

Open Storage Adoption

Sun Systems: Storage
White Paper

Contributing Authors:

Graham Lovell, Taylor Allis, Bruce Norikane, and Chris Ilg
Sun Microsystems
May 2008

Abstract

There are several new business opportunities that require vast amounts of inexpensive storage — and these opportunities cannot be realized with today's traditional, closed storage architectures. Sun believes that open storage products and services will emerge to meet these new opportunities and will represent just under 12 percent of the total storage market by 2011.

Table of Contents

Executive Summary	3
What Is Open Storage?	4
Need for a New Storage Architecture	6
New Digital Data Requirements	6
Web 2.0 Storage Requirements	7
Community Development Growth	8
Eco-Responsible IT Needs Open Storage.....	9
HPC and Open Storage	10
Server Virtualization and Open Storage	12
Open Storage Growth	14
Open Storage Vendor Landscape	16
Sun Microsystems.....	16
IBM.....	17
HP	18
EMC	19
Dell.....	19
NetApp.....	20
Sun Open Storage Case Studies	21
DigiTar	21
Nexenta	22
Sapotek Inc.	23
Gracenote.....	24
Texas Advanced Computing Center (TACC).....	24
University of Oxford	25
Sun Open Storage Offerings	26
Sun Open Storage Servers	26
Sun Open Storage Archive	27
Sun Open Storage HPC	28
Sun Open Storage Software	29
Conclusion	31

Sun believes that open storage products and services will represent just under 12 percent of the total storage market by 2011.

Chapter 1

Executive Summary

Sun is leading the open storage revolution, combining open-source software with industry-standard system components to reduce storage costs by up to 90 percent.

Years ago, servers were closed, proprietary, and expensive. Open systems hit the market and outgrew proprietary computer architectures by meeting the needs of high-scale computing in a much more affordable way. Today's server vendors, in fierce competition, drive rapid technical innovation at market-driven prices.

Using industry-leading server and operating system technology, Sun has opened the closed world of storage systems to innovation and market choice. By delivering industry-standard hardware and releasing a comprehensive suite of storage software to open source via the OpenSolaris™ project, Sun has broken the final barriers to building state-of-the-art storage systems with truly open architectures.

Sun believes that open storage products and services will represent just under 12 percent of the total storage market by 2011. This includes storage hardware, software, and services. Sun also believes that the rapid growth of new digital data demands new, open storage architectures at significantly better economics. Web 2.0, eco-responsible IT, high-performance computing (HPC), and virtualization trends are also creating demand for new storage architectures that can best be satisfied by industry-standard technologies and open-source software.

Open Storage Definition

Storage systems built with an open architecture using industry-standard hardware and software.

Open storage systems are similar to open-systems computers.

In an open architecture, customers can select the best hardware and software components to meet their requirements.

Despite the rise of open systems and open source, closed architectures continue to dominate the storage market.

Chapter 2**What Is Open Storage?**

As a general term, open storage refers to storage systems built with an open architecture using industry-standard hardware and open-source software. In an open architecture, customers can select the best hardware and software components to meet their requirements. For example, a customer who needs network file services can use an open storage filer built from a standard x86 server, disk drives, and OpenSolaris technology at a fraction of the cost of a proprietary NAS appliance, such as a NetApp fabric-attached storage (FAS) system.

Almost all modern disk arrays and NAS are closed systems. Examples include EMC Symmetrix, IBM System Storage DS8300, and HP StorageWorks Enterprise Virtual Arrays (EVAs). All the components of a closed system must come from the vendor. Customers are locked into buying disk drives, controllers, and proprietary software features from a single vendor at premium prices and typically cannot add their own drives or software to improve functionality or reduce the cost of the closed system.

For more than 20 years, storage system vendors have utilized more and more standard components in their products but have not passed along savings to their customers, because the products have remained closed and proprietary. Standard CPUs, memory, and disk drives are used by most storage vendors, but closed, proprietary storage systems can cost up to five times the market price for standard components such as disk drives.

During this decade, open-source software has radically altered the computing landscape. Many new storage systems use Linux or OpenSolaris as their base operating system. Vendors have turned open source into proprietary systems by augmenting basic Linux with their own storage-specific features such as snapshots, remote replication, and volume management. Ironically, most of these systems come to market as closed systems, and customers are not able to add software, substitute disk drives, or modify the vendor's software. Examples include IBM's XIV, EMC's Centera, and NEC's Hydrastor. Despite the rise of open systems and open source, closed architectures continue to dominate the storage market.

Sun Open Storage systems combine open architecture with sophisticated open-source storage software, freeing storage customers from proprietary lock-in. Sun has released a significant volume of storage software to many communities through open-source licensing, in order to enrich their code bases. Sun OpenSolaris open-source contributions include the world's best file system, Solaris ZFS, and two archive systems, the Sun StorageTek™ Archive Manager and Sun StorageTek™ 5800 system. In addition, the Sun StorageTek™ Availability Suite offers robust volume snapshot and remote replication features. NFS and CIFS servers providing NAS server

functionality have also been open sourced. Lastly, the innovative COMSTAR software framework adds a state-of-the-art SCSI target platform that uniquely separates the SCSI protocol handling from the transport protocols (such as FC, iSCSI, SAS, and tape). This facilitates higher performance and flexibility for a wide range of storage devices. These and the many other features in OpenSolaris technology enable end users, OEMs, and developers to build innovative and inexpensive storage systems with very little software development. By participating in the OpenSolaris project, developers can tap the expertise of world-class software engineers.

Google and Amazon would not exist if they hadn't built their own storage infrastructures — traditional storage architectures were simply too expensive and inflexible to accomplish what their online businesses required.

Chapter 3

Need for a New Storage Architecture

There are several new business opportunities that require vast amounts of inexpensive storage — and these opportunities cannot be realized with today's traditional storage architectures. Google and Amazon probably could not exist in their current forms if they hadn't built their own storage infrastructures. Traditional storage architectures built from proprietary products were simply too expensive and inflexible to accomplish the scale and economics demanded by their online business models.

The rapid growth of new digital data demands new storage architectures that offer more flexibility and radically different storage economics. Web 2.0 applications are growing at a tremendous rate and require highly scalable and affordable storage. Industry-standard hardware, open-source software, and community development trends also continue to grow, and they are key enablers to building a new, open storage architecture.

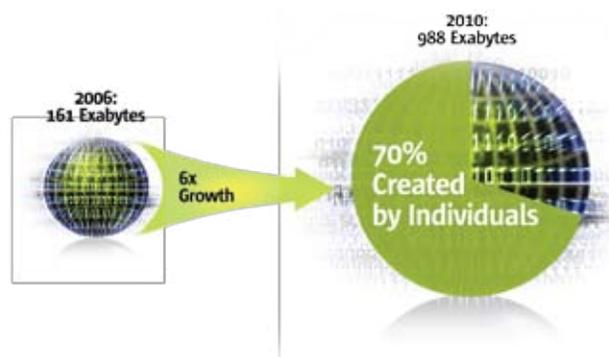
Additionally, there are many market segments and storage trends that are fast growing and can benefit greatly from a new, open storage architecture. Eco-responsible IT efforts can leverage open storage's lower energy consumption, economic, and consolidation advantages. HPC environments are almost exclusively built from open-source software and already utilize open storage architectures to efficiently manage vast storage pools, high I/O bandwidth, and low latency needs. Virtualized server environments can also leverage the flexibility and consolidation advantages of open storage.

IDC's "The Expanding Digital Universe: A Forecast of Worldwide Information Growth Through 2010":

- New digital data is set to grow by six times
- Digital data in emerging economies is growing 30 to 40 percent faster
- 70 percent of this data will be created by individuals
- Enterprises will be responsible for storing 85 percent of this new digital data
- 95 percent of this data will be unstructured

New Digital Data Requirements

In March 2007, IDC published "The Expanding Digital Universe: A Forecast of Worldwide Information Growth Through 2010." The report found that, in 2006, the amount of new digital information created, captured, or replicated was 161 exabytes and is expected to grow sixfold, or to 988 exabytes, in just four years. IDC predicts that 70 percent of this data will be created by individuals or nonenterprises. Just as important, enterprises and organizations will be responsible for storing, securing, and protecting 85 percent of this new digital data.



A good example of this trend is the video-sharing Web site YouTube, now a subsidiary of Google. The company was founded in February 2005, and by May 2006, 50,000 videos were being added every day.¹ At the time of this paper, a YouTube search returns about 72.1 million videos and 2.89 million user channels. This is a perfect example of individuals creating new digital data that an enterprise, YouTube, is responsible for storing and protecting.

This type of new digital data and growth requires a new storage architecture. Traditional architectures are simply too expensive when users attempt to scale to meet this type of storage requirement.

The emergence of Web 2.0 applications is fueling new digital data growth and enabling individuals to publish new content and data at an alarming rate.

Web 2.0 Storage Requirements

The emergence of Web 2.0 applications is fueling new digital data growth and is enabling individuals to publish new content and data at an alarming rate. Common Web 2.0 applications include Weblogs, wikis, podcasts, RSS feeds, mashups, and social-networking sites such as MySpace, Facebook, SmugMug, and LinkedIn. These sites have two things in common: 1) a growing community that is generating new digital content, and 2) personal responsibility for storing and protecting the data these communities create.

By May 2006, 50,000 videos were being added everyday, and in June 2006, 2.5 billion videos were watched on YouTube.

Web 2.0 growth: While the social-networking market is small compared to traditional IT markets, it is set to grow by 153 percent² just in 2008. At the time of this paper, there are currently 2,878 mashup applications on the Internet.³ In March 2008, there were 112.8 million blogs, with 175,000 new blogs being added every day. Bloggers update blog content with 1.6 million posts per day, or more than 18 updates per second.⁴ New data is being generated on a massive scale.

Analyst firm Forrester Research surveyed 2,200 IT decision makers from traditional enterprises, finding that 33 percent were planning on investing in Web 2.0 applications to support internal collaboration goals.⁵

Web 2.0 storage requirements: Web 2.0 storage requirements differ from traditional storage requirements. Given the growth cited previously, massively scalable and lower-cost systems are required. This is best illustrated by the fact that Web 2.0 users are willing to trade high availability in their storage for lower costs. Consider this comment from SmugMug's CEO Don MacAskill about Amazon's three-hour Simple Storage Service (S3) outage. SmugMug is a social photo-sharing service that stores data on Amazon's S3.

I believe there will probably be times where SmugMug is seriously affected, possibly even offline completely, because Amazon (or some other Web services provider) is having problems.... Nobody likes outages, especially not us, but we've decided the tradeoffs are worth it.⁶

1. "YouTube serves up 100 million videos a day online," USA Today, 7/16/06.

2. "Adoption of Digital Marketplace Applications," IDC, 12/07.

3. "The API Scorecard," ProgrammableWeb.com, 2008.

4. Technorati, 2008.

5. "The Web 2.0 Buyer Profile," Forrester Research, 2008.

6. "S3 outage — We weren't affected," Don MacAskill, SmugBlog: Don MacAskill, 1/15/08.

Another Web 2.0 storage requirement is flexibility. Web 2.0 applications are dynamic and can be customized. They depend on open standards and open-source software to reduce costs and give developers the ability to differentiate by adding their own custom software. In summary, the primary storage requirements for Web 2.0 applications are:

- Massive scalability
- Better storage economics
- Flexible and open systems

Open storage meets these requirements better than any other storage infrastructure or architecture available today. One of the most acute needs for open, scalable, and affordable systems is in the Web 2.0 application market.

Community Development Growth

The rapid success of open-source software has ushered in a new era of community development as well. To date, freely available and open-source software has been most successful in the Internet and Web server markets. The acronym LAMP (Linux, Apache, MySQL, Perl/Python/PHP) refers to a commonly used stack of open-source software deployed in Internet architectures. (On February 26, 2008, Sun announced that it closed the acquisition of open-source database provider MySQL.)

Primary storage requirements for Web 2.0 applications:

1. Massive scalability
2. Better storage economics
3. Flexible and open systems

Community innovation: The benefits of community development and open-source software have spread widely. To developers, open source means access to freely available software and the ability to customize software to meet their goals.

To IT users, open source means:

- Lower costs: No software licensing fees; low-cost community support
- Choice: Users do not need to wait for a single vendor to develop new software functionality; they can either find similar functionality already available through the development community or easily help develop the functionality they need through the community
- Innovation: Software and application innovation is no longer limited by a vendor's business goals or research and development budget
- No lock-in: Users are not held hostage by a single software supplier

Customers have also realized that open-source software can help them achieve much more than cost savings: Other benefits include innovation and faster time to market. Customers can participate in the collaboration of thousands of inventors in open-source communities to obtain creative features and functions sooner.

Some open-source users are also willing to pay for the convenience offered by certain open-source distributions, tools, and support in order to best utilize the innovation supplied by development communities. IDC states that worldwide revenue from standalone, open-source software will increase to \$5.8 billion in 2011. This would

To IT users, open source means:

- Lower costs
- Choice
- Innovation
- No lock-in

Sun's OpenSolaris project has more than 96,000 registered members.

More than 60,000 people have viewed or participated in the OpenSolaris storage community initiative.

represent an annual growth rate of 26 percent from 2006 to 2011 — three times the growth rate of commercial software.⁷

Sun Open Storage community: Sun's OpenSolaris project has more than 96,000 registered members.⁸ On April 10, 2007, Sun announced the donation of storage technologies for storage developers within the OpenSolaris project.⁹ More than 60,000 people have viewed or participated in the OpenSolaris storage community initiative, which started with a couple of open-source storage projects and now tracks more than 30 projects. Comprising more than 3,000 members and multiple communities, the OpenSolaris storage community is one of the fastest-growing open-source communities in the world.

In addition, storage industry leaders such as Hitachi Data Systems, QLogic, Emulex, and Brocade have contributed their software to the OpenSolaris project. Other community participants include storage solution vendors such as Nexenta, which leverages OpenSolaris technology and Solaris ZFS to deliver a NAS software solution, and Prominic.net, a hosting provider that delivers storage software appliances.

Conversely, LiveAmmo provides instructor-led training classes and products with a specific focus on computer security and enterprise virtualization. The company uses the COMSTAR software framework to provide customers with a low-cost FC SAN infrastructure.

Companies across the IT landscape are using OpenSolaris technology within their storage product offerings, including Sun; content management companies Drupal, Confluence, and Alfresco; and CRM provider SugarCRM.

Eco-Responsible IT Needs Open Storage

Open storage can also help businesses looking to reduce power and footprint costs. Storage currently accounts for up to 40 percent of overall datacenter energy usage from hardware, according to analyst firm StorageIO Group.¹⁰ Open storage architectures, and open storage servers in particular, not only reduce costs through open-source software and industry-standard components — they reduce power and footprint costs through server and storage consolidation.

Sun also offers the open-source Sun StorageTek Archive Manager (SAM), the only Solaris™ OS-based storage software that enables customers to take advantage of the ecological and economic benefits of tape in a tiered storage architecture. Sun StorageTek Archive Manager can migrate data to Sun StorageTek tape libraries that reduce power and energy costs as well as provide a more affordable tier of data storage.

7. "Has Open Source Sold Out?" Bill Snyder, *InfoWorld*, 1/24/08.

8. <http://www.opensolaris.org/os/community/marketing/metrics/latest/>.

9. <http://www.sun.com/aboutsun/pr/2007-04/sunflash.20070410.1.xml>.

10. "HP Gets Green with New Hardware," Chris Preimesberger, *eWeek*, 6/22/07.

Open storage can also help businesses looking to reduce power and footprint costs.

An example of how customers can also achieve storage savings through open storage is the Sun StorageTek Virtual Tape Library Value system. This system is built on the Sun Fire™ X4500 server and offers a self-contained server/storage device with 24 TB of storage in a 4 U rack.

A competitor to the Sun StorageTek Virtual Tape Library Value system is EMC's Disk Library 210 (DL210), which utilizes the same tape emulation software. EMC leverages the software, a single server, and 48 SATA drives to offer 24 TB of storage in an appliance package. Adding the server and disk drives, the EMC DL210 consumes a total of 1,315 W in a 15 U rack.¹¹

Compare this to the Sun StorageTek Virtual Tape Library Value system, which offers similar capacity and software benefits but only consumes 1,100 W and takes up 4 U of space.

See the following table comparison:

Product	Capacity	Software	Hardware Platform	Power	Rack Size
EMC DL210	24 TB	FalconStor	One server, 48 SATA drives	1,315 W	15 U
Sun StorageTek™ Virtual Tape Library Value System	24 TB	FalconStor	Sun Fire X4500 server	1,100 W (16 percent less)	4 U (73 percent less)

By leveraging an open storage server platform, the Sun StorageTek Virtual Tape Library Value system offers the same functionality as the EMC DL210 — but with 16 percent less power consumption and in 73 percent less space.

By leveraging an open storage server platform, the Sun StorageTek Virtual Tape Library Value system offers the same functionality as the EMC Disk Library 210 (DL210) — but with 16 percent less power consumption and in 73 percent less space.

HPC and Open Storage

According to IDC, the HPC server market crossed the \$10 billion threshold in 2006 and is predicted to exceed \$14 billion in 2011.¹² IDC estimates that HPC storage systems added about \$3.9 billion to the 2006 server revenue total and will undergo faster annual growth than HPC servers. HPC is yet another market that can benefit from new, open storage architectures.

HPC storage requirements: In the May 2007 IDC report “HPC Technical Computing Storage Trends,” 63.8 percent of HPC users surveyed stated they employed direct-attached storage (DAS) and file systems provided on dedicated servers attached to a compute cluster. Maximizing I/O bandwidth and minimizing latency while scaling storage capacity is obviously the top priority for HPC storage providers. This is why DAS and parallel file system architectures are favored. The top three desired data management capabilities of the HPC survey respondents were:

1. Parallel I/O support
2. Tuning and analysis tools
3. Managing data locality to support applications

11. <http://www.emc.com/products/detail/hardware/disk-library-dl210.htm>.

12. “HPC Storage and Data Management,” IDC, 2007.

IDC's "HPC Technical Computing Storage Trends, 2007":

- 18 percent of respondents use the Lustre file system
- 8 percent use the QFS file system

HPC services provider Instrumental, Inc., has collected HPC storage requirements from organizations such as the U.S. National Security Agency, the Department of Energy, and NASA. Instrumental elaborates on the issue of managing data locality:

Data locality is a big issue in some architectures. Sometimes you need to know where data is in memory to get the best performance. Locality issues are compounded by the enormous amount of software 'in the middle (OS, file system, volume management, failover, host bus adapters, and so on)'.

To manage issues such as data locality, an open architecture is needed. The one thing that HPC storage deployments have in common is that they are all custom built. HPC users need direct access to their storage components and software along with the flexibility to swap components and customize software to optimize their storage in order to meet their unique I/O bandwidth and latency needs. This is difficult to impossible to do with closed storage systems.

HPC open storage software: Parallel, shared, or clustered file systems that leverage global namespace technologies are used in most HPC storage environments. In the previous IDC survey, 18 percent of the respondents use the Lustre™ file system and 8 percent use the QFS file system — both are open-source offerings from Sun. HPC customer deployments of the Lustre file system support tens of thousands of nodes, petabytes of data, and billions of files in an object-based cluster. The Lustre file system is currently used in 15 percent of the top 500 supercomputers in the world and in six of the top 10 supercomputers.

Top three desired data management capabilities in HPC:

1. Parallel I/O support
2. Tuning and analysis tools
3. Managing data locality to support applications

An additional storage software requirement in HPC environments is long-term data retention with Hierarchical Storage Management (HSM) software. When the previous IDC survey asked HPC storage users what their general storage requirements were, their third-rated priority was tape storage. To understand why tape is a high priority in HPC storage environments, one need only look at the massive amounts of data that HPC applications generate. The San Diego Supercomputer Center runs several types of HPC applications and cites that earthquake simulations alone generate 47 TB of data per week. By 2011, the Center expects digital archival data to grow to more than 100 PB. HPC centers must leverage the economics of tape to store such massive amounts of data. And just as important as tape is the ability to efficiently move data from disk storage to tape archives. For this, an open HSM software is needed. The Sun StorageTek Archive Manager is available under open-source licensing with community support or as a Sun distribution with full support available. Sun StorageTek Archive Manager offers HPC users policy-based archiving services that automate data management between disk and tape storage systems. Sun StorageTek Archive Manager's striped disk access also enables multiple I/O streams to simultaneously write a file across multiple disks, improving performance.

For the benefits an open storage architecture delivers to HPC customers, see the following Texas Advanced Computing Center (TACC) case study.

Server Virtualization and Open Storage

An emerging market on which open storage can have a significant impact is server virtualization. Server virtualization is a technology that enables multiple applications to be consolidated onto a single server in such a way that each application believes it is running on dedicated hardware. Server virtualization will have a significant impact on storage requirements in general. Virtualization products such as VMware and Sun xVM™ software are enablers to one of the largest trends to hit the server market in decades.

Server virtualization and storage: According to the ESG report “The Impact of Server Virtualization on Storage,” 60 percent of the storage capacity supporting server virtualization is networked today, and this number will move to 74 percent in 2009.¹³ Because server virtualization management software typically enables applications to freely move from one system to the next without interruption, maintaining networked storage links becomes an important requirement. While networked storage can profit from the scalability and economic benefits of open storage, server virtualization will primarily benefit from the flexibility of open storage.

Server virtualization and open storage: Open storage introduces more flexibility and consolidation benefits to the server-virtualization market. This added functionality can be realized in two ways:

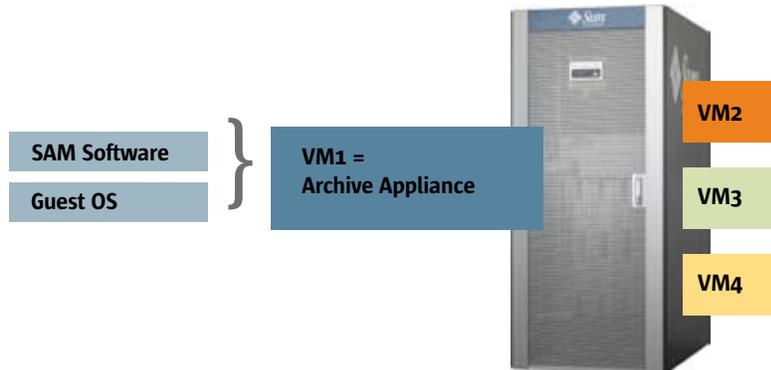
1. By running open storage software inside a virtual machine (VM)
2. By running any vendor’s storage software on an open storage server

In the first scenario, storage users can consolidate servers using offerings such as Sun xVM software or VMware. Each operating system instance on the server is a VM. However, one VM can deploy storage software in order to create a virtual storage appliance inside the server, providing fundamental economic, efficiency, and consolidation benefits. In the following diagram, VM1 is running Sun StorageTek Archive Manager, creating a virtual archive appliance inside the server. Storage users can now consolidate three servers and a storage appliance onto a single server. In a closed architecture appliance, storage software cannot be separated from the storage hardware, making this type of consolidation impossible.

13. “The Impact of Server Virtualization on Storage,” ESG, 12/31/07.

Open storage introduces more flexibility and consolidation benefits to server virtualization:

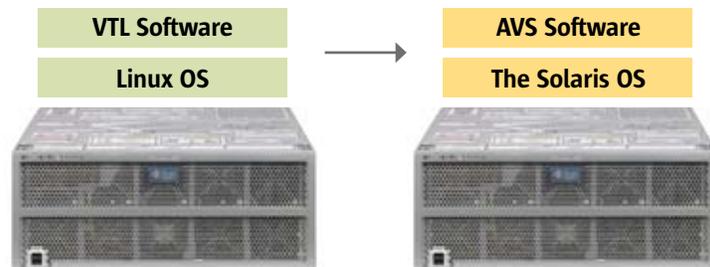
- By running open storage software inside a virtual machine (VM) on a single server
- By running any vendor's storage software on an open storage server



Storage users can now consolidate three servers and a storage appliance onto a single server.

In the second scenario, users can use an open storage server, such as the Sun Fire X4500 server, as a storage target or shared appliance. What's unique about this approach is that users can repurpose their storage appliance as their needs change. For example, customers can repurpose the same Sun Fire X4500 server into a NAS device, a Virtual Tape Library (VTL) appliance, or a data replication appliance without buying more hardware. This gives storage customers unparalleled investment protection.

In the following diagram, a customer has taken a Sun Fire X4500 server running Linux-based VTL software and has repurposed it into a remote replication appliance by leveraging server virtualization and Sun StorageTek Availability Suite software. Server virtualization enables users to utilize multiple software applications supported by different operating systems.



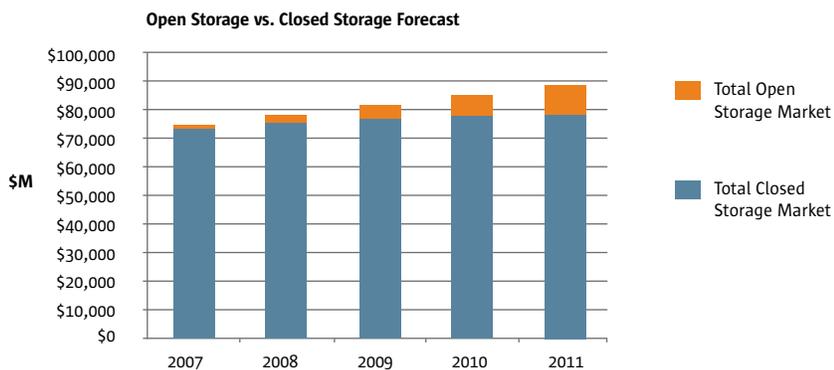
By leveraging open storage and server virtualization, users can realize greater consolidation, efficiency, economic, and reuse benefits than in closed storage systems.

Chapter 4 Open Storage Growth

As current market trends and segments illustrate, the need for an open storage architecture is clear and growing. But how fast will open storage architectures be adopted? What will the size and growth rate of open storage be, especially compared to that of closed storage architectures?

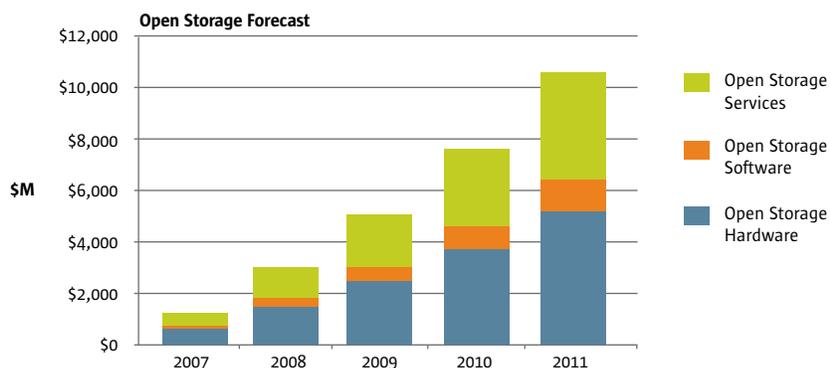
This is a difficult question to answer, as the market is still emerging and few to no analysts have actually attempted to size and scope open storage market adoption. The following data and charts show Sun’s best estimate of the size of the open storage market from 2007 through 2011. This information is based on storage industry forecast data (IDC, 2008) and is adjusted for open storage based on internally developed assumptions from Sun.

Sun believes that open storage products and services will represent just under 12 percent of the total storage market in 2011. With IDC estimating that the total storage market (hardware, software, and services) will generate approximately \$90 billion in 2011, the open storage portion is just over \$10.6 billion. Sun predicts that open storage will represent more than 20 percent of the external disk market by 2011, or approximately U.S. \$5 billion of the U.S. \$24 billion external disk market.



At \$5.2 billion, hardware represents the largest portion of the open storage market.

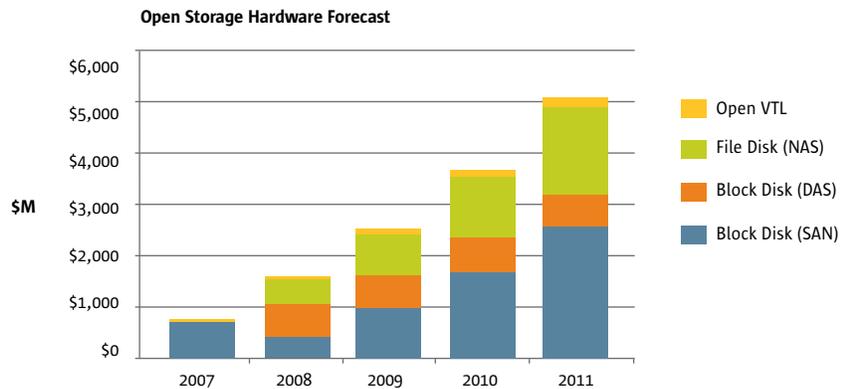
The open-storage market comprises hardware, software, and services segments:



At \$5.2 billion, hardware represents the largest portion of the open storage market. The open storage hardware forecast consists primarily of external disk. Tape storage and storage networking products are not included in the open storage forecast. Sun has further segmented the open storage hardware market into four categories: SAN-based block-level storage, DAS-based block-level storage, NAS-based file-level storage, and VTLs.

Sun believes the earliest adoption and fastest-growing hardware segments for open storage will appear in several different markets including NAS, unified storage (file, iSCSI, and FC storage), and object-based storage.

Sun believes that the earliest adoption and fastest-growing hardware segments for open storage will appear in several different markets including NAS; unified storage such as file, iSCSI, and FC storage; and object-based storage.



The open storage software segment is characterized by a significant discount versus traditional storage software. This discount is representative of the open-source nature of much of this software. Consequently, in open storage, software is a smaller portion of the overall market than it is in the traditional or closed storage market.

Chapter 5

Open Storage Vendor Landscape

Sun Microsystems

There are three areas that differentiate Sun Open Storage offerings from other market vendors’:

1. Innovative systems hardware (servers and storage)
2. OpenSolaris as a storage platform
3. Open-source storage software

Innovative systems hardware: Sun has invested heavily in innovative, efficient, open, and eco-friendly server and storage systems that leverage industry-standard components such as Intel® Xeon® and AMD Opteron™ processors and SAS and SATA disk drives. Sun’s hardware differentiation lies in design innovation. For example, the Sun Fire X4500 server combines a powerful, four-way x64 server with 48 TB of SATA disk in a 4 U rack space, offering the most innovative storage server in the industry with the highest storage density. This enables customers to accomplish more in less space while consuming less power. The Sun Fire X4600 server packs two Intel Xeon Processor-based servers into a compact, 4 U, energy-efficient system. The modular design makes upgrade to future processor technologies simple and nondisruptive. The Sun Blade™ 6000 modular system offers the most open blade platform in the industry — delivering the Solaris OS, Linux, Windows, or VMware running on single and multicore processors by Sun, AMD, and Intel in one chassis.

Sun offers the most dense, efficient, and open hardware platforms in the industry compared to IBM, HP, EMC, Dell, or NetApp.

Three areas that differentiate Sun Open Storage offerings:

1. Innovative systems hardware (servers and storage)
2. OpenSolaris as a storage platform
3. Open-source storage software

OpenSolaris as a storage platform: Sun’s open-source enterprise operating system and file systems continue to be the company’s largest asset and key differentiator from other industry vendors. OpenSolaris is one of the most robust, reliable, and innovative enterprise operating systems in IT, and Sun offers the most advanced open-source file system choices in the world today. Parallel NFS (pNFS), NFS, and Solaris ZFS can manage zettabytes of storage. pNFS can serve thousands of nodes. Solaris ZFS offers data services including volume management, data integrity, and software RAID for the storage industry. Sun’s QFS and Sun StorageTek Archive Manager combine a high-I/O file system with tiered storage-management software. Sun’s Lustre file system is a market leader in HPC storage and can move hundreds of GBps in order to support highly scalable solutions.

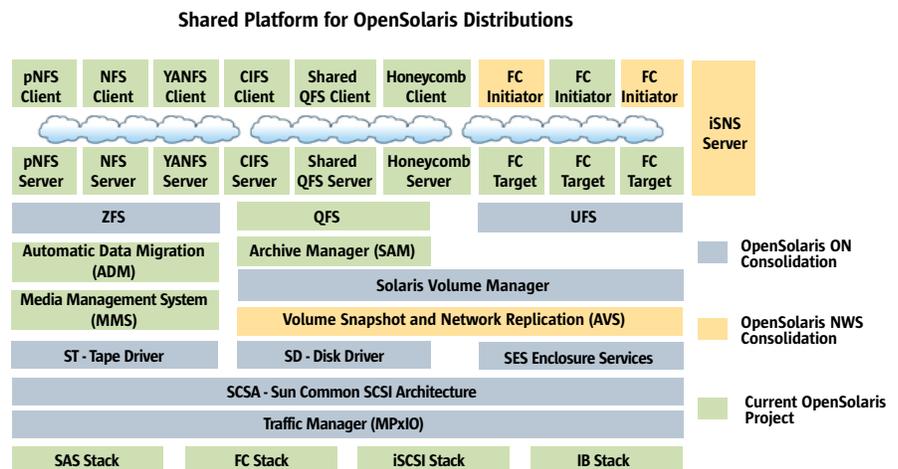
In storage, OpenSolaris technology offers more open and enterprise-class storage features than Windows and higher-level, more robust data services than Linux.

Open-source storage software: Sun has taken an early and clear leadership position in open-sourcing storage application software. Sun has now open-sourced more high-level storage application software than IBM, HP, and all other storage vendors. The last segment of the storage solution stack to be opened is storage software applications. Advanced storage application software such as remote-mirror-copy and point-in-time-copy is traditionally available through storage vendors at costly licensing fees. Sun became the first company to open-source data replication and mirroring applications when it launched the OpenSolaris storage community.

Sun is the only systems vendor to open-source a complete, end-to-end storage software stack, with software such as:

- Traffic management, disk, and tape drivers
- Volume snapshot and replication applications
- Media management and data migration applications
- Volume management and HSM software
- Fixed-content, archive applications
- Storage file systems
- FC, iSCSI, OSD, and object-based targets and initiators

The following diagram shows Sun's extensive list of open-source storage projects:



IBM

IBM sees the value in open source and is a large Linux supporter. However, Sun has more than 3,000 members and 30 open-source storage projects in development. Sun has even open-sourced its key, commercial software applications like the Sun StorageTek Availability Suite.¹⁴

14. <http://opensolaris.org/os/project/avs/>.

IBM's primary IP in storage products remains proprietary and includes the Storage Volume Controller (SVC), DS8000, and XIV NEXTRA, which all use custom components. IBM's largest investment in the storage market has been its recent acquisition of Israeli startup XIV. IBM's XIV NEXTRA does use industry-standard hardware, but its software is proprietary. In terms of industry-standard hardware, IBM sells Intel and AMD servers as well as SAS- and SATA-based disk and JBOD systems. The XIV NEXTRA product is an asymmetric RAIN cluster consisting of scalable interface and data nodes.¹⁵ It does not leverage RAID, as data is distributed across all nodes. IBM has realized that its customers need more than what traditional disk products offer today. The design points of XIV NEXTRA architecture are low cost and massive scalability; however, the technology is new. IBM's claims of low cost are to be determined as well.

IBM also sells Windows Storage Servers, described in more detail below.

HP

HP also sells Intel and AMD processor-based servers. HP is a market leader in industry-standard SAS, SATA, and SCSI JBOD arrays. HP's approach to open storage also involves its industry-standard servers running the Windows Storage Server or Linux operating systems, which can be clustered together for affordable, scalable storage. HP leverages industry-standard servers and a high-volume operating system to reduce storage costs; however, Windows remains a closed operating system, and HP's clustering software is also not open source.

HP is attempting to meet new customer and digital data demand with Microsoft's Windows Storage Server. HP's ProLiant servers are industry-standard servers that leverage the Windows Storage Server operating system. Windows Storage Server is a specialized server operating system built for file and print sharing storage in network attached storage (NAS) or storage area networks (SANs). Features include a distributed file system (DFS), support for SAN and iSCSI, a virtual disk service (VDS) that can manage JBOD or a group of individual storage devices as a single unit, and software RAID. Windows Storage Server is obviously based on a high-volume operating system, but it is an operating system that is proprietary. OpenSolaris is used in enterprise-scale Unix implementations and is open source.

HP announced that it would acquire startup PolyServe, Inc., in February 2007. PolyServe software can consolidate Linux and Microsoft Windows servers and storage into manageable utilities for databases and file serving. HP is using PolyServe software with HP ProLiant servers to offer clustered storage, or what the company calls HP StorageWorks Scalable NAS. Read the "HP Extreme NAS Competitive Edge" based on the recent HP announcement.

15. <http://www.xivstorage.com/>.

EMC

EMC primarily offers closed systems today — custom components and software that are available only through EMC. Sun is providing some of the most open and flexible storage offerings today. EMC does see the customer need for lower-cost storage, and that it will be difficult to compete with free and open storage software. As open storage adoption increases, demand for closed, expensive systems will wane.

EMC's does have the ability to identify business trends and adapt to them. In the past, EMC's revenue was dependent on storage hardware. The company saw that the market would demand more storage software, and through a series of acquisitions, EMC boasts more than 50% of its revenue from storage software today. EMC now sees the need to change the economics of storage. EMC has been investing in acquisitions and products in an attempt to meet this new market demand.

In January 2008, EMC announced its first Storage as a Service (SaaS) or “Cloud” storage platform, EMC Fortress, as well as an online backup service that the company gained through an acquisition. EMC's MozyEnterprise Backup charges users anywhere between \$0.70/GB and \$2.35/GB per month for online backup and storage.¹⁶

More significant is EMC's recent R&D investment in new storage products. At the time of this paper, EMC has not announced any detail on its new storage offerings, code-named “HULK” and “MAUI.” HULK has been reported to be a type of clustered NAS hardware offering. It has also been reported that MAUI is a software offering built on a clustered file system that will provide what EMC calls a “global repository.” EMC does offer Rainfinity global namespace technology, and it has been speculated that this technology will be included in MAUI. HULK and MAUI may be EMC's first ventures into the open-storage space — especially if the systems are able to work with other, third-party, industry-standard components. But the benefits to EMC's new offerings, and just how “open” they are, are yet to be seen.

Dell

Dell has built its business on industry-standard, volume-based products that are easy to configure and order. Dell allows customers to easily configure servers with industry-standard Intel and AMD processors, SATA disk drives, and various Red Hat Linux distributions. Dell also offers industry-standard SAS- and SATA-based JBOD and disk arrays.

Dell's strength in open storage is easy-to-order, configurable, industry-standard hardware. However, Dell has not had as much market-share success and penetration in the enterprise software, services, and support markets.¹⁷

16. <http://www.emc.com/mozyenterprise>.

17. <http://www.idc.com/>.

NetApp

NetApp sells proprietary hardware and develops its own operating system called Data OnTap. NetApp does not open-source its storage operating system. A large part of the industry has been moving to an open or high-volume operating system for storage, like Solaris, Linux, or even Windows. Customers of open-source storage platforms are able to benefit from the innovation and economic benefits that come from the large communities and projects that surround the Solaris and Linux storage platforms.

Chapter 6

Sun Open Storage Case Studies

The following are case studies about Sun customers who are currently using and deploying Sun Open Storage offerings and solutions. Sun Open Storage offerings leverage Sun's innovative systems hardware, OpenSolaris technology, and open-source storage software.

DigiTar¹⁸

DigiTar provides advanced messaging security and processing services over the Internet to customer organizations of all sizes. DigiTar's services include antivirus, antispam, antiphishing, firewall, and archiving. DigiTar's services enhance existing messaging systems with next-generation capabilities such as "DNA-based" spam filtering that uses contextual analysis to block spam with an unprecedented accuracy rate of more than 99 percent.

DigiTar is using OpenSolaris to further improve the performance and efficiency of its database servers. Solaris ZFS automates and simplifies database storage administration for DigiTar, reducing the administration time required for tasks such as identifying and fixing database corruption by days or even weeks. DigiTar participates in the OpenSolaris community to exchange tips and best practices with other users, and credits the community, along with SunSpectrum technical support, for helping it resolve technical issues.

Jason Williams is DigiTar's chief operating officer and chief technology officer. He describes transitioning to Sun Open Storage in his blog post titled *Democratizing Storage*.¹⁹

Mr. Williams blogs that what led DigiTar to implement OpenSolaris was Linux's unreliable SCSI and storage subsystems. He states that I/Os erroring out on their SAN array would be ignored — not retried — by Linux, which would create quiet data corruption. Even though DigiTar was full of "Linux zealots," OpenSolaris was brought into the company because it made for a superior storage platform — benefiting from 15 years in enterprise environments. Mr. Williams writes about OpenSolaris as a storage platform:

That's the really amazing thing about OpenSolaris as a storage platform. It has all of the features of an expensive array and, because it enables you to build reliable storage out of commodity components, you can build the storage architecture you need instead of being held hostage by the one you can afford.

18. <http://www.sun.com/customers/index.xml?c=digitar.xml>.

19. <http://blogs.digitar.com/jjww/index.php?itemid=62>.

Mr. Williams also talks about the flexibility benefits of OpenSolaris and Solaris ZFS:

When you've got rock-solid iSCSI, NFS, and I/O multipathing implementations, as well as a file system [ZFS] that loves cheap disks...and none of it requires licensing...you can suddenly do anything. Need to handle 3,600 noncached IOPs for under \$60K? No problem. Have an existing array but can't justify \$10K for snapshotting? No problem. How 'bout serving line-rate iSCSI with commodity storage and CPUs? No problem.

DigiTar implements OpenSolaris with Sun Fire X4500 storage servers for even more cost savings and breakthrough economics. Mr. Williams mentioned the cost reductions and flexibility DigiTar has been able to realize by deploying OpenSolaris over Sun Fire X4500 servers:

By using [Sun Fire] X4500s [servers], we get the same reliability and redundancy for about 85 percent less cost. That kind of savings means we can deploy 6.8 times more storage for the same price footprint and do all sorts of cool things such as:

- Create multiple data warehouses for data mining spam and malware trends
- Develop and deploy new service features whenever we want without considering storage costs
- Be cost competitive with competitors 10 times our size

Lastly, after realizing the data integrity and economic benefits of OpenSolaris and ZFS in an open storage implementation, Sun's open storage platform is DigiTar's platform of choice:

When it comes to storing data, you'll pry OpenSolaris [and ZFS] out of our cold dead hands. We won't deploy databases on anything else.

Nexenta²⁰

Billing its product as "Enterprise-class data storage for everyone!" Nexenta has built its NexentaOS and NexentaStor software appliance on Sun Open Storage products: OpenSolaris and ZFS. This is significant, as the Nexenta team developed the iSCSI stack that was adopted by the Linux community. Nexenta's team hosts some of the world's experts in storage and open-source software. Certain Linux distributions were limited in their enterprise storage functionality, and OpenSolaris' long history in enterprise environments was ultimately leveraged.

NexentaStor is a software-based NAS and iSCSI solution that boasts unlimited incremental backups or snapshots, snapshot mirroring (replication), and the inherent virtualization, performance, thin provisioning, and ease of use of Solaris ZFS. NexentaStor can be installed and provisioned in under 15 minutes, can be upgraded safely often without a reboot, and includes powerful data search and restore capabilities.

NexentaStor is optimized for use in second-tier NAS and iSCSI applications requiring open, low-cost, high-performance storage as well as dramatically simplified provisioning, expansion, backup, replication, and archiving. Nexenta lists the following NexentaStor benefits:

- Cost: Save 80 percent or more over proprietary legacy solutions. Leverage industry-standard x86/x64 servers and storage hardware. Unmatched price/performance and price/capacity.

20. <http://www.nexenta.com/corp/>.

- Freedom: Simplify storage deployment by running NexentaStor on x86/x64 hardware, server blades, or common virtualization platforms.
- Control: Open-source base, open standards, and community participation enable faster feature integration, better quality assurance, the ability to build your own custom solutions, and an end to legacy vendor lock-in.

NexentaStor enables a wide variety of disk technologies, such as local SCSI, SAS, SATA, and similar, to be directly employed, but also adds the flexibility of iSCSI, Fibre Channel, and newer storage interconnects such as InfiniBand.

NexentaStor derives its economic and innovation benefits by leveraging industry-standard hardware with the NexentaStor appliance, which is based on the OpenSolaris storage platform.

Sapotek Inc.²¹

Sapotek Inc. delivers on-demand solutions such as an online desktop in a software-as-a-service (SaaS) model, reaching approximately 200,000 users worldwide.

To help facilitate continued growth, Sapotek open-sourced its product, and an active free-software community — Sapodesk — is now expanding its capabilities. As of mid-2007, the challenge was to see if the company's infrastructure could scale as fast as the popularity of its service. Previously, Sapotek ran Red Hat Enterprise Linux on Dell servers and had been maxing out at five concurrent threads per server. Sapotek now has migrated to Sun Fire X4200 servers and the Sun Fire X4500 storage server running Solaris ZFS. The company deployed a single Sun Fire X4500 server to gain the highest storage density available, replacing four Dell/EMC storage systems. "That reduced our storage footprint by 75 percent, further lowering hosting costs," said Oscar Mondragon, chief technology officer at Sapotek. The company also uses the snapshot feature in Solaris ZFS and has reduced backup and recovery times by 99 percent — from hours or even days to just minutes. According to Mr. Mondragon:

The ZFS file system feature of the Solaris 10 OS is a marvel. It creates a common storage pool where all storage performs as fast as if it were local. Our administrators can grow, add, or remove storage on the fly in a single step. Just two people administer 24 TB.

21. <http://www.sun.com/customers/servers/sapotek.xml>.

“Sun and TACC are the ones to watch in HPC right now.”

— Addison Snell, Vice President and General Manager, Tabor Research

Gracenote²²

Gracenote is an established mobile technology leader powering mobile music services from the world’s leading handset manufacturers, including Nokia, Samsung, and Sony Ericsson.

To get the performance and reliability it needs, Gracenote has chosen Sun x64 AMD Opteron processor-based servers. To replace several of its existing rackmount servers and avoid purchasing extra external storage, Gracenote has deployed a Sun Fire X4500 server, which combines server functionality and ultradense storage. With Solaris ZFS, it’s also quick and easy to lay out the array groups for storage. Gracenote scales its infrastructure to handle peak, rather than average, traffic. Its two biggest days are “Christmas and New Year’s Day, when people look up a lot of music with their new MP3 players and CDs,” says Matthew Leeds, vice president of operations for Gracenote. “We maintain three collocation facilities, and can sustain the loss of one and still serve our customers. This means we have at least 150 percent of predicted peak capacity online at all times.”

According to Mr. Leeds:

Sun helps us deliver a scalable, reliable service that lets us sleep at night because we know that we can rely on the stability and performance of our Sun products.

Texas Advanced Computing Center (TACC)

Sun Open Storage Used in the World’s Largest Supercomputer

The Texas Advanced Computing Center (TACC) has deployed the largest HPC system in the world for open science research. TACC’s “Ranger” system will be used in computational science and technology research. Ranger went into production on February 4, 2008: 90 percent of the system is dedicated to the TeraGrid, a open scientific discovery infrastructure; 5 percent of the system is allocated to Texas higher-education institutions; and 5 percent of the system is allocated to TACC’s Science & Technology Affiliates for Research (STAR) Program.

Ranger runs 3,936 nodes and 62,976 processing cores. It boasts 123 TB of memory and 504 TFLOPS at peak performance. It uses 1.73 PB of shared disk and 31.4 TB of local disk.²³



22. <http://www.sun.com/customers/servers/gracenote.xml>.

23. <http://www.tacc.utexas.edu/resources/hpcsystems/#ranger>.

“Honeycomb [the Sun StorageTek 5800 system] is uniquely suited to our needs at Oxford. We plan to use it to store large digital objects as well as all the metadata associated with them. It’s like a digital bookshelf.”

— Neil Jeffries, R&D Manager,
University of Oxford

Ranger is built on the Sun Constellation System and incorporates Sun Open Storage servers and software.

For its compute engine data cache, Ranger uses the open-source Lustre file system running across 72 Sun Fire X4500 servers. For long-term data retention and archive, Ranger runs Sun StorageTek Archive Manager over six metadata servers. Leveraging Sun StorageTek Archive Manager and five Sun StorageTek™ SL8500 modular library systems with 48 Sun StorageTek™ T10000 tape drives, Ranger will scale to more than 3.1 PB of online storage and 200 PB of near-line storage.

By leveraging the performance and economic benefits of Sun Open Storage, TACC has built the world’s largest supercomputer.

University of Oxford

Sun Open Storage Used to Archive Nineteenth-Century Works

The University of Oxford’s main library, the Bodleian Library, was established more than 400 years ago and serves as the United Kingdom’s library of record (similar to the U.S.’ Library of Congress). Oxford’s collection includes some of the oldest and most valuable manuscripts in existence.

The University of Oxford’s Electronic Ephemera project is digitizing items from the world-famous John Johnson collection, which consists of nineteenth-century material that was not meant to be kept long term. Content includes entertainment (theater, performing arts) material; journals; crime, murder, and execution documents; and advertising material. The project will digitize more than 65,000 items and more than 150,000 images.²⁴

In order to store this digital archive content, the University of Oxford is deploying the Sun StorageTek 5800 system. The University of Oxford’s R&D manager, Neil Jeffries, explains:²⁵

Honeycomb [the Sun StorageTek 5800 system] is uniquely suited to our needs at Oxford. We plan to use it to store large digital objects as well as all the metadata associated with them. It’s like a digital bookshelf.

By leveraging Sun Open Storage systems, the University of Oxford can now affordably deploy a highly scalable and available digital archive repository to be used by students, faculty, and researchers for years to come.

24. <http://www.bodleian.ox.ac.uk/eejic/>.

25. Customer Snapshot: University of Oxford.

Chapter 7

Sun Open Storage Offerings

The promise of open storage is freedom from vendor lock-in with a global community sharing a passion to make storage better.

Sun's approach to open storage offers enterprise reliability and scalability at one-tenth the cost of closed, proprietary storage. Sun offers the open-source Solaris ZFS, which delivers incredible data integrity and can significantly reduce downtime. Additionally, Sun's innovative use of industry-standard hardware in systems such as the Sun Fire X4500 server delivers two to three times the density, using 50 percent less in power and cooling than competing closed storage.

Sun Open Storage also empowers developers to create services quickly for multiple platforms. Sun's open-source software is open, secure, and freely available. With the Solaris OS, developers, startups, and Web 2.0 companies can quickly develop highly scalable and secure storage services today — reliably, cost-effectively, and on the broadest set of platforms (UltraSPARC® and x64/x86) of any operating system. Unlike the competition, Sun remains active in the community after contributing code, so users can rely on Sun for expert support and service, even on certified third-party hardware.

Sun currently offers the following industry-standard hardware, software, and systems built from standard components and open-source software:

Sun Open Storage Servers

Sun Fire X4500 server: The Sun Fire X4500 server leverages industry-standard hardware and software in a unique package. It is a Dual-Core AMD Opteron™ processor-based server with 48 hot-swappable SATA drives in a single 4 U chassis that can achieve 48 TB of raw capacity with 1 TB SATA drives. It includes 4 GB NICs and 4 GB of RAM, and ships with the Solaris 10 OS. It can run multiple operating systems including OpenSolaris. The Sun Fire X4500 server recently won *InfoWorld's* 2008 Technology of the Year award for best storage server.

Sun industry-standard servers: In the x64 market, Sun offers a complete range of servers from one CPU to eight CPU and with a range of storage capacities. For example, an extensive range of modular blade systems: the Sun Blade™ 6000 and 8000 series, and the Sun Fire™ X4150 server with two CPUs, eight cores, and eight disks in a compact 1 U form factor, and the Sun Fire™ X4600 M2 server with up to eight CPUs in 4 U.

Sun Open Storage Archive

Sun StorageTek 5800 system: The Sun StorageTek 5800 system also leverages industry-standard components. The Sun StorageTek 5800 system is the first integrated, fixed-content archiving system built using open-source software.

In a recent *InfoWorld* product review, the Sun StorageTek 5800 system scored a 9.3 out of 10, with perfect 10s in reliability and scalability.⁵ According to Mario Apicella, the Sun StorageTek 5800 system provides:

Impressive resilience together with excellent performance and powerful administrative tools make “Honeycomb” [the Sun StorageTek 5800 system] one of the most interesting solutions in the emerging fixed-content archiving space. With a foot in the open source community, Honeycomb [the Sun StorageTek 5800 system] promises to deliver more software features faster than competing proprietary solutions, and customers that can’t wait have an easy and free alternative with a flexible SDK.

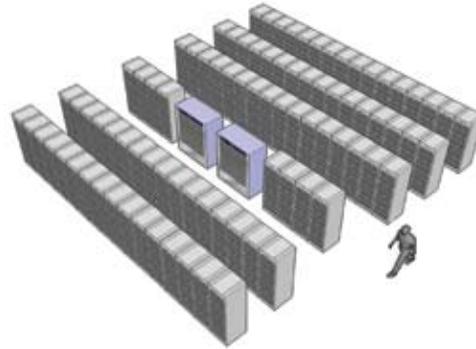
Sun StorageTek 5800 Open Edition: This freely downloadable binary of the Sun StorageTek 5800 system software fully implements a digital archive with fast, searchable content and rich metadata that runs on virtually any x86 hardware device. Like the Sun StorageTek 5800 system, it can store and manage large amounts of fixed content (videos, x-rays, digital books). The OpenSolaris project is focused on client and server implementations of this object-oriented storage system, with traditional Java™ technology and C interfaces to be later expanded by a “StorageBean” Java interface. Large data repository applications access the fixed content through these interfaces, which are designed to manage data collections that can total up to 100 million objects or petabytes of storage.

Another goal of the project is to add the Storage Networking Industry Association’s industry-standard eXtensible Access Method (XAM) specification to both the Sun StorageTek 5800 system and the Solaris OS.

Sun StorageTek Archive Manager and Sun StorageTek QFS software: Sun StorageTek QFS software used with Sun StorageTek Archive Manager provides a shared file system and storage archive management solution for tiered storage solutions in the HPC, data protection, and archive markets.

Sun Open Storage HPC

The Sun Constellation System builds on cost-effective, off-the-shelf components and state-of-the-art technologies to deliver an open, petascale architecture. Using a holistic approach that includes servers, software, storage, and services, Sun has created one of the most powerful HPC platforms in the world.



The Sun Constellation System requires less energy to operate than competitive solutions because of its power and cooling efficiencies. Applications can be created quickly, using open tools and interfaces in small environments, and then rapidly deployed to environments capable of providing up to 1.7 petaFLOPS of computing power.

The Sun Constellation System leverages the following open storage components:

- OpenSolaris technology: Offering key HPC functionality, including performance enhancements, system analysis tools, and high-performance file systems such as Solaris ZFS
- Open-source Lustre file system for unmatched scalability
- Open-source Sun StorageTek QFS software for maximum scalability, data management, and throughput
- Sun Open Storage Sun Fire X4500 servers: Delivering almost .5 PB of storage in a single rack, all accessible from the same IB network

The Sun Constellation System also supports other industry-standard and open-software components and interfaces including:

- Linux
- Intel Xeon and/or AMD Opteron processors
- Sun HPC ClusterTools™ software based on Open MPI
- CLI, IPMI, and SNMP protocols
- Fortress programming language

The Sun Constellation System leverages other industry-leading products from Sun as well, including:

- An ultradense 3,456-port InfiniBand switch — with 3,456 ports, the Sun InfiniBand switch helps reduce cost and overcomes the complexity typically associated with large-scale HPC clusters. (A single InfiniBand switch replaces hundreds of discrete switching elements.)
- UltraSPARC T1 processors with CoolThreads™ technology
- Sun StorageTek™ modular library systems
- Sun StorageTek™ tape drives and libraries for data backup and archive
- Factory integration, services, and support

The first customer install of the Sun Constellation System is the Texas Advanced Commuting Center (TACC) Ranger cluster. With more than 500 teraFLOPS of CPU power, the Ranger cluster is one of the most powerful general-purpose computing platforms in the world. (See Sun Open Storage Case Studies: TACC.)

Sun Open Storage Software

OpenSolaris technology: OpenSolaris technology is the cornerstone of Sun Open Storage offerings and provides a solid foundation as an open storage platform. The origin of OpenSolaris technology, the Solaris OS, has been in continuous production since September 4, 1991. OpenSolaris technology offers the most complete open-source storage software stack in the industry. Below is a list of current and planned offerings:

- At the storage protocol layer, OpenSolaris technology provides SCSI, iSCSI, iSNS, FC, FCoE, InfiniBand software, RDMA, OSD, SES, and SAS
- At the storage presentation layer, OpenSolaris technology offers Solaris ZFS, UFS, SVM, NFS, Parallel NFS, CIFS, MPxIO, Shared QFS, FUSE, and the Sun StorageTek 5800 system
- At the storage application layer, OpenSolaris technology offers MySQL™ software, Postgres, BerkeleyDB, AVS, SAM-FS, Amanda, and Filebench

OpenSolaris technology provides an end-to-end storage platform and includes these essential features:

Solaris ZFS

Another cornerstone of Sun's open storage platform is the Solaris ZFS file system. Solaris ZFS can address 256 quadrillion zettabytes of storage and handle a maximum file size of 16 exabytes. Solaris ZFS deploys several storage services including snapshots, point-in-time-copy, volume management, administration, and data integrity features such as copy-on-write and RAID.

Vendors of closed storage appliances typically charge customers extra software licensing fees for data management services such as administration, replication,

and volume management. The Solaris OS with Solaris ZFS moves this functionality to the operating system, simplifying storage management and eliminating layers in the storage stack. In doing this, Solaris ZFS changes the economics of storage. A closed and expensive storage system can now be replaced by a storage server running Solaris ZFS, or a server running Solaris ZFS attached to JBOD.

Solaris ZFS recently won *InfoWorld's* 2008 Technology of the Year award for best file system. In the *InfoWorld* evaluation, the reviewer stated, "Soon after I started working with [Solaris] ZFS (Zettabyte File System), one thing became clear: The file system of the next 10 years will either be [Solaris] ZFS or something extremely similar."²⁶

Solaris DTrace

Solaris DTrace provides an advanced tracing framework and language that enables users to ask arbitrary diagnostic questions of the storage subsystem, such as "Which user is generating which I/O load?" and "Is the storage subsystem data block size optimized for the application that is using it?" These queries place minimal load on the system and can be used to resolve support issues and increase system efficiency with very little analytical effort.

Solaris Fault Management Architecture

Solaris Fault Management Architecture provides automatic monitoring and diagnosis of I/O subsystems and hardware faults and facilitates a simpler and more effective end-to-end experience for system administrators, reducing cost of ownership. This is achieved by isolating and disabling faulty components and then continuing the provision of service through reconfiguration of redundant paths to data, even before an administrator knows there is a problem. The Solaris OS' reconfiguration agents are integrated with other Solaris OS features such as Solaris Zones and Solaris Resource Manager, which provide a consistent administrative experience and are transparent to applications.

Sun StorageTek Availability Suite

Sun StorageTek Availability Suite software delivers open-source remote-mirror-copy and point-in-time-copy applications as well as a collection of supporting software and utilities. The remote-mirror-copy and point-in-time-copy software enable volumes and/or their snapshots to be replicated between physically separated servers. Replicated volumes can be used for tape and disk backup, off-host data processing, disaster recovery solutions, content distribution, and other volume-based processing tasks.

Lustre file system

Lustre is Sun's open-source shared disk file system that is generally used for large-scale cluster computing. The Lustre file system is currently used in 15 percent of the top 500 supercomputers in the world, and six of the top 10 supercomputers. Lustre currently supports tens of thousands of nodes, petabytes of data, and billions of files. Development is underway to support one million nodes and trillions of files.

26. "Sun ZFS breaks all the rules," Paul Venezia, *InfoWorld*, 6/7/07. [ord](#).

Chapter 8

Conclusion

Open storage leverages industry-standard components and open software to build highly scalable, reliable, and affordable enterprise storage systems. Today's digital data, Internet applications, and emerging IT markets require new storage architectures that are more open and flexible, and that offer better IT economics.

Open storage architectures are competing with traditional storage architectures in the IT market, especially in Web 2.0 deployments and increasingly in other, more traditional storage markets. Open storage architectures won't completely replace closed architectures in the near term, but the storage architecture mix in IT datacenters will change over time.

Sun estimates that open storage architectures will make up just under 12 percent of the market by 2011, fueled by the industry's need for more scalable and economic storage.

Sun offers services that assist IT users with the opportunity to start evaluating open storage architecture strategies and Sun's offerings today — evaluating how an open storage architecture can better support their organizations' business and budgetary goals.

Sun is investing future resources in developing the most comprehensive set of products, systems, solutions, and services for customers who wish to deploy today's and tomorrow's data storage needs.

