Oracle Advanced Compression: Reduce Storage, Reduce Costs, Increase Performance

Bill Hodak
Principal Product Manager
The following is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. 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 for Oracle’s products remains at the sole discretion of Oracle.
Agenda

• Data Growth Challenges

• Advanced Compression Feature Overview
  • Relational Data Compression
  • Unstructured Data Compression
  • Backup Data Compression
  • Network Transport Data Compression

• Questions and Answers
Challenges

• Explosion in Data Volumes
  • Government Regulations (Sarbanes-Oxley, etc)
  • User Generated Content (Web 2.0)
  • Application Consolidation

• IT Managers Must Support Larger Volumes of Data with Limited Technology Budgets
  • Need to optimize storage consumption
  • Also maintain acceptable application performance

• Intelligent and Efficient Compression Technology can Help Address These Challenges
Oracle Advanced Compression Option

- Reduces resource requirements and costs
  - Storage System
  - Network Bandwidth
  - Memory Usage

**Relational Data Compression**
- OLTP Table Compression

**Unstructured Data Compression**
- SecureFiles Deduplication
- SecureFiles Compression

**Backup Data Compression**
- Data Pump Data Compression
- RMAN Fast Backup Compression

**Network Data Compression**
- Data Guard Redo Transport Compression
Oracle Advanced Compression Option

- Reduces resource requirements and costs
  - Storage System
  - Network Bandwidth
  - Memory Usage
Table Compression

- Introduced in Oracle Database 9i Release 2
  - Compression during bulk load operations (Direct Load, CTAS)
  - Data modified using conventional DML not compressed
- Optimized compression algorithm for relational data
- Improved performance for queries accessing large amounts of data
  - Fewer I/Os
  - Buffer Cache efficiency
- Data is compressed at the database block level
- Compression enabled at either the table or partition level
- Completely transparent to applications
OLTP Table Compression

• Oracle Database 11g extends table compression for OLTP data
  • Support for conventional DML Operations (INSERT, UPDATE)
• New algorithm significantly reduces write overhead
  • Batched compression minimizes impact for OLTP transactions
• No impact on reads
  • Reads may actually see improved performance due to fewer I/Os and enhanced memory efficiency
OLTP Table Compression Process

Empty Block | Initially Uncompressed Block | Compressed Block | Partially Compressed Block | Compressed Block

Legend

- Header Data
- Uncompressed Data
- Free Space
- Compressed Data
Block-Level *Batch* Compression

- Patent pending algorithm minimizes performance overhead and maximizes compression
- Individual INSERTs and UPDATEs do not cause recompression
- Compression cost is amortized over several DML operations
- Block-level (Local) compression keeps up with frequent data changes in OLTP environments
  - Competitors use static, fixed size dictionary table thereby compromising compression benefits
OLTP Table Compression

Employee Table

<table>
<thead>
<tr>
<th>ID</th>
<th>FIRST_NAME</th>
<th>LAST_NAME</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>John</td>
<td>Doe</td>
</tr>
<tr>
<td>2</td>
<td>Jane</td>
<td>Doe</td>
</tr>
<tr>
<td>3</td>
<td>John</td>
<td>Smith</td>
</tr>
<tr>
<td>4</td>
<td>Jane</td>
<td>Doe</td>
</tr>
</tbody>
</table>

Initially Uncompressed Block

Header

1•John•Doe 2•Jane•Doe 3•John•Smith 4•Jane•Doe

Free Space

INSERT INTO EMPLOYEE
VALUES (5, 'Jack', 'Smith');
COMMIT;
OLTP Table Compression

### Employee Table

<table>
<thead>
<tr>
<th>ID</th>
<th>FIRST_NAME</th>
<th>LAST_NAME</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>John</td>
<td>Doe</td>
</tr>
<tr>
<td>2</td>
<td>Jane</td>
<td>Doe</td>
</tr>
<tr>
<td>3</td>
<td>John</td>
<td>Smith</td>
</tr>
<tr>
<td>4</td>
<td>Jane</td>
<td>Doe</td>
</tr>
<tr>
<td>5</td>
<td>Jack</td>
<td>Smith</td>
</tr>
</tbody>
</table>

### Compressed Block

**Header**

<table>
<thead>
<tr>
<th>John=0</th>
<th>Doe=1</th>
<th>Jane=2</th>
<th>Smith=3</th>
</tr>
</thead>
</table>

**Free Space**

- 1
- 0
- 1
- 2
- 1
- 3
- 0
- 3
- 4
- 2
- 1
- 5
- Jack
- 3

**Local Symbol Table**
Table Compression Syntax

OLTP Table Compression Syntax:

```
CREATE TABLE emp (
    emp_id NUMBER
    , first_name VARCHAR2(128)
    , last_name VARCHAR2(128)
) COMPRESS FOR OLTP;
```

Direct Load Compression Syntax (default):

```
CREATE TABLE emp (
    emp_id NUMBER
    , first_name VARCHAR2(128)
    , last_name VARCHAR2(128)
) COMPRESS [BASIC];
```
Table Compression Advisor

Estimate Potential Storage Savings

- Available in 11g Release 2
- Available on OTN *
  - Supports Oracle Database 9i Release 2 through 11g Release 1
  - Shows projected compression ratio for uncompressed tables
  - Reports actual compression ratio for compressed tables (11g Only)

```sql
SQL> set serveroutput on
SQL> EXECUTE DBMS_COMPRESSION.GETRATIO(OWNERNAME=>'SH', TABNAME=>'SALES', SAMPLING_PERCENT=>10
Sampling table: SH.SALES
Sampling percentage: 10%
Expected Compression ratio with Advanced Compression Option: 2.96

PL/SQL procedure successfully completed.
SQL> 
```

Monitoring Table Compression

- View: DBA_TABLES, columns:
  - COMPRESSION (ENABLED / DISABLED)
  - COMPRESS_FOR (OLTP / BASIC)

```sql
SQL> select table_name, compression, compress_for from user_tables where compression = 'ENABLED';

TABLE_NAME        COMPRESS COMPRESS_FOR
------------------ ------ --------
GL_JE_LINES       ENABLED OLTP
GL_E BALANCES     ENABLED BASIC

SQL>  
```
Table Compression Results
Oracle’s Internal E-Business Suite DB

• Overall database storage savings: 3x
  • Table compression 4x
  • Index compression 2x
  • LOB compression 2.3x

• 95 TB of Total Storage Savings!
  • Primary, standby, test, dev, and backup

• Payroll, Order-2-Cash, AP/AR batch flows, Self-Service flows run without regression, Queries involving full table scans show speedup
Oracle’s Internal Beehive Email DB

- Average Compression Ratio: 2x
- Oracle Database 11g Release 1
- Exadata Storage Servers
- Storage savings add up with standby, mirroring, flash recovery area
  - Phase I in production
    - Email for 28K employees
    - 195 TB of storage savings with SecureFiles Compression
  - Phase II, Dec 2009
    - Migrate all 90K employees on this email server
    - 581TB estimated storage savings with SecureFiles Compression
- Performance improved by caching more data due to compression - reducing I/O latencies
• Compression on SAP databases at leading global company
  • Oracle Database 11g Release 2
    • SAP R/3 DB
      • 4.67TB Uncompressed
      • 1.93 TB Compressed
        • 2.4x compression ratio
    • SAP BW DB
      • 1.38 TB Uncompressed
      • .53 TB Compressed
        • 2.6x compression ratio
  • Leverage 11g compression for Tables, Indexes and LOB data

ORACLE
Real Networks

- Rhapsody Digital Music Subscription Service

- Compression results using 11g Release 1
  - Average Compression Ratio: 2.8x
  - Highest compression ratio: 8x
  - Total Savings: 3.5TB
CERN

- Oracle Database 11g Release 2
- Average Compression Ratio: 2x
- Highest Compression Ratio: 6x
- Scientific application supporting particle research
Oracle Advanced Compression Option

- OLTP Table Compression
- SecureFiles Deduplication
- SecureFiles Compression
- Data Pump Data Compression
- RMAN Fast Backup Compression
- Data Guard Redo Transport Compression
- Relational Data Compression
- Unstructured Data Compression
- Backup Data Compression
- Network Data Compression

- Reduces resource requirements and costs
  - Storage System
  - Network Bandwidth
  - Memory Usage
Introduction to SecureFiles

• Next-generation high performance LOB
  • Superset of LOB interfaces allows easy migration from LOBs
  • Transparent deduplication, compression, and encryption
  • Leverage the security, reliability, and scalability of database

• Enables consolidation of file data with associated relational data
  • Single security model
  • Single view of data
  • Single management of data
  • Scalable to any level using SMP scale-up or grid scale-out
SecureFiles Deduplication

- Enables storage of a single physical image for duplicate data
- Significantly reduces space consumption
- Dramatically improves writes and copy operations
- No adverse impact on read operations
  - May actually improve read performance for cache data
- Duplicate detection happens within a table, partition or sub-partition
- Specially useful for content management, email applications and data archival applications
SecureFiles Compression

- Significant storage savings for unstructured data
  - Three levels of compression (LOW/[MEDIUM]/HIGH) provide desired ratios
  - 2-3x compression for typical files (combination of doc, pdf, xml)
- Compression Level LOW (NEW in 11.2)
  - Compression algorithm optimized for high performance
    - 3x less CPU utilization than default SecureFiles Compression
    - Maintains 80% compression of default SecureFiles Compression
- Allows for random reads and writes to Compressed SecureFile data
- Can be specified at a partition level
- Automatically detects if SecureFile data is compressible
- Independent of table or index compression
SecureFiles Compression Syntax

Compression Syntax

```sql
CREATE TABLE t1 (a CLOB)
LOB(a) STORE AS SECUREFILE (COMPRESS CACHE);
```

Deduplication Syntax

```sql
CREATE TABLE t1 (a CLOB)
LOB(a) STORE AS SECUREFILE (DEDUPLICATE CACHE);
```
Oracle Advanced Compression Option

- Reduces resource requirements and costs
  - Storage System
  - Network Bandwidth
  - Memory Usage
Data Pump Compression

- Metadata compression available since Oracle Database 10g
- Oracle Database 11g extends compression to table data during exports
  - No need to decompress before import
  - `COMPRESSION={ALL | DATA_ONLY | [METADATA_ONLY] | NONE}`
- Single step compression of both data and metadata
  - Compressed data directly hits disk resulting in reduced disk space requirements
  - Internal tests reduced dump file size up to 75%
- Application transparent
  - Complete Data Pump functionality available on compressed files
Advanced Compression

*New in Oracle Database 11g Release 2*

- **RMAN Backup Compression**
  - Compression Level LOW (New in 11.2)
    - Fastest compression algorithm
    - Best suited when backup is constrained by CPU
  - Compression Level MEDIUM (New in 11.1)
    - Balance between CPU usage and compression ratio
    - *Formerly Fast RMAN Backup Compression*
      - 11.1 syntax supported in 11.2
  - Compression LEVEL HIGH (New in 11.2)
    - Best compression ratio and highest CPU utilization
    - Best suited when backup is constrained by network or I/O
Backup Compression Syntax

Data Pump Syntax

PROMPT> expdp hr DIRECTORY=dpump_dir1
       DUMPFILE=hr_comp.dmp COMPRESSION=ALL

---

FAST RMAN Backup Compression Configuration

RMAN> configure compression algorithm 'MEDIUM';

RMAN Backup Compression Syntax

RMAN> BACKUP AS COMPRESSED BACKUPSET DATABASE
       PLUS ARCHIVELOG;
RMAN Compression Results
RMAN Compression Overview

- Data from Oracle’s implementation of Oracle Applications
  - 3.5 GB Database
  - Oracle Enterprise Linux
  - Oracle Database 11g Release 1

- Test 1: Slow I/O (16 MB/s)
  - 11g RMAN without Compression
  - 10g RMAN with Compression
  - 11g RMAN with MEDIUM Compression

- Test 2: Fast I/O (200 MB/s)
  - 11g RMAN without Compression
  - 10g RMAN with Compression
  - 11g RMAN with MEDIUM Compression
Backup Compression Results

Backup Size Comparison

Compression reduced backup size by 6x
Backup Compression

Backup Speed Comparison
Slow I/O (Tape)

11g Compression is almost 2.5x faster than 10g Compression
Backup Compression

Backup Speed Comparison
Fast I/O (Disk)

11g Compression is almost 2.5x faster than 10g Compression
Oracle Advanced Compression Option

- Reduces resource requirements and costs
  - Storage System
  - Network Bandwidth
  - Memory Usage
Network Compression
Data Guard Redo Transport Services

• Compress network traffic between primary and standby databases

• Lower bandwidth networks (<100Mbps)
  • 15-35% less time required to transmit 1 GB of data
  • Bandwidth consumption reduced up to 35%

• High bandwidth networks (>100 Mbps)
  • Compression will not reduce transmission time
  • But will reduce bandwidth consumption up to 35%

• Syntax:
  \[ \text{LOG\_ARCHIVE\_DEST\_3} = '\text{SERVICE=denver SYNC}
  \text{COMPRESSION=ENABLE|DISABLE}' \]

• Ref. MetaLink 729551.1 “Redo Transport Compression in a Data Guard Environment”
Redo Transport Compression

OLTP workload

- More efficient bandwidth utilization, up to 5x compression ratio
- Compression did not impact throughput or response time

Validation performed by CTC in collaboration with Oracle Japan Grid Center

http://www.ctc-g.co.jp/en/
Summary

• Comprehensive data compression capabilities for all types of data
  • Structured, Unstructured, Backup, Network
• Reduces storage consumption by 2 to 4 times
• Improves read performance
• Enhances memory, buffer cache utilization
• Complete application transparency
• Benefits diverse application workloads
For More Information

search.oracle.com

advanced compression

Or

ORACLE IS THE INFORMATION COMPANY