Technical white paper

Best practices for Oracle and HP 3PAR StoreServ Storage

Includes HP StoreServ 7000 and 10000

Table of contents

Executive summary

HP 3PAR StoreServ Storage delivers the flexibility, efficiency, and high availability required in business-critical data centers. All HP 3PAR StoreServ models deliver simple yet powerful multi-tenant storage, and HP 3PAR StoreServ systems are supported by a powerful suite of software products and array features that enable ease of management, efficient storage utilization, thin provisioning, autonomic storage tiering, and leading availability features such as 3PAR Peer Persistence.

HP 3PAR StoreServ systems align with the growing trend towards consolidating Oracle databases into a scalable storage infrastructure. These systems manage mixed workloads in a single array, and they're also built with multi-tenancy in mind, so a single HP 3PAR StoreServ array can deliver high availability and performance to a data center with multiple Oracle environments. The HP 3PAR Gen4 ASIC in each controller provides on-the-fly optimization, whether the I/O is from Oracle Real Application Clusters (RAC) serving an online transaction database or a large data warehouse dependent on high throughput. Moreover, building on a common goal of high levels of resource utilization, Oracle and HP 3PAR StoreServ engineers have partnered to extend storage efficiency for Oracle database environments.

In an Oracle environment, HP 3PAR StoreServ systems enable data center consolidation, allowing customers to reduce administration overhead without compromising service levels. The HP 3PAR Thin Suite reduces costly stranded storage, while the HP Adaptive Optimization Suite tailors performance autonomically. HP 3PAR StoreServ systems are the ideal choice for Oracle installations that require a robust, scalable, and feature-rich storage solution.

Introduction

Oracle installations can include differing levels of service and reliability requirements. However, most deployments generally require high availability, excellent performance, and disaster tolerance. While many of these requirements can be met via host-enabled technologies, they often add significant host overhead that would be better utilized in application processing. A better option is to leverage the features of HP 3PAR StoreServ products to offload much of that Oracle host overhead while still meeting the most stringent enterprise demands.

This paper describes best practices for integrating HP 3PAR StoreServ with an Oracle environment. It provides technical descriptions of these integrated environments, including illustrations of solution architecture and implementation to provide context for making the best-informed design choices. The best practices in this document, along with the included references to more focused guides, will assist in building an optimal Oracle environment on top of HP 3PAR StoreServ solutions.

Intended audience: The information contained in this paper is intended for solution architects, database administrators, and system administrators who are involved in the design and deployment of Oracle database environments.

Overview of HP 3PAR StoreServ Storage

Shown in figure 1, HP 3PAR StoreServ elegantly meets the demands of the modern data center. With a wide range of models, HP 3PAR StoreServ delivers the efficiency and agility required by the most demanding virtual, cloud, and IT-as-a-Service (ITaaS) environments. HP 3PAR StoreServ customers spend less time managing storage, get more features for less money, and do it all without sacrificing performance or future scalability. HP 3PAR StoreServ is also the first product family with a common architecture that meets the needs of small- to medium-sized businesses (SMBs) as well as large global enterprises, enabling customers of all sizes to start small and grow without painful future upgrades.

The HP 3PAR StoreServ family includes the enterprise class HP 3PAR StoreServ 10000 and the midrange HP 3PAR StoreServ 7000. HP 3PAR StoreServ 7000 systems extend the innovative HP 3PAR StoreServ product line to the midrange and provide industry-leading performance and features at an attractive price. With "Thin Built In" to its HP 3PAR StoreServ family, HP remains the first storage vendor to incorporate thin capabilities into array hardware. The HP 3PAR Gen4 ASIC used in these systems provides hyper-efficient, silicon-based engines that drive simple, on-the-fly storage optimization while delivering high service levels.

**Figure 1**. The HP 3PAR StoreServ family
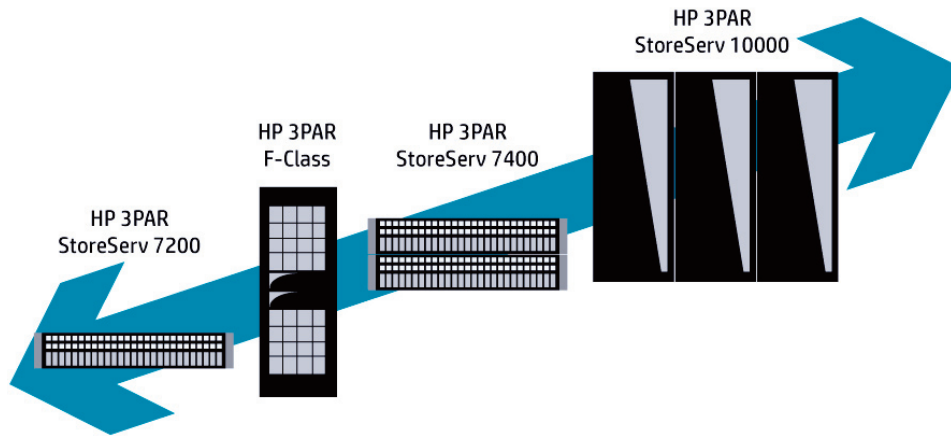
Key HP 3PAR StoreServ concepts

HP 3PAR StoreServ incorporates several improvements over conventional storage arrays. By making more effective use of all drive resources in the array, these enhancements allow higher performance with less hardware, which in turn leads to cost reduction.

HP 3PAR StoreServ storage arrays are a combination of redundant hardware and data structures designed to optimize performance and availability. As shown in figure 2, the HP 3PAR StoreServ array can be considered a three-layer stack:

- **Physical disks and chunklets**: Physical disks are the fundamental hardware layer of the 3PAR StoreServ storage array. When physical disks are configured in the array, they are formatted into small, equally sized allocations called "chunklets."

- **Chunklets and logical disks**: Logical disks (LDs) are collections of chunklets arranged into RAID sets.

- **Logical disks and virtual volumes**: Virtual volumes are created from logical disks. The makeup of a virtual volume is specified by a **common provisioning group (CPG)**.

Learn more about these structures on the following pages.

Figure 2. Hardware and data structures of the HP 3PAR StoreServ array
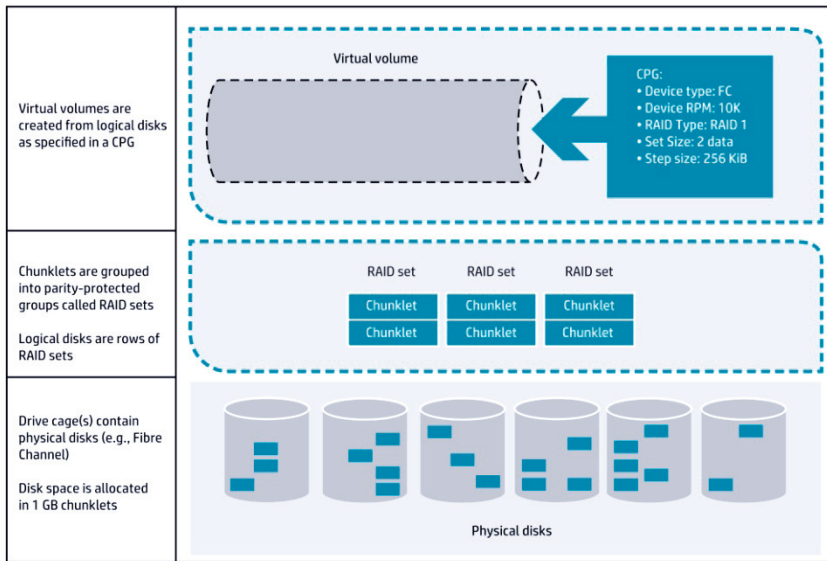
Physical disks

The HP 3PAR Operating System divides physical drives into several equally sized allocations called chunklets. Chunklet size is 1 GB for the HP 3PAR StoreServ 10000 and StoreServ 7000; chunklets are 256 MB in the F-Class. Each chunklet can be viewed as its own small disk.

With HP 3PAR StoreServ, RAID sets are constructed from chunklets on separate physical drives throughout the array, not from whole drives. Different chunklets on a physical drive can be used for volumes with different RAID levels.

RAID levels

The chunklet-based approach employed by HP 3PAR StoreServ allows all RAID levels to coexist on the same physical drives, allowing the use of the optimal RAID level for each volume. HP 3PAR StoreServ supports the following RAID types:

- RAID 10 (RAID 1)
- RAID 50 (RAID 5)
- RAID Multi-Parity (MP) or RAID 6

While all storage vendors offer most of these RAID levels in some form, the key difference is that in HP 3PAR StoreServ, RAID protection is not at the spindle level but at the chunklet level. Depending on the storage administrator's choice, the HP 3PAR array selects chunklets in such a way that the array continues to be available even if an entire disk cage goes offline.

RAID 5 Performance

HP 3PAR StoreServ combines the HP 3PAR ASIC, a battery-backed memory cache, and wide striping that reduces spindle contention. As a result, RAID 5 in HP 3PAR StoreServ arrays offers performance approaching that of RAID 1 in traditional arrays. In fact, for certain workloads, RAID 5 in an HP 3PAR StoreServ array can provide higher performance than RAID 1.

This minimizes the performance impact typical of RAID 5 on legacy storage architectures. The write-back cache in HP 3PAR StoreServ allows sequential writes to be collected until a full parity group can be written, reducing disk I/O traffic and possible back-end bottlenecks. RAID 5 is also appropriate for volumes that are dominated by read activity.

Best practice for RAID 5

HP 3PAR StoreServ allows selection of the number of data blocks per parity block (n+1) to suit different needs. For RAID 5, 3+1 is the default, but any value from 2+1 to 8+1 can be selected. Higher values of "n" result in higher storage efficiency but can reduce the performance of random writes.

Logical disks

Logical disks (LDs) are rows of chunklets in RAID-protected sets. Logical disks are mapped to virtual volumes. The size and number of logical disks in a virtual volume will vary depending upon the provisioning and size requirements of the volume.

Virtual volumes

The virtual volume is the only data layer that is visible to the host, and it is visible only after it has been exported to the host. Virtual volumes can be any size, from 256 MB up to 16 TB each. In HP 3PAR StoreServ, two types exist: the fully provisioned virtual volume (FPVV), and the thinly provisioned virtual volume (TPVV).

Common provisioning groups

The common provisioning group (CPG) is fundamental to administration and reporting of HP 3PAR StoreServ. It allows virtual volumes to share the CPG's resources and allocate space on demand. A CPG can automatically grow its space through mapping the logical disk space.

CPGs pre-define many of the attributes a virtual volume will have when it is created. Some of these attributes include:

- RAID level (for example, RAID 5, RAID 6, etc.)
- RAID set size
- Physical drive type
- Drive speed
- Availability (the level of hardware fault tolerance)
- Other attributes

When a virtual volume is created using a particular CPG, the virtual volume will include the properties specified in the CPG.

Availability

The combination of RAID level and physical drive type specified by the CPG determines the service level and availability level of a volume. Availability refers to the level of hardware fault tolerance; in a sense, availability is what occurs in case of failure. For example, a volume with cage-level availability can tolerate the failure of an entire drive cage.

Provisioning

CPGs automatically provision logical disk capacity. They enable fine-grained, shared access to pooled logical disk capacity.

A CPG can contain both fully provisioned virtual volumes (FPVVs) and thinly provisioned virtual volumes (TPVVs) that draw space from the CPG logical disk pool.

Instead of pre-dedicating logical disks to volumes, the CPG allows multiple volumes to share the buffer pool of logical disks. For example, when a TPVV is running low on user space, the system automatically assigns more capacity to the TPVV by mapping new regions from logical disks in the CPG that are associated with that TPVV. As a result, any large pockets of unused but allocated space are eliminated.

A fully provisioned virtual volume (FPVV) does not add capacity automatically. Instead, FPVVs are allocated when they are created. This differing allocation strategy does not prevent TPVVs and FPVVs from coexisting in the same CPG.

By default, a CPG is set to auto-grow new logical disks whenever the amount of available logical disk space falls below a configured threshold. The initial buffer pool of logical disks starts off at a fraction of the exported virtual capacity of mapped volumes and automatically grows over time as required by application writes.

Best practices for configuring CPGs for Oracle

In Oracle environments, CPGs are used to create virtual volumes that store Oracle databases. Follow these guidelines when configuring CPGs for an Oracle environment:

- Consider these factors to help determine the number and types of CPGs:
  - CPG availability settings define the level of hardware fault tolerance.
  - Use CPGs with solid-state drives for Virtual Volumes that require the highest performance.
  - CPGs with nearline (NL) drives in RAID 6 are ideal for rarely accessed archive redo logs.
  - CPGs with Fibre Channel physical disks are the best default choice for Oracle data files.
  - All members of an Oracle Automatic Storage Manager (ASM) disk group should be from the same CPG.
- Use CPGs for reporting on the storage space consumed by each Oracle database.

- When creating the CPGs in a real production environment, use CPG names that broadly express the key attributes of various groups. This will help the storage administrator to maintain the environment over time.

General recommendations for Oracle on HP 3PAR StoreServ

HP 3PAR StoreServ provides storage redundancy and performance options to the Oracle host. ***The redundancy is best performed at the array level, rather than using the redundancy feature of the Oracle ASM***. Consider the points below when designing your Oracle environment.

- Refer to the HP 3PAR StoreServ implementation guide for operating system details, multipath software specifics, and array port configuration.

- Create Oracle ASM disk groups with the external redundancy option, which offloads the work of maintaining volume availability to the array.

- Point archive log destinations to ASM disk groups created from NL drives if available. Partition hot tables onto solid-state disk (SSD) disk groups if I/O latency becomes an issue.

- Use RAID 5 for Oracle data files. In HP 3PAR StoreServ, RAID 5 approaches RAID 1 performance. Find more information about 3PAR StoreServ RAID in the HP 3PAR StoreServ RAID white paper.

- When HP 3PAR Virtual Copy is used for backup or replication, data files and archive logs should exist on separate volumes. Control files should reside on a volume separate from the archive logs and data files.

- HP 3PAR StoreServ arrays allow volumes to be placed in consistency groups. This grouping helps ensure that snap operations maintain the same point-in-time write consistency when virtual copies are created. Put all the volumes in an ASM disk group into the same HP 3PAR StoreServ consistency group when using HP Virtual Copy or Remote Copy. This will help ensure that the volumes are copied with the same point-in-time image.

Linux fdisk partitioning

Environments that require optimal random write performance, such as those for online transaction processing (OLTP), should align disk partitions to the RAID set.
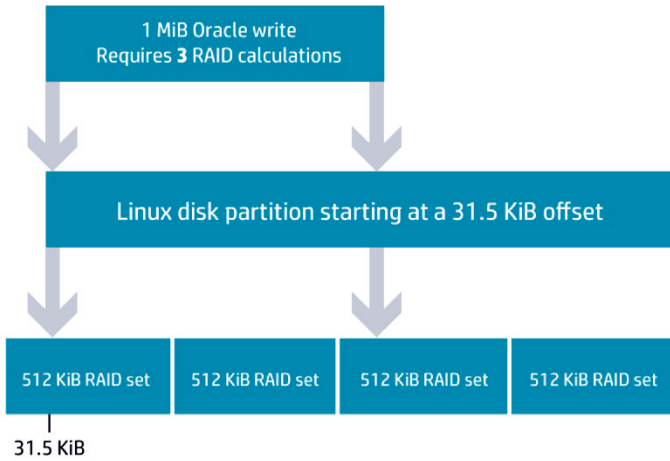
Background

All RAID devices, including HP 3PAR arrays, incur the overhead of calculating new parity after a write. A single write that spans two RAID sets will require two parity calculations. Aligning Linux disk partitions to the HP 3PAR RAID set reduces the frequency of RAID parity calculation, resulting in improved performance.

Best-practice example: Linux fdisk partitioning

Figure 3 (below) shows how a 1 MiB Oracle write spans three HP 3PAR StoreServ RAID sets because the Linux partition is offset by 31.5 KiB. The result of this misalignment is that one Oracle write requires three RAID sets to recalculate parity.

In figure 4, the Linux partition has been created at an offset equal to the size of one RAID set (512 KiB). This aligns the Oracle writes to the RAID set boundary—resulting in only two parity recalculations—which can improve random write performance by as much as 20 percent.

Figure 3. A partition table not aligned to the RAID set

1 MiB Oracle write
Requires **3** RAID calculations

Linux disk partition starting at a 31.5 KiB offset

512 KiB RAID set 512 KiB RAID set 512 KiB RAID set 512 KiB RAID set

31.5 KiB

Figure 4. A partition table aligned to the RAID set

1 MiB Oracle write
Requires **2** RAID calculations

Linux disk partition starting at a 512 KiB offset

512 KiB RAID set 512 KiB RAID set 512 KiB RAID set 512 KiB RAID set
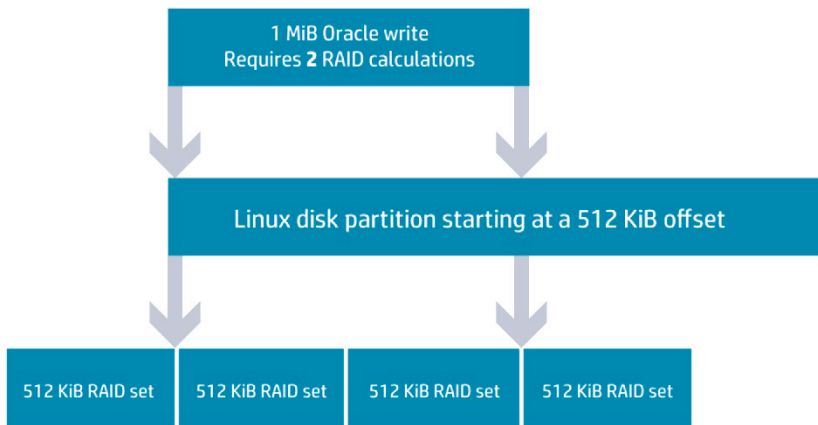
Procedure

Determine the best offset of the disk partition by multiplying the number of data chunklets in the RAID set by the step size. Then create the first OS disk partition at that offset.

1. Use the HP 3PAR StoreServ Management Console to find the Set Size and Step Size for the virtual volume. For example, in figure 5 there are four data chunklets in the set.
2. Find the size of the data stripe in bytes by multiplying the number of data chunklets by the Step Size. In the example, there are 131072 bytes in the step size (128 KiB x 1024). Thus, 4 x 131072 = 524288 bytes in a data stripe.

Figure 5. HP 3PAR StoreServ Management Console

3. Use the command `fdisk -lu <device file>` to identify the allocation unit for partitions. As shown in figure 6, the value is 512 bytes.
4. Then calculate the number of partition units in a data stripe. For example, 524,288/512 = 1,024.
5. Then run `fdisk` with the `-u` option. Create the first partition at the calculated offset (figure 7).

Figure 6. Fdisk allocation units

```
[root@winrac02 ~]#
[root@winrac02 ~]#
[root@winrac02 ~]# fdisk -lu /dev/mapper/mpath37

Disk /dev/mapper/mpath37: 16.1 GB, 16106127360 bytes
255 heads, 63 sectors/track, 1958 cylinders, total 31457280 sectors
Units = sectors of 1 * 512 = 512 bytes

                Device Boot      Start         End      Blocks   Id  System
[root@winrac02 ~]#
```

Figure 7. Creating a partition with offset

```
[root@winrac02 ~]#
[root@winrac02 ~]#
[root@winrac02 ~]# fdisk -u /dev/mapper/mpath37

The number of cylinders for this disk is set to 1958.
There is nothing wrong with that, but this is larger than 1024,
and could in certain setups cause problems with:
1) software that runs at boot time (e.g., old versions of LILO)
2) booting and partitioning software from other OSs
   (e.g., DOS FDISK, OS/2 FDISK)

Command (m for help): n
Command action
   e   extended
   p   primary partition (1-4)
p
Partition number (1-4): 1
First sector (63-31457279, default 63): 1024
Last sector or +size or +sizeM or +sizeK (1024-31457279, default 31457279):
Using default value 31457279

Command (m for help): w
The partition table has been altered!
```

Thin suite management tools

HP 3PAR Thin Suite is a comprehensive set of tools for managing thin storage. HP 3PAR Thin Suite software includes:

- Thin Provisioning—Allows for consumption of storage capacity only when needed
- Thin Conversion—Converts fat, legacy capacity to thin, efficient capacity
- Thin Persistence—Keeps thin capacity thin over time

Each of these software tools can be valuable in an Oracle environment.

Using 3PAR StoreServ Thin Provisioning with Oracle

HP 3PAR Thin Provisioning allows creation of large virtual volumes (VVs) without immediately dedicating physical storage. Instead, the array physically allocates the disk space only when the server writes data to the volume.

Best practices for thin provisioning

Use HP 3PAR Thin Provisioning to optimize storage space utilization. In an Oracle environment, this feature provides the highest value when used in the following cases:

- **Autoextend-enabled table spaces.** Oracle ASM-managed data files with the autoextend feature grow as the need for table space grows. Using such auto-extendable data files on thin provisioned volumes allows the system to allocate disk space to a database as it grows.

  However, Oracle's process of extending a table space is I/O intensive and can affect the performance of the database during the file extension.

- **Best practice**: Lessen the frequency of the performance impact by increasing the increment of space that Oracle adds (the `AUTOEXTEND ON NEXT` parameter of the `CREATE TABLESPACE` command). Set autoextend to a larger value for rapidly growing data files, and a smaller value if the data is growing at a slower rate.

- **ASM disk groups as archive log destinations**. The combination of Oracle ASM and HP 3PAR Thin Provisioning enables an expanding ASM archive log destination. Oracle databases become inaccessible when the archive log destinations fill up.

- **Best practice:** Put Oracle archive logs on thin provisioned ASM disk groups, which allows the underlying storage to self-tune, accommodating unexpected increases in log switch activity. After a level 0 backup, remove the archive logs and use the Oracle ASM Storage Reclamation Utility (ASRU, described later) to free up the storage that was allocated to the old logs.

When not to use thin provisioning

HP 3PAR Thin Provisioning is not always recommended for certain workloads or applications. The following examples illustrate cases in which the use of thin provisioning may not be the best option:

- **Data files residing on file systems.** During the formatting of a file system, the entire expanse of the underlying volume is written to. This causes a TPVV to provision all of its space, nullifying the value of thin provisioning.

- **Systems with high file system utilization.** If the file systems or ASM disk groups are full, then the benefits of thin provisioning are reduced. Consider any ASM disk group or file systems with utilization rates above 80 percent as inappropriate for use with thin provisioning. In this case, it may be more efficient to use fully provisioned virtual volumes to hold this data.

- **Oracle data files that are not in "autoextend" mode.** Oracle databases write format information to data files during table space creation. This has the same effect as provisioning file systems with high utilization and may be inefficient depending upon the ratio of provisioned storage to database size.

Using HP 3PAR Thin Conversion with Oracle ASM

HP 3PAR Thin Conversion migrates data from a fully provisioned virtual volume (FPVV) to a thinly provisioned virtual volume (TPVV). Such a migration can markedly increase storage efficiency.

Background

During migration, the zero-detect capabilities of the HP 3PAR StoreServ array allow only non-zero data on the TPVV. Then the TPVV is swapped for the FPVV in its presentation to the host OS. This process requires Oracle to stop I/O to the source FPVV, causing the data to be unavailable during the swap of the volumes.

Schedule a maintenance window that will accommodate these steps:

1. Stop Oracle processes from accessing the FPVV and exporting it from the host.

2. Create a physical copy of the FPVV on a newly create TPVV.

3. Export the TPVV to the host.

4. Configure the volume into the host OS.

5. Restart Oracle processes as needed.

Best practice for migrating data from an FPVV to a TPVV

- Use the UNIX `dd` command or a similar utility to write zeros to the unused space on the FPVV. Storage space that is populated with zeros is excluded from the FPVV-to-TPVV migration. Calculating the correct offset to begin writing zeros will vary by implementation.

- Consider creating a complete copy of the FPVV on the TPVV. Do not write zeros to the FPVV. Then swap the TPVV for the FPVV and bring the disk group online.

- To thin out the TPVV, use the Oracle ASRU discussed below. The ASRU is a safer option than manually zeroing the unused space. Writing zeros to an Oracle ASM volume without taking the proper preliminary steps, including reduction of the disk group size, will corrupt the disk group.

- Use a host-based utility or the HP 3PAR Physical Copy feature to copy data onto the TPVV. After creating the physical copy, promote it and swap it with the FPVV.

Using HP 3PAR Thin Persistence with the Oracle ASRU for online space reclamation

Oracle and HP 3PAR StoreServ offer the ability to improve storage efficiency for Oracle Database 11gR2 environments by reclaiming unused (but allocated) ASM disk space in thin-provisioned environments.

Background

This extension of storage efficiency is enabled by two recent innovations:

- **Oracle ASM Storage Reclamation Utility (ASRU)**—The Oracle ASRU is a standalone utility that enables reclamation of significant percentages of storage that were once allocated to dropped databases or data files. The ASRU is effective at space reclamation and can be used with HP 3PAR StoreServ with Thin Provisioning. The utility compacts ASM disks, writes zeroes to the free space, and resizes the ASM disks to original size. It accomplishes these tasks with a single command, online and without disruption. For a detailed description of this utility, refer to Oracle's ASRU-3PAR white paper.

- **HP 3PAR Thin Persistence**—HP 3PAR Thin Persistence software detects zero writes and eliminates the capacity associated with unused space in thin-provisioned volumes, also without disruption. HP 3PAR Thin Persistence leverages the unique, built-in, zero-detection capabilities of the HP 3PAR Gen4 ASIC in HP 3PAR StoreServ systems.

Best-practice example: Reclaiming storage from a dropped Oracle 11gR2 Database

To illustrate storage reclamation using the Oracle ASRU and the HP 3PAR StoreServ system, consider this example: a 4 TB Oracle ASM disk group, named 'TP_R5_FC_OLTP_DATA', created using eight 500 GB TPVVs on a 3PAR StoreServ HP 10400 Storage system.

The query of Oracle ASM in figure 8 shows that the three Oracle 11gR2 databases initially use 71 percent (2,935,092 MB) of the 4,096,000 MB of available storage in the ASM disk group. Table 1 shows the initial database configurations.

Figure 8. Disk utilization in example Oracle ASM Disk Group 'TP_R5_FC_OLTP_DATA'

```
SQL> select name, state, type, total_mb, free_mb from v$asm_diskgroup where name = 'TP_R5_FC_OLTP_DATA';

NAME                            STATE       TYPE    TOTAL_MB   FREE_MB
------------------------------- ----------- ------- ---------- ----------
TP_R5_FC_OLTP_DATA              MOUNTED     EXTERN    4096000    1160908
```

Table 1. Database configurations in Oracle ASM Disk Group 'TP_R5_FC_OLTP_DATA'

Database name	Size
db5	1 TB
db7	512 GB
db8	512 GB

To enable the reclamation of space Oracle is no longer using:
1. At the HP 3PAR StoreServ CLI, enable zero detection for the TPVVs as follows:
   cli% setvv –polzero_detect <*virtual volume name*>

2. Use the HP 3PAR StoreServ CLI command showvv to show the physical storage on the HP 10400 Storage system, as illustrated in figure 9.

Figure 9. Physical storage on the HP 10400 Storage system, using CLI command showvv

```
DSME-V400-GD04 cli% showvv -s *R5-OLTP*
                            ---Adm--- ---------Snp---------- ------------Usr-------------
                            --(MB)--- --(MB)--- -(% VSize)-- -----(MB)------ -(% VSize)-- ------(MB)------
 Id Name             Prov Type Rsvd Used Rsvd Used Used Wrn Lim    Rsvd    Used Used Wrn Lim Tot_Rsvd   VSize
311 AO-TP-FC-R5-OLTP.0 tpvv base  512  226    0    0  0.0  --  --  367616  366358 71.6  95   0   368128  512000
312 AO-TP-FC-R5-OLTP.1 tpvv base  512  226    0    0  0.0  --  --  367616  366396 71.6  95   0   368128  512000
313 AO-TP-FC-R5-OLTP.2 tpvv base  512  226    0    0  0.0  --  --  367616  366433 71.6  95   0   368128  512000
314 AO-TP-FC-R5-OLTP.3 tpvv base  512  226    0    0  0.0  --  --  367616  366374 71.6  95   0   368128  512000
315 AO-TP-FC-R5-OLTP.4 tpvv base  512  226    0    0  0.0  --  --  367616  366449 71.6  95   0   368128  512000
316 AO-TP-FC-R5-OLTP.5 tpvv base  512  226    0    0  0.0  --  --  367616  366407 71.6  95   0   368128  512000
317 AO-TP-FC-R5-OLTP.6 tpvv base  512  226    0    0  0.0  --  --  367616  366383 71.6  95   0   368128  512000
318 AO-TP-FC-R5-OLTP.7 tpvv base  512  226    0    0  0.0  --  --  367616  366378 71.6  95   0   368128  512000
-----------------------------------------------------------------------------------------------------------
  8 total                        4096 1808    0    0                2940928 2931178                 2945024 4096000
```

3. The 512 GB database db8 is dropped in Oracle (dropping db8 is not shown in this example). As shown in figure 10, a query in Oracle ASM verifies that the space allocated to the db8 database has been returned to ASM's free space.

Figure 10. Disk utilization in Oracle ASM Disk Group 'TP_R5_FC_OLTP_DATA', after dropping a 512 GB database

```
SQL> select name, state, type, total_mb, free_mb from v$asm_diskgroup where name = 'TP_R5_FC_OLTP_DATA';

NAME                            STATE       TYPE    TOTAL_MB   FREE_MB
------------------------------- ----------- ------- ---------- ----------
TP_R5_FC_OLTP_DATA              MOUNTED     EXTERN   4096000   1900216
```

4. Use showvv to again view the physical storage allocation on the HP 10400 Storage system. As shown in figure 11, the physical storage allocation has barely changed.

Figure 11. Physical storage on the HP 10400 Storage system, using showvv

```
DSME-V400-GD04 cli% showvv -s *R5-OLTP*
                            ---Adm--- ---------Snp---------- ------------Usr-------------
                            --(MB)--- --(MB)--- -(% VSize)-- -----(MB)------ -(% VSize)-- ------(MB)------
 Id Name             Prov Type Rsvd Used Rsvd Used Used Wrn Lim    Rsvd    Used Used Wrn Lim Tot_Rsvd   VSize
311 AO-TP-FC-R5-OLTP.0 tpvv base  512  226    0    0  0.0  --  --  367616  366358 71.6  95   0   368128  512000
312 AO-TP-FC-R5-OLTP.1 tpvv base  512  226    0    0  0.0  --  --  367616  366396 71.6  95   0   368128  512000
313 AO-TP-FC-R5-OLTP.2 tpvv base  512  226    0    0  0.0  --  --  367616  366433 71.6  95   0   368128  512000
314 AO-TP-FC-R5-OLTP.3 tpvv base  512  226    0    0  0.0  --  --  367616  366375 71.6  95   0   368128  512000
315 AO-TP-FC-R5-OLTP.4 tpvv base  512  226    0    0  0.0  --  --  367616  366449 71.6  95   0   368128  512000
316 AO-TP-FC-R5-OLTP.5 tpvv base  512  226    0    0  0.0  --  --  367616  366407 71.6  95   0   368128  512000
317 AO-TP-FC-R5-OLTP.6 tpvv base  512  226    0    0  0.0  --  --  367616  366383 71.6  95   0   368128  512000
318 AO-TP-FC-R5-OLTP.7 tpvv base  512  226    0    0  0.0  --  --  367616  366378 71.6  95   0   368128  512000
-----------------------------------------------------------------------------------------------------------
  8 total                        4096 1808    0    0                2940928 2931179                 2945024 4096000
```

5. As shown in figure 12, execute ASRU on the RHEL 6.2 host. This will reclaim the unused space from the TPVVs (with zero-detect enabled) on HP 10400 Storage system.

Figure 12. Execution of ASRU on the host to reclaim the unused space on the HP 10400 Storage system

```
[oracle@apsi144-127 ASRU]$ bash ASRU TP_R5_FC_OLTP_DATA
Checking the system ...done
Calculating the sizes of the disks ...done
Writing the data to a file ...done
Resizing the disks...done
Calculating the sizes of the disks ...done

/u01/app/oracle/product/11.2.0/grid/perl/bin/perl -I /u01/app/oracle/product/11.2.0/grid/perl/lib/5.10.0 /home/oracle/ASRU/zerofill 1 /dev/mapper/TP-R5-FC-OLTP.4 343089
 512000 /dev/mapper/TP-R5-FC-OLTP.7 343101 512000 /dev/mapper/TP-R5-FC-OLTP.6 343093 512000 /dev/mapper/TP-R5-FC-OLTP.5 343109 512000 /dev/mapper/TP-R5-FC-OLTP.2 343093
 512000 /dev/mapper/TP-R5-FC-OLTP.0 343081 512000 /dev/mapper/TP-R5-FC-OLTP.1 343089 512000 /dev/mapper/TP-R5-FC-OLTP.3 343097 512000
168911+0 records in
168911+0 records out
177116020736 bytes (177 GB) copied, 179.806 s, 985 MB/s
168899+0 records in
168899+0 records out
177103437824 bytes (177 GB) copied, 350.164 s, 506 MB/s
168907+0 records in
168907+0 records out
177111826432 bytes (177 GB) copied, 209.973 s, 843 MB/s
168891+0 records in
168891+0 records out
177095049216 bytes (177 GB) copied, 249.434 s, 710 MB/s
168907+0 records in
168907+0 records out
177111826432 bytes (177 GB) copied, 188.085 s, 942 MB/s
168919+0 records in
168919+0 records out
177124409344 bytes (177 GB) copied, 209.229 s, 847 MB/s
168911+0 records in
168911+0 records out
177116020736 bytes (177 GB) copied, 269.552 s, 657 MB/s
168903+0 records in
168903+0 records out
177107632128 bytes (177 GB) copied, 181.948 s, 973 MB/s

Calculating the sizes of the disks ...done
Resizing the disks...done
Calculating the sizes of the disks ...done
Dropping the file ...done
[oracle@apsi144-127 ASRU]$
```

6. After completion of ASRU, verify the reclamation of storage on HP 10400 Storage system, as shown in figure 13.

Figure 13. Physical storage on the HP 3PAR V-Class Storage system after reclamation

```
DSME-V400-GD04 cli% showvv -s *R5-OLTP*
                          ---Adm--- ---------Snp---------- ------------Usr-------------
                          --(MB)--- --(MB)--- -(% VSize)-- -----(MB)------ -(% VSize)-- ------(MB)------
 Id Name          Prov Type Rsvd Used Rsvd Used Used Wrn Lim    Rsvd    Used Used Wrn Lim Tot_Rsvd   VSize
311 AO-TP-FC-R5-OLTP.0 tpvv base 512  226    0    0  0.0  --  --  350464  342751 66.9  95   0  350976  512000
312 AO-TP-FC-R5-OLTP.1 tpvv base 512  226    0    0  0.0  --  --  350464  342681 66.9  95   0  350976  512000
313 AO-TP-FC-R5-OLTP.2 tpvv base 512  226    0    0  0.0  --  --  350464  342698 66.9  95   0  350976  512000
314 AO-TP-FC-R5-OLTP.3 tpvv base 512  226    0    0  0.0  --  --  350208  342763 66.9  95   0  350720  512000
315 AO-TP-FC-R5-OLTP.4 tpvv base 512  226    0    0  0.0  --  --  350208  342742 66.9  95   0  350720  512000
316 AO-TP-FC-R5-OLTP.5 tpvv base 512  226    0    0  0.0  --  --  350336  342623 66.9  95   0  350848  512000
317 AO-TP-FC-R5-OLTP.6 tpvv base 512  226    0    0  0.0  --  --  350592  342728 66.9  95   0  351104  512000
318 AO-TP-FC-R5-OLTP.7 tpvv base 512  226    0    0  0.0  --  --  349952  342751 66.9  95   0  350464  512000
-----------------------------------------------------------------------------------------------------------
  8 total                    4096 1808    0    0                 2802688 2741737               2806784 4096000
```

This example shows that 189,442 MB of storage was successfully reclaimed. Without ASRU, all of the space originally allocated to db8 would be stranded in this volume group. Running ASRU makes over a third of that space available for other applications. Using ASRU with 3PAR Store Serv's zero-detect enabled thin provisioning frees stranded storage, significantly improving efficient utilization of storage resources over the life of your Oracle Environment.

Adaptive optimization

An important business need in IT environments is the ability to optimize storage arrays and control Oracle database transaction times by utilizing storage tiers. This holds true for most applications and is particularly important in an Oracle environment, where the database may be running several mixed workloads that require specific storage optimization.

Mixing workloads can create performance challenges in a complex database environment. With OLTP workloads, for example, placing the most active data on the fastest tier and idle data on the slowest tier is very important. This allows control of the latency associated with database I/Os that cannot be tuned using Oracle tuning parameters alone (such as adjusting the database buffer cache). In decision support system (DSS) database environments, the goal is different: administrators want to execute queries within a minimum timeframe by effectively utilizing the available storage tiers.

HP 3PAR Optimization Suite Software allows moving frequently accessed data to the fastest tier and infrequently accessed data to the slower tier. The optimization suite delivers the powerful capabilities of HP 3PAR Adaptive

Optimization, Dynamic Optimization, and System Tuner software combined into a single suite. Using this software, a substantial amount of storage-level optimization can be achieved on HP 3PAR StoreServ systems. Adaptive Optimization can autonomically tune the running environment, even under the varied requirements of a mixed workload, delivering optimal performance for even the most complex combination of I/O performance demands.

Best practices for HP 3PAR Adaptive Optimization

The ability to establish policy-driven service levels and perform data movement at a sub-volume level enables flexibility in performance tuning. HP 3PAR Adaptive Optimization software works as an enabler to react swiftly to changing business needs while delivering service-level optimization over the entire application lifecycle. The software intelligently monitors sub-volume performance, then applies user-specified policies to autonomically and nondisruptively move data, aligning shifting quality of service (QoS) demands at a granular level. Support for multiple adaptive optimization profiles and for the co-existence of both tiered and non-tiered application volumes provides the flexibility to consolidate a wide range of applications onto a single array.[1]

Background

HP 3PAR Adaptive Optimization Software will work with any drive type in the array. These can be:

- SSDs—Solid-state disks; small capacity, high performance

- FC—Fibre Channel disks; medium capacity, medium performance

- NL—Nearline disks or storage; large capacity, low performance

A "tier" is a CPG of any combination of drive type and/or RAID type. HP 3PAR AO optimizes cost/performance using autonomic, sub-volume granular tiering, thereby using each drive tier where it is best suited.

For adaptive optimization, two or more tiers are required. The user can assign either two (minimum) or three (maximum) storage tiers (identified by a 0, 1, or 2) for data migration in a given configuration, where:

- Tier 0 is the fastest.

- Tier 1 is the intermediate speed.

- Tier 2 is the slowest.

Figure 14 illustrates the HP 3PAR Adaptive Optimization architecture, showing the different drive types and how chunklets from those drives are mapped in a typical volume. RAID sets made up of SSD chunklets are mapped to a Tier 0 logical disk, while RAID sets composed of NL storage chunklets are assigned to a less frequently accessed Tier 1 logical disk.

---

[1] For a more detailed discussion of tiering and adaptive optimization best practices, see the Autonomic Tiering with Oracle white paper.

Figure 14. HP 3PAR Adaptive Optimization architecture
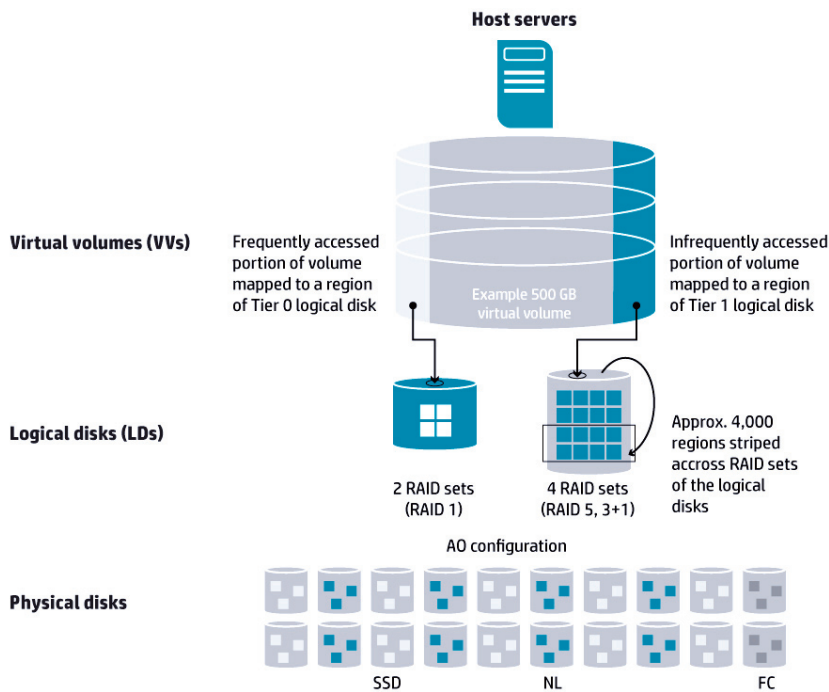
Adaptive optimization implementation

In order to implement tiering, HP 3PAR Adaptive Optimization needs to do several things:

1. Collect historical data of accesses for all regions in an array.
2. Analyze the data to determine the volume regions that should be moved between tiers.
3. Instruct the array to move the regions from one CPG (tier) to another.
4. Provide the user with reports showing the impact of adaptive optimization.

The separately licensed HP 3PAR System Reporter periodically collects detailed performance and space data from HP 3PAR StoreServ, stores the data in a database, analyzes the data, and generates reports. System Reporter enables HP implemented storage level optimization by collecting region-level performance data, performing tiering analysis, and issuing region movement commands to the array.

Tiering and data movement

Figure 15 shows how HP 3PAR Adaptive Optimization conducts autonomic tiering and data movement at the sub-volume level. In this illustration, several virtual volumes constitute the Oracle data files configured in an Oracle ASM disk group.
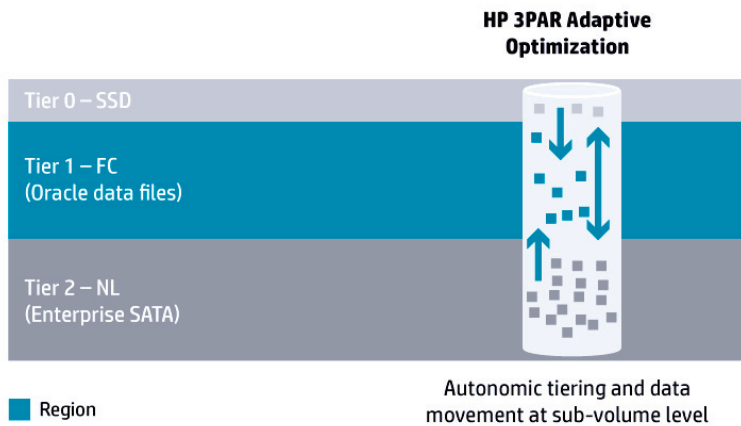
HP 3PAR Adaptive Optimization does not care about the virtual volumes. The software analyzes and moves regions independently of which VV they belong to.

Best-practice example: configuring Adaptive Optimization

Use the HP 3PAR System Reporter for configuring adaptive optimization. Multiple AO configurations can be defined for an HP 3PAR StoreServ array, in which each configuration corresponds to a related set of hosts or applications that require optimization. Figure 16 shows an example of configuration settings.

NOTE: This example is for a 3PAR StoreServ 10000 with 3.1.1 firmware.

Figure 16. HP 3PAR Adaptive Optimization configuration settings

Choosing the tier parameters

1. Choose the system.
2. Choose the CPG for each tier.
3. Choose the Tier Size. Tiers are defined by CPGs. All of the cost and performance characteristics of the tier are determined by the settings for the CPG, such as the RAID level, number of disks used, disk type, and speed. The user can also control the maximum space available for each tier.

   Setting a Tier Size limits the amount of space that the algorithm will use in each tier. Set the tier size field to 999999 for an unlimited tier size. Note that adaptive optimization will attempt to honor this size limit in addition to any warning or hard limit specified in the CPG.

4. Specify the schedule, including Date, Week Day, Hour, and Measurement Hours. The user can specify the schedule when an Adaptive Optimization configuration will execute during an OLTP or DSS workload, along with the measurement hours preceding the execution time. This allows scheduling data movement at times when the additional overhead of that data movement is acceptable (for example, during non-peak hours). The Schedule and Measurement Hours settings allow the AO policy to be fine-tuned for a specific database workload.

   Control of the measurement hours is important, especially for applications whose performance is not uniformly important throughout the day. For example, consider an Oracle OLTP environment whose performance matters only during the first few hours of a business day (Monday through Friday, 8 am through 11 am) and then is not particularly important throughout the rest of the day (though not necessarily idle). If we choose to measure the application throughout the day, the tiering algorithm will include the performance during times that are not important; as a result, the performance during the important times of the day may not be optimal.

   To achieve best results in this scenario, schedule HP 3PAR Adaptive Optimization execution for 11 am Monday through Friday with a measurement interval of 3 or 4 hours. Then only performance measurements during the important period for this OLTP workload will be considered.

Note

Measurement Hours also control hysteresis. A region will not move more than once in X hours = measurement hours. This prevents ping-ponging between tiers.

5. Choose the mode. Mode configuration can be set to one of three values:
   - Performance—Biased toward moving regions up to higher-performance tiers
   - Cost—Biased toward moving regions down to lower-cost tiers
   - Balanced—Between performance and cost

   The mode configuration parameter does not change the basic flow of the tiering analysis algorithm, but rather it changes certain tuning parameters that the algorithm uses.

6. To dynamically disable adaptive optimization, set Configuration Active to `false`. To re-enable AO, set Configuration Active to `true`. The false setting only disables analysis and data movement; it does not prevent System Reporter from sampling the region counters.

Note

Adaptive Optimization is not real time. It requires a minimum of 3 hours of region density information before attempting to optimize data placement.

Best practices for AO

- Use the HP 3PAR System Reporter (separate license) to collect region-level performance data, perform tiering analysis, and issue region movement commands to the array.
- Do not set the capacity limit at the AO configuration level. When creating the AO configuration through the HP 3PAR System Reporter policies, enter a value of 999,999 GiB (999 TiB) for each tier. This allows HP 3PAR OS provisioning to control the size of the CPG growth, leading to a much more manageable solution for the administrator.
- When using SSDs with AO, create an SSD CPG configuration that will use up to 95 percent of the SSD. This will deliver the highest performance throughput.
- SSDs are small capacity devices. Allocate space to SSD CPGs in small increments; this will prevent the CPG from allocating an excessive amount of valuable SSD storage.

- Modify the CPG grow size to reduce the amount an SSD CPG grows. Use the lowest possible value; for example, 8 GB per controller node pair as shown here:

```
ORA-T400-GC08 cli% setcpg -sdgs 16G SSD-R5-CPG
```

- Make certain that all CPGs used in an AO configuration have the same level of availability, which will help ensure that SLA (service-level agreement) availability requirements are met after data has migrated to a different tier.

- Do not set the growth limits on CPGs. If a warning threshold is required set a growth warning (warning in terms of capacity) and *not* an allocation warning (warning in %).

- The following combinations are acceptable within the same AO configuration policy:
  – SSD, FC, and NL
  – SSD and FC
  – FC and NL

- Using different RAID levels within the same policy is acceptable.

- Configurations that only contain SSD and NL are *not* recommended.

- CPGs for SSDs should use RAID 5, with the minimum growth increment supported (8 GB per controller node pair).

- CPGs for FC disks should use RAID 5, with the default growth increment.

- CPGs for NL disks should use RAID 6, with the default growth increment.

- A different AO configuration and its associated CPGs should be created for approximately every 100 TiB of data. AO supports a maximum of 125 TiB of data per AO configuration.

- Set the AO policy to measure only during the processing for which you would like to optimize, and schedule the data movement during times when the additional overhead of that data movement will not cause any significant impact—for example, during non-peak hours. HP 3PAR Adaptive Optimization will execute each policy serially, but will calculate what needs to be moved simultaneously.

- For an OLTP workload that requires a high level of performance, use tailored AO configurations that:
  – Preferably use all the available storage tiers
  – Are set to use 999,999 GB for each tier
  – Execute immediately at the end of the high activity period
  – Specify Measurement Hours that only cover the length of the high activity period
  – Have the Mode set to Performance

- Use a Mode of Balanced for most Oracle databases.

- HP 3PAR Adaptive Optimization (AO) is not recommended for volume hosting the Oracle database online redo log files.
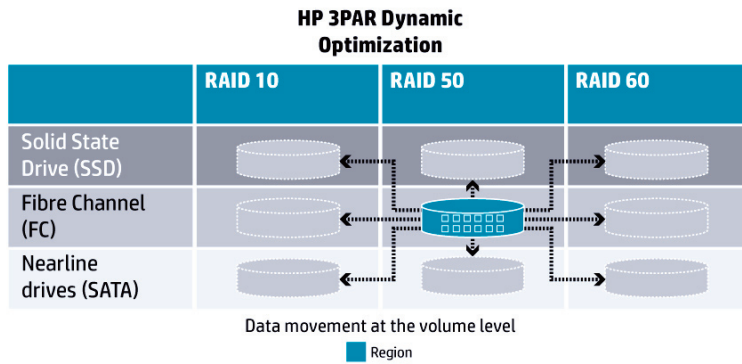
Dynamic optimization

HP 3PAR Dynamic Optimization (DO) is an optional licensed feature. It enables the storage administrator to adjust to changing application and infrastructure requirements, and to nondisruptively alter service levels associated with a storage volume RAID level. This software allows administrators to improve the performance of a virtual volume without interrupting access to the volume.

Background

Common Provisioning Groups define RAID level and disk type, and fast CPGs consisting of RAID 1 sets or SSDs are premium storage real estate. Dynamic optimization alters the service level of a virtual volume by moving it to a different CPG. In HP 3PAR StoreServ, DO provides the flexibility to place data on CPGs that meet current service level requirements without binding the data permanently to that configuration.

Figure 17 shows how dynamic optimization can be used to choose the service levels of a particular volume after it has already been created for the Oracle databases. Changing from FC to Serial ATA (SATA) is supported, for example, as well as different RAID set sizes. Performing this operation does not impact the volume, and changes can be performed while normal database processing is underway. This effectively gives the administrator the ability to alter the service levels when either incorrect performance information has been supplied or when Oracle Enterprise Manager (OEM) shows the volume either under- or over-utilizing the resources that have been provisioned.

General best practices for dynamic optimization

There are four general cases where DO may be desirable:

- **Volume layout changes after hardware upgrades**—Existing virtual volumes only take advantage of resources that were present at the time of volume creation. When a storage server is upgraded by adding cages or disks, the original volume layout may no longer be most effective. DO enables changing the layout of virtual volumes to take advantage of the new hardware.

- **Volume RAID level changes**—Because different RAID levels have varying capacity requirements and offer differing degrees of performance, an administrator might want to convert volumes from one RAID type to another as requirements change.

- **Volume availability level changes**—The availability of a virtual volume determines its level of fault tolerance. For example, a volume with cage-level availability can tolerate the failure of a drive cage because its RAID sets use chunklets from different drive cages. As application and business requirements change, it may be desirable to alter the availability characteristics of existing virtual volumes.

- **Volume service level changes**—In addition to nondisruptively altering RAID and availability levels for a given volume, it may also be useful to change volume parameters (such as disk filtering parameters) applied when the volume was created.

Best-practice example: using dynamic optimization for Oracle

To take advantage of DO in an Oracle environment, place data of like SLAs on the same ASM disk group. Dynamic optimization moves entire volumes, so the granularity of the migrated data will be at the ASM disk group level.

One example of how DO might be used to migrate ASM disk groups:

1. Use slower but low-cost NL disks to store an Oracle ASM disk group that contains data only accessed during month's-end processing.

2. In preparation for the intensive processing at month's end, use DO to migrate the ASM disk group to a faster but higher-cost Fibre Channel CPG in RAID 1.

3. After processing completes and the monthly data is no longer a hot spot, use DO to migrate it back to the archival NL CPG.

Business continuity

HP 3PAR StoreServ has a number of features to help maintain business continuity in Oracle environments. These include HP 3PAR Virtual Copy and HP 3PAR Recovery Manager for Oracle.

HP 3PAR Virtual Copy

HP 3PAR Virtual Copy Software provides instant, point-in-time (PIT) copies of data, allowing the organization to share and protect data from any application (including Oracle) simply and affordably.

Best practices for using Virtual Copy with Oracle

- Mount virtual copies elsewhere to create backup, test, or reporting databases.

- Ensure database write consistency across members of an ASM disk group or a file system by combining Oracle hot backup mode with Virtual Copy consistency groups.

- Guarantee data retention schedules for snapshots with HP 3PAR Virtual Lock Software.

HP Recovery Manager

HP 3PAR Recovery Manager software for Oracle is a highly efficient solution for automatically creating and managing snapshots of Oracle and Oracle RAC databases. The software can handle hundreds of these application-consistent, reservationless PIT snapshots for rapid online recovery.

Best practices for using Recovery Manager with Oracle

- Create and manage Oracle PIT snapshots.

- Present the snapshot images to other Oracle database instances.

- Manage the images from an easy-to-use graphical user interface (GUI) on the host.

- Recover an Oracle database to a known point in time quickly and simply.

- Also use Recovery Manager to speed up other operations, including recovery of the Oracle production server.

For more information, refer to HP 3PAR Recovery Manager for Oracle - Overview & Features

HP StoreOnce Backup

HP StoreOnce Backup systems deliver leading price performance for deduplicated backup of Oracle databases. The disk-based StoreOnce Backup system can be used to automate and consolidate the backup of multiple databases onto a single, rack-mountable device, while improving reliability by reducing errors caused by media handling. For business environments with remote offices or a disaster-recovery site, the StoreOnce Backup system can be used to replicate data to a central data center or remote facility.

Best practices for using HP StoreOnce with Oracle

- For business-critical Oracle databases, migrate or consolidate disparate backup targets to a scalable StoreOnce backup system.

- Migrate to a StoreOnce system to mitigate contention and performance issues found in I/O intensive backups, business, and other applications sharing the same storage array.

- Use existing processes and policies, and Oracle RMAN scripts, on the new platform.

- Use a virtual tape library (VTL) to migrate to a disk-to-disk backup environment while maintaining existing processes used with physical tape.

- Use the recently offered StoreOnce Catalyst to further speed data backup and recovery.

For more information about HP StoreOnce, see: http://www8.hp.com/us/en/products/data-storage/data-storage-products.html?compURI=1225909

HP Data Protector

HP Data Protector is a unified, hardware-based data protection solution. It utilizes an intelligent data management approach to seamlessly protect and harness data based on the meaning of that data, from edge to data center and across physical, virtual, and cloud environments. HP Data Protector offers the flexibility of tri-location deduplication—application source, backup server, and target device—helping organizations to meet their most demanding business SLAs while minimizing overall infrastructure cost. It interacts directly with media devices, and also provides scheduling, media management, network backups, monitoring, and interactive backup.

Best practices for using Data Protector with Oracle

- HP Data Protector can be tightly coupled and fully integrated into Oracle's RMAN backup utility via the Media Management Library (MML), a set of routines that allows reading and writing of Oracle file data to HP Data Protector.

- To test the media management software, use the `sbttest` utility. This utility does not perform an actual backup, but tries to communicate with the MML to determine whether the software is working correctly.

For more information on HP Data Protector, see: http://www8.hp.com/us/en/software-solutions/software.html?compURI=1175640

Summary

The design of the HP 3PAR StoreServ system provides configuration flexibility and high availability to Oracle environments. The HP 3PAR Thin Suite of products delivers the efficient use of storage space, while Adaptive Optimization and Dynamic Optimization provide flexible storage tuning, allowing administrators to seamlessly meet changing SLAs. Thin conversion enables several strategies to move fully provisioned volumes onto thinly provisioned volumes.

Thin provisioning, in combination with Oracle's autoextend feature, streamlines space allocation, enabling incremental, just-in-time scaling of the storage footprint. And the Oracle ASRU utility is an industry-changing tool that frees space that Oracle no longer needs, helping to reduce the overhead associated with stranded storage. Leveraging the best practices outlined in this document will help create an optimized Oracle environment with the flexibility and performance needed to meet the demands of even the most dynamic business environments.

For more information

To read more about HP 3PAR StoreServ and Oracle, refer to the links below:

- 3PAR Linux Implementation Guide
- http://bizsupport2.austin.hp.com/bc/docs/support/SupportManual/c03050683/c03050683.pdf
- http://h41111.www4.hp.com/event/uk/en/pdf/AO_Oct_2011.pdf
- HP 3PAR Dynamic Optimization Software
- http://h20195.www2.hp.com/V2/GetDocument.aspx?docname=4AA4-4392ENW&cc=us&lc=en
- http://www.oracle.com/technetwork/database/oracle-3par-wp-final-0-130057.pdf
- Oracle ASM Storage Reclamation Utility and 3PAR thin Persistence

Learn more at

hp.com/go/3PARStoreServ

Get connected

hp.com/go/getconnected

Current HP driver, support, and security alerts
delivered directly to your desktop