An Oracle Technical White Paper
December 2013; v1.1

Why Oracle ZFS Storage Appliance Optimizes Storage in Virtualized Environments

ORACLE

Introduction

Virtualization technology has helped organizations to dramatically reduce data center costs while at the same time improving the efficiency, flexibility, and utilization of existing assets. Virtualization's innovative management of computing resources permits easy consolidation of different workloads from hundreds or even thousands of physical servers into fewer physical servers and virtual machines. By providing an easy way to manage and deploy different operating systems and applications, virtualization technology has proven its flexibility and scalability, while also meeting the special demands of critical environments.

However, even with the technology running on fast computers, organizations still face serious performance issues related to storage devices that can penalize the virtualization environments, impacting both applications and databases. The architecture and features of Oracle ZFS Storage Appliance, including its integrated intelligent caching technology, improves the I/O performance for virtualized environments and simplifies storage management while reducing its cost.

This paper describes the most important features of Oracle's Sun ZFS Storage Appliance and how those features optimize storage in a virtualized environment.

NOTE: References to Sun ZFS Storage Appliance, Sun ZFS Storage 7000, and ZFS Storage Appliance all refer to the same family of Oracle ZFS Storage Appliance products. Some cited documentation or screen code may still carry these legacy naming conventions.

About Virtualization

The concept of virtualization was developed in the mid-1960s as a technique to optimize the utilization of expensive computing resources. Since then, this technology has been growing and changing the data center infrastructure, enabling server consolidation and new methods of application deployment. Before virtualization technologies, organizations used to work with an inflexible and expensive server infrastructure known as "one server one application," which dedicated a physical server to a single operating system. Partitioning or resource sharing of memory, CPU, disks or networks was unheard of. Today, virtualization has helped IT departments reduce the data center costs by lowering the amount of physical servers, cabling, energy and cooling required, while at the same time improving the efficiency, availability and flexibility of computing resources.

Though there are hundreds of different definitions as well as references for the term virtualization, a simple definition is a methodology that divides and shares the physical resources of a computer into multiple environments or instances, thus creating an abstraction layer between the hardware (physical layer) and software (logical layer). With virtualization, you can easily share and divide the computing resources of a single physical server - such as CPU, memory, and disks as well as multiprotocols such as Fibre Channel, iSCSI, NFS, and InfiniBand - with hundreds or even thousands of virtual machines (see Figure 1).

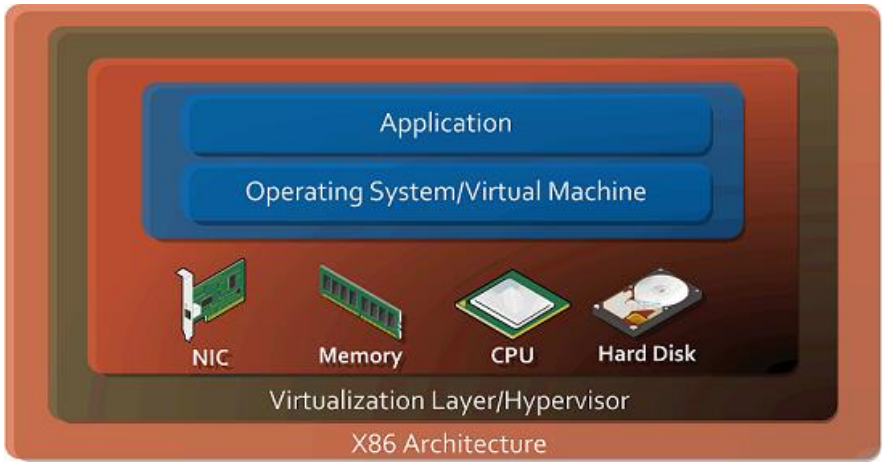

Figure 1. Virtualization architecture

About Oracle ZFS Storage Appliance

This section provides technical information about Oracle ZFS Storage Appliance architecture as well as its concept and terminologies.

Architecture Overview

The basic architectural features of Oracle ZFS Storage Appliance are designed to provide high performance, flexibility and scalability. The Oracle ZFS Storage Appliance provides multiple connectivity protocols, including Network File System (NFS), Common Internet File System (CIFS),

Internet Small Computer System Interface (iSCSI), InfiniBand (IB), and Fibre Channel (FC), for data access and also supports the Network Data Management Protocol (NDMP) for backing up and restoring data. The Oracle ZFS Storage Appliance architecture also offers the Hybrid Storage Pool feature, in which memory, flash and physical disks are integrated for efficient data placement (see Figure 2). A powerful monitoring tool called DTrace Analytics provides details about the performance of the various components, including network, storage, file systems, and client access. The tool also offers numerous drill-down options that allow the administrator to monitor specific rates of latency, size of transfer, and utilization of resources. Finally, the Oracle ZFS Storage Appliance offers a variety of RAID protections to balance the capacity, protection, and performance requirements of the applications, databases and virtualized environments.

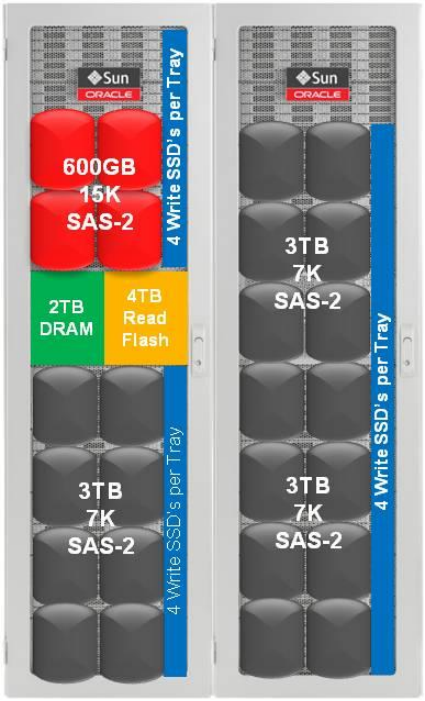

Figure 2. Oracle ZFS Storage Appliance – architecture overview

Additional Features

The following table provides some features of the Oracle ZFS Storage Appliance.

TABLE 1. ORACLE ZFS STORAGE APPLIANCE FEATURES	
DATA PROTOCOL	Fibre Channel, SCSI, InfiniBand over IP/RDMA, iSER, SRP, NFSv3 and v2, CIFS, WebDAV, HTTP/HTTPS, FTP/SFTP/FTPS, ZFDS NDMP V4.
DATA SERVICES	Hybrid Columnar Compression, Hybrid Storage Pool, Single, double, and triple parity RAID (RAIDZ, Z2, Z3), Mirroring and triple mirroring, End-to-end data integrity, Remote Replication, Snapshots and Clones Quotas, Compression, Thin Provisioning, Antivirus using ICAP protocol, Online data migration and Clustering.

MANAGEMENT	Browser and CLI interface, Management dashboard, Hardware/component view, Role-based access control, Phone home, Event- and threshold-based alerting, DTrace Analytics, Scripting, Workflow automation, Advanced networking, DFS root support and Source-aware routing.

For more comprehensive information on features, functions, and capacity of specific platforms, please refer to: http://www.oracle.com/us/products/servers-storage/storage/unified-storage/index.html

Oracle ZFS Storage Appliance and Virtualization

The architecture of the Oracle ZFS Storage Appliance works seamlessly with virtualization. Its features, detailed in the following sections, ensure that the Oracle ZFS Storage Appliance combines enterprise-class data services with massive scalability and flexibility while delivering significant cost savings and lower TCO. The Oracle ZFS Storage Appliance features:

- Provide excellent performance and throughput as well as broader bandwidth to meet the requirements of critical applications running in virtualized environments.
- Offer a new concept of architecture combining solid state drives (SSDs), dynamic random access memory (DRAM) and ZFS file systems, avoiding the typical saturation of resources caused by intense concentration of I/O in virtual desktops and virtual machine environments.
- Provide a realistically scaled amount of L1 cache based on DRAM that is complemented by a large L2 cache extension based on SSD disks.
- Use intelligent cache architecture for fast reading and writing operations (especially helpful for random I/O workloads originated by virtual machines and hypervisors).
- Are designed to handle heavy I/O load for virtualization environments, providing better performance and throughput than traditional NAS storage architecture.
- Are based on Hybrid Storage Pool architecture and efficient caching technology that allow running thousands of virtual machines on high capacity disks (3TB) without performance problems, bottlenecks, or I/O saturation.
- Are extremely fast for virtual machine operations, such as clones, linked-clones, snapshot, and migration.
- Offer different levels of compression and thin provisioning features that can be utilized in combinations to improve the storage efficiency.
- Use Intel multicore processors with features such as hyper-threading, Intel QuickPath (QPI) and Intel Turbo Boost, combined with Oracle Solaris, which is a multithreaded operating system. The Oracle ZFS Storage Appliance supports up to four Intel Xeon processors per controller, each of which can have up to 10 cores and 20 threads.
- Provide more security based on different RAID configuration levels that can be spread across single or multiple disk trays, as well as checksum and auto-correction.

- Provide LUN block size that automatically matches the block size of the volume, with a range from 512 bytes to 128kB.
- Provide native functionality and interoperability across multiple vendors and toolsets.
- Avoid virtual machine boot storm problems when booting hundreds or even thousands of virtual machines at the same time.

Using Hybrid Storage Pools for Efficiency Gains

Most of the storage available on the market today utilizes a small amount of NRAM devices (usually 1, 8 or even 16GB in size) as the first cache tier of the system and also hard disk drivers. NRAM devices are expensive, and disk drivers' performance is affected by seek operations, rotation and transfer times, which can result in I/O bottleneck and performance problems for operations requested by the application and virtualized environments. Oracle ZFS Storage Appliance implements a Hybrid Storage Pool architecture designed to work with multiple tiers of storage media to maximize the performance for the virtualized environment.

The Hybrid Storage Pools of the Oracle ZFS Storage Appliance can be divided into three tiers (shown in Figure 3):

- **First Tier:** DRAM (large L1 cache) – DRAM memory and high-optimized and low-latency solid state disks combined with ZFS file systems architecture (ReadZilla) accelerate read-cache operations for the virtualized environment. Unlike traditional NAS architecture, Oracle ZFS Storage Appliance and its Hybrid Storage Pools utilize DRAM devices as the main cache device of the system. DRAM devices are cheaper, faster and deliver higher performance than NVRAM, so they are well matched for random I/O workloads – the kind of workloads performed by hypervisors/virtualized environments. DRAM devices are also used by the Adaptive Replacement Cache (ARC), which is part of the ZFS file system architecture model and intelligently managed by instructions provided by multiple cache algorithms.
- **Second Tier:** SSDs (large L2 cache) – High-optimized and low-latency solid state disks combined with ZFS file systems log architecture (ZIL or LogZilla) to accelerate write-cache operations. These provide excellent performance and fast response for writing operations performed by applications and databases running in virtualized environment. The Hybrid Storage Pools are SSD devices designed to provide fast writing operations (100 times faster than traditional disk drivers) with low latency. Inside of the Hybrid Storage Pool architecture, SSD devices host the ZFS ZIL log (known as LogZilla or ZFS Intent Log), which is part of the ZFS file system architecture and mainly responsible for accelerating the synchronous writing operations requested by the critical applications and databases running in virtualized environments. Also, the SSDs are utilized by the Layer 2 Adaptive Replacement Cache (L2ARC), which is an extension of the ARC (main cache of the system) and hosts the read log devices for the ZFS architecture.
- **Third Tier:** Disk Pools – Disk pools are composed of high-performance (15000 rpm) and/or high capacity (7200 rpm) disks that are protected by different RAID levels and intelligently managed by the ZFS file system. Disk pools are designed to archive the application data, providing continuously high I/O rates for different types of workloads, even when utilizing high-capacity disks (7200 rpm).

Disk pools can optionally be configured with 15000 rpm disks, which provide the highest performance and thousands of IOPS for datastores typical of virtualized environments.

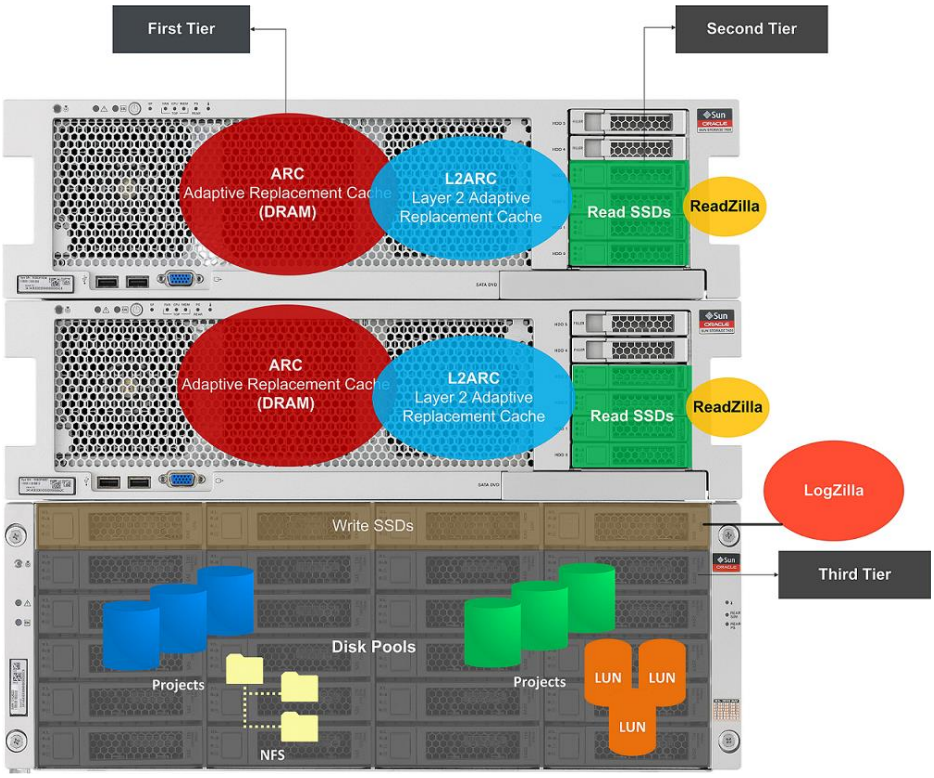

Figure 3. Oracle ZFS Storage Appliance Hybrid Storage Pool architecture

Using Intelligent Cache

A caching mechanism can dramatically influence the performance of critical applications running in virtualized environments. Oracle ZFS Storage Appliance provides a more sophisticated and "big" caching technology than traditional NAS storage solutions, as all devices that are part of the caching architecture (DRAM, SSDs) are intelligently managed by multiple algorithms for both monitoring and addressing the best performance and resource utilization. Scalable up to 2TB per head, DRAM memory additionally reduces the I/O contention and efficiently improves the cache utilization for virtualization environments, resulting in fast response time and low latency for virtualized applications. As the SSDs and DRAM work in combination as L1 and L2 cache devices, they maximize the I/O performance.

As part of the intelligent caching and ZFS file system architecture, two important technologies also deliver optimal performance for virtualization. They are ARC - Adaptive Replacement Cache - and L2ARC - Layer 2 Adaptive Replacement Cache. The ARC, located in DRAM, is the ZFS main cache

memory and designed to provide high performance. The ARC size is defined by the amount of physical RAM available on the system and is scalable up to 2TB. The L2ARC is the second cache layer of the Oracle ZFS Storage Appliance and sits in between the disks and the ARC cache; it consists of optimized low latency SSD disks. The L2ARC acts as an extension of ARC, and it is designed to accelerate read cache operations and improve the performance of random workloads. Figure 4 shows the difference between a traditional NAS storage architecture model and the Oracle ZFS Storage Appliance.

Oracle ZFS Storage Appliance utilizes a large L1 and L2 combined cache architecture:

- **Reading Cache**: Combination of big amount of L1 cache (DRAM memory =ARC) and a large L2 cache based on SSDs disks (extension of the main memory = L2ARC) – ZFS ReadZilla
- **Writing Cache**: Combination of SSD disks and ZFS ZIL – ZFS LogZilla

Traditional NAS storage architecture provides a small amount of L1 cache that is expensive and slow for virtualized environments:

- **Reading/Writing Cache** – Write-Back or Write-Through configuration modes based on NVRAM devices.

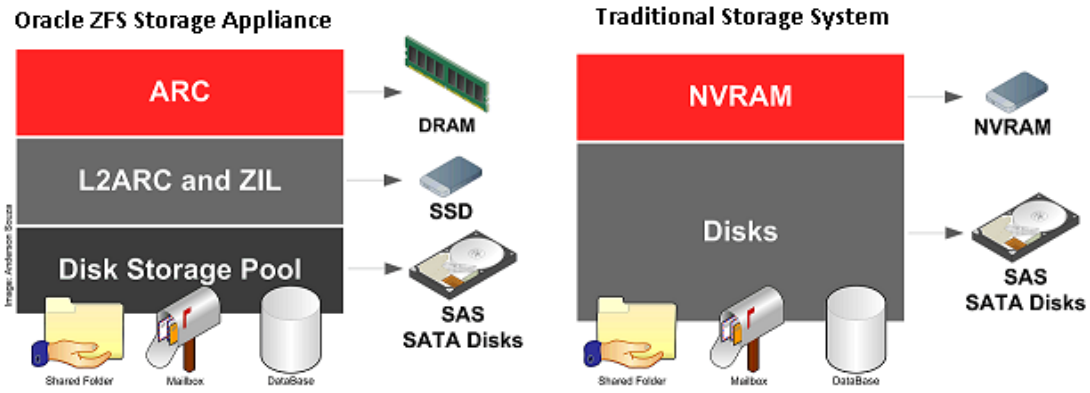

Figure 4. Comparing Oracle ZFS Storage Appliance and traditional storage architecture

Figure 5 represents the operational flow of the cache architecture performing in a virtual application. The schematic shows the following steps:

Step 1 – The application running in a virtual machine sends a "written" operation to the Oracle ZFS Storage Appliance – (random I/O).

Step 2 – The Oracle ZFS Storage Appliance writes the request on the cache, which is built of fast and optimized SSDs, then sends the answer back to the application running in the virtualized environment.

Step 3 – Afterward, in a sequential mode (intervals of 5 seconds), the Oracle ZFS Storage Appliance writes the data to disks.

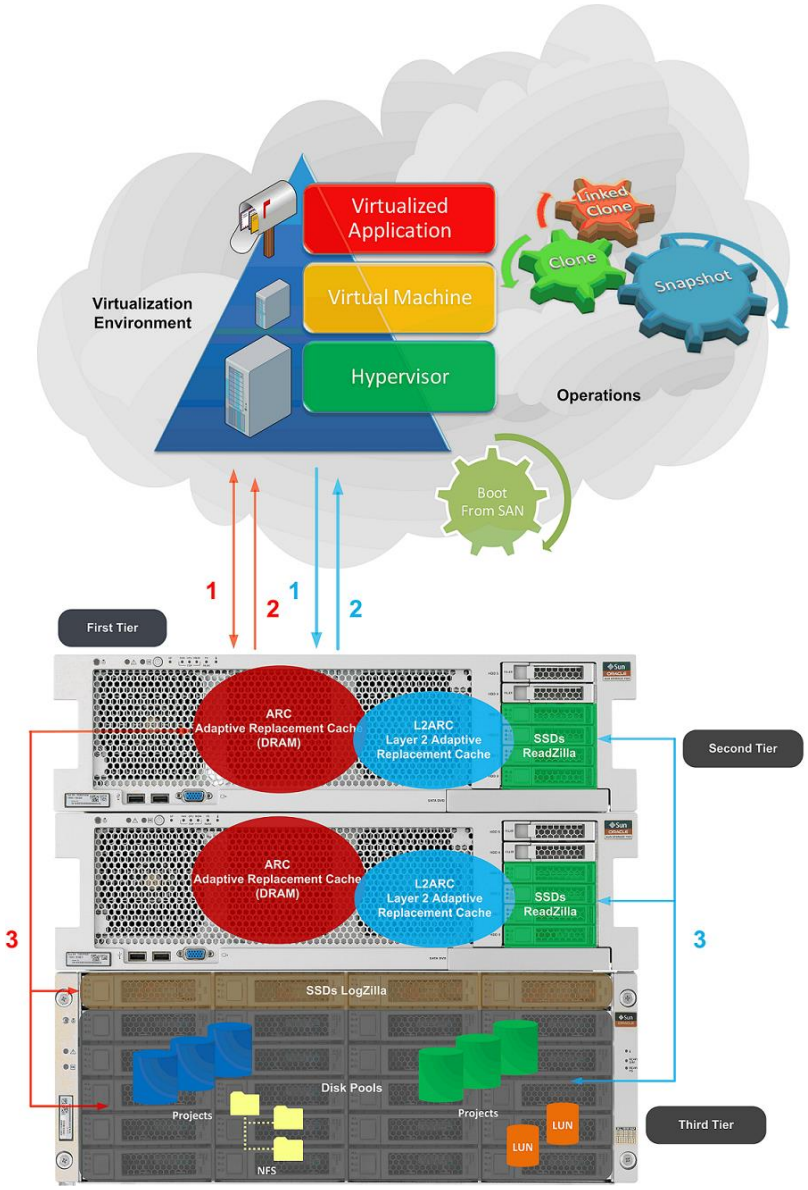

Figure 5. Oracle ZFS Storage Appliance – How intelligent cache works with virtualized applications

Using Enhanced Storage Features in Virtualized Environments

As previously mentioned, the Oracle ZFS Storage Appliance provides key features for virtualization environments – such as clones, snapshots, and shadow migration – as well as comprehensive multiprotocols – such as NFS, FC, iSCSI, CIFS, SMB, and IB. In addition, the Oracle ZFS Storage Appliance offers compression and thin provisioning features which can be leveraged to maximize storage capacity and efficiency. In combination, these features can significantly save disk space.

Oracle ZFS Storage Appliance offers four different levels of compression that allow administrators to address the specific compression needs of different applications. While data compression, which

optionally compresses data before writing to disk, mainly functions to reduce the amount of disk space utilized by the application, it can, in some cases, also improve system performance. As compression algorithms use fewer bytes of data, I/O operations to and from disks are reduced. In a virtualized environment, data compression levels offered by the Oracle ZFS Storage Appliance can be used to compress ISO images and virtual machine (VM) disks as well as images and VM templates, which maximizes storage efficiency.

Protecting Data in the Virtualized Environment

The Oracle ZFS Storage Appliance provides different levels of disk failure protection as well as data profile configurations, ranging from protection against single disk failures to mirroring data between disks in different shelves to protect against shelf failures. The data profiles available are double parity, mirrored, single parity narrow strips, striped, triple mirrored and triple parity RAID wide stripes.

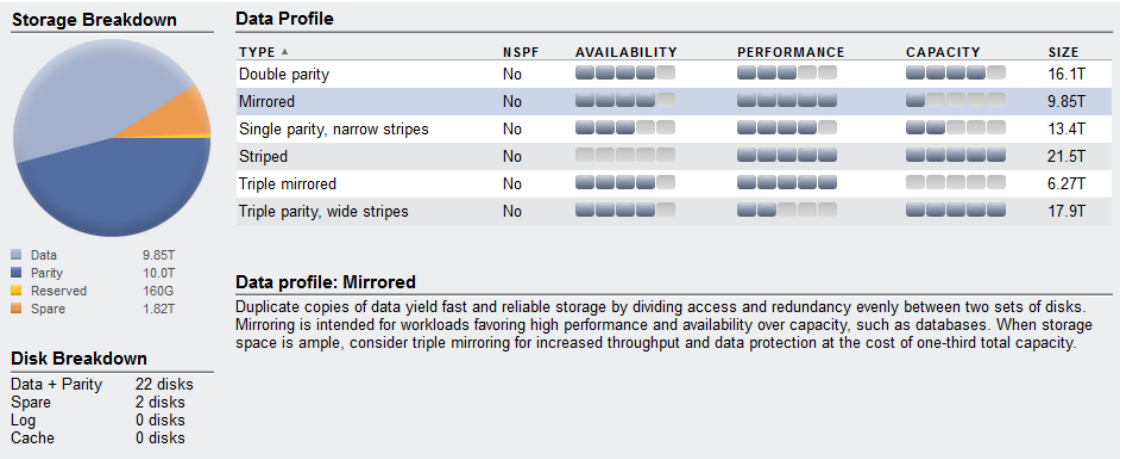

Figure 6. Disk date profile configuration shown in the Oracle ZFS Storage Appliance BUI

For more information regarding disk failure protection levels, please refer to:

Oracle ZFS Storage Appliance Administration Guide

Also, the Oracle ZFS Storage Appliance uses an advanced set of comprehensive data protection features providing snapshot, rollbacks, cloning and remote replication to protect application databases in the virtualized environment.

A snapshot is a point-in-time copy of a LUN or a file system and initially does not consume additional storage space until the data changes. With the snapshot capability of the Oracle ZFS Storage Appliance, you can manually create and schedule copies. These snapshots can be made visible and accessible by the Hypervisor in read-only mode, permitting re-allocation of VMs or entire datastores to different projects. The Oracle ZFS Storage Appliance also offers an unlimited number of snapshots from the source.

A clone is a writable copy of a snapshot that is treated as an independent resource. Like snapshots, clone operations do not consume additional storage space until the data is changed. In a virtualized

environment, you can use these snapshot and clone features to create different virtual machine releases, different versions of virtual images, as well as local migration of entire VMs and datastores.

The rollback feature is a fast mechanism for data restoration, and an effective solution for rolling back virtual machine and virtual datastores. With rollbacks, shares can be easily rolled back from a determined snapshot, restoring all data back to its initial point-in-time status.

With remote replication, data is asynchronously replicated from one Oracle ZFS Storage Appliance (source) to one or more remote sites (targets). With this feature the entire virtualized environment can be replicated or even utilized as a secondary site in a disaster recovery situation. Remote replication allows fast remote failover and failback operations as well as the ability to use remote sites as backup or a secondary solution in the events of failure of the primary site or a scheduled maintenance.

Monitoring in a Virtualized Environment

The Oracle ZFS Storage Appliance is easy to use because of the simplicity it offers in managing data, applications, and day-to-day data center tasks. Key to that management is monitoring, and the Oracle ZFS Storage Appliance provides detailed analytics as well as a robust alert system. The command-line interface also provides an accessible tool for managing storage resources.

DTrace Analytics

The advanced analytics tool called DTrace Analytics is designed to monitor and graph in detail a variety of different statistics of the operating system stack. DTrace Analytics provides unique visibility into bottlenecks and other issues that might exist in the virtual infrastructure. Examples are realtime statistics of throughput, capacity and utilization for items such as network, CPU, memory (ARC), flash (L2ARC) and storage (virtual disks' information as well as the exact IOPS and latency per virtual machine). These statistics allow administrators to quickly diagnose and resolve performance issues across the storage environment as well as record and play back storage workload events for further analysis. The following screenshots provide examples of DTrace Analytics utilization.

Figure 7 shows that the NFSv3 protocol is used for virtual machines' disk datastore. DTrace Analytics are monitoring the virtual machines' disk utilization in IOPS for each vmdk file.

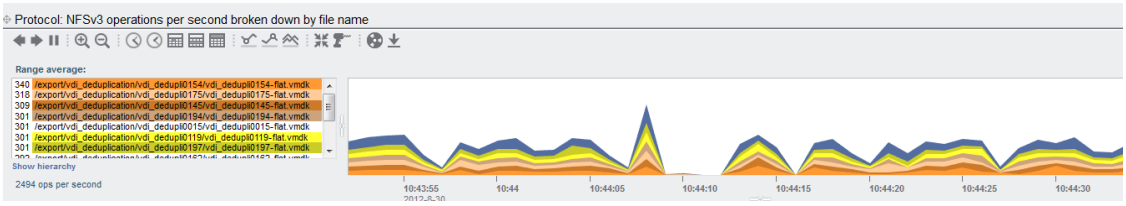

Figure 7. DTrace Analytics for protocol NFSv3 operations per second by file name (virtual disks)

Figure 8 shows the network utilization (in bytes per second) broken down by device.

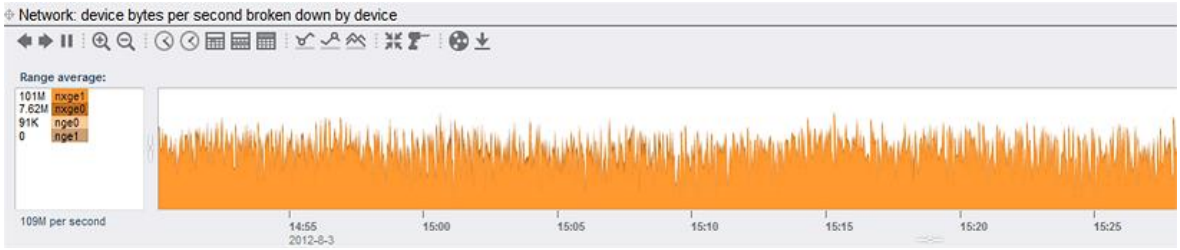

Figure 8. DTrace Analytics for network devices bytes per seconds

Figure 9 shows the disk I/O per second broken down by type of operation (read/write). DTrace Analytics offers different drill-down options; Figure 10 shows the "write" drill-down option that is monitoring disk I/O bytes per second of type write broken down by disk.

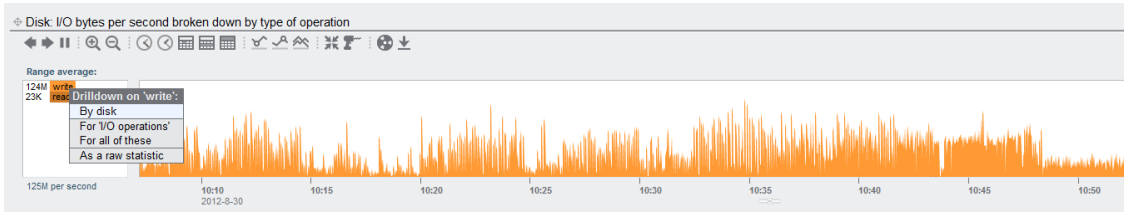

Figure 9. Disk I/O bytes per second broken down by type of operation

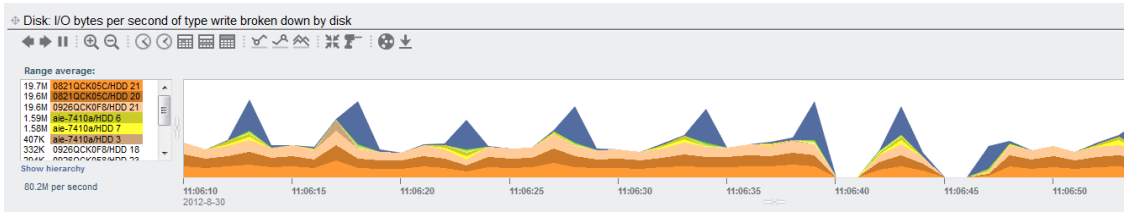

Figure 10. Drill-down option, by disk, of disk I/O bytes per second of type write

Figure 11 shows the graphical representation of CPU utilization.

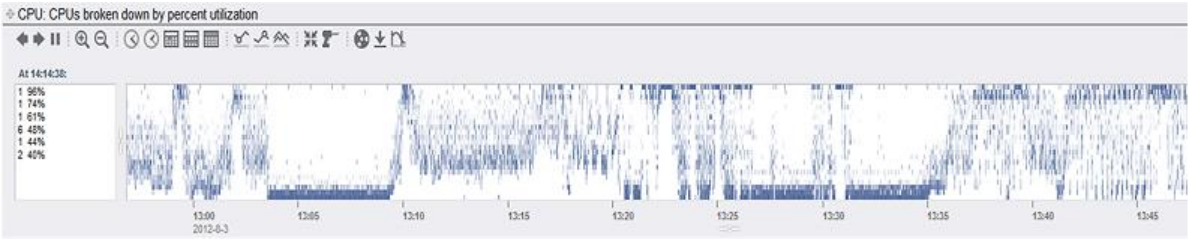

Figure 11. DTrace Analytics for CPU utilization percentage

Threshold and Alerts

Thresholds and alerts are important tools for monitoring statistics from the analytics and providing alerts about important metrics of the systems. In a virtualized environment, a threshold alert can be configured, for example, to monitor a virtual disk, defining a specific percentage of utilization as well as timing and alert options (such as send email, SNMP message, and syslog). Figures 12 and 13 respectively show the threshold and alert configuration screen in the Oracle ZFS Storage Appliance BUI. Alerts can be used to monitor different Appliance events such as: ZFS storage pool events including scrub and hot space activation, hardware and software failure and events, and cluster events that include link failures and peer errors.

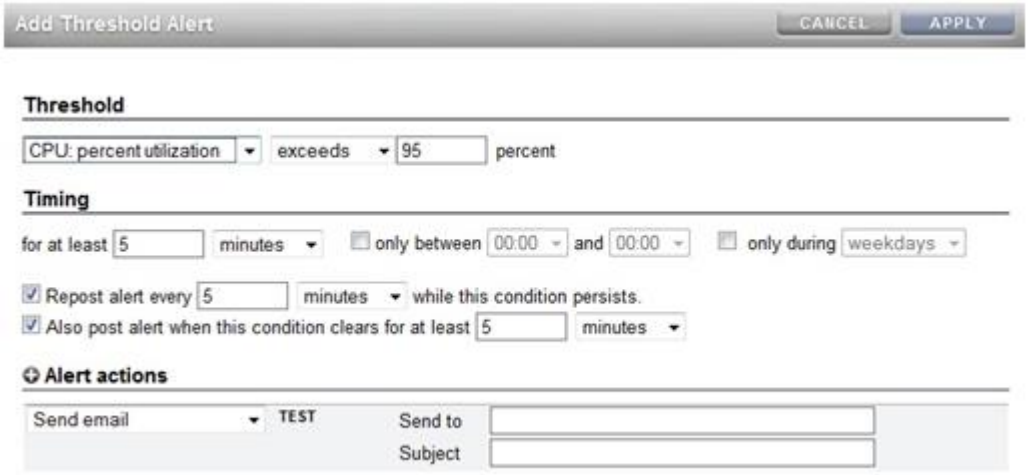

Figure 12. Threshold configuration screen

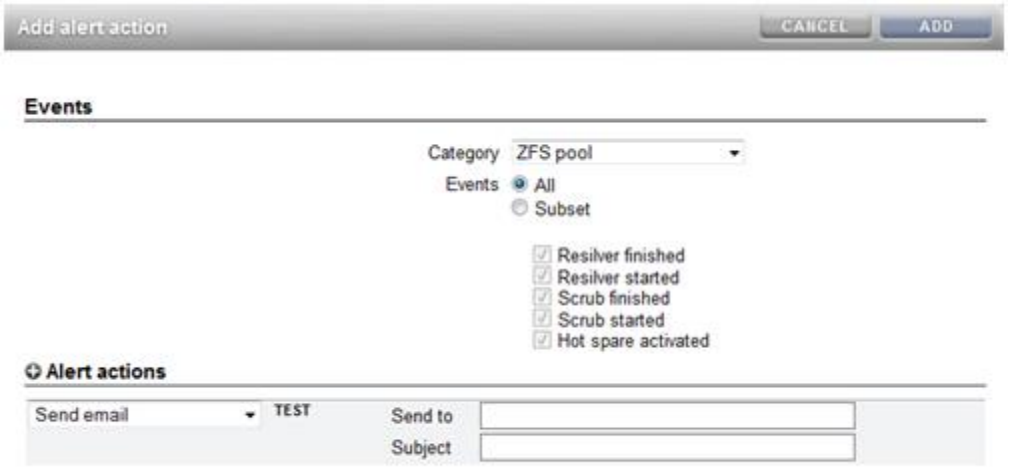

Figure 13. Alert configuration screen

In combination, these detailed monitoring features provide the accessible means to easily identify possible problems with the virtual environment as well as quickly address and remediate them.

Command Line Interface – CLI

The command line interface, or CLI, provides a powerful scripting environment in which to perform repetitive tasks utilizing scripts or batch jobs. The CLI can be accessed by SSH or through the serial port of the Oracle ZFS Storage Appliance, and provides an alternate route to administer the Oracle ZFS Storage Appliance in situations in which it is not reachable through the network. Designed to be a mirror of the Oracle ZFS Storage Appliance BUI, the CLI is useful for automating tasks for virtualized environments. Those tasks can include configuration of hundreds or thousands of LUNs at the same time, mapping of disks, execution of workflows, management of storage pools, and total administration of the system.

Interoperability with Other Virtualization Technologies

The Oracle ZFS Storage Appliance integrates seamlessly into virtualized environments and provides native functionality across multiple vendors as well as plug-ins for Oracle VM, Microsoft Volume Shadow Copy Service and VMware vCenter Site Recovery Manager. Supported plug-in technologies include:

- Asymmetric Logical Unit Access (ALUA) – Used to provide FC target multipathing support for virtualized environments
- Full interoperability with VMware vSphere4 and 5 utilizing VMW_SATP_ALUA
- Interoperability with Citrix Xenserver 6 and Microsoft Hyper-V

- Storage plug-in for Oracle VM – Oracle Virtual Machine Storage Connect Plug-in for Oracle ZFS Storage Appliance
- Network File System (NFS) plug-in for Oracle Solaris Cluster
- Volume Shadow Copy Services (VSS) plug-in for Microsoft operating systems
- Storage Replication Adapter (SRA) plug-in for VMware vCenter Site Recovery Manager (SRM) – Oracle ZFS Storage Appliance Storage Replication Adapter for VMware Site Recovery Manager 4.x

Conclusion

The Oracle ZFS Storage Appliance provides a storage concept and architecture that is optimized for virtualized environments. The Oracle ZFS Storage Appliance provides a user-friendly interface for easier management, a set of data services for business continuity/disaster recovery, and innovations in architecture and caching technology. To maximize storage efficiency, the Oracle ZFS Storage Appliance provides unique features such as compression and thin provisioning. For the performance side, a fast CPU, large storage capacity as well as better throughput and broader bandwidth help the Oracle ZFS Storage Appliance solidly meet the requirements of virtualized environments, delivering optimal performance for virtualization.

The Oracle ZFS Storage Appliance can play a key role in accelerating a virtualized infrastructure, enabling organizations to dramatically simplify their storage deployment, management and utilization while reducing costs and improving efficiency.

Appendix: References

NOTE: References to Sun ZFS Storage Appliance, Sun ZFS Storage 7000, and ZFS Storage Appliance all refer to the same family of Oracle ZFS Storage Appliance products. Some cited documentation or screen code may still carry these legacy naming conventions.

RESOURCE	LOCATION
Oracle ZFS Storage Appliance Administration Guide	http://download.oracle.com/docs/cd/E22471_01/index.html
Oracle ZFS Storage Appliances	http://www.oracle.com/goto/nas
Storage Performance Council benchmark results	http://www.storageperformance.org/benchmark_results_files/SPC-2/Oracle_SPC-2/B00058_Oracle_ZFS-7420/b00058_Oracle_Sun-ZFS_7420_SPC2_executive-summary.pdf
General Oracle information	http://www.oracle.com
Oracle virtualization technologies	http://www.oracle.com/us/technologies/virtualization/index.html
Oracle VM VirtualBox	http://www.oracle.com/us/technologies/virtualization/oraclevm/061976.html
Quality awards for Oracle's NAS solutions	http://www.oracle.com/us/products/servers-storage/storage/nas/storage-quality-awards-jan12-1521728.pdf
Open Solaris information	http://www.opensolaris.com

Why Oracle ZFS Storage Appliance Optimizes Storage in Virtualized Environments
December 2013, Version 1.1
Author: Anderson Souza

Oracle Corporation
World Headquarters
500 Oracle Parkway
Redwood Shores, CA 94065
U.S.A.

Worldwide Inquiries:
Phone: +1.650.506.7000
Fax: +1.650.506.7200

oracle.com

Oracle is committed to developing practices and products that help protect the environment

Hardware and Software, Engineered to Work Together