

Oracle Advanced Compressionの概念実証に 関するガイドライン、インサイト、 ベスト・プラクティス

2023年3月、バージョン21.2

Copyright © 2023, Oracle and/or its affiliates 公開

本書の目的

本書では、Oracle Database Release 21cの機能の概要と強化された点が説明されています。本書は、御社がOracle Database 21cへのアップグレードのビジネス上の利点を評価し、ITプロジェクトを計画することを支援することのみを目的としています。

免責事項

本文書には、ソフトウェアや印刷物など、いかなる形式のものも含め、オラクルの独占的な所有物である占有情報が含まれます。この機密文書へのアクセスと使用は、締結および遵守に同意したOracle Software License and Service Agreementの諸条件に従うものとします。本文書と本文書に含まれる情報は、オラクルの事前の書面による同意なしに、公開、複製、再作成、またはオラクルの外部に配布することはできません。本文書は、ライセンス契約の一部ではありません。また、オラクル、オラクルの子会社または関連会社との契約に組み込むことはできません。

本書は情報提供のみを目的としており、記載した製品機能の実装およびアップグレードの計画を支援することのみを意図しています。マテリアルやコード、機能の提供をコミットメント（確約）するものではなく、購買を決定する際の判断材料になさらないでください。本書に記載されている機能の開発、リリース、および時期については、弊社の裁量により決定されます。製品アーキテクチャの性質上、コードが大幅に不安定化するリスクなしに、本書に記載されているすべての機能を安全に含めることができない場合があります。

目次

本書の目的	2
免責事項	2
はじめに	4
高度な行圧縮	4
高度な行圧縮の有効化	5
高度な行圧縮が実行される場合	6
ダイレクト・パスと従来型パスによるバルク・ロード	7
高度な索引圧縮	8
Oracle RMANバックアップ圧縮	9
高度なLOB圧縮 (SecureFiles LOBの場合)	10
テスト開始前の考慮事項	10
圧縮オーバーヘッドについて	11
圧縮比の改善	11
概念実証の一般的な機能	12
無料の圧縮アドバイザ	13
付録A	13

はじめに

企業は、データ量の急増という重大な問題に直面しています。企業は、収益に影響を与えることなく変化の激しいビジネス状況に迅速に適応する必要があります。ITマネージャーは、既存のインフラストラクチャを効率的に管理してコストを制御しながら、優れたアプリケーション・パフォーマンスを提供し続けていく必要があります。

Oracle Advanced CompressionとOracle Databaseはともに、この複雑な環境で、ITマネージャーを成功に導く、一連の堅牢な圧縮、パフォーマンスおよびデータ・ストレージ最適化機能を提供します。

Oracleデータベースのデプロイメントがクラウドまたはオンプレミスのどちらであっても、Oracle Advanced Compressionにより、アプリケーションに変更を加えることなく、異なる環境において、堅牢な圧縮機能を実現できます。Oracle Advanced Compressionのメリットには、データベース・ストレージ・フットプリントの削減、バックアップでの時間とストレージの節約、システム・パフォーマンスの向上などがあります。

評価されるおもな機能

高度な行圧縮

あらゆるタイプのデータ操作中に表データの圧縮を可能にします。

高度な索引圧縮

サポートされるすべての一意索引と非一意索引のサイズを縮小します。

Oracle RMANバックアップ圧縮

Oracle Recovery Manager (Oracle RMAN) の使用時にバックアップ・データを圧縮します。

高度なLOB圧縮

SecureFiles LOBの圧縮機能です。

高度な行圧縮

Oracle Database 11g Release 1でOLTP表圧縮（現在の高度な行圧縮）が導入されました。この機能を使用することで、INSERTやUPDATEなどの従来のDMLを含むあらゆるタイプのデータ操作中に、データを圧縮できます。さらに、高度な行圧縮では、圧縮データに対する書き込み操作のオーバーヘッドが最小限に抑えられます。そのため、トランザクション環境やOLTP環境に加え、データウェアハウスにも適しており、圧縮の利点がすべてのアプリケーションのワークロードに拡張されます。

高度な行圧縮では、OLTP/DWアプリケーションで動作するよう特別に設計された、独自の圧縮アルゴリズムが使用されます。このアルゴリズムは、データベース・ブロック内や複数の列間の重複値を排除することによって動作します。圧縮されたブロックには、圧縮メタデータを維持する記号表と呼ばれる構造体が含まれます。ブロックが圧縮されると、最初に重複値のコピーが記号表に1つ追加されることにより、重複値が排除されます。そして、各重複値が、記号表内の適切なエントリへの短い参照に置き換えられます。

この革新的な設計では、圧縮されたデータを元の状態へ変換するために使用されるメタデータがブロック・ヘッダー内に保存されるため、圧縮されたデータはデータベース・ブロック内で自己完結します。グローバルなデータベースの記号表を維持する競合他社の圧縮アルゴリズムと比較すると、圧縮されたデータにアクセスする際、（グローバルな記号表が必要な）追加のI/Oが発生しないオラクル独自のアプローチでは、大幅なパフォーマンス上の利点が得られます。

このドキュメントについて

このドキュメントは、圧縮の概念実証を実行するためのステップ・バイ・ステップ・ガイドではありません。顧客のテストから学んだ圧縮のガイドラインおよびベスト・プラクティス、そして独自の圧縮の概念実証を計画し、その結果を理解するのに役立つ洞察を提供します。

ストレージの節約にとどまらない メリット

高度な行圧縮の利点は、ディスク上のストレージを節約できるだけではなく、問合せパフォーマンスのおもな利点の1つは、ブロックを解凍することなく、圧縮されたブロック（データと索引）を直接メモリ内で読み取れることです。そのため、I/Oが削減され、I/O操作に関連するシステム・コール数が削減されるため、パフォーマンスが向上します。さらに、メモリの追加を必要とせずにより多くのデータを保存することにより、バッファ・キャッシュの効率が向上します。

圧縮率は、圧縮されるデータ、特にデータのカーディナリティによって環境ごとに異なります。通常は、高度な行圧縮を使用することで、ストレージ領域の消費を1/2~1/4に削減できることが期待できます。つまり、非圧縮データ量が消費する領域量は、圧縮されたデータ量が消費する領域量の2~4倍になります。

高度な行圧縮の有効化

新しい表やパーティションに対して、高度な行圧縮を有効化する方法は簡単です。単純に表またはパーティションをCREATE構文で作成し、“ROW STORE COMPRESS ADVANCED”と指定します。たとえば、次のような文を使用します。

```
CREATE TABLE emp (emp_id NUMBER, first_name  
VARCHAR2(128), last_name VARCHAR2(128)) ROW STORE  
COMPRESS ADVANCED;
```

既存の表に高度な行圧縮を有効化する方法は多数あります。それらの方法を完全に説明するのは本ドキュメントの範囲外であるため、ここでは顧客が通常使用する方法の概要を示します。

それらの方法の詳細については、Oracle Databaseのドキュメントを参照してください。

ALTER TABLE ... ROW STORE COMPRESS ADVANCED

今後のすべてのDMLについて高度な行圧縮を有効化しますが、表の既存のデータについては非圧縮のままにします。

オンライン再定義 (DBMS_REDEFINITION)

この方法では、その後のDMLでの高度な行圧縮が有効化され、既存のデータが圧縮されます。DBMS_REDEFINITIONを使用すると、移行中に、読取り/書込みの両方のアクティビティに対して、表がオンライン状態で維持されます。最適なパフォーマンスを得るには、DBMS_REDEFINITIONをパラレルで実行します。

オンライン再定義では、操作の実行中に索引が仮表にクローンされます。クローンされたすべての索引の増分が同期（リフレッシュ）操作中に維持されるため、オンライン再定義の実行中も実行後も、索引の利用が中断されることはありません。

ただし、オンライン再定義をパーティションの再定義に使用する場合に限り、索引の利用は中断されます。この場合、グローバル索引がすべて無効化され、オンライン再定義の実行後にそのグローバル索引を再作成する必要があります。

ALTER TABLE ... MOVE ROW STORE COMPRESS ADVANCED

この方法では、その後のDMLでの高度な行圧縮が有効化され、既存のデータが圧縮されます。表の移行中、読取りアクティビティに対しては表がオンライン状態で維持されますが、排他（X）ロックがかかるため、移行コマンドが完了するまですべてのDMLがブロックされます。最適なパフォーマンスを得るには、ALTER TABLE ... MOVEをパラレルで実行します。

重要なドキュメント・ノート

これらの操作に関する詳細、使用例、および制限事項については、現行のOracle Databaseドキュメントを参照してください

ALTER TABLE ... MOVEにより、パーティションまたは表にある索引がすべて無効化されます。ALTER TABLE ... MOVEの実行後、これらの索引を再作成する必要があります。パーティションの移行の場合、ALTER TABLE ... MOVE PARTITIONをUPDATE INDEXES句とともに使用すると、索引が維持されます（排他（X）ロックがかかるため、移行コマンドが完了するまで、すべてのDMLがブロックされます）。パーティション化されていない表に対しては、この句は使用できません。

ALTER TABLE...MOVE文を使用すると、パーティション化されていない表のデータやパーティション化された表のパーティションのデータを新しいセグメントに再配置したり、オプションとして異なる表領域に再配置したりできます。

ALTER TABLE...MOVE ROW STORE COMPRESS ADVANCEDにより、圧縮データ用の新しいエクステントが移行先の表領域内に作成され、データが圧縮されます。ここで、新しいセグメントは、データファイルの末尾や先頭に配置されるとは限らず、あらゆる場所に配置される可能性がある点に注意が必要です。そのため、元のセグメントが解放されるとき、エクステントの位置によっては、データファイルが縮小されない場合もあります。

ALTER TABLE ... MOVE TABLE/PARTITION/SUBPARTITION ... ONLINE

この方法では、その後のDMLでの高度な行圧縮が有効化され、既存のデータが圧縮されます。

ALTER TABLE ...MOVE TABLE, PARTITION or SUBPARTITION ... ONLINEにより、移動中の表、パーティション、サブパーティションに対してDML操作を中断せずに実行し続けることができます。

索引は移行操作中に維持されるため、手動による索引の再作成は不要です。

高度な行圧縮が実行される場合

高度な行圧縮では、OLTPおよびデータウェアハウス・アプリケーションで動作するために特別に設計された独自の圧縮アルゴリズムを使用します。このアルゴリズムは、データベース・ブロック内や複数の列間の重複値を排除することによって動作します。圧縮されたブロックには、圧縮メタデータを維持する記号表と呼ばれる構造体が含まれます。

ブロックが圧縮されると、最初に重複値のコピーが記号表に1つ追加されることにより、重複値が排除されます。そして、各重複値が、記号表内の適切なエントリへの短い参照に置き換えられます。

圧縮のメリットや使用される圧縮技術は似ていますが、高度な行圧縮を使用すると、データベース内で異なる圧縮タイプを起動できます。以下に、異なる圧縮タイプの例を示します。これらの異なる圧縮タイプがどのような場合に使用されるかを理解しておく、概念実証の結果を分析する際に役立ちます。

インサート・ダイレクト・ロード圧縮

APPENDヒントやSQL*Loaderを使用した挿入など、ダイレクト・パス・ロード・メカニズムを使用してデータが挿入される場合に実行されます。この場合、データはセグメントの最高水位標（セグメントの最後に使用された仮想ブロック・マーカー）に挿入されるため、データ・ブロックに極めて効率的に書き出すことができます。圧縮エンジンでは大量の行が処理され、行をバッファして圧縮し、圧縮した行をデータ・ブロックに書き出すことができます。結果として、領域をすぐに節約できます。

インサート・ダイレクト・ロード圧縮では、行が非圧縮形式で書き込まれることはありません。

再帰的圧縮

単一行または配列の挿入や更新など、従来のDML操作で起動されます。この圧縮タイプでは、行は非圧縮形式で書き出され、データ・ブロックが内部ブロックの上限しきい値に達すると圧縮が起動されます。このような使用例では、データ・ブロックを再帰的トランザクションで圧縮でき、圧縮後は直ちにコミットされます。

圧縮によって節約された領域は直ちに解放され、他のトランザクションで使用できるようになります。圧縮はユーザーのDML操作（ユーザー・トランザクション）によってトリガーされますが、実際のデータ圧縮は再帰的トランザクションで行われるため、圧縮の結果はユーザーのトランザクションの結果に関係しません。

ダイレクト・パスと従来型パスによるバルク・ロード

バルク・ロード操作の実行時に、ダイレクト・パスまたは従来型パスのどちらかの方法を選択することで、ロード・パフォーマンスに重大な影響が及ぶ可能性があります。

バルク・ロード挿入操作の実行時、特に従来型パス・ロードを使用して大量の行を挿入する場合に、挿入パフォーマンスが低下する可能性があります。

大量の行を挿入する場合に従来型パス・ロードを使用するとパフォーマンスが低下します。その理由は、新しい行が既存の圧縮されたブロックに挿入されると、挿入は圧縮されずに実行され、その後同じブロックで追加の挿入が実行されるとブロックは埋まり始め、内部のしきい値に達するとブロックが圧縮されるためです（詳細は上述の再帰的圧縮を参照）。

圧縮後に別の領域が解放されると、ブロックで再び挿入が実行されます。そのため、同じ従来型パス・ロード操作の間に、同じブロックで複数回の圧縮が行われる可能性があります。

つまり、従来型パスによる挿入を使用すると、同じ操作の実行中に同じブロックが複数回圧縮され、CPUリソースや時間を消費する可能性があります。ワークロードがおもに従来型パスの挿入である場合、高度な行圧縮アルゴリズムの一部としてブロックが繰り返し再圧縮されると、I/Oが増加する可能性があります（ダイレクト・パス・ロードとの比較）。

大量の行で操作を実行する場合は、ダイレクト・パス・ロード操作が推奨されます。従来型パス・ロードとは異なり、ダイレクト・パス・ロードは最高水位標より上で実行されるため、ブロックは一度だけいっぱいになって圧縮され、ディスクに書き込まれるためです。これにより、バルク挿入が効率化され、従来型パス・ロードを使用してバルク挿入を実行する際に起こり得る、同じブロックでの複数の圧縮が回避されます。

AWRとダイレクト・パス/従来型パスのバルク・ロード

概念実証テストの実行中、ダイレクト・パス・ロードと従来型パス・ロードのどちらの方法がバルク・ロードで使用されるかに確信が持てない場合、(AWRで)以下の推奨される手順を使用することで、SQL操作中に生じる圧縮の量を特定できます。

従来型パス・ロード圧縮の特定

AWRには“Instance Activity Stats”セクションがあり、そこに正圧縮 (HSC OLTP positive compression) の総数と負圧縮 (HSC OLTP negative compression) の総数に関連する統計が表示されます。これら2つの統計を合計すると、試行された圧縮 (再圧縮またはそれ以外) の総数を把握できます。

$$\text{HSC OLTP positive compression} + \text{HSC OLTP negative compression} = \text{試行された圧縮と再圧縮の総数 (従来型パス・ロード)}$$

ダイレクト・パス・ロード圧縮の特定

“insert append”などのダイレクト・パス手法を使用してバルク・ロードを実行すると、データはデータ・ブロックに編成され、メモリに圧縮されます。つまり、バルク・ロード・データは一度だけ圧縮されます。データ・ブロックは、表のPCTFREE設定で指定されているポイントまで満たされます。Oracle DatabaseでのPCTFREEのデフォルト設定は10%です (PCTFREEにより、SQL UPDATE操作中の増加に備えてデータ・ブロックの領域を予約できます)。

Create Table as Select (CTAS) やinsert appendの場合など、最高水位標 (HWM) より上のブロック圧縮では、HSC IDL Compressed Blocksという統計があります。

$$\text{HSC IDL Compressed Blocks} = \text{HWMより上のブロック圧縮 (CTASやinsert appendなどのダイレクト・パス・ロード)}$$

HSC OLTP positive CompressionとHSC OLTP negative compressionの統計値のみが表示され、HSC IDL Compressed Blocksの統計値がまったく表示されないかほとんど表示されない場合は、実行されているすべての圧縮は従来型パス操作によるものです (特に、1秒あたりに実行されている圧縮の数を確認します)。

可能な場合は、同じ操作で従来型パス・ロードの代わりにダイレクト・パス・ロードが実行されるように、バルク挿入を変更することを検討する必要があります。そうした場合、HSC IDL Compressed Blocksの統計値が大きくなります。HSC IDL Compressed Blocksでラベル付けされた統計がない場合、HWMより上のブロック圧縮がなかったことを意味します。

高度な索引圧縮

Oracle DatabaseのOracle Advanced Compressionの一部である高度な索引圧縮を使用すると、索引圧縮を自動化できるため、DBAは圧縮で考慮するプリフィックス列の数を指定する必要がなくなります (索引キー圧縮では必要)。

高度な索引圧縮は、複数の圧縮レベル（低および高）を利用できるテクノロジーです。圧縮率は、実行する圧縮のレベルによって異なりますが、平均で2~5倍です。高度な索引圧縮による相当なストレージ節約によって、ITマネージャーはストレージの新規購入量を大幅に削減できます。

高度な索引圧縮の有効化

高度な索引圧縮は、CREATE/ALTER INDEX句のCOMPRESS ADVANCED副次句を指定することで有効化できます。圧縮時に新しい索引を自動的に作成できます。または、既存の索引を圧縮状態に再構築することもできます。

CREATE INDEX idxname ON tablename(col1, col2, col3) COMPRESS ADVANCED LOW/HIGH;

高度な索引圧縮では、プリフィックス・エントリの列数を指定する必要がないことに注意してください。リーフ・ブロックごとに自動的に計算されるためです。

高度な索引圧縮は、サポートされるすべての索引（プリフィックス・キー圧縮に適さない索引を含む）で問題なく使用できます。高度な索引圧縮を使用して索引を作成すると、すべての一意（または非一意）索引のサイズを減らすことができます（少なくとも負圧縮によるサイズ増大を防ぐことができます）。同時に、索引に効率的にアクセスしながら圧縮率を大幅に改善することができます。

高度な索引圧縮には以下の制限があります。

- ビットマップ索引では、高度な索引圧縮がサポートされません。
- 高度な索引圧縮を使用して索引構成表（IOT）を圧縮することはできません（索引キー圧縮を使用）。
- 高度な索引圧縮を使用してファンクション索引を圧縮することはできません。

高度な索引圧縮によって、Bツリー索引を圧縮できるようになりました。このとき、圧縮でメリットを得られるすべての索引リーフ・ブロックが自動的に圧縮され、ブロックごとに最適なプリフィックス列長が自動的に計算されます。そのため、索引圧縮がブロック・レベルの真にローカルなものになり、索引セグメント全体にとって最適な圧縮率の実現を目指して、圧縮プリフィックス表の作成とリーフ・ブロックの圧縮方法の決定がブロックごとにローカルで行われます。

Oracle RMANバックアップ圧縮

Oracle RMANがOracle Databaseと密接に統合されていることにより、圧縮済みのデータ/索引ブロックはOracle RMANバックアップ中も圧縮された状態が維持され、リカバリ前に解凍する必要がありません。そのため、ストレージ・コストが削減され、バックアップとリストアにかかる時間も大幅に短縮される可能性があります。

バックアップの圧縮に関しては、Oracle RMAN基本圧縮では非常に高い圧縮率が実現しますが、CPUに負荷がかかる場合があり、CPUの可用性がバックアップとリストアのパフォーマンスを制限する要因となる可能性があります。

Oracle Advanced Compressionに含まれるOracle RMANによるバックアップ圧縮には、LOW、MEDIUM、HIGHの3つのレベルがあります。ストレージが節約される量は、LOWからHIGHに向かって増加しますが、CPUリソースの消費量が増加する可能性があります。LOW / MEDIUM / HIGHの圧縮は、通常はOracle RMAN基本圧縮よりもCPUの使用量を減らしながら、さまざまなレベルの圧縮を提供するように設計されています。

I/Oに制約があるがCPUがアイドル状態の場合は、HIGHが最適な場合があります。HIGHの場合、CPUをより多く使用しますが、領域をもっとも多く節約するので、バックアップ・ファイルの書き込みに必要なI/O数をもっとも少なくなります。その一方で、CPUに制約がある場合は、LOWまたはMEDIUMがより適切かもしれません。CPUの使用量が少なめで、通常、約80%の領域が節約されます（Oracle RMAN基本圧縮と比較した場合）。

高度なLOB圧縮（SecureFiles LOBの場合）

通常、非構造化データをインラインで保存するのではなく、非構造化データをSecureFilesに移動（かつ高度なLOB圧縮を使用）することで表の圧縮率を改善できます。高度な圧縮では、ブロックのサイズを削減するために重複排除が使用されます。インラインで保存された非構造化データの場合、その非構造化データの複製が同じブロックに保存されることはあまりありません。つまり、ブロック内の非構造化データは、かなり大容量になる場合が多いですが、圧縮されません。このため、表の圧縮率が予想よりも全般的に低くなる可能性があります。ただし、高度なLOB圧縮では異なる圧縮アルゴリズムが使用されるため、多くの場合、インラインで保存されている場合には圧縮できない非構造化データを、SecureFiles LOBセグメントに保存されている場合に圧縮できます。

高度なLOB圧縮には、LOW、MEDIUM、HIGHの3つのレベルがあります。高度なLOB圧縮では、デフォルトでMEDIUMレベルが使用されます。MEDIUMレベルでは通常、3~5%という少ないCPUオーバーヘッドで適度な圧縮が実行されます。高度なLOB圧縮のLOWは、高パフォーマンス向けに最適化されています。高度なLOB圧縮のLOWでは、3分の1のCPU使用量で、MEDIUMによって達成される圧縮の約80%が維持されます。高度なLOB圧縮のHIGHでは、ストレージの節約が最大になりますが、CPUオーバーヘッドも最大になります。

テスト開始前の考慮事項

概念実証のテスト前計画の一部として、Oracleの圧縮で推奨される以下のベスト・プラクティスに留意し、（必要に応じて）実行してください。

- 最新リリースにアップグレードします（または重要なパッチを現行のリリースに適用）。以下のMOS Noteを参照してください。List of Critical Patches Required for Oracle 11g Table Compression (Doc ID 1061366.1)
- 概念実証の成功基準を定義します（データ、索引、およびバックアップ・ストレージ削減、問合せ/挿入/更新のパフォーマンス、バルク・ロード操作のパフォーマンス、アプリケーションのパフォーマンスなど）。
- Oracle E-Business Suiteを使用して概念実証を実行する場合は、Oracle MOS Note 2458161.1で詳細を確認してください。

Oracle RMANバックアップ圧縮の3つのレベル

3つのレベルは次のように分類されます。

HIGH - ネットワーク速度に制約がある、速度が遅めなネットワークを介したバックアップに最適

MEDIUM - ほとんどの環境に推奨。圧縮比と速度のバランスに優れている

LOW - バックアップ・スループットに一番影響が少なく、CPUリソースに制約がある環境に最適

SecureFilesデータの圧縮の可否の判断

Oracleでは、SecureFileデータが圧縮可能かどうかを検出され、業界標準の圧縮アルゴリズムを使用して圧縮されます。圧縮によって領域の削減が見込めない場合や、データがすでに圧縮されている場合、SecureFilesではそのようなLOBの圧縮は無効化されます。

- 圧縮列に“long”データ型が含まれていないことを確認します。このデータ型は、高度な行圧縮ではサポートされていません。
- 圧縮表/パーティションの列数が255未満であることを確認します（この制限はOracle Database 12c以降では解除されています）。詳しくは、Oracle MOS Note 1612095.1を参照してください。
- CPUオーバーヘッドは通常は最小限で済みますが、高度な行圧縮および索引圧縮は、CPUサイクルが空いているシステムに対して導入することが適しています。圧縮により、一部のDML操作では、非常に小さなものではありませんが、追加のオーバーヘッドが発生するためです。
- 各圧縮機能の最適なテスト環境は、本番環境に合わせて再現した環境です。この環境で、もっとも現実的な（圧縮前および圧縮後の）パフォーマンス比較と機能比較を行うことができます。
- 一般には、データベース内のアプリケーション関連のすべての表を圧縮することが推奨されますが、例外が1つあります。それは、表がキューとして使用される場合です。つまり、行が表に挿入された後、大部分またはすべての行が削除され、その後さらに多くの行が挿入された後で削除される場合です。このタイプのアクティビティは、圧縮のユースケースとしては適していません。本質的に一時的な行を絶えず圧縮するためのオーバーヘッドが生じるためです。
- 高度な行圧縮は、表領域レベルの暗号化によって適切に機能します。表は暗号化の前に圧縮されるため、圧縮率は暗号化の影響を受けません。列レベルの暗号化では、暗号化は圧縮の前に実行されるため圧縮率が低下します。

圧縮オーバーヘッドについて

概念実証を実行する前に、ユーザーは、圧縮解除のオーバーヘッドが問合せパフォーマンスに影響する可能性があるかと推測することがあります。しかし、実際のところ、そのような可能性は低いのが通常です。

高度な行圧縮、索引キー圧縮、および高度な索引圧縮が実行されたブロックは、ブロック・レベルで“圧縮解除”されることはなく、ほとんどの問合せでは、個々の行が圧縮解除されることもありません。ほとんどの問合せはメモリ内のデータベース・ブロックの圧縮形式で直接実行され、ほとんどの問合せ条件は圧縮されたデータ形式で直接実行されます。そして、問合せの後の段階で必要となる値のみが圧縮解除されます。

通常は、圧縮されたデータ/索引への問合せでのオーバーヘッドは増加せずに減少します。これは、一定量のユーザー・データに問合せを実行するためのI/Oが減るためです。たとえば、3倍の圧縮率でデータを圧縮する場合、圧縮の使用時にそのデータをディスクから読み取ってバッファ・キャッシュへ読み込むためのI/Oはわずか3分の1になります。確かに、圧縮されたデータ・ブロック内のポインタを間接参照して列値を抽出するために“追加の”命令サイクルがいくつか生じる可能性がありますが、通常はI/Oが減少することで相殺されます。

圧縮比の改善

特定の表やパーティションの圧縮率はおもに、その表やパーティションにブロック・レベルで存在する重複の量に関連します。

重複の量が多いほど圧縮率は高くなり、データの一意性が高いほど圧縮率は低くなります。データが一意である場合、表やパーティションは十分に圧縮されないか、まったく圧縮されない可能性が非常に高くなります。

特定の表の圧縮率を高めるために、できることがいくつかあります。通常どおりに、データ、アプリケーション、システムを使用して変更をテストし、変更が環境に及ぼす影響を特定する必要があります。

データのソート

データのロード時にデータを事前にソートしておくことで、表の圧縮率を改善できる場合があります。各列のデータのカーディナリティに基づいて、ソートする列を決定する必要があります。個別値が少ない列でソートできる場合は、圧縮率が向上する可能性があります。

ただし、事前にソートするには、データのロード前に追加の準備が必要となるため、追加される時間と圧縮率の向上を比較検討する必要があります。

概念実証の一般的な機能

前述したように、各圧縮機能の最適なテスト環境は、本番環境に合わせて再現した環境であることに留意することが重要です。この環境で、もっとも現実的な（圧縮前および圧縮後の）パフォーマンス比較と機能比較を行うことができます。

通常、圧縮概念実証の実行中にテストされるAdvanced Compressionの機能は以下のとおりです。

- 高度な行圧縮
- 高度な索引圧縮
- Oracle RMANバックアップ圧縮
- 高度なLOB圧縮/重複排除
- Oracle Data Guard REDO転送の圧縮

Advanced Compressionには他にも多数の機能が含まれますが、上記の機能が概念実証にもっとも一般的に含まれる機能です。Advanced Compressionの他の機能を含めることも、これらの機能の一部を含めないこともできます。

実際の場合、顧客は以下のように指摘しています。

- 概念実証テストの前に、構造化データ、索引、および非構造化データの圧縮率（ストレージ削減）を見積もります。圧縮アドバイザー（以下を参照）を使用して、高度な行圧縮、高度な索引圧縮、および高度なLOB圧縮の圧縮率を見積もることができます。
- テストを使用して、圧縮によるパフォーマンスの向上と、パフォーマンスへの考えられる影響を特定します。これを特定するには、アプリケーションを実行し、テストプラットフォーム（本番ハードウェアなど）上でデータを使用して、圧縮前と圧縮後のパフォーマンスをプロファイリングします。アプリケーション・テストには、アプリケーション問合せ、従来型パス・ロードとダイレクト・パス・ロードの両方を使用するバルク・ロード操作、単一行DML（従来型の挿入、更新、および削除操作）、およびOracle RMANバックアップを含めるのが理想です。

概念実証のテンプレート例

マルチステップ・プロセスでの圧縮概念実証の簡単な例については、後述の**付録A**を参照してください。

- 一般的にはすべての表を圧縮することが推奨されますが、中にはデータ・ストレージ要件のおよそ80 %以上を占める最大の表のみを圧縮することを選択している企業もあります。
- Data Guard Redo転送の圧縮を使用する場合の圧縮による節約の見積りについて詳しくは、MOS Note Doc ID 729551.1を参照してください。
- 使用可能な場合は、オラクルのReal Application Testing (RAT) 製品が圧縮概念実証に役立ちます。

無料の圧縮アドバイザー

Advanced Compressionを簡単に始める方法の1つは、圧縮アドバイザーを使用することです。“DBMS_COMPRESSION” PL/SQLパッケージ（一般的に圧縮アドバイザーと呼ばれる）が、データベース環境内の圧縮関連の情報を収集します。これには、非圧縮のパーティション化された表とパーティション化されていない表の双方の圧縮率の見積りや、以前に圧縮された表/パーティションの行レベルの圧縮情報の収集が含まれます。

圧縮アドバイザーにより、圧縮に関連する使用状況を判断するのに必要な、ストレージの節約情報を入手できます。

実行中の圧縮アドバイザーからは、圧縮アドバイザーのターゲットであった特定の表またはパーティションの推定圧縮率が出力されます。その出力で、“COMPRESSION_RATIO”は、2.1のような数値で表されます。

この特定の表またはパーティションで推定圧縮率が2.1倍である場合、この数値は、圧縮が有効になると、表またはパーティションのフットプリントが50 %程度削減されることを表しています。

Oracle Database 9i Release 2からOracle Database 11g Release 1までをサポートする圧縮アドバイザーのバージョンは、Oracle.comの高度な圧縮のページで無償で入手できます。以降のリリース向けの圧縮アドバイザーのバージョンは、Oracle Database Enterprise Edition 11g Release 2以降に付属しています。

付録A

マルチステップ・プロセスでの圧縮概念実証の例

関連するすべてのパッチを適用（任意で最新リリースにアップグレード）

- 該当する場合は最新のリリースにアップグレード
- パッチを適用

成功基準の定義

- データベースのパフォーマンス
- データベース・サイズ
- バックアップ領域の削減
- バックアップ時間/リストア時間
- アプリケーション・パフォーマンス

- Data Guard（該当する場合）
- Data Pump圧縮（該当する場合）

Compression Advisor -- DBMS_COMPRESSION

- データ/索引の圧縮率の見積もりを取得
- 圧縮される表/索引のリスト全体を確認

テスト環境での圧縮前のベースライン：本番ワークロード/データ

- データベースのパフォーマンス・データを収集（バルク・ロード操作、問合せ、挿入/更新など）
- バックアップ/リストア時間を収集
- Data Guardのパフォーマンス・データを収集（該当する場合）
- 表/索引のデータベース・サイズを収集
- バックアップ・サイズを収集
- Data Pumpのファイル・サイズを収集
- AWRLレポートを収集

テスト環境での圧縮の実装

- 推奨される方法を使用して特定された、候補となるすべての表/索引を圧縮（オンライン/オフライン）
- バルク・ロード操作を実行してベースラインと比較
- SQL文（問合せ/挿入/更新）を実行してベースラインと比較
- 非実行の問合せ（存在する場合）に対してSQLチューニングの調整を実行
- 本番ワークロードを実行してパフォーマンスを検証
- AWRLレポートを収集してベースラインと比較

本番カットオーバーの準備

- 得られた教訓
- 概念実証で得られた利点および課題/解決策をすべて文書化
- カットオーバー計画を定義

Connect with us

+1.800.ORACLE1までご連絡いただくか、[oracle.com](https://www.oracle.com)をご覧ください。北米以外の地域では、[oracle.com/contact](https://www.oracle.com/contact)で最寄りの営業所をご確認いただけます。

 blogs.oracle.com

 facebook.com/oracle

 twitter.com/oracle

Copyright © 2023, Oracle and/or its affiliates. All rights reserved. 本文書は情報提供のみを目的として提供されており、ここに記載されている内容は予告なく変更されることがあります。本文書は、その内容に誤りがないことを保証するものではなく、また、口頭による明示的保証や法律による黙示的保証を含め、商品性ないし特定目的適合性に関する黙示的保証および条件などのいかなる保証および条件も提供するものではありません。オラクルは本文書に関するいかなる法的責任も明確に否認し、本文書によって直接的または間接的に確立される契約義務はないものとします。本文書はオラクルの書面による許可を前もって得ることなく、いかなる目的のためにも、電子または印刷を含むいかなる形式や手段によっても再作成または送信することはできません。

本デバイスは、連邦通信委員会のルールに基づいた認可を未取得です。認可を受けるまでは、このデバイスの販売またはリースを提案することも、このデバイスを販売またはリースすることはありません。

OracleおよびJavaはOracleおよびその子会社、関連会社の登録商標です。その他の名称はそれぞれの会社の商標です。

IntelおよびIntel XeonはIntel Corporationの商標または登録商標です。すべてのSPARC商標はライセンスに基づいて使用されるSPARC International, Inc.の商標または登録商標です。AMD、Opteron、AMDロゴおよびAMD Opteronロゴは、Advanced Micro Devicesの商標または登録商標です。UNIXは、The Open Groupの登録商標です。0120

免責事項：データシートにこの免責事項の記載が必要かどうか分からない場合は、収益認識方針を参照してください。本書の内容と免責事項の要件についてさらに質問がある場合は、REVREC_US@oracle.com宛てに電子メールでご連絡ください。
