ORACLE
EXADATA

Storage Efficient Database Copy Options With Exadata

ORACLE

Introduction

Cloning a production database is an often needed procedure in customer environments to develop and test new applications or test changes to an existing application. When a new application or additional functionality is added, it calls for thorough testing using production data. Today this is achieved by making a copy of the production data files in a test environment. In addition to the test environment, copies of the production database are also made in one or more development environments where application developers are creating their new applications and testing them. Each developer usually wants to maintain his/her own copy of the production data. All of these potentially require a huge amount of storage space to be allocated and managed.

When a customer addresses the above requirements by creating multiple environments with full size production copies, they are faced with the following challenges:

(1) Each of these environments occupy the same amount of space as their parent database
(2) Creation of each environment consumes a lot of resources as the entire source database has to be read and written
(3) It takes longer to create a full copy

What is needed is a different approach to creating multiple read/write copies of the same database such that each copy can maintain their own changes without affecting other DBs and that does not suffer the above challenges. We will refer to this approach as *creating a storage efficient database copy*.

This ability to create storage efficient database copies is very useful to organizations. Many times different members of a team need to perform isolated changes to their own copy of a database without impacting others. In other words, there can be multiple DEV or TEST projects, each of them performing updates to the base database, requiring isolated environments.

Exadata customers have various methods of creating storage efficient database copies. These capabilities can be divided between:

» Capabilities that natively exist in Exadata
» Capabilities that can be used with Exadata but require additional hardware and/or licensing

Natively available capabilities in Exadata to create storage efficient database copies

The following options are natively available to Exadata customers to create storage efficient database copies.

(1) SPARSE disk group based database snapshots

SPARSE disk group based database snapshots is functionality included in Exadata Storage Software. It requires Exadata Storage Software version 12.1.2.1.0 and Oracle Database 12c version 12.1.0.2 with bundle patch 5 or later. This feature is designed to work on native Exadata ASM storage disk groups. It uses ASM SPARSE grid disk based thin provisioning where snapshot databases created on a SPARSE disk group need only the space for changes plus some metadata, thereby enabling storage efficient snapshot databases.

Snapshot technology as deployed on Exadata is "allocate on first write", not copy on write. Snapshots start with a shared read-only copy of the production database (or pluggable database (PDB)) that has been cleansed of any sensitive information. When any of the child snapshots issues a write, the write is issued to a private copy of that block inside that snapshot, preserving the base blocks for potential reads from other snapshots. If, for the child database, this is a write of a block it has never written, then a new block is allocated. Subsequent writes of the same block by the child will use the same, privately allocated, block. Thus the allocation only happens on first write - thus the term "allocate on first write".

Multiple users can create independent snapshots from the same base database, therefore multiple test and development environments can share space while maintaining independent databases for each task. The base database must remain read-only during the usable lifespan of snapshots that are created from that base database. If there is a need to create another set of snapshots from a more recent copy of production data, a new read-only base from a production database needs to be created. If there are space constraints for creating a new read-only base then existing snapshots can be dropped and the current base database can be refreshed from production. Exadata database snapshots are integrated with the Multitenant Database option to provide a simple interface for creating new PDB snapshots.

SPARSE disk group based snapshots on Exadata Storage Servers support all Oracle Exadata Storage Server Software features such as Smart Scan, storage indexes and IORM, as the snapshots are stored in ASM storage that is native to Exadata. This approach is recommended for Exadata customers as it is available natively, fully integrated with ASM, and customers will benefit from the future enhancements made to this functionality.

(2) ACFS Snapshots

Oracle ASM Cluster file system (ACFS) is a general-purpose POSIX, X/OPEN and Windows compliant file system designed for single node and cluster configurations. An Oracle ACFS snapshot is an online, point-in-time copy of an Oracle ACFS file system. ACFS snapshots based on Copy on Write capability are available on the Exadata platform when using Grid Infrastructure 12.1.0.2 or later.

Oracle ACFS supports all database files and general purpose files on Oracle Exadata Database Machine running Oracle Linux on database servers. Though the primary purpose of ACFS is to store general purpose files, it also supports storing database files. On Exadata Database Machine, Oracle ACFS supports the following database versions: Oracle Database 10g Rel. 2 (10.2.0.4 and 10.2.0.5), Oracle Database 11g (11.2.0.4 and higher) and Oracle Database 12c (12.1.0.1 and higher). In addition, all Oracle ACFS advanced functionality is supported on Exadata. This includes: snapshots, replication, tagging, security, encryption, audit and HANFS.

Oracle ACFS is managed using native operating system commands as well as the Oracle ACFS command line tool, acfsutil. Oracle ACFS is also fully integrated with Oracle Enterprise Manager Cloud Control. Database files stored in ACFS on Exadata storage are subject to the following guidelines and restrictions:

» Oracle ACFS replication or security/encryption/audit is only supported with general purpose files. These functions are currently not supported with database files in an Oracle ACFS file system.

» ACFS snapshots should not be used to snapshot the live, running prod database on Exadata for performance reasons.

» Oracle ACFS does not support the Exadata offload features (e.g. Smart Scan, Storage Indexes, IORM, Network RM).

» Exadata Smart Flash Cache will cache read operations. Caching of write operations is expected in a later release.

» No specialized cache hints are passed from the database to the Exadata Storage layer, which means the Smart Flash Cache heuristics are based on I/O size, similar to any other block storage caching technology.

» Exadata Smart Flash Logging is not supported.

» Hardware Assisted Resilient Data (HARD) checks are not performed.

Please refer to Oracle ACFS Support on Oracle Exadata Database Machine (Linux only) (Doc ID 1929629.1) for additional information.

This is the only approach available on Exadata that supports Oracle Database 10g.

Capabilities that can be used with Exadata but require additional hardware and/or licensing:

These capabilities for creating storage efficient database copies with Exadata require Exadata to be connected to an external storage system and are directly or indirectly based on the underlying storage system's Copy on Write (CoW) features. CoW operates on volumes (block devices), but have no native understanding of the database topology.

Copy on Write (CoW) can be conceptually described as following:

"Copy-on-write stems from the understanding that when multiple separate tasks use initially identical copies of some information (i.e., data stored in computer memory or disk storage), treating it as local data that they may occasionally need to modify, then it is not necessary to immediately create separate copies of that information for each task. Instead they can all be given pointers to the same resource, with the provision that on the first occasion where they need to modify the data, they must first create a local copy on which to perform the modification (the original resource remains unchanged). When there are many separate processes all using the same resource, each with a small likelihood of having to modify it at all, then it is possible to achieve significant resource savings by sharing resources this way."

Since CoW capability depends on external storage, storage efficient databases created using this technology on Exadata do not support Oracle Exadata Storage Server Software features such as Smart Scan, storage indexes and IORM.

The following options are available to Exadata customers who already have acquired or have plans to acquire additional infrastructure to connect with Exadata Database Machine.

(1) **Thin provisioning of PDBs using "Snapshot Copy"** feature of Oracle 12c Multitenant Option when Exadata is connected to ZFS Storage Appliance or NetApp:

ZFS Storage Appliance and NetApp are hybrid storage arrays designed to meet enterprise storage needs. Oracle takes advantage of CoW capabilities of ZFS Storage Appliance and NetApp to provide thin provisioning of PDBs using the "Snapshot Copy" feature of Oracle 12c multitenant option. In the following example, the `snapshot copy` clause causes the connected appliances to perform thin provisioning of the DEV44 PDB.

create pluggable database DEV44 from CRM_MASTER snapshot copy;

Instantaneous Database Clones with Oracle Multitenant explains specific details on configuring the ZFS STORAGE APPLIANCE or NetApp appliance for snapshot cloning.

(2) **Direct NFS Client CloneDB** feature can be used for creating storage efficient copies of Exadata database if the database backups are stored on an NFS share mounted on Exadata compute nodes.

The Direct NFS Client (dNFS) CloneDB feature utilizes CoW technology and was introduced in Oracle Database 11g Release 1.dNFS is an implementation of the NFS client that runs as part of the Oracle Database engine. Through this integration, Oracle Database optimizes the I/O access path to the NFS server to provide improved performance, scalability, and reliability. By tuning the protocol to match typical database I/O, dNFS provides faster performance than can be provided by the operating system's NFS driver.

The CloneDB feature makes it possible to instantly clone an existing backup of a database mounted over dNFS. The clone process uses Copy on Write technology; so only the changed data file blocks need to be stored separately; the unmodified data is referenced directly from the backup files.

Exadata's native backup solution is to back up Oracle databases to the Fast Recovery Area (FRA). When Exadata is connected to network attached storage, the customer can choose to back up Oracle databases to an NFS share mounted on the compute node. In this scenario, the dNFS CloneDB feature can be used to create storage efficient copies of Oracle databases. Use of this feature will also mean that database clones will be using non-ASM file storage.

For more information, see "**How to Accelerate Test and Development through Rapid Cloning of Production Databases and Operating Environments**" white paper on CloneDB

(3) **Enterprise Manager Snap Clone** functionality is built on top of the Enterprise Manager 12c Storage Management Framework (SMF) plug-in. The SMF plug-in provides a layer of abstraction to shield DBAs and users from the nuances of the different storage systems. At the storage level, Snap Clone accesses underlying storage technologies, such as Copy On Write or similar technologies, to perform the required tasks. Snap Clone functionality can use either a hardware specific or hardware agnostic solution:

a. Hardware vendor specific solution requires customers to connect a storage subsystem with Exadata and license clone and snapshot capabilities from the storage vendor.

This method, currently certified for ZFS Storage appliances and NetApp Storage appliances, enables storage administrators to register storage appliances with Enterprise Manager, which then connects directly to the storage appliance to perform all required snapshot and clone operations. A major advantage of using ZFS Storage Appliance is the ability to support HCC seamlessly with Exadata. No other vendor offers this. This approach requires you to license the relevant options on the storage appliance, but it is the easiest, most efficient and fault tolerant approach.

b. Vendor agnostic software solution requires Solaris 11 file system (ZFS) on physical or virtual server as a pre-requisite:

This method connects to the storage via an intermediate (ZFS) file system. This is a storage vendor agnostic solution and can be used by any storage vendor, NAS or SAN. To configure this solution the storage administrator mounts the volumes to a Solaris server and formats it with ZFS file system. The ZFS file system then controls all snapshots and clone operations. The benefit of this approach is that it does not require thin cloning options to be licensed on the storage since ZFS file system provides these required capabilities. For additional details, check out the EM Snap Clone home page.

SUMMARY MATRIX: COMPARISON OF VARIOUS OPTIONS AVAILABLE FOR CREATING STORAGE EFFICIENT DATABASE COPY ON EXADATA

	Keyword by which this feature is known	Natively Available on Exadata?	Core Technology Used	Additional Hardware or licensing needed	Can cloned DB support Exadata software features?	Minimum Software Requirements	Supported Database Versions
Natively available capabilities in Exadata							
SPARSE disk group based database snapshot	SPARSE database snapshot	Y	New type of ASM Disk group called SPARSE disk group introduced in Exadata S/W version 12.1.2.1.0	None	Yes	Exadata Storage Software version 12.1.2.1.0 Oracle RDBMS version 12.1.0.2	12c and above
ACFS Snapshots	ACFS Snapshot	Y	Database File support for Exadata on ACFS introduced in Oracle Grid Infrastructure 12.1.0.2	None	No	Exadata Storage Software Version 11.2.3.3.1 (12.1.1.1.1 is recommended) and Oracle Grid Infrastructure 12.1.0.2	10.2.0.4, 10.2.0.5, and 11.2.0.4 or higher
Capabilities that can be used with Exadata but requires additional hardware and/or licensing							
Thin provisioning of PDBs using "Snapshot Copy" feature of Oracle 12c Multitenant Option	Snapshot Copy	N	Copy On Write of the connected storage sub-system	ZFS Storage Appliance or NetApp Hardware connected to Exadata. License to use COW technology of the hardware vendor, multitenant option of Oracle RDBMS 12c	No	12c RDBMS with multitenant option	12c and above
dNFS ClondDB	dNFS CloneDB	N	11gR1 implementation of the NFS client that runs as part of the Oracle Database engine	NFS Share mounted on Exadata.	No	11gR1 RDBMS	11gR1 and above
Enterprise Manager Snap Clone	Snap Clone	N	Copy On Write of the connected storage sub-system	Cloud Management Pack for 12c, ZFS Storage Appliance or NetApp Hardware connected to Exadata if using the H/W solution or Solaris 11 file system on a physical or virtual server connected to Exadata if using a S/W solution	No	EM 12c + SMF Plugin	10g and above

CONNECT WITH US

blogs.oracle.com/oracle

facebook.com/oracle

twitter.com/oracle

oracle.com

Oracle Corporation, World Headquarters
500 Oracle Parkway
Redwood Shores, CA 94065, USA

Worldwide Inquiries
Phone: +1.650.506.7000
Fax: +1.650.506.7200

STORAGE EFFICIENT DATABASE COPY OPTIONS WITH EXADATA
March 2015
Author: Manish Shah

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