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

A Dell Technical White Paper

Dell | Database Solutions Engineering

Jisha J

Leena Basanthi

October 2010

October 2010

Contents

Introduction 2
Overview of FlexMem Bridge Technology 2
Benefits of Implementing Dell FlexMem Bridge for Databases 3
Test Methodology 4
OLTP Workload 5
OLAP Workload 7
Introducing FlexMem Bridge for a 2P Configuration 9
Summary 11
References 12

Tables

Table 1. Flex Memory Bridge Test Configuration Details 4
Table 2. Observed Response Time and Corresponding User Load 6
Table 3. OLAP Test Queries 7
Table 4. Performance Counters for Monitoring Memory Issues 10

Figures

Figure 1. FlexMem Bridge Architecture 3
Figure 2. OLTP Response Time 5
Figure 3. Additional Users Handled on Scaling up Memory 6
Figure 4. OLAP Query Response Time 8

Introduction

As data center size and complexity grows, the demand for resources like faster CPUs, larger memory capacity, and storage capacity continues to increase, requiring careful resource planning techniques to ensure that requirements are met.

At the same time, the nature of database workload 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 often means either moving to higher software licensing schemes or upgrading databases to higher form factor servers, which, in turn, consumes 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, and performance impact that FlexMem Bridge technology has on Microsoft® SQL Server® database performance.

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

- **Cost savings**: Adds flexiblity to inexpensively 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**: For memory scalability, FlexMem slugs must be inserted in the CPU3 and CPU4 slots, after which additional memory may be populated in DIMM slots C1-C8 and D1-D8. For processing scalability, the slugs may simply be replaced with extra processors.

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-processor 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 licensing cost.

FlexMem Bridge is designed to resolve these challenges as follows:

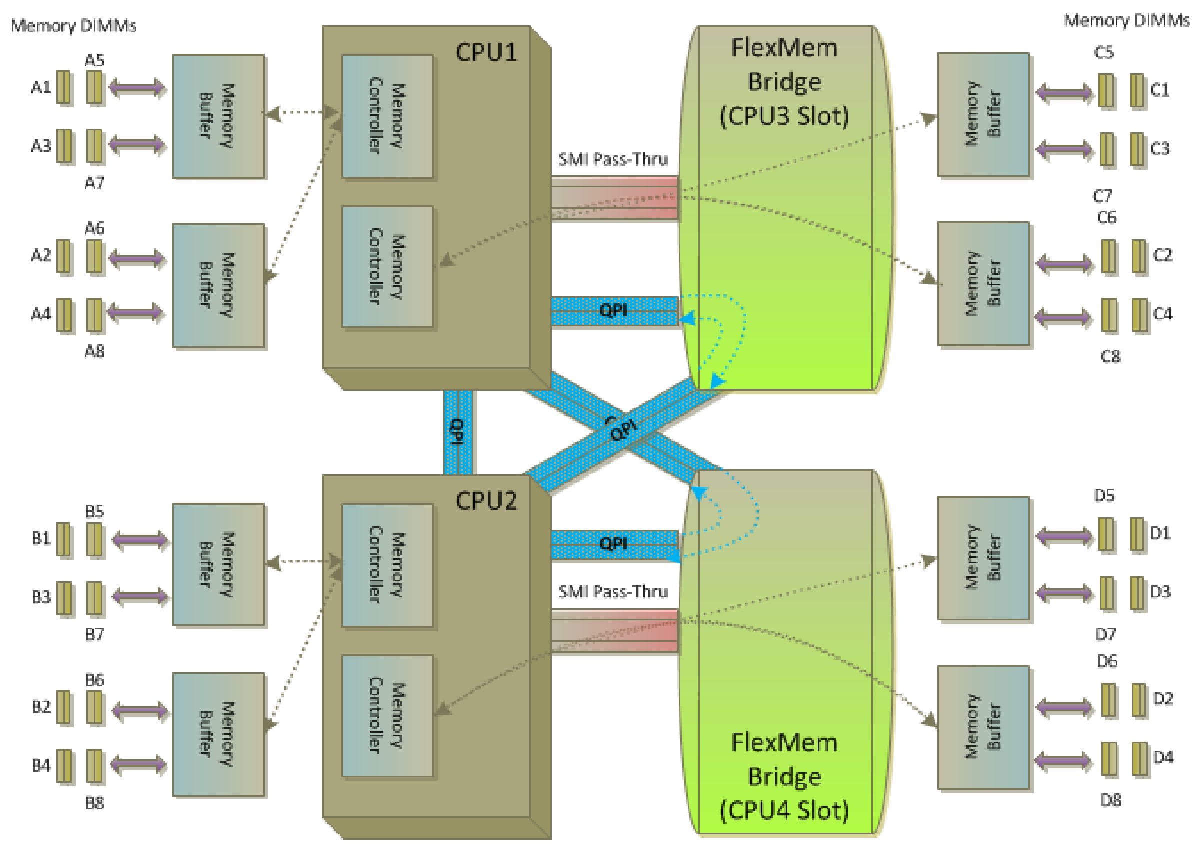
- **Extended Scalable Memory Interconnects (SMI):** SMIs from CPU1 and CPU2 are extended to the memory subsystems of CPU3 and CPU4 to ensure that the memory subsystems of CPU3 and CPU4 are accessible even when only CPU1 and CPU2 populated.
- **Better Quick Path Interconnect (QPI) Bandwidth:**   FlexMem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

PowerEdge R810 servers support a maximum of four processors. Figure 1 depicts the FlexMem Bridge layout with two processors populated. The pass-through SMI links introduced by the FlexMem Bridge enable the following features:

- CPU1 can access DIMMs C1-C8 (associated with CPU3), in addition to DIMMs A1-A8
- CPU2 can access DIMMs D1-D8 (associated with CPU3), in addition to DIMMs B1-B8

In other words, memory is internally scaled without an associated increase in CPU count in 1P and 2P configurations.

Figure 1. FlexMem Bridge Architecture



Benefits of Implementing Dell FlexMem Bridge for Databases

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

FlexMem Bridge technology can help solve memory bottleneck issues. Using FlexMem Bridge, memory capacity can be internally scaled without adding new processors or servers, allowing customers to maintain a steady software licensing cost scheme. Additionally, with more memory available per processor, customers can better manage database workloads, including the ability to scale memory usage without the need to populate all sockets within a server to access all available memory banks.

As a new four-socket server product, the PowerEdge R810 features high density without compromising memory space and I/O scalability. FlexMem Bridge technology makes 2U servers scalable to four sockets with large memory, and memory speed does not vary with the number of DIMMs populated in the server. PowerEdge R810 and M910 full-height blades can support a memory speed up to 1066MHz.

The Dell Flex Mem Bridge Technology makes the platform fit for workloads where there is more demand for memory per processor. This technology allows easy internal scalability of both memory and processor. This technology solves customers' concerns like memory-bottleneck issues and rack-density problems by allowing them to stay in the same software license cost scheme and get 32 DIMMs with two processors in a slim 2U form factor. In future, if the customer runs out of CPU resources, then 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 space.

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

- More memory scalability within the same licensing scheme
- No claim for extra data center space
- Additional processors needed only when more processing power is required

Test Methodology

Dell conducted several experiments to explore the performance benefits of FlexMem Bridge technology for database-specific workloads. Part of this exercise included evaluating FlexMem Bridge technology with both OLTP and OLAP workloads. The experiments were designed specifically to analyze the improvement in database performance on scaling up available database memory. Table 1 shows the hardware and software components used for the test configuration.

Table 1. Flex Memory Bridge 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 DDR3 1067GHz DIMMs (without FlexMem) 32 * 2GB DDR3 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	Microsoft Windows® 2008 R2
Database Software	Microsoft SQL Server 2008 R2
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)

Separate databases were configured for analyzing the OLTP and OLAP performance benefits. Quest Benchmark factory was used to simulate the OLTP and OLAP workloads. The database performance parameters were captured and analyzed using Windows performance monitor (perfmon).

All of the 16 DIMM slots for CPU1 and CPU2 were populated without introducing the FlexMem Bridge (32GB RAM). Test workloads were initially executed to introduce a memory bottleneck individually on the test databases. 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 similar DIMMs to scale up the memory (64GB RAM). The test workloads were rerun on the new configuration to analyze the performance effect. In all the cases, the SQL instance memory target was set to use the maximum available server memory.

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

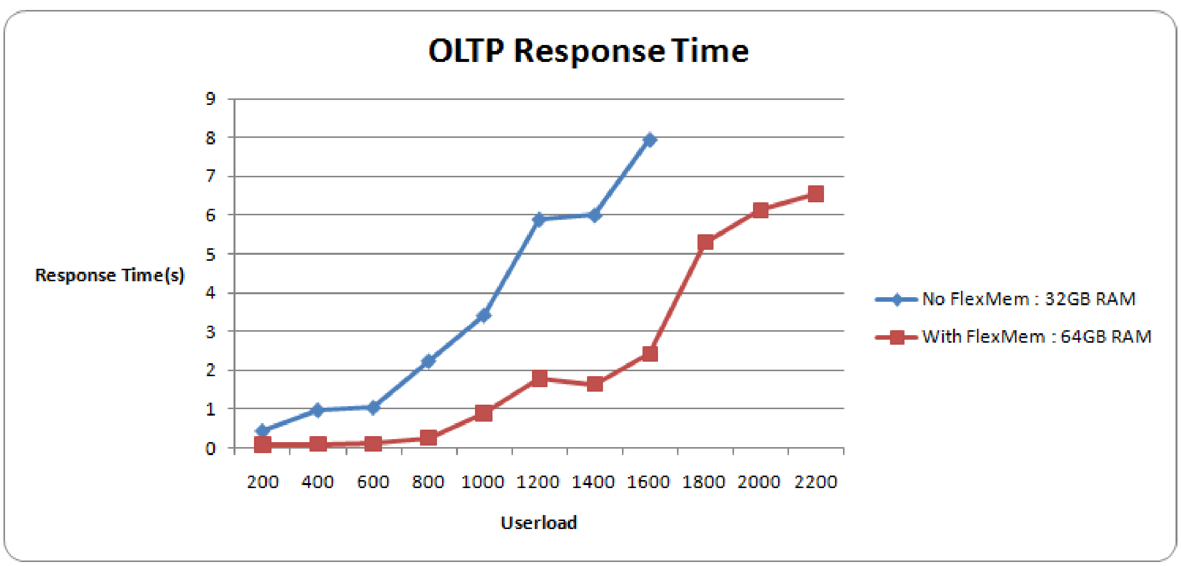
OLTP Workload

For an OLTP environment, one of the factors that decides the maximum number of concurrent users within the given response time is database memory: More database memory means a larger number of concurrent users can expect a given database response time.

Dell conducted several experiments to assess the performance improvement of using a larger memory configuration for an OLTP system. Steps were taken to ensure that none of the server components (CPU, memory, disk, network, etc.) were bottlenecked in the test configuration.

A test database size of 300GB was configured for analyzing the performance benefits of using FlexMem Bridge for an OLTP workload. A TPC-C workload was simulated, with user loads ranging from 200 to 2400 using Quest Benchmark factory tool. The tests were repeated on the configuration with and without FlexMem Bridge. Figure 2 shows the OLTP database response time with and without FlexMem Bridge enabled.

Figure 2. OLTP Response Time



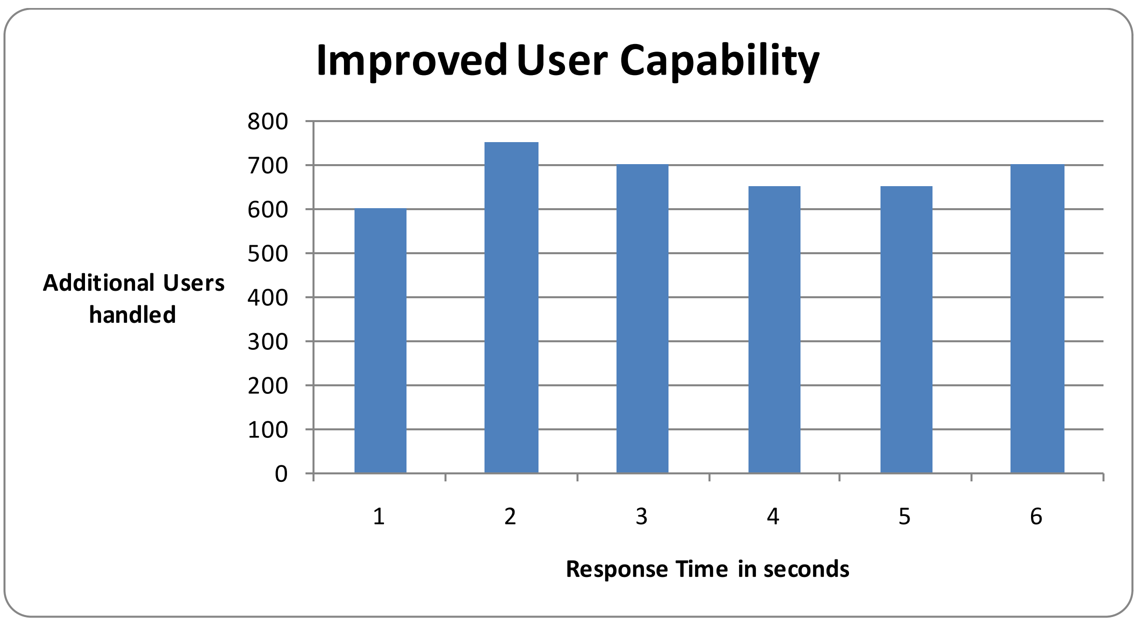
In Figure 2, it is observed that the query response time for each of the user load improved on adding more memory to the system. In other words, the database was able to support more users with a given response time, on adding more memory. As per the experimental workload, the number of users observed for different database response times may be summarized as in Table 2.

Table 2. Observed Response Time and Corresponding User Load

Response Time(s)	Users handled with the given response time	
	No FlexMem : 32GB RAM	With FlexMem : 64GB RAM
1	400	1000
2	750	1500
3	900	1600
4	1050	1700
5	1150	1800
6	1300	2000

Figure 3 depicts the number of additional users supported on scaling up the SQL server memory to 64GB from 32GB for the given workload.

Figure 3. Additional Users Handled on Scaling up Memory



Overall, the database with 64GB memory showed an improved capability to handle around 700 additional users with the given response time, compared to the database with 32GB memory.

When physical memory is insufficient for the number of concurrent SQL connections, it may result in an overall slowdown of the system. This may result in 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. This condition may 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 at this point. In case of insufficient database memory for user connections, FlexMem Bridge provides a flexible solution that adds more memory to the environment without having to change the database licensing scheme.

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, understanding the effects of large memory on the database performance is necessitated. 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, though it should be noted that real world OLAP databases are usually much larger. However, the size of the database doesn’t matter for this illustration — what really matters is the ratio of available memory to the amount of data queried from the database.

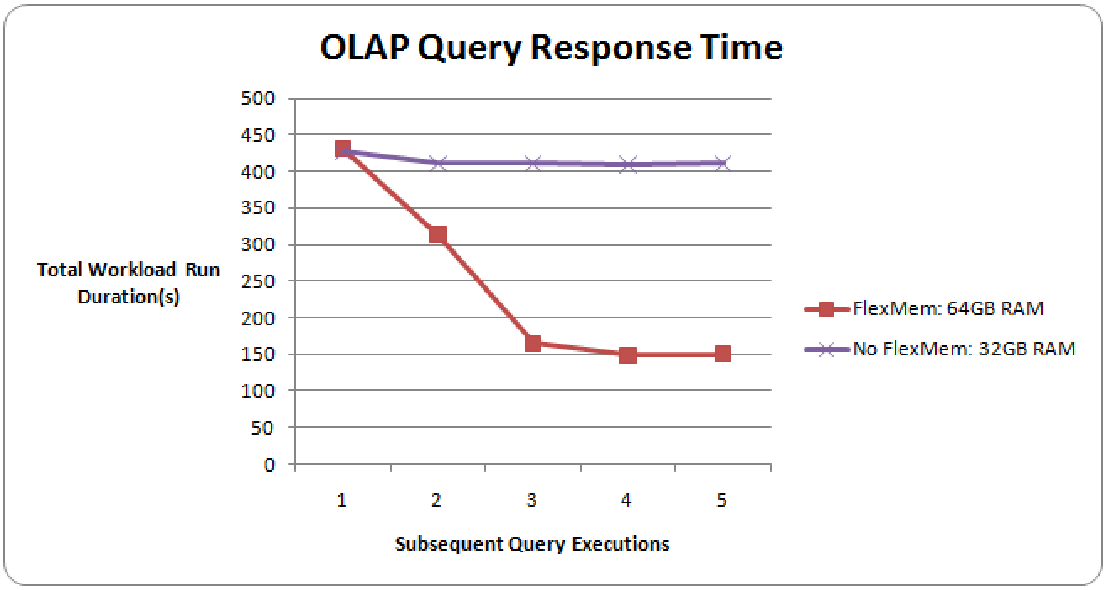
The TPC-H workload transaction mix used to stress memory is as given in Table 3. This transaction mix is designed to fetch around 50GB of data from the database. A TPC-H query workload was simulated using Quest Benchmark factory tool for a single user.

Table 3. OLAP Test Queries

Query	No. Of Executions
Promotion Effect Query (Q14)	1
Forecasting Revenue Change Query (Q6)	1
Pricing Summary Report Query (Q1)	1

To analyze the benefits of database buffering, the same set of queries was run five consecutive times. The query response time for each of the runs is as plotted in Figure 4.

Figure 4. OLAP Query Response Time



When a query is executed, SQL Server 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 SQL engine must perform a disk read, which is comparatively expensive. This would have a performance penalty. 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 overall maximization of the query response time.

In Figure 4, you may observe 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 the buffered data is not relevant. Subsequent query executions show significant improvement in the subsequent executions, because the SQL Server engine 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 the minimal access to the disk with the available memory configuration.

Ideally, if repeated queries are expected in an OLAP database, the query response time may be maximized by making the memory larger than the expected query data fetch. But 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. The excessive memory paging activities degrade 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 in better data mining and analysis, with minimal time spent waiting for reports to be generated.

Introducing FlexMem Bridge for a 2P Configuration

Database performance is directly impacted by the several components, some of which are listed below. If the database is experiencing performance issues, all the potential factors should be analyzed in detail to root cause the issue.

- Database design
- Indexing
- Query design
- Processor
- Memory
- Disk
- Network

More information on the analzing the potential bottlenecks on a database may be found in References section. In case of a memory bottleneck, scaling up the memory available for a database allows more buffering of the data for fast retreival, without the overhead of repeated disk acess. It may be noted that adding more memory for the database will not always be the right solution for all database performance problems.

FlexMem Bridge enables the ease and flexibility of internally adding more memory to the database server for a 2P database server. FlexMem Bridge may be helpful in solving the memory bottleneck issues in a database server, without changing the licensing scheme for a 2P system.

There are different performance monitor (perfmon) variables that may be considered in deciding whether the database is experiencing a memory performance issue. These counters must be continuously monitored to access SQL Server performance. Table 4 details some of the various performance monitor counters for the SQL Server buffer manager object that must be considered for database memory bottleneck analysis. Note that there are various operating system (OS) counters available to be monitored in addition to the SQL Server counters.

Table 4. Performance Counters for Monitoring Memory Issues

Performance Variables	Description	Desired Range
Free list stalls/sec	Number of requests per second that had to wait for a free page	< 2
Free pages	Number of memory pages available to receive database pages from disk	> 640 Not to be less than 4MB consistently for 2 seconds
Lazy writes/sec	Number of pages written per second by the lazy writer process	< 20
Page life expectancy	Number of seconds a page will stay in the buffer pool without references	> 300
Page Lookups/sec	Number of attempts made by the database to locate a page in the database buffer pool	(page lookups/sec)/(batch requests/sec) < 100
Batch requests/sec	Number of batch requests per second received by SQL Server	
Pages read/sec and writes/sec	Indicates the physical I/Os performed by the database	< 90
Buffer cache hit ratio	Frequency with which read requests are satisfied from the cache, compared to those processed from the disk	> 90 for OLAP and > 95% for OLTP

If the value of the performance counters is observed to be consistently out of the recommended range, there is a greater possibility of a SQL Server memory constraint. A detailed analysis of the counters should be performed to confirm the component being bottlenecked. More documents to aid on the SQL Server bottleneck analysis may be found in the References section.

Summary

Database server memory is a critical factor in database performance. Without FlexMem Bridge, scaling up database memory means increasing the number of processors, and, in turn, increasing the overall database licensing cost for an organization — even when there is no need to add additional processing power to the system. Alternatively, FlexMem Bridge can be used to scale memory internally within the server, without increasing the data center footprint.

A larger database memory allows more database buffering, which may enhance database performance. The experiments conducted showed up to 170% improvement in response time for the given OLAP workload on buffered queries. The sample OLTP workload showed more than 50% ability in handling more users, with the same response time. The average improvement in OLTP response time was also observed to be around 69%.

The benefits of using FlexMem Bridge for a database server with a PowerEdge R810 1P/2P configuration include:

- Easy Internal scaling of the memory within a slim 2U form factor
- No need to scale up the database licensing scheme
- Does not claim extra data center space
- Easy processor scalability

References

Dell SQL Server Solutions
www.dell.com\sql

Dell Services
www.dell.com\services

Dell Support
www.dell.com\support

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

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

OLTP and OLAP
http://datawarehouse4u.info/OLTP-vs-OLAP.html

SQL Server Performance Assessment and Optimization Techniques
http://www.demandtech.com/Resources/User%20Group%20meetings/CMG%202004/SQL%20Counters%20and%20Optimization%20CMG.pdf

Improving SQL Server Performance
http://msdn.microsoft.com/en-us/library/ff647793.aspx