An Oracle White Paper
January 2013

Oracle Cloud File System

Executive Overview

As storage requirements double every 18 months, Oracle customers continue to deal with complex storage management challenges in their data centers. Conventional methods for managing storage are complex and costly. The customer is often forced to be the integrator for multiple tools and software layers related to their storage hardware platforms, volume managers, file systems and clustering technologies. This problem is compounded when it comes to managing data across different types of platforms. The challenge is a result of using multiple management tools and managing support from multiple vendor support organizations.

As a result of these storage management challenges, organizations are seeking better ways to cost effectively manage enterprise data growth and data center complexity. Cloud computing has emerged as a way to help organizations provide a way to pool resources, unify network accessibility and deliver rapid elasticity.

To help organizations deploy their applications, databases and storage in private clouds, Oracle offers the Oracle® Cloud File System (Oracle CloudFS) providing a comprehensive volume and file management solution. This solution is designed to manage all general purpose and database files across multiple operating system platforms. Oracle CloudFS eliminates the need for 3rd party volume managers, file systems and clustering technologies. As a result, Oracle CloudFS simplifies the storage stack and reduces cost.

Oracle CloudFS delivers a storage cloud infrastructure that simplifies pooling of storage across file systems, middleware, and applications. It also provides network accessible storage with unified namespace for files and it supports rapid scaling through online storage provisioning. These capabilities are critical to cloud computing.

Oracle CloudFS management is streamlined through one management interface, one set of installation and configuration tools, one cluster framework and one vendor for support. The product provides a scalable solution designed to reduce the total cost of ownership, simplify system and storage management, and provide high performance without the need for custom tuning. Oracle CloudFS optimizes utilization through consolidation and ensures file system availability.

Introduction

Oracle CloudFS simplifies and automates storage management functions, increases storage utilization, uptime and agility to deliver predictable performance and availability for all general purpose and database files.

The Oracle CloudFS includes a cluster volume manager and a cluster file system:

- Oracle ASM Dynamic Volume Manager (ADVM), a general purpose cluster volume manager component built on Oracle's Automatic Storage Management. Oracle ADVM supports Oracle ACFS as well as 3rd party file systems.

- Oracle ASM Cluster File System (ACFS), a general purpose POSIX and Windows compliant cluster file system with advanced data services such as snapshot, replication tagging, security, encryption and auditing.
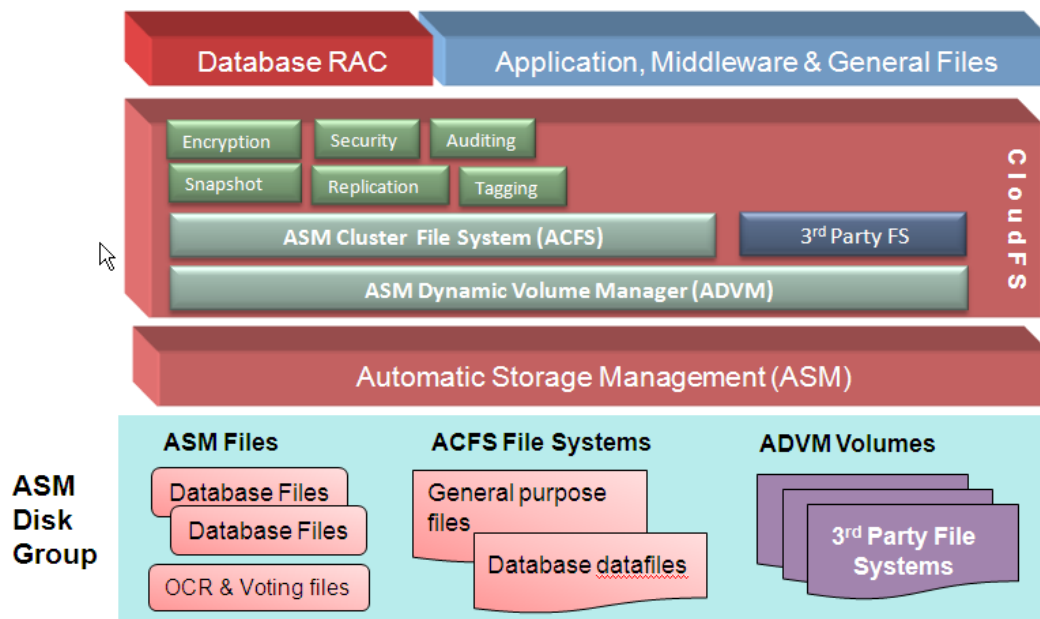

Diagram I

The combination of Oracle CloudFS, Oracle Automatic Storage Management and Oracle Clusterware form the Oracle Grid Infrastructure. This creates an integrated foundation for database and general purpose files as well as an infrastructure for cluster and single

node configurations. The Oracle Grid Infrastructure simplifies and streamlines management of volumes, file systems and cluster configurations, therefore eliminating the need for multiple 3rd party software layers, complexity and cost.

Leveraging the New Flex ASM Architecture

Oracle CloudFS leverages the new Flex ASM architecture introduced in Oracle Database 12c. Flex ASM removes the requirement to have an ASM instance on all nodes of the cluster. This allows for much higher degree of flexibility, scalability and availability for file services to clients.

Oracle Flex ASM introduces new components and methodology shown on diagram II. This is a 5 node cluster configuration that is running a 2 node RAC database on hub nodes 3 and 4. Three ASM instances have been configured on hub nodes 1, 3 and 5 that provide ASM services to all the nodes in the cluster. These file services use an internal ASM Network that facilitates all communications. The RAC databases use local ASM services if available and remote ASM services if not locally available. Therefore, database instance on node 3 uses local ASM where the database instance on node 4 uses ASM services from a remote note (i.e. ASM1 on node 5).
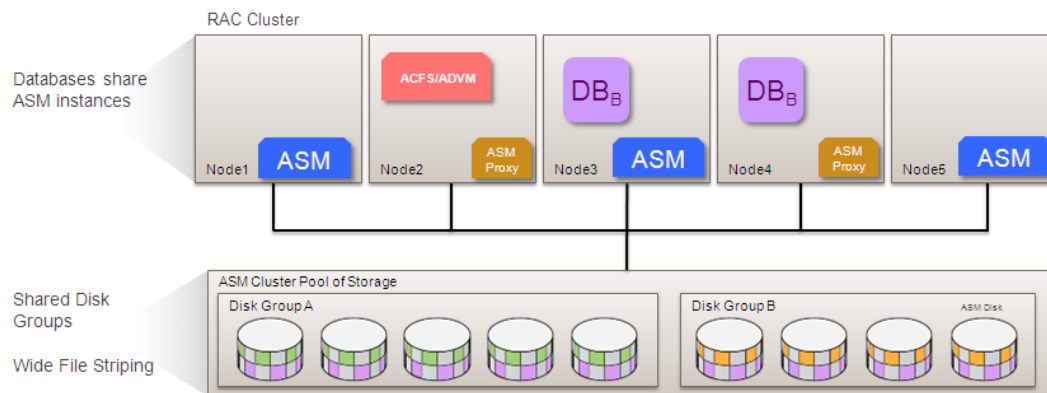

Diagram II

The ADVM volume management and ACFS file system services leverage the Flex ASM architecture. ADVM/ACFS utilize an ASM Proxy for satisfying metadata request to a remote ASM instance.

Oracle ASM Dynamic Volume Manager (ADVM)

Oracle CloudFS includes the Oracle ADVM that provides cluster volume management services and a standard device driver interface to its clients (ACFS, ext3, OCFS2…). Oracle ACFS and other 3rd party file systems can use ADVM as a volume management platform for creating and managing block devices for file systems. ADVM block devices leverages all the power and functionality of Oracle ASM. ADVM volumes are created in an ASM disk group.

An O/S logical block device is automatically created in the operating system when an ADVM volume is created in an ASM disk group. Oracle ADVM compliments Automatic Storage Management extent and striping algorithm to ensure the highest reliability and performance for general purpose and database files. It uses Dirty Region Logging (DRL) for mirror recovery after a node or instance failure.

Automatic Storage Management Cluster File System (ACFS)

Oracle CloudFS includes Oracle ACFS which is a general-purpose POSIX, X/OPEN and Windows compliant file system designed for single node and cluster configurations. This means that Oracle ACFS is managed by using native operating application programming interfaces and command line tools that administrators are already familiar with. An ACFS file system is created on top of an ADVM dynamic volume. These Dynamic Volumes are ASM files in a disk group, and therefore, leverage and benefit from all the powerful ASM features and functionality as an integrated solution.

Starting with Oracle Database 12c, Oracle ACFS supports Oracle database files in addition to general purpose files. This means that entire Oracle databases can be stored inside an ACFS file system. Oracle ACFS implements direct I/O bypassing the OS buffer cache to deliver high performance, similar to raw, for database files. However, general purpose files are still provided cached I/O suitable for non-database files that provide fast response time and high performance.

All databases stored on Oracle ACFS can now leverage Oracle CloudFS Advanced Data Services. Snapshots may be used for backup as well as provisioning sparse copies of databases for testing and other purposes. Snapshots can also be used to create read-write snapshots of Oracle Database Homes to simplify out-of-place patching. Oracle ACFS Replication compliments Oracle Data Guard and provides disaster recovery for

non-database files. The Tagging feature may be used to simplify aggregate operations based on assigned tags. And, the Security, Encryption and Auditing feature may be used for more granular access control and security overall.

Oracle ACFS supports large files with 64 bit file and file system data structures enabling exabyte file system capacities on 64 bit platforms. Oracle ACFS provides for variable extent-based storage allocation, direct path file I/O to ASM storage resources and the high performance directory architecture contributes to fast performance. File system integrity and fast recovery is achieved via metadata check-summing and journaling. ACFS is designed as a peer-to-peer, multi-node, shared file system model and delivers coherent, cached, direct file I/O to ASM pooled storage resources from each cluster node.

Oracle ACFS leverages Oracle Database ASM features that enable dynamic file system resizing, maximized performance through direct access to ASM disks, and balanced data distribution across the disks within an ASM disk group. Additionally, Oracle ACFS provides data redundancy through mirroring as implemented in ASM.

Oracle ACFS may be used to store Oracle database files, general purpose files like business data, application data and personal files. Oracle database home binaries, middleware and application binaries and configuration files are great candidate for using Oracle ACFS.

Oracle ACFS can be managed using command-line tools (OS native, ASMCMD and ACFSUTIL), Oracle Enterprise Manager, ASM Configuration Assistant and the SQL command interface. Oracle ACFS can also be configured for access by remote clients using industry standard NAS file access protocols: NFS and CIFS. Oracle ACFS provides for a single-node mount registry (similar to fstab in Linux) and the Oracle Grid Infrastructure cluster-wide mount registry. An Oracle ACFS file system is automatically mounted if a corresponding entry is inserted into the mount registry.

Leveraging Automatic Storage Management (ASM)

Both Oracle ADVM and Oracle ACFS leverage the Oracle ASM features of Oracle Database to consolidate and easily provision pooled storage resources for all volumes and file systems. Oracle ASM reduces overall storage costs and increases storage utilization while improving performance and availability over traditional file system and

volume management solutions. A fraction of the time is needed to manage your storage environment with Oracle ASM.

Oracle ASM improves manageability by simplifying storage provisioning, storage array migration, and storage consolidation. Oracle ASM provides flexible easy to manage interfaces including the SQL*Plus, Oracle Enterprise Manager GUIs and a UNIX-like asmcmd command line interfaces. It provides sustained and best in class performance because of its innovative rebalancing feature that distributes data evenly across all storage resources, providing an even distribution of IO and optimal performance. It scales to multi-TB databases efficiently without compromising functionality or performance. Oracle ASM provides the following benefits:

- Simplifies and automates storage management
- Increases storage utilization, uptime and agility
- Delivers predictable performance and availability

Advanced Data Services

Oracle ACFS High Availability NFS Services (HANFS)

High Availability Network File Storage (HANFS) for Oracle Grid Infrastructure provides uninterrupted service of NFS V2/V3 exported paths by exposing NFS exports on Highly Available Virtual IPs (HAVIP) and using Oracle Clusterware agents to ensure that the HAVIPs and NFS exports are always online. If a cluster node fails the HAVIP and NFS exports are automatically migrated to a surviving node.

The HANFS feature enables highly available NFS servers to be configured using Oracle ACFS clusters. The HANFS cluster configurations may be built from your existing infrastructure or commodity servers and storage. This provides network services similar to 'network filer' at the fraction of the cost.
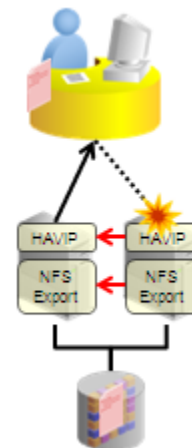
Oracle ACFS Snapshots

Oracle ACFS Snapshots may be used in test and development and recover from inadvertent modification or deletion of files from a file system. A snapshot can also be used as the point-in-time source for a file system backup, as it can be created on demand to deliver a current, consistent, on-line view of an active file system.

An Oracle ACFS Snapshot is a read-write or read-only on-line, space efficient, point in time copy of a file system. The snapshot copy is initially sparse and merely references the storage allocation information maintained by the source file system and therefore, requires minimal storage capacity because it maintains only the changed blocks. An Oracle ACFS Snapshot utilizes a Copy-On-Write (COW) technology to maintain point in time view of the file system by copying the current extent to the snapshot before modifying or deleting the source file extent.

Oracle ACFS Snapshots are immediately available for use after they are created. Snapshots are always on-line while the file system is mounted.

Oracle ACFS Tagging

Oracle ACFS Tagging is a feature that allows a user to associate one or more files together as a group by assigning a unique 'tag name' attribute. Group operations can be performed based on tagged files that may span across different directories within an Oracle ACFS within a single node or a cluster. Different groups of tagged files (e.g. 'medical imaging', 'photo album', etc) may be replicated as groups of related files, complementing the replication of database files using Oracle Data Guard.

Oracle ACFS Replication

Oracle ACFS Replication is a Disaster Recovery (DR) solution that enables replication of an Oracle ACFS file system across the network from a primary to a standby site, providing disaster recovery capability for the file system. Either the primary or the standby sites can be a single node or a cluster of nodes. A primary site for one file system can be the standby site for a different file system and vice-versa.

Oracle ACFS Replication captures file system changes on the primary site and records the changes in files called replication logs. These logs are transported to the standby site asynchronously where background processes read the logs and apply the changes recorded in the logs to the standby file system. After the changes are recorded in a

replication log and have been successfully applied to the standby file system, the replication log is deleted on both the primary and standby sites.

Oracle ACFS Replication compliments Oracle Data Guard and provides the customer with an end-to-end DR solution for all files. You no longer need to rely on 3rd party replication software for external files and Oracle binaries. Oracle ACFS Replication is very easy to configure and implement.

Oracle ACFS Security

Oracle ACFS Security can be used to protect sensitive data from internal and external threats by providing fine grained access control on top of the access control provided by the operating system.

Oracle ACFS Security provides realm-based security for Oracle ACFS file systems enabling users to create realms and specify security policies for users and groups and therefore control access on file system objects. Oracle ACFS Security uses realms, rules, rule sets, and command rules to enforce security policies. The Oracle ACFS realm is virtual container of files and directories; access defined by security filters (command rules and rule sets).

An Oracle ACFS security administrator manages security for all Oracle ACFS file systems with a Security administration password that is different from his/her OS password. An Oracle ACFS Security administration password is stored in an Oracle Wallet for additional security, and all security commands are protected by the realm management password.

Oracle ACFS Encryption

Encryption is another feature that may help meet regulatory compliance requirements by keeping data in an Oracle ACFS file system in an encrypted format and hence protecting against unauthorized use of such data in the case of data loss or theft. Oracle ACFS Encryption enables users to encrypt data stored on disk (data-at-rest) and provides secured encryption keys to decode data.

Oracle ACFS Encryption can be applied to the entire file system or just individual files and directories. It is completely transparent to authorized users, and applications work unchanged with encrypted files. Encrypted and unencrypted files can co-exist on the

same file system. System administrators and Oracle ACFS security administrators can initiate encryption operations.

Each file is protected by two keys, the File Encryption Key (FEK) and the Volume Encryption Key (VEK). File data is encrypted using a unique FEK; the FEK is stored on disk, and is encrypted using a VEK. VEKs are stored in an Oracle Wallet and can be password protected.

Oracle ACFS Auditing

Oracle ACFS Auditing provides auditing capabilities for Oracle ACFS security and encryption. This auditing framework produces a separate audit trail, or audit source, for each Oracle ACFS file system on each individual node, and enforces separation of duties regarding the management and review of the audit source.

Audit sources, can be generated from Oracle Database audit trail tables, database operating system audit files, and database redo logs. Both Oracle ACFS security and encryption are also audit sources, and these sources can be enabled and disabled by an Oracle ACFS audit manager.

Installation and Configuration Tools

Installation and configuration tools have been developed to allow Oracle ASM, Oracle CloudFS and Oracle Clusterware installation and configuration independent of Oracle databases. This provides the flexibility for aligning roles and responsibilities for organizations where system administrators are tasked with managing infrastructure software such as Oracle ASM and Oracle Clusterware while DBAs continue to have responsibility for Oracle Database installation and configurations.

Oracle Universal Installer (OUI) and ASM Configuration Assistant (ASMCA) (diagram III) are system admin friendly tools and are used to install and configure Oracle ASM, Oracle CloudFS and Oracle Clusterware in a single Grid Infrastructure home 'Grid'. The OUI installation has the option to minimally configure Oracle ASM and Oracle Oracle ACFS. It also automates creation of an Oracle Database home on Oracle ACFS in preparation for a database home to be installed in the database installation phase.
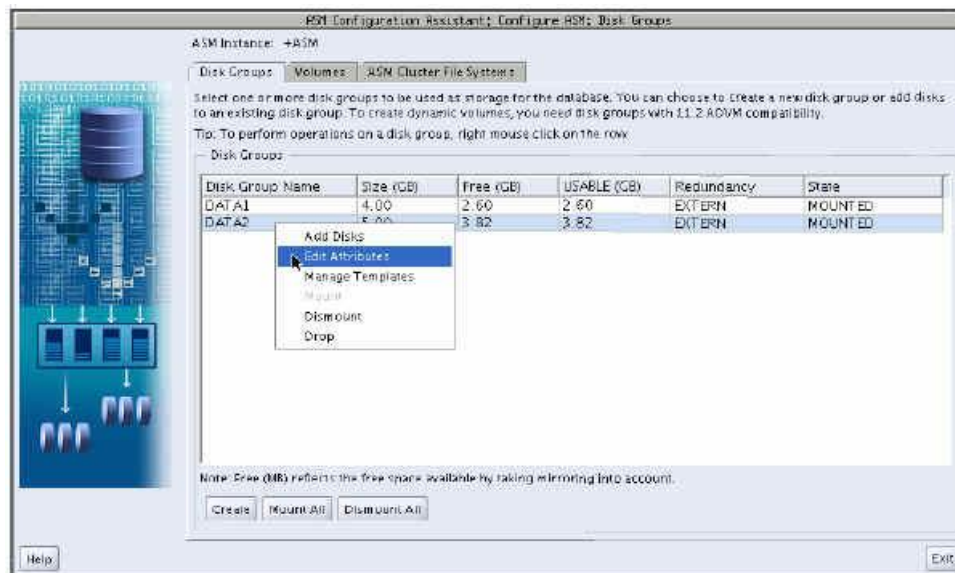
Diagram III

Oracle ASM Configuration Assistant (ASMCA) is a graphical tool that is designed to simplify Oracle ASM configuration management especially for the system administrators. ASMCA simplifies initial configuration and management for Oracle ASM disk groups, Oracle ADVM and Oracle ACFSs.

Conclusion

Oracle customers benefit from a single integrated solution designed to manage general purpose data as well as database data across multiple operating system platforms with one management interface, one set of installation and configuration tools, one clusterware framework and a single vendor for support and troubleshooting. Oracle ASM and CloudFS eliminates the need for 3rd party volume managers and file systems.

Oracle CloudFS and Oracle ASM are designed to be always-on-line, increasing file system and database availability despite storage hardware failures and storage configuration changes. Oracle CloudFS continues to bring innovations in automation to further simplify file and storage management. Oracle CloudFS when combined with Oracle ASM features of the Oracle Database provide for the highest performance and

continuous balanced I/O distribution without tedious complex administration. They provide a scalable solution that makes them suitable for very small as well as multi-hundred terabyte file systems. Oracle ASM disk groups function as storage containers that are ideal for consolidation of databases and file systems that result in optimal storage utilization and reduce the storage waste in computing environments.

ORACLE