
Storage	EqualLogic™ PS6010 10GB array (fully populated) Firmware : 4.3.6
Operating System	RHEL 5.5
Database Software	Oracle 11gR2 (11.1.0.7)
Dell EqualLogic HIT Software	HIT 3.4.1
Benchmarking Tool	Quest Benchmark factory 6.1.1
Test Workload	TPC-C(OLTP) and TPC-H(OLAP)

The database performance parameters were captured and analyzed using the following Linux tools:

Dstat – to monitor CPU, memory and network
iostat - to collect and show operating system storage input and output statistics
netstat - to gather network statistics

The test scenarios followed for the OLTP and OLAP workloads were different because of the varied nature of the workloads on the database. The specific test scenarios and results for each of the workloads are explained in subsequent sections of this paper.

OLTP Workload

Dell engineers conducted several experiments to assess the performance improvement of using a larger memory configuration for an OLTP system. Steps were taken to ensure that there was no bottleneck on the server CPU, disk, network, etc. A total database size of 300 GB was used for the exercise. Using Quest Benchmark factory, TPC-C transactions were simulated with user loads ranging from 200 to 5000 concurrent users, in increments of 100 users. During the workload, all the simulated users would be randomly running transactions against the database. A Dell EqualLogic 6010XVS was configured as the backend storage subsystem.

The test methodology used was as follows:

- All of the 16 DIMM slots for CPU1 and CPU2 were populated with 2GB DIMMs (a total of 32GB RAM).
- Test workload (TPC-C) was executed to introduce a memory bottleneck on the test database using increased user load.
- Once the memory bottleneck was achieved on the system, FlexMem Bridge was placed on the CPU3 and CPU4 slots. All of the additional 16 slots for CPU3 and CPU4 were then populated with 2GB DIMMs to scale up the memory (64GB RAM). The test workload was then rerun on the new configuration to analyze the performance effect. In all the cases, the memory target was set to use the maximum available server memory.

Figure 2 shows the TPS and the number of concurrent users with and without FlexMem Bridge enabled.

Figure 2. **TPS Graph With and Without FlexMem Bridge**

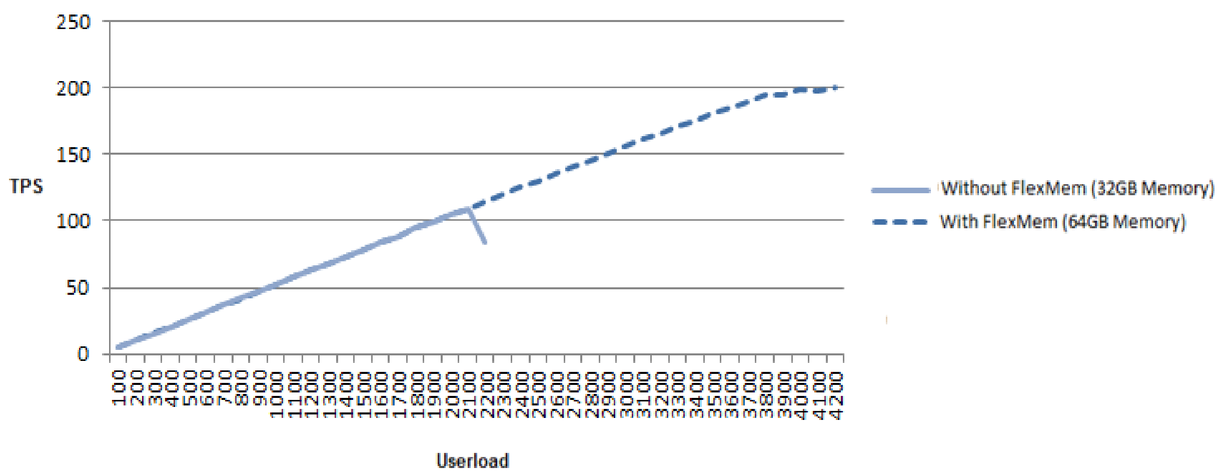


Figure 2 shows that there was an 100% increase in the supported number of concurrent users with FlexMem Bridge enabled. In other words, the database could continue to support more users by adding additional memory. There was also an 90% increase in TPS because of the additional memory.

Figure 3. **CPU Utilization Graph Without FlexMem Bridge Enabled**

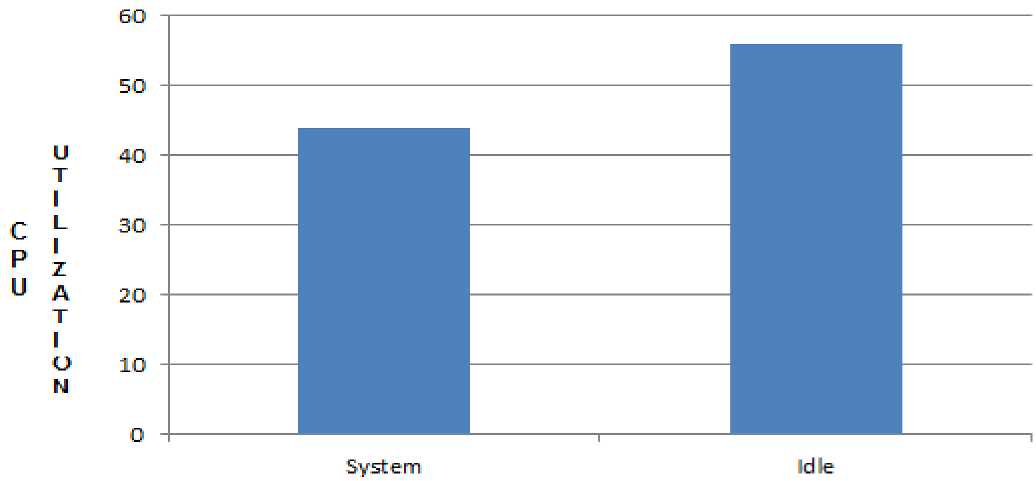


Figure 3 is a snapshot of CPU utilization at the saturated user load of test without FlexMem (16 DIMMS of 2GB). By analyzing the CPU usage it is clear that the system had over 55% free CPU cycles. This means that the system could not scale-up because of the memory bottleneck.

Figure 4. **Response Time Comparison With and Without FlexMem Bridge**

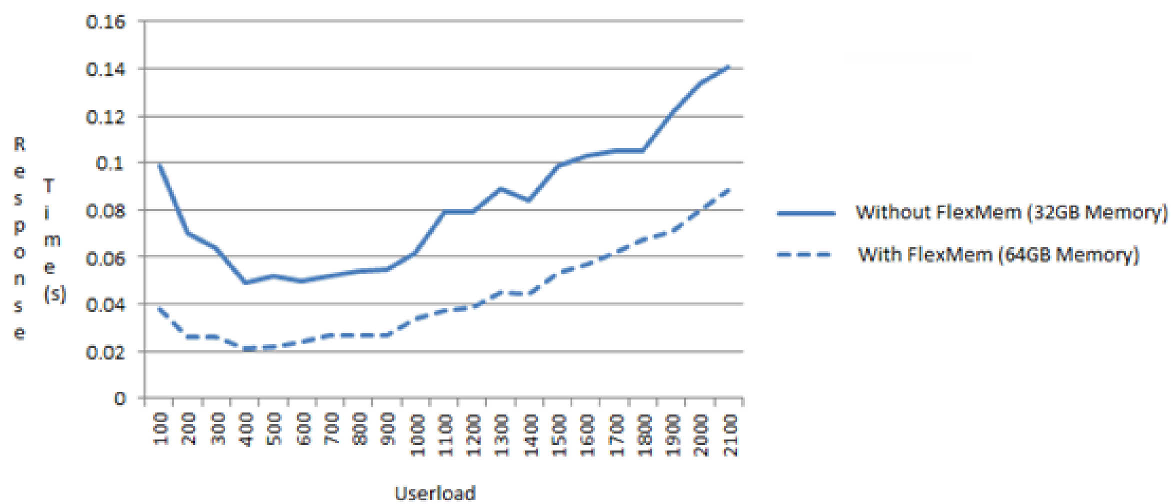


Figure 4 illustrates that the query response time for each of the user load improved on adding more memory to the system. At the same time, the database was able to support more users within a given response time, on adding more memory. There was an almost 60% improvement in response time with FlexMem bridge enabled.

Overall, the database with 64GB memory showed an improved capability to handle around 2100 additional users with the given response time, compared to the database with 32GB memory. When physical memory is insufficient for the number of concurrent user connections, it may lead to an overall slowdown of the system. This may be due to excessive paging of the working data set of connected users, leading to an increased number of disk writes and reads. When the memory is too low, it may result in low space in the system page files, which in turn causes the system to fail memory allocations as it is unable to page out currently allocated memory. Insufficient memory may eventually result in the whole system responding very slowly, or it may even bring the system to a halt. This makes it necessary to add more memory for the database. In case of insufficient database memory for user connections, FlexMem Bridge provides a flexible solution to scale up the maximum users supported for a given response time, or achieve a better response time for a given user load.

OLAP Workload

FlexMem bridge enables more available memory per processor to be used for a database, allowing more data to be cached in the database buffer. Since OLAP applications contain large amounts of data processing, understanding the effects of large memory demands on database performance is essential. As part of this experiment, OLAP exercises were designed to reveal the performance benefit of having more memory for buffering data for an OLAP workload.

A test database of approximately 100GB was used for this exercise. Although real world OLAP databases are usually much larger, the size of the database is not relevant for this illustration — what really matters is the ratio of available memory to the amount of data queried from the database.

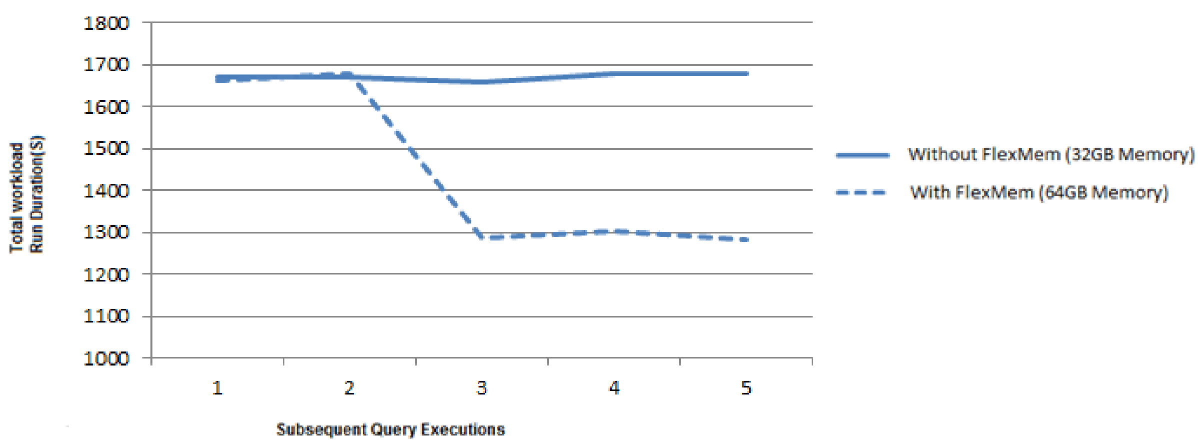
The TPC-H workload transaction mix used to stress memory is shown in Table 2. This transaction mix is designed to fetch around 50GB of data from the database. A TPC-H query workload was simulated using Quest Benchmark factory tool for a single user.

Table 2. OLAP Test Queries

Query
Promotion Effect Query (Q14)
Forecasting Revenue Change Query (Q6)
Pricing Summary Report Query (Q1)

To analyze the benefits of database buffering, OLAP workloads with the queries shown in Table 2 were executed for five iterations. For each iteration, the above mentioned queries were run once sequentially. The query response time for each run is plotted in Figure 5.

Figure 5. **OLAP Query Response Time**



When a query is executed, the Oracle database engine checks whether the data is available in the data buffer. If the data is found in the cache/buffer, it is retrieved directly from the cache. Otherwise the engine must perform a disk read, which is comparatively expensive and has a performance penalty. However, when the query is executed repeatedly, the database engine optimizes the query plan to process most of the data from the buffer and minimize the disk access. The end result is an overall optimization of the query response time.

It is evident that the first execution of the workload, with and without FlexMem Bridge, does not show any difference in the overall duration to run the query. This is because almost all of the data rows must be fetched from the disk. In this case, the benefit of buffered data is not relevant. Subsequent query executions show significant improvement in the response time because the Oracle Server optimizes the query plan at each level of execution.

The query response time becomes a stable state starting at the third execution because the query execution plan is optimized to run with minimal access to the disk with the available memory configuration.

Ideally, if repeated queries are expected in an OLAP database, query response time may be maximized by making the memory larger than the expected query data fetch. However, the real need for more memory arises when available memory is much less than the required amount of data to be fetched by the query. This causes swapping of the database memory pages between the disk and the cache. This excessive memory paging activity degrades the database performance to a great extent.

For organizations, an improvement in the OLAP database response time basically means faster retrieval of business reports. This is beneficial for better data mining and analysis, with minimal time spent waiting for reports to be generated. FlexMem helps the OLAP environment to improve the response time by means of database buffer scale-up.

Summary

NOTE: The results we have provided are intended only for comparison of the two database environments consisting of specific configurations in a lab environment. These results do not portray the maximum capabilities of any system, database software, or storage.

In a server subsystem memory is a critical component in terms of database performance. In these tests, even though there was sufficient processing power available, the database could not scale-up because of memory limitations. Traditionally, scaling up memory can be only achieved by adding additional processors. With the help of Dell FlexMem Bridge technology on the new R810 and M910 11G servers, memory scalability can be achieved without adding additional CPUs to the system, in the case of 1 CPU and 2 CPU configurations.

In addition, these tests demonstrated that for both OLTP and OLAP workloads, when more memory is available to the database through FlexMem technology, more data is buffered. This in turn helps to improve database performance by reducing the response time.

The experiments conducted showed up to 25% - 30% improvement in response time for a given OLAP workload and an average of 60% improvement in OLTP response time. The sample OLTP workload showed an improvement of over 100% in terms of user load for a given response time.

In addition, FlexMem Bridge technology helps in terms of scaling up the memory in case processing power is not the limitation. This helps the database admin to maintain a reduced number of database licenses.

References

Dell PowerEdge 11th Generation Servers: R810, R910, and M910 Memory Guidance

http://www.dell.com/downloads/global/products/pedge/en/poweredge-server-11gen-whitepaper-en.pdf

FlexMem Bridge Technology-The Keyword is Flex-ibility

http://en.community.dell.com/dell-blogs/enterprise/b/inside-enterprise-it/archive/2010/04/19/flexmem-bridge-technology-the-keyword-is-flex-ibility.aspx

Impact of Dell FlexMem Bridge on Microsoft® SQL Server® Database Performance

http://content.dell.com/us/en/enterprise/d/business~solutions~whitepapers~en/Documents~flexmem-bridge-sql-database.pdf.aspx

Dell oracle Solutions

http://www.dell.com/oracle