The Oracle ZS3 Storage Appliance

The Industry’s First Convergence of Oracle Database and Networked Storage

By John Webster

September, 2013
A New Storage System Choice: General Purpose or Optimized for the Database Application?

Years ago, there were a limited number of choices when it came to selecting a high performance, production data center quality storage array. The enterprise data center storage user could choose from a handful of disk-based SAN-attached and NAS arrays that were designed for general purpose applications. Enterprise storage buyers are now offered abundant choices. Storage platforms now available include NAS and SAN, hybrid, scale-out, object, all solid state, software defined, and the most recent entry: storage that is optimized for an application or set of applications—database applications for example.

The emergence of storage that is optimized or “engineered” by the storage vendor for a specific purpose or set of applications—applications supported by a common database for example—shifts the IT buyer’s decision making process. We note that it is common for users to buy general purpose storage knowing full well that it will be used during its entire productive life for one application such as MS Exchange, SQL Server, or Oracle Database. Knowing that, users will then “tweak” the system for performance and availability within that specific application environment and layer manual management processes that are unique to the application environment on top.

The application and workload are always top of mind when defining storage requirements. Application-focused engineering of the external storage system by the storage vendor simply moves that thinking to the center of the decision making process with an eye toward maximizing performance and automating the IT management processes commonly used in these environments.

Here we look at a hybrid storage system that has been engineered by Oracle for Oracle Database applications—the new Oracle ZFS Storage Appliance, ZS3 Series—where we note a first time convergence of external, networked storage and database. The result is a hybrid storage platform that offers unique, automated management features that can simplify the lives of both storage administrators and database administrators.

The ZS3 Appliance—A Storage Platform for Oracle Applications

As a result of the Sun acquisition, Oracle inherited the open systems storage appliance based originally on Sun servers running the Zettabyte File System (ZFS). As Oracle now rolls-out the next new release of the ZFS storage appliance—the ZS3—it counts over 4,000 customers and over 13,000 ZFS Storage appliances installed.

In some respects, the ZFS Storage Appliance was an early example of software defined storage when it was first announced in 2008. Hardware consisted simply of a standard Solaris-based server with disk in a
JBOD configuration. The critical features that made it enterprise-grade storage were defined in ZFS. These included RAID-based data protection (RAID Z); a number of data replication options including cloning, snapshotting, and remote copy; as well as advanced features now found in modern storage arrays including thin provisioning, clustering, and deduplication. At the time it was introduced, the original ZFS Storage Appliance offered the first use of a storage system-based analytics tool called DTrace Analytics software—a lead that some competitors are just now starting to follow.

The preceding ZFS Storage Appliance to the ZS3 experienced strong growth in mixed workload environments where its performance results across SPC-2 and SPECsfs have been notable. Oracle now reports record SPECsfs/NFS results for the ZS3. It has also gained traction in highly virtualized environments with SMP workloads, such as VMware, while avoiding the addition of multiple NAS filers to accommodate growth.

ZS3 represents a new direction for external hybrid storage arrays—the first convergence of Oracle database with the ZFS appliance. Not surprisingly, Oracle has unencumbered visibility into the entire information flow—from creation at the application level, to storage on the platform and on to long-term retention in the archive. This gives Oracle a unique ability to use an external storage system as a way to improve the experience of the application user experience and optimize information workflow. Oracle refers to this as “application engineered” storage.

The Oracle ZS3—Doubling Down on Oracle Application Optimization

The previous ZFS Appliance release (November, 2011) built on the extensive feature set of ZFS by including Hybrid Storage Pool (HSP) technology to integrate and manage multiple storage performance and capacity tiers that included high-performance FC disk, high capacity SATA (now SAS) disk, and SSD for the first time. It also introduced the notion of building-in storage-based features that were specifically aimed at enhancing the Oracle environment as experienced by the application user and the manageability of Oracle databases from the standpoint of DBAs and storage administrators. Perhaps foremost among these was Hybrid Columnar Compression (HCC—discussed in more detail below) that greatly improved the storage efficiency of the ZFS Appliance by compressing infrequently accessed data by up to a factor of 50.

ZS3 doubles down on the Oracle application engineered storage strategy. It includes new capabilities that, when integrated with Oracle Database 12c-based applications, will enhance their performance and manageability in numerous ways that general purpose NAS storage currently can’t address. In addition,

1 Contact Oracle for more information.
these features, discussed in detail below, will be of as much if not more interest to Oracle DBAs as they will to enterprise storage administrators.

**Multiple Hybrid Storage Pools (HSPs)**

The Hybrid Storage Pool is a feature of ZFS that was included in the first ZFS Appliance and has been carried forward with enhancements since then. The HSP concept uses file system-based software to manage data within a single storage system to control placement of application data on different classes of storage disk devices for performance and cost reasons. Alternatively, the Automatic Data Optimization feature in Oracle Database 12c automatically discovers and stores data access patterns and represents them in a heat map. Information contained in the heat map can be used to automatically optimize data placement within the ZS3 HSP.

A single ZS3 can have multiple distinct Hybrid Storage Pools (HSPs) that share DRAM and Flash caches. Use of ZS3 HSPs by storage and database administrators offers an up to 2X improvement in response time over the previous version through improved caching algorithms. In fact, Oracle states that as a result of the new caching algorithms, 70-90% of I/O can now be serviced from the ZS3’s DRAM Cache—a factor that no doubt contributes significantly to the ZS3’s SPECsfs/NFS record performance benchmark results mentioned earlier. We also note that Oracle has implemented in-memory deduplication to maximize DRAM utilization.

Using HSPs offers both storage and database administrators a range of options. Here, it’s helpful to think of HSPs as composed of spinning disk (performance/capacity) in differing data protection configurations (mirrored, RAID Z2, etc.) as what defines each storage pool, while DRAM and read/write flash are shared among HSPs. For example, an administrator could define a high-performance disk pool for actively used OLTP data which could be mirrored high performance drives. Data is staged from HSP disk to DRAM-based cache during periods of peak demand. Afterward, it ages back out into the Flash-based read cache at the mid-level. Under this caching architecture, sometimes referred to by Oracle as Dynamic Storage Tiering, the data has a single physical home (disk) which is consistent. Data access is accelerated by the use of DRAM and Flash as read and write cache that is shared among the HSPs.

Example HSP configurations include:

**Maximum IOPS Pool**

- Mirroring
- Performance Optimized Disks

**Maximum Streaming Pool**

- Mirroring or Single-parity RAID-Z
• Capacity Optimized Disks

Balanced Performance & Capacity

• Single-parity RAID-Z
• Capacity Optimized Disks

Max Capacity & Availability Pool

• Double-parity RAID-Z2 or triple parity RAID-Z3
• Capacity Optimized Disks

Using HSPs enables storage administrators to put actively updated portions of an Oracle Database on “Maximum IOPS” storage pools while those portions used for reporting and archive could be placed on “Balanced” and “Maximum Capacity” pools. DBAs could then optimize capacity utilization by assigning different levels of compression to each of these pools, such as no compression for the active data, HCC query compression for reporting data, and maximum HCC archive compression for archive data.

The Oracle ZS3 combines multiple storage pools (HSPs), multiple levels of HCC compression (see below), and automates data placement among HSPs using Oracle Database 12c Automatic Data Optimization (see below). Storage and database administrators can now automate data placement among storage pools optimized either for performance or storage efficiency.

HSPs and Pluggable Oracle 12c Databases

In addition to the introduction of multiple HSP support, Oracle announced that independent databases can have access rights to read-flash and write-flash resources dynamically controlled within the ZFS Storage Appliance HSP. Oracle Database 12c Enterprise Edition introduced multi-tenancy in the form of “pluggable” databases. A “container” 12c database is implemented that holds multiple pluggable databases. The ZFS Appliance also supports the assignment of 12c pluggable databases to individual HSPs and can simplify the administration of them on a single ZFS Storage Appliance. For example, a storage administrator can fine-tune the performance of pluggable databases with DTrace Storage Analytics (see below). This gives the storage administrator greater ability to control the performance parameters for a given Oracle application or user group.

Oracle Intelligent Storage Protocol (OISP)

Oracle Database 12c also introduces Oracle Intelligent Storage Protocol (OISP) to enable the database to communicate directly with the Oracle ZS3 to optimize configurations and auto-tune storage systems without administrative intervention. Presently OISP enables the automation of logbias and record size tuning. Since there is a direct communications path from Oracle Database 12c to Oracle ZS3, we expect
that future releases will take additional advantage of this protocol and provide deeper provisioning capabilities and more advanced database-to-storage optimization. OISP impacts overall IT administration positively by reducing the number of manual tasks required to optimize the database and supporting storage system for the application environment.

**Snapshots and Clones**

The ZS3 Appliance supports the creation of point-in-time snapshots for data protection or development as well as non-disruptive cloning of production or backup database images. The ZS3’s snapshots are managed by Snap Manager Utility for Oracle DB. Cloning can be also used to establish multiple full copies of the Oracle Database for development and test purposes. A full backup of the production database is first created and stored on an Oracle ZFS Storage Appliance using Oracle Recovery Manager (RMAN). The backup copy can be kept current by using Oracle Data Guard to apply updates on a continuous basis. The full database backup on the Oracle ZFS Storage Appliance can then be cloned as many times as needed for development and test environments. ZS3 cloning stores only changed data, allowing developers to each utilize their own copy of the database while enhancing the efficiency of ZS3 storage capacity.

**Hybrid Columnar Compression (HCC) and Automated Data Optimization**

Oracle Hybrid Columnar Compression (HCC) is supported on the ZS3 Appliance and can be used to automatically compress query and archive data 10-50x, dramatically reducing the cost of storing and managing the ZS3’s less frequently accessed data and increasing application performance when querying historical data.

As mentioned earlier, Oracle Database 12c Automatic Data Optimization takes HCC two steps further. First, it intelligently tracks data usage and creates heat maps based on active (hot), less active (warm), inactive (cool) data categories. Then, it enables policies to be set so that data can automatically move from no compression to HCC compression as a background process called Heat Map Tiering (HMT) without administrative intervention to track the data usage and change compression settings. Additionally, it does so on an ongoing basis so that DBAs and application users take advantage of HCC compression and performance benefits throughout the lifecycle of the database.

Unique to the ZS3 is the fact that there is no need to decompress or “rehydrate” data for secondary processing. Database and storage administrators can create clone and snapshot copies of HCC-compressed data for test, development and quality assurance and can keep HCC compression through backups as well.
DTrace Analytics

For real-time visualization of performance-related metrics and to troubleshoot and resolve bottlenecks, ZS3 features a built-in application and workload analysis tool called DTrace Analytics. DTrace measures ZS3 CPU, cache, protocol, disk, memory, networking, and system-related data, and presents the results graphically and in real-time. Administrators have mouse-click drill down access to areas of concern from application to storage at the I/O and virtual machine-level. DTrace also supports multiple simultaneous application and workload analysis in real-time to help compare various aspects of system stress. Analysis can be saved, replayed and exported to other administrative tools for further analysis. Administrators can also remotely monitor their ZS3 environments from their Apple iPhones or iPads with a secure application, ZFS Appliance Monitor, that enables monitoring of storage service logs, I/O statistics, real-time analytics, component status, faults and recommended repairs. We see DTrace as a groundbreaking use of real time analytics at the storage system level that can save the significant time that is often spent by IT administrators when application users complain of sub-par performance.

Enterprise Manager Integration

Not surprisingly, given the level of Oracle Database/ZFS Appliance integration already demonstrated, the ZS3 includes a software plug-in that enables Oracle IT management groups to monitor and
troubleshoot the ZS3. Enterprise Manager will now support “single pane of glass” management to display and respond to ZS3 storage alerts, incidents and problems, to view metrics and analytics generated by DTrace, and to create on-demand reports. This enables DBAs and storage administrators to monitor and manage performance from the application level through the Oracle database, servers, and storage so as to more quickly identify the root cause of application performance issues and respond to them. An iPhone app called the ZFS Appliance Monitor (ZAM) is also available that supports remote monitoring of multiple ZS3 Appliances.

Available ZS3 Configurations and Transparent Data Migration

Upon first release, the ZS3 will be available as a mid-range storage appliance (ZS3-2) and a highly scalable enterprise-class model (ZS3-4) as follows:

**ZS3-2**

- Single or Dual Controllers
- 512GB DRAM
- 15.2 TB Cache
- 12TB Read Flash
- 2TB Write Flash
- 8 Disk Enclosures/768 TB raw uncompressed storage capacity
- 12 PCIe slots for 10Gbe, 40Gb IB, SAS, and/or 16Gb FC connectivity

![The Oracle ZS3-2 Storage Appliance](image)
ZS3-4

- Single or Dual Controllers
- 2TB DRAM
- 25.3 TB Cache 12TB Read Flash
- 2TB Write Flash
- 36 Disk Enclosures/3.5 PB raw uncompressed storage capacity
- 22 PCIe slots for 10Gbe, 40Gb IB, SAS, and/or 16Gb FC connectivity

The Oracle ZS3-4 Storage Appliance

In order to migrate data to the ZS3 without disruption to normal production operations, Oracle offers the ZFS-SA Shadow Migration feature. ZFS-SA supports transparent data migration from any NFS source to an Oracle ZFS-based storage appliance. Migrations of data to the ZS3 as a target could include earlier ZFS Appliances or NAS filers from other vendors. Oracle also offers a service related to Shadow Migration in cases where customers want assistance in setting-up the feature.

Conclusion—Database-Aware Storage meets Storage-Aware Database

With the introduction of the ZS3 Storage Appliance, Oracle’s hardware/software integration strategy that was announced with the acquisition of Sun Microsystems becomes ever clearer with very tangible customer accessible business and technology benefits. Oracle Database 12c, released in July, 2013, includes a number of features aimed at optimizing the storage environment including OISP, Automatic Data Optimization, and pluggable databases. Converging the Oracle ZS3 with Oracle 12c Database results in a unique and more efficient management environment than that which would be experienced using more general purpose storage platforms from the likes of EMC, NetApp, and IBM etc. that adapts to the changing needs of the database, DBAs, and high-performance businesses that need to get the most out of their IT investments.

Given the significant level of convergence and cross-stack automation with Oracle 12c, enterprise IT storage administrators managing the ZS3 can become significantly more responsive to the needs of
Oracle applications and application users. Oracle DBA’s will also have new avenues upon which they can make the Oracle storage environment more responsive to their needs. And while Oracle ZFS Storage Appliances are well suited to most workloads, Oracle application and database administrators will experience a more responsive storage environment with ZS3, while Oracle database and storage administrators will get simplified and more efficient storage management.

About Evaluator Group
Evaluator Group Inc. is dedicated to helping IT professionals and vendors create and implement strategies that make the most of the value of their storage and digital information. Evaluator Group services deliver in-depth, unbiased analysis on storage architectures, infrastructures and management for IT professionals. Since 1997 Evaluator Group has provided services for thousands of end users and vendor professionals through product and market evaluations, competitive analysis and education. www.evaluatorgroup.com Follow us on Twitter @evaluator_group

Copyright 2013 Evaluator Group, Inc. All rights reserved.
No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying and recording, or stored in a database or retrieval system for any purpose without the express written consent of Evaluator Group Inc. The information contained in this document is subject to change without notice. Evaluator Group assumes no responsibility for errors or omissions. Evaluator Group makes no expressed or implied warranties in this document relating to the use or operation of the products described herein. In no event shall Evaluator Group be liable for any indirect, special, inconsequential or incidental damages arising out of or associated with any aspect of this publication, even if advised of the possibility of such damages. The Evaluator Series is a trademark of Evaluator Group, Inc. All other trademarks are the property of their respective companies.