# Understanding IOPS

By Jamon Bowen
Texas Memory Systems


July 2007

# Contents

Section 1

# Introduction

Texas Memory Systems (TMS) takes great pride in the I/O performance of its systems. This document is intended to explain how IOPS (I/Os per second) numbers are achieved and provide some practical advice to companies evaluating disks, RAIDs, and solid state disks.

TMS I/O performance results can be demonstrated at customer and test sites. In other words, they can be produced outside of our lab. It often takes multiple host systems in order to saturate TMS Fibre Channel or InfiniBand links. This aspect of the RamSan allows it to easily service multiple hosts concurrently, with multiple hosts using the same port on a RamSan (over a fabric connection) or with multiple hosts using separate RamSan ports.

Section 2

# Burst vs. Sustained

The numbers for the RamSan-400 are the same for burst rates and sustained rates. We never publish burst only rates. Be wary of hardware vendors that publish burst rates, as these are not sustainable in a realistic high traffic environment. Similarly, many storage systems will publish high IOPS rates "from cache," which cannot reflect real-world application performance.

Section 3

## Random vs. Sequential

The numbers for the RamSan-400 are always based on 100% random performance. Because the RamSan-400 uses DDR RAM as the storage media, sequential and random performance values are almost identical. Be wary of hardware vendors publishing IOPS that are based on sequential reads and writes. These numbers are not generally representative of real-world data traffic.

Section 4

## Read vs. Write

The performance of a storage device, particularly flash disks, can vary dramatically based on whether the accesses that are being serviced are reads or writes. The RamSan is based on DDR memory and can service reads, writes, or a combination of the two at the same rate. The numbers that are reported for the RamSan can be sustained for any combination of read and write accesses.

Section 5

# Multiple LUNs vs. One LUN

Many storage system use dedicated cache on their storage controllers to increase performance. This makes it difficult for the storage to support concurrent access to the same LUN from multiple controllers all the cache has to be kept consistent between all the controllers. Many storage vendors will only allow a LUN to be actively accessed on one port to simplify this problem, and then report their maximum IOPS by presenting a different LUN on each port. The RamSan doesn't use a controller cache as the back end storage is all high speed DDR memory. The brochure numbers for the RamSan-400 are based on accessing a single LUN presented across all of the interface ports on the RamSan. This allows the RamSan-400 to scale performance as more interface cards are added.

Section 6

# Effect of Transfer Size on I/O Performance

The following table provides an example of how I/O performance changes depending on the block size used, and the read/write percentage. These results were gathered with a single, dual ported interface card attached to the RamSan-400. The RamSan can hold up to four of these cards, with performance scaling linearly for each card.

| **Block Size** | **100 % Read** | | **100% Write** | |
|---|---|---|---|---|
| | IOPS | MB/s | IOPS | MB/s |
| 512 bytes | 124k | 60 | 118k | 58 |
| 1k bytes | 120k | 117 | 116k | 113 |
| 2k bytes | 113k | 221 | 116k | 227 |
| 4k bytes | 102k | 398 | 91k | 354 |
| 8k bytes | 76k | 598 | 61k | 470 |
| 16k bytes | 46k | 727 | 37k | 579 |
| 32k bytes | 24k | 759 | 21k | 653 |
| 64k bytes | 12k | 777 | 11k | 697 |
| 128k bytes | 6k | 781 | 6k | 723 |
| 256k bytes | 3k | 784 | 3k | 737 |

A good rule of thumb is that as frame size increases, the number of IOPS decrease and MB/s increases. Therefore you are likely to see the best IOPS performance with small frame sizes and the best bandwidth (MB/s) performance with large frame sizes.

When storage manufacturers design interfaces they tend to optimize the hardware and software for 512 byte transfers, to maximize their advertised IOPS rate. As the table shows, the RamSan-400 is optimized for performance closer to the 4 and 8KB size-- a far more common transfer size in real world applications. It's at this 4 or 8 KB level that performance "jumps" to the most efficient level balancing IOPS and bandwidth. This ensures the best real world performance for TMS customers.

Section 7

# Troubleshooting Performance

Duplicating peak RamSan performance is difficult without multiple, fast host bus adapters; preferably in multiple servers.

It is often the case, especially in older systems, that a single processor is not capable of driving a system to generate sufficient IOPS to saturate the RamSan. This is especially true if multiple HBAs are used in a single system. Additionally, while multi-processor systems improve overall throughput, it does not scale linearly. Therefore, we frequently see multi-processor systems that cannot provide the throughput of an equivalent number of separate host servers. The ability of the RamSan-400 to saturate host servers allows the same RamSan-400 to be effectively used as faster processors become available and servers are upgraded.

Section 8

# Our recommendations

- Use the fastest host bus adapter available.
- Use the fastest servers available.
- Contact TMS support with help on maximizing performance for your particular application. Sometimes a simple setting tweak in a given application can "open the floodgates" and allow full RamSan utilization.