High-Performance Cloud Computing with IN-MEMORY OLTP
Executive Summary


The bet paid off: Broncos win over Falcons 34-19.

That same year, a small Internet company bet big: $2 million on a Super Bowl ad; nearly half of their annual revenue. Minutes after the ad’s fictional hero got his dream job, the servers at HotJobs went off-line under crushing demand.

Internet-enabled companies everywhere got serious about scaling to meet consumer demand and maintaining high availability.

“A few seconds after the piece aired, the traffic spiked on all the machines, flat lined, then all the machines shut themselves off. This could be referred to as a kernel panic.

We had an office panic.”

This paper examines several issues of interest to IT leaders responsible for the performance and availability of mission-critical databases, particularly those with requirements for high input/output speeds. In-memory (RAM) OLTP is examined closely as an important strategy for accommodating heavy throughput workloads. Private cloud and hybrid configurations add to the economic advantages of a database migration to a new infrastructure.
Why worry?

No extensive market research is required to illustrate the need for high-performance computing (HPC) in business:

• Complexity and velocity of data expands algebraically, which can be useful to business intelligence purposes, but must not diminish the performance of day-to-day (or second-to-second) business operations.

• Seasonality, or event-driven user requirements, make an argument for burst capacity to meet performance standards even during episodic peaks in utilization.

• Increasingly, business operations depend on workloads relying on heterogeneous data sources. For example, Internet travel services draw pricing and flight schedules from multiple providers.

• Highly dynamic scheduling workloads such as just-in-time inventory for a large manufacturer may require multiple high-trust inputs to manage real-time operations management.

• With the average consumer possessing several devices and the perceived value of on-demand services anytime, anywhere, brands that deliver satisfying compute experiences can build loyalty over competitors that present slow, sluggish Internet services. In a competitive market, meeting user expectations of instant-response under any conditions can make the difference between a sale and a “drop” or “bounce” statistic in web analytics.

• Whether balancing a checkbook or making high-speed stock trades, there is no tolerance for mediocre performance of a database, especially where fractions of a second can have a profound impact on profit opportunity.

In short: Whether servicing consumers, integrating vendors, or enabling employees in the enterprise, “waiting is death” and SLAs for very high-speed computing are keeping many IT planners and managers awake at night.
The trouble with latching.

“What the heck is this application DOING?”

Whether on disk or in memory, database performance can be stymied by internal latching. Latching is a method by which a data point is locked during a query or process, forcing data updates or changes to occur only in sequence, regardless of the demand for that data. Securing the data to prevent concurrent access ensures the integrity of a data element.

However, system response and reporting performance can be compromised by the necessity to ensure serial execution of queries and commands.

Frequently used queries will already have the data in RAM, so latching is less often a problem. This is especially true when changes aren’t being made to the data, so protecting it is not important.

Yet, as a business exploits the value of their data to innovate, analyzes it in new ways, or otherwise subjects it to unusual queries, the latch/unlatch speed becomes a performance issue. New, infrequent calls require the system to locate and/or assemble the data on disk, read it into memory, and then latch prior to execution so that the integrity is maintained.
At the core of it all.

For several years, scaling compute performance has been accomplished by designing multiple-core hardware solutions. This solution is problematic on a number of levels:

First, serial code (where a given execution can only be performed on one core) limits the true scaling advantage of the added cores. The more serial code present in the program, the less actual increase in compute performance can be realized.

Similarly, maintaining concurrency over a number of cores can be very difficult and requires extreme complexity in database architecture. The cost-benefit of this programming rarely makes the investment worthwhile, further limiting the additive impact of multi-core processing. There have been exceptions in the gaming industry where popular titles have been designed with multi-core scaling as a basis for extreme performance and user satisfaction.

As important as any other consideration: manufacturers need to overcome limitations relating to power consumption (compute per watt) and clock speed limitations of the processors themselves. To put the issue simply: chip technology hasn’t produced an inherently faster processor that doesn’t exceed the thermal cooling system required to dissipate the resulting heat.

Taking a step forward.
(And sleeping better at night.)

Microsoft has made a material advance in processing performance with SQL Server 2014 with In-Memory. SQL 2014 introduces processing of mission-critical data in the RAM of a dedicated cloud or hybrid infrastructure. Reliance on comparatively slower hard disks is vastly reduced. Microsoft claims that processing speed gains as high as 100x can be realized.

Database administrators can select the data tables to process In-Memory, eliminating the requirement of other database solutions to move the entire data set into memory. Minimal adjustments to the application are required to enable In-Memory processing—typically less than one day of developer support, and often managed entirely by the DBA themselves. Businesses that operate with custom line-of-business applications that have become sluggish can realize an extension of their application life cycle with accelerated performance.

When configured with the AlwaysOn features of active secondary instances, very complex reporting and data analysis can be performed on the secondary instances without compromise to the application performance, which continues on the primary instance. Since data tables stay in memory, recovery from scheduled and unscheduled downtime is accelerated. Beginning with the release of SQL Server 2012, processing of internal, external, and unstructured data is enabled and now vastly faster with In-Memory processing of SQL 2014.
Rackspace: The Managed CLOUD Company.

Rackspace currently powers over 10,000 physical servers running Microsoft SQL Server on its private and public clouds, and is the first managed cloud provider to support Microsoft’s In-Memory OLTP. This accelerated platform for dedicated and hybrid configuration is now available to the enterprise more interested on working with their data rather than nurturing their infrastructure.

The persistent effort required to out-pace obsolescence of capital investment coupled with the operating efficiencies of Rackspace’s Managed Cloud makes a compelling argument for selecting Rackspace as your managed hosting provider.

Recognized four times as the Microsoft Hosting Partner of the Year and as a Leader in Gartner’s Magic Quadrant for Cloud-Enabled Managed Hosting, Rackspace delivers a world-class hosting experience for businesses of all sizes.

Economies of expertise.

For a fraction of the cost of a full-time database administrator, Rackspace offers dedicated DBAs available 24x7x365 to enable a rapid migration, upgrade, and ongoing maintenance of In-Memory OLTP. By combining SQL Server 2014 In-Memory OLTP with managed high availability, Rackspace can help customers to speed up OLTP workloads 30x while ensuring maximum uptime and recovery.

When uptime is directly tied to revenue, high-performing databases are crucial. DBA Services by Rackspace extends our partnership with customers, becoming extensions of their existing team of administrators. Our team can handle the company’s day-to-day care and feeding of your database platform, as well as consulting and providing advanced design, architecting, planning, and support services. Rackspace DBA Services can help you rapidly scale up performance and drive greater database consolidation density and reduce total cost of ownership.

For more information, please contact your Rackspace sales consultant or visit: rackspace.com/dba-services.
About Rackspace.

Rackspace® (NYSE: RAX) is the global leader in managed cloud and founder of OpenStack®, the open-source operating system for the cloud.

Hundreds of thousands of customers look to Rackspace to deliver the best-fit infrastructure for their IT needs, leveraging a product portfolio that allows workloads to run where they perform best—whether on the public cloud, private cloud, dedicated servers, or a combination of platforms.

The company’s award-winning Fanatical Support® helps customers successfully architect, deploy, and run their most critical applications. Headquartered in San Antonio, TX, Rackspace operates data centers on four continents. Rackspace is featured on Fortune’s list of 100 Best Companies to Work For.

For more information, visit WWW.RACKSPACE.COM, or learn more from our cloud team at 844-855-7568.

Rackspace supports Windows® operating systems.