Lab Validation Report

Virident FlashMAX

Storage Class Memory with Sustainable, Multi-dimensional Performance

By Brian Garrett and Mike Leone

December 2011

Contents

ESG Lab Reports

The goal of ESG Lab reports is to educate IT professionals about data center technology products for companies of all types and sizes. ESG Lab reports are not meant to replace the evaluation process that should be conducted before making purchasing decisions, but rather to provide insight into these emerging technologies. Our objective is to go over some of the more valuable feature/functions of products, show how they can be used to solve real customer problems and identify any areas needing improvement. ESG Lab's expert third-party perspective is based on our own hands-on testing as well as on interviews with customers who use these products in production environments. This ESG Lab report was sponsored by Virident.

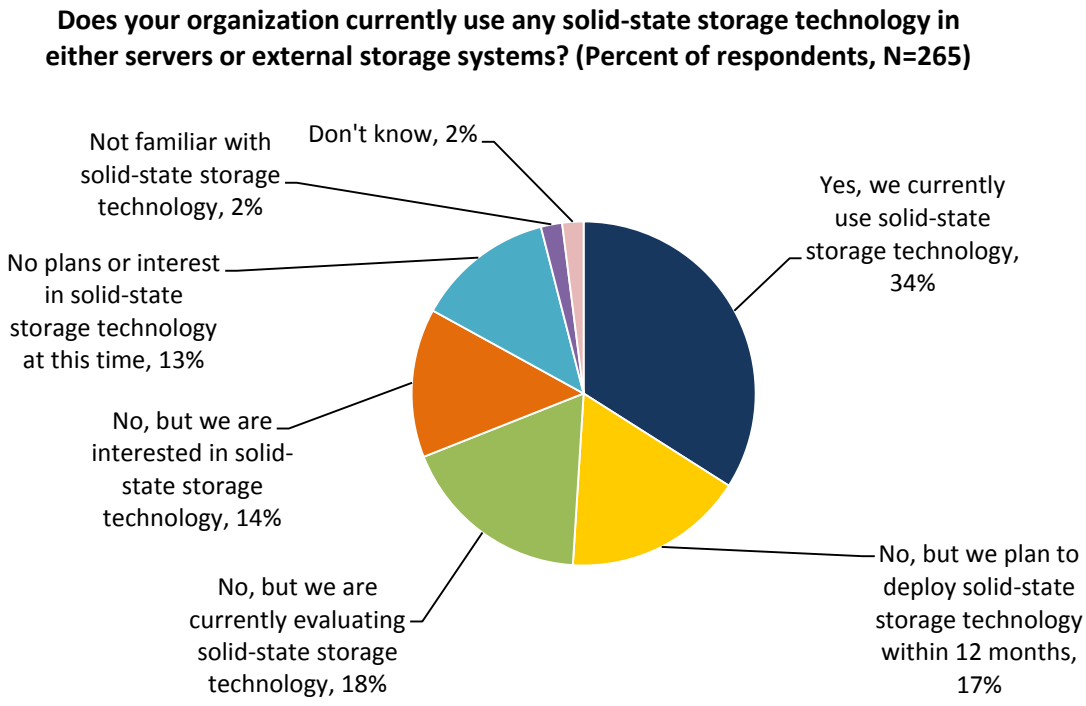
Introduction

Virident Storage Class Memory (SCM) solutions are designed to meet the extreme IO performance needs of business-critical database, Web 2.0, high performance computing (HPC), and data center tier-0 application workloads. This report documents ESG Lab's hands-on validation testing of Virident FlashMAX PCI Express (PCIe) adapters with a focus on the multi-dimensional performance capabilities of SCM from Virident.

Background

A growing number of organizations are using solid-state storage solutions in the data center. As shown in Figure 1, 34% of respondents to a recent ESG survey are currently using solid-state storage technology in either servers or external storage systems, and another 35% are currently evaluating it or have plans to do so in the next 12 months.[1] While early adopters report that improved performance was their primary reason for deploying a solid-state storage solution, they've achieved a number of additional benefits, including improved power and cooling efficiency, increased environmental tolerance, enhanced longevity, and improved reliability.

Figure 1. Current Usage of Solid-state Storage Technology



Source: Enterprise Strategy Group, 2011.

The first wave of widespread solid-state storage adoption began about four years ago, when flash memory became available as a solid-state disk drive tier in enterprise-class disk arrays. More recently, a growing number of organizations have installed PCI flash adapters in servers to create a low-latency pool of primary disk or an extended disk cache. As matter of fact, 21% of respondents to a recent ESG survey indicate that they are currently using a flash memory solution in servers, and 15% plan on doing so in the next 12 months.

Early adopters report that a growing number of performance-critical applications are accelerated with solid-state storage, including OLTP database, ERP financial, OLAP business intelligence, supply chain management, and high-performance computing (HPC) applications. Solid-state storage is also being used to accelerate the performance of consolidated virtual server and virtual desktop infrastructures.

[1] Source: ESG Research Report, *Solid-state Storage Market Trends*, November 2011.

Virident FlashMAX

FlashMAX is a small form factor PCIe adapter that delivers high-speed, flash-based storage in capacities ranging from 300 GB to 1.4 TB. FlashMAX is designed to bridge the ever-growing performance gap between server CPU cores and traditional storage solutions. Typical performance issues common to real-world workloads are eliminated as FlashMAX delivers extremely high levels of predictably fast and sustainable performance for mixed application workloads.

Figure 2. Virident FlashMAX Adapter

Virident uses unique software and hardware to leverage the benefits of flash technology inside of the FlashMAX SCM solution. The SCM architecture provides storage-like capacity and resilience while delivering memory-like performance in a small, common form factor. The software layer, called vFAS (Virident Flash management with Adaptive Scheduling), serves as a gatekeeper, granting access to the flash media as efficiently as possible at all times. Without the need for slower storage protocols or interconnects, major improvements in application performance are realized. These improvements occur due to vFAS's virtualization of the primary flash media, which is accessible to applications via a standard block device interface.

Concerns regarding reliability and data availability are put to rest by vFAS's support of flash-aware RAID. Each flash module is spread across a RAID group, allowing failures to occur non-disruptively while maintaining application data access and operational continuity. FlashMAX with vFAS maximizes flash technology life by providing global wear-leveling techniques. When necessary, less-used parts of the flash media are populated with relocated data to maximize flash media lifetime.

FlashMAX provides a high level of consistency across all application workloads, whether the device is brand new or fully utilized. The challenges associated with many first-generation PCIe flash adapters have been addressed with the multi-dimensional performance capabilities of FlashMAX, which offers:

- High throughput for small, medium, and large IO block sizes.
- Similar levels of performance for random and sequential access patterns.
- High levels of performance for reads, writes, and a mix of reads and writes.
- Sustainable performance over time.
- Extreme performance scalability with multiple FlashMAX adapters in a single server.
- Exceptionally fast response times for real-world application workloads.

ESG Lab Validation

ESG Lab performed hands-on evaluation and testing of the FlashMAX SCM solution at Virident corporate headquarters in Milpitas, California. Testing was designed to demonstrate the multi-dimensional performance capabilities using the industry-standard FIO utility.[2]

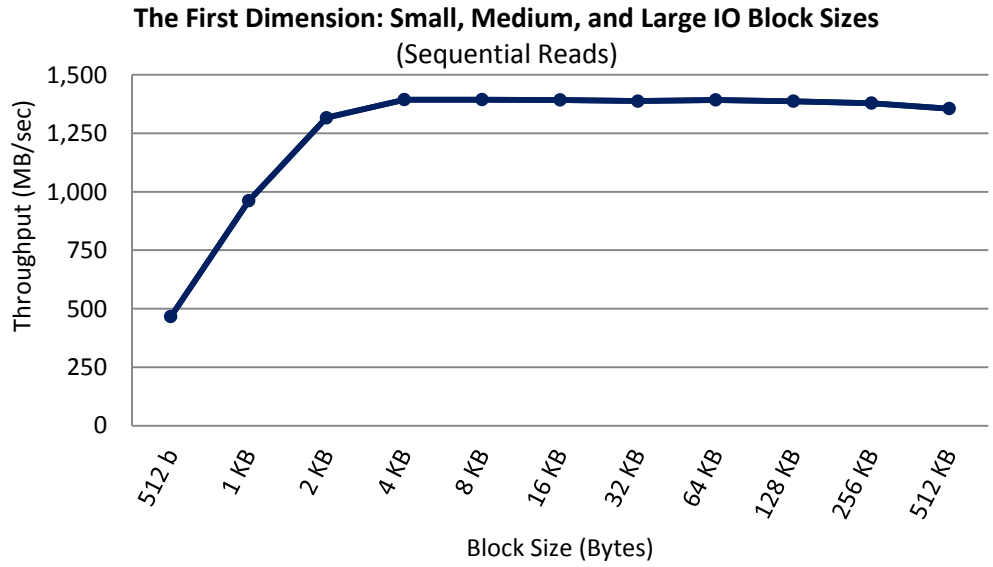
First Dimension: Small, Medium, and Large Block Sizes

Performance-sensitive applications that benefit from solid-state storage often have high throughput requirements (e.g., an HPC application processing a large machine-generated data set with 512 KB IOs). Others require fast performance for relatively small IO requests (e.g., an OLTP database application with 4 KB IOs). A third class of applications requires high performance for a mix of block sizes, with large IOs being used for data requests and small IOs used for metadata requests. FlashMAX SCM optimizes performance for each of these workloads, delivering predictably fast performance for a mix of IO block sizes.

ESG Lab Testing

ESG Lab used the FIO utility to test the sequential read throughput capabilities of a single FlashMAX adapter as it processed IO requests with block sizes ranging from 512 bytes to 512 kilobytes. The results are shown in Figure 3.

Figure 3. Predictable, Scalable Performance



What the Numbers Mean

- Much like the horsepower rating of a car, the aggregate throughput of a storage solution is a good indicator of the underlying power of that storage solution's engine.
- Storage throughput is a measure of the bandwidth available to the system. Throughput can be measured on a stream or aggregate basis. A stream is represented by one application or user communicating through one IO interface to one device. Aggregate throughput is a measure of how much data the storage solution can move, as a whole, for all applications and users.
- Aggregate FlashMAX throughput scaled in a near-linear fashion for the smaller block sizes, shown toward the left side of Figure 3.

[2] http://freecode.com/projects/fio

- Sequential read throughput reached a peak of 1,394 MB/sec at a 4 KB block size.
- A peak aggregate throughput of 1.394 GB/sec (1,394 MB/sec) is an excellent result for a single PCI flash adapter.
- Performance remained predictably high as IO block sizes increased up to 512 KB, shown toward the right side of Figure 3.

Why This Matters

Database, HPC, and virtualized application workloads often have strict performance demands. A performance bottleneck in any of the systems between the application and the data can lead to lost revenue and dissatisfied customers. Meeting the performance demands of IO-intensive workloads using traditional disk-based architectures often leads to over-provisioning, wasted capital costs, increased complexity, and excessive demands on data center infrastructure.

ESG Lab verified that FlashMAX SCM delivers high levels of predictably high aggregate throughput for a mix of IO block sizes. Aggregate throughput performance for a single FlashMAX adapter scaled in a near linear fashion to an extremely high level of 1.394 GB/sec as block sizes scaled up to 4 KB and remained steady at more than 1.25 GB/sec for larger block sizes.

Second Dimension: Sequential and Random Access Patterns

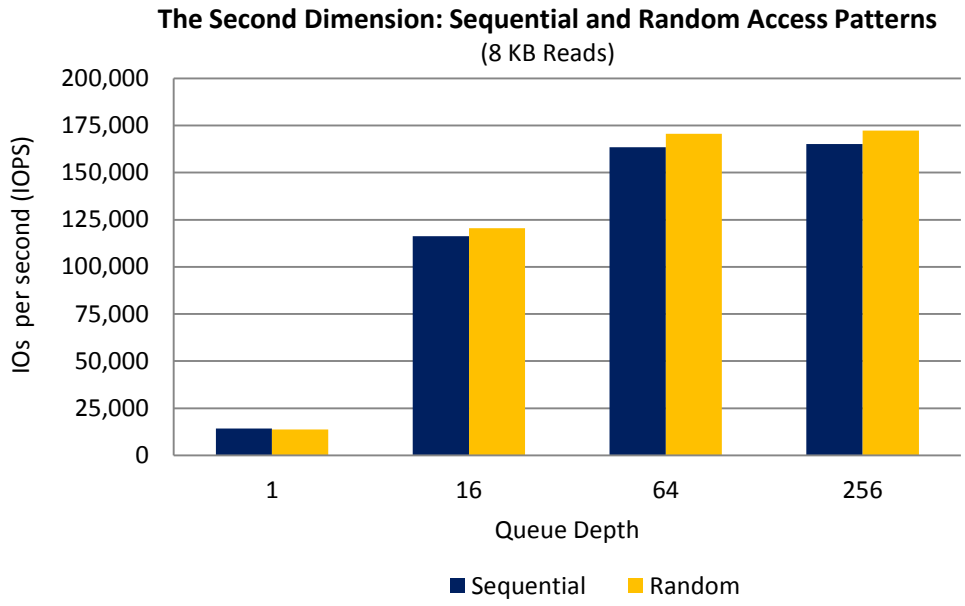
HPC applications can have a variety of workload access patterns, including random, sequential and a mix of random and sequential access. Regardless of the IO access pattern, one key advantage of PCI flash solutions is the ability to perform thousands of times more IOs per second (IOPS) than a traditional spinning disk drive.

While early adopters of flash storage solutions in the HPC market were initially focused on throughput-intensive sequential workloads, broader adoption in the wider (more horizontal) database, server virtualization, and desktop virtualization markets has begun to take off. Multi-user database and virtualization applications tend to have more random IO access patterns. First-generation flash solutions tended to have different performance characteristics for random and sequential access patterns. FlashMAX SCM provides similar levels of high performance for random and sequential workloads.

ESG Lab Testing

As shown in Figure 4, ESG Lab ran both random and sequential 8 KB reads at four different queue depth sizes to show not only how the access patterns perform, but also how they scale.

Figure 4. Random and Sequential Performance



What the Numbers Mean

- Online database applications, including those that rely on the latest version of Microsoft SQL Server, are typically composed of a mix of random and sequential IO access patterns, with a block size of 8 KB as tested during this phase of the ESG Lab Validation.
- Extremely fast sub-millisecond response times of 38 and 64 microseconds were recorded for 4 KB random write and read workloads respectively.
- The extremely fast response times recorded during ESG lab testing are significantly faster than a drive form factor SSD. SSDs are slower due to the additional overhead of an IO protocol (e.g., SAS) vs. the low latency of a PCIe bus with FlashMAX.
- The extremely fast response times recorded during ESG lab testing are 25 to 200 times faster than a traditional disk drive.

- The total number of IOPS processed at a single queue depth was slightly higher for the sequential access pattern (14,286 IOPS) compared to the random access pattern (13,716 IOPS).
- At a queue depth size of 16, the random access slightly outperformed the sequential access.
- As queue depth continued to increase, performance eventually leveled out at a little over 170,000 IOPS for random 8 KB reads and 165,000 IOPS for sequential 8 KB reads. It would take more than 1,000 power-hungry disk drives to deliver 170,000 random IOPS.

Why This Matters

For many transactional databases and business-critical applications, faster application performance means more revenue, better customer satisfaction, and greater productivity. While solid-state storage can clearly be used to accelerate application performance, one of the challenges with first-generation solutions is the significant difference in performance between random- and sequential-access IO patterns.

ESG Lab confirmed that FlashMAX delivers similar levels of performance for random and sequential read workloads. With excellent response times of less than 100 microseconds, significantly faster than a drive form factor SSD, FlashMAX performance varied no more than 5% between random and sequential 8 KB reads as up to 256 IOs were queued.

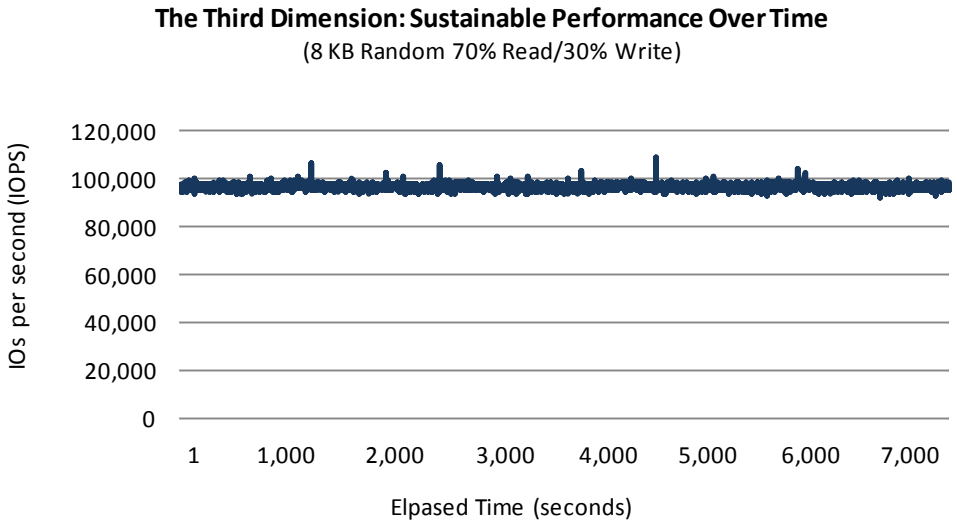
Third Dimension: Sustainable Performance over Time

Solid-state storage solutions have historically had problems maintaining performance over time. As space is needed to service incoming write requests, a "garbage collection" process needs to be run in the background. The goal of the garbage collection process is to free-up necessary space. Because this process happens along with all of the other user requests, enterprise application performance can be greatly impacted. This can be seen by a severe drop in throughput and large response-time spikes. The phenomenon is often referred to as a "write cliff," which aptly describes how write performance seems to fall off a cliff over time. FlashMAX SCM was designed with a goal of providing sustainable performance over time and avoiding the write cliff problem.

ESG Lab Testing

ESG Lab tested the ability of a FlashMAX to sustain performance over time. An online database workload was emulated using a mixed 8 KB random workload with a mix of 70% reads and 30% writes. The duration of the test was set for more than two hours to allow ample time for the flash device to reach full capacity and potentially get impacted by the garbage collection process. The results are shown in Figure 5.

Figure 5. Mixed Workload Sustainability



Performance varied minimally over the entire test, but the sustainability is clear. The FlashMAX was able to deliver consistent performance of just under 100,000 IOPS throughout the full duration of the test.

Why This Matters

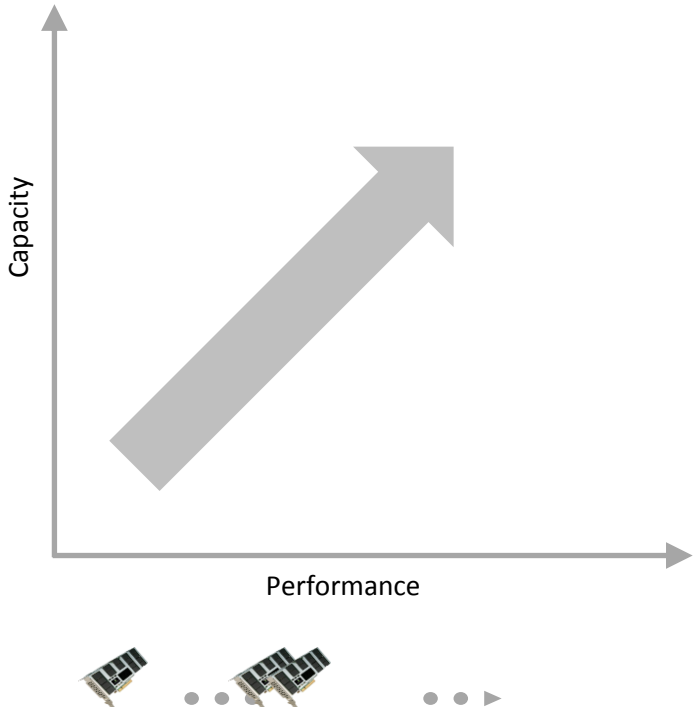
Companies continuously face challenges in cost-effectively meeting service level agreements for business-critical applications—especially for IO-intensive database applications with strict performance requirements. Failure to meet performance requirements can result in lost productivity and costly loss of services. Over-provisioning in an attempt to avoid potential performance problems with flash storage over time (e.g., a write cliff) is a waste of money.

ESG has confirmed that FlashMAX delivers predictable performance over time. Performance varied less than five percent (0.96% standard deviation) as a simulated OLTP database workload exercised all of the capacity within a single FlashMAX adapter.

The Fourth Dimension: Scaling Performance with Additional FlashMAX

Scalability, put very simply, is the ability to elegantly handle more work or to physically grow to accommodate that additional work. In this case, scalability applies specifically to the near-linear performance increase that can be achieved with more than one FlashMAX adapter installed in a single server. Illustrating this concept, Figure 6 shows how performance and capacity increase in a near-linear fashion with each additional FlashMAX adapter that is installed in a server.

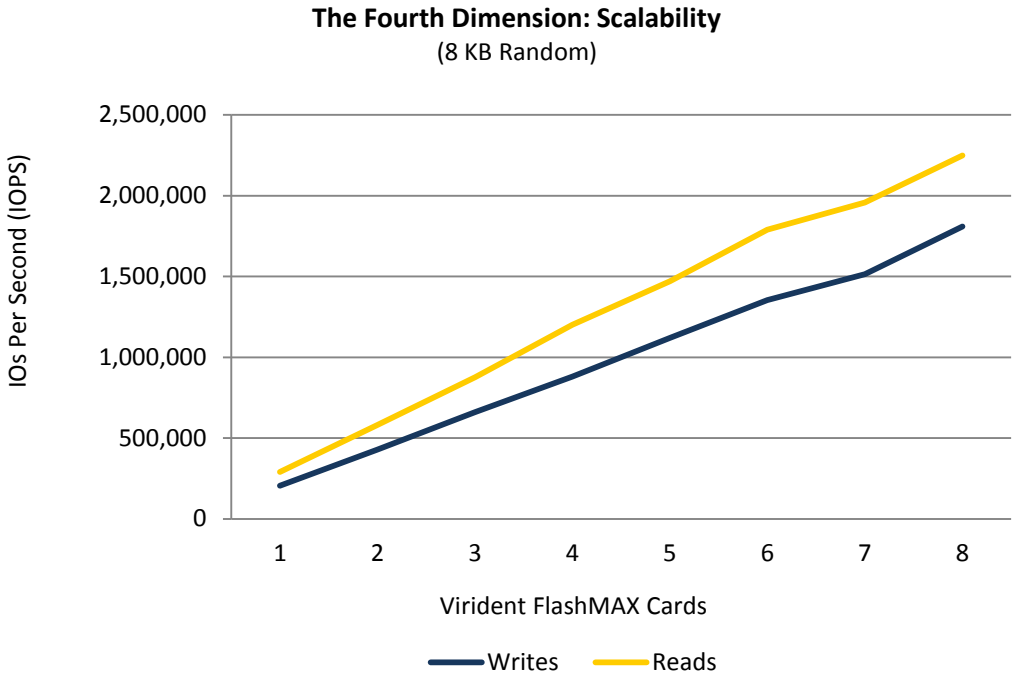
Figure 6. Scaling Capacity and Performance with Multiple FlashMAX Adapters



ESG Lab Testing

Up to eight FlashMAX adapters were installed in a powerful NEC GX server with sixteen PCI slots during this phase of testing. The FIO utility was used to test 8 KB random read and write workloads. Results are shown in Figure 7.

Figure 7. FlashMAX Performance Scalability



What the Numbers Mean

- Performance scaled in a near-linear fashion for random read and write workloads as up to eight FlashMAX adapters were installed in a single server.
- Performance peaked at an extremely high level of 2.2 million IOPS and 10.6 GB/sec of throughput on a single server.

Why This Matters

As data continues to grow, so to do performance requirements. Adding a bunch of underutilized disk drives to solve performance problems with IO-bound applications decreases the efficiency of the IT infrastructure by increasing acquisition, maintenance, power, cooling, and data center floor space costs.

Scalability testing with up to eight FlashMAX adapters delivered extremely high levels of near-linear performance that peaked at more than 10 GB/sec of throughput and two million IOPS with a single server.

Fifth Dimension: Real-World, Mixed-Application Workloads

Having looked at the throughput, IOPS, and response-time ratings of the turbo-charged FlashMAX engine, here's where ESG Lab found "the rubber meets the road" when examining FlashMAX performance with real-world application workloads.

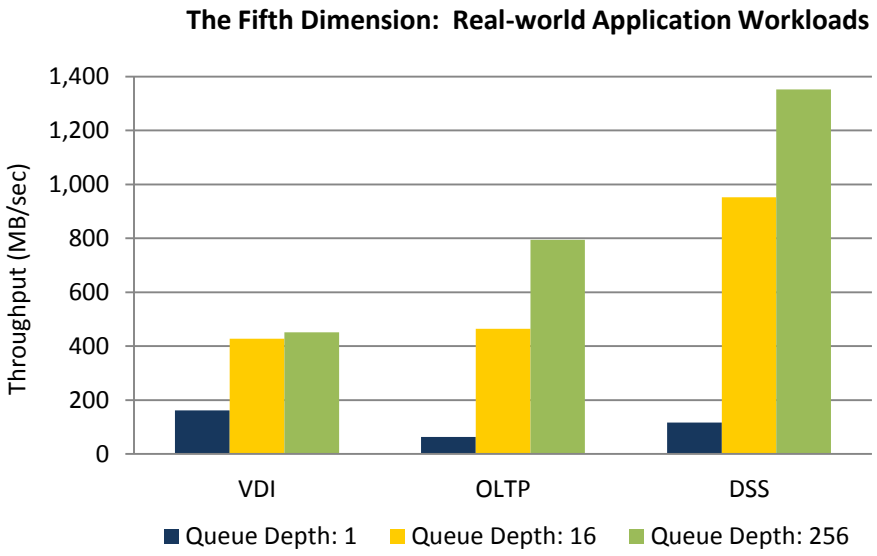
ESG Lab Testing

ESG Lab used the FIO utility to measure the performance of a single FlashMAX card for three application workloads:

- Virtual Desktop Infrastructure (VDI): Designed to emulate a virtual desktop environment composed of heavy knowledge-worker users sharing a common gold image (a.k.a., a linked clone). This workload is composed of 80% 16 KB random writes and 20% 16 KB random reads.
- Online Transaction Processing (OLTP): Order entry and reservation systems are two examples of OLTP applications. Oracle and Microsoft SQL Server are two examples of database applications used to create such OLTP applications. OLTP applications are characterized by a number of users accessing a shared system in parallel. This workload was composed of mostly random (70%) reads.
- Decision Support System (DSS): This workload, also referred to as data mining, emulates a database application that is doing a large-scale random query with a block size of 4 KB. An end-of-month analysis of the effect of a coupon-redemption program on same-store sales is an example of a decision support application.

Figure 8 shows the throughput scalability of these applications as queue depth increases.

Figure 8. Mixed Real-World Application Performance



What the Numbers Mean

- Having one outstanding request at a time, the VDI workload was able to achieve 162 MB/sec throughput, and it eventually scaled up to more than 400 MB/sec, with as little as 16 outstanding requests.
- The OLTP simulation reached a maximum of 794 MB/sec throughput at a high queue depth of 256.
- Showing the largest scaling factor, the DSS workload scaled from 117 to 1,352 MB/sec as queue depth increased.

- The performance of a single FlashMAX adapter that was recorded during simulated VDI workload testing can be used to support more than 1,000 heavy desktop users.[3]
- A traditional disk array with more than 1,000 disk drives would be needed to deliver the performance recorded during the OLTP testing with a single FlashMAX adapter.

Why This Matters

Companies continuously face challenges in cost-effectively meeting service level agreements for business-critical applications, especially for IO-intensive VDI, OLTP, and DSS applications with strict performance requirements. Attempting to over-provision to avoid performance problems is a waste of money. Yet, a failure to meet the performance requirements can result in a costly loss of productivity or services.

ESG Lab has confirmed that FlashMAX is designed to comfortably meet the demanding performance requirements of these business-critical applications. ESG Lab confirmed that FlashMAX delivers high levels of mixed read/write performance scalability for VDI and OLTP workloads (up to 800 MB/sec) and extremely high levels for a read-only DSS workload (up to 1.3 GB/sec).

[3] Using a conservatively high estimate of 20 IOPS per user

ESG Lab Validation Highlights

- ☑ Predictable performance scalability with a variety of IO block sizes that peaked at 1.39 GB/sec from a single FlashMAX adapter (4 KB sequential reads).
- ☑ Predictably fast performance for reads, writes, and a mix of reads and writes for workloads simulating real-world OLTP, VDI, and DSS applications.
- ☑ Nearly identical levels of high performance for sequential and random read workloads.
- ☑ Sustained performance of 96,825 IOPS over two hours for an 8 KB OLTP workload.
- ☑ Up to 345,046 IOPS from a single FlashMAX adapter for a 4 KB random read workload.
- ☑ Extremely fast sub-millisecond response times (38 and 64 microseconds for 4 KB random write and read workloads, respectively).
- ☑ Near-linear performance scalability as up to eight FlashMAX cards delivered up to 2.2 million IOPs and 10.6 GB/sec of throughput on a single server.

Issues to Consider

- ☑ The high cost of solid-state capacity compared with traditional hard drive capacity has focused early adoption mainly among businesses whose revenue depends strictly on application performance (e.g., trading applications within the financial industry). As a matter of fact, ESG research indicates that the high cost of solid-state capacity is the number-one objection by organizations that have not yet deployed a flash-based storage solution. As the cost of flash capacity decreases over the next three to five years, ESG expects that adoption will grow in the broader horizontal enterprise IT market; especially within server virtualization and desktop virtualization environments.
- ☑ Early adopters considering using a flash-based PCI adapter in a server to solve a performance problem with a business-critical, high-performance application should consider the extra costs of installing FlashMAX adapters in multiple clustered servers for high availability and failover. Solid-state disk drives installed at the other end of the wire (in a SAN-attached storage array) are a viable alternative for more cost-effective sharing and failover. However, they are typically slower than a PCI flash adapter. A FlashMAX adapter that's engineered to work within a SAN-attached disk array could be used to create a simply elegant, highly available alternative that cost-effectively accelerates the performance of tier-0 applications and consolidated virtual server environments.
- ☑ ESG Lab ran performance scalability tests on a high-end server that could support multiple FlashMAX adapters in a single server. The server being tested needed to have enough PCIe slots to support eight FlashMAX adapters, and enough processing power to support the necessary requests to drive each FlashMAX to its limit. To achieve the performance scalability results documented in this report, a high-end server with similar specifications is recommended.[4]

[4] The specifications for the server used during ESG Lab testing are listed in the Appendix.

The Bigger Truth

The growing gap between the speed of servers and traditional disk-based storage solutions is causing a number of problems in the data center. Though flash memory solutions have filled this void by serving as an answer for some of the most performance-critical application workloads in recent years, issues still exist. First-generation flash-based storage solutions often have challenges maintaining predictably high levels of performance, over time, for real-world applications with mixed-IO patterns. Sustainable performance over the life of an adapter is a particularly vexing challenge due to the impact of background garbage collection processes.

ESG Lab has confirmed that Virident FlashMAX is a next-generation PCI flash adapter that leverages intelligent algorithms to provide high levels of sustained, multi-dimensional performance. Extremely low latencies and high levels of performance were recorded with a variety of workloads. ESG Lab was most impressed with the paucity of "saw-tooth" and "drop-off" performance patterns associated with first-generation PCI flash adapters. Scalability testing with up to eight FlashMAX devices delivered extremely high levels of near-linear performance scalability that sustained more than 10 GB/sec of throughput and two million IOPS with a single server.

With a proven ability to deliver predictably fast real-world application performance over the life of the device, FlashMAX is well suited for the growing number of performance-sensitive OLTP, OLAP, DSS, HPC, and VDI workloads that are migrating from traditional disk drives to high-speed flash memory. ESG Lab believes that Virident, with the FlashMAX family of storage-class memory solutions, has unlocked the potential for affordable, large-scale deployment of flash technology in the modern data center.

Appendix

Table 1. ESG Lab Test Bed

Server	NEC Express5800/A1080a(GX) with 16 PCIe slots, 40 Intel Xeon CPU cores, and 64 GB RAM
Storage Class Memory	Virident FlashMAX SLC, 800 GB
Workload Generator	FIO, version 1.58

Enterprise Strategy Group | **Getting to the bigger truth.**

20 Asylum Street | Milford, MA 01757 | Tel: 508.482.0188 Fax: 508.482.0218 | www.enterprisestrategygroup.com