ORACLE
DATABASE

Oracle Advanced Compression Helps Global Fortune 500 Company Meet Storage Savings Initiative

ORACLE WHITE PAPER | FEBRUARY 2017

ORACLE

Table of Contents

Introduction

This whitepaper provides an overview of the evaluation process a Global Fortune 500 corporation completed to test the various features included in Oracle Advanced Compression option, as well as the impact on space and performance they experienced.

Using actual customer data, Oracle Database 11g compression and space optimizations reduced the overall size of the company's disk consumption, on their proof-of-concept test systems, by 50% for structured data and up to 98% for unstructured data. As the customer's testing demonstrates, the benefits from Advanced Compression can be significant, yet some organizations may still have uncertainty regarding the performance impact that compression may have on their systems.

This customer's demonstration confirms that Oracle Advanced Compression option provides *storage savings and increased performance* with no significant increase in overall CPU consumption. Following the completion of their proof-of-concept, the company purchased the Advanced Compression option and began to quickly roll out their Advanced Compression deployment.

Organizations have growing requirements to run business applications and analytics faster and faster in spite of the growth in storage volumes. The capabilities provided by Oracle Advanced Compression option enable many customers to run the database in an in-memory mode by allowing the entire working set of the application to fit in memory after 2x to 4x compression – giving applications a significant boost in performance.

Oracle Advanced Compression option, introduced with Oracle Database 11g, provides tremendous space savings as well as performance improvements. In addition, given the growth experienced by today's databases, where operations like backups are increasingly taking longer to complete, Advanced Compression can help make those operational tasks faster, thereby providing both storage efficiency and performance as benefits.

Oracle Advanced Compression

Oracle Advanced Compression is an option introduced in Oracle Database 11g which includes a comprehensive set of compression capabilities to help customers maximize resource utilization and reduce costs. It allows IT administrators to significantly reduce their overall database storage footprint by enabling compression for all types of data – be it relational (OLTP table compression), unstructured (SecureFiles deduplication and compression), backups (RMAN and Data Pump) or network traffic (Data Guard redo transport compression).

Although storage cost savings are often seen as the most tangible benefit of compression, innovative technologies included in the Advanced Compression option are designed to reduce resource requirements and technology costs for all components of an IT infrastructure, including memory and network bandwidth, while also improving performance.

About the Company

This case study describes the results from a Proof-of-Concept conducted by a Global Fortune 500 company.

Data Growth Challenges / Company Expectations

Exponential increases in data volumes in recent years have put enterprise IT infrastructures under severe pressure – from a cost, performance, scalability and manageability perspective. It has become imperative to employ more efficient ways of storing and managing data to meet the growing demands being placed on IT systems.

The rapid growth in data volumes, and associated costs, were of significant concern to the company – so much so that the company started a global initiative to identify and drive a reduction in storage costs.

The company's IT organization identified data compression as an intelligent and efficient solution to address their data storage growth and costs challenges. As the company began to perform their due diligence on Oracle Database 11g Advanced Compression option their initial expectations began to take shape, which included:

- The company expected their storage costs to drop as the result of any compression deployments, and
- The company expected no query performance degradation and only minimal Data Manipulation Language (DML) performance impact from compression

Proof-of-Concept (POC) Test Environment

The customer performed Proof-of-Concept testing and analysis of the Oracle Database 11g Advanced Compression option to determine if the capabilities provided by the option matched their goals, namely to minimize storage space usage/costs across Oracle database environments – with minimal or no performance impact.

The test environment infrastructure included:

- Sun Sparc 64-bit w/ Solaris (96 GB RAM, 12 CPU)
- Oracle Database 11.1.0.7
- Quest Benchmarking Utility

The company's Proof-of-Concept tested the following capabilities of Oracle Advanced Compression:

- Structured data compression (OLTP Table Compression)
- Unstructured data compression (SecureFiles Deduplication and Compression)
- Backup data compression (Data Pump Export and RMAN Compression)

Proof-of-Concept Testing Goals, Results and Conclusions

Compression Advisor Testing

Compression Advisor (available for free on the Oracle Technology Network) provides an estimate of potential storage savings that can be realized through the use of Oracle Advanced Compression OLTP Table Compression feature. The estimate is based on analysis of a sample of data and provides a good estimate of the actual results that will be experienced when OLTP Table Compression is utilized.

The compression ratio achieved in a given environment depends on the nature of the data being compressed, specifically the cardinality of the data. In general, Oracle Advanced Compression customers can expect to reduce their storage space consumption by a factor of 2x to 4x by using Oracle Advanced Compression OLTP Table Compression feature. That is, the amount of space consumed by compressed data will be two to four times smaller than that of the uncompressed data.

The company tested Compression Advisor against approximately 20 to 30 tables with the goal of determining the expected compression ratios for the tables and then later comparing the estimates against the actual compression ratios, to determine the overall reliability of the compression advisor tool.

Below are the results of running Compression Advisor -- the test was performed on a table with about 60 million records in an Oracle Database 11g Release 1 database taken from a production OLTP database.

Compression Advisor Output Example

```
exec DBMS_COMP_ADVISOR.getratio('TEST_USER','TEST_TABLE','OLTP',10);

Sampling table: TEST_USER.TEST_TABLE

Sampling percentage: 10%

Compression Type: OLTP

Compression ratio: 1.96[1]
```

The Compression Advisor output above shows, based upon a sampling of 10% of the TEST_TABLE data, the customer can expect a compression ratio of approximately 1.96x, or in other words, compression should reduce the storage requirements of the table by nearly 50%.

The actual compressed size of the table, after being built, proved to be a close match with advisor estimates, see below.

Compression Physical Testing (after table is compressed)

```
SQL>
SEGMENT_NAME                       MB      BLOCKS  COMPRESS  COMPRESS_FOR
---------------------------------- ------- ------- --------- ------------------
TEST_TABLE                         4405    563840  DISABLED
TEST_TABLE_COMP                    2268    290304  ENABLED   FOR ALL OPERATIONS
```

The compression estimate, produced by Compression Advisor based upon a 10% sampling, asserted a compression ratio of 1.96%, which means the 4405MB table would have compressed to 2237MB based upon the compression estimation. In reality, once the table was physically compressed, the actual measured compression of the table saw the table compress from 4405MB to 2268MB, nearly identical to Compression Advisor's estimate.

[1] The compression ratio achieved in a given environment depends on the nature of the data being compressed, specifically the cardinality of the data.

The compression ratio achieved in this case indicates that the table has duplicate values that OLTP Table Compression was able to compress -- low cardinality data usually results in higher compression ratios. Further, the small difference between the Compression Advisor estimate and the actual OLTP Table compression enabled the customer to use Compression Advisor with a reasonable expectation that the estimates would be consistent with the actual compression as additional tables are compressed.

Structured Data Testing (OLTP Table Compression)

OLTP Table Compression often improves read operations, because database blocks stay compressed when they're read into the buffer cache. There is additional work performed while writing data, making it impossible to eliminate performance overhead for write operations. However, Oracle Advanced Compression minimizes this overhead with the OLTP Table Compression feature.

With Oracle Advanced Compression's OLTP Table Compression feature, blocks are compressed in batch mode rather than compressing data every time a write operation takes place. A newly initialized block remains uncompressed until data in the block reaches an internally controlled threshold. When a transaction causes the data in the block to reach this threshold, all contents of the block are compressed in one operation. Subsequently, as more data is added to the block and the threshold is again reached, the entire block is recompressed to achieve the highest level of compression.

This process repeats until Oracle Advanced Compression determines that the block can no longer benefit from further compression. Only transactions that trigger the compression of the block will experience the slight compression overhead. Therefore, a majority of OLTP transactions on compressed blocks will have the exact same performance as they would with uncompressed blocks.

Test Scenario and Goals

No industry-specific test scenario was developed by the company for Oracle Advanced Compression structured data testing. Instead, the company utilized a pre-packaged, inventory management system test environment (open order, process order, delete order and then close order) that was included with their benchmark tool.

- Data was loaded connecting first to a user with the default tablespace uncompressed; after that test was finished the configuration was changed to a user with the default tablespace compressed, and the test was run again
- Testing consisted of loading 9 tables with indexes starting with 10 users and ending with 100

The company had no pre-set expectations prior to beginning the proof-of-concept in terms of any overall performance impact as a result of compression -- but a primary goal of the proof-of-concept testing was to determine what, if any, performance impact would result from Advanced Compression.

To ensure that the company tested for, and understood any performance implications on their OLTP system, the performance included not only QUERY testing, but UPDATE and DELETE testing as well.

Test Results and Conclusions

Note: All times in minutes/seconds

Table Conditions	SELECT operation (select count (*))	DELETE Operation	UPDATE Operation (one column update)	UPDATE Operation (selected records update)
Not Compressed (60 Million Records)	28.04			12:06.36

Compressed (60 Million Records)	9.92			15.33
Not Compressed (100K)		5.93	4.51	
Compressed (100K)		7.56	2:13.38	

Source: Customer Tests

- Full table scan on "select count (*)" shows excellent results with compression. It took almost 1/3rd less time to do a record count of a 60 million record table (28.04 seconds vs. 9.92 seconds)
- DELETE DML on compressed data was ~20% slower than on uncompressed table (5.93 seconds vs. 7.56 seconds)
- One Column UPDATE DML showed slower performance with one column update for the whole table, while doing a selected records update shows 48 times better performance (last column in table) (4.51 seconds vs. 2:13.38 minutes). **Note:** After their Proof-of-Concept testing Oracle helped identify some issues with bulk updates to OLTP table-compressed tables and these issues have been resolved with 11.2.0.2. There are some patches that can be applied to 11.2.0.1 or 11.1.0.7 that address these issues. See MOS Note #1061366.1 for the specific patches.
- Selected Records UPDATE DML showed improved performance of 48 times better performance (12:06.36 minutes vs. 15.33 seconds)
- **Note:** Although not included on the table above, load tests showed no performance impact with Oracle Advanced Compression enabled.

Key Takeaways

Although the test scenario was deemed sufficient for the proof-of-concept, overall the company determined that the best test environment is where you can most closely duplicate the production environment with a test environment – but they acknowledge that due to time, budget and infrastructure limitations this would have proved difficult to provide. The company felt that sufficient application testing is required to fully understand the performance characteristics of their data when using compression since performance cannot be 100% predicted.

Overall, the customer reported that the majority of their performance testing proved that the use of Oracle Advanced Compression's OLTP Table Compression feature had no, or very little, impact on their overall OLTP performance – in fact their test queries took 3x less time, deletes weren't significantly impacted and selected record updates showed 48x improvement on compressed data. The testing also helped them identify some areas, rarely encountered in their environment, as having a longer operations window with compression than without -- such as a one column update across the entire 60 million record table.

Additional key takeaways realized by the company, based upon their performance testing, include:

- Oracle Advanced Compression's OLTP Table Compression feature often improves read performance because database blocks stay compressed when they're read into the buffer cache (reduces memory use in the buffer cache).
- The company noted that Redo generation could be up to ~30% more (based on the environment). This could result in more archive logs to be backed up, shipped to standbys and increase recovery time with Oracle Advanced Compression.

- Oracle Advanced Compression's OLTP Table Compression feature will have additional overhead for some DML operations. The overhead can vary based on the cardinality of the data, the CPU available and the Input/Output Operations per Second (IOPS) available.

Unstructured Data Testing (SecureFiles Deduplication and Compression)

Oracle Advanced Compression's SecureFiles feature introduced in Oracle Database 11g, offers a 'best-of-both-worlds' architecture for storing unstructured content, such as documents, spreadsheets and XML files. Oracle's SecureFiles feature is specifically engineered to deliver high performance for file data comparable to that of traditional file systems while retaining the advantages of Oracle Database.

Oracle Advanced Compression's SecureFiles feature has two storage optimization options that can be leveraged. SecureFiles Deduplication is an intelligent technology that eliminates duplicate copies of SecureFiles data. SecureFiles Compression utilizes compression algorithms to further minimize the storage requirements of SecureFiles data.

Test Scenario and Goals

No industry-specific test scenario was developed by the company for Oracle Advanced Compression Unstructured Data performance testing. Instead, the company utilized a small Java test (downloaded from Oracle Technology Network[2] (OTN)) for read/write comparison between Oracle Advanced Compression's SecureFiles feature for Deduplication and Compression versus BasicFiles.

- BasicFiles is the Oracle feature, provided prior to SecureFiles, that was used to store large blocks of unstructured data. The tests involved comparing BasicFiles (no compression) storage to SecureFiles (with Compression/Deduplication)
- Oracle's SecureFiles feature and LOB performance was measured using database driver protocol JDBC Thin driver (pure Java, Type IV driver)
- **Note** that Test One and Test Two (see Test Results and Conclusions chart below) used the same SQL operations executed against two different data sets – one dataset provided by Oracle OTN website and the other an internal company dataset

The company had no pre-set expectations prior to beginning the proof of concept in terms of overall performance impact as a result of compression/deduplication, but a primary goal of the testing was to determine what, if any, performance impact would result from the use of Advanced Compression and to compare the effectiveness of compression/deduplication with SecureFiles to similar data stored in BasicFiles (without compression).

Test Results and Conclusions

- Space utilization utilizing Oracle Advanced Compression' s SecureFiles Compression and Deduplication feature (vs. BasicFiles) was reduced by ~98%
- WRITE DML tests showed results of both faster and slower performance – additional testing would be needed to further define the specific performance characteristics of both
- READ operations show similar performance, or virtually no performance differences, between compressed and non-compressed tables.

[2] The Oracle Technology Network is the world's largest community of application developers, database admins, system admins/developers, and architects using industry-standard technologies in combination with Oracle products. (See here)

Source: Customer Tests

Note: All times in seconds

Test Conditions	Compression and Deduplication	Used Space (bytes)	WRITE Operation	READ Operation
Test One				
BasicFiles	NO	4423680	1.17	0.13
SecureFiles	YES	49152	0.69	0.14
Test Two				
BasicFiles	NO	1048576	1.47	0.32
SecureFiles	YES	131072	2.14	0.26

Key Takeaways

Overall, the customer reported that the majority of their performance testing showed no, or very little, impact on their overall OLTP performance when using Oracle Advanced Compression's Secure Files Compression and Deduplication feature. Additional analysis performed by the company showed that they were able to achieve significant levels of deduplication – which not only surprised the proof-of-concept test team but also provided significant, and unexpected, storage savings.

Advanced Compression with SecureFiles compression and deduplication is expected to enable the company to reduce their unstructured data (large objects) storage requirements by approximately 98% (50x reduction) with no significant impact on performance – although the company noted compression did have some overhead associated with some WRITE operations.

Backup Data (DataPump and RMAN Testing)

In Oracle Database 11g, the capabilities of Data Pump compression has been extended so that table data can be compressed on export. Oracle Advanced Compression's Data Pump compression feature is an inline operation, so the reduced dump file size results in a significant savings in disk space.

Unlike operating system or file system compression utilities, Oracle Advanced Compression's Data Pump compression feature is fully inline on the import side as well, so there is no need to uncompress a dump file before importing it. The compressed dump file sets are automatically decompressed during import without any additional steps by the Database Administrator.

Oracle Advanced Compression also includes RMAN compression technology that can dramatically reduce the storage requirements for backup data. Due to RMAN's tight integration with Oracle Database, backup data is compressed before it is written to disk or tape and doesn't need to be uncompressed before recovery – providing an enormous reduction in storage costs.

There are three levels of RMAN Compression: LOW, MEDIUM, and HIGH. The amount of storage savings increases from LOW to HIGH, while potentially consuming more CPU resources.

Test Scenario and Goals

No specific scenarios were tested as part of the backup compression testing. Instead, the proof-of-concept was intended to compare RMAN and Data Pump compression against the compression levels possible with 3rd party products like gzip.

- The test compared RMAN compression (included with RMAN) and RMAN with Oracle Advanced Compression, running backup and recovery on uncompressed and compressed data – for Oracle Advanced Compression the company tested using the MEDIUM level of RMAN compression.
- Prior to Oracle Database 11g, Oracle Advanced Compression's DataPump compression feature allowed compression of metadata; with Oracle Database 11*g*, the Advanced Compression's DataPump compression feature has been extended to the data in the dump file.

Test Results and Conclusions (DataPump)

Data Pump Test Tables	Compression= ALL	Total Space to Export	Export dmp size (GB)	Time (Min)	GZIP (GB)	Time to GZIP (Min)
Test One						
Uncompressed	NO	1.637	1.420	1.23	0.901	4.55
Uncompressed	YES	1.637	0.901	8.00		
Compressed	NO	1.518	1.420	1.22		
Compressed	YES	1.518	0.901	8.25		
Test Two						
Uncompressed	NO	4.301	3.808	1.40	0.831	5.20
Uncompressed	YES	4.301	0.672	9.40		
Compressed	NO	2.214	3.808	4.36		
Compressed	YES	2.214	0.672	14.54		

Source: Customer Tests

- Testing showed that Oracle Advanced Compression's Data Pump compression feature compresses data slightly better than OS compression (GZIP) ~5% better.
- The company noted that Oracle Advanced Compression's Data Pump compression feature compresses data as it does the export and does not need to have large space dedicated to uncompressed export files -- and that the import process understands compression in the export file.
- The testing also showed that Oracle Advanced Compression's Data Pump compression feature was slower than the OS (GZIP) compression (22-25%).
- **Note**: By design Oracle Advanced Compression's Data Pump compression feature uses SQL to extract data from the database and export it, so it doesn't benefit from compressed data within the database. If COMPRESSION=ALL is not listed as a parameter for export, the dump file will be created without compression.

Test Results and Conclusions (RMAN)

Tests were executed for RMAN Compression (included with RMAN) and RMAN with Advanced Compression (MEDIUM level) running backup and recovery on uncompressed and compressed tables. The company had no pre-set expectations regarding the performance of RMAN Compression versus RMAN with Advanced Compression; however, the company believed it was important to monitor this test for any significant CPU overhead associated with backup compression.

Test results for previously **uncompressed data**:

- Compression factor was 1.37x better using RMAN's compression instead of RMAN w/ Oracle Advanced Compression.
- Performance time was 2.22x better using RMAN with Oracle Advanced Compression instead of RMAN Compression.
- Database Restore/Recover Time was 2.63x better using RMAN with Oracle Advanced Compression instead of RMAN Compression.
- **Note:** Test One utilized RMAN Compression, Test Two utilized RMAN with Oracle Advanced Compression

RMAN Test Conditions	Channels	DB Size (GB)	Backupset Size (GB)	Compression Ratio	Time (Min)	DB GB/Hr
Test One						
RMAN Backup	8	105.9	18.04	5.9x	49.03	129.6
RMAN Restore	8	105.9	18.04	NA	43:42	
Test Two						
ACO Backup	8	105.9	24.6	4.3x	22:03	288.2
ACO Restore	8	105.9	24.6	NA	16:48	

Source: Customer Tests

Test results for previously **compressed data:**

- Compression factor was 1.50x better using RMAN Compression instead of RMAN with Oracle Advanced Compression.
- Performance Time was 2.26x better using RMAN with Oracle Advanced Compression instead of RMAN Compression.
- Database Restore/Recover Time was 1.98x better using RMAN with Oracle Advanced Compression instead of RMAN Compression.
- Note: Test One utilized RMAN Compression, Test Two utilized RMAN with Oracle Advanced Compression.

RMAN Test Conditions	Channels	Compressed DB Size (GB)	Backupset Size (GB)	Compression Ratio	Time (Min)	DB GB/Hr
Test One						
RMAN Backup	8	74.3	13.04	5.7x	33.46	132.2
RMAN Restore	8	74.3	13.04	NA	32.10	
Test Two						
ACO Backup	8	74.3	19.7	3.8x	15.50	282.6

ACO Restore	8	74.3	19.7	NA	16:21	

Source: Customer Tests

Key Takeaways

Overall, the customer reported that the majority of their performance testing showed that the use of Oracle Advanced Compression had no, or very little impact on their overall backup performance – instead showing that RMAN backups/recovery with Oracle Advanced Compression (MEDIUM level) was (on average) 1.4x faster for backups and 2.67x faster for recovery compared to RMAN backing up or recovering uncompressed datasets. (see below)

Specifically, their Proof-of-Concept testing demonstrated:

- With a compressed dataset the end result backupset size with RMAN and Oracle Advanced Compression was almost the same as the backupset size of RMAN compression on uncompressed dataset.
  - RMAN Compression (uncomp) – 18G / Oracle Advanced Compression(comp) – 19.7G
- With a compressed dataset the time to back up with RMAN and Oracle Advanced Compression was 1.42 times better than backing up the uncompressed dataset with RMAN Oracle Advanced Compression.
  - Oracle Advanced Compression(uncomp) – 22:03m / Oracle Advanced Compression(comp) – 15:50m (initially ~105G DB size versus ~74G to backup)
- With a compressed dataset the time to recover with RMAN and Oracle Advanced Compression was 2.67 times better than recovering uncompressed dataset with RMAN compression.
  - RMAN Compression – 43:42m / Oracle Advanced Compression – 16:21m

Suggested Best Practices

Oracle Advanced Compression can be implemented for existing database servers and new designs. Below are some best practices identified by the company in their Oracle Advanced Compression Proof-of-Concept:

- Each Oracle Advanced Compression feature requires thorough testing – the best test environment is where you can most closely duplicate the production environment with a test environment – but this is often difficult to do given budget, staffing and infrastructure limitations.
- Space usage reduction with Oracle Advanced Compression enabled gives the best results where the most duplicate data is stored (low cardinality). This is especially true for backups -- greater compression will result in less data backed up and hence shorter recovery time.
- Although CPU overhead is typically minimal, the company determined that implementing Oracle Advanced Compression was preferred primarily on systems with available CPU cycles -- as compression will have additional overhead for some DML operations.
- Tables can be redefined online, while queries continue against tables being converted to a compressed format.
- Compression Advisor is a PL/SQL package that is used to estimate potential storage savings (before implementing Oracle Advanced Compression) based on analysis of a sample of data, and provides a good estimate of the actual results that may be obtained after implementing Oracle Advanced Compression's OLTP Table compression feature. Compression Advisor, which supports Oracle Database 9*i* Release 2 through 11g Release 1, is available for free on the Oracle Technology Network website. Compression Advisor is built in to Oracle Database 11g Release 2.

Conclusion

Following the completion of their Proof-of-Concept, the company purchased the Advanced Compression option and began to quickly roll out their Oracle Advanced Compression deployments on Oracle Database 11.2.0.1.

The first application to utilize Oracle Advanced Compression was an OLTP deployment for archiving/content management. The company converted 40 database objects to utilize compression, with an average compression ratio of 2x (approximately a 50% reduction in space) -- which resulted in a reduction of the overall database size from approximately 8.1TB to 3.9TB. The company expects that the use of Oracle's Advanced Compression option will play a significant role in helping the company meet their storage savings initiative with no, or minimal performance impact.

Oracle Database 11g's Advanced Compression option provides a robust set of compression capabilities that enable IT managers to succeed in this complex environment. Using Oracle Advanced Compression, enterprises can efficiently manage their increasing data requirements throughout all components of their data center – minimizing costs while continuing to achieve the highest levels of application performance.

ORACLE

Oracle Corporation, World Headquarters
500 Oracle Parkway
Redwood Shores, CA 94065, USA

Worldwide Inquiries
Phone: +1.650.506.7000
Fax: +1.650.506.7200