ORACLE

25. hroug
godišnja konferencija

Essential Concepts for Data Lifecycle Aware Compression

Database Storage Optimization

Gregg Christman
Oracle Product Management
Core Database Product Development

October 2021

Why is Data Lifecycle Aware Compression Important?

- **Exponential increases in data volumes are putting enterprise IT infrastructures under severe pressure**
  - Including: storage costs, performance, scalability and manageability
- **Regulatory requirements are changing how and why data is being retained**
  - Organizations are now required to retain and control much more information for much longer periods – often for 7-10 years
- **Many data lifecycle solutions often have no knowledge of the use, or value, of the data under Oracle Database management**
  - Making these “database-unaware” technologies virtually useless

Copyright © 2021, Oracle and/or its affiliates | Confidential: Public

Database Compression Levels

Copyright © 2021, Oracle and/or its affiliates | Confidential: Public

Compression Included With Oracle Advanced Compression

- **Advanced Row Compression**

  Data Compression

  - Optimized for both OLTP and Data Warehouse
  - No downtime required
  - No SQL or application changes
  - Maintains compression transparently
  - 50% (or more) data storage reduction typical

- **Advanced Index Compression**

  Index Compression

  - One of the most overlooked features
  - Optimized for both OLTP and Data Warehouse
  - No downtime required
  - No SQL or application changes
  - 50% (or more) data storage reduction typical

- **RMAN Backup Compression**

  Backup Compression

  - Data/indexes already compressed remain compressed during backup/recovery
  - Three levels of RMAN backup compression – Low, Medium and High
  - Backup data is compressed before it is written to disk or tape and doesn't need to be uncompressed before recovery

- **Advanced LOB Compression**

  Unstructured Data Compression

  - Works with Oracle Database SecureFiles feature
  - Maintains compression transparently
  - 50% (or more) data storage reduction typical

Copyright © 2021, Oracle and/or its affiliates | Confidential: Public

Compression Included With Oracle Database Enterprise Edition

- ## HCC Warehouse (Query) Compression

  ### Data Compression

  - For applications with no, or very limited DML operations
  - Optimized to increase scan query performance for query-mostly tables
  - Maximizes storage savings and query performance benefits
  - No SQL or application changes

- ## HCC Archive Compression

  ### Data Compression

  - For applications with no, or very limited DML operations
  - Optimized to maximize storage savings
  - Intended for tables or partitions that store cold historic/archive data that is rarely accessed
  - No SQL or application changes

Requires: Exadata, SuperCluster, Pillar Axiom, ZFSSA storage, FS1 or Oracle Database Appliance (ODA)

- ## Prefix Key Compression

  ### Index Compression

  - Optimized for both OLTP and Data Warehouse
  - No downtime required
  - No SQL or application changes
  - User must run Analyze Index to determine which columns to include
  - 50% (or more) data storage reduction typical
  - Requires DBEE

- ## Basic Table Compression

  ### Data Compression

  - Optimized for Data Warehouse
  - No downtime required
  - No SQL or application changes
  - Does NOT compress DML Insert or Updates
  - 50% (or more) data storage reduction typical
  - Requires DBEE

Copyright © 2021, Oracle and/or its affiliates | Confidential: Public

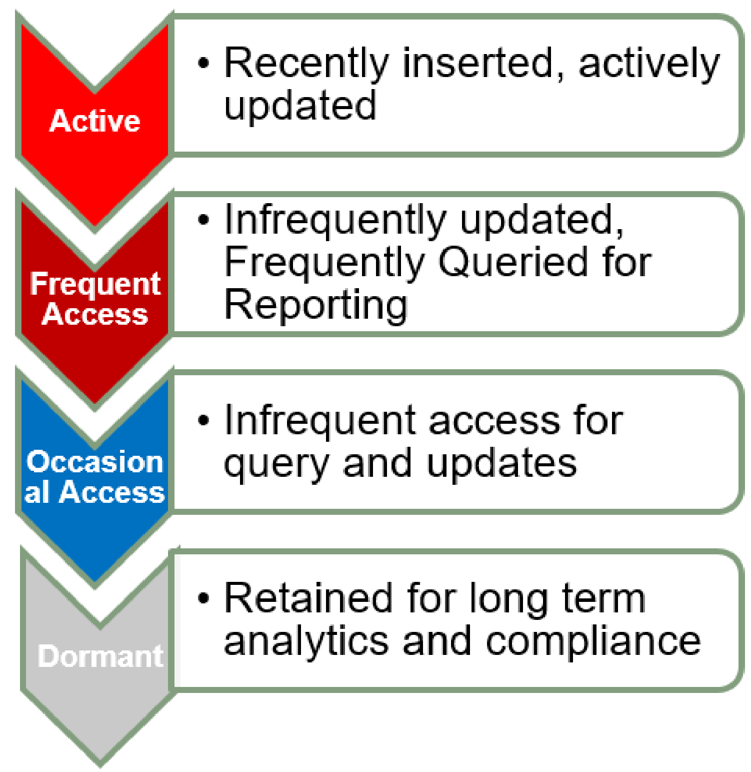
Tracking Data Usage

Copyright © 2021, Oracle and/or its affiliates | Confidential: Public

Heat Map

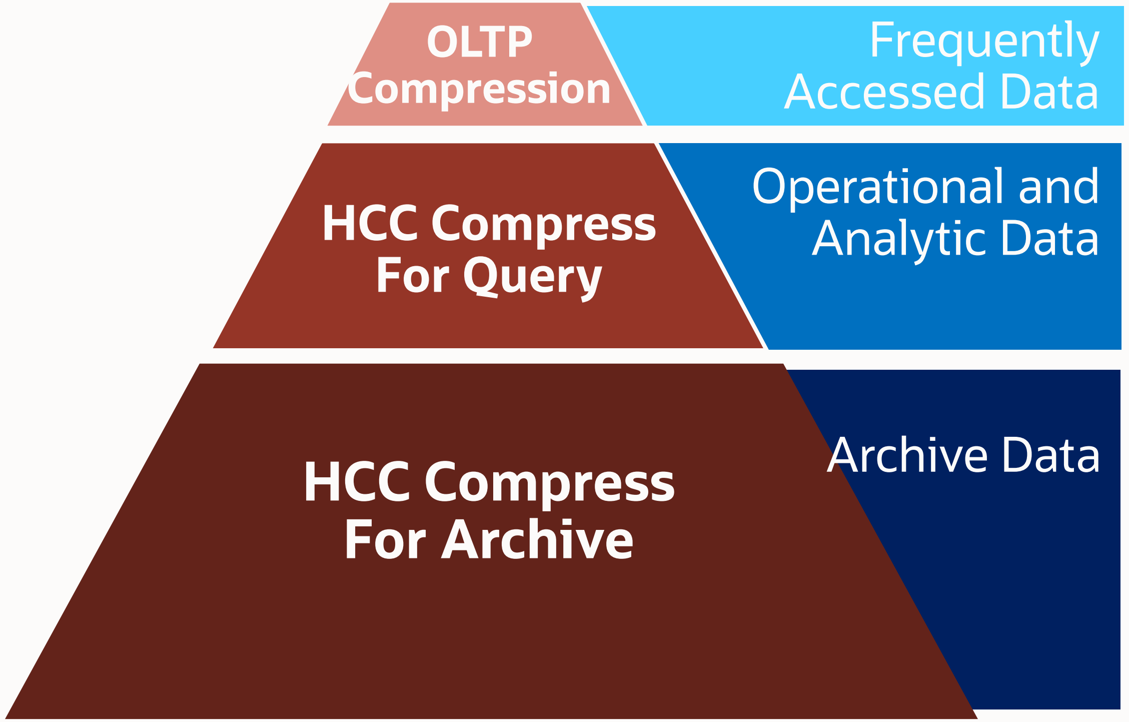
Tracking data usage in a database. enables organizations to understand…

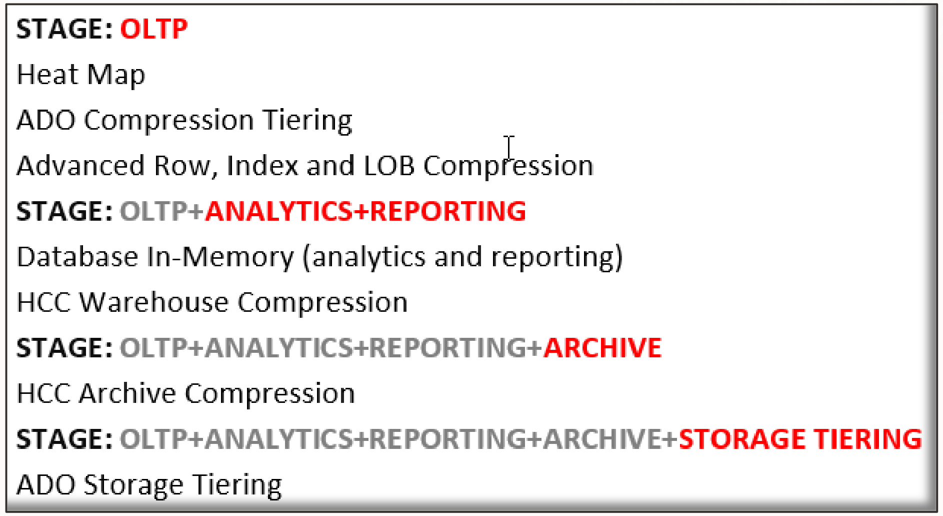
- Where tables/partitions are in regards to their data lifecycle (active, in-active or historic)
- How data is accessed (for queries and/or modification)
- How access patterns change over time
- The granularity of the database object (segment vs. row)


Active
• Recently inserted, actively updated
Frequent Access
• Infrequently updated, Frequently Queried for Reporting
Occasion al Access
• Infrequent access for query and updates
Dormant
• Retained for long term analytics and compliance

Copyright © 2021, Oracle and/or its affiliates | Confidential: Public

Data Lifecycle Aware Compression




STAGE: OLTP

Heat Map

ADO Compression Tiering

Advanced Row, Index and LOB Compression

STAGE: OLTP+ANALYTICS+REPORTING

Database In-Memory (analytics and reporting)

HCC Warehouse Compression

STAGE: OLTP+ANALYTICS+REPORTING+ARCHIVE

HCC Archive Compression

STAGE: OLTP+ANALYTICS+REPORTING+ARCHIVE+STORAGE TIERING

ADO Storage Tiering

Copyright © 2021, Oracle and/or its affiliates | Confidential: Public

Frequently Accessed Data (OLTP)

Data Lifecycle Usage

OLTP/Transactional

Data Compression (Advanced Row Compression)

- **Compression specifically designed to work with OLTP/DW applications**
  - Everything is faster: table scans, backups, database cloning, etc.
  - Buffer cache becomes more efficient by storing more data without having to add memory
  - Data remains compressed in memory
- **Compression during all types of data manipulation operations, including conventional DML such as INSERT and UPDATE**
  - The compression ratio achieved depends on the data being compressed, specifically the cardinality of the data
  - ***Customer Experience: 2x-4x compression ratios***

Copyright © 2021, Oracle and/or its affiliates | Confidential: Public

Data Lifecycle Usage

OLTP/Transactional

Index Compression (Advanced Index Compression)

Low Index Compression

- **The correct and most optimal numbers of prefix columns are computed automatically to produce the best compression ratio**
  - Possible to have different index leaf blocks compressed with different prefix column count or not be compressed at all, if there are no repeating prefixes
  - Customer Experience: 2x-3x compression ratios

High Index Compression

- **Utilizes additional complex compression algorithms on a potentially larger number of index keys to achieve higher levels of compression**
  - Customer Experience: 4x-5x, highly compressible indexes 15x-20x

Copyright © 2021, Oracle and/or its affiliates | Confidential: Public

Data Lifecycle Usage

OLTP/Transactional

Prefix Compression (Index Key Compression)

- **Eliminates duplicate copies of pre-defined number of index prefix columns at the index leaf block level**
  - Effective way to permanently reduce the index size, both on disk and in cache
- **The number of prefix columns to consider for compression is specified by the DBA at the index create time (or rebuild time) and is constant for all index leaf blocks**
  - Compression can be very beneficial when the prefix columns of an index have many repeated rows within a leaf block
  - ***ANALYZE INDEX*** will give advice on whether / how many columns to choose
  - **Customers experience: 2x compression ratios**

Copyright © 2021, Oracle and/or its affiliates | Confidential: Public

Data Lifecycle Usage

OLTP/Transactional

Unstructured Data Compression (Advanced LOB Compression)

- **Detects if SecureFiles LOB data is compressible and will compress using industry standard compression algorithms**
  - If the compression does not yield any savings, or if the data is already compressed, SecureFiles will turn off compression for such LOBs
- **Random access reads and writes to compressed SecureFiles LOBs are achieved without the need to decompress the entire file**
  - Only the sub portion of the compressed file needs to be decompressed thus saving CPU and I/O
- **Setting data or index compression does not affect SecureFiles LOB compression or vice versa**

Copyright © 2021, Oracle and/or its affiliates | Confidential: Public

Analytics and Reporting Data

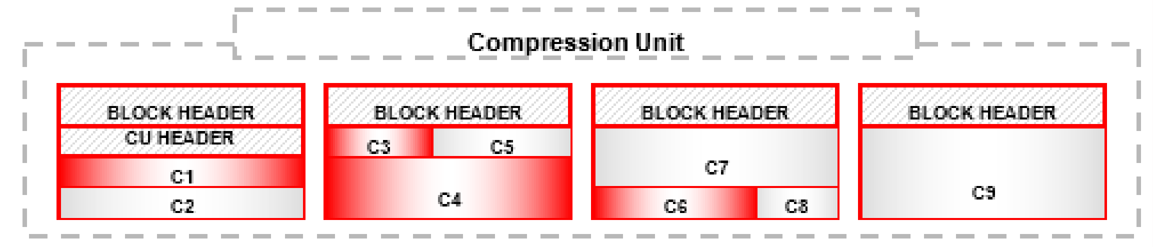
Copyright © 2021, Oracle and/or its affiliates | Confidential: Public

Data Lifecycle Usage

Analytics/Reporting

Data Compression (Hybrid Columnar Compression – Query Level Compression)

- **Optimized to increase scan query performance for query-mostly tables**
  - Maximizes storage savings and query performance benefits
- **Tables are organized into Compression Units (CUs) – CU's comprised of multiple database blocks**
  - Within Compression Unit, data is organized by column instead of by row.
  - Column organization brings similar values close together, enhancing compression – 6x to 10x compression ratio typical



Copyright © 2021, Oracle and/or its affiliates | Confidential: Public

Archive and Historic Data

Copyright © 2021, Oracle and/or its affiliates | Confidential: Public

Data Lifecycle Usage

Archive/Historic

Data Compression (Hybrid Columnar Compression – Archive Level Compression)

- Optimized to maximize storage savings, typically achieving a compression ratio of 10x to 15x
- In contrast to Warehouse Compression, Archive Compression is a pure storage saving technology
  - Intended for tables or partitions that store cold historic/archive data that is rarely accessed
  - No need to move data to tape - data is always online and always accessible
- Tables also organized into Compression Units (CUs) – CU's comprised of multiple database blocks

Copyright © 2021, Oracle and/or its affiliates | Confidential: Public

Data Lifecycle Management with ADO Compression Tiering and Storage Tiering

Copyright © 2021, Oracle and/or its affiliates | Confidential: Public

Partitioning, Heat Map and Automatic Data Optimization

Data Lifecycle Management Key Features

- Oracle Partitioning
  - Powerful functionality that allows tables, indexes and index-organized tables to be subdivided into smaller pieces
  - Multiple partitioning strategies allow aligning data subdivision with business requirements and current data usage
  - Decreases storages costs by aligning compression level to current partition usage
  - Requires no changes to applications and queries

- Oracle Heat Map
  - Data and index access usage tracking
  - Determines where tables/partitions are in regards to their data lifecycle (active, in-active or historic)
  - Tracks how data is accessed (queried/modified)
  - Provides information regarding how access patterns change over time
  - Automatically excludes maintenance tasks

- Oracle Automatic Data Optimization
  - Enables organizations to create policies for data compression – transparently change to most optimal compression level
  - Enables organizations to create policies for data movement – enables cold tables to be moved to lower-cost storage
  - All operations are executed automatically and in the background -- no user intervention or application changes required

Copyright © 2021, Oracle and/or its affiliates | Confidential: Public

ADO Compression Tiering and Storage Tiering

Data Lifecycle Management

- Automate Compression Tiering
  - Enables organizations to create policies for data compression – allows tables to transparently switch from row to columnar compression
    - Transparently change to most optimal compression level
  - Oracle Database evaluates policies during the DBA-defined database maintenance window, and uses the information collected by Heat Map to determine which policies to execute
  - All operations are executed automatically and in the background -- no user intervention or application changes required

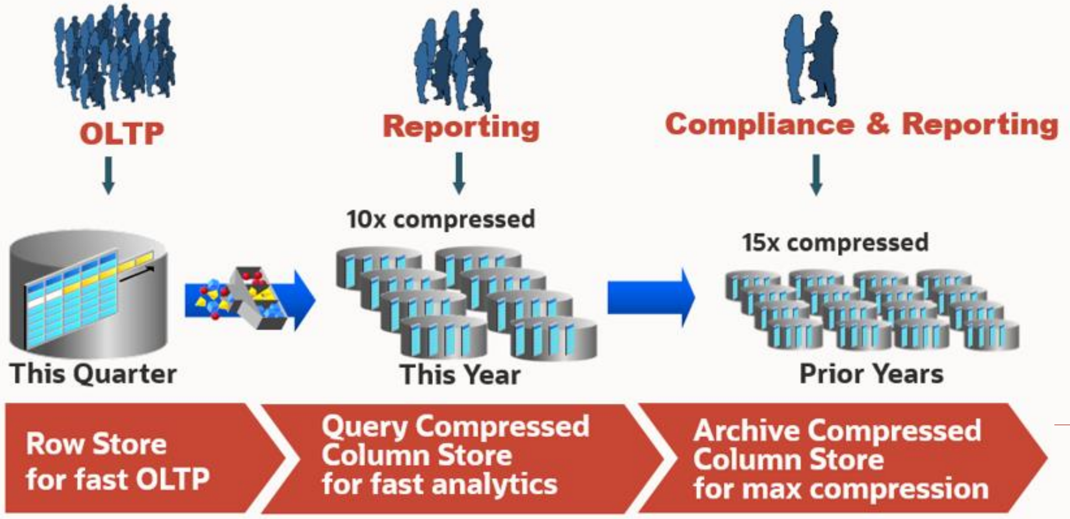
- Automate Storage Tiering
  - Enables organizations to create policies for data movement – enables cold tables to be moved to lower-cost storage
    - Automatically move to cost effective tier 2 storage based upon tier 1 storage pressure
    - Compress and move data in same operation
  - Oracle Database monitors tier 1 storage space availability and automatically moves “cold” tables to lower cost tier 2 storage
  - All operations are executed automatically and in the background -- no user intervention or application changes required

Copyright © 2021, Oracle and/or its affiliates | Confidential: Public

Automatic Data Optimization Best Practice Workflow

Data Lifecycle Management

ADO automatically converts *COLD* tables/partitions into HCC Archive Compressed once data cools down further and is no longer frequently queried (10x – 15x compression)



ADO automatically converts *WARM* tables/partitions into HCC Query High Compression once the modifications cool down, and is used mainly for reporting (6x – 10x compression)

ADO compresses *HOT* tables/partitions with Advanced Row Compression (2x-4x compression)

Copyright © 2021, Oracle and/or its affiliates | Confidential: Public

ADO Compression Tiering Example – Advanced Row Compression

Data Lifecycle Management

In this example, a segment-level ADO policy is created to automatically compress a table using **Advanced Row Compression** after there have been **no modifications for 30 days**

```
ALTER TABLE orders ILM ADD POLICY
ROW STORE COMPRESS ADVANCED SEGMENT
AFTER 30 DAYS OF NO MODIFICATION;
```

ADO policy specifies Advanced Row Compression
– Can specify segment or row level compression

ADO policies conditions include: *no access*, or ***no modification*** or ***creation time***
and *when* the policy will take effect – for example, after "*n*" days or months or years

Copyright © 2021, Oracle and/or its affiliates | Confidential: Public

ADO Compression Tiering Example – Hybrid Columnar Compression

Data Lifecycle Management

In this example, a segment-level ADO policy is created to automatically compress the same table using **Hybrid Columnar Query High Compression** after there have been **no modifications for 90 days**

```
ALTER TABLE orders ILM ADD POLICY
COLUMN STORE COMPRESS FOR QUERY HIGH SEGMENT
AFTER 90 DAYS OF NO MODIFICATION;
```

In this example, a segment-level ADO policy is created to automatically compress the same table using **Hybrid Columnar Archive High Compression** after there have been **no modifications for 180 days**

```
ALTER TABLE orders ILM ADD POLICY
COLUMN STORE COMPRESS FOR ARCHIVE HIGH SEGMENT
AFTER 180 DAYS OF NO MODIFICATION;
```

Automatic Index Optimization

Data Lifecycle Management

Compression and optimization for indexes using existing Automatic Data Optimization (ADO) framework

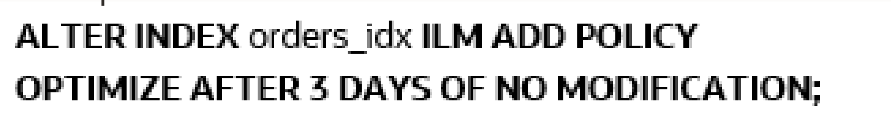
- Existing Heat Map capability collects activity statistics on the index
- Database automatically chooses best way to "optimize" index

Index optimizations include:

- **Compress:** Compresses portions of the key values in an index segment. (3x compression ratio typical)
- **Coalesce:** Merges the contents of index blocks where possible to free blocks for reuse
- **Rebuild:** Rebuilds index to improve space usage and access speed

Automates movement of indexes to tier 2 storage when tier 1 storage under space pressure

Example



```
ALTER INDEX orders_idx ILM ADD POLICY
OPTIMIZE AFTER 3 DAYS OF NO MODIFICATION;
```

Copyright © 2021, Oracle and/or its affiliates | Confidential: Public

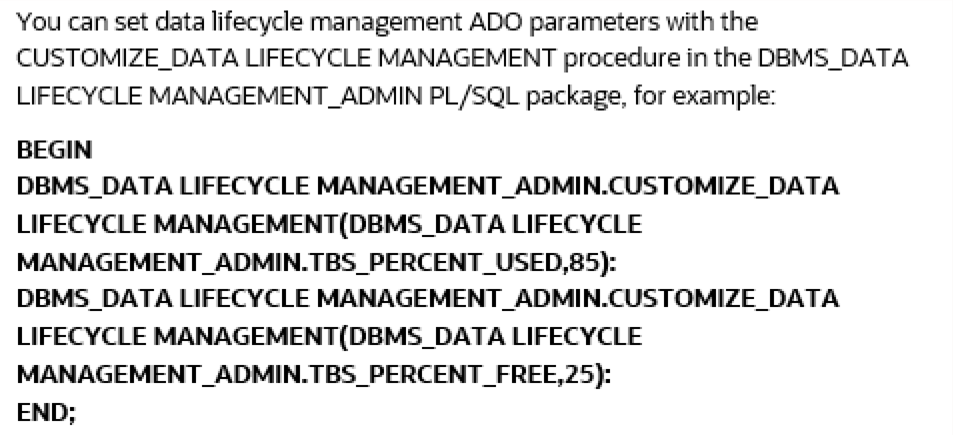
ADO Storage Tiering Example

Data Lifecycle Management

In this example, a tier-to ADO policy is created to automatically move the table to lower cost tier two storage upon tier one storage pressure

Must be an Oracle Tablespace

```
ALTER TABLE orders ILM ADD POLICY TIER TO lowercosttbs
```



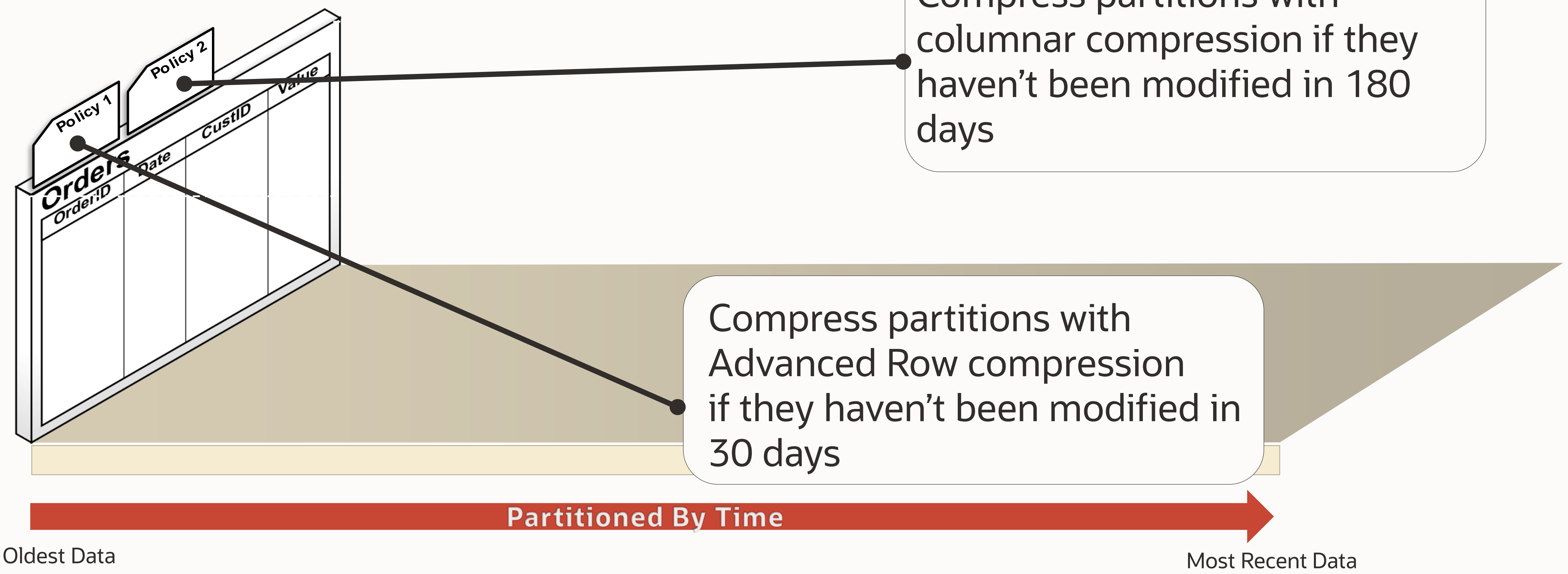
You can set data lifecycle management ADO parameters with the CUSTOMIZE_DATA LIFECYCLE MANAGEMENT procedure in the DBMS_DATA LIFECYCLE MANAGEMENT_ADMIN PL/SQL package, for example:

```
BEGIN
DBMS_DATA LIFECYCLE MANAGEMENT_ADMIN.CUSTOMIZE_DATA
LIFECYCLE MANAGEMENT(DBMS_DATA LIFECYCLE
MANAGEMENT_ADMIN.TBS_PERCENT_USED,85):
DBMS_DATA LIFECYCLE MANAGEMENT_ADMIN.CUSTOMIZE_DATA
LIFECYCLE MANAGEMENT(DBMS_DATA LIFECYCLE
MANAGEMENT_ADMIN.TBS_PERCENT_FREE,25):
END;
```

Usage Example

Create Automatic Data Optimization Compression Policies

Step-by-Step Usage Example



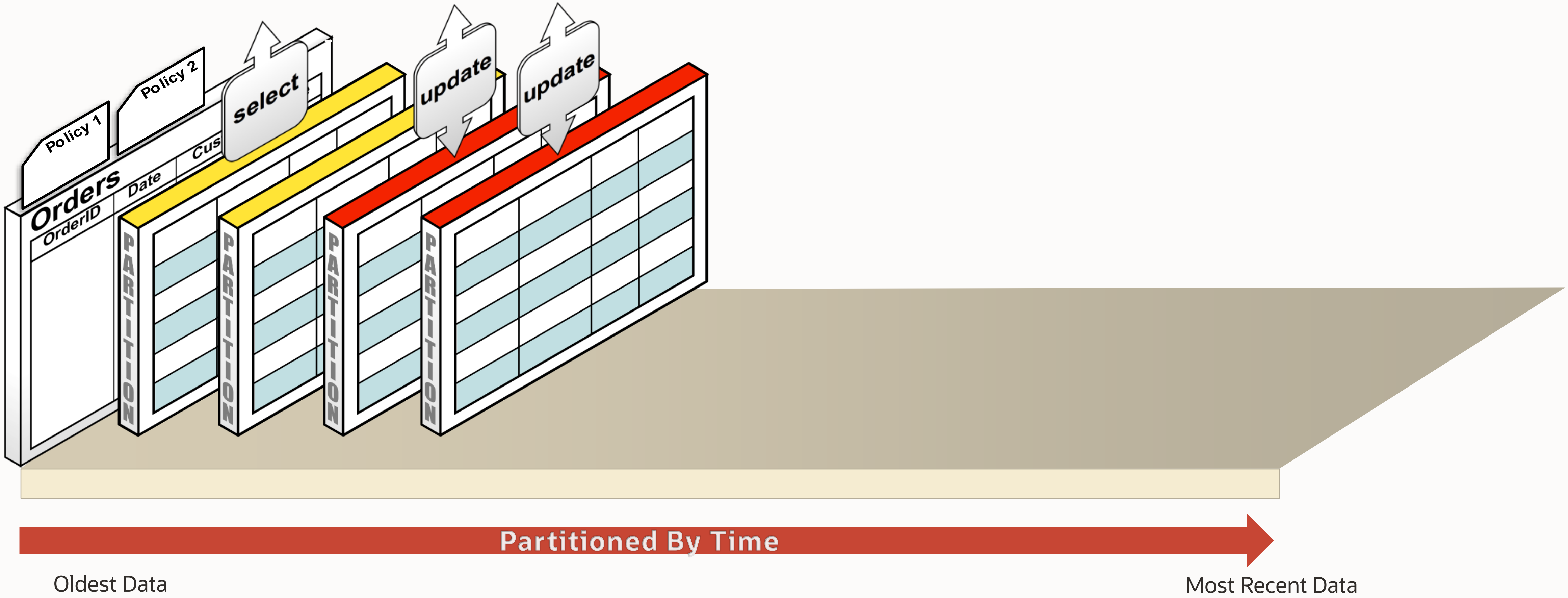
Compress partitions with columnar compression if they haven’t been modified in 180 days

Compress partitions with Advanced Row compression if they haven’t been modified in 30 days

Copyright © 2021, Oracle and/or its affiliates | Confidential: Public

Heat Map Automatically Tracks Data Usage

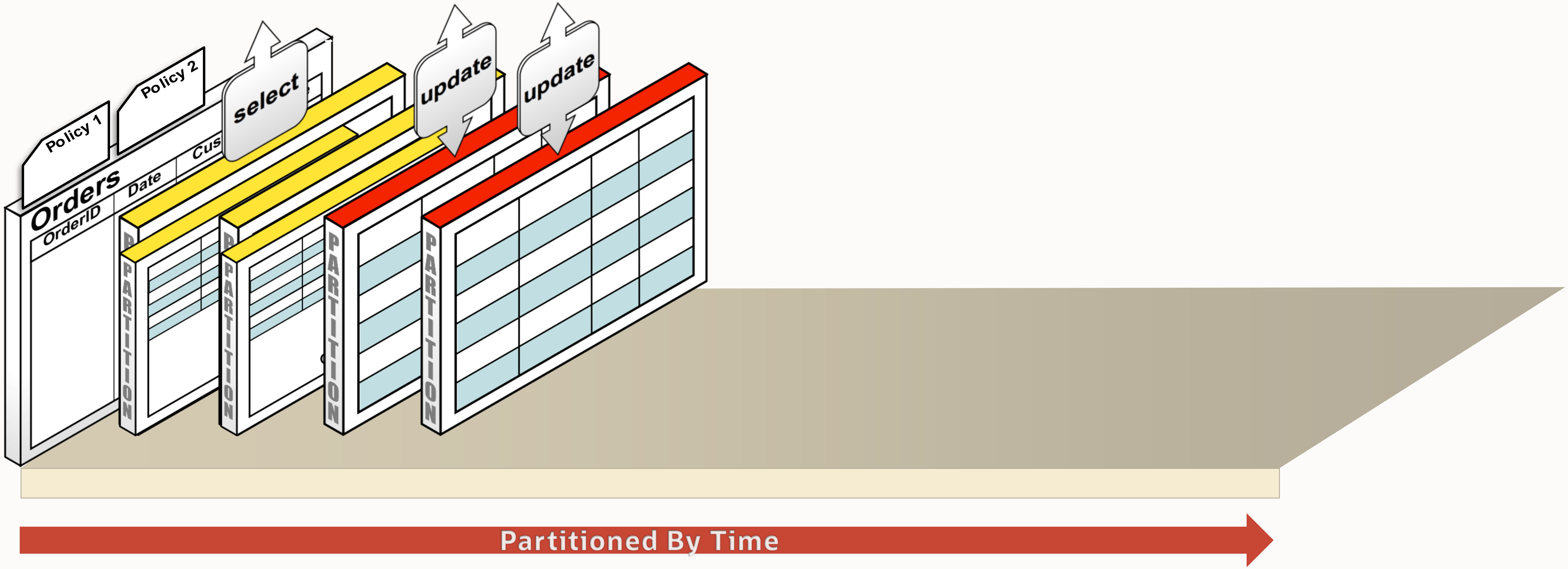
Step-by-Step Usage Example



Copyright © 2021, Oracle and/or its affiliates | Confidential: Public

Policies Automatically Applied to Tables and Partitions

Step-by-Step Usage Example



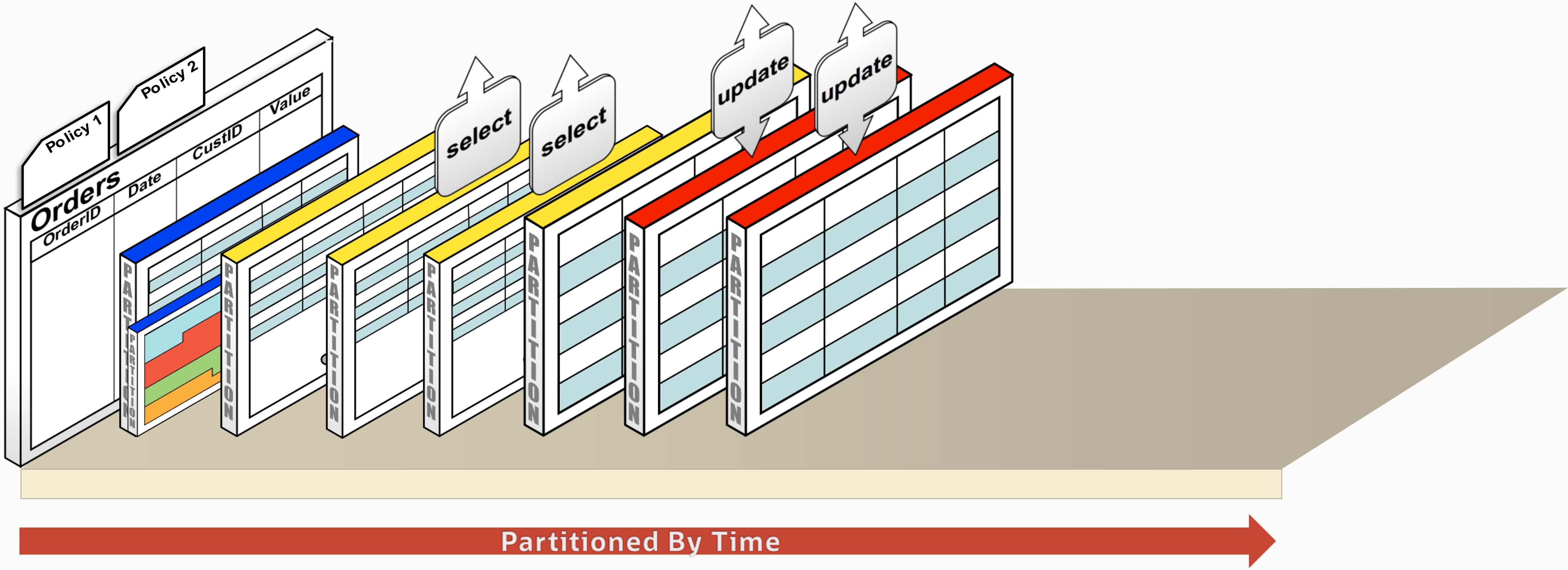
Oldest Data

Most Recent Data

Copyright © 2021, Oracle and/or its affiliates | Confidential: Public

Policies Automatically Applied to Tables and Partitions

Step-by-Step Usage Example



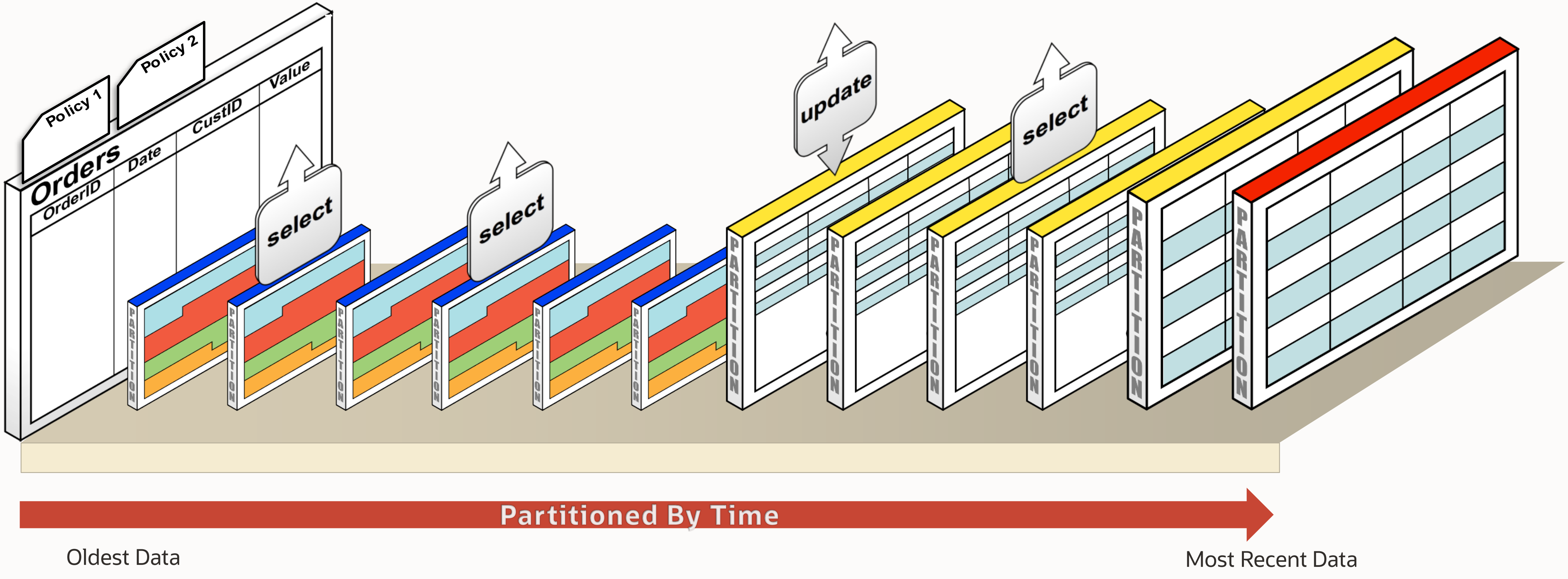
Oldest Data

Most Recent Data

Copyright © 2021, Oracle and/or its affiliates | Confidential: Public

Compression Optimized for Current Data Usage

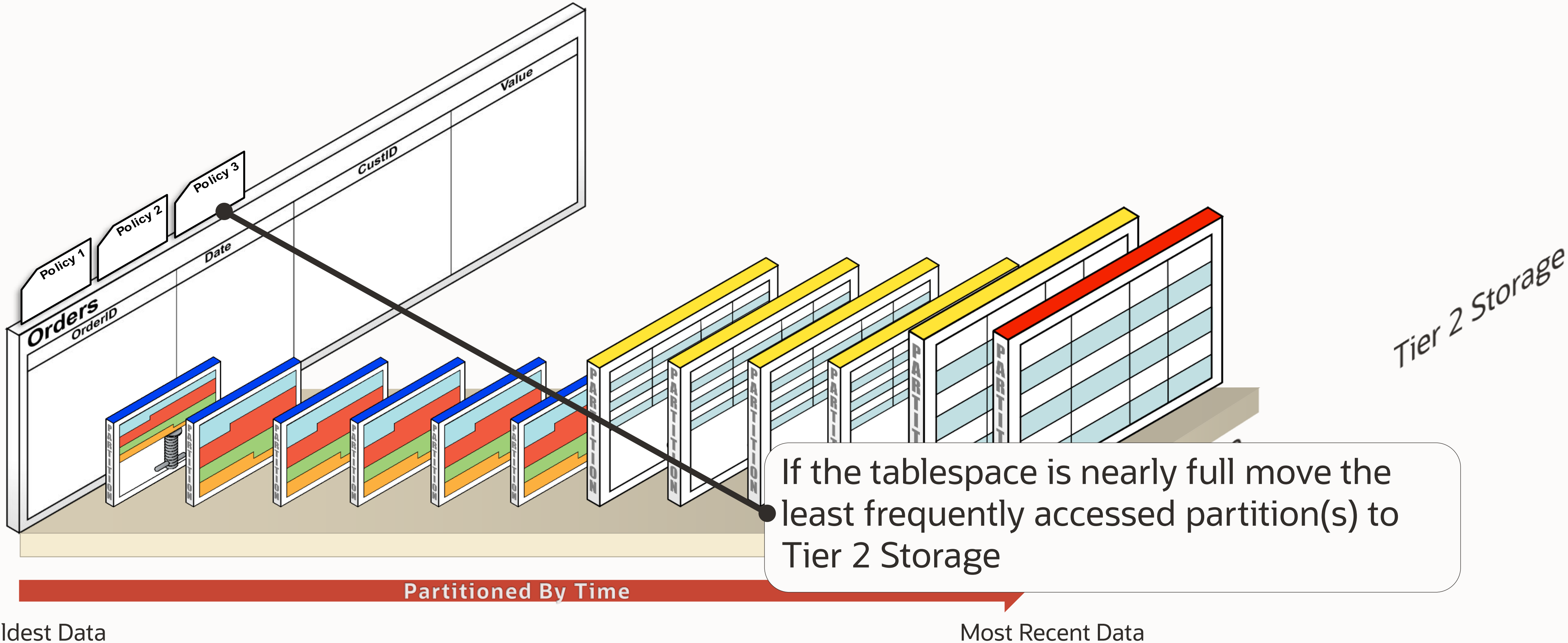
Step-by-Step Usage Example



Copyright © 2021, Oracle and/or its affiliates | Confidential: Public

Automatically Move Cold Tables to Low Cost Storage

Step-by-Step Usage Example



If the tablespace is nearly full move the least frequently accessed partition(s) to Tier 2 Storage

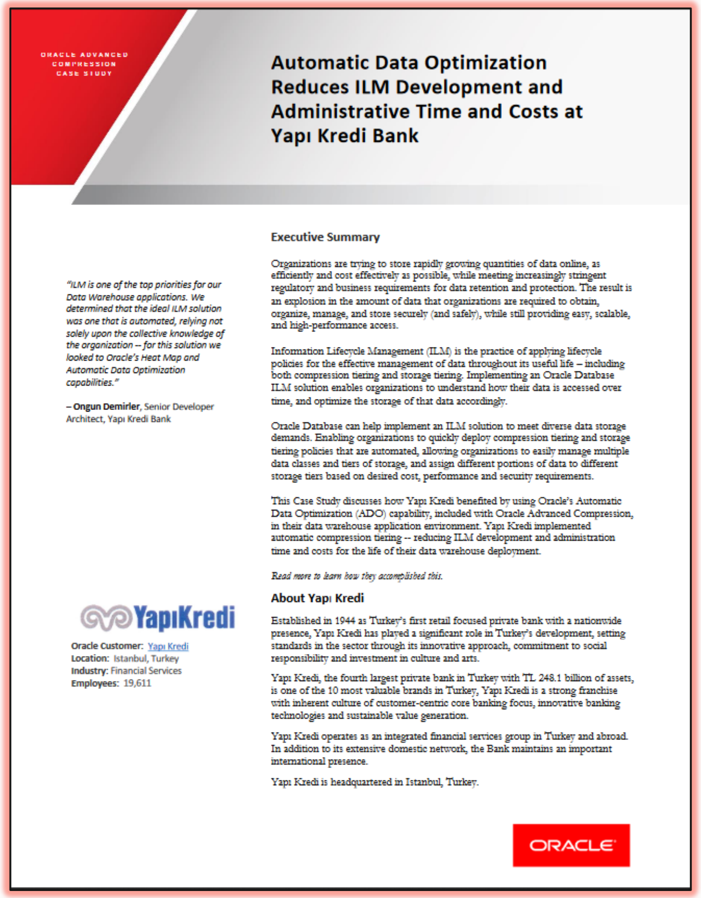
Copyright © 2021, Oracle and/or its affiliates | Confidential: Public

More Information

Copyright © 2021, Oracle and/or its affiliates | Confidential: Public

Customer Case Study

Data Lifecycle Management



ORACLE ADVANCED COMPRESSION CASE STUDY

Automatic Data Optimization Reduces ILM Development and Administrative Time and Costs at Yapı Kredi Bank

> *"ILM is one of the top priorities for our Data Warehouse applications. We determined that the ideal ILM solution was one that is automated, relying not solely upon the collective knowledge of the organization -- for this solution we looked to Oracle's Heat Map and Automatic Data Optimization capabilities."*
>
> – **Ongun Demirler**, Senior Developer Architect, Yapı Kredi Bank

Executive Summary

Organizations are trying to store rapidly growing quantities of data online, as efficiently and cost effectively as possible, while meeting increasingly stringent regulatory and business requirements for data retention and protection. The result is an explosion in the amount of data that organizations are required to obtain, organize, manage, and store securely (and safely), while still providing easy, scalable, and high-performance access.

Information Lifecycle Management (ILM) is the practice of applying lifecycle policies for the effective management of data throughout its useful life – including both compression tiering and storage tiering. Implementing an Oracle Database ILM solution enables organizations to understand how their data is accessed over time, and optimize the storage of that data accordingly.

Oracle Database can help implement an ILM solution to meet diverse data storage demands. Enabling organizations to quickly deploy compression tiering and storage tiering policies that are automated, allowing organizations to easily manage multiple data classes and tiers of storage, and assign different portions of data to different storage tiers based on desired cost, perfomance and security requirements.

This Case Study discusses how Yapı Kredi benefited by using Oracle's Automatic Data Optimization (ADO) capability, included with Oracle Advanced Compression, in their data warehouse application environment. Yapı Kredi implemented automatic compression tiering -- reducing ILM development and administration time and costs for the life of their data warehouse deployment.

Read more to learn how they accomplished this.

About Yapı Kredi

Established in 1944 as Turkey's first retail focused private bank with a nationwide presence, Yapı Kredi has played a significant role in Turkey's development, setting standards in the sector through its innovative approach, commitment to social responsibility and investment in culture and arts.

Yapı Kredi, the fourth largest private bank in Turkey with TL 248.1 billion of assets, is one of the 10 most valuable brands in Turkey, Yapı Kredi is a strong franchise with inherent culture of customer-centric core banking focus, innovative banking technologies and sustainable value generation.

Yapı Kredi operates as an integrated financial services group in Turkey and abroad. In addition to its extensive domestic network, the Bank maintains an important international presence.

Yapı Kredi is headquartered in Istanbul, Turkey.

YapıKredi

Oracle Customer: Yapı Kredi
Location: Istanbul, Turkey
Industry: Financial Services
Employees: 19,611

ORACLE

YapıKredi

Oracle Customer: Yapı Kredi
Location: Istanbul, Turkey
Industry: Financial Services
Employees: 19,611

"ILM is one of the top priorities for our Data Warehouse applications. We determined that the ideal ILM solution was one that is automated, relying not solely upon the collective knowledge of the organization -- for this solution we looked to Oracle's Heat Map and Automatic Data Optimization capabilities."

– **Ongun Demirler**, Senior Developer Architect, Yapı Kredi Bank

"Heat Map and Automatic Data Optimization clearly met our expectations. It enables us at Yapı Kredi to save development and administration time and costs for the life of the deployment. A win all around for us!"

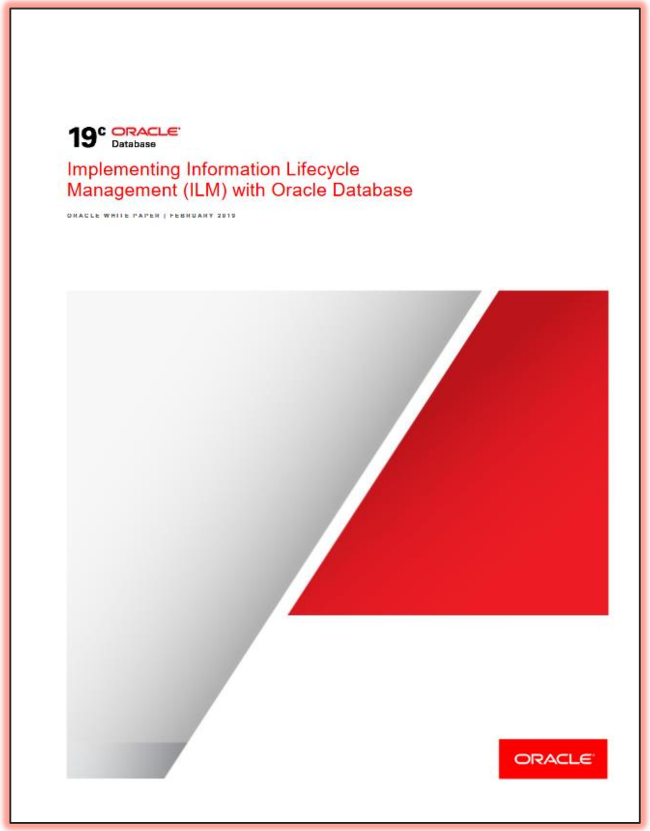
– **Oktay Elmas**, Senior Database Administrator, Yapı Kredi Bank

Copyright © 2021, Oracle and/or its affiliates | Confidential: Public

Additional Resources

Join the Conversation

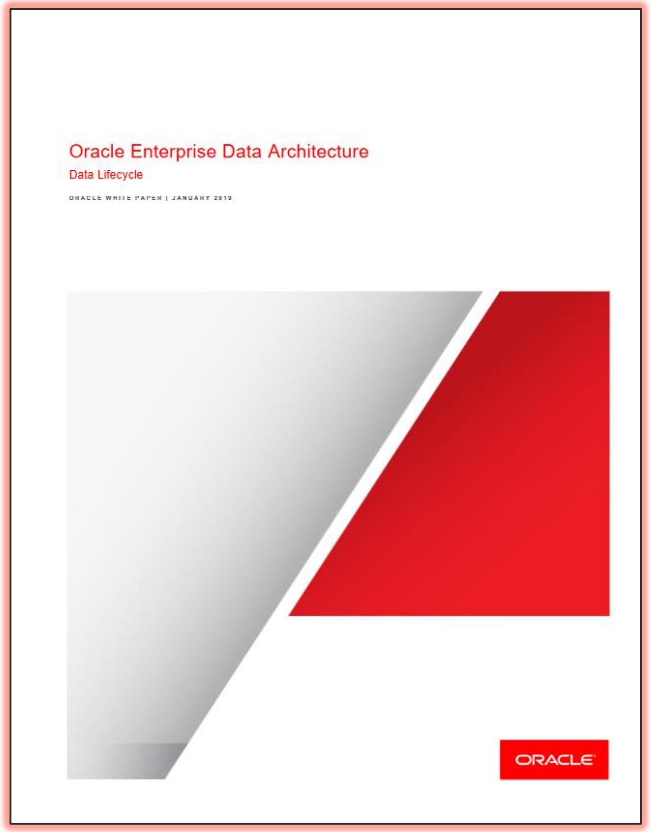
- https://twitter.com/aco_gregg
- https://blogs.oracle.com/DBStorage/
- http://www.oracle.com/database/advanced-compression/index.html



19c ORACLE Database

Implementing Information Lifecycle Management (ILM) with Oracle Database

ORACLE WHITE PAPER | FEBRUARY 2019



Oracle Enterprise Data Architecture

Data Lifecycle

ORACLE WHITE PAPER | JANUARY 2019

Copyright © 2021, Oracle and/or its affiliates | Confidential: Public