
Ultimate Guide to Oracle Storage

Presented by George Trujillo

George.Trujillo@trubix.com

www.trubix.com



George Trujillo

- Twenty two years IT experience with 19 years Oracle experience.
 - Advanced database solutions such as RAC, Data Guard, Streams, etc.
 - Oracle Fusion Middleware - Application Server, J2EE, ADF, Web Services, SOA.
 - Focus on enterprise management of all tiers of heterogeneous environments.



George Trujillo - Perspective

- IOUG - Board of Directors
- IOUC - Oracle Fusion Council
- Oracle Beta Leadership Council
- President of Trubix Inc. - www.trubix.com
 - Largest library of Oracle education materials in the world outside of Oracle Corporation.
 - Leading-edge education courses in Oracle database and Oracle Fusion (middleware).



Objectives

- Oracle 10g storage options
- Oracle 10g ASM
- Oracle Cluster File system
- Understanding the role of ASM, OCFS and Oracle Clusterware

Evolution of IT Major Costs

- Hardware
- People
- Software
- Storage



Storage Priorities

- Important storage priorities:
 - Capacity
 - Speed
 - Availability
 - Scalability
 - Storage location
 - Backup and recovery requirements
 - Business requirements
 - Flexibility



Oracle 10g Storage Options

- Disk volume managers providing striping and mirroring.
 - Third party vendors Veritas, EMC, ...
 - Oracle 10g Automatic Storage Management
- Network Attached Storage (NAS)
- Standard file systems
- Raw devices

Network Attached Storage

- Network Attached Storage (NAS) is a storage system that uses standard file access protocols like FCP, NFS, CIFS, HTTP and iSCSI.
 - Storage is considered remote from the local systems.
- Features like Dual-parity RAID can use snapshots to support recovery of large systems in hours.



Storage Area Network

- Storage Area Networks (SAN) attach remote storage devices (disk array controllers, etc.) so the operating system sees the devices as local devices.
- SANs can replicate data to a secondary storage array.
 - Secondary arrays can be local (SCSI) or remote (Fiber Channel, FCIP, ...).
- SANs support cloning and snapshotting which allows real-time duplication of LUNs. Cloning create a copy of a LUN in the background. Snapshotting stores the original state of blocks that get changed.



Basics

- Disk controllers provide block level access to high speed, reliable logical hard drives (or LUNs).
- Storage virtualization abstracts logical storage from physical storage.

Storage Best Practices

- Minimize the number of vendors and products. Keep the number of vendors to two or three.
 - Hardware independence, portability will help maintain flexible purchasing and management decisions.
 - Hardware independence and the ability to support new innovations is key.
 - Storage technology is constantly changing.
- Design a simple layout with standard naming conventions. Consistent methods for striping and mirroring will help diagnostics.



Storage Best Practices (2)

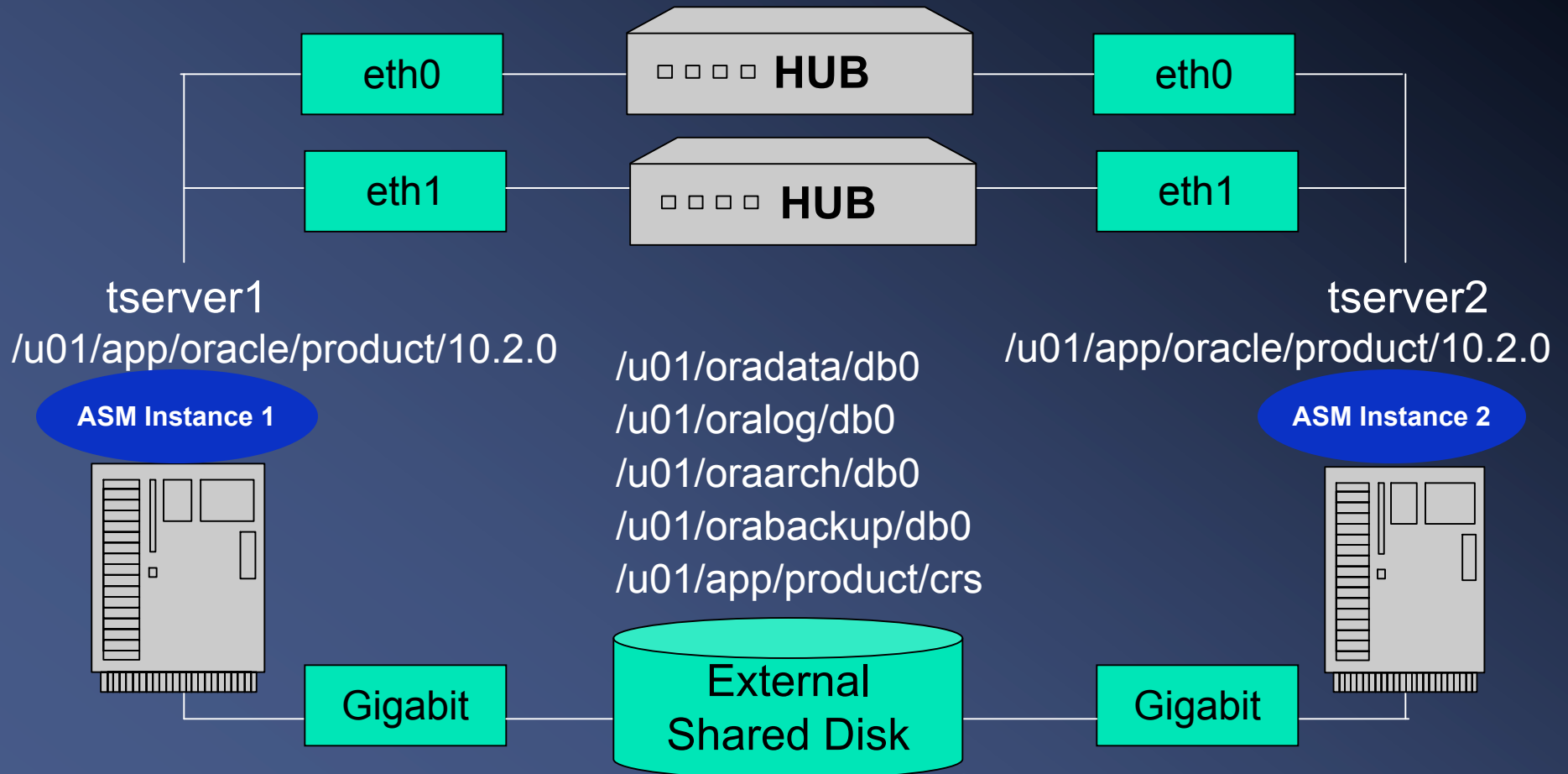
- Maintain standards and best practices. Training is necessary to maintain skill sets.
 - Proper capacity planning is essential.
- Leverage technology for automation, virtualization, replication, Continuous Data Protection (CDP) and disk-to-disk copy and snapshots



Oracle Database Storage

- Oracle database storage requirements include:
 - Striping and mirroring solutions
- Oracle Real Application Clusters (RAC) require shared disk storage.
 - Vendors offer shared file solutions dependent on the operating system version and platform.
 - Oracle offers the Oracle Cluster File System (OCFS) for Linux and Windows platforms.

RAC Environment



Oracle 10g Logical Volume Mgr

- Oracle10G provides its own Logical Volume Manager (LVM) called Automatic Storage Management (ASM) for striping and mirroring.
- An ASM instance is a LVM for Oracle storage.
- ASM uses disk groups to manage raw disks or LUNs.
 - The storage devices are assigned to disk groups.



Benefits of an Oracle LVM

- Benefits of an Oracle LVM (ASM).
 - Improves performance by eliminating O.S. overhead (replaces file systems).
- Free, eliminates need for a 3rd party LVM.
- ASM provides automatic I/O tuning.
- ASM performs online storage configuring.
- Reduced storage maintenance.

ASM Instance

- ASM instance is very lightweight.
 - The database buffer cache (DB_CACHE_SIZE) stores metadata information.
 - The shared pool areas (SHARED_POOL_SIZE) stores extent maps and standard control structures.
 - Use a large pool area (LARGE_POOL_SIZE) for large configurations.

ASM Instance Parameters

- Sample of initialization parameters for an ASM instance.

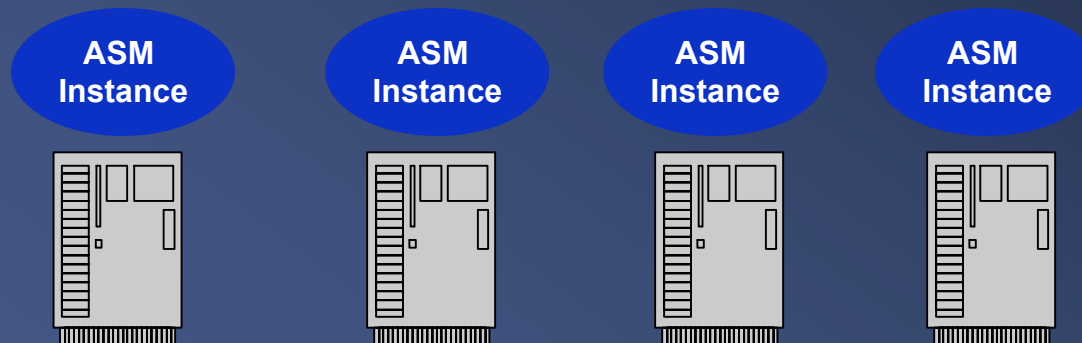
```
*.instance_type=ASM  
*.asm_diskgroups=+DATA  
*.asm_diskstring='/dev/rdisk/c3t19d*s4'  
*.large_pool_size=16M  
*.background_dump_dest='/u04/ora/app/admin/+ASM/bdump'  
*.core_dump_dest='/u04/ora/app/admin/+ASM/cdump'  
*.user_dump_dest='/u04/ora/app/admin/+ASM/udump'
```

ASM Logical Volume Manager

- The ASM instance (logical volume manager) is a shared memory area managed by background processes.
 - Disk groups and their storage are identified during startup.
 - The disk groups are mounted and an extent map is passed to the Oracle instances.
 - Disk groups can be added and removed without taking them offline. Online rebalancing then occurs.
 - Management tools include the Oracle Enterprise Manager (OEM), Database Configuration Assistant (DBCA) and ASMCMD.

ASM Logical Volume Manager (2)

- The ASM instance does not perform I/O. Its responsibility is managing disk groups.
- There is one ASM instance per hardware platform.
 - There can be multiple Oracle database servers per hardware platform.



Disk Groups

- Database files are stored in a disk group.
- Multiple databases can share a disk group.
- Disk group types
 - External - No mirroring.
 - Normal - Two-way mirroring.
 - High Redundancy - Three-way mirroring.



Creating a Disk Group

Create in ASM instance.

```
SQL> CREATE DISKGROUP dk_grp_001  
  FAILGROUP fail_group_01  
    DISK '/dev/disk001', '/dev/disk002',  
  FAILGROUP fail_group_02  
    DISK '/dev/disk011', '/dev/disk012';
```

Create in database instance.

```
SQL> CREATE TABLESPACE fin_data  
  '+dk_grp_001/user_data_01' SIZE 2048M;
```



Using Disk Groups

```
SQL> ALTER database add logfile  
      group 3  
      '+dg3/log3.redo' size 50M;
```

```
SQL> ALTER database datafile  
      '+Dg4/user1.dbf' resize 30M;
```



Who Should Manage a LVM?

- Skill set:
 - Basic skills for managing an Oracle instance.
 - Strong storage fundamentals and a good understanding of the storage environment.

DBA



Storage Mgmt



Unix SA



T R U B I X

www.trubix.com

How do we avoid this?



What skills are needed?

- Storage managers and Unix administrators will need to learn the basics of the Oracle logical volume manager and how to execute commands to manage Oracle storage.
- Oracle DBAs will need to learn more about storage management.

ASM Managing Disk I/O

- ASM writes directly to the storage devices eliminating the operating system overhead.
- ASM writes in one-megabyte stripes versus standard 32K or 64K stripe sizes.
 - Oracle has found this provides high levels of data transfer rates.
- ASM identifies hot spots and moves data extents to balance I/O across all the devices in a disk group.
 - Rebalancing can be managed to ensure it occurs during acceptable time frames. Only one extent is locked at a time.



ASM Disk Groups

- A disk group is collection of raw disks or LUNs (from a storage array) made up of different storage devices.
- Database files are stored in a disk group.
 - Files cannot be spread across disk groups.
 - Files are evenly spread across a disk group.
 - Files can be striped and mirrored across the disks in a disk group.
- A few disk groups can be used to manage Oracle storage versus managing hundreds or thousands of individual data files.



ASM Disks

- ASM disks are raw disks or LUNs assigned to ASM.
- ASM Disks are divided into Allocation Units (AU). ASM files are a collection of AUs (one megabyte).
- Frequently accessed AUs are cached in the Oracle kernel.
- Block zero of each ASM disk contains the disk header. Disk group information is replicated to other ASM disks.

ASM Files

- ASM files can include control files, datafiles, redo logs, flashback logs, archived logs and RMAN backups.
- ASM currently does not support:
 - Operating system files.
 - Oracle software.
 - Trace files and alert logs.
 - RAC OCR and Voting disks.

- Example file name:

`+DATA/orcl/datafile/system.252.363840`



ASM Virtual and Physical Space

- Virtual address space refers to all the AUs (bytes) in an ASM file.
- ASM maps virtual space to the physical storage through extent maps.
- Physical blocks are not spread across allocation units.
- A virtual data extent maps to 1 to 3 data extents dependent on the mirroring.



ASM Striping

- ASM rebalances extents across the storage devices.
- The ASM disks within a Disk Group should be consistent.

ASM Mirroring

- Failure groups are a set of storage devices that define a mirror.
- Reads are performed on the primary mirror group. The primary group should be physically located for performance.
- When an ASM disk fails, The extents from the failed disk are copied from good ASM disks to a free area on the available disks.
- The invalid disk is then dropped from the disk group.

The Cluster Software

- Oracle Clusterware provides the infrastructure that integrates multiple servers so that they operate as a single system.
- The cluster software and Oracle software must be installed in different directories.
 - ORA_CRS_HOME
 - ORACLE_HOME
- The Oracle Clusterware requires a voting disk that records node information and the Oracle Cluster Registry (OCR) that records cluster configuration information. Both require shared storage.



Cluster Software

- Cluster software allows multiple machines to share the same disk.
 - Some vendors provide their own cluster software. This software is vendor and platform specific.
 - Oracle also offers the Oracle Cluster File System that can be used on Linux and Windows platforms.
- Oracle 10g RAC offers Oracle Clusterware previously Cluster Ready Services (CRS).
 - Clusterware software eliminates the need to license cluster software.
 - This provides a standard cluster interface for all platforms.



Oracle Clusterware

- Oracle Clusterware software contains two files that must be shared by all files in a RAC cluster.
 - The Oracle Cluster Registry (OCR) file.
 - The Voting Disk file.
- These two files must be on shared storage.
 - A shared file system using vendor clusterware, OCFS or Oracle Clusterware is recommended.
 - Raw devices can be used but are not recommended.
 - ASM cannot be used.



Cluster Synchronization Service

- Cluster Synchronization Service (CSS)
- ASM uses CSS to synch its data with the DB
- CSS facilitates that synch.
- On RAC, CSS is installed by default.
- On single instance, you have to enable CSS – only do this once, after that it starts up automatically.

```
$ $ORACLE_HOME/bin/localconfig add
```

- Check status to make sure it is happy.

```
$ $OH/bin/crsctl check cssd
```



Creating an OCFS Partition

- Formatting an OCFS partition

```
# mkfs.ocfs -b 128 -g root -L  
oraclevol1 -m /ocfs -p 0755 -u  
root /dev/sda1 -F
```

- **Once the partition is then formatted the DBA may mount the drive as the root super user.**

- **Example command to mount OCFS drive:**

```
# mount -t ocfs /dev/sda1 /u01
```



Oracle Clustered File System

- Independent, open source code.
- Available for Linux and Windows.
- Provides cluster support for a file system.
- DirectIO and asynchIO (OCFS2) support.
- Regular file system – can be for anything.



Trubix - Hot classes

- Ultimate Guide to Oracle Storage
- MySQL Admin for Oracle DBAs
- SQL Server Admin for Oracle DBAs
- Oracle Fusion Admin Workshop I
- Oracle Fusion Admin Workshop II
- Oracle Fusion Development Workshop I
- Oracle Fusion Development Workshop II
- Oracle 10g Grid Admin Workshop



Summary

- Oracle 10g storage options
- Oracle 10g ASM
- Oracle Cluster File system
- Understanding the role of ASM, OCFS and Oracle Clusterware