Oracle ASM
Letting Oracle Manage Your Storage

Simon Pane
First4 Database Partners
Simon.Pane@first4db.com
March 21, 2013

Bio

- Partner at First4 Database Partners
- Working with the Oracle Database since version 6
- Oracle Certified Database Administrator
  - Oracle 8, 8i, 9i, 10g, 11g
- MOS Communities & OTN
  - Username: **`Simon_DBA`**

What is ASM

- Simply put: Oracle's own filesystem
- A filesystem that is not visible outside of Oracle tools
- Not only for RAC – but definitely recommended for RAC

Overview

- First introduced with Oracle 10g (SE and EE)
  - GA: January 2004 - Beta (OSM) versions available for almost a year before
  - No additional licensing
  - Fundamental for the Exadata
- Based on raw devices or unformatted disks
  - A physical disk or partition
  - A disk or partition from a storage array (SAN/NAS)
  - A logical volume
  - Network attached file system (NFS)

ASM Provides

- Simplified storage administration
  - Based on the 9i concept of OMF
  - Software mirroring and striping
  - Software fault tolerance
  - Fast failed disk recovery
  - Hot swappable storage
  - Performance optimization and I/O load balancing
  - Stretch cluster I/O localization
  - Scalability (thousands of disks – Petabytes)

What the SAN Doesn't Give

- Custom RAID & Striping based on file type
- Query configuration & performance through SQL
  - Some views populated on the ASM instance only – others on RDBMS instance as well
    - **`v$asm_disk, v$asm_diskgroup, v$asm_file, v$asm_disk_iostat`**
- Ability to hot swap SANs!
  - Without issuing a single DB command

Usage

- Can (but doesn't have to) be stored in ASM:
  - Database files, redo log files, control files
  - Archived redo log files / FRA
  - Server Parameter File (spfile)
  - RMAN backups & Data Pump Exports
- Can't be stored in ASM (unless ACFS):
  - Trace files / Trace Directories / ADR
  - Oracle software / binaries / oratab
  - DBA scripts

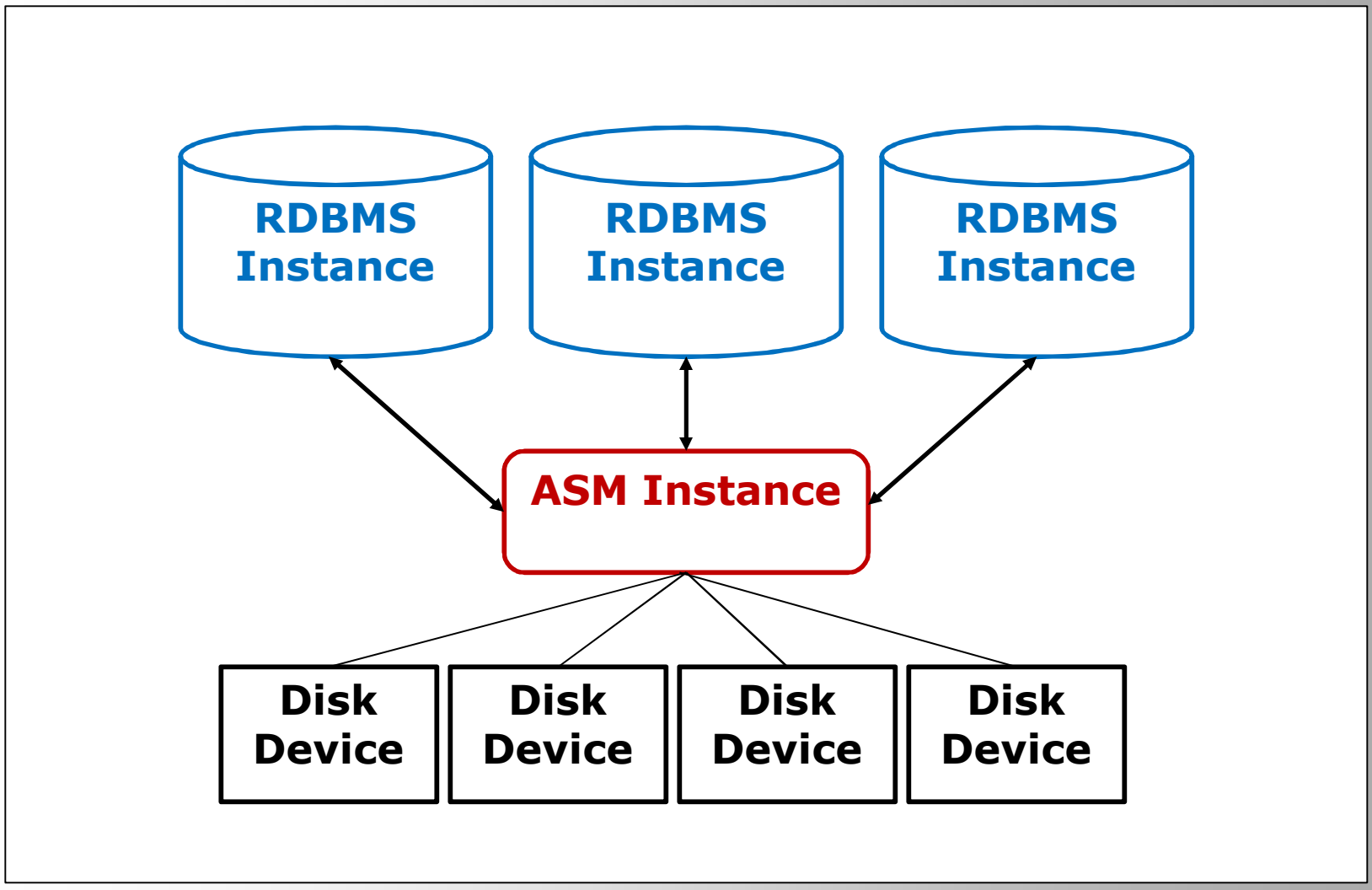
Concepts

- Initialization Parameter: **`INSTANCE_TYPE`**
  - Values: **`RDBMS`** or **`ASM`**
- ASM Instance
  - Small instance specifically for managing ASM disks and files
  - No physical files or controlfiles
  - Does have an initialization file (spfile)
  - Very few parameters – some ASM specific

The ASM Instance

- Runs on each server or RAC node
- SID: **+ASM** or **+ASM<RAC node ID>**
- Practical Parameter Settings:
  - **INSTANCE_TYPE = ASM**
  - **SGA_TARGET = 1250M**
  - **LARGE_POOL_SIZE = 12M**
  - **PROCESSES = 1024**
  - **ASM_DISKSTRING = '/dev/asm-disk*'**

Server Logical Configuration



The RDBMS Instance

- No special parameters or changes required
- Can have any mixture of ASM and non-ASM files
- Some ASM specific views are visible through the ASM instance
- Parameters that can reference ASM:
  - **DB_CREATE_FILE_DEST, DB_CREATE_ONLINE_LOG_DEST_m, DB_RECOVERY_FILE_DEST, CONTROL_FILES, LOG_ARCHIVE_DEST_n, STANDBY_ARCHIVE_DEST**

Diskgroups

- A grouping of ASM Disks
  - What is presented to the RDBMS instances
  - Equivalent of your traditional mount point or drive or partition
  - Starts with a “+”
  - Oracle recommends only two: DATA & FRA
- ASM automatically creates a hierarchy
  - **<DISK GROUP>/<DB NAME>/<FILE TYPE>/<FILE NAME>**

Diskgroups and OMFs

- Usually just need to specify the Disk Group name in place of full file names:
  - **CREATE TABLESPACE test DATAFILE '+DATA' SIZE 500M;**
- Can still see the full file names in the normal views (i.e. **DBA_DATA_FILES**) and use in the same way as any other file

Failure Groups

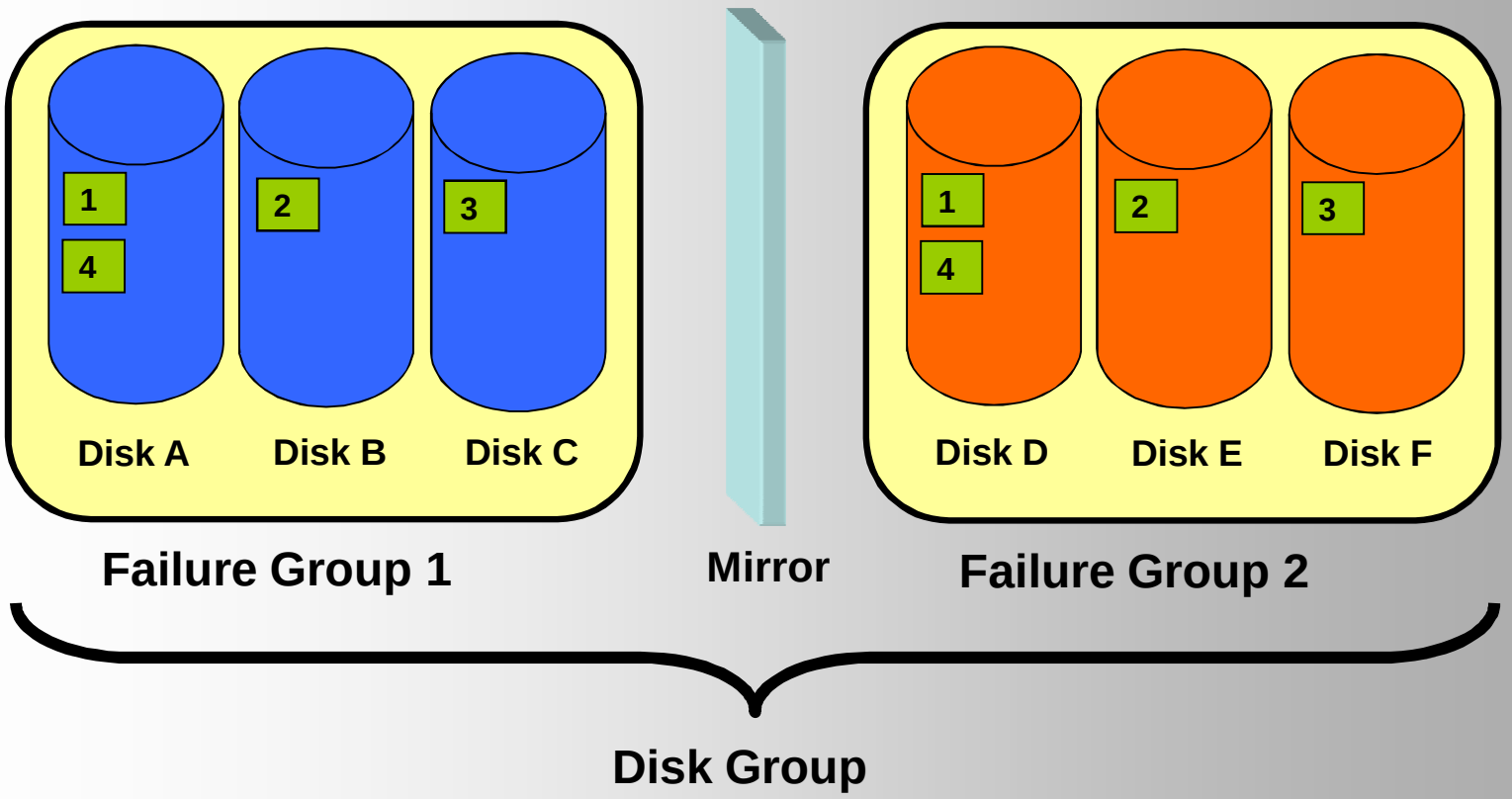
- Disks that share common components
  - Could all potentially fail together (i.e. common disk controller)
  - Used for determining which disk groups can be used for redundancy (i.e. mirroring)

Disk Group Redundancy

- NORMAL
  - 2-way mirroring (2 failure groups)
- HIGH
  - 3-way mirroring (3 failure groups)
- EXTERNAL
  - No Oracle mirroring – relying on external hardware configuration (RAID)
  - Typical option for SANs

Striping and Mirroring

- Stripe data among disks within the failure group and mirror to a second failure group



ASMCMD

- Command line utility to allow ASM navigation and commands similar to an OS CLI
  - Introduced with 10gR2
  - Allows commands against disks, diskgroups, and files
  - Can be used to copy files including into and out of ASM

Security

- New OS security group: **`osasm`**
- New SYS role: **`sysasm`**
- Oracle recommends installing ASM software under a different OS user but this can add complexity
- In most cases the DBA manages ASM

New in 11g: G.I.

- New “Grid Infrastructure” (GI) home
  - No longer runs from the RDBMS home
  - Brings the management concepts of RAC
    - “**`srvctl`**” commands are mandatory
    - High Availability Services (HAS)
    - Clusterware (CRS & CSS)
  - New concepts: “Oracle Restart” and the “Oracle Local Registry” (OLR)
    - Traditional “**`dbstart`**” and “**`dbshut`**” scripts cannot be used

Complexities

- Must understand how to start, stop, re-configure, and troubleshoot HAS and CRS/CSS
  - Commands must run as root
    - Does the DBA have root access or sudo?
  - MOS:
    - **Top 5 Grid Infrastructure Startup Issues [ID 1368382.1]**
    - **How to Proceed from Failed Upgrade to 11gR2 Grid Infrastructure on Linux/Unix (Doc ID 969254.1)**

Learning & Experimentation

- Easy to learn & experiment using a VM
  - Oracle VirtualBox (free)
    - Install OEL, GI, and RDBMS software
  - Add disks to VM: **`fdisk`**, **`udev`**, **`chown`**

--- *OR* ---

  - Hidden Parameter:
    - **`"_asm_allow_only_raw_disks" = FALSE`**
  - Create empty files on top of cooked FS:
    - **`dd if=/dev/zero of=asm_1 bs=1048576 count=100`**

Other Considerations

- Backup ASM separately (metadata only)
  - **`md_backup`**
    - exports configuration into text files: uses perl module **`data::dumper`**
  - **`md_restore`**
- Additional software (home) to maintain and patch
  - Patching the GI home will require all RDBMS databases to be stopped

Performance Comparison Issues

- Watch out for OS buffer caching – “Double Buffering”
  - Is the cooked FS physical I/O really a physical I/O?
  - Raw devices (including ASM) will always bypass OS file system caching
  - Use that memory for the SGA instead – regardless of whether raw devices (ASM) are used
  - MOS:
    - **File System's Buffer Cache versus Direct I/O [ID 462072.1]**

New in 12c

- Flex ASM
  - Only one RAC node requires an ASM instance
  - Nodes can use a non-local ASM instance
- ASM Disk Scrubbing
  - Logical corruption detection and automatic repair
- Can specify DG type
  - DATA / RECOVERY / SYSTEM
- Oracle Database File System (ODFS)

Further Reading

- Oracle Press:
  - “Oracle Automatic Storage Management”
- OTN Whitepaper:
  - “Oracle Database 11g Release 1 and 10g Release 2 - Automatic Storage Management - Overview and Technical Best Practices”

Conclusions

- Initial Complexities
  - Disk provisioning
  - Learning curve on “Clusterware”, “Oracle Restart”, and “ASM”
- Many Benefits
  - Management simplicity
  - Performance, redundancy, and resiliency
  - Disk space optimization
  - No need to purchase a volume manager

Questions or Comments?

Simon.Pane@first4db.com