Analysis of IBM® FlashSystem™ 840 with VMware VMmark 2.5 benchmark

Evaluation report prepared under contract with IBM

Introduction

IBM commissioned Demartek to evaluate the performance of IBM FlashSystem 840, an all-flash storage array, by running the VMware VMmark2 benchmark against systems under test using FlashSystem 840 as the virtual machine storage.

Key Findings

IBM FlashSystem 840 achieved a near perfect linear scaling for VMmark2 scores across a range of tiles, comparable to or better than similar published results. Latency values from array performance were in the sub-millisecond range, resulting in excellent Quality of Service scores from the benchmark.

The Need for More Bandwidth and Lower Latency

Virtualization is a key technology for consolidating compute environments from lightly utilized servers to more highly utilized, shared computing platforms. While some enterprises may be able to consolidate homogenous workloads to take advantage of complementary I/O and processing profiles, many others must support a diverse grouping of applications that make differing demands on both processing and storage.

Sequential operations and high volume workloads demand increasingly more bandwidth in the form of MB and GB delivered per second and high IOPS. However, highly transactional workloads also have a strong dependency on the time required to access data, referred to as latency.

A diverse workload, such as often found in virtualized environments, will make all these demands on back-end storage. Even if a storage system can deliver very high numbers on MB/s or GB/s and I/O operations per second, if the time required to access that data is above a certain threshold, the overall performance will be perceived as poor by the user.

Flash technology, properly implemented, drives up bandwidth and IOPS while lowering latency.

IBM FlashSystem 840

IBM FlashSystem 840 is an all-flash storage system designed for very low, sub-millisecond latency, trademarked as IBM MicroLatency™.

IBM states that FlashSystem 840 is capable of 135µs read, 90µs writes, and up to one million IOPS to support the diverse and intensive workloads likely to be found in a heavily consolidated, virtualized environment. This is accomplished in a high density, 2u chassis of 2TB to 48TB of eMLC flash, drawing an efficient 625 Watts of power. The system has no single point of failure, and flash modules, power supplies, fans, batteries, and canisters are hot swappable and accessible from the front or back of the system. Software and firmware updates can be completed with the system up and running.

FlashSystem utilizes IBM's Variable Stripe RAID™ at the flash module level, as well as traditional RAID5 at the system level. AES 256 hardware-based encryption is available for data at rest.

The storage system offers 16 ports of 8Gbit FC, 8 ports of 16Gbit FC, 8 ports of 40Gbit Infiniband, 16 ports of 10Gbit iSCSI, or 16 ports of 10Gbit FCoE. For this paper, the 16Gbit FC target was implemented.

IBM FlashSystem 840

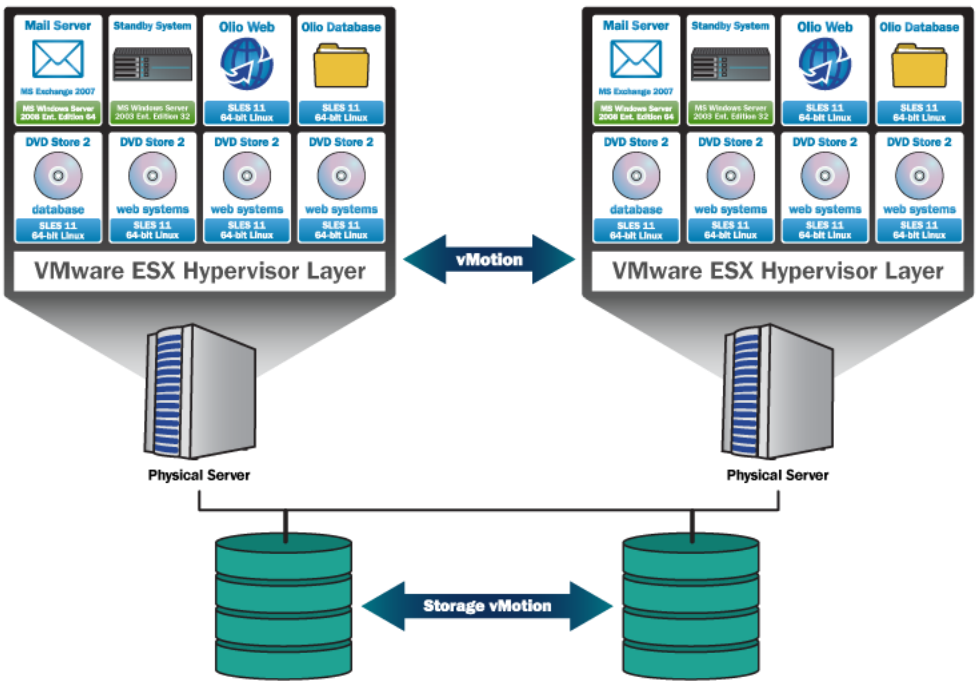
Test Configuration and Procedures

VMware VMmark is a multi-host data center virtualization benchmark which reproduces a complex virtualization environment running diverse workloads.

The benchmark is deployed on virtualization hosts in one or more "tiles" to scale workload intensity. Each tile is composed of eight virtual machines representing a diverse application load across the tile, made up of the following:

- Mail Server
- Web 2.0 load simulation
- E-commerce simulation
- Standby VM

Infrastructure workloads such as server and storage vMotion are included as infrastructure loading components. A deployment example is provided below.



Each tile's workload and performance is managed and recorded by a dedicated client external to the vSphere cluster.

For information on the setup and execution of the VMware VMmark2 benchmark, please refer to http://www.vmware.com/products/vmmark.

The environment for this analysis was conducted on the following compute resources. Specific host and VM setup was compliant with VMmark2.5.2 requirements.

Test Cluster

- 1x PCIe Gen2 Server, 2x Intel Xeon E5-2690, 2.9 GHz, 16 total cores, 32 total logical processors, 192 GB RAM
- 2x PCIe Gen3 Servers, 2x Intel Xeon E5-2690 v2, 3.0 GHz, 20 total cores, 40 total logical processors
- Emulex LPe 16000 series 16Gb FC adapters
- Emulex OCE 14000 series NIC,
- ESXi 5.5
- 1x Windows Server 2008 R2 64 bit with MS Exchange 2007 R2 VM (MailServer)
- 1x Windows 2003 R2 32 bit VM (Standby)
- 6x SLES 11.2 VMs (Olio and DVDstore2)
- IBM FlashSystem 840, 40TB drive space, 8 16GbE ports
- IBM 2498 16Gb FC switch
- 48 HDD array (2 shelves plus controller) for comparison

Prime Client\Client0 Standalone Host

- 1x PCIe Gen3 Server, 1x Intel Xeon E3-1275 v3, 3.5GHz, 4 total cores, 8 total logical processors, 32 GB RAM, Samsung SSD 840 PRO internal drive
- Windows Server 2008 R2
- Exchange Loadgen 2007
- Cygwin64
- STAF 3.4.17
- Other client software as bundled with VMmark2.5 client from VMware

Client ESXi Hosts\Virtual Clients\vCenter

- 1x PCIe Gen2 Server, 2x Intel Xeon X5690, 3.46GHz, 12 total cores, 24 total logical processors, 144GB RAM, OCZ ZR4 PCIe SSD
- Intel 82599EB 10Gb adapter

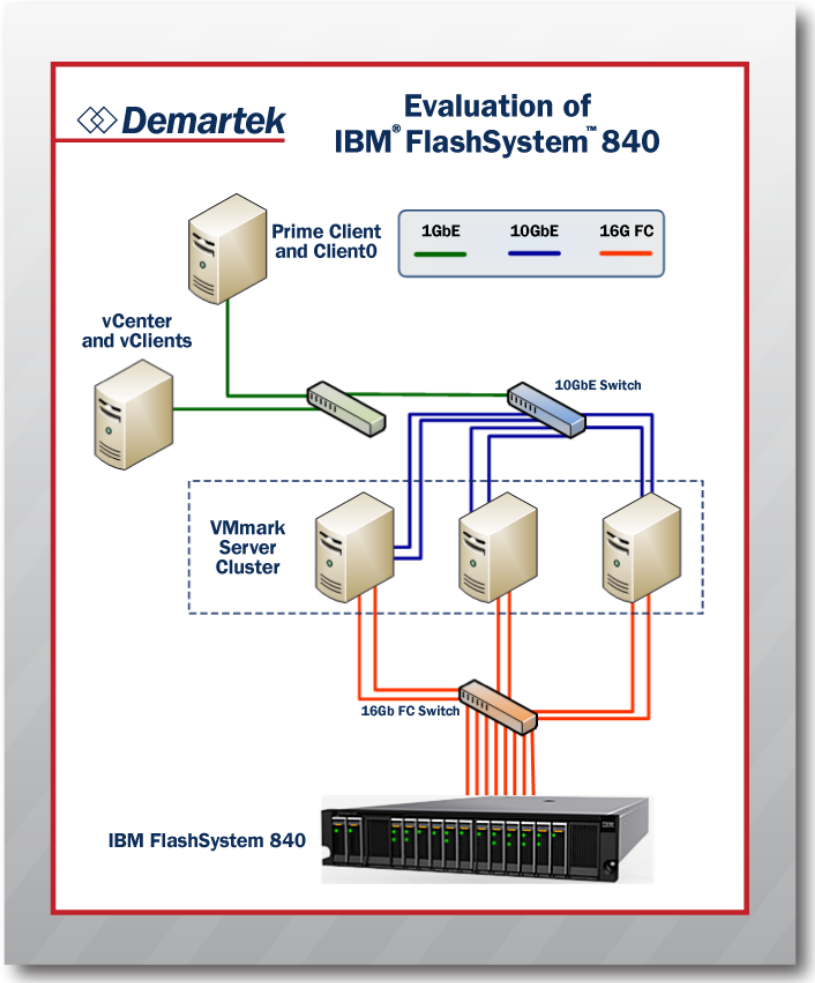
- ESXi 5.1
- 8x Windows Server 2008 R2 64 bit with MS Loadgen 2007 VM(Client VMs)
- 1x Windows Server 2008 R2 64 bit VM (vCenter 5.5 server)

Network Infrastructure

- Cisco 5020 10GbE switch (vMotion, Olio database/Web data networks)

A three node ESXi server cluster was created with IBM FlashSystem 840 array as storage for all VMmark tiles. Separate 10GbE networks were created for VM communication and vMotion.

A single standalone host was provisioned for the Prime Client, with all other clients deployed as virtual machines on a single ESXi 5.1 host.



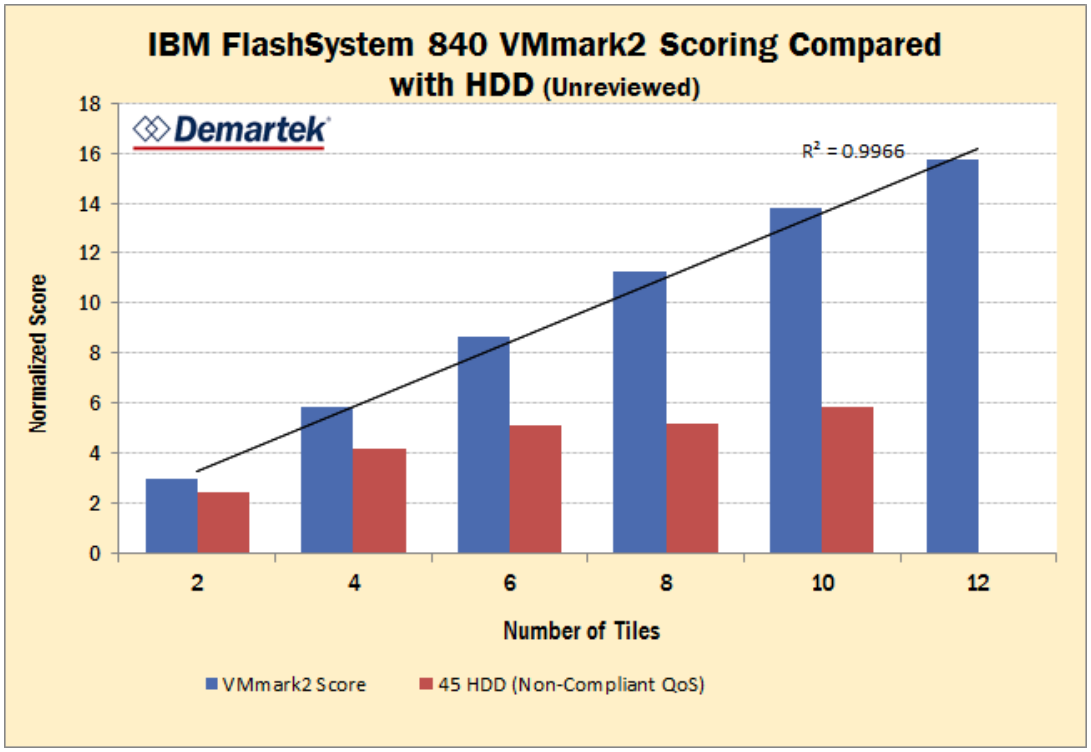
Two 250 GB volumes were provisioned from the IBM FlashSystem 840 array for each tile and deployed as datastores to the cluster. This provided the drive space required for each tile and a datastore for storage vMotion.

Performance Results

IBM FlashSystem 840 clearly supported the number of VMmark tiles tested in this analysis. Performance bottlenecks were intended to be the result of storage limitations. The intent was to continue to scale up sufficient compute and network infrastructure so that any performance throttling would ultimately come from the storage. However, the FlashSystem 840 performance exceeded the processing capacities of the ESXi compute environment provisioned for this project.

When running on IBM FlashSystem 840 storage, the normalized VMmark2 scores generated by the benchmark demonstrate almost no plateau such as would be expected if performance were reaching any kind of bottleneck. The R-squared value of nearly 1.0 shows just how nearly perfectly linear these data points are as the workload scales up in tiles.

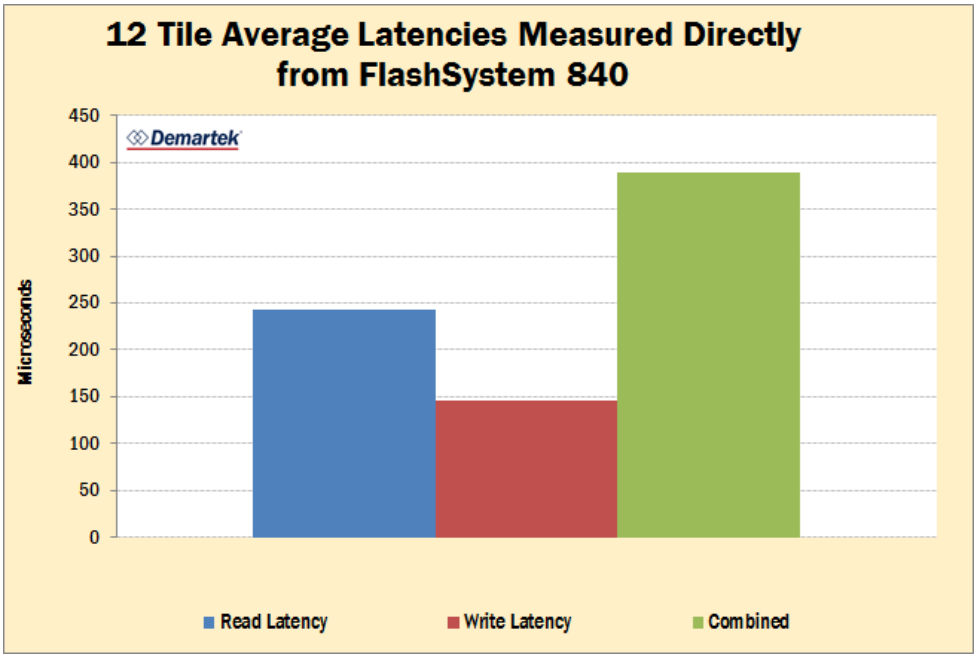
This contrasts sharply with the scores seen when the same VMs are deployed on four times the number of spinning hard disk drives. The hard disk drive array displays both a low VMmark2 score and very prominent flattening of its trending. It is worth noting that there is no 12-tile score for the HDD storage. This is due to the benchmark workload consistently failing with twelve tiles deployed on the 48 HDDs.



From the trending displayed here, it is clear that the IBM FlashSystem 840 array did not reach the limit of VMmark2 tiles it can support with acceptable Quality of Service (QoS) scores. Because there is no indication of even a slight tapering off of the score, it is reasonable to assume that if CPU, memory, and network resources were scaled up, the FlashSystem 840 array would support a greater number of tiles.

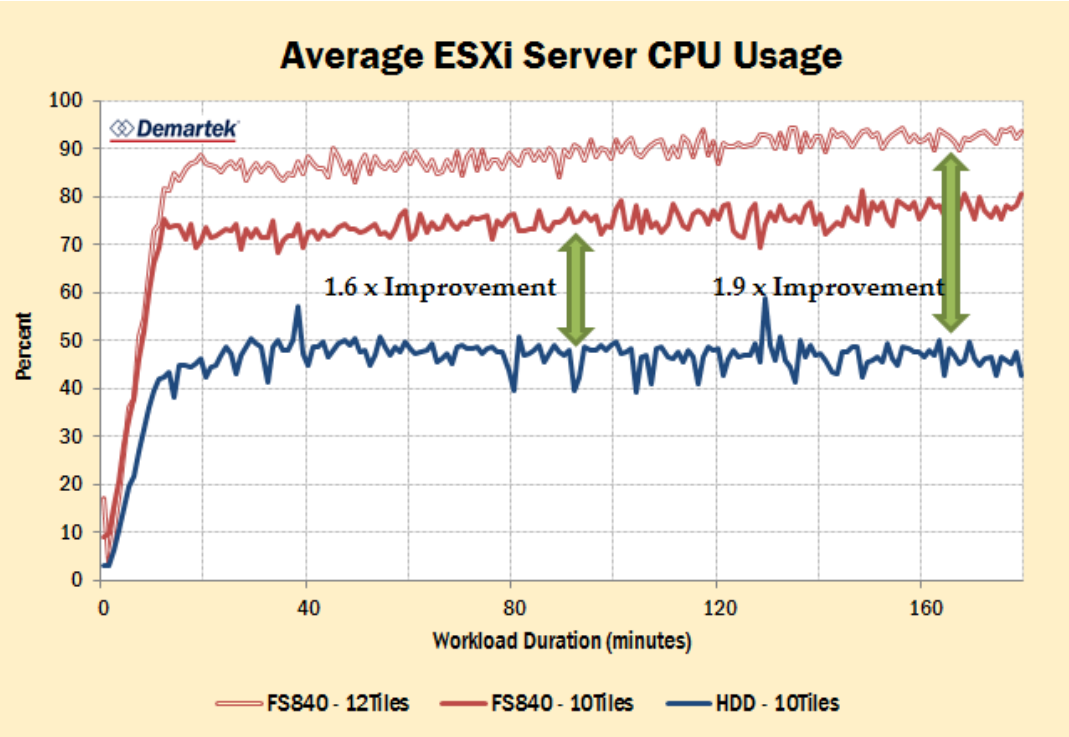
The scores achieved with this ESXi environment compare favorably, or exceed, scores for similar numbers of tiles published at VMware's VMmark2 site at http://www.vmware.com/a/vmmark.

Storage latency is another crucial metric to consider. User experience is highly dependent on latency values. The VMmark2 QoS scores include an aggregate of latency measured from the host. We can go directly to the storage system to query latency at the device level to get a picture of storage-side latency. For flash devices such as SSDs, low single digit millisecond latency values are generally considered to be good performances. IBM FlashSystem 840 as billed achieves extremely low latencies. The graphic below demonstrates sub-millisecond read, write, and combined latencies for the 12-tile test scenario.



The low latencies produced by the FlashSystem 840 array have the added benefit of improving the value of the compute environment by increasing overall utilization. Because the FlashSystem 840 array is able to service more I/O requests, server CPUs spend

less time in an I/O wait state and more time in active processing. The benefit of this is clearly seen by comparing the VMmark 2 scores.



Keeping in mind that the hard drive environment had four times the number of drives, the compute environment is 1.6 times more productive with FlashSystem 840 at ten VMmark2 tiles. There is a 1.9 times improvement when we consider the 12-tile configuration supported by the FlashSystem 840 array. As mentioned earlier, there is no corresponding 12-tile data from the HDD backed environment because that workload was unable to complete without erring.

With FlashSystem 840 providing storage for the test environment, we see the processing capacity fully utilized, clearly demonstrating that the IBM storage system is not a performance chokepoint in this benchmark environment.

Summary and Conclusion

Raw bandwidth is becoming easier to achieve with flash-based storage and fast storage targets. However, a combination of high bandwidth with low latency is becoming increasingly important to maximize the user experience. This is true of highly transactional, standalone applications, but even more so in highly consolidated virtualization environments with diverse workloads.

The storage administrator needs to be cognizant of this, while also understanding the expected return on the hardware investment, including the total cost of operation, to make an informed recommendation on storage technology.

The VMmark2 benchmark mimics a busy virtualized data center in which it becomes difficult to plan for predictable I/O patterns. This increases the importance of a high performance, flexible backend storage system that provides the bandwidth to service high throughput applications, while also simultaneously providing low enough latency to maintain the QoS of strongly transactional or otherwise time sensitive applications.

IBM FlashSystem 840 provides both these features. The raw capacity and throughput performance clearly exceeded the requirements of the twelve tile VMmark2 virtual data center, while the very low latencies offer an exceptional user experience. The VMmark2 benchmark scores compare quite favorably with similar published results available for review from VMware at http://www.vmware.com/a/vmmark.

The single FlashSystem 840 array tested with the VMmark2 benchmark in this analysis appears more than capable of scaling well beyond the compute limits of this three-server test environment.

From a practical perspective, deploying IBM FlashSystem 840 storage into an ESX/ESXi environment creates a more efficient compute environment. This translates into greater consolidation and a higher VM density, or better support for applications with a very strong processing demand.

The most current version of this report is available at http://www.demartek.com/Demartek_IBM_FlashSystem_840_VMmark_Evaluation_2014-09.html on the Demartek website.