IBM

Technical report:

Using Oracle Database 10g Automatic Storage Management with IBM System Storage N series

Best Practices

Document NS3329-0

January 2, 2008

Table of contents

Abstract

This paper seeks to educate the Oracle DBA on the combined benefits of using IBM N series storage for Oracle Databases and database layered applications and includes best practices recommendations for deployment of Automatic Storage Management with IBM N series networked storage solutions.

Introduction

The purpose of this joint white paper is to share best practices for use of the Oracle® Database 10*g*™ feature Automatic Storage Management (ASM) with IBM® System Storage™ N series NFS NAS, Fibre Channel, iSCSI SAN storage products. ASM provides a new way to manage the storage underlying the database. IBM N series filers simplify the deployment and management of enterprise data. Both ASM and IBM N series storage lower total cost of ownership, and they complement each other to offer even greater cost savings. This paper shares some of the considerations as to how the combined technology can best be configured.

ASM provides an integrated file system and volume manager for the database files, built into the Oracle Database kernel. With this capability, ASM provides an alternative to some of the third-party file-system and volume management solutions for database storage management tasks, such as creating/laying out databases and managing the use of disk space.

This paper seeks to educate the Oracle DBA on the combined benefits of using IBM N series storage for Oracle Databases and database layered applications and includes best practices recommendations for deployment of ASM with IBM N series networked storage solutions. Like Oracle on IBM N series NAS storage, Oracle ASM on IBM N series SAN and iSAN storage gives customers an alternative capability for volume management on the Oracle server host using familiar create/alter/drop SQL statements, simplifying the job of DBAs with regard to database storage provisioning.

The following topics summarize how features of ASM and IBM N series storage complement and combine with one another:
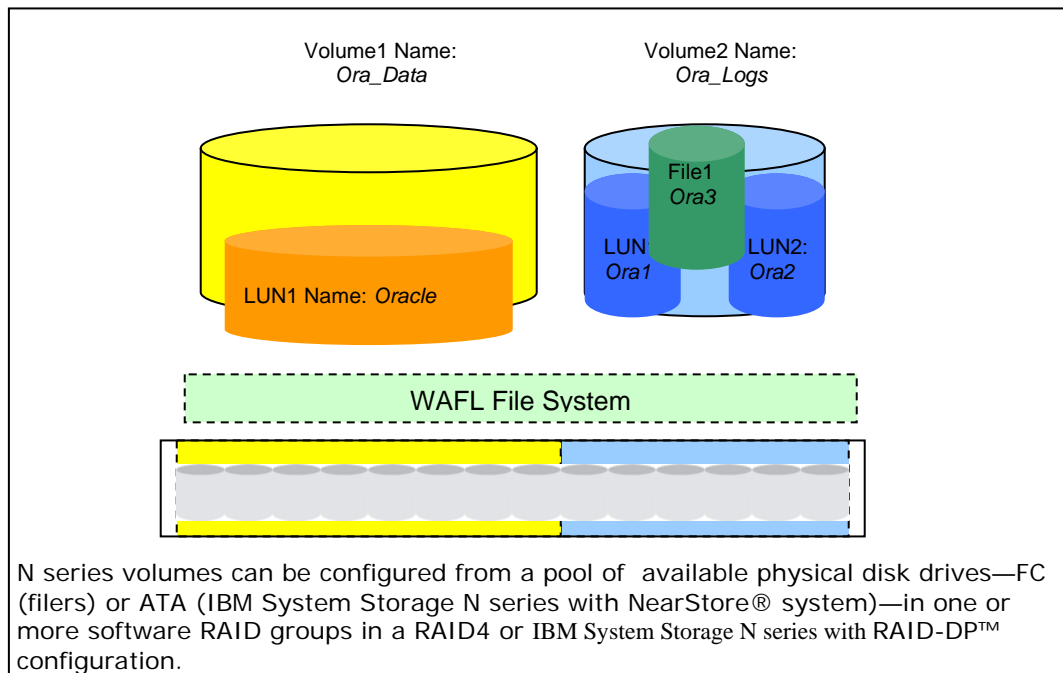
I/O Load Balancing

Load balancing avoids performance bottlenecks by assuring that the I/O workload utilizes all available disk drive resources.

- IBM N series storage automatically load-balances I/O among all the disk drives in a IBM System Storage N series with WAFL$^{®}$ (write anywhere file layout) volume (see Figure 1). All logical unit numbers (LUNs) and files placed within a single WAFL volume can be assured of utilizing all the volume's disk drives in a balanced fashion.

- ASM provides load-balanced I/O across all LUNs or files (see Appendix 1) in an ASM disk group by distributing the contents of each data file evenly across the entire pool of storage in the disk group based on a 1MB stripe size.

When used in combination, IBM N series load balancing allows multiple LUNs and file-system data to share common disk drives, while reducing the number of LUNs per ASM disk group for improved manageability. ASM further allows load balancing across multiple WAFL volumes or filers. The WAFL file system and RAID4 are optimized for random small I/O. WAFL ensures the data is evenly spread across all the disk drives of a RAID group in a volume. This ensures optimal read/write performance in high-transaction database environments.

This figure shows IBM N series with WAFL file-system volumes, each residing on a distinct set of physical disk drives. The volumes contain files and LUNs, which are distributed across all the disk drives in the volume.

Figure 1. IBM N series storage overview.

Adding/Removing Storage Capacity

Support for online addition and removal of storage capacity maximizes data availability while making necessary storage configuration changes (refer to Figure 1 for an illustration of these principles).

- ASM allows online resize, addition, or removal of LUNs or files within each ASM disk group, automatically rebalancing data distribution for new configurations. Entire disk groups can be migrated from one set of LUNs or files to another (e.g., on different volumes or filers) by adding then removing LUNs or files from a disk group without database downtime.

- IBM N series storage allows online addition and replacement of disk drives within each WAFL volume and rebalances the load to utilize new disk drives as writes occur.

N series I/O rebalancing across physical disk drives only occurs when disk drives are added to expand an IBM N series volume nondestructively. Adding LUNs or files to a volume does not cause any additional I/O rebalancing to occur as long as there is enough space in the volume to add a new LUN/file or grow an existing LUN/file. The key benefit to volume and LUN/file resizing is that it is a nondestructive process that is rapid and easy to accomplish. This allows you to grow your ASM disk group easily and rapidly with no downtime, allowing you to provision storage dynamically in sync with your database growth needs.

Data Protection and Availability

Enterprises require very high levels of protection and availability for their data, regardless of faults or component failures in storage subsystems.

- IBM N series storage implements RAID4 or RAID-DP™ protection on all WAFL volumes. RAID4 assures full data availability in spite of the loss of any given disk drive in a RAID group. RAID-DP implements double-parity protection to guard against the loss of any two disk drives. The WAFL write anywhere data layout assures excellent RAID write efficiency without incurring the storage costs implied by full mirroring. IBM N series cluster failover protects against failure of other filer hardware components, such as within the filer head itself. Additionally, IBM N series mirroring solutions can be implemented for additional protection for business continuance and disaster recovery requirements.

- ASM optionally implements two-way or three-way extent-based mirroring, offering an additional means for protection against storage failures. ASM also provides the ability to mirror between sets of ASM disks (implemented as LUNs or files on IBM N series storage), called failure groups, and can keep data available after a component failure causes all storage in one of those failure groups to become unavailable.

Combining the data protection technologies from IBM N series and Oracle offers customers several ways to address their data protection requirements. The following best practices address how these might best be configured.

Shared Storage for RAC Clustered Nodes

Oracle10*g* Real Application Clusters (OracleRAC10*g*) provides high availability and scalability for Oracle Databases implemented across clustered database servers. OracleRAC10*g* requires that all servers have full, direct read/write access to the database files.

- ASM implements simple and Oracle RAC-optimized cluster file system and volume manager functionality for shared storage.

- IBM N series networked storage also provides the shared access RAC requires. In SAN and NAS environments, ASM provides the cluster-capable file system and volume management for OracleRAC10*g*.

ASM and IBM N series unified storage (SAN/NAS) combine to create a simple and scalable storage facility for OracleRAC10*g* . Together, IBM N series and Oracle toolsets provide a rich assortment of data management options for use separately or in combination. Customer selection of techniques will depend on their unique requirements and the context of their enterprise's data infrastructure.

Best Practices Recommendations

This section discusses guidelines for configuring ASM disks, creating ASM disk groups, and creating databases within ASM, discussing some best practices along the way. Please refer to Figures 2 through 4 for further illustration of some of the best practices concepts and recommendations proposed.

Common Recommendations

The following recommendations apply regardless of which of the two high-level deployment models above is selected:

- Use host file systems or NFS to store Oracle binaries and other non-ASM-enabled files. In OracleRAC10*g* environments, NFS can be used to simplify deployment using a shared Oracle Home.

- Use either IBM N series unified storage (Fibre Channel, iSCSI SANs or NFS/TCPIP NAS) protocols to access the storage for ASM disk groups.

- Combine each filer's disk drives into few large WAFL volumes to minimize administrative overhead.

- Make sure that the LUNs or files within each disk group are balanced in terms of throughput per capacity, so that the I/O throughput available to a disk group from each LUN or file is proportional to the LUN's (or file's) size. In more concrete terms, all LUNs/files within the disk group should supply roughly equivalent I/O operations per second per gigabyte. Generally this will be best accomplished using disk drives with very similar capacity and performance properties throughout the underlying WAFL volume(s).

- Configure several (e.g., four) LUNs or files per ASM disk group per WAFL volume. Configuring multiple LUNs/files can maximize I/O concurrency to the disk group, overcoming any per-LUN limits of the host driver stack (operating system, host bus adapter driver, and multipath driver). Even if initially not required, this provisioning can avoid unneeded data movement that occurs if LUNs within the same volume are later added to the disk group.
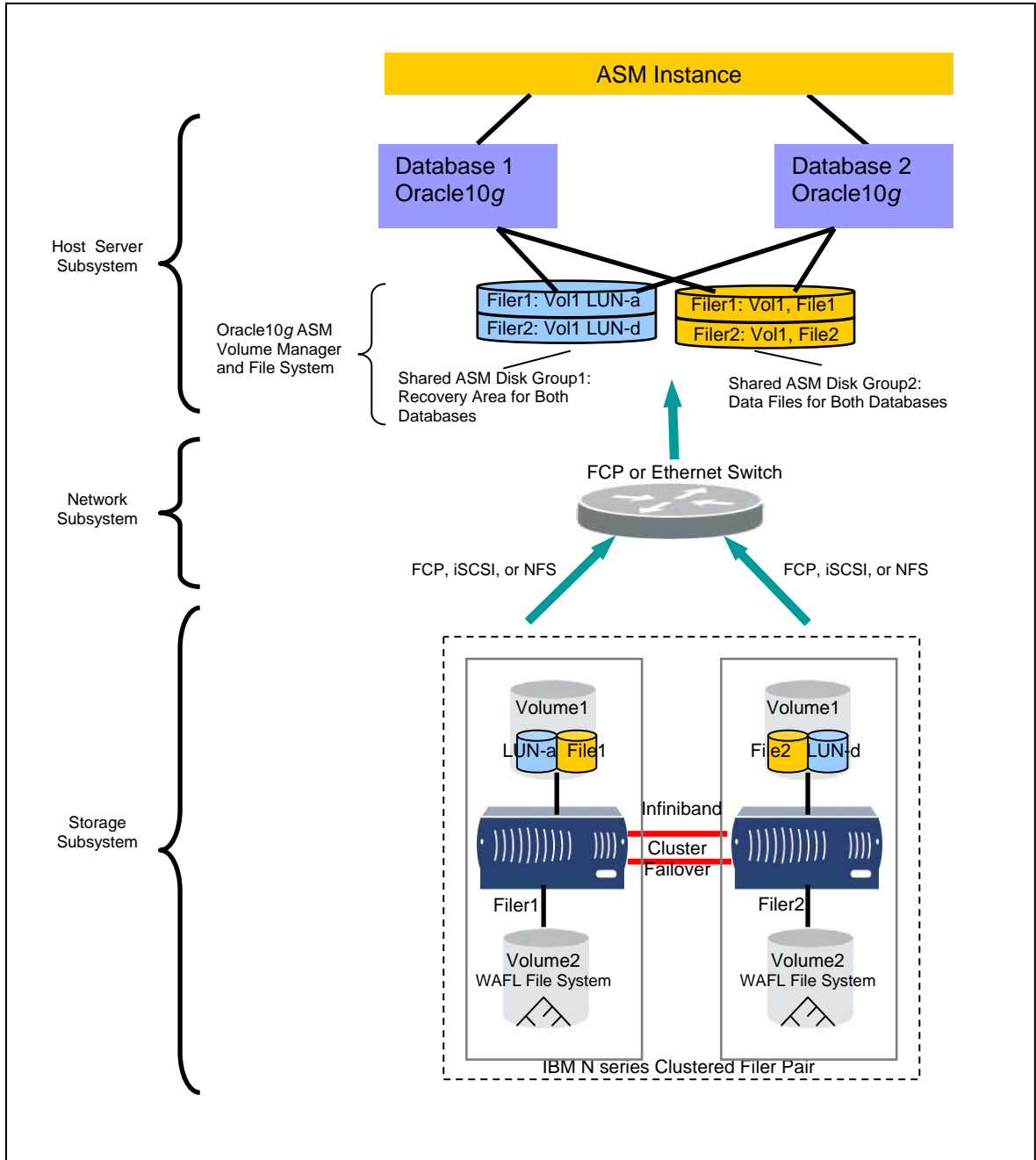
Deployment Models

Either of two high-level deployment models may be appropriate for ASM on IBM N series, depending on the intended use of data management facilities at the storage subsystem (filer) level:

- If the ASM deployment will use filer-based features for data management (such as filer-based facilities to implement full database backup/restore, cloning, mirroring, and/or migration), then separate ASM disk groups should be established for each database deployed, and each disk group should be stored within a single WAFL flexible volume. This deployment model is termed "**Separate Disk Groups**" below.

- If filer-based data management will not be used, then it will be simpler to group multiple databases into shared disk groups. This deployment model is termed "**Shared Disk Groups**" below.

Shared Disk Groups

When all data management methods will be based exclusively on Oracle tools, the easiest deployment model will group multiple databases into shared disk groups that provide a common shared pool of storage for several databases. The following additional recommendations are specific to this model. See Figure 2 for an illustration of these recommendations:

- Use two or more Oracle ASM disk groups to store the databases. The database work area should be stored in separate disk groups from their respective flashback recovery areas. For databases that may require independent control over the underlying storage resources (e.g., to assure quality of service or independent data migration capabilities), assignment into distinct ASM disk groups may be considered.

- Each ASM disk group may include storage from one or more WAFL volumes or filers. Spreading a disk group across multiple WAFL flexible volumes will distribute the disk group's storage and throughput demands across volumes based on the combined size of the LUNs or files that each volume contributes to the disk group to improve performance.

- When adding disk drives to existing volumes for greater capacity or throughput, expand all the WAFL flexible volumes and the disk group's LUNs (or files) within them to maintain uniform throughput-per-capacity ratios and to minimize unneeded data movement due to load balancing.

- If adding storage from a new filer within a disk group, configure the new filer's LUNs (or files) with similar throughput-per-capacity properties to those already in the disk group. If the new filer contains multiple LUNs/files for host I/O concurrency, then all LUNs/files should be added to the disk group at once to limit ASM's load-balancing movement to that required.

Figure 2. Multiple Oracle Databases running on shared ASM disk groups that are distributed across multiple N series filers.
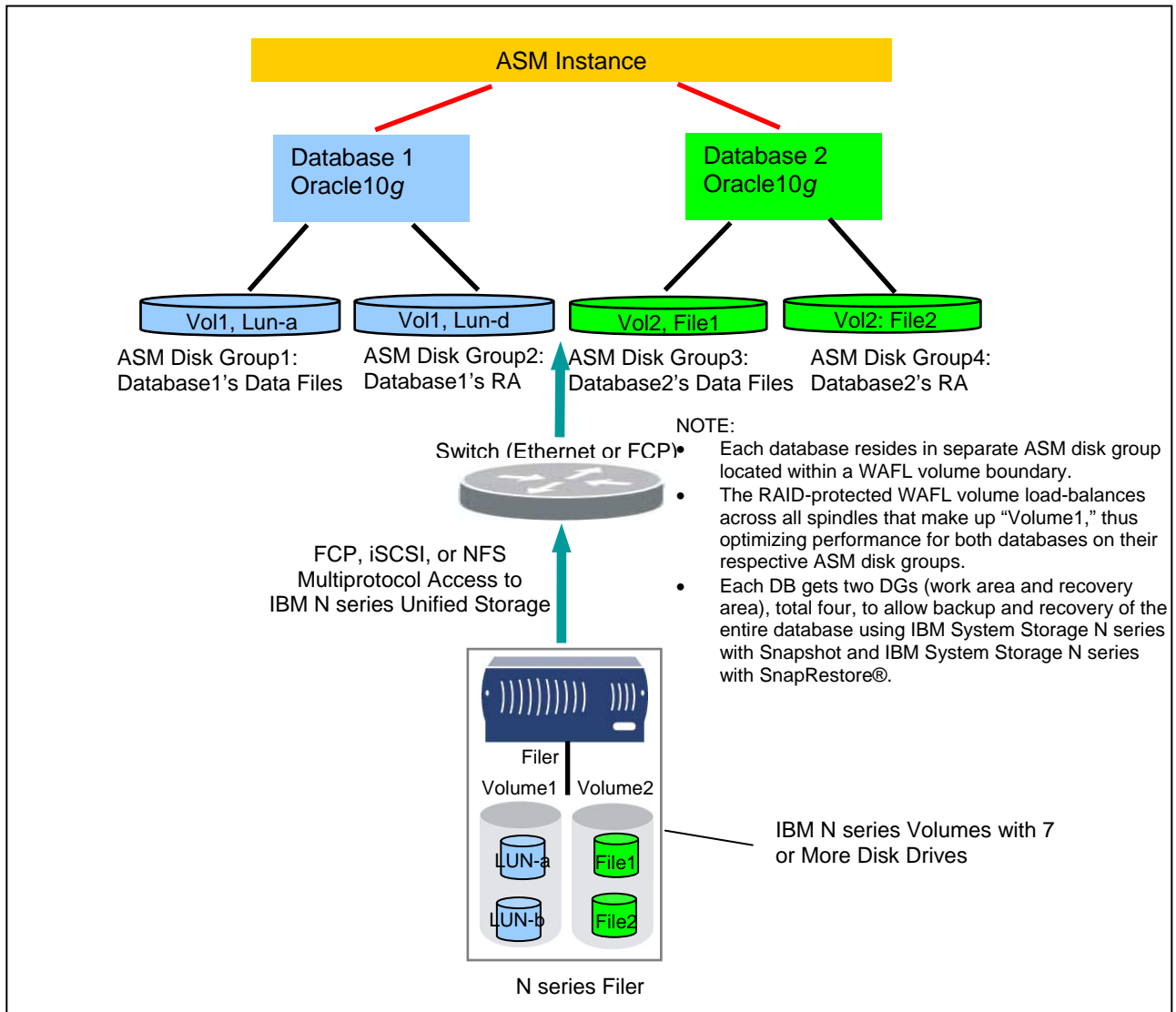
Separate Disk Groups

In order to leverage the filer's ability to implement data management, each database should be stored in distinct ASM disk groups with each ASM disk group stored in a single WAFL flexible volume. If that is not possible, an ASM disk group that spans WAFL flexible volumes requires that the disk group be unmounted when creating IBM System Storage N series with Snapshot™ copies to assure ASM metadata consistency among the Snapshot copies. See Figure 3 for an illustration of these recommendations.

- Use two Oracle ASM disk groups per database, one containing database work area data files and the other containing the flash recovery area.

- Place all LUNs within a given disk group in a separate WAFL qtree (quota tree directory) to maximize data management flexibility.

- When adding disk drives for greater capacity or throughput, expand the WAFL flexible volume and then resize the LUNs on the filer, in the host operating system, and finally in ASM to expand the ASM disk groups. Perform all the ASM add disk and drop disk commands after you have set the power limit parameter to 0 in order to pause the rebalance process until you are ready to increase the power limit and have the rebalance done in one pass.

ASM Instance

Database 1
Oracle10*g*

Database 2
Oracle10*g*

Vol1, Lun-a

Vol1, Lun-d

Vol2, File1

Vol2: File2

ASM Disk Group1:
Database1's Data Files

ASM Disk Group2:
Database1's RA

ASM Disk Group3:
Database2's Data Files

ASM Disk Group4:
Database2's RA

NOTE:
- Each database resides in separate ASM disk group located within a WAFL volume boundary.
- The RAID-protected WAFL volume load-balances across all spindles that make up "Volume1," thus optimizing performance for both databases on their respective ASM disk groups.
- Each DB gets two DGs (work area and recovery area), total four, to allow backup and recovery of the entire database using IBM System Storage N series with Snapshot and IBM System Storage N series with SnapRestore®.

Switch (Ethernet or FCP)

FCP, iSCSI, or NFS
Multiprotocol Access to
IBM N series Unified Storage

Filer

Volume1 Volume2

LUN-a File1

LUN-b File2

IBM N series Volumes with 7
or More Disk Drives

N series Filer

Figure 3. Multiple Oracle Databases running on separate ASM disk groups that map to separate IBM System Storage N series with WAFL file-system volumes.

Failure Groups and Redundancy

Generally, ASM's external redundancy option should be configured for ASM disk groups on IBM N series storage. This choice minimizes the amount of storage hardware required to realize data protection for a given level of database capacity and throughput. With external redundancy, the IBM N series LUNs that comprise the ASM disk groups can be RAID parity protected. Standard RAID4 parity protection allows for continued operation with a single disk drive failure in a RAID group. IBM System Storage N series with double-parity (RAID-DP option) protection allows for continued operation even after a simultaneous double disk drive failure in a RAID group, dramatically reducing the risk of data loss due to simultaneous double disk drive failure.

The IBM N series cluster failover (CFO) option should generally be configured to assure availability in case of unscheduled filer downtime. As an alternative to using CFO, ASM disk groups could be configured with two-way mirroring (i.e., without the external redundancy option) to provide this assurance,

for a cost of up to double the required disk drive hardware. If ASM mirroring is used to protect against filer failure, all LUNs in a disk group hosted by one filer should be placed within the same failure group.
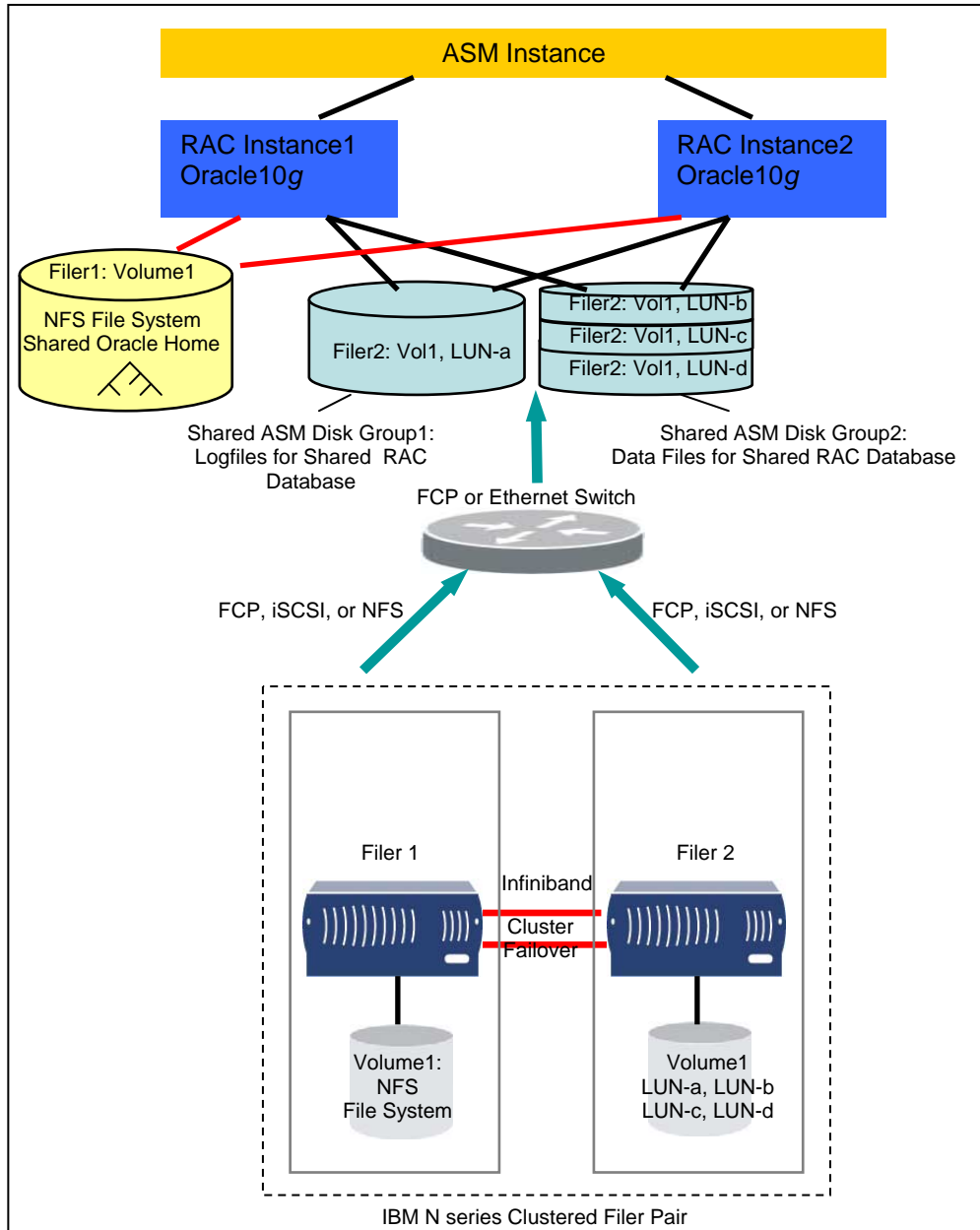
ASM Disks

IBM N series LUNs have the same semantics on the host side as any other LUNs, and as such they are incorporated into ASM disk groups as documented in the 10*g* ASM documentation. The best practice recommendation is to use LUNs (or files) of the same size and performance capacity when configuring an ASM disk group. In order to achieve good load balancing for a disk group stored on multiple WAFL volumes, the total capacity from each WAFL volume should be proportional to the throughput available to the disk group from that volume.

Running RAC10g on ASM Disk Groups Derived from IBM N series

When OracleRAC10*g* is deployed with ASM on storage derived from IBM N series, we recommend placing the shared Oracle Home on an IBM N series NFS volume for ease of manageability. In order to install and enable cluster synchronization services in Oracle10*g* one needs to follow the directions in the product documentation. Alternatively, some clustered host file systems can support a shared Oracle Home on SAN storage, or each node can have its own Oracle Home on a private host file system. See Figure 4 below.

With the release of OracleRAC10*g*, Oracle has made it extremely easy, as compared to OracleRAC9*i*[TM], to share a single Oracle Home on a NAS volume among multiple nodes of an OracleRAC10*g* cluster. The ability to share the Oracle Home binaries and configuration files between multiple RAC instances adds tremendous benefits to manageability and administration of a multinode OracleRAC10*g* cluster. The shared instance specific admin and configuration files (TNSnames.ora, initSID.ora, etc.) have separate parameter entries for each cluster node. When deploying ASM on IBM N series SAN storage, it is generally preferable to implement the OracleRAC10*g* CRS files (voting disk and OCR device) on SAN storage as well; consult the configuration details used for the applicable OracleRAC10*g* validation.

Figure 4. Multiple Oracle10*g* RAC instances running on shared ASM disk groups.

Conclusion

Automatic Storage Management simplifies the configuration and management of storage management for the Oracle10*g* database data files. ASM complements IBM N series Fibre Channel and iSCSI SAN storage products, and they can be combined in ways that deliver the full value of both sets of technologies. Although several possible combinations of redundancy protection and data management options exist, IBM and Oracle provide best practices for how these can be combined to meet the business requirements of our joint customers. Depending on the high-availability and performance requirements, there are some powerful combinations of our joint technology that meet those demands. Although the increased number of possible combinations may seem complex at first, this paper sets the criteria in place to help choose the most cost-effective solution that fits the specific requirements.

Summary of Best Practices

- ASM is the recommended file system and volume management solution in an IBM N series NAS or SAN/iSAN unified storage environment for Oracle Databases.

- Configure ASM disk groups with external redundancy for use with filers.

- Use IBM N series cluster failover to assure no single point of failure.

- Use ASM for management of Oracle Database files.

- Use the IBM N series clustered file system for shared Oracle Home, alert logs, and trace files.

- Build ASM disk groups with LUNs or files of similar performance/capacity properties.

- Distributing I/O across the front end with ASM and striping within the filer are complementary and can further simplify the combined configuration versus using either approach in isolation.

Appendix 1: Implementing ASM on NFS-based Storage

ASM can be configured with NFS-based files as ASM disks. As described in the paper above, when used in this manner, a large file is configured to be a storage container for an ASM 'Disk'. When using ASM on top of NFS, the NFS files are used as ASM disks, so the provisioning and data management model is similar to that for ASM on SAN storage, while NFS is used to perform I/O. ASM on IBM N series NFS can be used to load balance across multiple filers just as you are able to with ASM on IBM N series SAN (FCP, iSCSI). This is easily achieved by building an ASM disk group with ASM disks located on volumes from multiple filers. See Figure 2.

There may be circumstances where ASM on NFS is applicable, whether for uniformity of operating procedures or to satisfy product, host OS, and/or other environmental constraints. The following points discuss such configurations further.

NFS-based storage from IBM N series and ASM provides some common functions. Oracle DBAs can directly employ the NFS-based file system functionality to implement simple, scalable, RAC-ready file storage. An Oracle DBA can directly provision storage from an IBM N series filer using NFS storage and the familiar create/alter/drop SQL statements. ASM on NFS similarly implements scalable, RAC-ready storage and further extends scalability at the high end, but provisioning additional filer storage to Oracle files within Oracle will require some additional steps when using ASM on NFS storage.

When implementing ASM on NFS, one should combine the guidelines for deploying Oracle directly on NFS with those above for implementing ASM on IBM N series SAN storage:

- Use NFS mount options as recommended for non-ASM Oracle files on NAS. See the applicable IBM N series technical reports (see Appendix 3) for guidance on configuring specific NFS clients for use with Oracle.

- Create NFS-based files according to the guidelines for LUNs discussed in the "ASM Disk Groups" best practices above. Either the "shared disk groups" or "separate disk groups" model may be followed as appropriate. Multiple files and/or mount points from a single WAFL volume can be used to overcome I/O concurrency limits, just as multiple LUNs were discussed for this purpose in the SAN best practices above.

- Use *mkfile* on UNIX® platforms (or *dd* on Linux platforms) to size the files to be incorporated into the ASM disk groups (as if they were LUNs). If later resizing is required, establish up-to-date backups, then alter the files' sizes as required without truncating or overwriting any live data; the tools needed to do this vary among host environments.

- Set the permissions on the files to allow "oracle:dba" read/write access.

- Use an ASM disk string that allows the ASM instance to find the file, for example, /nfs/<host>/oradata/<db_name>/*. Symbolic links can be used to make such a disk string work for files spread across multiple NFS mount points.

Appendix 2: Using IBM N series with SnapShot Copies with ASM-based Data

Backup and recovery methods are of course critical elements of any Oracle production deployment. Refer to Figures 5a and 5b for an illustration of the concepts presented in this section.

- IBM N series with Snapshot copies implement inexpensive point-in-time copies, which can be used to accelerate media recovery, making them a very popular feature for Oracle environments. Snapshot copies also provide the basis for various other backup/recovery methods and can also be used to clone databases nearly instantaneously. In order to effectively apply Snapshot copies with ASM-based Oracle data, it is important to focus on the ASM disk group as the unit of Snapshot copy–based recovery, copying, or cloning.

- ASM also works with the Oracle full suite of data management tools, including RMAN, Data Guard, Flash recovery, and Flashback queries. These technologies can be used when databases span multiple WAFL volumes or more than one filer.

RMAN is the only interface able to take hot and cold backups of Oracle Databases on ASM disk groups, and, more importantly, RMAN is the only interface for single file restore capability from a backup set. However, you can also do hot and cold backups of Oracle using the IBM N series volume-level point-in-time Snapshot copy and IBM System Storage

N series with SnapRestore to back up and restore entire LUN(s) and hence entire ASM disk group(s) to a point in time. In order to preserve ASM disk group integrity, all disks within the ASM disk group must be snapped atomically with the current IBM N series technology; this practically implies following the "separate disk groups" guidelines discussed in the best practices section earlier, including placement of all LUNs for a given disk group into a single WAFL volume. ASM maintains all disk group metadata within the disk group itself, so a disk group can be restored to a previous state by dismounting, restoring, and remounting the disk group. A disk group's metadata will be fully consistent for any atomic point-in-time copy of the disk group, so there is no need for flushing and freezing the metadata as is required prior to creating Snapshot copies of some traditional host file systems. All that is required for doing a hot backup is to put the individual tablespaces of the database in online or hot backup mode and then create a Snapshot copy of the volume containing the LUN(s).
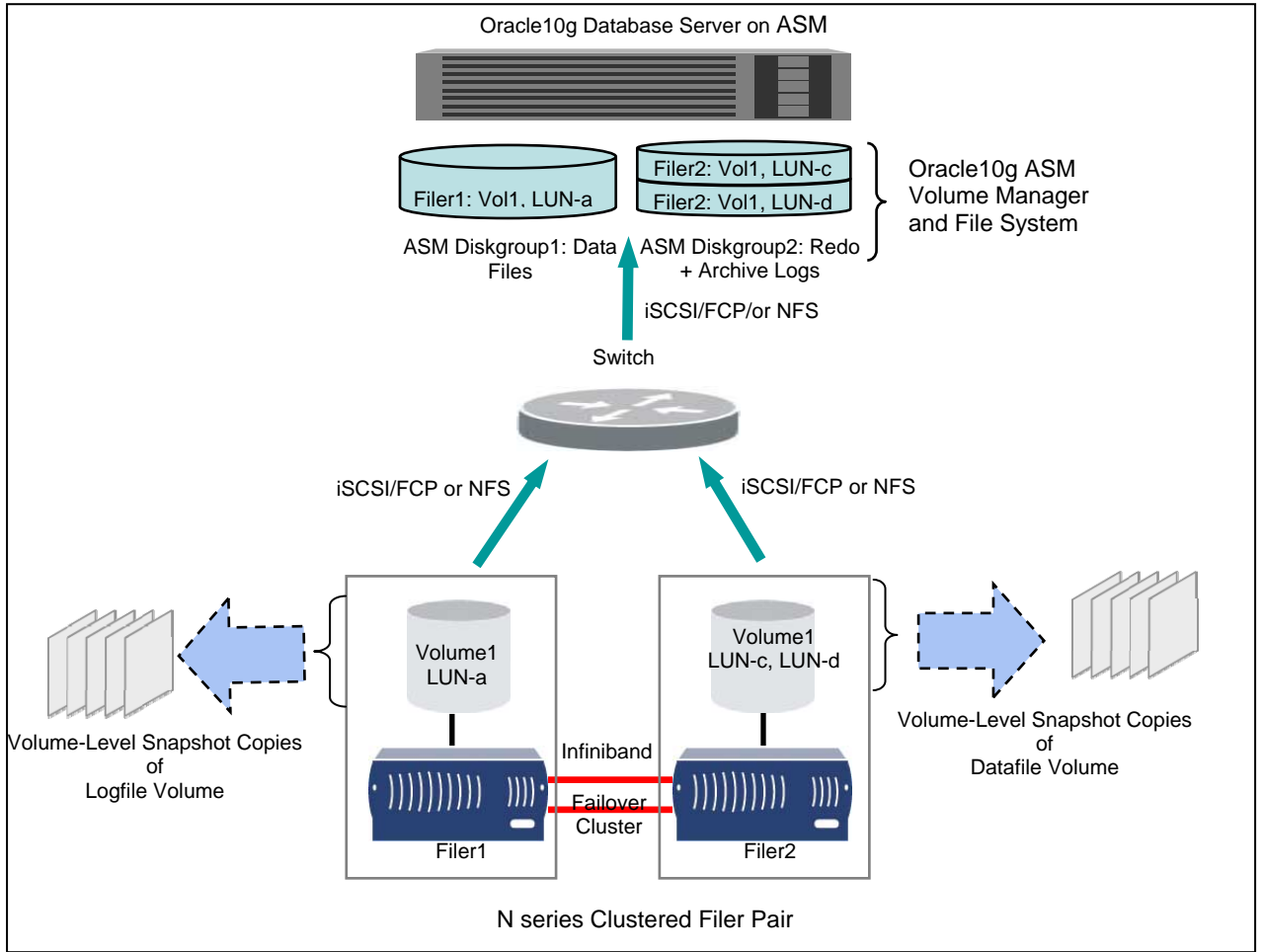
Figure 5a. Overview—Backing up an Oracle10*g* Database on ASM with IBM N series.
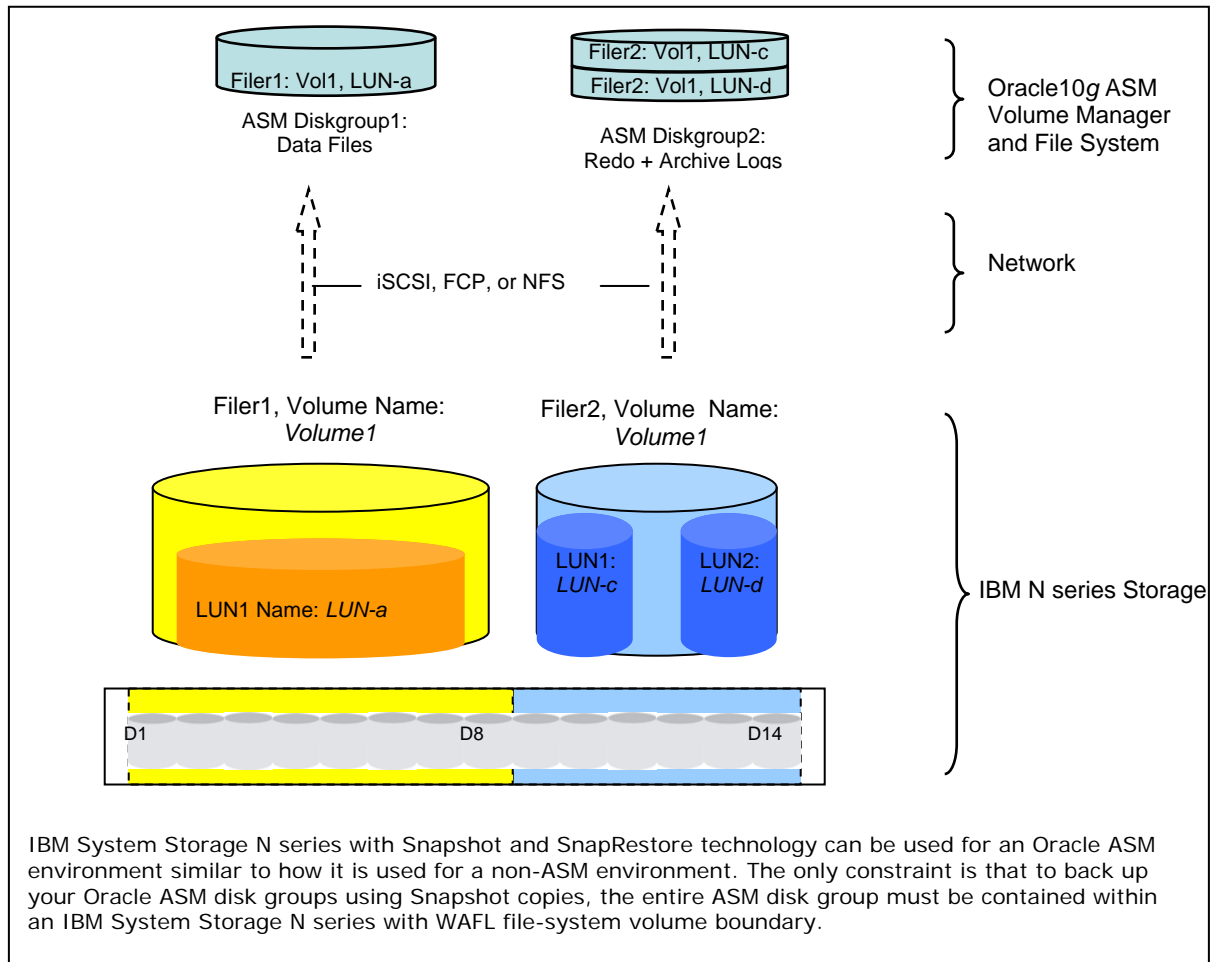
Figure 5b. Backing up an Oracle Database on ASM with IBM N series storage.

Within the figure:

Filer1: Vol1, LUN-a

ASM Diskgroup1:
Data Files

Filer2: Vol1, LUN-c
Filer2: Vol1, LUN-d

ASM Diskgroup2:
Redo + Archive Logs

Oracle10*g* ASM
Volume Manager
and File System

iSCSI, FCP, or NFS

Network

Filer1, Volume Name:
Volume1

Filer2, Volume Name:
Volume1

LUN1 Name: *LUN-a*

LUN1:
LUN-c

LUN2:
LUN-d

IBM N series Storage

D1   D8   D14

IBM System Storage N series with Snapshot and SnapRestore technology can be used for an Oracle ASM environment similar to how it is used for a non-ASM environment. The only constraint is that to back up your Oracle ASM disk groups using Snapshot copies, the entire ASM disk group must be contained within an IBM System Storage N series with WAFL file-system volume boundary.

Trademarks and special notices