

# **An In-depth Look at the RamSan-500 Cached Flash Solid State Disk**

---

Woody Hutsell  
Texas Memory Systems



March 2008

# Contents

---

|   |           |
|---|-----------|
| <b>Introduction</b> .....   | <b>1</b>  |
| <b>Performance</b> .....  | <b>3</b>  |
| <b>Reliability</b> .....  | <b>5</b>  |
| <b>Availability</b> .....   | <b>7</b>  |
| <b>Manageability</b> .....  | <b>8</b>  |
| <b>Scalability</b> .....  | <b>10</b> |
| <b>Connectivity</b> .....   | <b>11</b> |
| <b>Interoperability</b> .....   | <b>13</b> |
| <b>For More Information</b> .....   | <b>14</b> |
| <br>  |           |
| <b>Figure 1: The RamSan-500</b> .....   | <b>2</b>  |
| <b>Figure 2: Back Connectivity of the RamSan-500</b> .....  | <b>11</b> |
| <b>Figure 3: A Sample of Texas Memory Systems Fibre Channel<br/>Interoperability Partners</b> ..... | <b>13</b> |

## Introduction

The RamSan-500, produced by Texas Memory Systems (TMS) is the world's first cached Flash solid-state disk (SSD) designed from the chip up to meet enterprise requirements for performance, scalability, and reliability. The RamSan-500 is a fusion of RAM and Flash memory. The system takes advantage of the strengths of each technology:

- Blazing fast write performance (RAM)
- Fast reads (Flash)
- High density (Flash)
- Low power consumption (Flash)
- Lower cost per capacity (Flash).

The RamSan-500 benefits from TMS' fastest ever storage backplane (over 4GB/second) and our fastest I/O controllers (Fibre Channel). The system includes brand new caching algorithms optimized to take advantage of our ultra low latency DDR RAM cache (under 15 microseconds) and the massively parallel array of Flash memory with sub 200 microsecond read performance. The RamSan-500 has lower latency and higher peak random I/Os per second than any other Flash SSD solution on the market.

The RamSan-500 is a non-volatile storage system. The flash memory in the system is inherently non-volatile. The RAM memory, used as cache in the system, is backed up by redundant batteries and data in the cache is flushed to Flash if external power fails.

Traditional Flash drives have trivial amounts of cache, as they are solely geared toward low power rugged environments. Without significant cache to permit parallel Flash operations, the write performance of these drives tends to be lower than hard disk drives. Without a large cache, the drives are not isolated from small writes that over time can wear out Flash media. Prior to the introduction of the RamSan-500 there were very few credible Flash RAID solutions on the market. The options that did exist had very small cache and were built with controllers not designed to accommodate the capability of a Flash storage system (low latency reads).

The RamSan-500 meets the needs of customers who have large read-intensive applications that need high I/O and low latency. Based on our many years of experience assisting customers with application performance challenges, we believe the RamSan-500 will be an excellent fit for read intensive OLTP and Data Warehouse applications. The RamSan-500 will offer 10x the performance of most

enterprise RAID solutions at a price that is nearly comparable per capacity. One RamSan-500 will be able to replace a rack of enterprise hard disk drives (HDD).

The RamSan-500 will help customers decrease power consumption. A 2TB RamSan-500 providing 100,000 read I/Os per second will use at least one third less power than most enterprise HDD RAID solutions.

The RamSan-500 has the following key features:

- **The World's Fastest Storage™.** Read I/O rates higher than 100,000 IOPS and sustained, random bandwidth up to 2 GB/second.
- **4Gbit Fibre Channel.** The RamSan-500 comes with new dual ported 4Gbit Fibre Channel controllers for even greater performance and higher throughput.
- **Large RAM Cache.** 16 to 64GB of DDR RAM cache to accelerate writes and cache frequently accessed reads.
- **RAID Protected Flash Modules.** The RamSan-500 includes nine RAID-3 protected Flash modules that are hot swappable. After RAID protection, the RamSan provides 1 or 2TB of usable capacity.
- **Wear Leveling.** The RamSan-500 includes wear-leveling technology that protects the underlying Flash media by spreading out writes and maximizing the write life of the system.
- **Chipkill Technology.** Robust memory parity technology prevents cache data corruption even if an entire memory chip is lost.
- **Completely Non-Volatile.** Flash memory is inherently non-volatile. Additionally, if external power fails, internal redundant batteries power the unit long enough to flush dirty cache blocks from the RAM cache to the internal Flash RAID.
- **Unparalleled Scalability.** A single 4U chassis upgrades to 64GB of cache, 2 TB of Flash storage, and 8 Fibre Channel ports. Multiple units can be added to meet any needed capacity or performance requirement.
- **Low Power.** The RamSan-500 operates with 300 watts of power, less than any comparable HDD RAID system.
- **Interoperable.** The RamSan is "just another disk" to the network, meaning it can be implemented in any way a conventional disk can.



Figure 1: The RamSan-500

## Performance

Performance is key to quality SSD. The RamSan-500 is the latest entry to the TMS World's Fastest Storage™ product line. The RamSan-500 has the following key performance metrics:

- Cache reads/writes:
  - 15 microsecond access time
- Cache miss reads (reads from Flash):
  - 100,000 random IOPS
  - 2GB/second sustained bandwidth
  - 200 microsecond latency

### **Bandwidth**

The RamSan delivers its bandwidth to hosts through up to eight 4-Gb Fibre Channel interface ports. Its high bandwidth provides two main benefits:

1. High bandwidth enables administrators to share the RamSan-500 across multiple hosts without affecting performance.
2. Some applications require high bandwidth, including video on demand and non-linear editing of HDTV.

By comparison, competing SSDs that have less bandwidth than the RamSan-500 cannot support all of the applications the RamSan supports, and cannot be shared with as many hosts.

An important note, as you compare SSD to HDD RAID: our solid state disk can sustain its bandwidth numbers with random data streams. Conversely, HDD RAID systems can only sustain high bandwidth numbers with sequential data streams.

### **I/Os Per Second (IOPS)**

The RamSan-500 can deliver 100,000 random IOPS (this performance is sustained even with cache misses) to your applications. A single port on the RamSan-500 can provide 100,000 random read IOPS and 25,000 random write IOPS. We stress "random IOPS" because HDD RAID manufacturers occasionally cite their IOPS performance, and discerning buyers should be aware that these quoted numbers are almost always sequential IOPS performance. Similarly, if you see Flash SSD performance quoted, beware that many times these numbers only reflect read performance, are often based on cache hits, and are not sustainable without extremely high thread counts to the system.

The problem with sequential IOPS is that almost no "real world" applications actually result in sequential small block disk access. So why do you need IOPS? They are needed to process database transactions. Database transactions have two main characteristics: they are small (averaging around 8K) and they are random. Small random file accesses thrash HDD. In fact, even really good HDD can only provide around 300 random IOPS. Fast HDD RAID can handle 5,000 random IOPS. The fastest cache built in to expensive storage systems can handle at most 150,000 IOPS. This brings us back to the original question – is it important to support a high number of IOPS because your servers can produce a high number of IOPS?

When your processor is faster than your storage, then your processor literally waits on storage for data needed to perform calculations – this is called I/O wait time. If your processor is waiting, then your users are waiting, too. In addition, you are wasting a significant investment in the latest processors and software licenses.

By supplying very high random IOPS, the RamSan-500 eliminates I/O wait time for all of its stored files. Because its random IOPS capabilities are so amazing, a single RamSan-500 can provide I/O acceleration for multiple host servers.

### **Latency and I/O**

Peak application performance is affected by the combination of latency and peak I/O at certain thread counts, block sizes, and read/write patterns. HDD RAID systems typically offer 4-8 millisecond response times. Cached HDD RAID systems offer around 300 microsecond access times for a cache hit and 4-8 millisecond response times for a cache miss. The RamSan-500 offers 15 microsecond access times for cache hits. This is 20 times faster than most HDD RAID systems. The RamSan-500 provides 200 microseconds access times for a read cache miss. This is at least 20 times faster than most HDD RAID systems as well.

For some customers, this combination of low latency and sustainable I/O is the key to improving application performance.

When you consider our combined bandwidth, random IOPS performance, and extremely low latency, it is clear why we call the RamSan "The World's Fastest Storage™".

## Reliability

It's commonly believed that Flash-based SSD is not ready for the enterprise, partly due to issues of data reliability. Until the RamSan-500, this was true. Most Flash SSD is designed for consumer electronics, military applications, and laptops. These drives are designed to be rugged and low power. Write performance and write endurance are not major requirements for these embedded applications. The RamSan-500, on the other hand, is designed from the ground up to store mission critical data in the enterprise data center.

In order to meet enterprise reliability requirements, the RamSan-500 includes many key reliability features:

### **Two Layers of Flash Memory Protection**

#### **Layer 1: ECC for Flash Memory Chips**

A proprietary ECC implementation detects and corrects single bit errors and detects multi-bit errors for each 64 byte chunk of data. This error correction algorithm is far more aggressive than industry standard solutions that provide this level of protection across 512 byte chunks.

#### **Layer 2: RAIDed Flash Memory Modules**

The nine Flash memory modules each provide 128GB or 256GB (available) capacity. After RAID-3 is applied, the system presents either one or two terabytes of usable capacity. In the event a problem is observed with a Flash module the system continues to operate, the problem module can be hot swapped, and the RAID controller will rebuild the Flash array from the remaining modules.

#### **Flash Wear Leveling**

Wear leveling is a common feature in Flash SSD solutions. Wear leveling is required because the underlying Flash memory chips are specified to handle a limited number of writes. The RamSan-500 uses high quality SLC (single layer chip) NAND Flash memory. This memory is specified to handle 10x more programs/erases (writes) than the lower quality MLC (multi layer chip) NAND Flash used in inexpensive consumer devices. Despite the use of SLC Flash memory, it is still important to implement wear leveling strategies in order to prolong the write endurance of the system. The RamSan-500 includes algorithms to maximize its write life and is designed to operate for over four years even with a constant stream of writes.

## **ChipKill**

The RamSan is the only enterprise SSD to implement Chipkill™ technology. Chipkill is implemented for the RAM cache in the RamSan-500. Most SSD implements only limited ECC or other error-protecting code. With these systems, failure of a memory chip often equals corrupted data. Chipkill™ is active on every RamSan-500 cache memory board, protecting data from single bit memory errors, multi-bit memory errors, and even entire chip failures.

## **Soft Error Scrubbing**

The RamSan-500 provides soft error scrubbing for its cache memory. All SSD can correct a single bit error before sending data to the server as a part of its ECC protection. The RamSan-500, however, goes a step further by scrubbing: re-writing the corrected data to memory and then verifying the data to determine if a memory chip has a true failure (a hard error) or if radiation transients caused a soft error. Systems that do not scrub single bit memory errors will either a) report errors to a system log that encourage replacement of a memory board or b) hide these errors, thus leaving a potentially unsafe memory board in the system. Research on this topic suggests that 90% of single bit errors are soft errors. In these cases, the RamSan-500 will correct the error through the scrubbing process and prevent unnecessary downtime to replace the memory board.

## **Battery Backed Cache**

The RamSan-500 offers the largest RAM cache available in a cached Flash solution. This large cache helps isolate the Flash memory from small random writes and provides a working area for the Flash memory RAID controller to optimize and manage its wear leveling algorithm. In order to avoid losing data stored in cache, the RamSan-500 includes redundant batteries that power the system in the event of external power failure. The system runs on these internal batteries long enough to flush all dirty cache lines to the internal Flash RAID, making the RamSan-500 data storage completely non-volatile.



## Section 4

---

### Availability

The RamSan-500 is designed to carry the most vital data for a network or server. This means that it must be more reliable than any other SSD; it must also be highly available for the applications and users that depend on that data. Together, the redundancy and hot swappable features of the RamSan make it the only *true* enterprise-grade Flash SSD on the market.

The RamSan includes redundant, hot swappable power supplies and redundant, hot swappable RAID-protected Flash modules. A failure of one does not affect the RamSan's functionality. SNMP alarms, the browser-based management console, and the system's front panel simultaneously alert administrators to any such failure, and a new component can be swapped in without any disruption of service. To assist in this process, Texas Memory Systems offers a variety of support options, including next day or 4-hour onsite support, advance parts replacement, 24x7 phone support and onsite spare kits for all serviceable components.

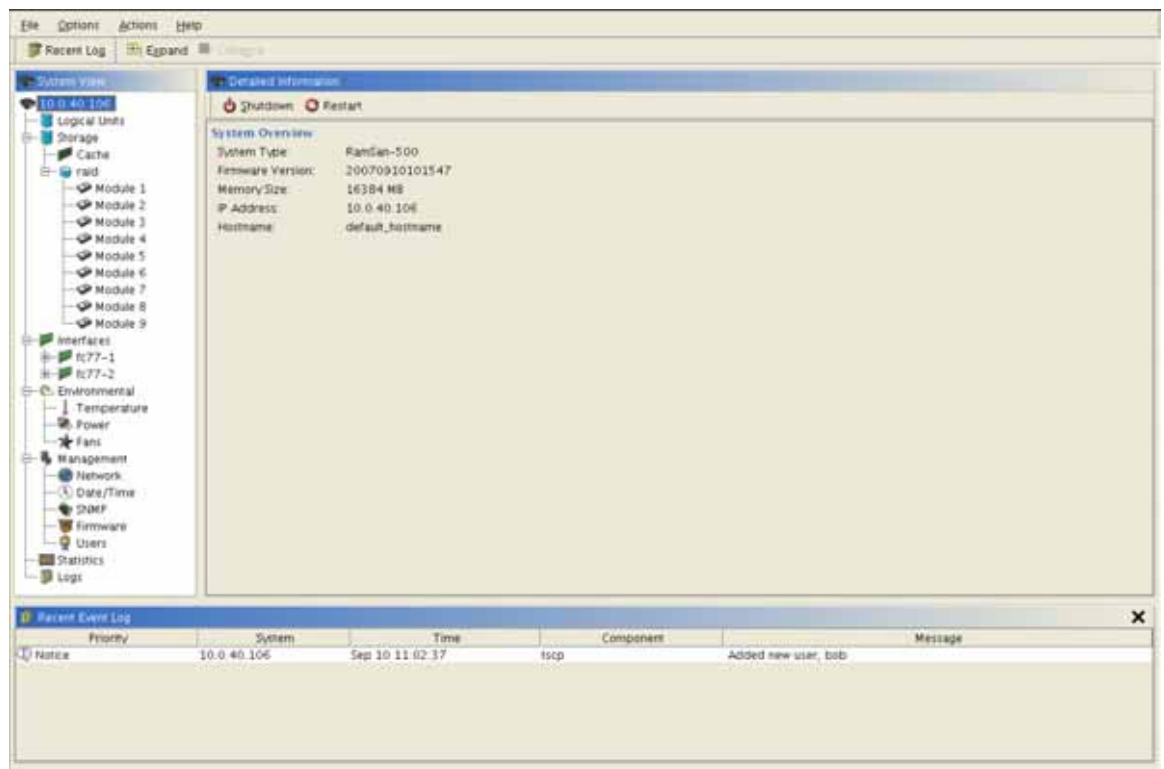
For more information on Texas Memory Systems support options, visit [www.superSSD.com/warranty.htm](http://www.superSSD.com/warranty.htm).

# Section 5

## Manageability

The RamSan-500 includes strong management features to ensure the most effective and efficient use of your investment.

Administrators can access the RamSan's management capabilities over the network or directly attached to the unit. From either a Telnet or a Browser interface (all password protected), administrators can monitor the RamSan's various sensors and status alerts and fully configure the system.



**Figure 1: Screenshot from the Browser Configuration Interface (system information)**

From these interfaces, administrators can

- Configure from 1 to 1024 LUNs of variable capacity. The LUNs can be used by a single server or shared across multiple servers
- Manage LUN masking lists for each interface port. The LUN masking lists can be used to grant universal access or limited access to specific worldwide port names
- Configure Fibre Channel controller settings
- View all SNMP alerts, system temperatures, and the status for every major component
- Specify the IP address for the RamSan
- Configure the RamSan's caching features
- Monitor system health, including the health of the Flash RAID array.

Unlike many other storage appliances, the RamSan-500 also incorporates a front panel display and basic control scheme. The front panel display shows Fibre Channel port activity during normal operation and displays any system-generated warnings. Using the buttons next to the front panel display, the user can configure the RamSan's IP address and initiate a manual shutdown of the unit.

The RamSan-500 is fully compatible with existing SNMP monitors.

## Section 6

---

# Scalability

The RamSan-500 provides 2TB of capacity, and users can choose capacity either 1TB to 2TB. The RamSan-500 can present a single 2TB LUN for application use. Users starting with a 1TB configuration can upgrade storage capacity as their database grows.

In addition to its large Flash storage, the RamSan's RAM cache can be scaled from 16GB to 64GB in 16GB increments. Larger RAM cache will increase read cache, preserve application performance, and increase write performance.

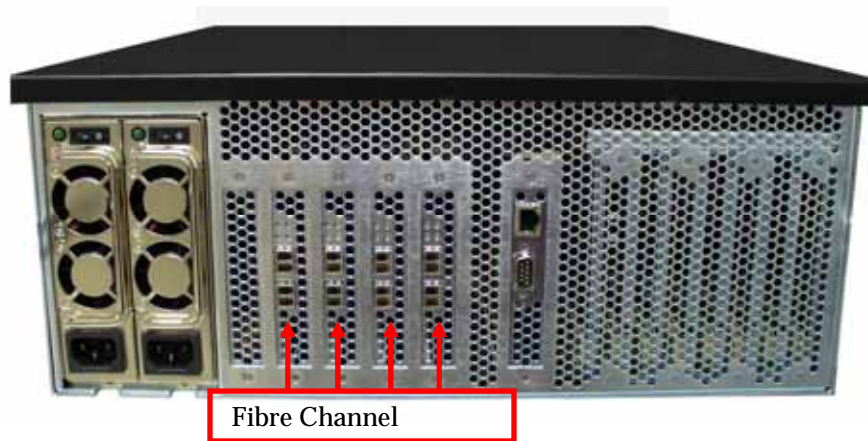
For customers with much larger data sets, multiple RamSan-500 systems can be arrayed. For example, a single rack containing 10 RamSan-500 units can provide 20TB of usable cached Flash storage capacity and sustain 20GB/second of random data transfers.

The RamSan-500 includes slots for four dual ported 4-Gbit Fibre Channel interfaces. Some of our customers use this capability to provide active: active multipathing configurations, while others use additional slots to directly connect to multiple servers. Each controller added to the RamSan-500 increases the performance available to performance-starved applications.

## Connectivity

### Fibre Channel

The RamSan-500 uses Fibre Channel interfaces to connect to up to eight host devices directly and to storage networks for connection to a virtually unlimited number of devices.



**Figure 2: Back Connectivity of the RamSan-500**

Our newest Fibre Channel interface, the FC-77, includes two 4-Gbit Optical (LC) Fibre Channel interface ports. The FC-77 supports both 4- and 2-Gbit devices and auto-negotiates to operate in the proper mode. The RamSan-500 can support up to four FC-77s in a single chassis. All eight ports can be active, enabling LUN sharing across servers and various multipathing and clustering schemes. The FC-77 supports all three of the topologies available under the Fibre Channel protocol:

1. Point-to-point. The RamSan-500 properly implements the point-to-point (n-port to n-port) topology for Fibre Channel. Through the point-to-point topology, the RamSan-500 can be attached directly to up to eight host servers.
2. Switched Fabric. The most popular way to attach a RamSan-500 is with the switched fabric topology. The switched fabric topology implies that there is a storage network switch (such as those from Brocade and Cisco) between the host and the RamSan-500. Switched fabric topologies can be used to provide unparalleled scalability and redundancy for storage networking. Because Fibre Channel switches tend to add very little latency to storage transactions, the switched fabric topology is still an excellent choice for high performance storage.

3. **Arbitrated Loop.** In many ways, arbitrated loop is the legacy topology for Fibre Channel. Arbitrated loop operates similarly to the old token ring local area networks and tends to be somewhat slower than the other topologies. In addition, arbitrated loop can support far fewer devices than the switched fabric topology. This functionality allows the RamSan-500 to be connected directly to host bus adapters and storage devices that work best with arbitrated loop.

## Interoperability

### Fibre Channel

The RamSan-500 is interoperable in virtually any environment and application that supports Fibre Channel. It appears as "just another disk" (or, if desired, up to 1,024 disks) to any connected network or server. This means that the RamSan-500 can go anywhere a Fibre Channel disk can go.

Our customers have deployed RamSans successfully in Microsoft, AIX, HP-UX, TRU-64, Linux, IRIX, OpenVMS, HP Non-Stop, and Solaris environments. Texas Memory Systems partners with a wide variety of companies to ensure continuous interoperability support and testing of Fibre Channel connectivity hardware.

If you have followed TMS over the years, you know that we have invested heavily to ensure that our Fibre Channel interfaces are interoperable with other storage network components.



Figure 3: A Sample of Texas Memory Systems Fibre Channel Interoperability Partners

## For More Information

The RamSan-500 is the fastest Flash-based enterprise storage on the market. It is completely non-volatile and incorporates a suite of reliability and error-correction strategies, from traditional RAID to Chipkill technology. The RamSan-500 offers a faster, less power-hungry, and more reliable alternative to traditional HDD RAID solutions.

Now that SSD at this price, performance, and capacity has arrived, the days of the exclusively HDD-powered enterprise data center are over.

Texas Memory Systems specialists are available to discuss what the RamSan-500 can do in your particular application or environment. Call the main office in Houston, Texas, at 713-266-3200 or do one of the following:

- For more in-depth information, visit [www.superSSD.com](http://www.superSSD.com)
- Existing customers contact [support@superSSD.com](mailto:support@superSSD.com).
- Potential customers contact [sales@superSSD.com](mailto:sales@superSSD.com).