Building a Highly Scalable Database Platform

Using Microsoft SQL Server 2000 Enterprise Edition (64-bit)

Microsoft® SQL Server™ 2000 Enterprise Edition (64-bit) is designed to offer improvements in memory availability and parallel-processing performance compared with 32-bit SQL Server software. This article describes the capabilities of Microsoft SQL Server 2000 Enterprise Edition in the 64-bit Intel® architecture environment—highlighting differences from the 32-bit environment and discussing some of the applications and usage scenarios that can benefit from a SQL Server platform optimized for the 64-bit environment. Additionally, guidance is provided on identifying potential applications and setting appropriate expectations for deployment and performance.

By Will Claxton

To make timely and informed business decisions in dynamic environments, organizations store and analyze massive amounts of business data. Always a critical part of the IT infrastructure, the database is at the center of three converging IT trends: growth in application size and complexity, development of high-end database management system (DBMS) capabilities, and consolidation of data center servers.

Because of their size, some applications may be reaching the limits of the 32-bit platform, specifically regarding number of processors and addressable memory. At the same time, high-end DBMS capabilities have become a business necessity as organizations gather and analyze data from numerous databases, and serve that data to growing numbers of business users. Further, many organizations are moving to consolidate servers to simplify critical data center operations. Consolidation can reduce management complexity and cost while reducing physical space requirements in the data center.

In response to these trends, organizations can take three general paths:

- Scale up without 64-bit Intel architecture (IA-64) migration: Upgrade the existing 32-bit server environment to include new and more powerful Intel® Xeon™ processors with expanded cache capability. Upgrading 32-bit servers also enables administrators to deploy storage with improved disk performance, which can help alleviate the disk I/O bottleneck in database applications.
SCALABLE ENTERPRISE

- **Scale up with IA-64 migration**: Replace 32-bit servers with 64-bit servers designed to provide applications with greater memory addressability and storage capability.
- **Scale out**: Restructure the database using a variety of methods to provide the required performance.1

This article explores the reasons why and circumstances under which organizations may consider 64-bit database technology using Microsoft® SQL Server™ 2000 Enterprise Edition software as well as how IA-64 architecture can help remove certain performance bottlenecks cost-effectively.

**IA-64 components in SQL Server 2000**

Microsoft SQL Server 2000 Enterprise Edition (64-bit) components include:

- **Database server**: Core database functionality
- **Server agent**: Alerts and management
- **Analysis server**: Online analytical processing (OLAP) and data mining

The preceding IA-64 components are code-compatible with the 32-bit version of SQL Server 2000, allowing administrators to integrate an IA-64 server with other SQL Server 2000 database servers.

**Capabilities of IA-64 architecture**

The enhanced scalability and performance capabilities of SQL Server 2000 (64-bit) is enabled by several features of the 64-bit Intel Itanium® architecture. Key benefits are as follows.

- **Memory addressability.** Generally, 32-bit systems can address only a 4 GB address space. See the section “Comparison of SQL Server 2000 (32-bit) with extension technology versus 64-bit IA-64 architecture” in this article for more information about options available with Address Windowing Extensions (AWE) and Physical Address Extension (PAE) on 32-bit platforms. The Microsoft Windows Server™ 2003 operating system running on Intel Itanium IA-64 architecture supports up to 1,024 TB of physical memory and 512 GB of addressable memory.

- **Parallel-processing support.** Intel Itanium chips include several features that are designed to enhance parallel-processing performance compared to 32-bit Intel chips.2 The Intel Itanium 2 chip offers a wider system bus, more registers,3 and Explicitly Parallel Instruction Computing (EPIC) technology, which is designed to enable a processor to execute up to six instructions simultaneously. Such performance improvements in parallel processing can help enhance many SQL Server 2000 parallel operations, including parallel query resolutions, index builds, backup and restore operations, data loads, and maintenance operations.

- **Enhanced bus architecture.** The bus architecture on current IA-64 chip sets can provide greater throughput than in 32-bit environments.4 More data can be passed to the cache and processors quickly—an enhancement somewhat analogous to the improvement that broadband offers over dial-up connections.

**Comparison of SQL Server 2000 (32-bit) with extension technology versus 64-bit IA-64 architecture**

Microsoft SQL Server 2000 Enterprise Edition (32-bit) uses the Microsoft Windows® 2000 Server AWE and PAE application programming interfaces (APIs) to support very large amounts of physical memory in applications and the operating system, respectively. For some applications, using AWE to enhance SQL Server 2000 (32-bit) may be a viable alternative to upgrading to the IA-64 platform.

Standard 32-bit systems can map 4 GB of memory at most, limiting the addressable memory space for Windows 2000 systems to 4 GB. With 2 GB reserved for the operating system, only 2 GB of memory remain for the application—in this case, SQL Server 2000. Administrators can increase the amount of addressable application memory to 3 GB by setting the /3GB switch in the Windows boot.ini file. However, slight performance degradations may occur when using AWE and PAE in this way, and not all aspects of SQL Server 2000 (and many other applications) can take advantage of these extensions.

In contrast to the 32-bit limitations, SQL Server 2000 (64-bit) makes extended memory available to all database processes and operations. When using the 64-bit version of SQL Server 2000 on Itanium 2-based hardware like the Dell™ PowerEdge™ 7250 server, a SQL Server instance can address up to 512 GB,5 which is the current maximum memory supported by Windows Server 2003. (The theoretical addressable limit is 18 exabytes.6) This memory is available to all components of SQL Server and to all operations within

1 For more details on the scale-out option, see The Definitive Guide to Scaling Out SQL Server by Don Jones at http://www.dell.com/sql/ebook.
5 One exabyte ~ 2^54 (1,051,209,600,864,024) bytes, or 1,024 petabytes.
the database engine. As a result, SQL Server 2000 (64-bit) is designed to enhance performance of a wide range of memory-intensive database applications.

**Usage scenarios for SQL Server 2000 (64-bit)**

Although SQL Server 2000 (64-bit) is designed to offer significant scalability and performance for many applications—including SAP®, PeopleSoft®, Siebel®, and other applications that require frequent disk caching—not every application will benefit from the 64-bit version of SQL Server 2000. This section is designed to help organizations determine whether it is more appropriate to use existing SQL Server instances or develop new applications based on the IA-64 architecture.

The improved memory and parallel-processing capabilities of SQL Server 2000 (64-bit) compared with SQL Server 2000 (32-bit) are compelling in several usage scenarios, including:

- **Enhancing performance for memory-constrained relational database applications**: Helping to alleviate memory constraints allows a larger percentage of the database—or possibly the entire database—to reside in memory.
- **Creating or accelerating large OLAP systems with rapid response-time requirements**: Fast databases provide decision makers with quick access to simplified views of complex data.
- **Consolidating multiple Windows-based databases and applications onto fewer, larger systems**: By hosting multiple databases on a single IA-64 platform, organizations can simplify management, improve storage utilization, and generally improve operational efficiency.
- **Scaling up current applications that are experiencing significant growth**: Migrating existing database servers that are outgrowing their current platform does not affect the other tiers of multi-tiered applications.
- **Replacing UNIX systems and applications**: The IA-64 platform offers a powerful alternative to UNIX® systems for high-end database servers.

The following section discusses factors to consider when evaluating specific applications for the IA-64 platform.

**Relational database performance factors**

Memory-intensive SQL Server relational database workloads are good candidates for SQL Server 2000 (64-bit). Many SQL Server resources are restricted to a 3 GB limit in the 32-bit environment, resulting in systems that are starved for memory. Such systems degrade performance because applications wait for resources and experience delays while processors compile stored procedures that have been evicted from cache. These systems may also experience excess disk activity when writing objects such as hash tables—which cannot fit into the available memory—to disk. Moving to SQL Server 2000 (64-bit) can help improve the performance of applications experiencing the following memory-related performance problems:

- **Recompilation of stored procedures that were evicted from memory**: The IA-64 environment provides a large plan cache for high-volume transaction applications using large numbers of stored procedures. This helps reduce the need to compile stored procedures that have been evicted from memory—reducing CPU utilization and reducing query latency.
- **Resource semaphore waits associated with queries awaiting memory grants**: Multiple queries utilizing large-scale hash joins—especially those executing against a data warehouse and spanning large data sets—can be adversely affected by resource semaphore waits.

Other relational database operations that can benefit from extended memory include:

- **Index creation**: In a 32-bit environment, index creation, including full-text indexing, is restricted to a 3 GB workspace.
- **Complex queries**: Operations that use sorting, large hash joins, or hash aggregates to construct complex queries can benefit from extended memory. Memory-intensive hash joins are very efficient, but when memory is under pressure, these queries may be removed from cache in favor of slower memory-conserving query plans.
- **Active stored procedures**: Benefits can be achieved for large numbers of active stored procedures through improved cache capacity. The IA-64 architecture can help substantially reduce overall CPU utilization and latency by helping to eliminate the need to evict procedures from cache and compile repeatedly.
- **Server cursors**: As database objects that applications can use to manipulate data sets, server cursors can more readily be kept in memory, helping to improve performance because memory access is faster than hard disk access.

Many I/O-intensive applications can potentially benefit from loading a larger working data set into memory, which is possible in the IA-64 environment. Although applications requiring more memory than the 64 GB limit supported by AWE are rare, certain applications or workloads can benefit from the speed of extracting database pages from extremely large cache memory instead of the disk subsystem.

Additionally, the improved in-processor parallelism capabilities of IA-64 chip sets as compared to 32-bit chips benefits SQL Server 2000 (64-bit) in situations where context switching degrades performance. Administrators can configure SQL Server 2000 to use fibers...
instead of threads for more efficient parallel operations. Using fibers is helpful in three cases: when most CPUs consistently run at or near capacity, when the application executes across many CPUs, and when a high level of context switching occurs.

**Analysis Services considerations**

SQL Server 2000 Analysis Services requires that all dimensions for all cubes—OLAP data structures that offer richer information by providing a 3-D view of a database—be held in memory simultaneously. This is true in both 32- and 64-bit environments; however, 64-bit environments can hold much larger cubes in memory. Because Analysis Services cannot take advantage of the memory extensions of AWE, its memory is limited to 3 GB in a 32-bit environment, even if more memory is actually available.

The additional memory available to IA-64 platforms gives Analysis Services the capability to support very large dimensions or numerous large dimensions. Consider a 64-bit environment for OLAP applications that require:

- Very large dimensions—SQL Server 2000 (64-bit) has demonstrated support for dimensions consisting of more than 50 million members7
- A large number of sizeable dimensions
- Large memory use of process buffers
- Very large cubes—significant performance benefits for very large cubes can be achieved through the use of the file-system cache, helping to reduce the need for physical disk access for base-cube or aggregate data during queries; this can benefit Analysis Services deployments that have extensive fact partitions and a large number of aggregates, even if the dimensions themselves fit into the memory limits of the 32-bit environment
- Fast cube-processing requirements—memory availability helps reduce the need for writing to temporary files on the disk subsystem
- A large number of concurrent users

Before adopting an IA-64 environment for data analysis, administrators must ensure that the following components are available:

- **OLE DB providers**: IA-64 object linking and embedding database (OLE DB) providers are necessary for all data sources used to populate a database in the IA-64 environment. SQL Server 2000 (64-bit) includes an OLE DB provider for accessing SQL Server. When using other data sources from other database vendors, administrators should verify the availability of a suitable OLE DB.

Alternatively, administrators can configure a 32-bit server utilizing Data Transformation Services (DTS) to pump data from other OLE DB sources to the IA-64 SQL target.

- **User-defined functions (UDFs)**: Analysis Services UDFs or other components written in the Microsoft Visual Basic® 6 development system are not supported in the IA-64 environment. Administrators must verify whether these components exist and, if necessary, rewrite them using C++ and compile them using the IA-64 software development kit (SDK).

**Alternatives to IA-64 architecture**

Although an IA-64-based system can offer significant performance increases, it is unrealistic to assume that a 64-bit system will automatically double the performance of a 32-bit system. In some situations, upgrading to IA-64 may not be the best alternative. Instead, for example, administrators may achieve a greater improvement in application performance by increasing the number of CPUs or the speed of the CPUs.

For many SQL Server workloads, the 2 to 3 GB of virtual address space available with a 32-bit platform is sufficient. If a workload performs well using 2 to 3 GB of memory (or when using AWE) and does not require scaling up beyond four sockets, the benefits of fast CPUs in a 32-bit architecture may outweigh the advantages of moving to an IA-64 platform.

Many operations within SQL Server 2000 can benefit from the fast CPU speeds currently available on the modern Intel Xeon architectures, which now have clock speeds exceeding 3 GHz, than from the memory and scalability benefits of the IA-64 environment, which is currently limited to 1.5 GHz. Examples include single-threaded query response times on systems that are not busy—or aggregations, hash joins, string comparisons, and other common operations that have adequate memory on 32-bit systems and reside comfortably on two- to four-socket servers. These applications may be better served by using the fastest 32-bit chip sets, and administrators should be aware of the CPU speed trade-off in such circumstances.

Administrators must also consider the implications of using very large amounts of memory on the IA-64 platform. For example, a system configured with very large amounts of memory could take a long time to shut down because the system checkpoint must flush a significant amount of data. As a result,
IT staff may choose to perform system checkpoints at small intervals, such as every minute. This can be achieved through the recovery-interval server configuration option.

Limitations of SQL Server 2000 (64-bit)
When deciding on an IA-64 platform and SQL Server 2000 (64-bit), administrators should also evaluate potential platform limitations. At press time, they were as follows:

- The IA-64 platform is still a relatively new architecture.
- SQL Server tools such as Enterprise Manager, Query Analyzer, and SQL Profiler are not yet unavailable on IA-64. Best practice is to run these from 32-bit SQL clients.
- No support exists for the execution of DTS packages. DTS packages can be saved on an IA-64 instance but not executed on an IA-64 instance. Best practice is to provide the 32-bit server hosting DTS with a high-speed connection (Gigabit Ethernet) to the IA-64 server that is the source or target of the transformations in the DTS packages, as well as to other data sources.
- No Microsoft-provided Oracle database or IBM DB2 drivers for the IA-64 platform exist, and third-party database vendor support is limited for drivers on IA-64-based Windows platforms. This restricts the capability to define linked-server connections from IA-64 instances to non-Microsoft databases.
- Microsoft Operations Manager (MOM) is not currently supported on the IA-64 platform, and availability of third-party systems management tools is limited. Best practice is to capture the performance counters and events generated by Windows and SQL Server 2000 (64-bit) over the network to a 32-bit instance, and use the tools that support a 32-bit instance.
- Device drivers for I/O, storage area networks (SANs), and other components in the environment may not yet be available for an IA-64 platform.
- The Microsoft .NET framework for IA-64 is not yet available.

Advantages of SQL Server 2000 (64-bit)
SQL Server 2000 (64-bit) addresses the need to provide a highly scalable database platform for memory intensive, performance-critical applications. The IA-64 version of SQL Server 2000 provides massively scalable performance for large, complex queries through:

- Support for up to 64 processors in symmetric multiprocessor (SMP) systems
- Enhanced parallelism

While many SQL Server workloads perform well in a 32-bit environment, the additional memory and processors available to the IA-64 environment are valuable in several situations, including:

- Scale-up scenarios requiring 16-processor or larger SMP servers
- Workloads with large-scale sorting, hash joins, and query memory such as complex relational data warehouse queries
- Analysis Services applications with very large dimensions or large volumes of data that can leverage file system cache
- Other applications that may be memory-constrained in a 32-bit environment

SQL Server 2000 (64-bit) is optimized for the Intel Itanium 2 processor, and uses advanced memory addressing capabilities for essential resources such as buffer pools, caches, and sort heaps—helping reduce the need to perform multiple I/O operations to bring data in and out of memory from disk. Great processing capacity without the penalties of I/O latency can provide a mechanism to achieve new levels of application scalability. Using Itanium 2–based servers with large amounts of memory, SQL Server 2000 (64-bit) is designed to load and process multigigabyte databases in significantly less time than that required in a 32-bit environment.

Moreover, the 64-bit version of SQL Server 2000 can achieve performance and scalability gains while maintaining integration with existing products and applications and offering a simple migration path. SQL Server 2000 (64-bit) can integrate easily into a database server cluster with 32-bit databases. Organizations can use 32-bit application servers connecting to IA-64-based database servers, phasing in the IA-64 technology as required.

Will Claxton (will_claxton@dell.com) is the Dell alliance manager for enterprise applications, focusing on oversight and strategic direction for Microsoft applications. Will has a bachelor's degree from Texas A&M University and a graduate degree from The University of Texas at Austin.