Solaris™ ZFS™ Enables Hybrid Storage Pools—Shatters Economic and Performance Barriers

Modern servers are a powerhouse of processing. Intel's latest technology innovations, including 6-core Intel® Xeon® 7400-series (formerly code-named Dunnington) processors only highlight a significant gap in the IT ecosystem: multi-socket, multi-core servers have far outpaced the performance limitations of traditional disk drives. The result is expensive and complex architectures—which use massive amounts of expensive DRAM and/or disk drives—designed to maximize CPU utilization. The result is an IT infrastructure that is costly to buy and operate.

Highlights

- Solaris™ ZFS™ seamlessly integrates solid state drives (SSDs) to improve application performance and operating efficiency.
- Architectures based on Hybrid Storage Pools can consume 1/5 the power and 1/3 the cost of standard monolithic storage pools.
- Hybrid Storage Pools deliver unprecedented efficiencies today, using Solaris ZFS, Sun Fire x64 servers, and Intel® SSD based on NAND flash memory.
- Intel® High-Performance SATA SSDs are uniquely engineered to increase performance by 100x and reliability by 2x over mechanical disk drive technology.

Solaris ZFS and Intel® High-Performance Solid State Drives (SSDs) combine to provide breakout innovation—Hybrid Storage Pools. Intel flash-based SSDs provide 100x I/O performance improvement compared to mechanical disk drives. Solaris ZFS can utilize a handful of Intel SSDs as high-speed disk cache, keeping up with modern, multi-core systems, turbocharging application performance. Using Hybrid Storage Pools can provide maximum performance while slashing capital and operating expenses.

The ZFS file system transparently manages data placement, holding copies of frequently used data in fast SSDs while less-frequently used data is stored in slower, less expensive mechanical disks. The application data set can now be completely isolated from slower mechanical disk drives, unlocking new levels of performance and higher ROI. This 'Hybrid Storage Pool' approach provides the benefits of high performance SSDs while still saving money with low cost high capacity disk drives.

The CPU-storage imbalance

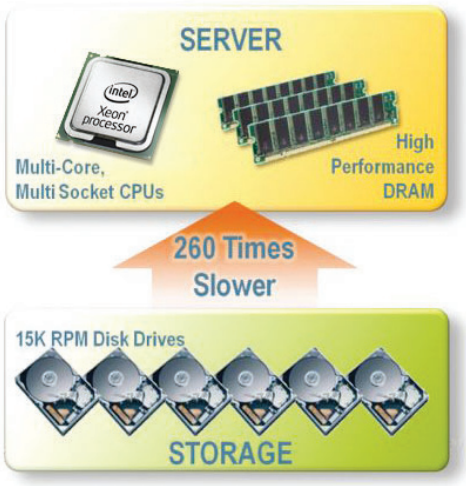

Figure 1: Even the fastest HDDs are 100s of times slower than the processing capabilities of multi-core, multi-socket servers. One solution is keeping the working set in expensive DRAM.

Modern servers are more powerful than ever. Cost-effective 4-socket systems that use 6-core Xeon CPUs, such as the Sun Fire™ X4450 Server, are now available and can process data like never before. Average servers are fast approaching processing capabilities in excess of one million Input/Output Operations Per Second (IOPS). However, mechanical disk drives have failed to keep up with Moore's Law in terms of serving data. Today's fastest drives are capable of 300-400 IOPS.

Why Hybrid Storage Pools?

Hybrid Storage Pools address cost and performance issues. For example, a large online service provider hosts private-branded web-based reservation systems, each branded system serves millions of customers. To meet the end-user response times as outlined in the service level requirement with traditional technologies, each system requires a storage array with 250 SAS 15K drives. Unfortunately, the provider has filled his existing data center—it is out of space, power, and cooling. Hybrid Storage Pools can provide a way to address these restrictions, allowing the service provider a way to expand storage without building a brand new and costly data center.

Addressing this dramatic imbalance is an increasingly expensive proposition. Depending on an application's requirements, large working sets—which typically double in size every two years—can be kept in DRAM; or, IT architects can resort to costly high-end storage systems with hundreds or thousands of 15K RPM drives to deliver data as quickly as possible. Either of these options is expensive to buy and operate, especially with rising electricity and cooling costs.

The Hybrid Storage Solution

Mechanical disk drives are very cost-efficient for storing data, with today's prices approaching $2/GB (bare 15K drives). However, they are relatively slow, offering 300-400 IOPS, and expensive to operate, typically consuming 15-20 watts of power. Solid state drives are fast (as much as 3,300 write IOPS and 35,000 read IOPS) and consume little power (2.5 watts), but are expensive at $30/GB. While SSD performance and operating costs are appealing, clearly it is not cost-effective in every case to substitute SSDs for mechanical drives in a storage array.

At today's prices, SSDs should be viewed not as a replacement for existing storage, but rather as a means to enhance it. Conventional storage systems mix DRAM and hard drives; SSDs are interesting because they fall in a cost and performance sweet spot between those two components. SSDs are significantly cheaper and denser than DRAM and also significantly faster than disk.

Until now, SAS 15K drives represented the dominant building block for storing data—the only way to add more capacity or IOPS was to add more drives. With a Hybrid Storage Pool, storage architects can use the best resource—SSDs, SAS, and SATA drives—to accommodate read and write performance, and capacity, with the most effective component.

Hybrid Storage Pools use a small number of enterprise SSD drives in combination with traditional HDDs to dramatically increase data throughput and CPU utilization. Integrating flash devices into the storage pool as a data cache can result in significant performance gains for I/O intensive workloads. In a Hybrid Storage Pool, SSDs are used to store actively used data, making the entire storage infrastructure appear as fast as flash storage to applications.

ZFS transparently caches data on SSDs, overcoming the need to modify applications. This functionality was introduced in the OpenSolaris 2008.05 OS release, and will continue to be enhanced in future OpenSolaris OS updates. Support is planned to be integrated in a future Solaris 10 update as well.

Solaris ZFS

Solaris ZFS is ready to seamlessly take advantage of SSDs today. Unlike less sophisticated file systems, ZFS recognizes different media types and will optimize how it handles each type to maximize system throughput. For example, ZFS can take advantage of the performance characteristics of high-speed SAS drives when they are present; now ZFS can also leverage SSDs where available, for even more significant performance gains.

ZFS is the native filesystem in the OpenSolaris OS, so increasing system throughput is as easy as adding Intel Enterprise SSDs into the server's disk bays. ZFS is designed to dynamically recognize and add new drives, so SSDs can be configured as a cache disk without dismounting a file system that is in use. Once this is done, ZFS automatically optimizes the file system to use the SSDs as high-speed disks that improve read and write throughput for frequently accessed data, and safely cache data that will ultimately be written out to mechanical disk drives.

The Benefits of Hybrid Storage

A simplified example that focuses on IOPS shows the benefits of Hybrid Storage Pools. If an application environment with a 350 GB working set needs 30,000 IOPS to meet service level agreements, 100 15K RPM HDDs would be needed. If the drives are 300GB drives, consume 17.5 watts, and cost $750* each, this traditional environment provides the IOPS needed, has 30TB capacity, costs $75,000 to buy, and consumes 1.75 kWh of electricity.

Using a Hybrid Storage Pool, six 64 GB SSDs (at $1,000 each) provide the 30,000 IOPS required, and hold the 350GB working set. Lower cost, high-capacity drives can be used to store the rest of the data; 30 1TB 7200 RPM drives, at $689* each ($20,670) and consuming 13 watts, provide cost-effective HDD storage.The savings are dramatic:

- Purchase cost is $26,670, a 64-percent savings
- Electricity consumed is 0.392 kWh, a 77-percent savings

** Sun List Price October, 2008.*

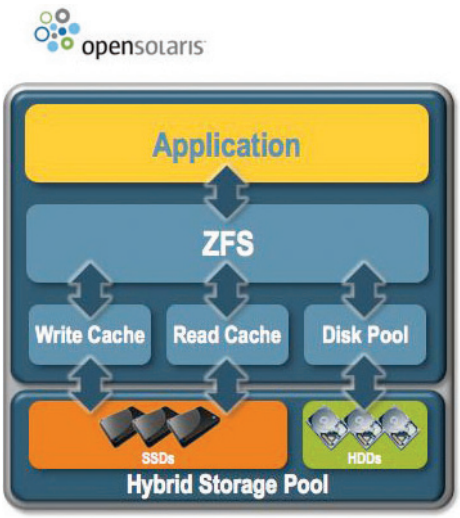

Figure 2: ZFS distinguishes between SSDs and mechanical HDD, placing data where it can be used most effectively.

As shown in Figure 2, ZFS and SSDs combine to form a Hybrid Storage Pool. ZFS automatically:

- Writes new data to the write cache located on an SSD
- Determines data access patterns and stores frequently accessed data on an SSD
- Bundles I/O into sequential writes—all data is written out to mechanical HDDs

Solaris ZFS is dynamically configurable. SSDs can be added without taking down the file system, or modifying the applications using it. Overall, server-integrated SSDs act as a high-performance data cache to existing NAS or SAN, making applications more resilient to SAN/NAS storage latency. Frequently accessed data are distributed to the edge of the computing network, where low latency data delivery offers the highest ROI.

Intel solid state drives

Intel® Solid-State Drives are a quiet, cool storage solution that also offers significantly higher performance than traditional server drives. One Intel® Extreme SATA Solid-State Drive (X25-E) can provide the same IOPS as up to 50 high-RPM hard disk drives (HDDs)—handling the same server workload in less space, with no cooling requirements and lower power consumption. Unlike hard drives that use a motor to spin magnetic media and a read/write head that must move to perform operations, enterprise SSDs contain no moving parts—data are stored on integrated circuits (flash memory) that can withstand significant shock and vibration.

In addition, enterprise SSDs operate in a wider thermal operating range and wider operational vibration range than hard disk drives to deliver significantly higher Mean Time Between Failure (MTBF), approximately 2.0 million hours versus 1.2 million hours.

Intel® High-Performance SATA SSDs deliver higher IOPS and throughput performance than other SSDs on the market today—and drastically outperform traditional hard disk drives. Drawing from decades of memory engineering experience, Intel® SATA SSDs are designed to deliver outstanding performance, featuring the latest-generation native SATA interface with an advanced architecture employing 10 parallel NAND Flash channels equipped the latest generation (50nm) of NAND Flash memory. With powerful Native Command Queuing to enable up to 32 concurrent operations, Intel SATA SSDs deliver the performance needed for multi-core, multi-socket servers while minimizing acquisition and operating costs.

Intel High-Performance SATA SSDs feature sophisticated “wear leveling” algorithms that maximizes SSD lifespan, evening out write activity to avoid flash memory hot spot failures. These Intel drives also feature low write amplification and a unique wear-leveling design for higher reliability, meaning Intel drives not only perform better, they last longer. The result translates to a tangible reduction in your TCO and dramatic improvements to system performance.

Committed to the Open Storage revolution

Data can be the most critical competitive weapon your business has, but its effectiveness is reduced if it is not accessible or costs too much to store. Hybrid Storage Pools are just one highlight of Sun's and Intel's investment in OpenSolaris and the open source community to address this. Innovation from Sun and Intel results in solutions that save time and money for companies seeking to use technology to find opportunity. With an unwavering commitment to technological innovation and Open Storage solutions, Sun and Intel continue to drive storage technology forward. With ZFS and the introduction of enterprise SSD technology, Sun and Intel help enterprises reduce the risk, cost, complexity, and deployment time of multitiered storage environments—all while providing the right data at the right time at the right cost.

Learn More

For a technology brief on Enterprise Flash Technology and more information on how to join the Open Storage revolution, visit sun.com/openstorage, or contact your local Sun sales representative.

For more information on ZFS visit sun.com/solaris/zfs/

For more information on OpenSolaris, visit opensolaris.org or opensolaris.com

More information on Intel Enterprise SSDs can be found at: intel.com/design/flash/nand

Sun Microsystems, Inc. 4150 Network Circle, Santa Clara, CA 95054 USA **Phone** 1-650-960-1300 or 1-800-555-9SUN (9786) **Web** sun.com