Xiotech®

The New Foundation of Storage

Intelligent Storage Element Technology

For the last several decades the underlying elements of data storage—disk drives, drive enclosures, and controllers—have remained more or less the same, and have been unable to keep pace with dramatically changing business requirements. Xiotech offers a fresh approach with its Intelligent Storage Element technology: A storage foundation that sets new standards for reliability, performance, scalability, and value.

Happy Anniversary!

This year the channel-attached disk drive turns 54 years old (IBM RAMAC in 1954), and the internal-bus-attached disk drive turns 35 years old (Winchester 30/30 in 1973). RAID had its beginnings 30 years ago (Norman Ken Ouchi of IBM was awarded a patent for what would later be termed RAID 5 by Patterson, Gibson, and Katz in 1988), and the SCSI interface celebrates its 22nd year (published as an industry standard by ANSI in 1986). This turns out to be a very long run for what are still the fundamental foundations of today's storage systems.

Since the advent of the Small Computer System Interface (SCSI) and the realization of a Redundant Array of Inexpensive Disks (RAID), data center disk storage has matured to its present form that we know as modular or monolithic disk arrays. These arrays are available with various interconnect protocols and in various capacity points and performance levels. However, the essential elements of data center storage (disk drives, drive enclosures, and controllers) have changed little over the years. In relation to this foundation, neither IT operations nor fundamental element characteristics have changed significantly.

We still perform SCSI I/O to a set of disks fastened in an unintelligent enclosure. We still rebuild disks when they are flagged as failed. We perform RAID as a one-dimensional stripe of data over a given set of disks.

In addition, the fields of architecture and physics teach us that a structure is only as strong as its foundation. This principle, when applied to modern disk arrays, means that we cannot progress much further (outside of capacity increases) unless we completely redesign the foundation to become much faster, more reliable, and more intelligent than it is today. Only in this manner can we progress beyond the decades-old disk storage paradigms in place.

Three Issues with the Current Foundation

1. Fixing What Isn't Broken

The structure of the current foundation forces IT staff and vendor service staff alike to fix what isn't broken—namely, disk drives. According to Seagate Technology, the world's leading provider of disk drives, nearly 75 percent of drives returned to it as "failed" are examined and found to be without failure, a finding commonly called "no trouble found." Of course, one would think that modern array controllers could discriminate between transient failure and true failure—but because of the nature of SCSI I/O and RAID, they cannot.

Nearly 75 percent of drives returned as failed are found to be without failure, or no trouble found.

This forces a "service event"—i.e., several layers of error-prone manual intervention, coupled with significant time and money—to occur. Ironically, this service event is completely unnecessary nearly three-fourths of the time. In terms of both business interruption and financial impact, clearly, the unnecessary service event is a significant issue.

2. Disks Are Too Large... And Too Small

Current disk drives are too large and too small—at the same time. Drives that are too large—e.g., a capacity of 1 terabyte on a single drive—require inordinate amounts of time to rebuild given a failure declaration by an array controller. Testimonials abound on the fact that it may require tens of hours to rebuild a 1 terabyte drive via well-known RAID techniques.

In addition, the prospect of reading 4 or more terabytes just to rebuild 1 terabyte (e.g., in a 5-drive RAID 5 set) begins to push the physical limits of disk drives' ability to read without error, a metric known as unrecoverable read error rate (UER). For example, it is well known that a rebuild of a 1 terabyte drive in a 5-drive RAID 5 set with a drive UER of 1 in 100 trillion (10^{14}) bits will fail 40 percent of the time. A drive that does not rebuild means a potential catastrophic array failure.

At the same time, drives that are too small—e.g., a capacity of 73 gigabytes on a single drive—require dozens or even hundreds of disks to reach performance points necessary for today's transactional applications and other high I/O per second (IOPS)/quick response time (RT) workloads. Even though adequate performance levels may be reached, the prospect of a farm of hundreds of these small drives, forcing a very small mean time between service events (MTBSE) is becoming a poor business proposition. The mean time between failure (MTBF) mathematics is inescapable in today's drives-in-unintelligent-enclosure designs. The more drives you spin, the more failures you will see over time.

It may require tens of hours to rebuild a 1 terabyte drive via well-known RAID techniques.

3. Storage Controllers Are Becoming Bottlenecks

Current storage controllers are quickly becoming bottlenecks to application processing due to the inherent nature of their design. They must perform RAID; manage a static cache; and attempt to manage processes such as drive sparing, drive rebuild, parity placement, diagnostics, Fibre Channel Arbitrated Loop (FC-AL) protocol handling and repair, and so on. Truth be told, there are myriad complex calculations and built-in assumptions in the storage controllers of today.

This leaves little headroom for intelligence that would actually enhance the value of disk storage to modern business organizations. In short, storage controllers have an ever-impossible job to perform—trying to "herd cats," i.e., hundreds to thousands of disk drives that can and do fail in a plethora of ways, expected and unexpected.

Given these fundamental and progress-inhibiting issues with the foundation of disk storage arrays, what can be done to resolve them? Can the decades-old principles of array design be kept in place for more decades without serious redesign? The answer, in short, is no.

ISE: The New Foundation

To overcome these issues, in 2007, Xiotech acquired the Advanced Storage Architecture (ASA) group from Seagate Technology. This group of talented engineers and designers was informally known as the "Skunk Works" in respect to the legendary design group at Lockheed Martin. Like the original Skunk Works, the ASA group designed and built a revolutionary technology—the Intelligent Storage Element (ISE™).

ISE incorporates innovative and disruptive thinking around the foundation of storage—the disk drive and its enclosure. With its game-changing design, ISE delivers reliability, performance, and scalability that no other storage system today can match.

ISE (Fig. 1) is an integrated environment of:

- Single or dual sealed DataPacs, each with up to 8 terabytes of capacity—up to 16 terabytes in just 3U of rack space
- Dual *Managed Reliability Controllers*, which locally manage cache, data recording processes, and more
- Dual power and cooling units
- Dual battery modules

Unprecedented Reliability

ISE reaches reliability levels impossible for standard drives and enclosures, providing more than 100 times the reliability of a regular disk drive enclosed in a typical storage drive bay—significantly reducing service events and their impact on IT organizations.

Because of such reliability, Xiotech provides an industry-exclusive five-year hardware warranty with all its ISE-based Emprise™ systems. This can save organizations thousands of dollars over typical systems that offer only one- to three-year warranties followed by hefty maintenance renewal charges.

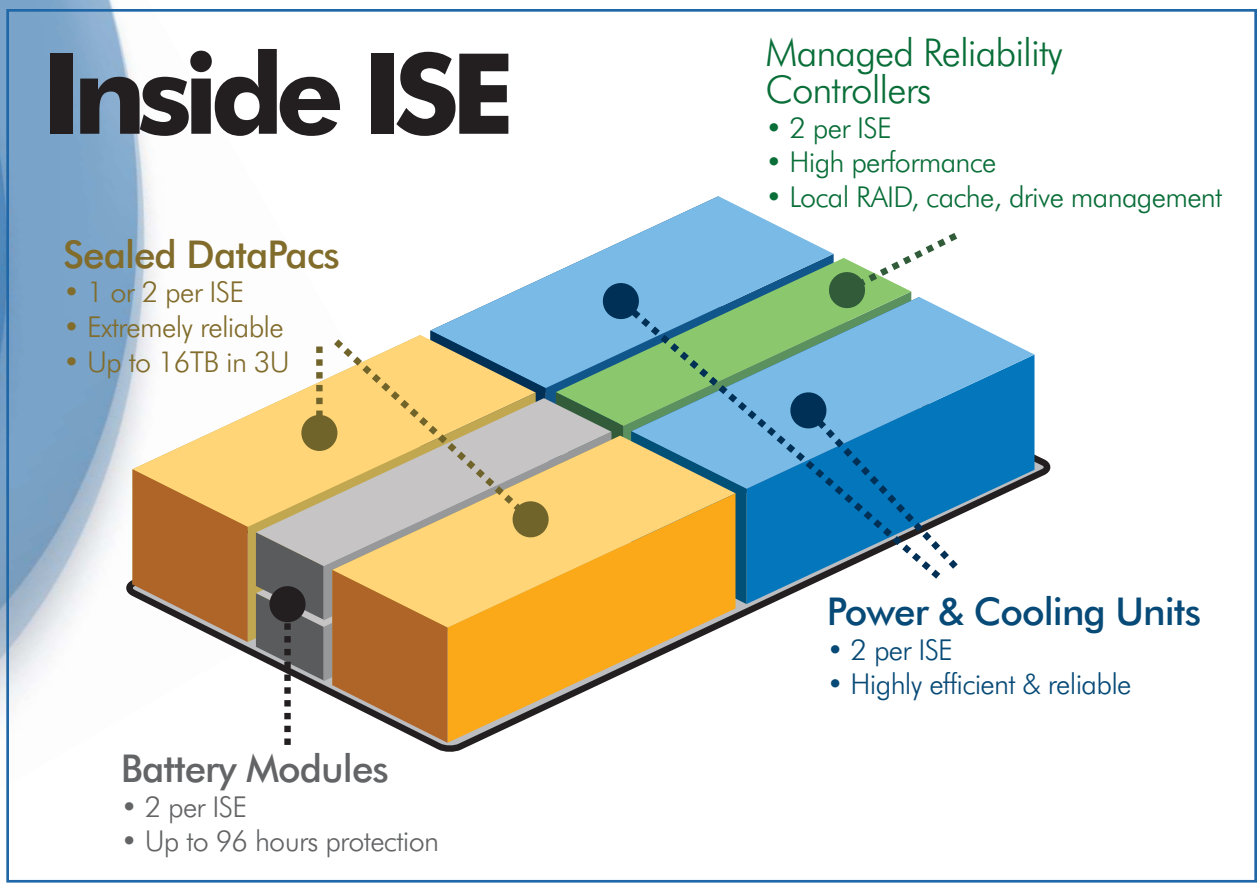

Fig. 1. In just 3U, an ISE module incorporates dual Managed Reliability Controllers; 1 or 2 sealed DataPacs (up to 16 terabytes); and redundant power, cooling, and battery supplies.

How is such reliability achievable?

- **Significantly improved environmentals**. ISE dramatically reduces heat and vibration—the two greatest causes of drive failure.
- **ANSI T10-DIF**. ISE is one of the first storage technologies to incorporate ANSI T10-DIF for end-to-end error correction and detection and the prevention of silent data errors. ANSI T10-DIF provides:
  - » Logical block guard for comparing the actual data written to disk.
  - » Logical block application tag to ensure writing to the correct logical unit.
  - » Logical block reference tag to ensure writing to the correct virtual block.

The result is the equivalent of a fresh, factory-remanufactured drive. Only the components that are irreparable are taken out of service. All others are restored to full activity.

- **Granular recovery**. ISE can recover data down to a level of granularity impossible with traditional RAID controllers that can only "see" an entire drive.
- **Self-Sparing**. Each sealed array includes spare capacity, which overcomes the need for service if there is a failure within the DataPac.

Industry-Leading Performance and Performance Efficiency

Over the decades, storage systems have evolved such that the three key components—drives, drive enclosures, and system controllers—are typically manufactured separately and independently. Each component has a significant "compatibility burden," where it must interoperate across multiple manufacturers and multiple generations of other system components.

This results in significant processing overhead to achieve compatibility, an accumulation of "Band-Aid" fixes as components evolve, and inherent inefficiencies that significantly affect system performance.

ISE delivers industry-leading performance and low price-performance, validated by the SPC.[1]

- **Self-healing technology**. ISE automatically performs preventive and remedial repair of components within its sealed DataPacs—requiring no intervention to pull failed drives.

  When ISE detects a potential drive error, it:
  - Migrates data from the suspect drive to a another drive within the DataPac
  - Resets and/or power cycles the drive
  - Performs a complete factory remanufacturing process
  - Recalibrates heads
  - Rewrites servo tracks
  - Performs a low-level format

Additionally, most systems are designed with centralized, controller-based intelligence that must manage:

- High-level storage processing: Virtualization, FC-AL protocol handling, host connectivity, overall system management.
- Drive management: RAID and cache management, drive rebuilds, drive sparing.

As systems grow, the compatibility burden and centralized management limit overall performance and

reduce the intelligence available for higher-level storage functionality.

ISE overcomes these issues by tightly integrating the drive, enclosure, and controller software and firmware, and by distributing much of the processing and cache into the enclosure.

In addition, unlike most other storage offerings, ISE performance does not degrade as capacity utilization increases. With other systems, organizations often are forced to buy additional capacity (disk drives) simply to meet performance demands. This requirement and expense are eliminated with ISE.

ISE offers linear scale of capacity and performance.

ISE software also incorporates technologies to reduce processing complexity and provide more robust functionality with no loss in reliability:

- Improved data organization for high-speed rebuilds, maintenance operations, and replication (45 percent faster rebuilds than any known array).
- Optimal actuator profiles for each DataPac.
- More efficient striping.
- Less qualifying for bug-free interaction between controllers and groups of disks.

ISE not only delivers impressive performance, but delivers it at a lower cost. The Storage Performance Council (SPC) validated that ISE (Emprise 5000) provided one of the industry's lowest costs for both I/Os per second (IOPS) and megabytes per second (MBps).[1]

Ultimately, this means that organizations can achieve the performance they require with less disk—significantly lowering costs.

Maximum Scalability

In the architecture of storage systems, the key to achieving scalability is to reduce complexity and linearly scale such things as processing and cache to avoid large system bottlenecks. ISE achieves this simplified scalability by pooling storage to create a singular element that is 10 to 20 times the capacity of a typical disk drive, with the internal processing power and cache it needs to handle all drive management functions.

Storage solutions can be created using ISE in a building block fashion, starting as small as one terabyte and growing, as needed, up to a petabyte within a single system. And ISE provides the industry's maximum *usable* storage capacity, so organizations can meet their capacity needs at a lower cost.

ISE can be:

- Part of a direct- or switch-attached system, (Emprise 5000, Fig. 2)—with or without a logical volume manager, such as DataCore SANmelody or IBM SVC
- Used with a network-attached storage (NAS) solution, like Xiotech's Magnitude® Storage Services Gateway
- Utilized in a controller-based storage area network system (Emprise 7000)

[1] *Best protected disk array price/performance for SPC-1 IOPS; industry-leading price/performance for SPC-2 MBps—composite mirroring and large file mirroring—as of May 12, 2008. Audited reports are available at: www.storageperformance.org/results/benchmark_results_spc1#a00064 and www.storageperformance.org/results/benchmark_results_spc2#b00031.*

Fig. 2. ISE-based Emprise 5000 can be used as direct- or switch-attached storage, part of a NAS solution, or the foundation of an Emprise 7000 SAN system.

The environment even can evolve, using an Emprise 5000 system as the foundation of an Emprise 7000 SAN solution.

Finally, the ISE foundation enables tremendous performance scalability. Because ISE includes cache and processing power, each time capacity is added, performance grows—in a linear fashion for Emprise 5000.

This scalability and incremental growth enable organizations to buy only the storage required, manage very large capacities from fewer systems, and grow from very small to very large without forklift upgrades and model family changes.

A Platform for the Future

Computing has changed significantly in the last 20 to 30 years. Adherence to Moore's Law has led to exponential growth in processing power, network throughput, and storage densities. Business practices have driven an ever-increasing demand for storage in terms of capacity, availability, and protection. At the core of it all are the decades-old disk drive, RAID, and SCSI technologies, all wrapped in an accumulated protective layer of complex and expensive fixes to make up for their shortcomings.

ISE is a fresh new approach to the fundamental aspects of storage, setting new standards for reliability, performance, and scalability. Rather than shoring up the foundation that has been with us for decades, ISE is the new foundation that promises to fulfill storage needs well into the future.

Contact Xiotech today to learn more about Emprise storage systems, built on revolutionary ISE technology. Visit www.xiotech.com, email info@xiotech.com, or call toll free 1.866.472.6764.

Important Notice

By accepting, reviewing, or using this document, you, as the recipient, agree to be bound by the terms of this notice. Information in this document is subject to change without notice. Names and data used in examples are fictitious unless otherwise noted. No association with any real company, organization, product, domain name, email address, logo, person, place, or event is intended or should be inferred. Xiotech and/or its licensors may have patents or pending patent applications, trademarks, copyrights, or other intellectual property rights covering the subject matter in this document. The configuration(s) tested or described in this document may not be the only available solution(s). This document is not intended (nor may it be construed) as an endorsement of any product(s) tested, as a determination of product quality or correctness, or as assurance of compliance with any legal, regulatory, or other requirements. This document provides no warranty of any kind, including, but not limited to, any express, statutory, or implied warranties, whether this document is considered alone or in addition to any product warranty (limited warranties for Xiotech products are stated in separate documentation accompanying or relating to each product). No direct or indirect damages or any remedy of any kind shall be recoverable by the recipient or any third party for any claim or loss in any way arising or alleged to arise from or as a result of this document, whether considered alone or in addition to any other claim.

 PN 070601-0709

Corporate Headquarters:
6455 Flying Cloud Drive
Eden Prairie, MN 55344-3305
1.866.472.6764
www.xiotech.com