# Oracle Exadata Hybrid Columnar Compression

Morana Kobal Butković
Principal Sales Consultant
Oracle Hrvatska

Oracle Database Compression Overview
Compress All Your Data

Compression Feature	Application Fit	Availability
Warehouse Compression	• Data Warehouses • Scan Oriented Access	Exadata

Warehouse Compression

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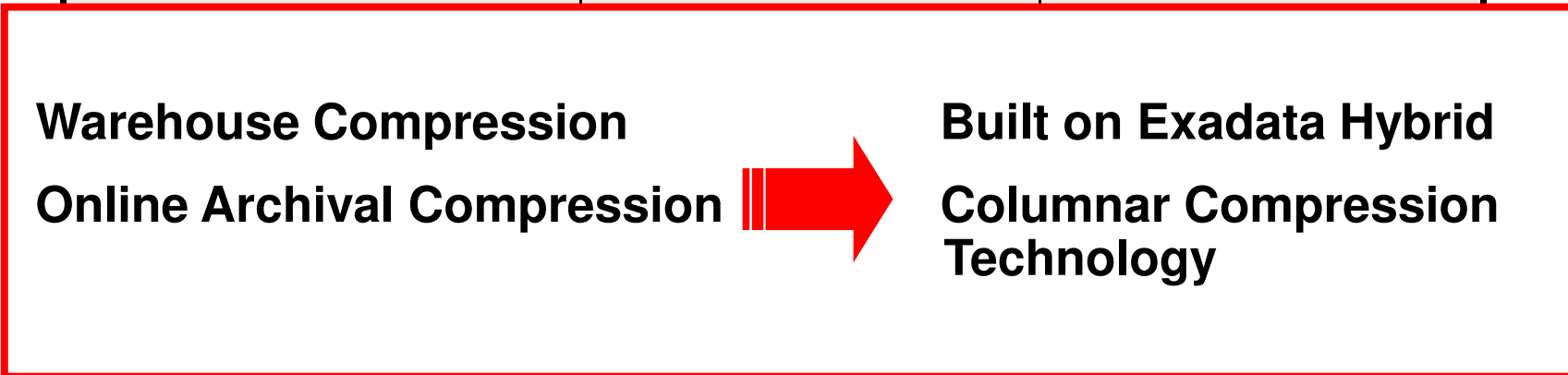
|		
SecureFiles Deduplication	All Applications • Unstructured (File) Data	Database 11g Advanced Compression
Backup Compression	• RMAN Compression • Data Pump Compression	Database 11g Advanced Compression
Network Compression	• Data Guard Redo Transport Compression	Database 11g Advanced Compression

Why Hybrid Columnar Compression

- **Traditionally: data is organized in 'row' format within database block**

- **Alternative approach: store data in a 'columnar' format, data is organized and stored by column**
  - Storing column data together drastically increases storage savings achieved from compression
- **Exadata storage has been optimized to maximize performance of queries when accessing HCC tables**
  - Takes advantage of processing power, memory and Infiniband network that are integral parts of Exadata storage server
  - Hybrid Columnar Compression extended to Pillar Axiom and Sun ZFS Storage Appliance (ZFSSA) storage

Compression Unit

Column 1 · Column 2 · Column 3 · Column 4 · Column 5
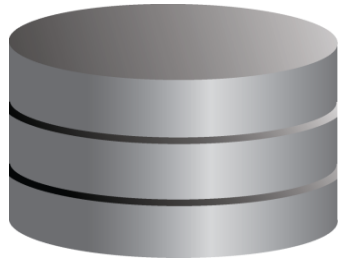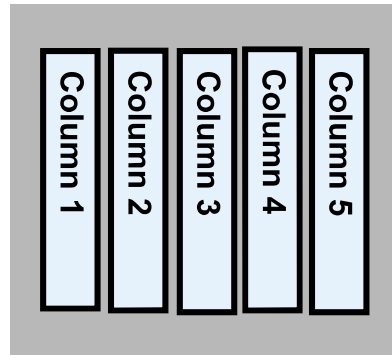
10x to 15x
Reduction

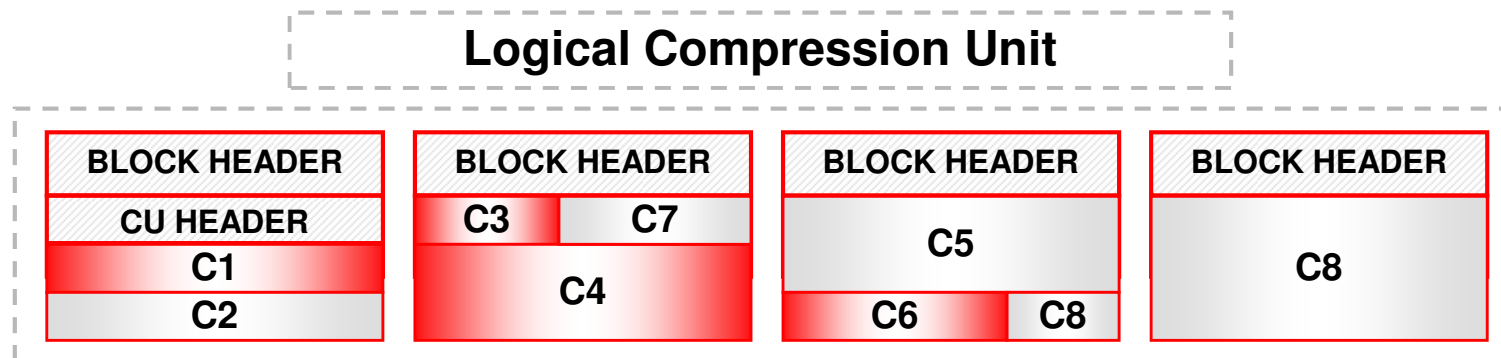- # Hybrid Columnar Compressed Tables
  - For data that is BULK loaded and queried
    - Designed for data that is NOT frequently updated
    - Designed for LOW concurrency environments: transaction modifying a single row in a CU locks the entire CU
    - Compressed tables allow using conventional DML (INSERT/UPDATE/DELETE)
    - All index types are supported (B-Tree, Bitmap)

- # How it Works
  - Tables are organized into Compression Units (CUs)
  - Within Compression Unit, data is organized by column instead of by row
    - Column organization brings similar values close together, enhancing compression

ORACLE

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

Logical Compression Unit

BLOCK HEADER	BLOCK HEADER	BLOCK HEADER	BLOCK HEADER
CU HEADER	C3 C7	C5	C8
C1	C4	C6 C8	
C2			

ORACLE®

Example

Table 7–17   Sample Table daily_sales

Item_ID	Date	Num_Sold	Shipped_From	Restock
1000	01-JUN-07	2	WAREHOUSE1	Y
1001	01-JUN-07	0	WAREHOUSE3	N
1002	01-JUN-07	1	WAREHOUSE3	N
1003	01-JUN-07	0	WAREHOUSE2	N
1004	01-JUN-07	2	WAREHOUSE1	N
1005	01-JUN-07	1	WAREHOUSE2	N

Conceptually compressed value can be represented as follows:
WAREHOUSE1WAREHOUSE3WAREHOUSE2

Exadata Hybrid Columnar Compression
Warehouse and Archive Compression

Warehouse Compression

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

Optimized for Speed

**Smaller Warehouse
Faster Performance**

Archive Compression

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

Optimized for Space

**Reclaim 93% of Disks
Keep Data Online**

Can mix OLTP and hybrid columnar compression by partition for ILM

ORACLE

Warehouse Compression

- **Warehouse Compression: LOW and HIGH**
  - HIGH typically provides a 10x reduction in storage
  - LOW typically provides a 6x reduction

- **Both levels optimized to increase scan query performance by taking advantage of fewer number of blocks reads**

- **To maximize storage savings and query performance use default level - HIGH**
  - LOW should be chosen for environments where load times are more critical than query performance

ORACLE

Archive Compression

- **Archive Compression: LOW and HIGH**
  - HIGH typically provides a 15x reduction in storage
  - LOW typically provides a 10x reduction

- **Best approach for ILM and data archival**
  - Minimum storage footprint
  - Minimal access and update requirements
  - No need to move data to tape or less expensive disks
  - Data is always online and always accessible
    - Run queries against historical data (without recovering from tape)
    - Update historical data
    - Supports schema evolution (add/drop columns)

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

ORACLE

EHCC DDL

- CTAS (create table as select)
  - create table foo compress for query as select * from bar1;
- IAS (insert direct load)
  - create table foo compress for archive low;
  - insert /*+APPEND*/ into foo select * from bar2;
- Compression can be specified at segment level
  - Each partition can have different compression type
  - create table orders (cid, pid, sid, price, discount, odate)
    partition by range (cid)
    (partition p1 values less than (100000) nocompress,
     partition p2 values less than (200000) compress for archive low,
     partition p3 values less than (300000) compress for query high,
     partition p4 values less than (maxvalue) compress for query low)
    enable row movement
   as select * from prev_orders;

EHCC DDL

- Existing tables/partitions can be converted to use EHCC
  - alter table bar move compress for query high
  - alter table orders modify partition p1 compress for archive low

  Tables are locked for DML during move
- New data loaded in existing tables/partitions can go to EHCC blocks, keeping existing data as is
  - alter table bar compress for query high
  - alter table orders modify partition p3 compress for query low
- EHCC can be disabled
  - alter table bar nocompress
  - alter table bar move nocompress
- Online redefinition package can be used for enabling/disabling EHCC: DBMS_REDEFINITION
- Import Data Pump compressed table as uncompressed using: TRANSFORM:SEGMENT_ATTRIBUTES=n

Exadata Hybrid Columnar Compression
Business as Usual

- Fully supported with…
  - B-Tree, Bitmap Indexes, Text indexes
  - Materialized Views
  - Exadata Server and Cells including offload
  - Partitioning
  - Parallel Query, PDML, PDDL
  - Schema Evolution support, online, metadata-only add/drop columns
  - Data Guard Physical Standby Support
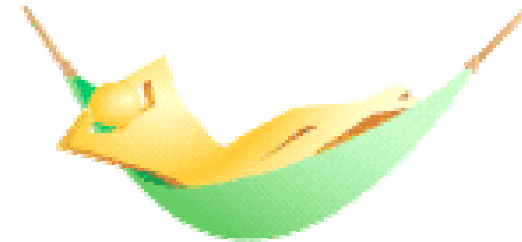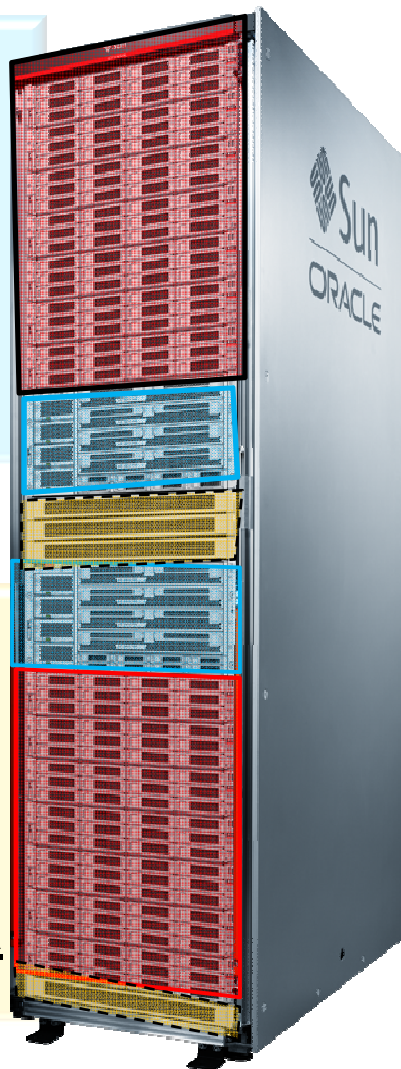
Exadata Hardware Architecture

Database Grid
- 8/4/2 Database Servers
- 2 x 6-core Intel Xeon processors per server
- 96GB DRAM per server
- Oracle Database 11.2
- Oracle Linux or Solaris

InfiniBand Network
- 3 x 36-port 40Gb/s switches
- **Unified** server & storage network

1 or 10 Gb Ethernet to Data Center

Intelligent Storage Grid
- 14/7/3 Storage Servers

- 12 x 600GB High Performance or
  12 x 2TB High Capacity disks

- 2 x 6-core Intel Xeon processors per server

- 4 x 96GB Flash PCIe per server

Intelligent Exadata Storage Server Software

Efficient Data Storage

- EHCC is a feature of Exadata Storage
  - Decompression, selection and projection performed on storage
  - Data is stored compressed on disk and compressed in the Flash Cache, frequently accessed data cached on Flash
    - Table can be forced to be on Flash by setting *cell_flash_cache_keep*

- EHCC is tightly integrated with Oracle DB 11gR2
  - Data is stored compressed in buffer cache (DRAM)

- With EHCC, entire databases can now run in memory
  - DRAM can hold 5TB of a compressed database
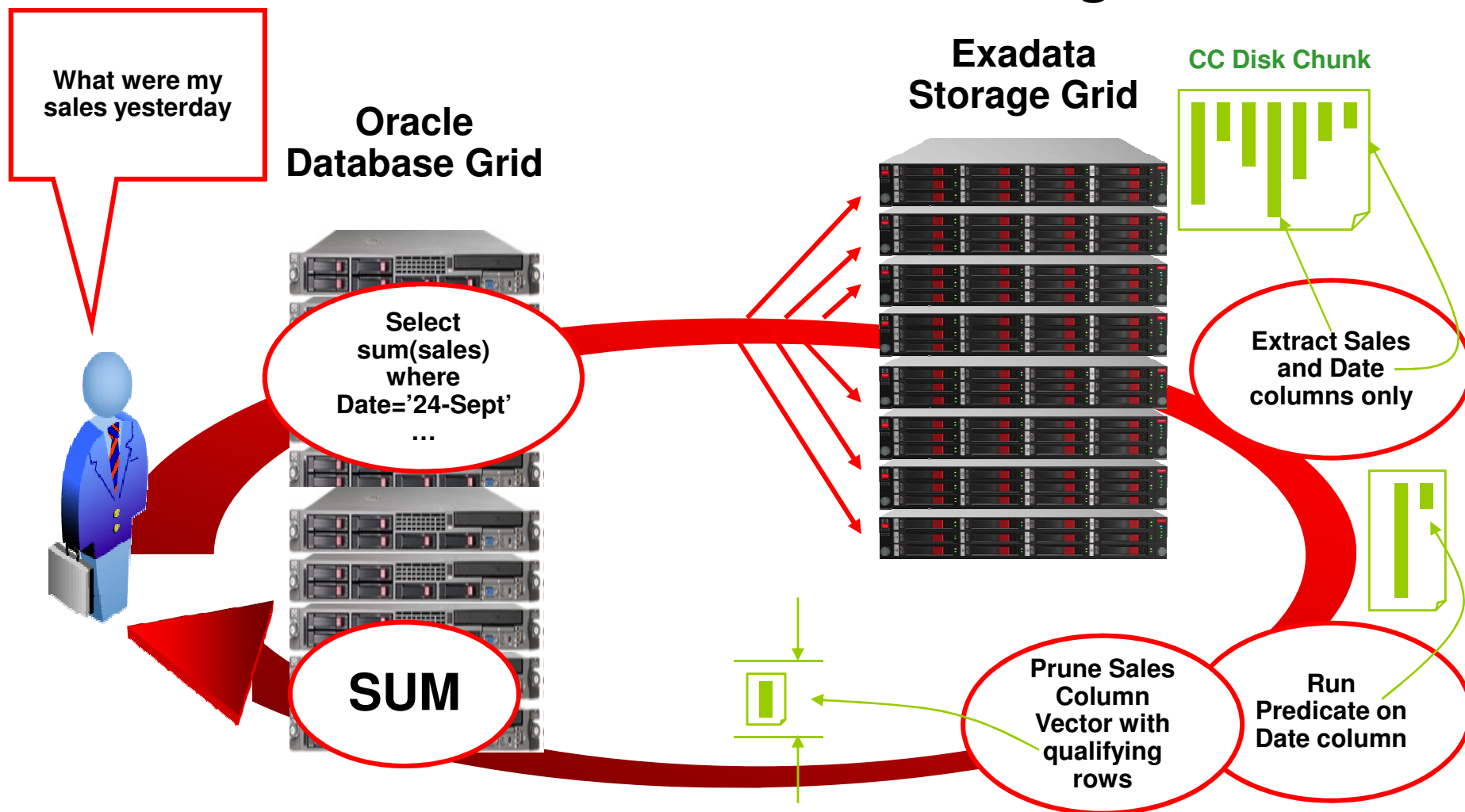 - Flash can hold 50TB of a compressed database

Efficient Data Movement

- Read/Write Compressed data to disk
  - Write Compressed data to ASM mirrors

- Read/Write Compressed data in Flash Cache
  - 10x improvement for Flash price performance

- Send Compressed data over Infiniband

- Write Compressed data to Redo Logs

- Send Compressed data to Standby
  - 10x reduction in WAN bandwidth cost: makes ADG appealing for DW

- Write Compressed data to Backups

Efficient Data Processing

- Specialized columnar query processing engine runs in Exadata Storage Server to run directly against compressed data
  - Column optimized processing of query projection and filtering
  - Vector processing techniques used to fully leverage columnar format
- 10x smaller subset of qualifying data returned over Infiniband to database server for further query processing
- Optimized single row lookups to perform efficient I/O of a contiguous set of blocks that form a Compression Unit

ORACLE

Smart Scans of Columnar Compressed Tables offloaded to Exadata Storage

What were my sales yesterday

Oracle Database Grid

Exadata Storage Grid

CC Disk Chunk

Select sum(sales) where Date='24-Sept' ...

Extract Sales and Date columns only

SUM

Prune Sales Column Vector with qualifying rows

Run Predicate on Date column

ORACLE

Table 7–18   Table Compression Methods

Table Compression Method	Compression Level	CPU Overhead	Applications
Basic compression	High	Minimal	DSS
OLTP compression	High	Minimal	OLTP, DSS
Warehouse compression	Higher[1]	Higher[2]	DSS
Online archival compression	Highest[1]	Highest[2]	Archiving

[1] Compression level depends on compression level specified (LOW or HIGH)

[2] CPU overhead depends on compression level specified (LOW or HIGH)

Table 7–19 Table Compression Characteristics

Table Compression Method	CREATE/ALTER TABLE Syntax	Direct-Path Insert	DML
Basic compression	`COMPRESS [BASIC]`[1]	Yes	Yes[2]
OLTP compression	`COMPRESS FOR OLTP`	Yes	Yes
Warehouse compression	`COMPRESS FOR QUERY [LOW|HIGH]`	Yes	Yes[3,4]
Online archival compression	`COMPRESS FOR ARCHIVE [LOW|HIGH]`	Yes	Yes[2,4]

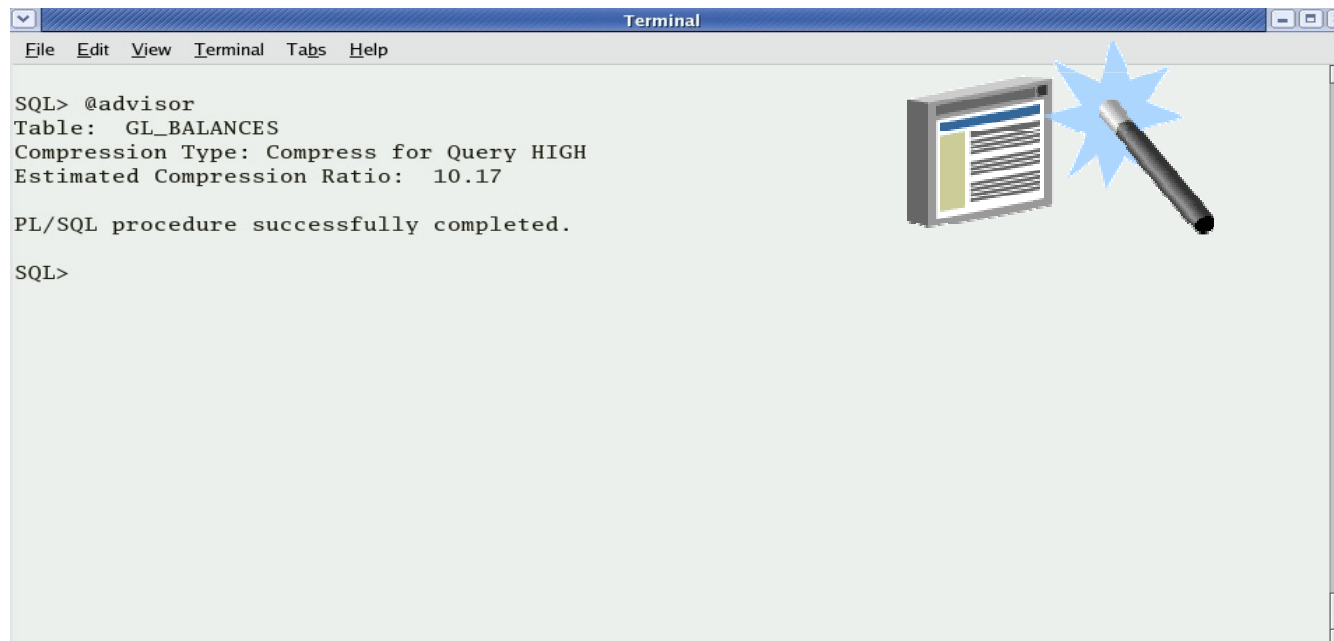[1] `COMPRESS` and `COMPRESS BASIC` are equivalent

[2] Inserted and updated rows are uncompressed

[3] High CPU overhead

[4] Inserted and updated rows go to a block with a less compressed format and have lower compression level

Compression Advisor

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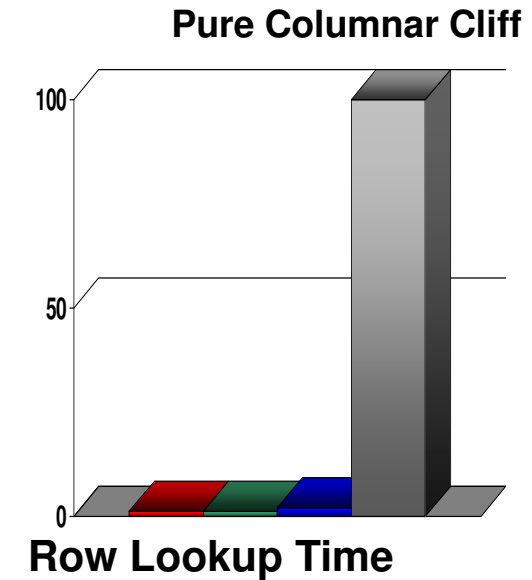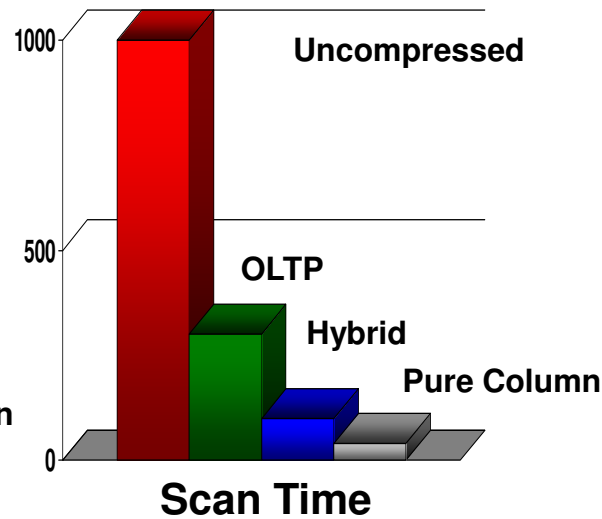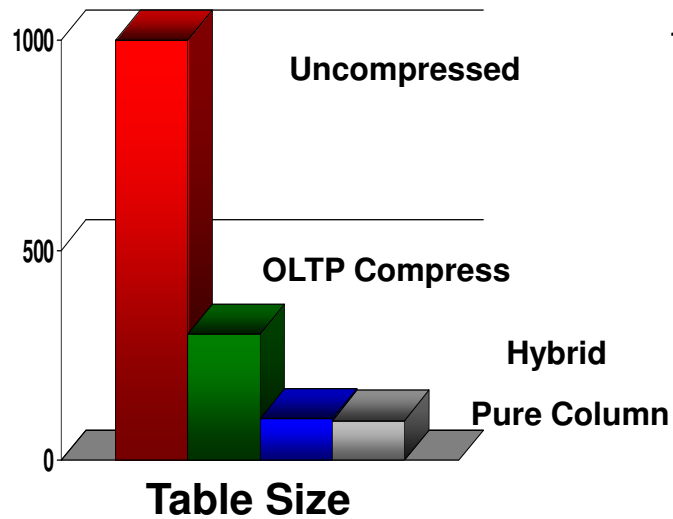 - Requires Patch # 8937922

```
SQL> @advisor
Table:  GL_BALANCES
Compression Type: Compress for Query HIGH
Estimated Compression Ratio:  10.17

PL/SQL procedure successfully completed.

SQL>
```

Hybrid Columnar Comparisons

Pure Columnar Cliff

Chart 1 — **Table Size**: Uncompressed, OLTP Compress, Hybrid, Pure Column

Chart 2 — **Scan Time**: Uncompressed, OLTP, Hybrid, Pure Column

Chart 3 — **Row Lookup Time**

- 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 – Over 90% of full columnar benefits
  - Good single row lookup – no full columnar "cliff"

ORACLE

Real World EHCC storage savings on customer datasets

- 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