ORACLE

Oracle Advanced Compression Frequently Asked Questions

January 2023, Version 21.2

Public

Purpose statement

This document provides an overview of features and enhancements included in release Oracle Database 21c. It is intended solely to help you assess the business benefits of upgrading to Oracle Database 21c and to plan your I.T. projects.

Disclaimer

This document in any form, software, or printed matter, contains proprietary information that is the exclusive property of Oracle. Your access to and use of this confidential material is subject to the terms and conditions of your Oracle software license and service agreement, which has been executed and with which you agree to comply. This document and information contained herein may not be disclosed, copied, reproduced, or distributed to anyone outside Oracle without prior written consent of Oracle. This document is not part of your license agreement nor can it be incorporated into any contractual agreement with Oracle or its subsidiaries or affiliates.

This document is for informational purposes only and is intended solely to assist you in planning for the implementation and upgrade of the product features described. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. The development, release, and timing of any features or functionality described in this document remains at the sole discretion of Oracle. Due to the nature of the product architecture, it may not be possible to safely include all features described in this document without risking significant destabilization of the code.

Table of contents

What are the features of Advanced Compression

Advanced Compression includes a comprehensive set of compression features designed to maximize resource utilization and reduce costs by enabling compression for structured data (Advanced Row Compression), Indexes (Advanced Index Compression), unstructured data (Advanced LOB Deduplication and Compression), database backups (RMAN and Data Pump compression) and network compression (Data Guard redo log transport and Advanced Network Compression).

Advanced Network Compression compresses data transferred between Oracle Database and client applications. It is transparent to client applications and can improve SQL response times while saving network bandwidth. Advanced Compression also includes an optimization for Flashback Data Archive (FDA) history tables, which reduces the storage requirements for FDA History tables when using FDA to track transactional changes.

Several features of Advanced Compression enhance the data lifecycle management capabilities of Oracle Database. Heat Map automatically tracks data modification times at the row and segment level, and access times at the segment level, providing unprecedented insights into how data is being accessed. Automatic Data Optimization (ADO) provides declarative syntax to automatically move and compress data based on usage statistics collected by Heat Map.

What are the benefits of Advanced Row Compression

Advanced Row Compression uses a unique compression algorithm specifically designed to work with OLTP (and data warehouse) applications. The algorithm works by eliminating duplicate values within a database block, even across multiple columns.

Compressed blocks contain a structure called a symbol table that maintains compression metadata. When a block is compressed, duplicate values are eliminated by first adding a single copy of the duplicate value to the symbol table. Each duplicate value is then replaced by a short reference to the appropriate entry in the symbol table.

Through this innovative design, compressed data is self-contained within the database block, as the metadata used to translate compressed data into its original state is stored in the block header. When compared with competing compression algorithms that maintain a global database symbol table, Oracle's approach offers significant performance benefits by not introducing additional I/O when accessing compressed data.

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

However, the benefits of Advanced Row Compression go beyond just on-disk storage savings. One significant advantage is Oracle's ability to read compressed blocks directly without uncompressing the blocks. This helps improve performance due to the reduction in I/O, and the reduction in system calls related to the I/O operations. Further, the buffer cache becomes more efficient by storing more data without having to add memory.

Why not just use storage-based compression and deduplication

An alternative to initiating compression, at the database level, is to use deduplication and compression at the storage level to address capacity utilization concerns. With a storage-based solution, duplicate blocks are eliminated from incoming data streams, then many storage systems can analyze what remaining data can be compressed as a secondary function. The combination of deduplication and compression can be used to minimize the amount of data stored on the storage system.

With deduplication and compression enabled at the storage level, the two operations are executed with no knowledge of the use or value of incoming data. Without true application awareness, the deduplication and compression functions are not applied strategically to the incoming data but are reducing data by any means necessary to minimize use of limited storage resources. Without awareness of the data value and use, duplicate data is deduplicated, and any compressible data is compressed, without regard for rehydration or decompression impacts on data retrieval. If there is duplicate, or compressible data, that should remain in duplicate or uncompressed form for speed of access, the deduplication and compression functions must be turned off. This is also true for compression-only appliances that may be used to minimize bandwidth usage between remote sites.

Because an effective storage deduplication/compression strategy requires fully hydrated and uncompressed data to be sent to the storage device, the database either would have to be not compressed or would have to be uncompressed before going to storage. If the database is not compressed, then all the advantages of compression (reduced storage, reduced I/O, reduced memory use in the buffer cache etc.) will not be available on the database servers. If the database must be uncompressed before storage, then additional CPU and I/O overhead will be required to perform the decompression.

An additional consideration when evaluating storage-based solutions is to consider the impact on network traffic and available bandwidth. Where Oracle Advanced Compression is driven by the database and can reduce the amount of network traffic between server and storage, an effective storage deduplication or compression strategy requires fully hydrated and uncompressed data to be sent to the storage device.

Why use database compression if disk is getting cheaper

Enterprises are experiencing an explosion in the volume of data required to effectively run their businesses. This trend in data growth can be attributed to several key factors. Recent changes in the regulatory landscape, such as

Sarbanes-Oxley and HIPAA, are contributing to this trend by mandating that enterprises retain large amounts of information for long periods.

Mass distribution of rich and multimedia content over the Internet, made possible through advancements in broadband technologies, also contributes to the growth in overall data volume. Various estimates indicate that data volume is almost doubling every 2-3 years.

This sudden explosion in data volume presents a daunting management challenge for IT administrators. First and foremost are the spiraling storage costs even though the cost of storage has been declining dramatically in the last few years, the enormous growth in the volume of data that needs to be retained online makes storage one of the biggest cost elements of most IT budgets. In addition, application scalability and performance must continue to meet the demands of the business – even as data volumes explode.

Advanced Compression helps organizations cope with these challenges. Innovations in Oracle compression technologies help organizations reduce the resources and costs of managing large data volumes. Another important benefit is in database performance. A major bottleneck for many systems is I/O bandwidth. Advanced Compression can help alleviate that bottleneck in several cases by reducing the amount of data that needs to be transferred across I/O channel and further boost performance through improved memory efficiencies.

What types of organizations use Advanced Compression

Financial, government, education, healthcare, utilities, insurance, retail, manufacturing, and more use oracle compression….

How do I enable compression on existing tables

There are three recommended approaches to enabling Advanced Row Compression on existing tables:

Online Redefinition (DBMS_REDEFINITION)

This approach will enable Advanced Row Compression for future DML and compress existing data. Using DBMS_REDEFINITION keeps the table online for both read/write activity during the migration. Run DBMS_REDEFINITION in parallel for best performance.

Online redefinition will clone the indexes to the interim table during the operation. All the cloned indexes are incrementally maintained during the sync (refresh) operation so there is no interruption in the use of the indexes during, or after, the online redefinition. The only exception is when online redefinition is used for redefining a partition -- any global indexes are invalidated and need to be rebuilt after the online redefinition.

ALTER TABLE … MOVE

This approach will enable Advanced Row Compression for future DML and will compress existing data. While the table is being moved, it is online for read activity but has an exclusive (X) lock – so all DML will be blocked until the move command completes. Run ALTER TABLE…MOVE in parallel for best performance.

ALTER TABLE… MOVE will invalidate any indexes on the partition or table; those indexes will need to be rebuilt after the ALTER TABLE… MOVE. For partition moves, the use of ALTER TABLE… MOVE PARTITION with the UPDATE INDEXES clause will maintain indexes (it places an exclusive (X) lock so all DML will be blocked until the move command completes) – not available for non-partitioned tables.

The ALTER TABLE… MOVE statement allows you to relocate data of a non-partitioned table, or of a partition of a partitioned table, into a new segment, and optionally into a different tablespace. ALTER TABLE…MOVE compresses the data by creating new extents for the compressed data in the tablespace being moved to -- it is important to note that the positioning of the new segment can be anywhere within the data file, not necessarily at the tail of the file or head of the file. When the original segment is released, depending on the location of the extents, it may or may not be possible to shrink the data file.

ALTER TABLE … MOVE TABLE/PARTITION/SUBPARTITION … ONLINE

This approach will enable Advanced Row Compression for future DML and will compress existing data. ALTER TABLE … MOVE TABLE/PARTITION/SUBPARTITION … ONLINE allows DML operations to continue to run uninterrupted on the table/partition/subpartition being moved. Indexes are maintained during the move operation, so a manual index rebuild is not required.

For more detailed information, please see the Oracle Database documentation.

What compression ratio can I expect

The compression ratio achieved in each environment depends on the nature of the data being compressed; specifically, the cardinality of the data.

In general, organizations can expect to reduce their storage space consumption by a factor of 2x to 4x when using Advanced Row Compression. A 2x compression ratio represents approximately a 50% reduction in the storage footprint.

Can partitioned tables be compressed

Yes. Compression is enabled at a tablespace, table, or partition level.

Can compression be set as the default for a tablespace

Yes. Compression can be enabled as the default when any new table is created or a table is rebuilt.

Do all tables need to be compressed

No – You can pick-and-choose which tables/partitions you want to compress and you do not have to compress all tables/partitions.

Should compression be enabled by default at the tablespace level

For custom applications, we recommend compressing at the tablespace level, but users should consider turning off compression on very high traffic tables, such as tables used as queues.

For commercial packaged applications, where typically the number of objects can be very large, the recommended approach is object selection instead of exclusion. Often the largest tables and indexes consume much of the database space. Compressing those objects, while excluding high traffic objects like tables used as queues, will give most of the compression benefits. Other objects can be compressed over time as needed.

Can the compression ratio, for my tables, be estimated before enabling compression

Yes. Compression Advisor is a free PL/SQL package (included with DBEE) used to estimate potential compression ratios, for Advanced Row Compression, based on analysis of a sample of data. It provides a good estimate of the actual compression ratio that may be obtained after implementing Advanced Row Compression.

A version of Compression Advisor, which supports Oracle Database 9*i* Release 2 through 11g Release 1, is available (free) on the Advanced Compression page on the Oracle.com website. Compression Advisor (DBMS_COMPRESSION) is included free with Oracle Database 11g Release 2 and above.

Can SecureFiles LOB segments be compressed

Yes. The Advanced LOB Compression and Advanced LOB Deduplication features of Advanced Compression are intended to reduce the amount of storage required for SecureFiles LOBs.

Can Index-Organized Tables (IOTs) be compressed

Yes. Index-Organized Tables (IOT's) are essentially indexes, so they cannot be compressed with Advanced Row or Basic Compression. However, IOT's can be compressed with Prefix Key Compression (included with DBEE).

Should all tables be compressed

The general recommendation is to compress all the tables in the database with one exception: if the table is used as a queue, i.e., rows are inserted into the table, then later most or all the rows are deleted, then more rows are inserted then deleted, then you should not compress the table.

Should test data and test apps be used to test Advanced Compression

No. The best test environment for each Advanced Compression capability is where you can most closely duplicate the production environment (using your actual data and applications) – this will provide the most realistic (pre- and post-compression) performance and functionality comparisons.

What is the CPU overhead associated with Advanced Row Compression

Although CPU overhead is minimal (3% to 5% typically), implementing Advanced Row Compression is ideal on systems with available CPU cycles, as compression will have additional, although minor, overhead for some DML operations.

What data types are not supported by Advanced Row Compression

Advanced Row Compression is NOT supported for use with tables that have LONG data types.

Will a larger block size improve the compression ratio

It is possible a larger block size will have a better compression ratio if the larger block has more duplicate data on the block. However, larger blocks do not always ensure higher Advanced Row Compression compression ratios.

Test with your own data to determine if larger block sizes will have an impact on your Advanced Row Compression compression ratio.

Doesn't Basic Table Compression have the same benefits as Advanced Row Compression

Oracle Database 9i introduced Basic Table Compression, which only compressed data that was loaded using bulk load operations. Advanced Row Compression, a feature of Advanced Compression, allows data to be compressed during bulk loads, as well as during all types of data manipulation operations, including conventional DML such as INSERT and UPDATE.

In addition, Advanced Row Compression reduces the associated compression overhead of write operations making it suitable for transactional/OLTP environments. Advanced Row Compression, therefore, extends the benefits of compression to all application workloads.

How is Advanced Compression RMAN Backup Compression different from RMAN Basic Compression

Advanced Compression provides 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.

Compression Level LOW provides the fastest compression algorithm and is best suited when backup is constrained by CPU.

Compression Level MEDIUM provides a balance between CPU usage and compression ratio and finally,

Compression LEVEL HIGH provides the best compression ratio and highest CPU utilization and is best suited when backup is constrained by network or I/O.

What are Advanced LOB Compression and Advanced LOB Deduplication

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

Advanced Compression has two storage optimization features that can be leveraged with SecureFiles. The first feature, Advanced LOB Deduplication, is an intelligent technology that eliminates duplicate copies of SecureFiles data. The second feature, Advanced LOB Compression, utilizes industry standard compression algorithms to further minimize the storage requirements of SecureFiles data.

There are three levels of Advanced LOB compression available: LOW, MEDIUM, and HIGH. By default, Advanced LOB Compression uses the MEDIUM level, which typically provides good compression with a modest CPU overhead of 3%-5%. SecureFiles Compression LOW is optimized for high performance and maintains about 80% of the compression achieved through MEDIUM, while utilizing 3x less CPU. Finally, Advanced LOB Compression HIGH achieves the highest storage savings but incurs the most CPU overhead.

How does Advanced Row Compression compress data

Advanced Row Compression uses a unique compression algorithm specifically designed to work with OLTP/DW applications. The algorithm works by eliminating duplicate values within a database block, even across multiple columns.

Compressed blocks contain a structure called a symbol table that maintains compression metadata. When a block is compressed, duplicate values are eliminated by first adding a single copy of the duplicate value to the symbol table. Each duplicate value is then replaced by a short reference to the appropriate entry in the symbol table. Through this innovative design, compressed data is self-contained within the database block as the metadata used to translate compressed data into its original state is stored in the block.

When compared with competing compression algorithms that maintain a global database symbol table, Oracle's unique approach offers significant performance benefits by not introducing additional I/O when accessing compressed data.

What optimizations has Oracle done to reduce compression overhead

Advanced Row Compression has no adverse impact on read operations. There is additional work performed while writing data, making it impossible to eliminate performance overhead for write operations. However, Oracle has put in a significant amount of work to minimize this overhead for Advanced Row Compression.

Oracle compresses blocks 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. 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 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, most transactions on compressed blocks will have the exact same performance as they would with uncompressed blocks.

Do data, and indexes, stay compressed in memory

Yes. Oracle Database can read compressed blocks directly without having to first uncompress the block. Therefore, there is no measurable performance degradation for accessing compressed data.

In fact, in many cases performance may improve due to the reduction in I/O since Oracle will have to access fewer blocks. Further, the buffer cache will become more efficient by storing more data without having to add memory.

What is the performance impact on DML writes

For DML operations on a compressed table, Advanced Row Compression's specialized batch algorithm keeps the performance overhead to a minimum. Internal tests at Oracle showed a minimal overhead of less than 5% (CPU) for a DML workload.

It is important to note that Oracle compresses blocks in batch mode rather than compressing data every time a write operation takes place. When a transaction causes the data in the block to reach an internal threshold, all contents of the block are compressed. 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 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, most transactions on compressed blocks will have the exact same performance as they would with uncompressed blocks.

Do chained rows get compressed

Yes. Before Oracle Database 12*c*, blocks containing many types of chained rows could not be compressed. This limitation was removed starting with Oracle Database 12*c* Release 2.

Does Advanced Compression include data lifecycle management capabilities

Yes. Data lifecycle management is the practice of applying policies for the effective management of information throughout its useful life. Data lifecycle management includes every phase of a "row" from its beginning to its end, and consists of the policies, processes, practices, and tools used to align the business

value of information with the most appropriate and cost-effective IT infrastructure from the time information is created through its final disposition.

The Advanced Compression features, Heat Map and Automatic Data Optimization (ADO) are used to create data lifecycle management policies to automate compression tiering and storage tiering as part of a data lifecycle management solution.

Depending on your data lifecycle management requirements, the use of Oracle Partitioning and Oracle Hybrid Columnar Compression may also be a best practice.

More information about ADO and data lifecycle management is available on the Advanced Compression page on the Oracle.com website.

How to get started with the free Compression Advisor

An easy way to get started, with Advanced Row Compression, is by using compression advisor. The "DBMS_COMPRESSION" PL/SQL package (commonly called compression advisor) gathers compression-related information within a database environment. This includes estimating the compressibility of both uncompressed partitioned, and non-partitioned tables, and gathering row-level compression information on previously compressed tables/partitions. Compression advisor provides organizations with the storage reduction information needed to make compression-related usage decisions.

The output of running compression advisor is an estimation of the compression ratio for the specific table or partition that was the target of compression advisor. The output indicates the "COMPRESSION RATIO" presented as a number such as 2.1. This number indicates that, for this specific table or partition, the estimated compression ratio is 2.1x, which represents about a 50% reduction in the footprint of the table or partition should compression be enabled.

A version of Compression Advisor, which supports Oracle Database 9*i* Release 2 through 11g Release 1, is available free on the Advanced Compression page on Oracle.com. For later releases, a version of Compression Advisor is included with Oracle Database Enterprise Edition 11g Release 2 and above.

Connect with us

Call +**1.800.ORACLE1** or visit **oracle.com**. Outside North America, find your local office at: **oracle.com/contact**.

blogs.oracle.com facebook.com/oracle twitter.com/oracle