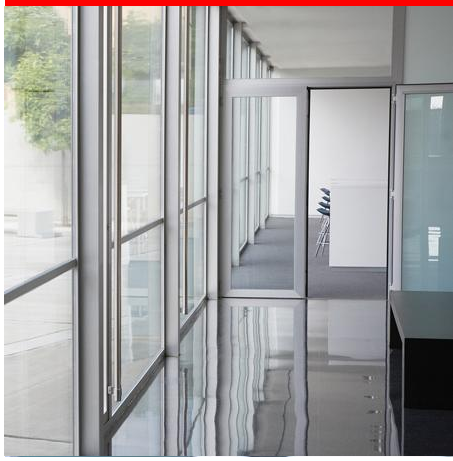

ORACLE®

ORACLE®

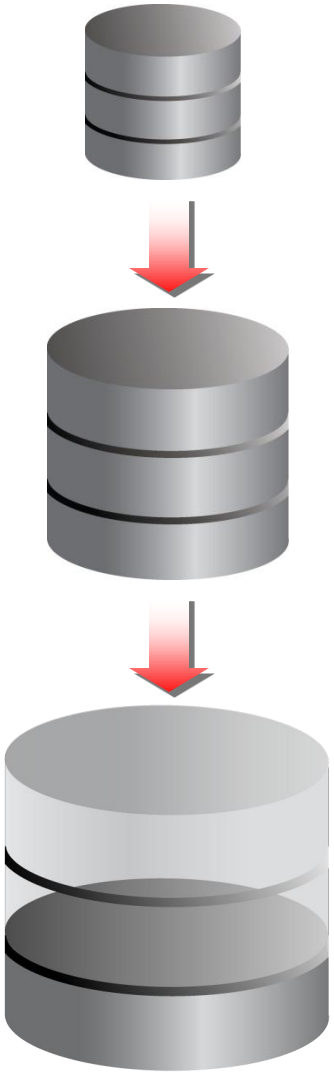
Exadata Hybrid Columnar Compression: The Next-Generation Compression Technology

Program Agenda



- Data Growth Challenges
- Oracle Compression Overview
- Exadata Hybrid Columnar Compression
  - Warehouse Compression
  - Archive Compression
- Questions and Answers

Data Growth Challenges



- IT must support exponentially growing amounts of data
  - Explosion in online access and content
  - Government data retention regulations
- Performance often declines as data balloons
- IT budgets are flat or decreasing
- Need to grow data
  - Without hurting performance
  - Without growing cost
- Powerful and efficient compression is key

Oracle Database Compression Overview

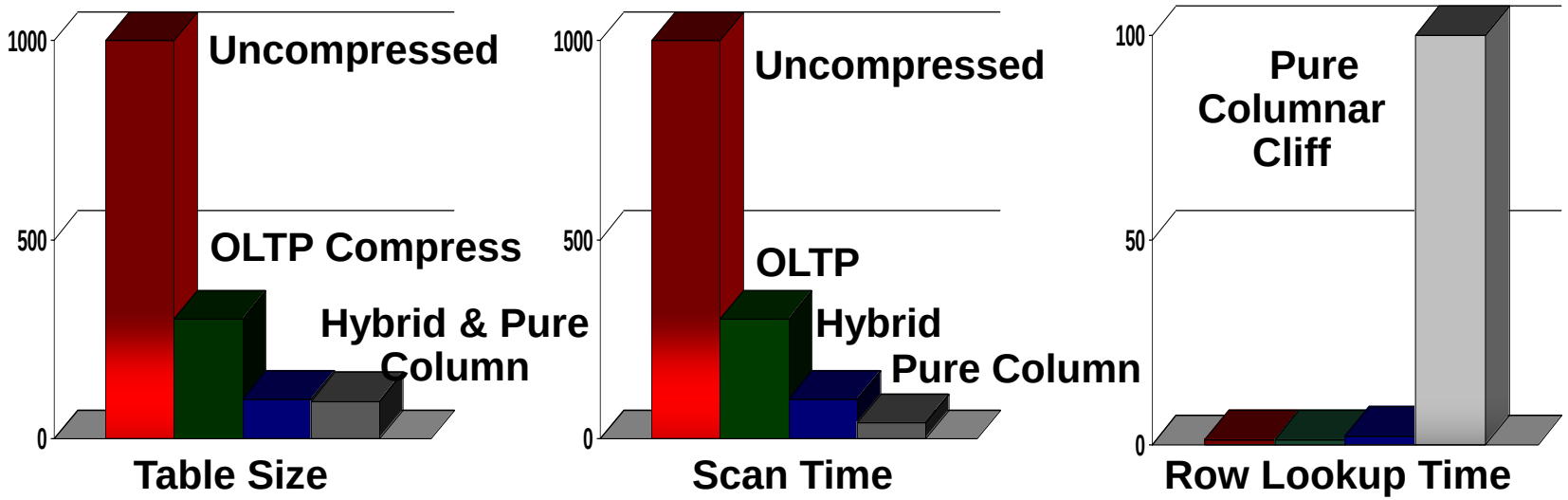
Compress All Your Data

Compression Feature	Application Fit	Availability
Warehouse Compression **Online Archival Compression**		**Built on Exadata Hybrid Columnar Compression Technology**
SecureFiles Compression	• Unstructured (File) Data	Database 11g Advanced Compression
SecureFiles Deduplication	• All Applications • Unstructured (File) Data	Database 11g Advanced Compression
Backup Compression	• RMAN Compression • Data Pump Compression	Database 11g Advanced Compression
Network Compression	• Data Guard Redo Transport Compression	Database 11g Advanced Compression

Why Hybrid Columnar Compression

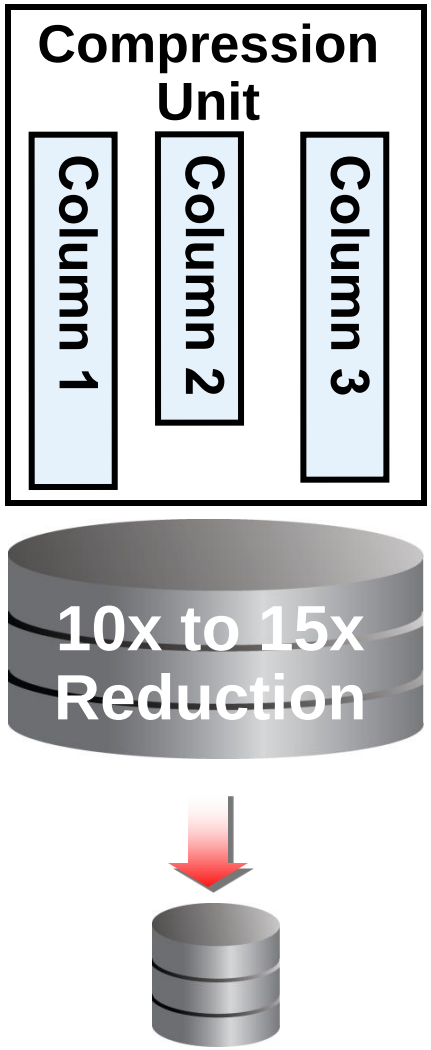
- Traditionally, data has been organized within a database block in a ‘row’ format, where all column data for a particular row is stored sequentially within a single database block
- An alternative approach is to store data in a ‘columnar’ format, where data is organized and stored by column.
  - Storing column data together, with the same data type and similar characteristics, drastically increases the storage savings achieved from compression
- Oracle’s Exadata Hybrid Columnar Compression (EHCC) utilizes a combination of both row and columnar methods for storing data
  - This hybrid approach achieves the compression benefits of columnar storage, while avoiding the performance shortfalls of a pure columnar format

Hybrid Columnar Comparisons



- Exadata Hybrid Columnar Compression is a second generation columnar technology combining the best of row and column formats
  - Best compression – matching full columnar
  - Excellent scan time – 93% as good as full columnar
  - Good single row lookup – no full columnar “cliff”
- Row format remains best for workloads with updates or trickle feeds

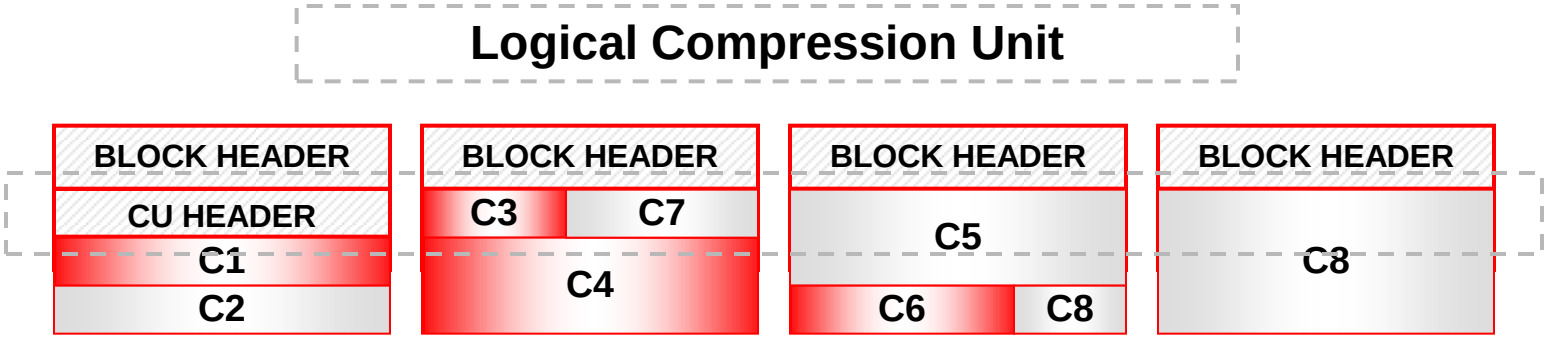
About Exadata Hybrid Columnar Compression



- Hybrid Columnar Compressed Tables
  - New approach to compressed table storage
  - Compressed tables can still be modified using conventional DML operations, such as INSERT and UPDATE
  - Useful for data that is bulk loaded and queried
    - Update activity is light
- How it Works
  - Tables are organized into Compression Units (CUs)
    - CUs are larger than database blocks
  - Within Compression Unit, data is organized by column instead of by row
    - Column organization brings similar values close together, enhancing compression

Compression Units

- Compression Unit
  - Logical structure spanning multiple database blocks
  - Data organized by column during data load
  - Each column compressed separately
  - All column data for a set of rows stored in compression unit



Decompression

- Queries run directly on Hybrid Columnar Compressed data does not require the data to be decompressed
- Data that is required to satisfy a query predicate does not need to be decompressed; only the columns and rows being returned to the client are decompressed in memory
- The decompression process typically takes place on the Oracle Exadata Storage Server in order to maximize performance and offload processing from the database server

Exadata Hybrid Columnar Compression

Warehouse and Archive Compression

Warehouse Compression

- 10x average storage savings
- 10x reduction in Scan IO

Optimized for Speed

Archive Compression

- 15x average storage savings
  - Up to 70x on some data
- For cold or historical data

Optimized for Space

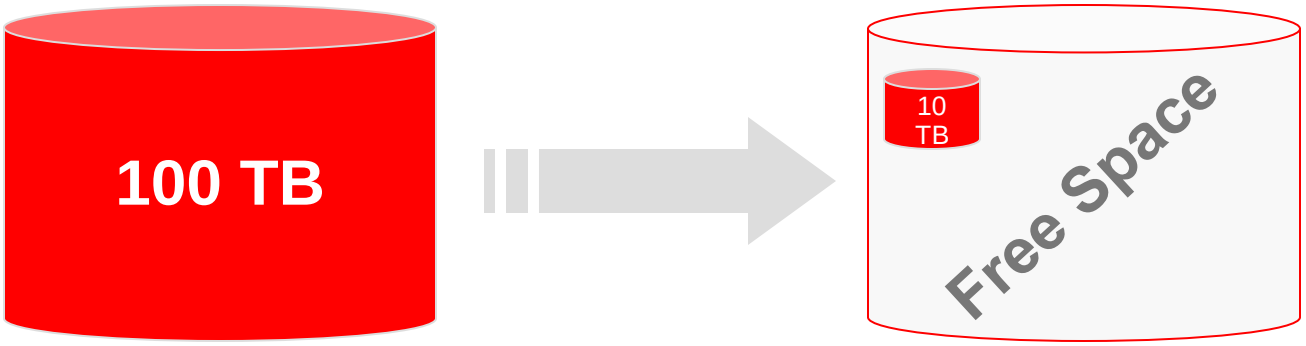
Can mix OLTP and hybrid columnar compression by partition for ILM

Warehouse Compression Levels

- Warehouse Compression provides two levels of compression: LOW and HIGH.
  - HIGH typically provides a 10x reduction in storage
  - LOW typically provides a 6x reduction
- Both levels have been optimized to increase scan query performance by taking advantage of the fewer number of blocks on disk
- To maximize the storage savings and query performance benefits of Warehouse Compression, the default level is HIGH.
  - LOW should be chosen for environments where load time service levels are more critical than query performance

Warehouse Compression

- 10x average storage savings
  - 100 TB Database compresses to 10 TB
  - Reclaim 90 TB of disk space
  - Space for 9 more ‘100 TB’ databases
- 10x average scan improvement
  - 1,000 IOPS reduced to 100 IOPS



Archive Compression

- Compression algorithm optimized for max storage savings
- Benefits any application with data retention requirements
- Best approach for ILM and data archival
  - Minimum storage footprint
  - No need to move data to tape or less expensive disks
  - Data is always online and always accessible
    - Intended for tables or partitions that store data that is rarely accessed
    - Run queries against historical data (without recovering from tape)
    - Update historical data
  - Supports schema evolution (add/drop columns)

Archive Compression

- Optimal workload characteristics for Archive Compression
  - Any application (OLTP, Data Warehouse)
  - Cold or Historical Data
  - Data loaded with bulk load operations
  - Minimal access and update requirements
- 15x average storage savings
  - 100 TB Database compresses to 6.6 TB
  - Keep historical data online forever
  - Up to 70x savings seen on production customer data

Archive Compression

ILM and Data Archiving Strategies

- OLTP Applications
  - Table Partitioning
  - Heavily accessed data
    - Partitions using OLTP Table Compression
  - Cold or historical data
    - Partitions using Online Archival Compression
- Data Warehouses
  - Table Partitioning
  - Heavily accessed data
    - Partitions using Warehouse Compression
  - Cold or historical data
    - Partitions using Online Archival Compression

Table Compression Syntax

Warehouse Compression Syntax:

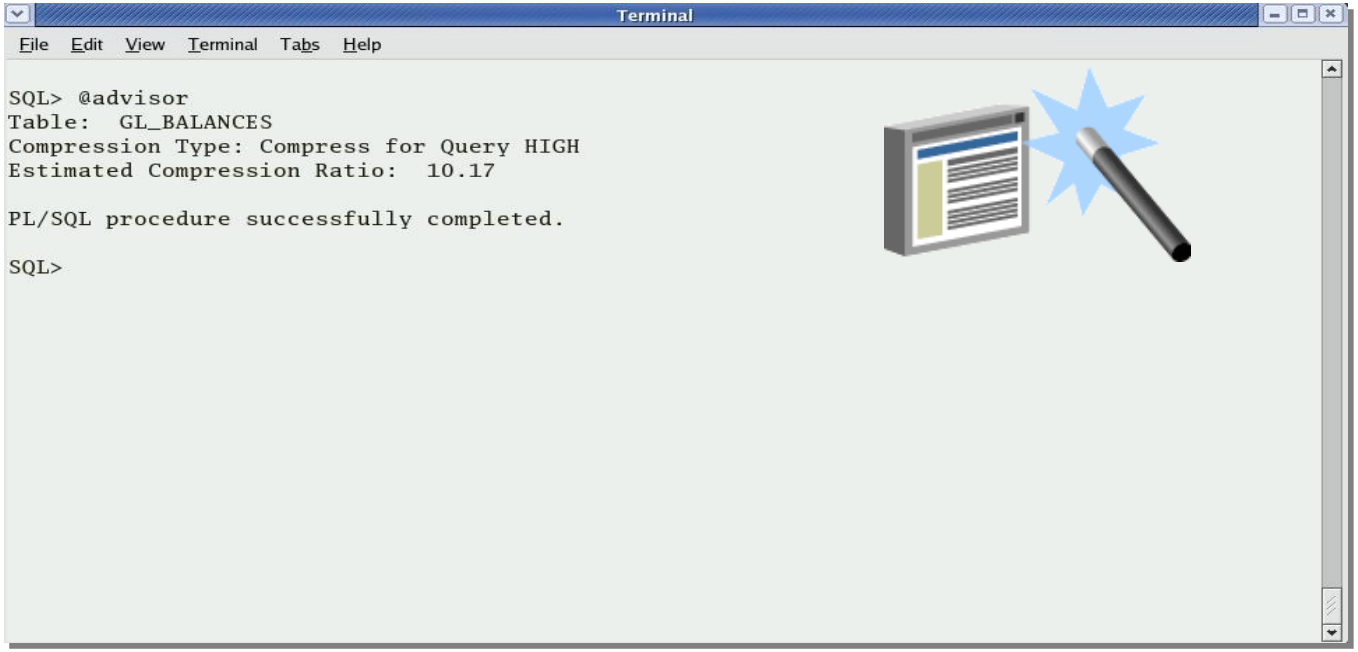
```
CREATE TABLE emp (…)
 COMPRESS FOR QUERY [LOW | HIGH];
```

Online Archival Compression Syntax:

```
CREATE TABLE emp (…)
 COMPRESS FOR ARCHIVE [LOW | HIGH];
```

Compression Advisor

- New Advisor in Oracle Database 11g Release 2
  - `DBMS_COMPRESSION` PL/SQL Package
  - Estimates Hybrid Columnar Compress storage savings on non-Exadata hardware
  - Requires Patch # 8937922



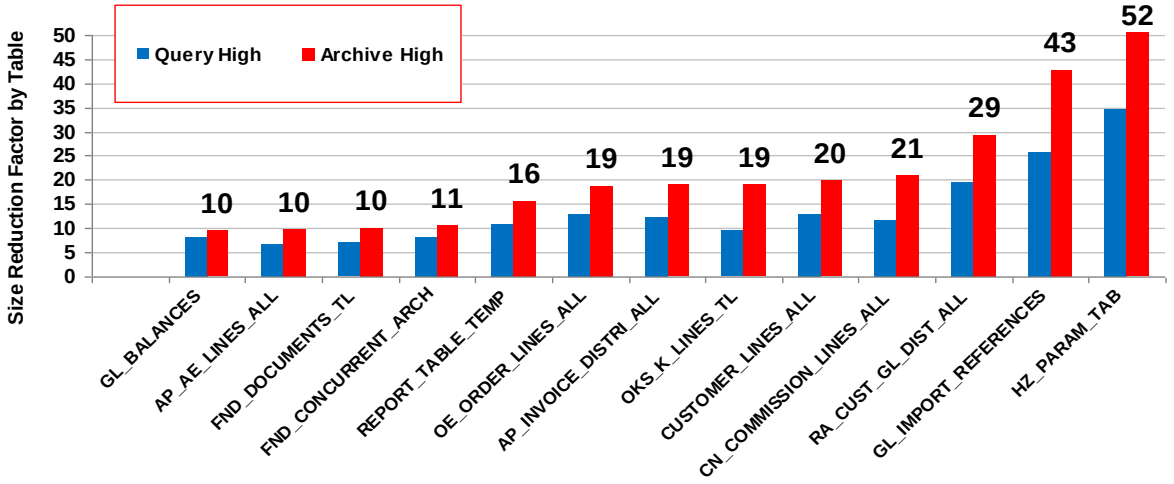
Exadata Hybrid Columnar Compression

Customer Success Stories

- Data Warehouse Customers (Warehouse Compression)
  - Top Financial Services 1: 11x
  - Top Financial Services 2: 24x
  - Top Financial Services 3: 18x
  - Top Telco 1: 8x
  - Top Telco 2: 14x
  - Top Telco 3: 6x
- Scientific Data Customer (Archive Compression)
  - Top R&D customer (with PBs of data): 28x
- OLTP Archive Customer (Archive Compression)
  - SAP R/3 Application, Top Global Retailer: 28x
  - Oracle E-Business Suite, Oracle Corp.: 23x
  - Custom Call Center Application, Top Telco: 15x

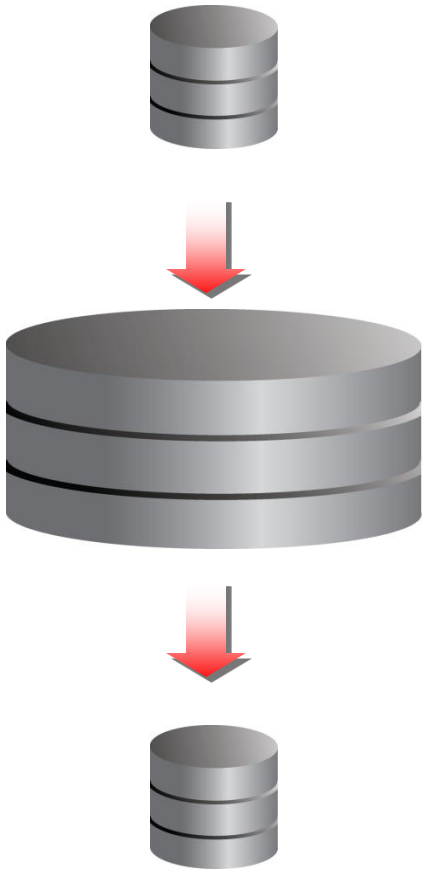
Oracle's Production EBS Dataset

- Archive High
  - Overall: 23x
  - Largest Table: 52x
- Query High
  - Overall: 15x
  - Largest Table: 35x
- OLTP Compression
  - Overall: 3.6
  - Largest Table: 4x



Up to 52x reduction in table size for EBS

Summary



- Average storage savings can range from 10x to 15x
- Need to grow data
  - Without growing cost
  - Without hurting performance
- Exadata Hybrid Columnar Compression
  - Extreme Storage Savings
    - Warehouse Compression
    - Archive Compression
  - Improve I/O Scan Rates

Oracle Products Available Online

ORACLE STORE

Oracle Store

Buy Oracle license and support online today at
oracle.com/store

SOFTWARE. HARDWARE. COMPLETE.


ORACLE®