

# How Autonomous is the Oracle Autonomous Data Warehouse?

Christian Antognini / Dani Schnider



@chrisantognini



[antognini.ch/blog](https://antognini.ch/blog)



@dani\_schnider



[danischnider.wordpress.com](https://danischnider.wordpress.com)

BASLE ■ BERN ■ BRUGG ■ DÜSSELDORF ■ FRANKFURT A.M. ■ FREIBURG I.BR. ■ GENEVA  
HAMBURG ■ COPENHAGEN ■ LAUSANNE ■ MUNICH ■ STUTTGART ■ VIENNA ■ ZURICH

**trivadis**  
makes IT easier. ■ ■ ■

# ■ Agenda

1. Introduction
2. Set Up
3. Connectivity
4. Loading Data
5. ETL Performance
6. Query Performance
7. Monitoring Performance
8. Miscellaneous
9. Conclusion

# Introduction

# ■ Larry Ellison, ATP Announcement, 7th August 2018

“There is nothing to learn,  
there is nothing to do”



# ■ Autonomous Data Warehouse Cloud

## Automated Database Administration

«...you do not need to configure or manage any **hardware**, or install any **software**. Autonomous Data Warehouse handles creating the data warehouse, **backing up** the database, **patching and upgrading** the database, and **growing or shrinking** the database.»

## Automated Performance Tuning

«When you use Autonomous Data Warehouse, no tuning is necessary. You do not need to consider any details about **parallelism**, **partitioning**, **indexing**, or **compression**. The service automatically configures the database for high-performance queries.»

Source: Using Oracle Autonomous Data Warehouse, Chapter 1

<https://docs.oracle.com/en/cloud/paas/autonomous-data-warehouse-cloud/user/getting-started.html#GUID-4B91499D-7C2B-46D9-8E4D-A6ABF2093414>

# Set Up

# ■ Creating

Create Autonomous Data Warehouse [help](#) [cancel](#)

[Click here](#) to enable compartment selection for your Autonomous Data Warehouse.

DISPLAY NAME

DATABASE NAME

The name must contain only letters and numbers, starting with a letter. 14 characters max.

CPU CORE COUNT  STORAGE (TB)

The number of CPU cores to enable. Maximum cores per database: 128. Available cores are subject to your tenancy's service limits. The available storage, up to 128 TB.

**Administrator Credentials**

Set the password for your Autonomous Data Warehouse ADMIN user here.

USERNAME READ-ONLY  
ADMIN

PASSWORD

CONFIRM PASSWORD

LICENSE TYPE

MY ORGANIZATION ALREADY OWNS ORACLE DATABASE SOFTWARE LICENSES  
Bring my existing database software licenses to the database cloud service ([details](#)).

SUBSCRIBE TO NEW DATABASE SOFTWARE LICENSES AND THE DATABASE CLOUD SERVICE

TAGS

Tagging is a metadata system that allows you to organize and track resources within your tenancy. Tags are composed of keys and values that can be attached to resources.

[Learn more about tagging](#)

TAG NAMESPACE TAG KEY VALUE

None (apply a free-for ↕)

+ Additional Tag

Create Autonomous Data Warehouse

Name



Storage (1..128 TB)



Number of CPU (1..128)



PDB admin password



License type



Tags (optional)



# ■ Initialization Parameters That Can Be Modified

APPROX\_FOR\_AGGREGATION

PLSCOPE\_SETTINGS

APPROX\_FOR\_COUNT\_DISTINCT

PLSQL\_CCFLAGS

APPROX\_FOR\_PERCENTILE

PLSQL\_DEBUG

AWR\_PDB\_AUTOFLUSH\_ENABLED<sup>1</sup>

PLSQL\_OPTIMIZE\_LEVEL

OPTIMIZER\_CAPTURE\_SQL\_PLAN\_BASELINES<sup>2</sup>

PLSQL\_WARNINGS

OPTIMIZER\_IGNORE\_HINTS

TIME\_ZONE<sup>2</sup>

OPTIMIZER\_IGNORE\_PARALLEL\_HINTS

NLS\_\*

<sup>1</sup>System level only <sup>2</sup>Session level only



# ■ Resource Manager

A plan (DWCS\_PLAN) with three consumer groups is pre-configured

Consumer Group	CPU	Session Pool	PX Server Limit	DOP Limit
HIGH	4	Unlimited	50	CPU_COUNT <sup>1</sup>
MEDIUM	2	Unlimited	84	4
LOW	1	2*CPU_COUNT <sup>1</sup>		1
OTHER_GROUPS	1	Unlimited		1

<sup>1</sup>When several instances are used, CPU\_COUNT <> “CPU Core Count”

# Connectivity

# Service Management via Browser

Used for

- Service Console
- Start/Stop DB
- Scale up/down
- Restore
- Management of credential

The screenshot displays the Oracle Cloud Infrastructure console interface. The browser address bar shows the URL: <https://console.eu-frankfurt-1.oraclecloud.com/a/db/adw/ocid1.autonomousdwarehouse.oc1-eu-frankfurt-1.abthejszv2qzboj5qtsl3an4kzt7...>. The page title is "Autonomous Data Warehouses » Autonomous Data Warehouse Details".

The main content area features a large green square with the text "ADW" and the status "AVAILABLE" below it. To the right, the title "TPC-DS" is displayed with a "Help" link. Below the title are several action buttons: "DB Connection", "Service Console", "Scale Up/Down", "Stop", and "Actions".

The "Autonomous Data Warehouse Information" section is expanded, showing the following details:

Property	Value
Display Name	TPC-DS
Database Name	TPCDS
Database Version	18.4.0.0
CPU Core Count	4
Storage (TB)	2
Created	Wed, 18 Jul 2018 16:04:17 GMT
Compartment	oc2018 (root)/BDS
OCID	...bedjta <a href="#">Show</a> <a href="#">Copy</a>
License Type	License Included
Lifecycle State	Available

Below the information section, there are "Resources" and "Backups" sections. The "Backups" section indicates that backups are automatically created daily and includes a "Create Manual Backup" button.

At the bottom of the page, there are links for "Terms of Use and Privacy" and "Cookie Preferences", and a copyright notice: "Copyright © 2019, Oracle and/or its affiliates. All rights reserved."

# ■ Connection with SQL Developer

Download client credential file

■ Via Service Console

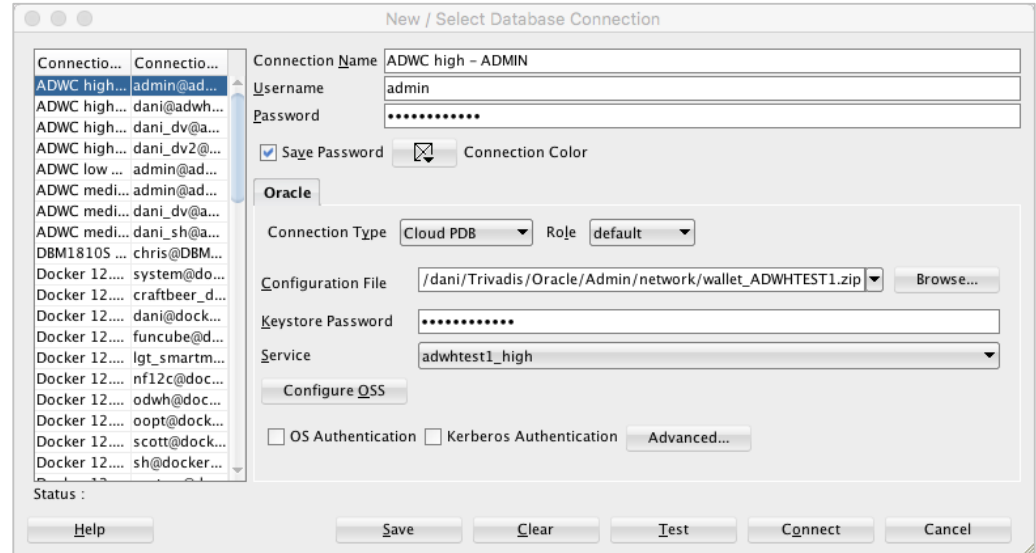
Create new connection

■ Type “Cloud PDB”

■ Import credential file

■ Enter keystore password  
(only for versions < 18.2)

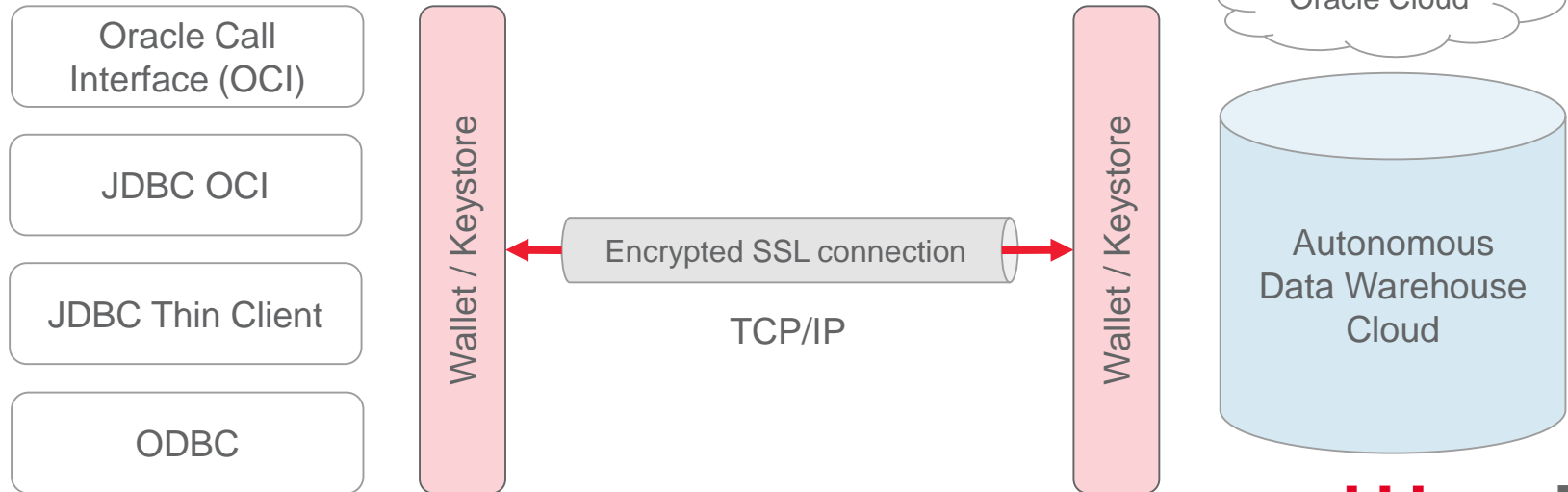
■ Select service level  
(LOW, MEDIUM, HIGH)



# Manual Connection Configuration

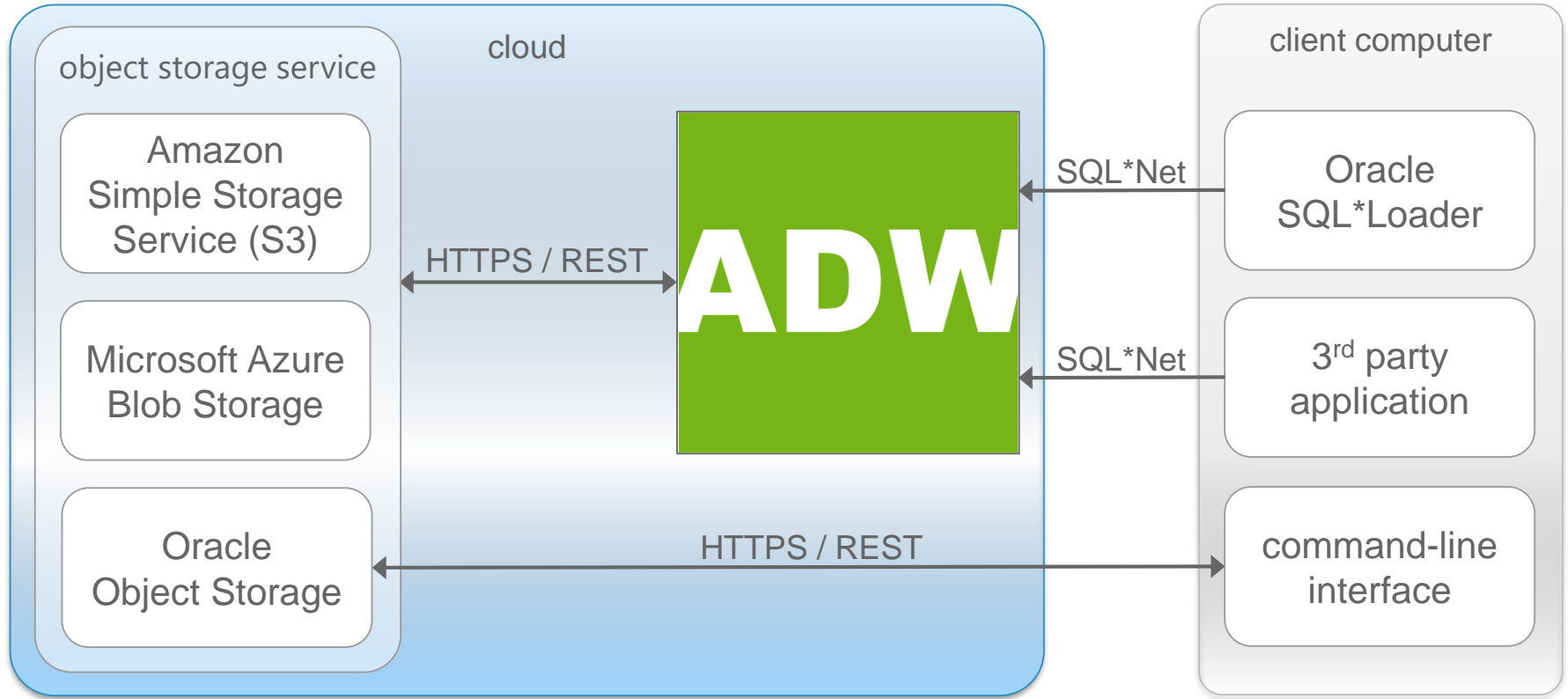
Connections using Oracle Net Services (SQL\*Net)

- For SQL\*Plus, SQLcl, Toad, ...
- ETL and BI tools, 3<sup>rd</sup> party tools



# Loading Data

# ■ Loading Data



# ■ DBMS\_CLOUD

It provides features to

- manage credentials to access an OSS
- manage external tables that can be used to query data stored in an OSS
- handle objects stored in an OSS
- handle files stored on ADW in DATA\_PUMP\_DIR

Data Pump 12.2 supports importing (but not exporting) a file stored in an OSS

```
ALTER DATABASE PROPERTY SET default_credential = 'ADMIN.CHRIS'
```

```
impdp dumpfile=default_credentials:https://.../test.dmp
```



# ■ Example – Loading Data into a 1TB TPC-DS Schema (1)

Step 1: Upload the text files from the file system of a server hosted in the cloud (same region) to the Oracle's object storage

```
oci os object bulk-upload --bucket-name tpcds
                           --src-dir /data/tpcds
                           --include "*.dat"
```

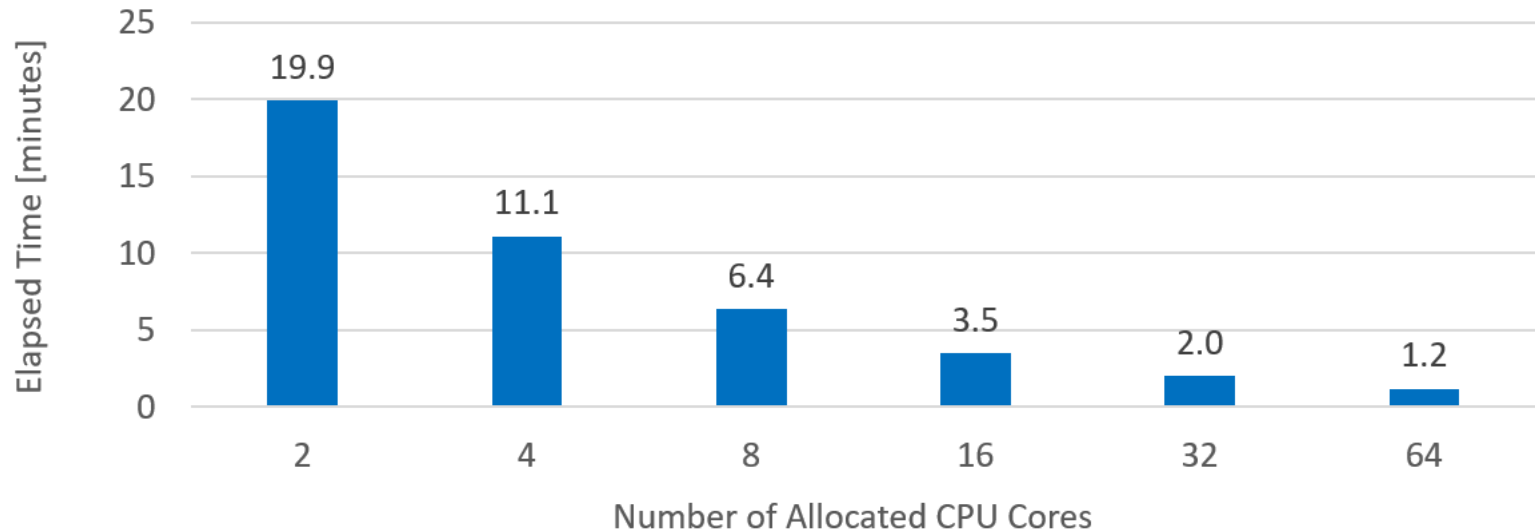
- By default, such a statement uses 10 parallel threads
- The load took 6 hours

Step 2: Load the data into the database through DBMS\_CLOUD

- The number of allocated CPU cores determines how fast data is loaded
- With 32 cores it took 67 minutes

## ■ Example – Loading Data into a 1TB TPC-DS Schema (2)

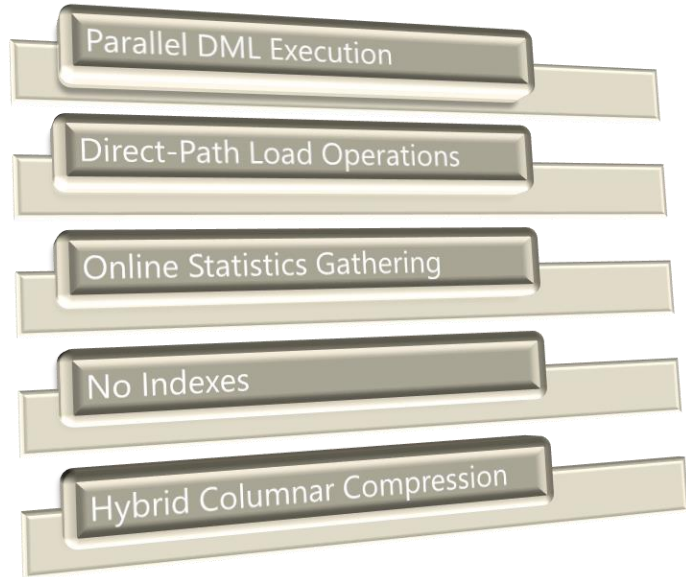
Example of scalability: load 22GB into the CATALOG\_RETURNS table



source: <https://antognini.ch/2018/07/observations-about-the-scalability-of-data-loads-in-adwc/>

# ETL Performance

# ■ Prepared for High ETL Performance



Good ETL performance is possible

- Setup is optimized for DWH and ETL
- But details must be considered

# ■ Parallel DML Execution and Direct-Path Operations

**Parallel DML** (PDML) is enabled by default

- Only if CPU core count > 1 and consumer group is *MEDIUM* or *HIGH*
- Hint */\*+ parallel \*/* can be added (usually not needed)
  - Set *optimizer\_ignore\_parallel\_hints = FALSE* (default: *TRUE*)

**Direct-Path INSERT** is used

- For parallel DML
- If hint */\*+ append \*/* is added

# Parallel DML Execution and Consumer Groups

Id	Operation	Name	TQ	IN-OUT	PQ Distrib
0	INSERT STATEMENT				
1	LOAD AS SELECT	CUSTOMERS_1			
2	PX COORDINATOR				
3	PX SEND QC (RANDOM)	:TQ10000	Q1,00	P->S	QC (RAND)
4	OPTIMIZER STATISTICS GATHERING		Q1,00	PCWC	
5	PX BLOCK ITERATOR		Q1,00	PCWC	
6	TABLE ACCESS STORAGE FULL	CUSTOMERS	Q1,00	PCWP	

HIGH

MEDIUM

- automatic DOP: Computed Degree of Parallelism is 8

Id	Operation	Name
0	INSERT STATEMENT	
1	LOAD TABLE CONVENTIONAL	CUSTOMERS_1
2	TABLE ACCESS STORAGE FULL	CUSTOMERS

LOW

- automatic DOP: Computed Degree of Parallelism is 1 because of no expensive parallel operation
- PDML disabled because object is not decorated with parallel clause
- Direct Load disabled because no append hint given and not executing in parallel

# Parallel DML / Direct-Path and Constraints

Restrictions must be considered:

- If FK constraints are defined, PDML / direct-path is disabled
- Conventional load is used

Recommendation:

- Define reliable constraints

```
-----  
| Id | Operation                                | Name          |  
-----  
|  0 | INSERT STATEMENT                        |               |  
|  1 | LOAD TABLE CONVENTIONAL              | CUSTOMERS_1  |  
|  2 | PX COORDINATOR                          |               |  
|  3 | PX SEND QC (RANDOM)                      | :TQ10000     |  
|  4 | PX BLOCK ITERATOR                       |               |  
|  5 | TABLE ACCESS STORAGE FULL              | CUSTOMERS    |  
-----
```

Note

- ```
-----  
- automatic DOP: Computed Degree of Parallelism is 8  
- PDML disabled because parent referential constraints  
  are present
```

```
ALTER TABLE customers_1  
ADD FOREIGN KEY (country_id) REFERENCES countries  
RELY DISABLE NOVALIDATE
```

# ■ Online Statistics Gathering for Direct-Path Loads

Statistics are gathered automatically

Unlike 12c, this works also

- for non-empty tables
- for histograms

| Id | Operation                             | Name     |
|----|---------------------------------------|----------|
| 0  | INSERT STATEMENT                      |          |
| 1  | LOAD AS SELECT                        | TARGET   |
| 2  | PX COORDINATOR                        |          |
| 3  | PX SEND QC (RANDOM)                   | :TQ10000 |
| 4  | <b>OPTIMIZER STATISTICS GATHERING</b> |          |
| 5  | PX BLOCK ITERATOR                     |          |
| 6  | TABLE ACCESS STORAGE FULL             | SOURCE   |

Two new undocumented parameters

- `_optimizer_gather_stats_on_load_all` (default: TRUE)
- `_optimizer_gather_stats_on_load_hist` (default: TRUE)



# ■ Statistics Gathering for Conventional Loads

**Attention: statistics are not gathered automatically**

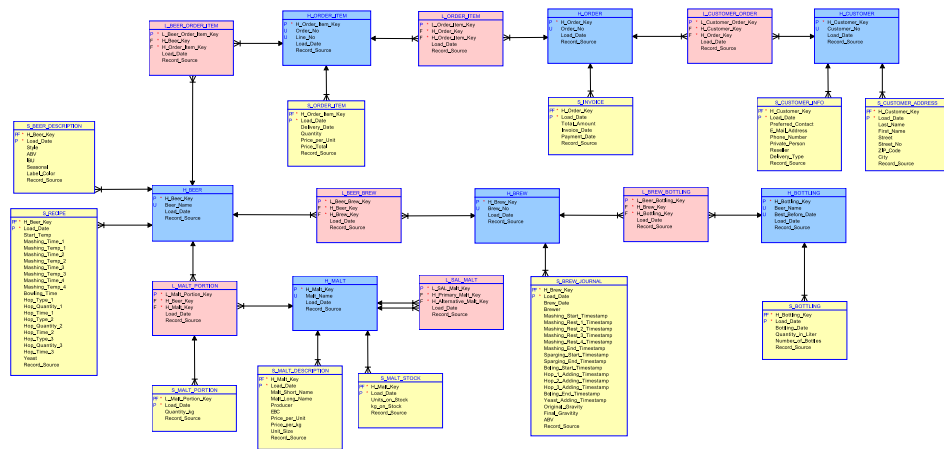
Call DBMS\_STATS with default values

Automatic Statistics Gathering job is enabled, but maintenance windows are disabled

# Example: Loading Data Vault Schema

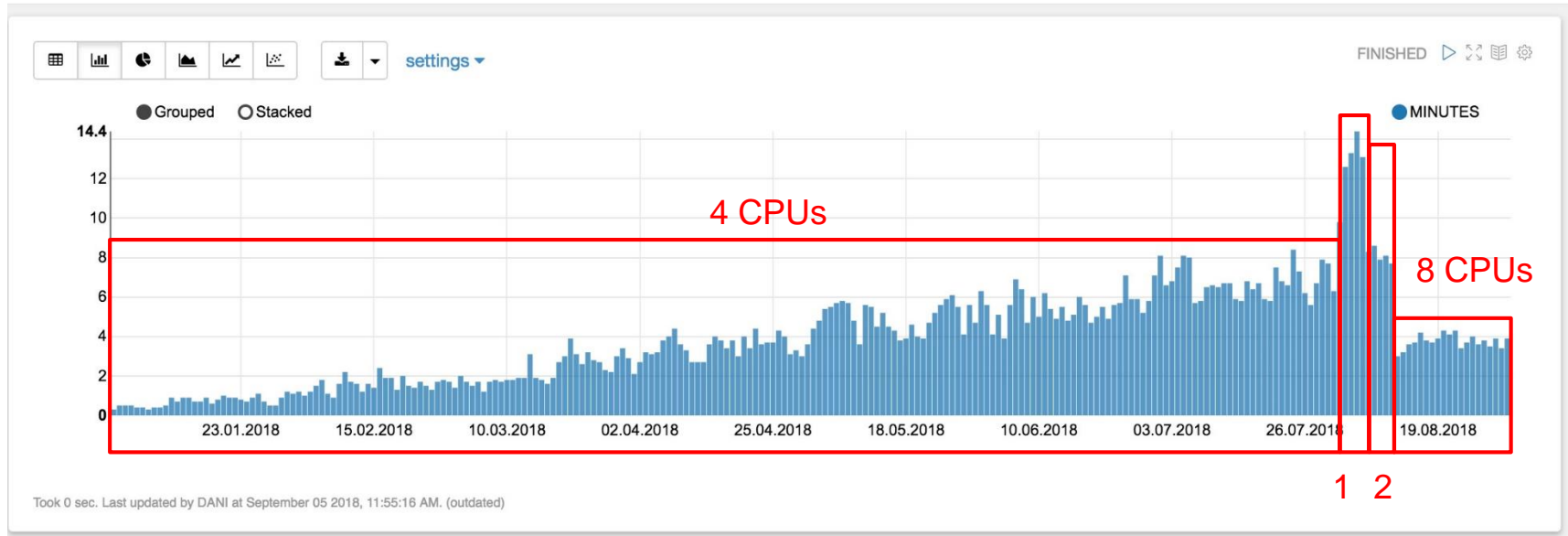
## Data Vault Model of Craft Beer Brewery

- 25 target tables (Hubs, Links, Satellites)
- 20 beers, 272K customers
- 16M orders, 173M order items
- “Daily loads” for 8 months
  - Random data generator
  - Data Vault load patterns
  - Delta detection and versioning
  - 1.4M order items / day



# Example: Loading Data Vault Schema

## ETL Monitoring



# Query Performance

# ■ Result Cache

It is enabled by default on ADW

■ `RESULT_CACHE_MODE = FORCE`

For queries being re-executed, it can lead to a tremendous performance improvement

To avoid caching, use the `NO_RESULT_CACHE` hint

■ Even if `OPTIMIZER_IGNORE_HINTS = TRUE` (default on ADW)

# ■ Improve Query Performance

## What cannot be done

- Use In-Memory Column Store

## What should not be done

- Partition tables
- Create materialized views
- Create indexes

## What can be done

- Scale up the number of CPU cores
- Use different service
- Use constraints to enable query transformations (e.g. join elimination)

# ■ How Are Queries on a Star Schema Optimized?

The star transformation is disabled (`star_transformation_enabled=FALSE`)

The query optimizer can use the vector transformation

- Introduced in 12.1.0.2 for *In-Memory Aggregation*
- In many situations, faster than star transformation

Note

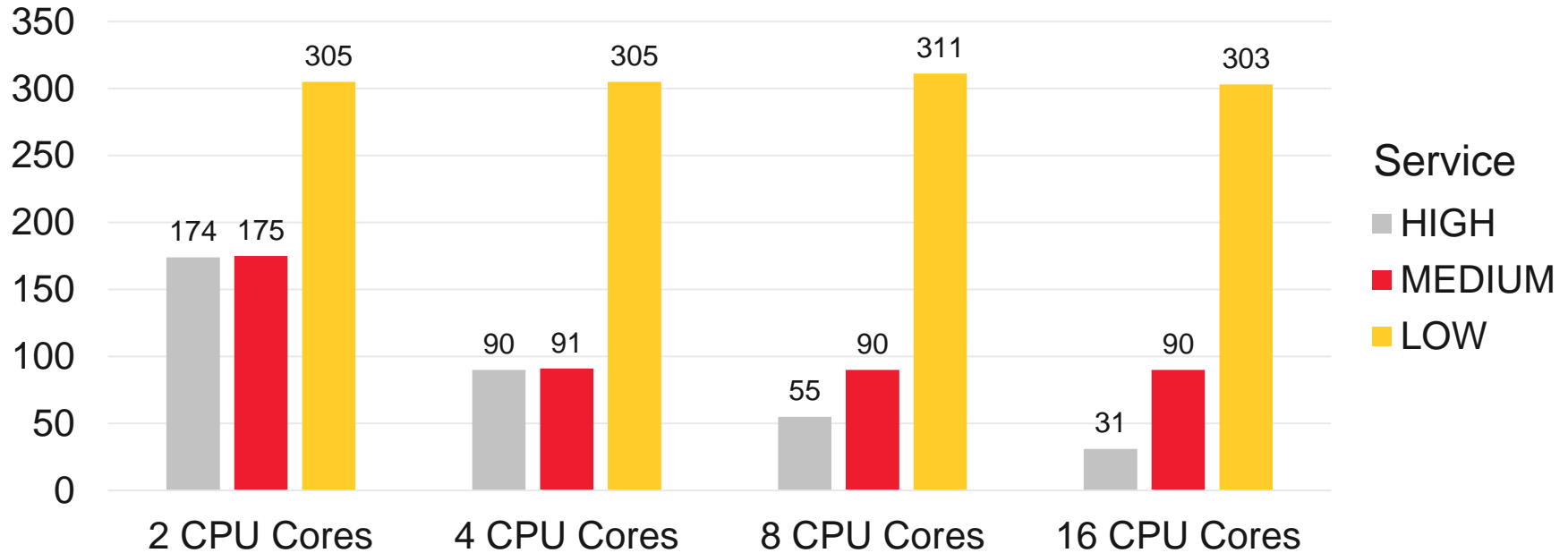
-----

- vector transformation used for this statement

- Works even if tables are not populated in IMCS
- In-Memory is enabled (but not used) in ADW
  - `INMEMORY_SIZE = 1073741824`

## ■ Example – TPC-DS Queries

### Query 38 – Execution Time in Seconds





# ■ Automated Tuning?

Exadata  
storage  
indexes?

## Industry-Leading Performance

- Integrated machine-learning algorithms drive automatic caching, adaptive indexing, advanced compression, and optimized cloud data-loading to deliver unrivaled performance.
- Automatic adaptive performance tuning delivers faster analytics.

Result cache?  
(in addition to disk I/O caching)

Not  
observed ☹️

source: <https://www.oracle.com/database/data-warehouse.html>

# Monitoring Performance

# ■ Monitoring Performance and SQL Statements

Service Console with monitoring capabilities:

## ■ Activity and utilization

- Storage and CPU usage
- Real-time or time period

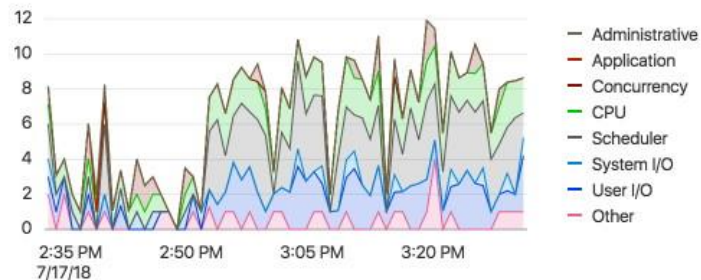
Default 8 days, can be changed with AWR settings

## ■ Running SQL statements

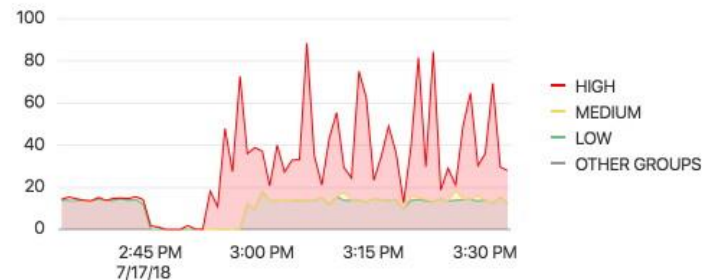
- Time & wait statistics, I/O statistics
- Runtime execution plan
- Parallel processes
- Downloadable real-time SQL Monitoring report

[Overview](#)**[Activity](#)**[Administration](#)

### Database Activity



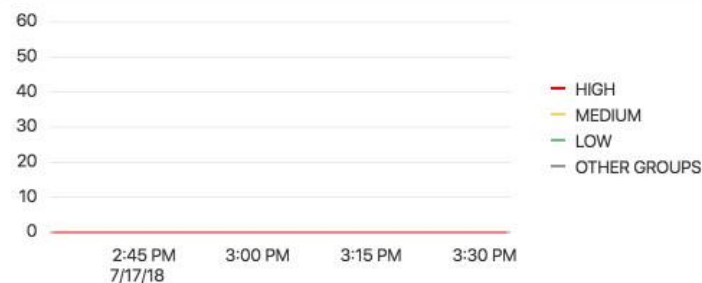
### CPU Utilization (%)



### Running Statements



### Queued Statements



# ■ Monitoring Performance with AWR and ASH Reports

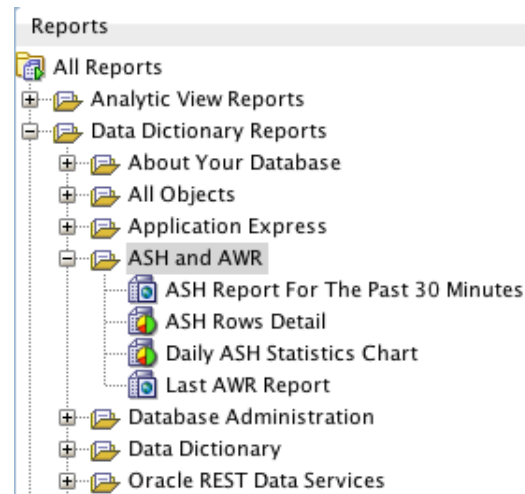
AWR reports can be generated (**only on PDB level**)

- With DBMS\_WORKLOAD\_REPOSITORY
- With SQL Developer reports

```
SELECT output FROM TABLE(  
    DBMS_WORKLOAD_REPOSITORY.AWR_REPORT_HTML  
    (3951758934, 7, 3524, 3535)  
)
```

ASH reports can be generated based on

- V\$ACTIVE\_SESSION\_HISTORY / DBA\_HIST\_ACTIVE\_SESS\_HISTORY
- SQL Developer reports



# Miscellaneous

# ■ Backup & Recovery

Automatic (incremental) backups take place daily

- The start time cannot be set
- The retention period for automatic backups is 60 days

Manual backups can be initiated through the console

- They are stored in the object store
- The `DEFAULT_CREDENTIAL` and `DEFAULT_BUCKETS` database properties must be set

Recovery at any point-in-time can be initiated through the console

# ■ Patching

Oracle is patching the service on a regular basis

- No announcement

- No downtime

There is no way to “schedule” when the installation of the patches takes place



# ■ Oracle Support

As soon as a problem cannot be solve because of missing privileges, an SR has to be opened

Support has (almost) no visibility

- E.g. no access to the alert.log of the PDB

- E.g. no OS access

Support relies on Operations to fix things

The issues we experienced were fixed in a time frame going from one days to one week

# Conclusion

# ■ Conclusion

- Appropriate setup for DWH
- Better than many manual configured DWHs
- Easy (limited) administration, ready to use
- No In-Memory Column Store at the database-server level
- Support in case of a service request can take too long
- Sometimes “shaky” (no control about patches / changes)
- Not as simple as it looks at first sight
- Knowhow about physical DB design still important



## ■ Further Information in Blog Posts

- DBMS\_CLOUD Package – A Reference Guide  
[https://antognini.ch/2018/07/dbms\\_cloud-package-a-reference-guide/](https://antognini.ch/2018/07/dbms_cloud-package-a-reference-guide/)
- Which Privileges Are Required to Use the ADWC Service Console?  
<https://antognini.ch/2018/07/which-privileges-are-required-to-use-the-adwc-service-console/>
- Observations About the Scalability of Data Loads in ADWC  
<https://antognini.ch/2018/07/observations-about-the-scalability-of-data-loads-in-adwc/>
- External Tables in Autonomous Data Warehouse Cloud  
<https://danischnider.wordpress.com/2018/07/04/external-tables-in-autonomous-data-warehouse-cloud/>
- Gathering Statistics in the Autonomous Data Warehouse Cloud  
<https://danischnider.wordpress.com/2018/07/11/gathering-statistics-in-the-autonomous-data-warehouse-cloud/>
- 10 Tips to Improve ETL Performance – Revised for ADWC  
<https://danischnider.wordpress.com/2018/07/20/10-tips-to-improve-etl-performance-revised-for-adwc/>
- Star Schema Optimization in Autonomous Data Warehouse Cloud  
<https://danischnider.wordpress.com/2018/09/13/star-schema-optimization-in-autonomous-data-warehouse-cloud/>

# Questions and Answers...

Christian Antognini  
Dani Schnider



@chrisantognini



[antognini.ch/blog](http://antognini.ch/blog)



@dani\_schnider



[danischnider.wordpress.com](http://danischnider.wordpress.com)