

An In-depth Look at the RamSan-400 Solid State Disk

Woody Hutsell
Texas Memory Systems



August 2006

Contents

Introduction	3
Performance	4
Reliability	6
Availability	8
Manageability	9
Scalability	11
Connectivity	12
Interoperability	14
For More Information	16

Figure 1: The RamSan-400: The World's Fastest Storage®	3
Figure 2: Screenshot from the Browser Configuration Interface	9
Figure 3: Back Connectivity of the RamSan-400	12
Figure 4: A Sample of Texas Memory Systems Fibre Channel Interoperability Partners	14
Figure 5: A Sample of Texas Memory Systems InfiniBand Interoperability Partners	15

Introduction

Accelerate OLTP, batch processing and other intense applications with the RamSan solid state disk from Texas Memory Systems—the only solid state disk available with the performance, software, and features that critical enterprise applications demand. It is 2x-10x faster than cache included in popular RAID storage systems, yet it also contains the most high-availability features in the solid state disk industry and fits the most GB of RAM storage per rack unit. The RamSan is the only solid state disk that allows users two distinct modes of data retention, customizable for each LUN on the unit. Placing hot data on a RamSan allows maximum return on existing critical applications and hardware.

The World's Fastest Storage®. I/O rates higher than 400,000 IOPS and sustained, random bandwidth up to 3 GB/second.

4Gbit Fibre Channel. The RamSan-400 comes with the new FC-400 dual ported fibre channel controller for even greater performance and higher throughput.

The Only Solid State Disk with 4x InfiniBand. The RamSan-400 can come with one to four 4x InfiniBand ports.

The Only Solid State Disk with Active Backup™. The RamSan-400 Active Backup™ mode constantly backs up data from memory to the internal hard disks, helping mission critical data survive even catastrophic problems.

Chipkill technology. Robust memory parity technology prevents data corruption even if an entire memory chip is lost.

Completely Non-Volatile. Data backs up to four redundant hot swap disks in one of two user-configurable methods.

Power Protection. Redundant hot swappable power supplies and redundant internal UPS systems ensure power remains long enough to completely backup data from memory to four hot swap disks if power fails.

Unparalleled Scalability. A single 3U chassis upgrades to 128 GB of RAM storage and up to 8 FC ports. Multiple units can be added to provide any needed capacity or performance requirement.

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-400: The World's Fastest Storage®

Performance

Performance is key to a quality solid state disk. The RamSan-400 is the World's Fastest Storage®. The system delivers up to 3 GB/sec of sustained, random bandwidth and over 400,000 random I/Os per second.

Bandwidth

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

1. High bandwidth enables administrators to share the RamSan-400 across multiple hosts without affecting performance.
2. Some applications require high bandwidth, including video on demand and non-linear editing of HDTV.

Competing solid state disks that have less bandwidth than the RamSan-400 cannot support all of the applications that the RamSan supports and cannot be shared with as many hosts without compromising performance.

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

I/Os Per Second (IOPS)

IOPS are the single best indicator of how well a solid state disk or any other storage device supports database traffic. Solid state disk drive manufacturers like to talk about low latency— the "lag time" in between a drive receiving a request for data and providing that data. In fact, most solid state disk manufacturers claim latency numbers from 10 microseconds to 50 microseconds. The truth of these numbers is borne out in the IOPS. If all latencies are so similar, then performance for solid state disks should be similar. In fact, nothing could be further from the truth. The highest performing solid state disk competitor claims 60,000 random IOPS; a number that is still six times less than the RamSan-400.

The RamSan-400 can deliver 400,000 random IOPS to your applications. A single port on the RamSan-400 can provide 100,000 random IOPS. We stress "random IOPS" because RAID manufacturers occasionally cite their IOPS performance. Discerning buyers should be aware that these quoted numbers are almost always sequential IOPS performance. 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? IOPS 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 hard disk drives. In fact, a really good disk drive can provide around 300 random IOPS. A fast 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 -- it is 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.

By supplying incredibly high random IOPS, the RamSan-400 eliminates I/O wait time for all of its stored files. Because its random IOPS capabilities are so amazing, a single RamSan-400 can provide I/O acceleration for multiple host servers.

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

Reliability

It's a common myth that solid state disks are more "risky" than conventional storage. In truth, a true enterprise-grade solid state disk, such as the RamSan-400, is safer and more reliable than most storage, serves up critical data 24x7x365, and ensures the persistence and recoverability of that data.

The RamSan-400 is a non-volatile solid state disk. Redundant (N+1) batteries power the system in the event of external power failure. Redundant and hot-swap hard disk drives provide persistence in the event of an extended power failure. The RamSan product line is the only solid state disk to offer four N+1 redundant, hot swappable hard disk drives. Most competing solid state disks offer only one backup hard disk per storage subsystem. By building in four hard disks and backing up to them in a parity RAID configuration, Texas Memory Systems ensures that the unit is always non-volatile, even if one of the hard disks fails. Additionally, this multiple hard disk setup means that the RamSan-400 backs up and restores data from those hard disks at the fastest rate in the industry.

Most other solid state disks and "destage-to-disk" cache have only a single method of backing up volatile memory – they copy all data to hard disk on power loss or manual shutdown. The RamSan-400 has two distinct modes to match to particular needs of the application and risk aversion of the administrator. These modes can be administered independently on a system wide or per LUN basis. Every application on the RamSan can receive a custom approach to its persistence needs.

1. Active Backup™ Mode – available only on the Texas Memory Systems RamSan-400. This is the ideal mode for most users. In active backup mode, the RamSan uses available internal bandwidth (i.e. performance that applications are not utilizing) to constantly backup memory to the hard disks. The system automatically optimizes this operation to backup the maximum amount of data at any given time without affecting performance. This means that at any one time, 60-100% of the data already resides on the hard disks. Upon emergency or shutdown, the unit need only finish the backup that is already nearly completed.
2. Data Synch Mode – the conventional method available on most solid state disks. If the unit is shut down or it loses power, internal batteries keep it powered long enough to flush all data to redundant hard disk drives. On power up, the system brings this data back to memory, and normal operation resumes.

ChipKill™

The RamSan is the only enterprise solid state disk to implement Chipkill™ technology from IBM. Most solid state disks implement only limited ECC or other limited error-protecting code. With these systems, failure of a memory chip equals corrupted data. Chipkill™ is active on every RamSan-400 memory board protecting data from single bit memory errors, multi-bit memory errors, and entire chip failure.

Technically, Chipkill™ works hand-in-hand with the RamSan's ECC. Every bit in a memory "word" write is written on a different memory chip. Thus, a memory chip failure causes a single bit correctable error across multiple memory words. The system's built-in ECC easily corrects these single bit errors. Further information on this technology is available at:

<http://www-1.ibm.com/servers/eserver/pseries/campaigns/chipkill.pdf>

Section 4

Availability

The RamSan-400 is designed to carry the most vital data for a network or server. This means that it must be more reliable than any other solid state disk; it must also be 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 solid state disk on the market.

The RamSan is the only solid state disk to incorporate redundant, hot swappable power supplies and redundant, hot swappable RAID-protected backup hard disks. A failure of either of these does not affect the RamSan's functionality. SNMP alarms, the browser-based management console, and the system's front panel simultaneously alert the administrator 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, and onsite spare kits for all serviceable components.

For more information on Texas Memory Systems support options, visit www.superSSD.com/warranty.htm.

Section 5

Manageability

The RamSan-400 includes the most management features of any solid state disk 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), the administrator can monitor the RamSan's various sensors and status alerts and fully configure the system.



Figure 2: Screenshot from the Browser Configuration Interface

From these interfaces, the administrator can

- Configure from 1 to 64 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 or InfiniBand controller settings.
- View all SNMP alerts, system temperatures and the status for every major component.
- Specify independent backup methods for each LUN.
- Specify the IP address for the RamSan.

Unlike many other storage appliances, the RamSan-400 also incorporates a front panel display and basic control scheme. The front panel display shows Fibre Channel or InfiniBand 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-400 is fully compatible with existing SNMP monitors.

Section 6

Scalability

The RamSan-400 provides an industry leading 128GB of capacity in each 3U chassis. Users can scale capacity in the RamSan-400 from 32GB to 128GB in 32GB increments. For the smaller database or for large databases with huge hot files, a single RamSan-400 can present a single 128GB LUN for application use. The user that starts out with a 32GB configuration can expand storage capacity as their database grows. Storage administrators can add memory to the RamSan easily.

An added benefit of having higher capacity per chassis is decreased total cost of ownership for large solid state disk configurations. Total cost of ownership comes from the lower price per capacity available at high capacities in a single chassis and from the lower management costs associated with having fewer discrete systems to manage. For example, Texas Memory Systems' Tera-RamSan (one Terabyte of solid state disk) configuration can be assembled with eight 128GB units and requires only 24U of rack space versus other solutions in the market, which are much more cumbersome. The Fibre Channel controllers are field upgradeable.

Texas Memory Systems also recognizes that as servers, applications and storage networks improve the ability to scale I/O and bandwidth is as crucial to delivering a product that is future proof as the ability to scale capacity. Therefore, the RamSan-400 includes slots for four dual ported 4-Gbit Fibre Channel ports or four 4x InfiniBand ports. Some of our customers use this capability to provide active:passive configurations, while others use additional slots to augment performance. Each controller added to the RamSan-400 linearly increases the performance available to performance-starved applications.

Connectivity

Fibre Channel

The RamSan-400 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.

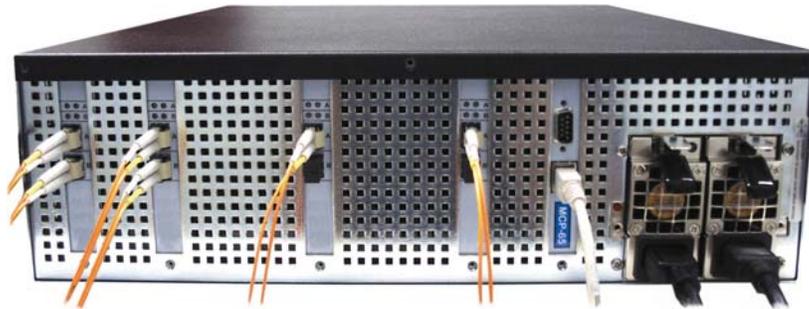


Figure 3: Back Connectivity of the RamSan-400

Our newest Fibre Channel interface, the FC-400, includes two 4-Gbit Optical (LC) Fibre Channel interface ports. The FC-400 supports both 4- and 2-Gbit devices and auto-negotiates to operate in the proper mode. The RamSan-400 can support up to four FC-400s in a single chassis. All eight ports can be active, or it is possible to have pairs of FC-400 controllers in an active:passive configuration. In this configuration, one controller (both ports) is active and the other controller is passive. In the event of a link, port, or controller failure, the other controller takes over. This approach requires a switch. The FC-400 supports all three of the topologies available under the Fibre Channel protocol:

1. Point-to-point. The RamSan-400 properly implements the point-to-point (n-port to n-port) topology for Fibre Channel. Through the point-to-point topology, the RamSan-400 can be attached directly to up to eight host servers. Point-to-point connections are faster than the other methods for attaching storage.
2. Switched Fabric. The most popular way to attach a RamSan-400 is with the switched fabric topology. The switched fabric topology implies that there is a storage network switch (such as those from Brocade and McData) between the host and the RamSan-400. 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 very similar 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-400 to be connected directly to host bus adapters and storage devices that work best with arbitrated loop. The RamSan-400 will also work when connected to a private arbitrated loop hub.

InfiniBand

The RamSan-400 can also support one to four single ported 4x InfiniBand controllers (IB-900). Any single port can provide close to 800 MB/second of random data access or over 70,000 random I/O's per second. Both ports can be active simultaneously resulting in bandwidth of 3 GB/second or over 280,000 random I/O's per second.

The IB-900 controller has been tested with HCAs and with InfiniBand switches.

Section 8

Interoperability

Fibre Channel

Storage hardware is notorious for interoperability problems. Texas Memory Systems built the RamSan to help alleviate this situation. It bucks industry trends by providing interoperability for virtually any environment and application that supports Fibre Channel. It appears as "just another disk" (or, if desired, up to 64 disks) to any connected network or server. This means that the RamSan-400 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, 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 Texas Memory Systems over the years, you will know that we have invested heavily to ensure that our Fibre Channel interfaces are interoperable with other storage network components. In fact, our RamSan-520 was the first solid state disk to pass Brocade's Fabric Aware interoperability certification; the RamSan-210, was the first solid state disk to be both certified Solaris Ready and Designed For Windows 2000; and the RamSan-320 was the first solid state disk to be certified IBM Totalstorage Proven.



Figure 4: A Sample of Texas Memory Systems Fibre Channel Interoperability Partners

InfiniBand

The RamSan has been tested with HCAs from Mellanox and Qlogic. Currently, the IB-900 has only been tested with switches from Qlogic (SilverStorm).



Figure 5: A Sample of Texas Memory Systems InfiniBand Interoperability Partners

For More Information

Texas Memory Systems specialists are available to discuss what the RamSan-400 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
- Existing customers contact support@superSSD.com.
- Potential customers contact sales@superSSD.com.

Also, consider these other articles and white papers, available online or through your Texas Memory Systems representative:

- What the Tera-RamSan can do for Business
- Faster Sybase/Oracle/SQL performance with Solid State Disk
- Understanding IOPS
- The Storage Performance Dilemma
- And many more...ask your Texas Memory Systems rep!