

DataCore Lowers Server and Storage Costs with Parallel I/O Processing Technology

Situation

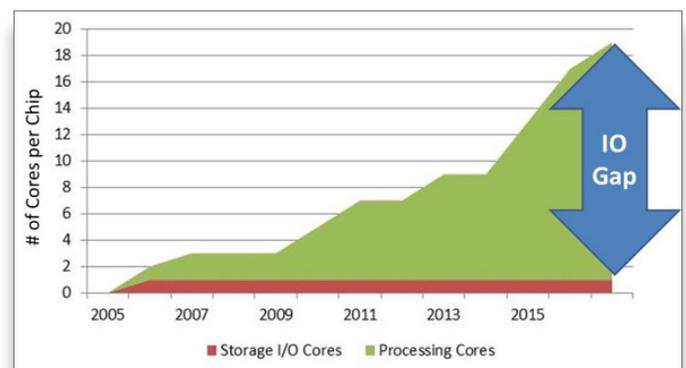
As predicted by Moore's Law, server-processing power has increased by 100% every two years over the last several decades. Until 2007, this improvement was primarily achieved through an increase in the clock rate of the CPU. The continued performance improvement has enabled applications to get more work done in shorter periods of time. Initially, each server hosted a single operating system instance, and there was a tight coupling of hardware and software. Running multiple applications on a server was avoided to prevent a conflict of resources. As processing continued to increase in speed, single applications could no longer consume the processing potential of a server, leaving the server underutilized. This situation drove the development and introduction of server virtualization in the late 1990s. Virtualization leverages the increased processing power of servers by consolidating multiple underutilized servers onto a single physical server. Through abstraction, virtual machines (VMs) replace the original individual physical servers, with each VM allocated its own virtual compute, network, and storage resources to avoid conflict. Using fewer physical servers reduces capital and operational expenses and the VM abstraction provides greater flexibility in the allocation of resources and mobility of workloads.

Following the introduction of server virtualization, servers continued to grow in processing capability, fueled by ever-faster CPU clock speeds. Around 2007, however, CPU chips reached a critical point. As CPU clock rates increased to achieve faster processing speeds, the power consumption and thermal output of the CPU chips exceeded what was operable within an enclosed server chassis. To address the thermal problem while continuing to increase performance, processor vendors began to deploy multiple cores within the same CPU chip in 2008. Multiple cores (independent processing units that can execute instructions simultaneously) enable parallel processing within a single physical package. This multicore approach elevated processing power without requiring an increase in the clock rate and the associated increase in thermal output.

As two-core CPUs entered the market, the onward progression of server performance in accordance with Moore's Law was re-established. However, as the number of CPU cores has grown from two to four to eight and beyond, a new performance issue has surfaced. The root of the problem is that in these virtualized environments only a single CPU core is assigned by the hypervisor to process input/output (I/O) operations. This restriction creates an "I/O gap" between application processing and I/O processing. The aggregation of mixed workloads from virtualized servers coupled with the I/O gap creates a performance bottleneck for applications. The I/O queue depth lengthens as more and more data awaits serial I/O processing before writing to and reading from storage. This situation results in competing workloads running slower delayed by the serial I/O backlog.

This I/O gap will only worsen as the number of CPU cores per server continues to expand. The chart reflects the growth of CPU cores over time and the resultant expansion of the I/O gap.

IT organizations traditionally take two approaches to work around the long I/O queues. Both, however have proven to be expensive and fail to get at the root cause.



The first approach is to add flash or another fast storage device to improve the throughput at the end of the serial I/O process. Even with the faster storage, all I/O processing is still limited to a single core and storage performance still lags behind concurrent application processes. The second approach, often used in concert with the first, is to place fewer workloads per server and spread the serial I/O processing over several physical servers; each taking some percent of the I/O load.

This combination ends up creating severe server underutilization with many idle CPU cores- the very issue that virtualization was developed to resolve. It is clear the I/O gap problem needs a new solution.

Event

[DataCore](#), a leader in software-defined storage and hyper-converged virtual SAN solutions, has introduced Adaptive Parallel I/O software. The software multiplies the performance of virtualized and hyper-converged systems that operate with multiprocessors. This performance benefit is achieved by leveraging all of the power of multicore CPUs for both computational and I/O processing. On January 12, 2016, DataCore announced a new world record for price performance (\$0.08 / SPC-1 IOPS™) using the industry's highly recognized and peer-reviewed storage benchmark, the SPC-1 benchmark from the Storage Performance Council (SPC). SPC is a vendor-neutral, industry standards body with a comprehensive benchmark portfolio utilizing I/O workloads that represent the “real world” storage performance. This real-world behavior is simulated by using both *online transaction processing (OLTP)* and sequential workloads. The industry holds the SPC-1 benchmark results in the highest regard because of consistent, repeatable workloads and vendor-independent testing and validation. DataCore™ SANsymphony™ software-defined storage platform and DataCore™ Hyper-converged Virtual SAN software employ Adaptive Parallel I/O technology to attain these extraordinary price performance advantages.

How the Adaptive Parallel I/O Software Works

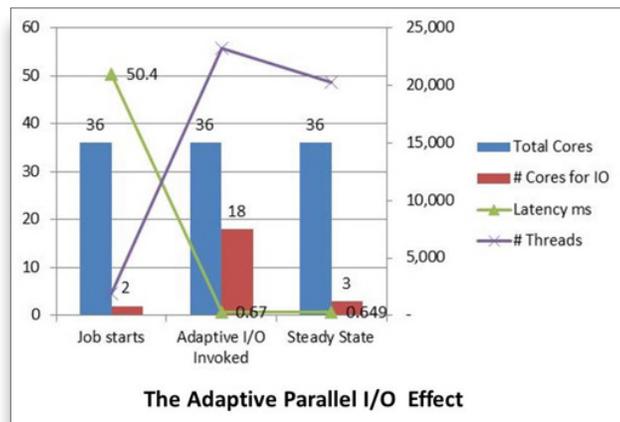
The premise of DataCore's Adaptive Parallel I/O software is simple. Leveraging the parallel I/O work that the founding members of DataCore did back in the 1990s, Adaptive Parallel I/O software utilizes multiple cores for I/O processing, balancing the ratio between computational and I/O processing. The Adaptive Parallel I/O software automatically allocates the number of core resources needed to eliminate the mismatch between computational and I/O processing. There is no need for an engineer or administrator to monitor performance and manually allocate cores to I/O and to continue to tune the software. This effective use of core resources for I/O processing results in two primary benefits: 1) significantly reduces the number of physical servers, and 2) achieves faster application response using lower-cost, commodity-based storage hardware. In fact, DataCore recorded the fastest response time ever measured on the SPC-1 benchmark with an incredible 0.320 milliseconds at 100% load (459K SPC-1 IOPS) on a hyper-converged system costing only \$38,400. That's 3x to 10x faster than competing systems costing several hundreds to millions of dollars.

The first benefit of Adaptive Parallel I/O software is the reduction in the number of physical servers it enables. This reduction in servers increases in importance as the market transitions from traditional enterprise storage (NAS, SAN, DAS) to enterprise server SAN storage, sometimes called hyper-converged. In a benchmark study run by DataCore with a server utilizing two Intel Xeon E5-2696 V3 processors, each with 18 physical cores / 36 logical cores with hyperthreading, significant reductions in latency were achieved with no other changes to storage or the server. In the test, hundreds of Java Virtual Machines (JVMs) representing OLTP applications were started with only two cores dedicated for I/O processing. Response time was 50.4 milliseconds at 315K IOPS. When the Adaptive Parallel I/O software was engaged, up to 34 previously underutilized (5% to 25% utilization) cores were brought to bear to eliminate the I/O backlog.

Latency was reduced to 0.670 *milliseconds*, an improvement of 75X! Once the I/O backlog was cleared and response time goals met, no additional cores were allocated, leaving another 30+ cores for new workloads.

After the Adaptive Parallel I/O software was invoked and reached steady state, the number of threads that could be processed increased from 1,932 to 20,273, an increase of over 10X. When the benchmark stopped generating I/Os, all but the 2 original cores were freed up for whatever processing might come next.

Clearly, the baseline performance of 50.4 milliseconds without Parallel I/O would have dictated distributing the JVMs across multiple physical servers to arrive at a satisfactory I/O response, even though most of the server's cores were idle. Instead of throwing more servers at the problem, DataCore fulfilled the service level agreement (SLA) with sub-millisecond response by putting the available resources in the one server to work in parallel. Results will always be dependent on the number of cores per CPU and the I/O intensity of the workload.



The second benefit is the savings in storage costs. By treating the root cause of the problem, performance requirements can be met or exceeded with less costly storage resources. This benefit is quantified by the aforementioned SPC-1 benchmark results, which can be found on the [DataCore website](#) along with information on how the DataCore solution compares to other storage vendors on cost and performance. What is clear from the results is that storage that costs significantly less than traditional, network-based systems can achieve better results in terms of both IOPS and latency when powered by DataCore Parallel I/O technology.

EMA Perspective

A leader in software-defined storage (SDS), DataCore has been delivering significant cost savings with comprehensive storage services for heterogeneous environments long before the creation of the term “SDS.” By putting multi-core servers to work using Adaptive Parallel I/O technology, DataCore has added a capability to save on capital and operational expenses for both storage and servers. The Adaptive Parallel I/O software enables IT organizations to reclaim the savings of consolidated servers that virtualization delivered, but which had become diminished due to the I/O gap with multicore computing.

The SPC-1 benchmark independently quantifies the performance capability of the software, with a price/performance benefit of five to one over the market share leader. In fact, newer results for DataCore’s streamlined Parallel Server product that includes a modest increase (\$11K) of hardware elevated performance from 459 thousand to 1.5 million IOPS, while reducing latency from 320 microseconds to 100 microseconds. As noted above, a significant savings opportunity also lies in the reduction in servers that is possible with the software. Reducing the number of physical servers from five to one offers a significant savings in capital, software licensing, administrative effort, and environmental expense. Lenovo, Fujitsu, Huawei and Dell offer solutions that include DataCore software with their servers. As Adaptive Parallel I/O technology gains traction, EMA believes that other server vendors will follow suit and offer the DataCore software as an add-on feature. Not to do so would put these vendors at a significant disadvantage.

EMA applauds DataCore for approaching the virtualized server performance problem in a new, more efficient way. DataCore was recently awarded “Best Enterprise Solution for Software-Defined Storage” in the newly released EMA Radar™ Report for Enterprise Software-Defined Storage. For parallel I/O technology, DataCore currently has no competition.

About EMA

Founded in 1996, Enterprise Management Associates (EMA) is a leading industry analyst firm that provides deep insight across the full spectrum of IT and data management technologies. EMA analysts leverage a unique combination of practical experience, insight into industry best practices, and in-depth knowledge of current and planned vendor solutions to help EMA’s clients achieve their goals. Learn more about EMA research, analysis, and consulting services for enterprise line of business users, IT professionals and IT vendors at www.enterprisemanagement.com or blogs.enterprisemanagement.com. You can also follow EMA on [Twitter](#), [Facebook](#) or [LinkedIn](#).

3331.030216