Cloudmark Slays Spam with Fusion ioMemory™ Solutions

Security company improves performance in several areas by an order of magnitude and cuts server footprint by more than half.

Summary of Benefits

- **5x improvement** to database replication performance
- **5x improvement** to data-intensive queries
- **10x improvement** to analysis server performance that enabled 5x improvement to analysis routines
- **Simplified** maintenance by eliminating at least 210 disk node failure points from its system
- **Improved reliability and availability** with full system redundancy
- **Recovered 27U** of rack space
- **Greatly lowered** power and cooling expenses

The Challenge

Cloudmark, a messaging security provider, is a veteran in the arms race against spam, phishing, and viruses. Spammers are constantly looking for ways to beat the system, and Cloudmark's reaction time to new threats is of paramount importance to fighting new attacks. It invested in a high-performance disk solution, but data growth ate up every performance upgrade within a few months.

Always on the lookout for innovative ways to improve service at a lower cost, Ryan White, director of operations, began looking for a cost-effective and scalable performance solution that did not require continual outlays on disks.

Cloudmark had a couple of systems that required perpetual upgrading. The first system hosted its messaging security application and required database replication. As its data load grew, so did its performance requirements and disk investment. In order to get the top-notch performance Cloudmark's customers required, continuous investments in capital were needed. The second system was an analysis server for which they wanted a more scalable solution to analyze ever-growing amounts of data.

After speaking with SanDisk®, Ryan decided to give some Fusion ioMemory™ ioDrive® cards a try.

The SanDisk Solution

Benchmark Testing

Before implementing the Fusion ioMemory solution, Ryan ran benchmarks for various load patterns. Following are the results of a benchmark with 8KB records[1] and a MySQL simulation.[2]

8KB Record Benchmark

	2x 160GB ioDrive cards	4x10K SAS (512MB write cache)	24x15K SAS (512MB write cache)	ioDrive card's Improvement over 10K SAS	ioDrive card's Improvement over 15K SAS
MBytes/sec					
Read	823.70	68.00	151.00	1,211.32%	545.50%
Random Read	811.80	10.00	57.00	8,118.00%	1,424.21%
Random Write	523.50	7.00	34.00	7,478.57%	1,539.71%
IOPS					
Read	100,549.32	8,300.78	18,432.62	1,211.32%	545.50%
Random Read	99,096.68	1,220.70	6,958.01	8,118.00%	1,424.21%
Random Write	63,903.81	854.49	4,150.39	7,478.57%	1,539.71%

1. The test compared two 160GB ioDrive cards against a 4-disk 10,000 RPM SAS array and a 24-disk 15,000 RPM SAS array. It is worth noting that bandwidth improved at exactly the same percentage as IOPS. This is expected in a simulation, where the threads and records are of equal size. Also noteworthy: Ryan began a benchmark with 16KB records, but the test was taking so long on the disk array systems that he gave up.
2. This run consisted of 1 thread at 150GB per thread and 16KB records. The benchmark compared two 160GB write-optimized ioDrive cards to a 4-disk 10,000 RPM SAS array.

MySQL Simulation Benchmark

	2x 160GB ioDrive cards	4x10K SAS (512MB write cache)	ioDrive card's Improvement over 10K SAS
MBytes/sec			
Read	308.5	121.7	253%
Random Read	218.8	2.6	8,415%
Random Write	211.4	14.7	1,438%
IOPS			
Read	18,829.3	7,428.0	253%
Random Read	13,354.5	158.7	8,415%
Random Write	12,902.8	897.2	1,438%

Ryan also ran some benchmark tests on Cloudmark's analysis server.[3]

Analysis System Benchmark

	1x 320GB ioDrive	4x10K SAS (512MB write cache)	24x15K SAS (512MB write cache)	ioDrive card's Improvement over 10K SAS	ioDrive card's Improvement over 15K SAS
MBytes/sec					
Read	236.7	44.00	87.00	538%	272%
Random Read	211.2	6.00	41.70	3,520%	506%
Random Write	259.2	4.00	29.00	6,480%	894%
IOPS					
Read	57,788.10	10,742.19	21,240.23	538%	272%
Random Read	51,562.50	1,464.84	10,180.66	3,520%	506%
Random Write	63,281.30	976.56	7,080.08	6,480%	894%

Supercharging Database Replication

The system supporting Cloudmark's back-end consisted of a MySQL master server that replicated to slave (or mirrored) databases at three sites. The slaves at these sites were supported by massive disk arrays taking up a whopping 49U of rack space (see page 6 for detailed system information).

Just two ioDrive cards in each slave server completely eliminated the replication I/O bottleneck, improving Cloudmark's average system reads and writes over five times. It also eliminated 210 spinning disks supporting the slave servers and recovered 27U of rack space.

Needless to say, Ryan was thrilled. "SanDisk has given us the edge in the security/spammer arms race," he said. "The ioDrive card performance improvements improved our replication times immeasurably, effectively shortening the window spammers have to work in before we block them."

Creating Opportunities to Improve Application Performance

The ioDrive card performance improvements had a side-effect—it allowed Cloudmark's messaging security application to fully utilize the CPUs. After adding the ioDrive cards, the application maxed the system's processors. "Once we installed Fusion ioMemory products, we found that we ran into CPU walls before we ran into I/O walls," Ryan said. "I've never had a database system that hit the CPU wall."

The Cloudmark team began digging into ways to more efficiently use its system's processors and found the application was making some inefficient queries whose impact wasn't felt when disk I/O was the bottleneck. For example, one of its main applications had been issuing inserts and updates one query at a time. With a

SanDisk®
a Western Digital brand

3. The test compares one 320GB ioDrive card to a 4-disk 10,000 RPM SAS array and a 24-disk 15,000 RPM SAS array. This run consisted of 32 threads at 256MB per thread and 8KB records.

small amount of effort, fixing this problem produced amazing results. Ryan told us, "Without the I/O bottleneck, we could now update 1,000 or more records per query. This work improved the performance of our data-intensive queries by about five times. We went from our disk-based slaves having constant trouble keeping up, with constant 90-100% disk I/O saturation on the P800 controller with MSA70's, to the SanDisk-based slaves always keeping up at 5% or less I/O utilization."

Improving Analyst Productivity and Problem Turnaround Time

In addition to improving its data replication performance, Cloudmark was also interested in finding a more scalable solution for problem data analysis. This analysis consisted of investigating spam messages that should have been caught as well as authentic messages getting blocked. The investigation often involved searching massive amounts of data, including many large files (500MB-1GB in size), and could take hours.

The ideal solution would improve analyst productivity by reducing the time the system took to read the data from its sources, analyze it, manipulate it, and update the results in the master database to be propagated to customers with the next update.

SanDisk answered the call. The new system consisted of a server with two RAID0-striped 320GB ioDrive cards that provided sufficient capacity to hold the last three to four days worth of data, which covered well over 90% of the problem data cases. Cloudmark relegated its legacy 4TB system to searches of older data.

"People love it because they can do whatever they want, just bash away and it never slows down," Ryan told us. "Our analysts are much more productive and we can get customer fixes for problem data much more quickly. This freed our analysts to improve analysis routines to leverage ioDrive cards to crunch through data nearly five times faster."

Enhancing the Datacenter

In addition to achieving Cloudmark's primary goal of improving data replication and analysis performance while curbing future outlay on disks, it also improved the reliability and availability of the system, reduced its server footprint, and lowered its operational costs.

Improving Availability and Reliability

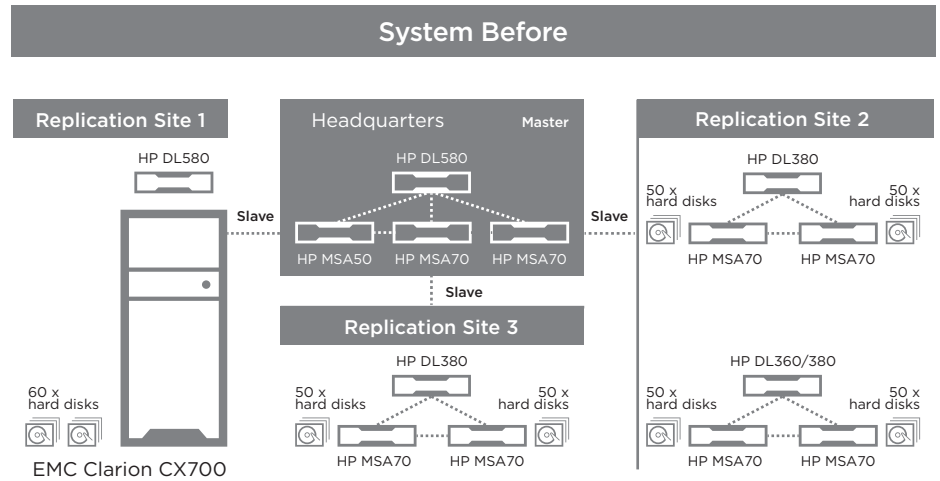Cloudmark implemented full system redundancy, improving both data reliability and availability.

"We RAID0 striped the ioDrive cards to get the capacity we needed, which meant that if we lost the data on one card, we would lose it on both." Ryan told us, "However, we saved so much money on disks that we were able to purchase additional SanDisk powered servers to give us system-level clustering and redundancy."

Lowering Operational Costs

As if these benefits weren't enough, Cloudmark lowered future operating costs on several fronts.
- First, it reclaimed valuable real estate. Even with the addition of full system redundancy, Cloudmark reduced the size of its system by more than half, from 49U to 22U.
- Also, by replacing the spinning disks with ioDrive cards, Cloudmark eliminated 210 mechanical failure points in its system.

Application and Database System

System Before

Replication Site 1

HP DL580

Slave

60 x hard disks

EMC Clarion CX700

Headquarters Master

HP DL580

Slave

HP MSA50   HP MSA70   HP MSA70

Slave

Replication Site 3

HP DL380

50 x hard disks   50 x hard disks

HP MSA70   HP MSA70

Replication Site 2

HP DL380

50 x hard disks   50 x hard disks

HP MSA70   HP MSA70

HP DL360/380

50 x hard disks   50 x hard disks

HP MSA70   HP MSA70

Master at Headquarters (9U)

- 1 HP DL785 G5, 8 x Quad Core, 128GB RAM (7u)
- HP DL580 G5, 4x Quad Core Xeon, 128GB RAM
- OS: RHEL5 x64
- Database Software: MySQL 5.0.x
- Disks:
  - 8x 73GB 10K RPM SFF internal drives
  - 10x 146GB 10K RPM SFF drives, inside and HP MSA 50, connected via P600 controller
  - 50x 73GB 15K RPM SFF drives, inside 2x HP MSA 70s connected via P800 controller

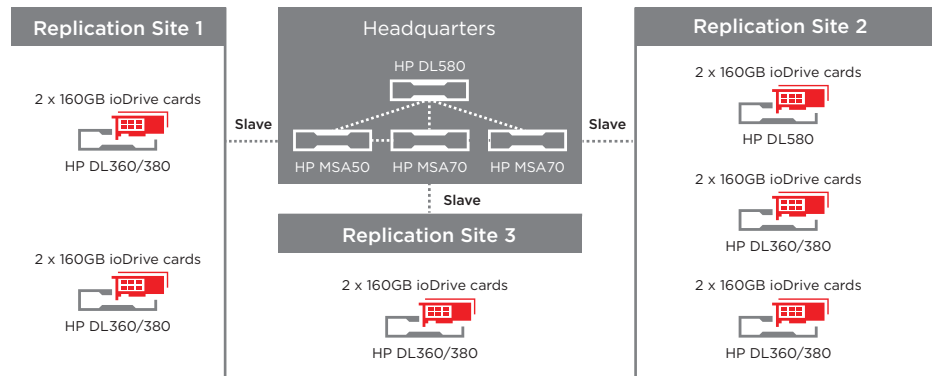### *1 Slave at Replication Site 1 (22U)*

- HP DL580 G5, 4x Quad Core Xeon, 128GB RAM
- OS: RHEL5 x64
- Database Software: MySQL 5.0.x
- Disks:
  - 8x 73GB, 10K RPM SFF internal drives
  - EMC Clariion CX700, connected via QLogic FC HBAs with 60x 36/73GB, 15K RPM FC drives

2 Slaves at Replication Site 2 & 1 at Site 3 (18U total)

- HP DL360/380 G5, 2x Quad Core Xeon, 32GB RAM
- OS: RHEL5 x64
- Database Software: MySQL 5.0.x
- Disks:
  - 6-8x 73GB 15K RPM SFF internal drives
  - Two MSA70s connected via P800 controller with 50x 73GB 15k RPM SFF drives

## Changes to Application and Database System

System After

Replication Site 1

2 x 160GB ioDrive cards

HP DL360/380

2 x 160GB ioDrive cards

HP DL360/380

Headquarters

HP DL580

Slave

HP MSA50 HP MSA70 HP MSA70

Slave

Replication Site 3

2 x 160GB ioDrive cards

HP DL360/380

Replication Site 2

2 x 160GB ioDrive cards

HP DL580

2 x 160GB ioDrive cards

HP DL360/380

2 x 160GB ioDrive cards

HP DL360/380

Replication Sites

- Eliminated EMC Clariion at replication site 1 and added two ioDrive card-equipped HP ProLiant servers.
- Replaced MSA70 servers at replication sites with ioDrive cards

Performance per Rack Unit (Density)

With SanDisk

Without SanDisk
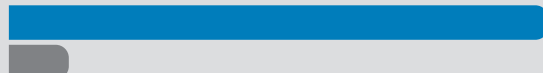
Consolidated servers from 49U to 27U = 1.81 times. Improved average database operations by 5 times.
1.81*5 = 9.05x greater performance density.

9X IMPROVEMENT

| Analyst System Before	Analyst System After

Supermicro Whitebox, 1x Quad Core Xeon E5345, 4GB RAM (3U), 3Ware 9650SE RAID Controller with battery-backed write cache
 	
• Disks:
  – 16x 750GB, 7200K RPM SATA drives
	HP DL380 G5, 1x Quad Core Xeon E5430, 14GB RAM (2U)
	• Disks:
	  – 8x 146GB, 10K RPM SFF internal drives
	  – 2x 320GB MLC ioDrive cards in RAID 0
	• Moved Supermicro to an archive search server

Performance per Rack Unit (Density)

With SanDisk

Without SanDisk
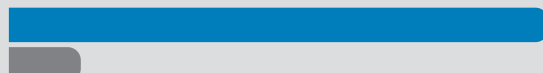
Consolidated servers from 3U to 2U = 1.5 times. Improved analysis routines by 5 times.
1.5*5 = 7.5x greater performance density.

7.5X IMPROVEMENT

SanDisk®
a Western Digital brand
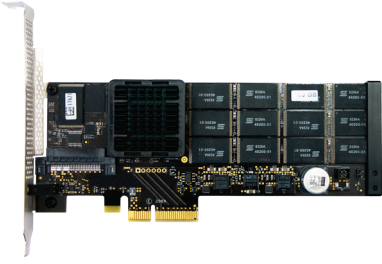
Fusion ioMemory™ - ioDrive®

Summary

SanDisk helped Cloudmark stop the continual outlay on disks previously necessary to provide its customers with industry leading performance. This investment improved its data center as follows:

- 5x improvement in database replication performance
- 5x improvement to data intensive queries
- 10x improvement in analysis server performance
- Enabled 5x improvement to analysis routines
- Simplified maintenance by eliminating at least 210 disk node failure points from its system
- Improved reliability and availability by implementing full system redundancy
- Recovered 27U of rack space
- Greatly lowered power and cooling expenses

Ryan has the following to say about SanDisk: "After about a year and a half of using Fusion ioMemory products, when anyone asks me for a disk array, my first question is 'Why?' There is no reason for a disk array here. It's just a box that I have to support with 25 individual failure points."

While the previous statement is a testament to SanDisk solutions' superiority to disk, the following is the most ringing endorsement: "We spend a lot less time now worrying about disk I/O and our databases keeping up, and we can focus on other things in the business."

About the Customer

Cloudmark provides comprehensive messaging security solutions that protect subscribers and messaging infrastructure from spam, phishing, viruses and other harmful content. Founded by pioneers in messaging anti-abuse, Cloudmark's solutions were architected from the ground up to stop current and emerging threat vectors with high efficiency and scalability.

As a result of its unrivaled accuracy and performance, Cloudmark has emerged as the most widely deployed messaging security solution in the world today. Cloudmark's customers include over 100 of the world's largest service providers, including the majority of tier-1 operators in North America, Japan and Europe.

Contact information

fusion-sales@sandisk.com

Western Digital Technologies, Inc.
951 SanDisk Drive
Milpitas, CA 95035-7933, USA
T: 1-800-578-6007

Western Digital Technologies, Inc. is the seller of record and licensee in the Americas of SanDisk® products.

SanDisk Europe, Middle East, Africa
Unit 100, Airside Business Park
Swords, County Dublin, Ireland
T: 1-800-578-6007

SanDisk Asia Pacific
Suite C, D, E, 23/F, No. 918 Middle
Huahai Road, Jiu Shi Renaissance Building
Shanghai, 20031, P.R. China
T: 1-800-578-6007

For more information, please visit:
www.sandisk.com/enterprise

SanDisk®
a Western Digital brand

At SanDisk, we're expanding the possibilities of data storage. For more than 25 years, SanDisk's ideas have helped transform the industry, delivering next generation storage solutions for consumers and businesses around the globe.