Oracle White Paper
November 2012

Expanding the Storage Capabilities of the Oracle Database Appliance

Executive Overview

The Oracle Database Appliance is an engineered system offering that saves time and money by simplifying the deployment, maintenance, and support of a high availability database solution - all supported by a single vendor – Oracle. It is a fully integrated system with software, servers, storage and networking in a single box capable of supporting a wide range of home grown, packaged OLTP and Data Warehousing applications.

The Oracle Database Appliance contains two internal servers and all of the components required to deploy standalone or clustered Oracle databases. From the factory, the hardware contains 4TB of usable hard drive space for storing the contents of the databases[1]. When the space requirements exceed 4TB, it is possible to expand the storage using external storage devices such as the Oracle Sun ZFS Storage Appliance. This paper discusses the concepts, techniques, and best practices that should be used to create a scalable storage solution.

Introduction

Oracle Database Appliance is a self-contained solution for standalone or clustered databases. It contains server, storage, and network hardware, combined with network, cluster, and database software and templates. The hardware and software are engineered together to be simple to configure and maintain, and to be preconfigured for database workloads. Oracle fully supports all hardware and software components. Oracle Database Appliance is designed to minimize cost, time, and risk in database deployment and maintenance.

Customers can use Oracle certified NFS-attached storage with Oracle Database Appliance to store database files for both read and write operations. This external storage can be used to further extend the storage offered by Oracle Database Appliance.

Some use cases of why to expand the Oracle Database Appliance storage include:

- Expand beyond the 4TB[1] limit: place additional data on NFS attached storage.
- Backups: create database disk backups external to the Oracle Database Appliance.

[1] Starting with ODA Software Release 2.4, usable space can be 6TB when using Normal ASM Diskgroup redundancy.

- Storage Tiering: place frequently accessed tables on internal SAS disk, and less frequently used tables on NFS. You can partition the large tables and move the older partitions to an NFS attached tablespace based on usage

- Enable Hybrid Columnar Compression (HCC): see dramatic compression and performance gains using HCC on ZFS Storage Appliance

Storage

The Oracle Database Appliance uses twenty hard drives for storing user data. These disks are 600 GB SAS hard drives, allowing for a total of 12TB of raw storage. They are hot-pluggable, front mounted, and are accessible to each of the two servers in the Oracle Database Appliance.

The Oracle Database Appliance is designed to tolerate hardware component failures. The onboard storage subsytem is designed for maximum availability, where each server has its own disk backplane. If one server loses access to the disks, the other server will still have access.

When a new Oracle Database Appliance is configured for use, Oracle Automatic Storage Management (ASM) is used to create and manage the underlying tablespaces. The ASM +DATA and +RECO diskgroups are created during the install of Oracle Database Appliance. When configuring the appliance, you have the options of "External Backups" or "Internal Backups". The +RECO tablespace will be larger if the "Internal Backup" option is chosen during the installation process. For more information on Oracle Database Appliance backup, please review the white paper, Backup and Recovery Strategies for the Oracle Database Appliance.

All of the ASM diskgroups are configured to use "high redundancy" (triple mirroring) to ensure high availability and survival in the case of disk loss. ASM handles the recreation of data whenever a disk is lost.

Since all user data written to the onboard disk is triple mirrored, the net result is 4TB of usable space[2]. For some implementations, the available usable space is insufficient. To address this matter, customers may expand their storage to one or more network attached storage devices using Oracle dNFS. Use of Oracle dNFS is the only supported method.

[2] Starting with ODA Software Release 2.4, usable space can be 6TB when using Normal ASM Diskgroup redundancy.

Networking

The Oracle Database Appliance has various options for attaching network based storage. It currently contains a total of six 1GbE ports and two 10GbE ports. Ports are bonded at the kernel level to provide high availability in an active/passive configuration. If one of the bonded ports were to fail, the network connections would failover to the passive network port and continue operating. The Oracle Database Appliance uses four bonded network configurations:

- Net0 and Net1 - Network ports for 1GbE public network.
    - Commonly used for the database public network interface
    - Eth2 and Eth3, bonded as "bond0"
- PCIe 0 – Network ports for 10GbE public network
    - Commonly used for the database public network interface when 10GbE is required
    - Eth8 and Eth9, bonded as "Xbond0"
- PCIe 1 – Additional 1GbE network interface ports
    - Commonly used for NFS or backup networking
    - Eth4 and Eth5 bonded as "bond1"
    - Eth6 and Eth7 bonded as "bond2"
- Eth0 and Eth1 – internal interconnect

Determining the best Network Port for NFS

When selecting a network interface to use with your network based storage, you'll need to choose one of the available bonded interfaces and plan for two network cables; one for each network port in the bond. Ideally, these cables would connect to two independent switches for optimal high availability. Additionally, you can establish multiple paths to the storage using multiple bonded interfaces for example, bond1 and bond2. You should chose a network interface based on the specific needs of your implementation whether it be for performance or port availability. If bond0 is used for public traffic, you'll need to choose one of the other available interfaces. If 10GbE is required, Xbond0 must be used. For 1GbE, bond1 or bond2 would be the conventional choice.

As a general rule, the storage traffic should be isolated from the database traffic which is typically using bond0 or Xbond0. If database and storage traffic are combined, it becomes more difficult to troubleshoot network performance issues. To ensure proper routing of storage traffic, static routes can be applied to the bond interface you select.

For more details on the Oracle Database Appliance physical specification, please refer to the Oracle Database Appliance Documentation Owner's Manual or Getting Started Guide here: http://www.oracle.com/technetwork/server-storage/engineered-systems/database-appliance/documentation/index.html

Expanding Storage Considerations

While Oracle Database Appliance provides 4TB (6TB possible with ODA Software Release 2.4) of storage to use for your databases, there are considerations when expanding storage to NFS devices concerning high availability configured with the network and latency when it comes to accessing data or backups located on external storage.

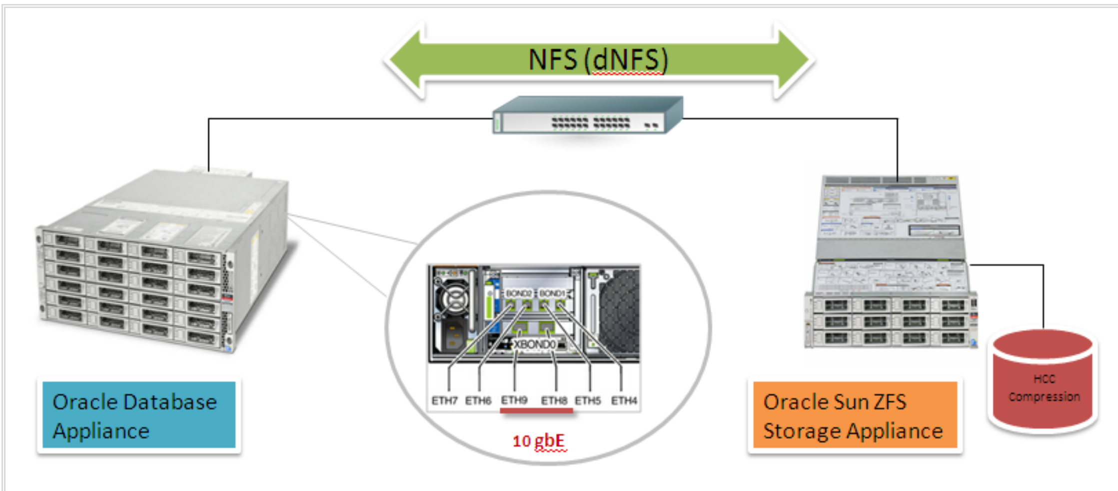

Figure 1 Oracle Database Appliance using dNFS with Oracle Sun ZFS

High Availability

The Oracle Database Appliance is an engineered system that is designed from the hardware through the software stack to ensure high availability of the shared disk subsystem. Each Oracle Database Appliance server has redundant paths to the disks in the event of a hardware component failure. Additionally, each server can access the disks independently of the other, which means one server can be down while the other server still has access to the disks.

When an external storage device is added to this type of pre-engineered environment, the high availability characteristics of the storage must be evaluated carefully. The network or the external storage device itself could become a "point of failure" if there was an outage. Should the network or device become unavailable, you would lose access to those datafiles.

Performance

The performance characteristics of network based storage must also be considered since it will likely add a measurable amount of latency that doesn't exist when accessing user data on the onboard disks. For that reason, you will want to think strategically about where you place your user data. Data that requires optimal response time and the highest levels of availability should be placed in the ASM diskgroups on the appliance hard drives. Data placed on the external storage will not have the same performance characteristics as the data that resides in the appliances ASM diskgroups. Archived data or backups for example, would be an ideal choice for external storage.

Oracle ASM diskgroups are created on the appliance at install time. The +DATA tablespace contains user data. For Oracle Database Appliance, ASM is not supported on network based storage. ASM may only be used for datafiles that reside directly on the appliance HDDs. This means you cannot expand an ASM managed tablespace to network based storage. You also cannot create a new ASM managed tablespace on network based storage.

For expanding the storage capabilities of the appliance for your database, you must create filesystem-based tablespaces on one or more NFS mounted directories. These new tablespaces are not limited in any way. For example, they may be Read-only, Read-write, Oracle-managed, or use non-standard block sizes. When creating tablespaces that will reside on network storage, you should be sure to consider all of your applications' data access requirements.

Oracle Direct NFS

Oracle offers the ability to manage NFS using a feature called Oracle Direct NFS (dNFS). Oracle Direct NFS implements NFS V3 protocol within the Oracle database kernel itself. Oracle Direct NFS client overcomes many of the challenges associated with using NFS with the Oracle Database with simple configuration, better performance than traditional NFS clients, and offers consistent configuration across platforms.

Configuring Direct NFS

In order to use Oracle Direct NFS, the NFS file systems must first be mounted and available over regular NFS mounts using Oracle's recommended NFS mount options.

To configure dNFS, you create a configuration file called oranfstab. This file contains the options, attributes, and parameters that enable Oracle Database to use Direct NFS. The default location for this file is $ORACLE_HOME/dbs directory. Direct NFS determines mount point settings for NFS storage devices based on the configuration information in oranfstab file.

You can also place the oranfstab file under the /etc directory in order to make the dNFS settings globally available to database instances. This needs to be done on both nodes.

```
server: MyNFSServer
path: 142.48.25.12          -- xbond0 interface
export: /storage/sharedvol  mount: /u02/oradata/nfs
```

Figure 2 Example oranfstab file

The next step in configuring Direct NFS involves configuring the dNFS option within the Oracle database kernel. The following steps need to be executed on both nodes.

1. Shutdown the Oracle Database instance.
2. cd $ORACLE_HOME/rdbms/lib
3. mk –f ins_rdbms.mk dnfs_on
4. Start the Oracle Database instance.

After starting the instance, database instance alert.log will show the following:

```
Oracle instance running with ODM: Oracle Direct NFS ODM Library
Version 3.0
```

Figure 3 Alert.log output with dNFS running

Placing Data on NFS Storage

After configuring dNFS, you can create tablespaces on the NFS storage.

```
SQL> create tablespace archive_data datafile
'/u02/oradata/nfs/archive_data01.dbf' SIZE 500M;

Tablespace created.

SQL> select file#, name, status from v$datafile;

 FILE# NAME                                          STATUS
------ --------------------------------------------- -------
     1 +DATA/pbrb/datafile/system.272.781351553      SYSTEM
     2 +DATA/pbrb/datafile/sysaux.273.781351557      ONLINE
     3 +DATA/pbrb/datafile/undotbs1.274.781351561    ONLINE
     4 +DATA/pbrb/datafile/undotbs2.276.781351569    ONLINE
     5 +DATA/pbrb/datafile/users.277.781351571       ONLINE
     6 +DATA/pbrb/datafile/users.bigfile             ONLINE
     7 /u02/oradata/nfs/archive_data01.dbf           ONLINE
```

Figure 4 Create Oracle tablespace on NFS

Verifying Oracle Direct NFS Usage

Subsequent startup of the database instance alert log will show the following information during subsequent startup:

```
ALTER DATABASE OPEN
Direct NFS: attempting to mount /storage/sharedvol on filer mynfsserver defined in
oranfstab
Direct NFS: channel config is:
     channel id [0] local [] path [142.48.25.12]
Direct NFS: mount complete dir /storage/sharedvol on mynfsserver mntport 770 nfsport
2049
…
Direct NFS: channel id [0] path [142.48.25.12] to filer [mynfsserver] via local [] is UP
Direct NFS: channel id [1] path [142.48.25.12] to filer [mynfsserver] via local [] is UP

You can also query the V$DNFS_SERVER view.

SQL> select * from v$dnfs_servers;

ID         SVRNAME      DIRNAME                  MNTPORT   NFSPORT      WTMAX      RTMAX
-----------------------------------------------------------------------------------------
1         mynfsserver  /storage/sharedvol         770      2049         0           0
```

Figure 5 Create Oracle tablespace on NFS

Oracle Sun ZFS Storage Appliance

The Oracle Sun ZFS Storage Appliance is a network attached storage (NAS) device that provides robust data storage solution that offers Oracle integration, simplicity, efficiency, performance, and Total Cost of Ownership (TCO).

There are several features of the Oracle Sun ZFS Storage Appliance that make it the ideal choice for expanding the storage capabilities of the Oracle Database Appliance. These key features include intuitive management tools, real-time performance analysis, data compression, multiprotocol integration, deduplication, cloning etc.

Benefits of using the Sun ZFS Storage Appliance

- Simplified storage management.
- Highly reliable data updates with data corruption protection.
- High availability and high performance.
- Oracle Hybrid Columnar Compression (HCC) for Oracle Databases
- Optimal use of network resources using compression because fewer blocks are transferred across the network.
- Ability to scale I/O throughput, processor performance, and storage capacity as storage needs change.

Storage Device Options

- Sun ZFS Storage 7120 - 3.3 TB to 177 TB of raw capacity
- Sun ZFS Storage 7320 - up to 432 TB raw
- Sun ZFS Storage 7420 - can expand to more than 1.7 PB of raw

For configuring ZFS Appliance for an Oracle database, please refer to the following whitepaper:

http://www.oracle.com/technetwork/articles/systems-hardware-architecture/ss700-config-oracle-db-163902.pdf

ZFS Appliance Overview

http://www.oracle.com/us/products/servers-storage/storage/nas/zfs-appliance-software/overview/index.html

Mount Options for Oracle

https://support.us.oracle.com/oip/faces/secure/km/DocumentDisplay.jspx?id=359515.1

Compression

Innovations in Oracle compression technologies help customers reduce the resources of managing large data volumes. It allows IT administrators to significantly reduce their overall database storage footprint by enabling compression for all types of data – be it relational (table), unstructured (file), or backup data. Compression benefits include:

- Reduced Live OLTP data size
- Reduced Backup Size
- Reduced Disaster or Standby database size
- Reduced Export Dump size
- Reduced disk I/O while reading data blocks (without overhead while reading)
- Reduced network traffic while sending archivelogs to DR site.

Hybrid Columnar Compression

With Oracle Database Appliance version 2.2 and higher and Sun ZFS Storage Appliance (ZFSSA), data can now be compressed using Hybrid Columnar Compression (HCC). This is significant because with HCC, you can store your user data in significantly less space and retrieve data with less scan IO. With HCC warehouse compression, you can usually have 6-10x in storage savings and with HCC archive compression you can have 15-70x storage savings. For both archive and warehouse compression, there are LOW and HIGH settings from which you can choose. The increased storage savings may cause data load-times to increase modestly. Therefore, LOW should be chosen for environments where load time service levels are more critical than query performance.

HCC will allow you to maximize your storage capabilities and account for accelerated data growth without sacrificing performance. The I/O savings you'll see with HCC will actually improve query performance, sometimes significantly, in both OLTP and data warehousing environments.

Traditionally, Oracle organizes the data within a database block in a 'row' format where all column data for a particular row is stored sequentially within a single database block. HCC stores the data in a nontraditional Oracle format. HCC utilizes a combination of both row and columnar methods for storing data (hence the term "hybrid"). Since data in columns often has more duplicate values, this type of compression can result in significant increases in storage savings and improved performance.

When using HCC, the Oracle database blocks are organized into logical constructs called "compression units", which is used to store a set of hybrid columnar-compressed rows. Each compression unit consists of multiple data blocks. HCC provides different levels of compression, focusing on query performance or compression ratio respectively. To make effective use of HCC, data must be loaded using data warehouse bulk loading techniques.

```
SQL> create table mydata (…) compress for query [ LOW |
HIGH ];

Archive compression:

SQL> create table mydata (…) compress for archive [ LOW |
HIGH ];
```

Figure 6 Example to use HCC when creating new tables

With HCC, you can create new compressed tables or alter your existing tables to use compression using the *ALTER TABLE MOVE* command. Use of HCC is also supported on partitioned tables. For example, newer data can be placed on partitions that utilize warehouse compression while older data can be placed on partitions using archive compression.

HCC can be very valuable in an environment where the database files are stored on a network storage appliance (ZFSSA in this case). Since the data is compressed, fewer blocks will be transferred across the network, which allows the database appliance processors to receive blocks more quickly and spend less time waiting for I/O. It also puts less pressure on the storage server because it can retrieve a smaller number of blocks more quickly. The storage server's cache can also be used more efficiently.

Advanced Compression

Oracle offers Advanced Compression option for compressing data. This is also called "OLTP Table Compression". This supports compressing data using all DML activities (INSERT, UPDATE, and DELETE) as well as direct path inserts. This differs from HCC in that the compression occurs at the block level and it can span columns.

Typically, the storage savings range from 2x to 4x using this type of compression. The compression algorithm runs in the background once a certain number of DML activities have occurred on a table. This works well in both OLTP and Data Warehousing environments because the compression does not occur during the DML activity. There is no overhead for writing compressed data and minimal overhead for reads.

```
SQL> ALTER TABLE … COMPRESS FOR OLTP
SQL> ALTER TABLE … MOVE COMPRESS FOR OLTP

To determine if compression is in use:

SQL> SELECT TABLE_NAME, COMPRESSION, COMPRESS_FOR FROM USER_TABLES;

TABLE_NAME             COMPRESSION    COMPRESS_FOR
----------------       -----------    ------------------
T1                     DISABLED
T2                     ENABLED               BASIC
T3                     ENABLED               OLTP

SQL> SELECT TABLE_NAME, PARTITION_NAME, COMPRESSION, COMPRESS_FOR FROM
USER_TAB_PARTITIONS;

TABLE_NAME     PARTITION_NAME   COMPRESSION  COMPRESS_FOR
-----------    ---------------- -----------  -------------------------
SALES                 Q4_2008    ENABLED            OLTP
SALES                 Q1_2009    ENABLED            OLTP
SALES                 Q2_2009    ENABLED            OLTP
```

Figure 7 Oracle Advance Compression example

For g:

http://www.oracle.com/technetwork/database/storage/advanced-compression-whitepaper-130502.pdf

http://www.oracle.com/technetwork/database/options/compression/faq-092157.html

Basic Compression

Basic table compression was introduced in Oracle9*i*. This was effective at compressing data that was loaded into the database using bulk insert methods.

Expanding Storage Summary – Step by Step

Best Practices for Expanding Storage

- Isolate storage network traffic from general database traffic.
- Use multiple cables, switches, and network paths to the storage device.
- Use 10GbE for the best possible network performance.
- Place the least accessed tables on external storage verses the onboard hard drives.
- Use Oracle Direct NFS (dNFS) for managing NFS mounts.
- Use Hybrid Columnar Compression (where appropriate) to best utilize storage, networking, and CPU resources.

Prepare the Network Storage Device

- Identify space requirements for your environment. Determine if usable space is adequate space to store your databases.
- Procure the necessary storage hardware.
- Identify networking requirements between the Oracle Database Appliance and the storage devices.
- Is 10GbE or 1GbE network needed?
- How many switches are available?
- Identify which ports are available for use on the Oracle Database Appliance.
- Run the cables that will be used to attach the networked storages.
- Create the necessary filesystem(s) that will be used for storing Oracle database files on the storage device.
- Export the filesystems on the storage devices.

Prepare the Oracle Database Appliance

- Identify the tables/partitions that will reside on the network storage.
- Configure Oracle dNFS to manage the network shares.
- Create tablespaces for external storage by placing the new datafiles on the newly created NFS mount points.
- Create or Move tables to the new tablespaces you just created.

Conclusion

In this document, we have discussed options to expand the storage capabilities of your Oracle Database Appliance. With this knowledge, you should be able to design a storage infrastructure that will scale with your business needs. Using NFS and Oracle dNFS, you can create a storage system that is easy to manage using Oracle's best practice guidelines. You've learned that using Hybrid Columnar Compression on the Sun ZFS Storage Appliances can help you maximize storage capacity and performance in a network storage environment. Lastly, we've listed the basic steps required to build an NFS based storage infrastructure from the ground up.

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

ExpandngOracle Database Appliance Storage
July 2012

Authors: Ravinder Ransi, Josh Ort
Contributing Authors:
RACPack / A-Team

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