

An Oracle White Paper
January 2015

Benefits of Oracle ACFS

Introduction	1
Performance.....	2
Oracle ACFS Advanced Functionality.....	4
Simple Point-in-Time Read-Write Snapshots.....	4
Oracle ACFS High Availability NFS Services.....	4
Oracle ACFS Tagging	5
Oracle ACFS Replication.....	5
Oracle ACFS Security	5
Oracle ACFS Encryption.....	6
Oracle ACFS Auditing	6
Self-Tuning and Resilience.....	6
Disk Space Scalability and Performance Factors.....	7
Automatic Resource Management and Automation	7
Oracle ACFS on Oracle Database Appliance	8
Oracle ACFS on Linux Exadata Database Machines.....	9
Oracle ACFS Use Cases.....	10
Conclusion	11

Introduction

Oracle ASM Cluster File System (ACFS) provides unprecedented simplicity, automation and functionality in storage and file management for all file types including Oracle database and general purpose files.

Oracle ACFS is a general-purpose industry standard POSIX, X/OPEN and Windows compliant cluster file system supporting multiple operating systems and server platforms, including Oracle Exadata Database Machine, Oracle Database Appliance (ODA) and traditional server environments. Oracle ACFS includes advanced features such as file system snapshot, replication, tagging, security, encryption, auditing and highly available NFS (HANFS) services.

Oracle ACFS/ASM software stack is bundled with a general purpose cluster volume manager, the Oracle ASM Dynamic Volume Manager (Oracle ADVM), which provides common volume management services and a standard device driver interface for system administrators to manage volumes across different nodes of a cluster with the same functionality on multiple platforms.

Oracle Database and Oracle ACFS both are clients of Oracle Automatic Storage Management (ASM). Oracle ACFS makes use of ASM files and inherits ASM file features including striping, mirroring, rebalancing, intelligent data placement, preferred read, fast resync, even read, flex ASM and other features.

The combination of Oracle ACFS, ADVM, Oracle ASM and Oracle Clusterware make up the Oracle Grid Infrastructure (GI). Oracle GI is a tightly integrated foundation to manage databases, volumes and file systems in cluster environments. Oracle GI simplifies storage and cluster management, and greatly reduces complexity and costs by eliminating the need for 3rd party software.

Oracle ACFS 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, to simplify system and storage management, and to provide high performance without the need for custom tuning. Oracle ACFS optimizes utilization through consolidation and ensures file system availability.

Performance

Oracle ACFS is designed to deliver the best possible performance with Oracle databases and general purpose files. Oracle ACFS obtains an extent map list and disk list information from the Oracle ASM instance whenever it opens an ASM file. This means, neither Oracle ASM nor Oracle ACFS are in the I/O path when accessing files on disk. Oracle ACFS caches ASM disk mapping information and uses that information on subsequent ACFS I/O operations to directly read and write to ASM disk group disks.

Oracle ACFS provides a direct I/O path from the database through to the ASM disk group devices and bypasses the traditional OS VM page caching mechanism used for file system applications. This model enables databases configured to use ACFS to achieve performance metrics comparable to databases configured to use Oracle ASM directly.

Comprehensive internal benchmarks using OLTP and DSS workloads have proven the performance of Oracle ACFS and ASM to be similar. The following is a summary of the performance data and the benchmark environment used for those tests:

Hardware: Two node Linux cluster with 72 GB RAM and 12 CPUs

Software: Swingbench 2.4.0.873, Oracle Database 12c

Storage: Two ASM disk groups with normal redundancy setup

- one disk group for data and one for the recovery area
- For data, 5 shelves of the disk array were used, total of 80 physical disks, configured as RAID10. In case of DSS 120 disks on 5 shelves were used.
- For recovery area, 1 shelf was used, total of 16 physical disks, configured as RAID 10.
- The disk groups were created with enough space to accommodate 2 databases: one of them using ACFS and the other using ASM directly.
- One ACFS file system was created on each of the above disk groups.

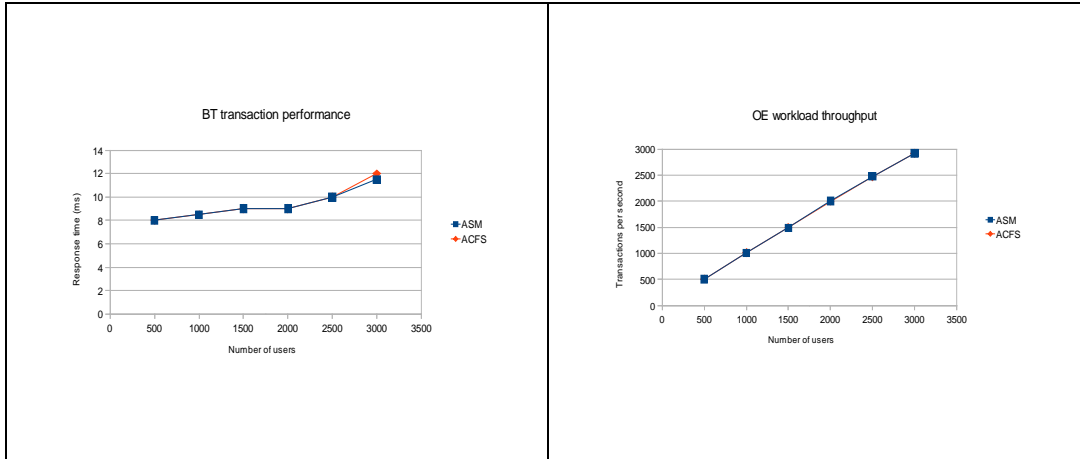
Database:

- Two identical databases were configured, with files stored in ACFS and ASM.
- Two database instances per database were used.
- Archivelog and flashback database enabled (disabled for DSS).

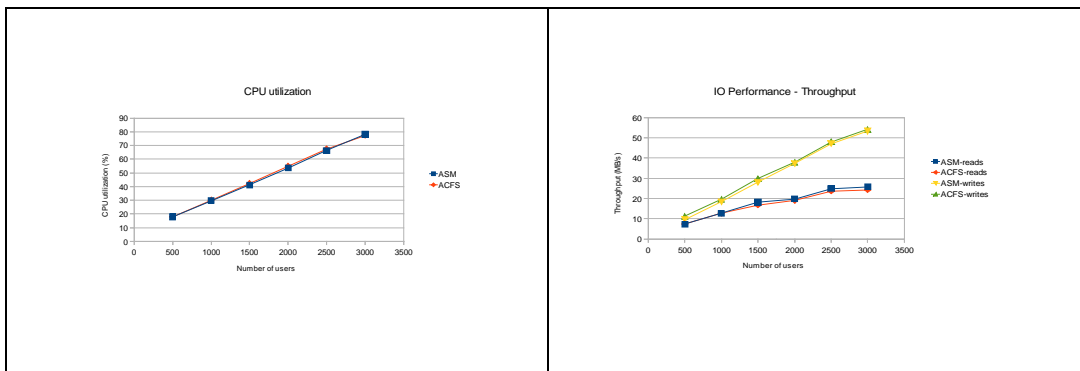
Swingbench Schema:

- OE Schema (OLTP) created with version 2, no partitioning, no compression, bigfile tablespace, all indexes, sizing factor 6 (= 19 GB Tablespace)
- SH Schema (DSS) created with range partitioning, no compression, bigfile tablespace, all indexes, sizing factor 41 (= 70 GB Tablespace)

The following chart shows performance based on OLTP workload. As you can see, the throughput performance and transaction response time are almost identical regardless of whether the database use ACFS or ASM directly.



The operating system and CPU loads show no significant difference between Oracle ACFS and ASM; neither did the DSS workload as shown in the following:



	ASM	ACFS	% diff
SH throughput (TX/min)	3.29	3.26	-1%
IO throughput (Reads / sec)	3501	3312	-5%
IO throughput (MB Read / sec)	724	712	-2%
CPU usage (%)	16.5	16.7	+1%

DSS workload response time:

	ASM	ACFS	% diff
Response time (sec)	24.0	23.6	-2%

Oracle ACFS Advanced Functionality

Simple Point-in-Time Read-Write Snapshots

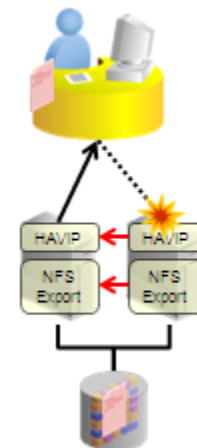
An Oracle ACFS snapshot is an online, read-only or read-write, point in time copy of an Oracle ACFS file system. The snapshot copy is space-efficient and uses Copy-On-Write functionality. Before an Oracle ACFS file extent is modified or deleted, its current value is preserved in the snapshot to maintain the point-in-time view of the file system.

Oracle ACFS supports 1023 RO/RW snapshots and snaps of snapshots allowing a complete hierarchy of snapshots for any file system, with the ability to arbitrarily delete or add a snapshot anywhere in the hierarchy and support full inheritance.

Oracle ACFS snapshot storage is maintained within the file system, eliminating the management of separate storage pools for file systems and snapshots. All ACFS snapshots are always available when the ACFS file system is mounted. Oracle ACFS file systems can be dynamically resized online to accommodate additional file and snapshot storage requirements.

Oracle ACFS High Availability NFS Services

High Availability Network File Storage (HANFS) for Oracle Grid Infrastructure provides un-interrupted service of NFS v2/v3/v4 exported paths by exposing NFS exports on Highly Available Virtual IPs (HAVIP). Oracle Clusterware agents ensure that the HAVIPs and NFS exports are always online. If a cluster node fails, the HAVIP and NFS services are automatically taken over by 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 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 file system and 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 updates in real-time and transfers changes asynchronously to the standby file system. The standby file system is available for reading and RW snapshot of the RO file system can be created. The standby file system is available in read-only mode during replication, however, read-write snapshots of the file system can be created for other applications.

Oracle ACFS Replication complements Oracle Data Guard and provides the customer with an end-to-end DR solution for all files. The transaction time stamps are recorded within the log files so that recovery can be synchronized with Oracle Data Guard for fail over or role reversal.

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 offered 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 thereby 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 a virtual container of files and directories for which access is defined by security filters.

Oracle ACFS Encryption

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

Oracle ACFS Encryption can be applied to the entire file system or just individual files and directories. It is completely transparent to authorized users. In addition, 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.

Using Oracle ACFS Encryption, 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.

Self-Tuning and Resilience

Oracle ACFS supports large files with 64-bit file and file system data structure sizes leading to exabyte capable file and file system capacities on 64 bit platforms. Variable extent-based storage allocation and high-performance directories contribute to fast performance and shared disk configurations that provide direct storage paths to Oracle ACFS data from each cluster member. File system integrity and fast recovery is achieved with Oracle ACFS metadata checksums and journaling.

Oracle ACFS is designed as a multi-node, shared file system model that delivers coherent, cached, direct storage paths to Oracle ACFS file system data. It presents single system file access semantics across cluster configurations. Applications and users on all cluster members are always presented with the same view of shared Oracle ACFS file data, supported by cluster-wide user and metadata cache coherency mechanism.

Disk Space Scalability and Performance Factors

Oracle ACFS supports 2^{40} (1 trillion) files in a file system. Oracle ACFS supports 256 mounted file systems on 64-bit systems. However, more file systems can be mounted if there is adequate memory. ACFS makes use of 4K block size allocation units and up to 128MB extent sizes when managing ACFS file storage. Benchmarks have shown that the 4k default block size works best in all cases and therefore eliminates the need for tuning. ACFS automatically issues storage disk I/Os in sizes up to the maximum supported by the storage I/O subsystem eliminating the need for per-OS consideration and configuration.

Oracle ACFS pre-allocates large user files to improve performance when writing data. Oracle ACFS also allocates local metadata files when nodes mount the file system for the first time. Oracle ACFS keeps local bitmaps available to reduce contention on the global storage bitmap when searching for free space. This disk space is reported as “in-use” by tools such as the UNIX `df` command, although some of the space may not actually be allocated as of yet. This local storage pool can be as large as 128 megabytes per node and can allow space allocations to succeed, while commands such as `df` report less space available than what is being allocated.

Automatic Resource Management and Automation

Oracle Clusterware resources support Oracle ACFS, Oracle ADVM, and Oracle Kernel Services Driver (OKS) startup, cluster mount registry, and single file system startup, shutdown, and steady-state actions. These resources are configured automatically at installation time and therefore do not require any management. The drivers are managed as a single resource set.

Whenever Oracle Clusterware is started on a cluster node, the Oracle ACFS startup operations for the node consult the cluster mount registry and attempt to mount all Oracle ACFS file systems that are registered for this node.

In order to manage dependent resources as a single item, an Oracle ACFS file system configured as the underlying file system for an Oracle Database home, includes a

resource dependency to the Oracle Database. This dependency causes the stack to be started and the file system to be automatically mounted prior to the Oracle Database being started.

Oracle ACFS on Oracle Database Appliance

The 12.1.2 Oracle Database Appliance (ODA) release(s) has adopted Oracle ASM Cluster File System (ACFS) as the primary cluster file system to store database files and general purpose data. Three default file systems are created automatically in the Oracle Database Appliance ASM disk groups (DATA, RECO, and REDO) when using the 12.1.2 release. When new databases are created, the associated database files are created in these three ACFS file systems instead of in the ASM disk groups directly.

Note that when you are upgrading from a previous release to 12.1.2, the databases that were previously created on ASM are not automatically moved to ACFS, but can co-exist with the new databases that are created on ACFS. Customers can choose to migrate existing databases from ASM to ACFS post-upgrade by following the steps outlined in the white paper titled 'Steps to Migrate Non-CDB databases to ACFS on ODA 12.1.2'.

Oracle ACFS was chosen as the primary file system for Oracle Database Appliance because it provides:

- Increased functionality without requiring additional management
- Equivalent performance to Oracle ASM
- Industry standard and simple user interface
- Database snapshots to quickly and easily provision test and development environments at a fraction of the disk storage required with full database copies
- Advanced functionality for general purpose files such as replication, tagging, encryption, security, and auditing

Oracle ACFS file systems benefit from all the Oracle ASM functionality such as striping, mirroring, rebalancing and dynamic capacity management that make both Oracle ASM and ACFS highly performing, scalable and highly available platforms.

Database creation on ACFS follows the same Oracle Managed Files (OMF) principles that are used with Oracle ASM. For every database creation, new directories are created in

DATA, RECO, and REDO file systems for the new database files (similar to Oracle ASM). Oracle ACFS in the Oracle Database Appliance requires no administration. The Appliance Manager automatically manages the storage; including ACFS file systems and the underlying ASM disk groups.

Internal testing of database workloads has show similar performance for Oracle ACFS and ASM. Both Oracle ACFS and ASM provide direct I/O to disk resulting in high performance for database files. In addition, Oracle ACFS performs buffered I/O for non-database files that benefit from cached data in the system memory. The workload performance details are described in the 'performance' section of this paper.

Oracle ACFS on Linux Exadata Database Machines

Starting with Oracle Grid Infrastructure 12c Release 12.1.0.2, Oracle ACFS supports all database files and general purpose files on Oracle Exadata Database Machine running Oracle Linux on database servers. This makes ACFS an attractive storage alternative for data loads, extracts, migrations, and other use cases for storing general purpose files. In addition, Oracle ACFS supports database files on Oracle Exadata Database Machines, specifically; Oracle ACFS supports all database file types for the following database versions:

- Oracle Database 10g Rel. 2 (10.2.0.4 and 10.2.0.5)
- Oracle Database 11g (11.2.0.4 and higher)
- Oracle Database 12c (12.1.0.1 and higher)

Please note that Oracle ACFS does not currently support the Oracle Smart Scan (predicate processing) and is not able to push database operations directly to storage. For best performance, store database files on the Exadata Storage using ASM. Since Exadata Storage Servers do not support Oracle Database 10g, Oracle ACFS provides a storage option for customers wishing to host older databases on their Exadata server. Also, while Oracle ACFS snapshots and tagging are supported for database files, Oracle ACFS replication and security/encryption/audit are only supported with general purpose files, and cannot be used with database files.

Oracle ACFS supports all database file types and general purpose files. All Oracle ACFS advanced functionality on the Exadata platform are supported for general purpose files including snapshots, replication, tagging, security, encryption, audit and HANFS.

Oracle ACFS Use Cases

Oracle ACFS is a true cache coherent cluster file system supporting concurrent read-write operations guaranteeing that clients on all nodes of the cluster have exactly the same consistent view of data. There is no extra effort to install ACFS since it is installed during the Oracle Grid Infrastructure deployment and is easily configurable for different use cases in your environment.

Oracle ACFS is the ideal file system for Oracle RAC cluster and standalone deployments as well as Oracle application and middleware environments, while it can be used with other any application vendor. The typical use cases for ACFS are:

- Support 10g Oracle Databases on Linux based Oracle Exadata Database Machines. This makes Exadata a more flexible platform enabling consolidation and easy migration for customers.
 - Benefit: Consolidation
- Store Business data and unstructured data such as scanned documents, image files, and BLOB data types in an Oracle ACFS file system as an alternative to storing them inside a database.
 - Benefits: Ease of use and manipulations
- Use Oracle ACFS for shared Oracle Database Homes and as a data store for all administrative, log files, trace and audit files.
 - Benefit: Manage and share one set of binaries and administrative files
- Facilitate Oracle ACFS Replication to complement Oracle Data Guard (DR for Oracle Databases) as a DR solution for all non-database files. This allows for Database recovery to be synchronized with the Oracle ACFS file system by leveraging the new 'time stamp' feature introduced in Oracle ACFS 12.1.0.2.
 - Benefit: No need to use 3rd party solutions
- Consolidate monitoring of audit files and log files from many systems in your data center using Oracle ACFS Replication (audit and log files can be configured to be replicated into one Monitoring server).
 - Benefits: Centralize your monitoring and simplify management
- Use Oracle ACFS snapshots for creating point in time snapshots. The read-write snapshots allow you to present data for test and development environment and

also allow for taking backups on demand. The 'gDBClone' script that is downloadable from the Oracle Technology Network together with a best practice white paper facilitates ACFS snapshots for an easy deployment and management of test and development environments.

- Benefits: Eliminate 3rd party snapshot solutions and automate database provisioning
- Use Oracle ACFS NFS/CIFS network protocols to deliver NFS/CIFS exported file systems and leverage the HANFS feature to provide highly available network file systems to clients.
 - Benefits: Highly available and simple sharing of file systems across the network
- Oracle ACFS Security and Encryption features are complimentary solution to Oracle Advanced Security Option (ASO) and Data Vault for the general purpose files. This feature may be used for a highly granular file access control and encryption for data at rest.
 - Benefits: Tighter file access control and data file protection through encryption
- Oracle ACFS Tagging feature enables group operations through 'tags'. A great application of Tagging is to use it in conjunction with ACFS Replication and allow only the tagged files to be replicated.
Benefits: Ease of management

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 including Oracle Exadata Database Machine, Oracle Database Appliance and traditional non-engineered servers with one management interface, one set of installation and configuration tools, one clusterware framework and a single vendor for support. Oracle ASM and ACFS eliminate the need for 3rd party volume managers and file systems on Oracle platforms greatly simplifying the storage management stack.

Oracle ACFS 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 ACFS continues to bring innovations in automation to further simplify file and storage management. Oracle ACFS, when combined with Oracle ASM, 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.



Benefits Oracle ACFS
January 2015 Update
Author: Ara Shakian
Oracle Corporation
World Headquarters
500 Oracle Parkway
Redwood Shores, CA 94065
U.S.A.

Worldwide Inquiries:
Phone: +1.650.506.7000
Fax: +1.650.506.7200
oracle.com



Oracle is committed to developing practices and products that help protect the environment

Copyright © 2014, Oracle and/or its affiliates. All rights reserved. This document is provided for information purposes only and the contents hereof are subject to change without notice. This document is not warranted to be error-free, nor subject to any other warranties or conditions, whether expressed orally or implied in law, including implied warranties and conditions of merchantability or fitness for a particular purpose. We specifically disclaim any liability with respect to this document and no contractual obligations are formed either directly or indirectly by this document. This document may not be reproduced or transmitted in any form or by any means, electronic or mechanical, for any purpose, without our prior written permission.

Oracle is a registered trademark of Oracle Corporation and/or its affiliates. Other names may be trademarks of their respective owners.