CERN

# Research Analytics at Scale: CERN's Experience with Oracle's Cloud Solutions

Oracle OpenWorld London, 16-17th January, 2019
https://indico.cern.ch/event/789142/

Eric Grancher, Manuel Martin Marquez, Sébastien Masson

# Outline

Introduction

Cloud solutions

ADWC Setup

Moving Data to ADWC

Insights

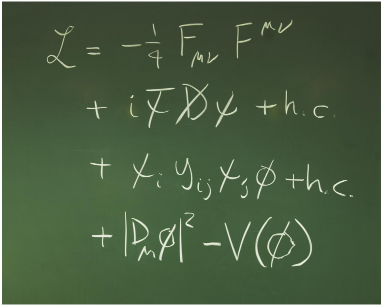
Conclusion

# CERN

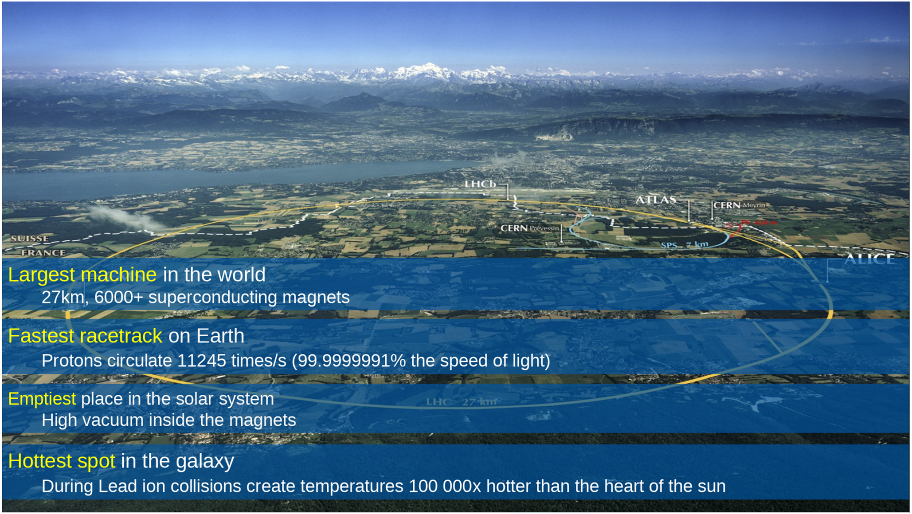
- CERN - European Council for Nuclear Research
- Founded in 1954 by 12 countries for fundamental physics research in the post-war Europe
- Today 22 members states (see `https://home.cern/about/member-states`) and world-wide collaborations, 2 300 CERN personnel.
- More information at `https://home.cern/about`
- Can be visited, see `https://visit.cern/` (and 14-15 September 2019 CERN Open Days).

# Fundamental Research

- What is 95% of the Universe made of?
- Why do particles have mass?
- Why is there no antimatter left in the Universe?
- What was the Universe like, just after "Big Bang"?



$$
\begin{aligned}
\mathcal{L} = &-\frac{1}{4} F_{\mu\nu} F^{\mu\nu} \\
&+ i \bar{\psi} \not{D} \psi + h.c. \\
&+ \psi_i y_{ij} \psi_j \phi + h.c. \\
&+ |D_\mu \phi|^2 - V(\phi)
\end{aligned}
$$



Largest machine in the world

27km, 6000+ superconducting magnets

Fastest racetrack on Earth

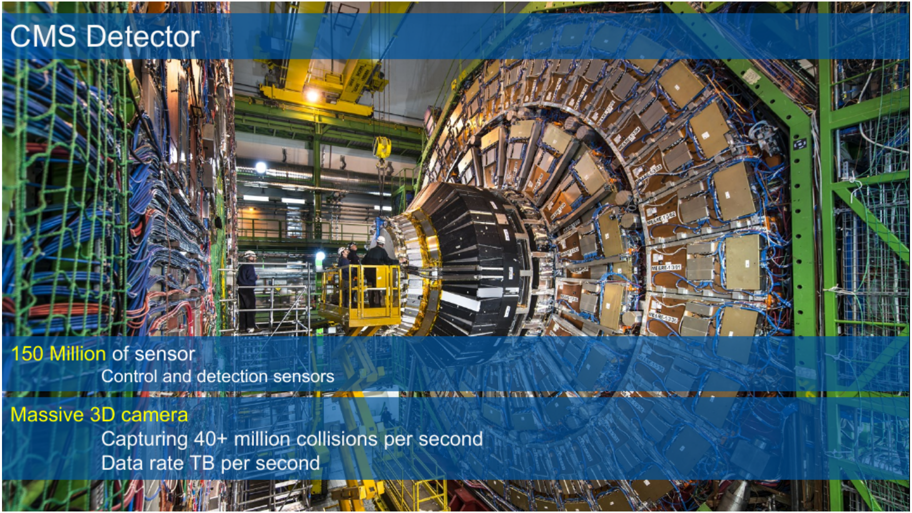
Protons circulate 11245 times/s (99.9999991% the speed of light)

Emptiest place in the solar system

High vacuum inside the magnets

Hottest spot in the galaxy

During Lead ion collisions create temperatures 100 000x hotter than the heart of the sun



# CMS Detector

150 Million of sensor

Control and detection sensors

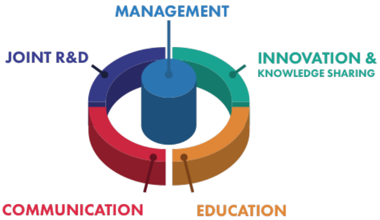
Massive 3D camera

Capturing 40+ million collisions per second

Data rate TB per second

# CERN openlab

- Public-private partnership, through which CERN collaborates with leading ICT companies and other research organizations.
  - Evaluate state-of-the-art technologies in a challenging environment and improve them.
  - Train the next generation of engineers/researchers.
  - Promote education and cultural exchanges.
  - Communicate results and reach new audiences.
- Oracle is a member since 2003.

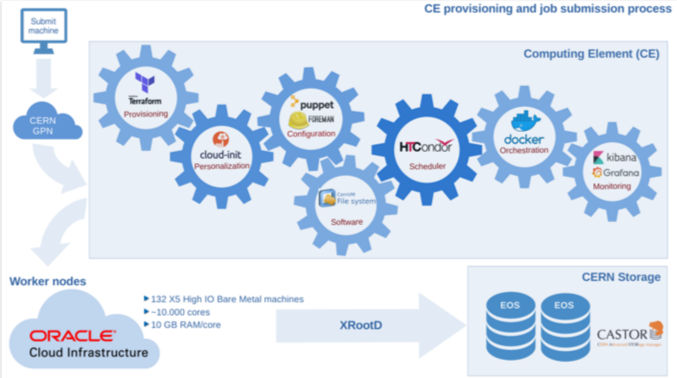


# Oracle Managed Cloud Service

- Previously "online-DBA".
- Manages CERN's E–Business Suite HR system.
- "At customer", in our computer center in Meyrin (CH).

# Oracle Cloud Infrastructure - WLCG

- Proof of concept, 2017Q3, integration of OCI inside WLCG.
- 132 X5 High IO Bare Metal Machines. Around 10 000 cores. 10GB RAM per core.



# Autonomous database - Industrial IoT

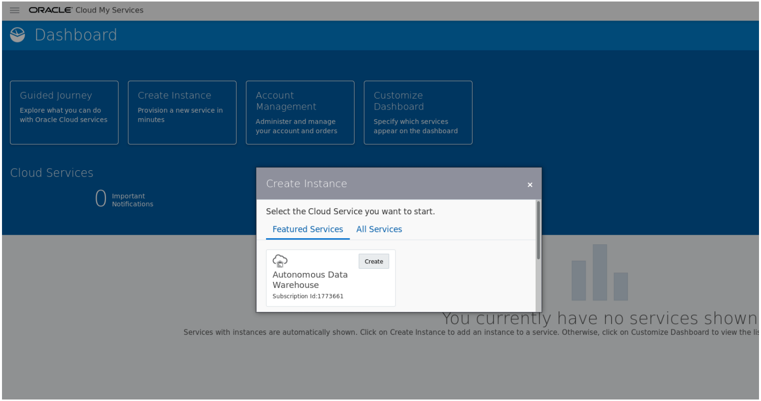
- In addition to the vast amount of physics data created at CERN, the control systems serving the accelerator complex and the supporting technical infrastructure generate very large amounts of data themselves.

# PSEN schema - the first step

- 750GB of SCADA data about Electrical Network
- Contains IOT partitioned tables
- One big IOT partitioned table of 620G

# ADWC Setup

# Creation of your ADWC instance: GUI



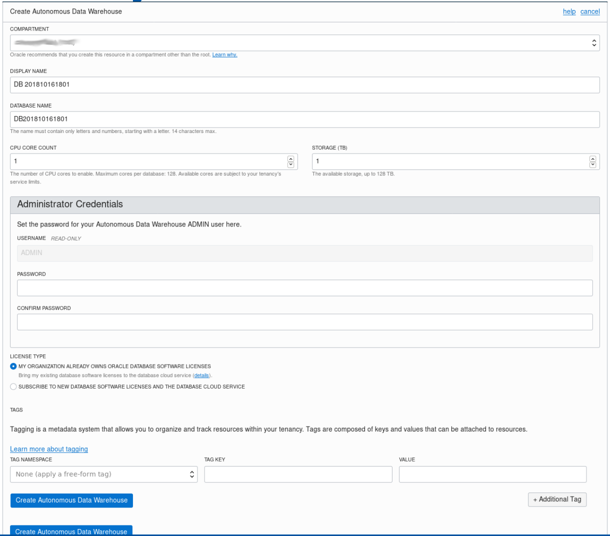
ORACLE Cloud My Services
Dashboard
Guided Journey
Explore what you can do with Oracle Cloud services
Create Instance
Provision a new service in minutes
Account Management
Administer and manage your account and orders
Customize Dashboard
Specify which services appear on the dashboard
Cloud Services
0 Important Notifications
Create Instance
×
Featured Services
All Services
Create
You currently have no services shown

Select the Cloud Service you want to start.

Autonomous Data Warehouse
Subscription Id:1773661

Services with instances are automatically shown. Click on Create Instance to add an instance to a service. Otherwise, click on Customize Dashboard to view the lis

# Creation of your ADWC instance: GUI



Create Autonomous Data Warehouse

COMPARTMENT

[redacted]

Oracle recommends that you create this resource in a compartment other than the root. Learn why.

DISPLAY NAME

DB 201810161801

DATABASE NAME

DB201810161801

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

CPU CORE COUNT

1

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

STORAGE (TB)

1

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

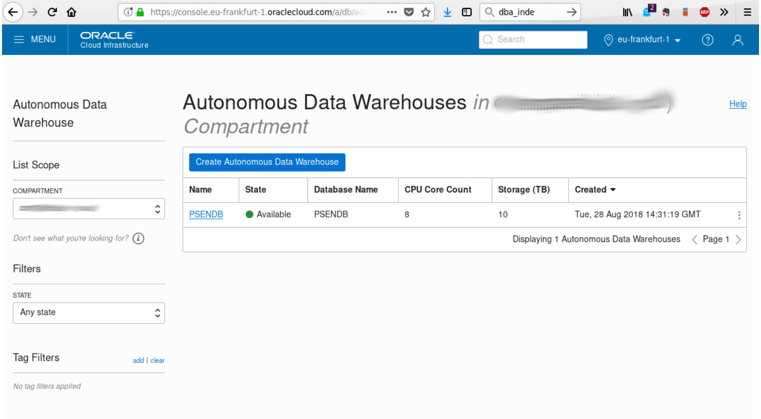
Tagging is a metadata system that allows you to organize and track resources within your tenancy. Tags are composed of keys and values that can be attached to resources.

Learn more about tagging

| TAG NAMESPACE | TAG KEY | VALUE |
|---|---|---|
| None (apply a free-form tag) | | |

Create Autonomous Data Warehouse
+ Additional Tag
Create Autonomous Data Warehouse

# Creation of your ADWC instance: GUI


https://console.eu-frankfurt-1.oraclecloud.com/a/db/ad
dba_inde
MENU
ORACLE Cloud Infrastructure
Search
eu-frankfurt-1

Autonomous Data Warehouse

List Scope

COMPARTMENT

[redacted]

*Don't see what you're looking for?*

Filters

STATE

Any state

Tag Filters add | clear

*No tag filters applied*

## Autonomous Data Warehouses *in* [redacted] *Compartment*

Help

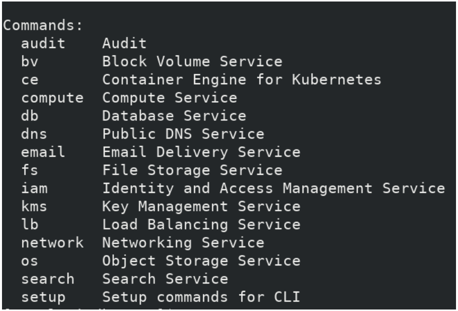
Create Autonomous Data Warehouse

| Name | State | Database Name | CPU Core Count | Storage (TB) | Created ▾ |
|---|---|---|---|---|---|
| PSENDB | ● Available | PSENDB | 8 | 10 | Tue, 28 Aug 2018 14:31:19 GMT |

Displaying 1 Autonomous Data Warehouses ‹ Page 1 ›

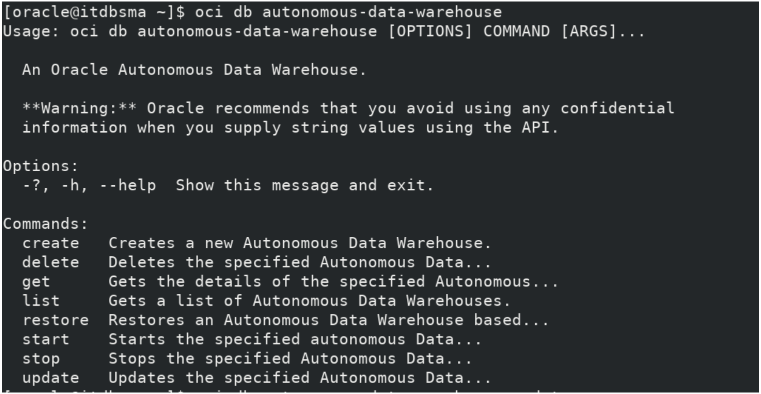
# Creation of your ADWC instance: OCI

- You can also use Oracle Cloud Infrastructure CLI or OCI CLI command-line to perform this kind of operation.
- OCI REST APIs.



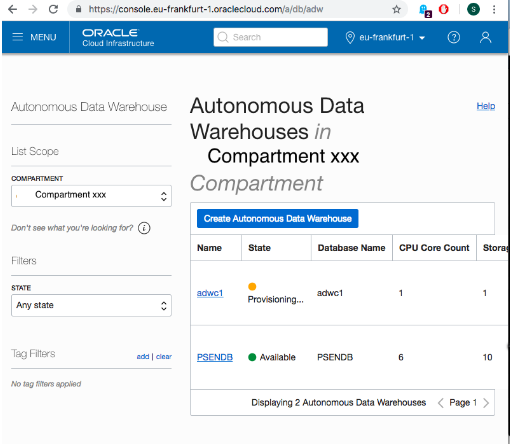
```
Commands:
  audit    Audit
  bv       Block Volume Service
  ce       Container Engine for Kubernetes
  compute  Compute Service
  db       Database Service
  dns      Public DNS Service
  email    Email Delivery Service
  fs       File Storage Service
  iam      Identity and Access Management Service
  kms      Key Management Service
  lb       Load Balancing Service
  network  Networking Service
  os       Object Storage Service
  search   Search Service
  setup    Setup commands for CLI
```

# Creation of your ADWC instance: OCI



```
[oracle@itdbsma ~]$ oci db autonomous-data-warehouse
Usage: oci db autonomous-data-warehouse [OPTIONS] COMMAND [ARGS]...

  An Oracle Autonomous Data Warehouse.

  **Warning:** Oracle recommends that you avoid using any confidential
  information when you supply string values using the API.

Options:
  -?, -h, --help  Show this message and exit.

Commands:
  create   Creates a new Autonomous Data Warehouse.
  delete   Deletes the specified Autonomous Data...
  get      Gets the details of the specified Autonomous...
  list     Gets a list of Autonomous Data Warehouses.
  restore  Restores an Autonomous Data Warehouse based...
  start    Starts the specified autonomous Data...
  stop     Stops the specified Autonomous Data...
  update   Updates the specified Autonomous Data...
```

# Creation of your ADWC instance: OCI



https://console.eu-frankfurt-1.oraclecloud.com/a/db/adw
MENU
ORACLE Cloud Infrastructure
Search
eu-frankfurt-1

Autonomous Data Warehouse

List Scope

COMPARTMENT

Compartment xxx

*Don't see what you're looking for?*

Filters

STATE

Any state

Tag Filters add | clear

*No tag filters applied*

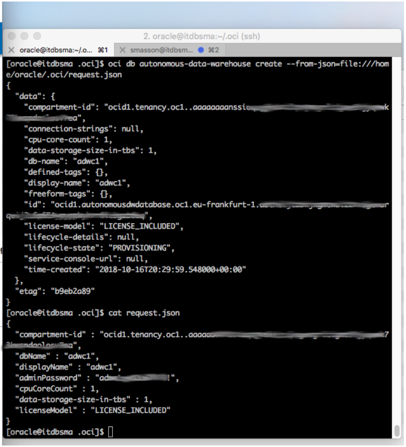
Autonomous Data Warehouses *in* Compartment xxx *Compartment*

Help

Create Autonomous Data Warehouse

| Name | State | Database Name | CPU Core Count | Storag |
|---|---|---|---|---|
| adwc1 | ● Provisioning... | adwc1 | 1 | 1 |
| PSENDB | ● Available | PSENDB | 6 | 10 |

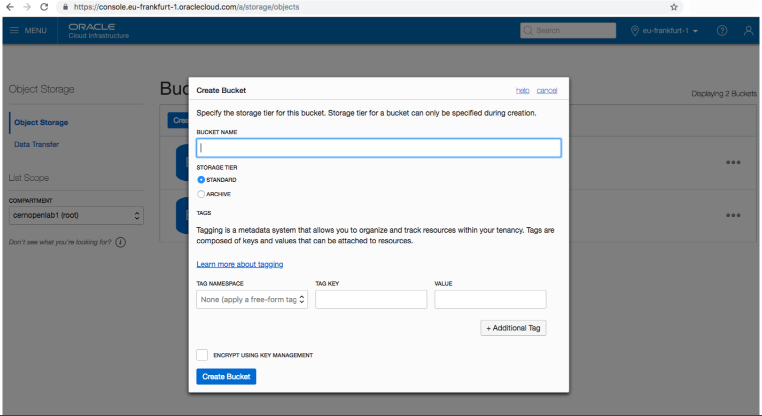
Displaying 2 Autonomous Data Warehouses < Page 1 >



```
[oracle@itdbsma .oci]$ oci db autonomous-data-warehouse create --from-json=file:///hom
e/oracle/.oci/request.json
{
  "data": {
    "compartment-id": "ocid1.tenancy.oc1..aaaaaaaanssi[redacted]",
    "connection-strings": null,
    "cpu-core-count": 1,
    "data-storage-size-in-tbs": 1,
    "db-name": "adwc1",
    "defined-tags": {},
    "display-name": "adwc1",
    "freeform-tags": {},
    "id": "ocid1.autonomousdwdatabase.oc1.eu-frankfurt-1.[redacted]",
    "license-model": "LICENSE_INCLUDED",
    "lifecycle-details": null,
    "lifecycle-state": "PROVISIONING",
    "service-console-url": null,
    "time-created": "2018-10-16T20:29:59.548000+00:00"
  },
  "etag": "b9eb2a89"
}
[oracle@itdbsma .oci]$ cat request.json
{
  "compartment-id" : "ocid1.tenancy.oc1..aaaa[redacted]",
  "dbName" : "adwc1",
  "displayName" : "adwc1",
  "adminPassword" : "ad[redacted]",
  "cpuCoreCount" : 1,
  "data-storage-size-in-tbs" : 1,
  "licenseModel" : "LICENSE_INCLUDED"
}
[oracle@itdbsma .oci]$
```

# Creation of your Object Storage Bucket: GUI

Now, you need to configure cloud object storage to upload your data pump export files.



https://console.eu-frankfurt-1.oraclecloud.com/a/storage/objects
MENU
ORACLE Cloud Infrastructure
Search
eu-frankfurt-1
Object Storage
Object Storage
Data Transfer
List Scope
Buc
Crea
Displaying 2 Buckets
Don't see what you're looking for?

Create Bucket help cancel

Specify the storage tier for this bucket. Storage tier for a bucket can only be specified during creation.

BUCKET NAME

STORAGE TIER

● STANDARD

○ ARCHIVE

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-form tag | | |

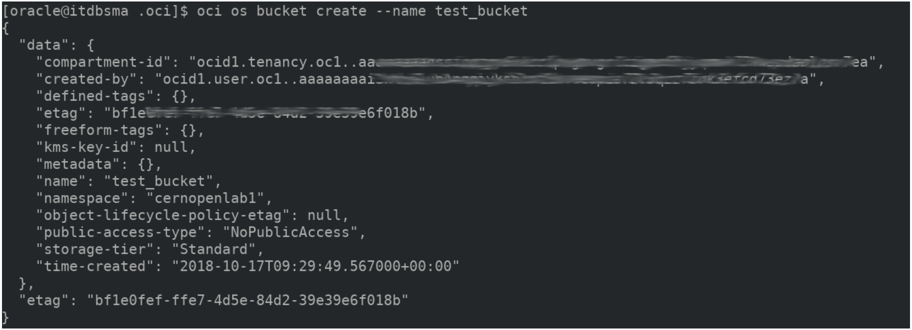
+ Additional Tag

□ ENCRYPT USING KEY MANAGEMENT

Create Bucket

COMPARTMENT: cernopenlab1 (root)

# Creation of your Object Storage Bucket: OCI



```
[oracle@itdbsma .oci]$ oci os bucket create --name test_bucket
{
  "data": {
    "compartment-id": "ocid1.tenancy.oc1..aa[redacted]ea",
    "created-by": "ocid1.user.oc1..aaaaaaaai[redacted]a",
    "defined-tags": {},
    "etag": "bf1e[redacted]e6f018b",
    "freeform-tags": {},
    "kms-key-id": null,
    "metadata": {},
    "name": "test_bucket",
    "namespace": "cernopenlab1",
    "object-lifecycle-policy-etag": null,
    "public-access-type": "NoPublicAccess",
    "storage-tier": "Standard",
    "time-created": "2018-10-17T09:29:49.567000+00:00"
  },
  "etag": "bf1e0fef-ffe7-4d5e-84d2-39e39e6f018b"
}
```

# Moving Data to ADWC

# Moving Data to ADWC

- ADWC principle: data is imported in its simplest form:
  - No indexes
  - No partitions
  - No IOTs
  - No materialized views
  - ...
- Data pump allows the needed transformations

# Data pump export

- Oracle documentation recommends the following expdp parameters:

```
exclude=index, cluster, indextype, materialized_view, materialized_view_log,
materialized_zonemap, db_link
data_options=group_partition_table_data
parallel=n
schemas=schema name
dumpfile=export%u.dmp
```

# Data pump export

- data_option=group_partition_table_data :
  - Modifies storage parameters for partitions to allow faster import afterwards
  - **Requires 12.2 data pump client and database !** (not available in 11.2.0.4)

# Send data to object storage

- OCI is a very convenient way to send expdp files to object_storage
- OCI can split files and upload them in parallel. Files are automatically merged afterwards.

```
oci os object bulk-upload --bucket-name PSEN_BUCKET_1 --src-dir
/mnt/oci/ --part-size 64 --parallel-upload-count 10
```

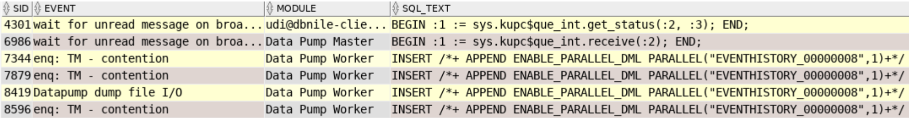
# Data pump import: parameters

```
parallel=4
partition_options=merge
transform=segment_attributes:n
transform=dwcs_cvt_iots:y
transform=constraint_use_default_index:y
exclude=index, cluster, indextype, materialized_view, materialized_view_log, materialized_zonemap, db_link
```

- Parallel set to the number of CPUs you have
- Partitioned tables are converted to non-partitioned tables
- All segment attributes are ignored
- IOTs are converted to regular tables
- PK and unique indexes renamed to constraint name
- Same exclusion as during export

# Data pump import

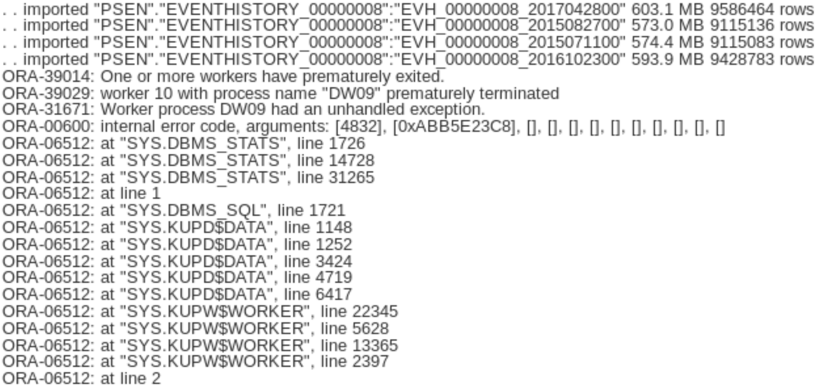
In our case, we could not use data_option=group_partition_table_data so we did not get any parallelism:



| SID | EVENT | MODULE | SQL_TEXT |
|---|---|---|---|
| 4301 | wait for unread message on broa... | udi@dbnile-clie... | BEGIN :1 := sys.kupc$que_int.get_status(:2, :3); END; |
| 6986 | wait for unread message on broa... | Data Pump Master | BEGIN :1 := sys.kupc$que_int.receive(:2); END; |
| 7344 | enq: TM - contention | Data Pump Worker | INSERT /*+ APPEND ENABLE_PARALLEL_DML PARALLEL("EVENTHISTORY_00000008",1)+*/ |
| 7879 | enq: TM - contention | Data Pump Worker | INSERT /*+ APPEND ENABLE_PARALLEL_DML PARALLEL("EVENTHISTORY_00000008",1)+*/ |
| 8419 | Datapump dump file I/O | Data Pump Worker | INSERT /*+ APPEND ENABLE_PARALLEL_DML PARALLEL("EVENTHISTORY_00000008",1)+*/ |
| 8596 | enq: TM - contention | Data Pump Worker | INSERT /*+ APPEND ENABLE_PARALLEL_DML PARALLEL("EVENTHISTORY_00000008",1)+*/ |

# Data pump import

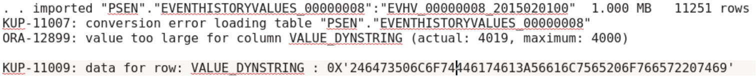
But not there yet...



```
. . imported "PSEN"."EVENTHISTORY_00000008":"EVH_00000008_2017042800" 603.1 MB 9586464 rows
. . imported "PSEN"."EVENTHISTORY_00000008":"EVH_00000008_2015082700" 573.0 MB 9115136 rows
. . imported "PSEN"."EVENTHISTORY_00000008":"EVH_00000008_2015071100" 574.4 MB 9115083 rows
. . imported "PSEN"."EVENTHISTORY_00000008":"EVH_00000008_2016102300" 593.9 MB 9428783 rows
ORA-39014: One or more workers have prematurely exited.
ORA-39029: worker 10 with process name "DW09" prematurely terminated
ORA-31671: Worker process DW09 had an unhandled exception.
ORA-00600: internal error code, arguments: [4832], [0xABB5E23C8], [], [], [], [], [], [], [], [], [], []
ORA-06512: at "SYS.DBMS_STATS", line 1726
ORA-06512: at "SYS.DBMS_STATS", line 14728
ORA-06512: at "SYS.DBMS_STATS", line 31265
ORA-06512: at line 1
ORA-06512: at "SYS.DBMS_SQL", line 1721
ORA-06512: at "SYS.KUPD$DATA", line 1148
ORA-06512: at "SYS.KUPD$DATA", line 1252
ORA-06512: at "SYS.KUPD$DATA", line 3424
ORA-06512: at "SYS.KUPD$DATA", line 4719
ORA-06512: at "SYS.KUPD$DATA", line 6417
ORA-06512: at "SYS.KUPW$WORKER", line 22345
ORA-06512: at "SYS.KUPW$WORKER", line 5628
ORA-06512: at "SYS.KUPW$WORKER", line 13365
ORA-06512: at "SYS.KUPW$WORKER", line 2397
ORA-06512: at line 2
```

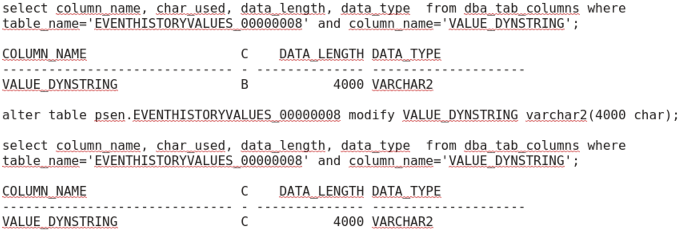
# Data pump import

- During this second attempt we got:



```
. . imported "PSEN"."EVENTHISTORYVALUES_00000008":"EVHV_00000008_2015020100"  1.000 MB   11251 rows
KUP-11007: conversion error loading table "PSEN"."EVENTHISTORYVALUES_00000008"
ORA-12899: value too large for column VALUE_DYNSTRING (actual: 4019, maximum: 4000)

KUP-11009: data for row: VALUE_DYNSTRING : 0X'246473506C6F74446174613A56616C7565206F766572207469'
```

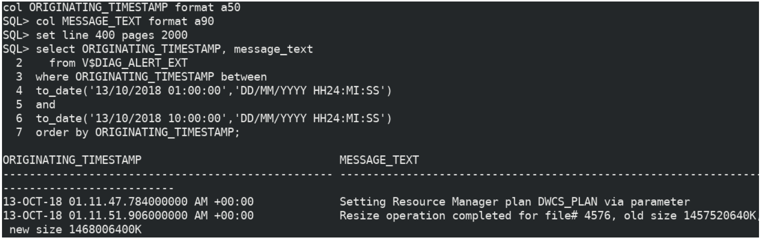
- This is due to the migration from a single-byte character set to a multi-byte one.
- The fix was to recreate the table, change column definition from byte to char and import data again:



```
select column_name, char_used, data_length, data_type  from dba_tab_columns where
table_name='EVENTHISTORYVALUES_00000008' and column_name='VALUE_DYNSTRING';

COLUMN_NAME                    C    DATA_LENGTH DATA_TYPE
------------------------------ - -------------- -------------------
VALUE_DYNSTRING                B           4000 VARCHAR2

alter table psen.EVENTHISTORYVALUES_00000008 modify VALUE_DYNSTRING varchar2(4000 char);

select column_name, char_used, data_length, data_type  from dba_tab_columns where
table_name='EVENTHISTORYVALUES_00000008' and column_name='VALUE_DYNSTRING';

COLUMN_NAME                    C    DATA_LENGTH DATA_TYPE
------------------------------ - -------------- -------------------
VALUE_DYNSTRING                C           4000 VARCHAR2
```

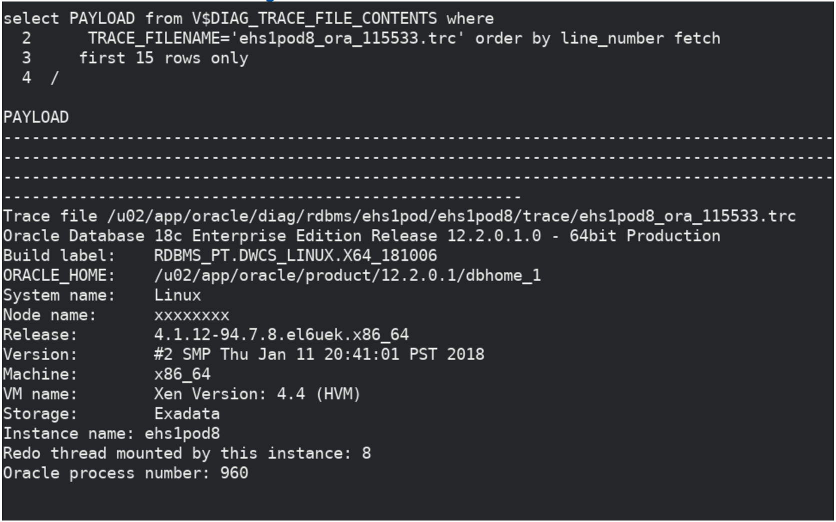
# Where is my alert.log?

Something that is a bit confusing when you start using ADWC.



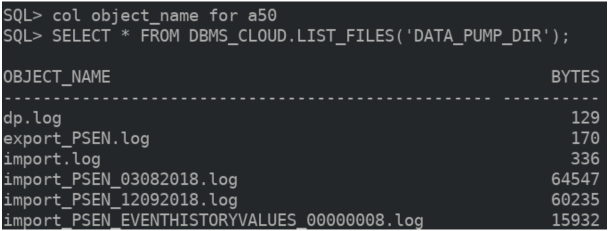
```
col ORIGINATING_TIMESTAMP format a50
SQL> col MESSAGE_TEXT format a90
SQL> set line 400 pages 2000
SQL> select ORIGINATING_TIMESTAMP, message_text
  2    from V$DIAG_ALERT_EXT
  3  where ORIGINATING_TIMESTAMP between
  4  to_date('13/10/2018 01:00:00','DD/MM/YYYY HH24:MI:SS')
  5  and
  6  to_date('13/10/2018 10:00:00','DD/MM/YYYY HH24:MI:SS')
  7  order by ORIGINATING_TIMESTAMP;

ORIGINATING_TIMESTAMP                              MESSAGE_TEXT
-------------------------------------------------- ---------------------------------------------------------------
--------------------------
13-OCT-18 01.11.47.784000000 AM +00:00             Setting Resource Manager plan DWCS_PLAN via parameter
13-OCT-18 01.11.51.906000000 AM +00:00             Resize operation completed for file# 4576, old size 1457520640K,
 new size 1468006400K
```

# Where are my trace files?



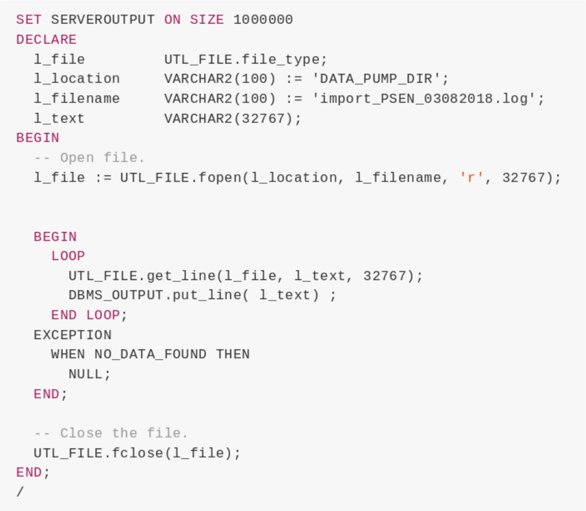
```
select PAYLOAD from V$DIAG_TRACE_FILE_CONTENTS where
  2      TRACE_FILENAME='ehs1pod8_ora_115533.trc' order by line_number fetch
  3     first 15 rows only
  4  /

PAYLOAD
----------------------------------------------------------------------------------------
----------------------------------------------------------------------------------------
----------------------------------------------------------------------------------------
-------------------------------------------------------
Trace file /u02/app/oracle/diag/rdbms/ehs1pod/ehs1pod8/trace/ehs1pod8_ora_115533.trc
Oracle Database 18c Enterprise Edition Release 12.2.0.1.0 - 64bit Production
Build label:    RDBMS_PT.DWCS_LINUX.X64_181006
ORACLE_HOME:    /u02/app/oracle/product/12.2.0.1/dbhome_1
System name:    Linux
Node name:      xxxxxxxx
Release:        4.1.12-94.7.8.el6uek.x86_64
Version:        #2 SMP Thu Jan 11 20:41:01 PST 2018
Machine:        x86_64
VM name:        Xen Version: 4.4 (HVM)
Storage:        Exadata
Instance name: ehs1pod8
Redo thread mounted by this instance: 8
Oracle process number: 960
```

# And what about my data pump logs?



```
SQL> col object_name for a50
SQL> SELECT * FROM DBMS_CLOUD.LIST_FILES('DATA_PUMP_DIR');

OBJECT_NAME                                             BYTES
-------------------------------------------------- ----------
dp.log                                                    129
export_PSEN.log                                           170
import.log                                                336
import_PSEN_03082018.log                                64547
import_PSEN_12092018.log                                60235
import_PSEN_EVENTHISTORYVALUES_00000008.log             15932
```

# And what about my data pump logs?



```
SET SERVEROUTPUT ON SIZE 1000000
DECLARE
  l_file         UTL_FILE.file_type;
  l_location     VARCHAR2(100) := 'DATA_PUMP_DIR';
  l_filename     VARCHAR2(100) := 'import_PSEN_03082018.log';
  l_text         VARCHAR2(32767);
BEGIN
  -- Open file.
  l_file := UTL_FILE.fopen(l_location, l_filename, 'r', 32767);


  BEGIN
    LOOP
      UTL_FILE.get_line(l_file, l_text, 32767);
      DBMS_OUTPUT.put_line( l_text) ;
    END LOOP;
  EXCEPTION
    WHEN NO_DATA_FOUND THEN
      NULL;
  END;

  -- Close the file.
  UTL_FILE.fclose(l_file);
END;
/
```

# Insights

# Direct comparison between on premise and ADWC is difficult...

- ADWC and on premise schemas are very different:
    - Transformations of partitioned, compressed IoTs to partitioned, compressed tables + PK index.
    - Execution plans changed a lot.
    - Our case is perhaps too specific to give a conclusion (But that is real life).

# Statistics

Some statistics were missing after data pump import:

- So we needed to gather them on our schema
- A check to the documentation told us it was expected if you do not use recommended parameters (partition merge related parameters in our case).

# Compression

Interesting case of EVENTHISTORY_00000008 table:

- This table is 620GB Index Organized Table (IOT) partitioned, compressed in our local database
- Transformed as non-IOT (but still partitioned and compressed) + Primary Key index on ADWC
  - Thanks to Hybrid Columnar Compression the table is now 70GB: full scans are smaller!
  - The ratio is the same for most of our tables.

# Indexes

Default scenario is to not use indexes on ADWC:

- HCC reduces the Full Scan workload
- Exadata Smart Scan
- Storage Indexes. Are useful to quickly locate data

# Operations

- CPU and storage can be adjusted online, at anytime in few seconds!

```
oci db autonomous-data-warehouse update –cpu-core-count 8
```

- So you can start with few resources and grow only if needed
- Automatic access to Oracle optimization features.

# Conclusion

# Some takeaways

- CLOUD is a significant change / opportunity in the way to do operations. Will help to focus on what really matters, in our case research.
- Automation frees time from some of the routine tasks (no prior expertise needed on some of the parts).
- Fast and flexible deployments. Minimize the risks and costs.
- Networking is a key element.
- However, you cannot move to the cloud from one day to another – careful planning, data transfers, etc.

# Thank you!

Questions, suggestions most welcome, now or via email.

Thanks for the support of many at Oracle for the ADWC work (Cris Pedregal, Pauline Mahrer, Cemil Alper, Sebastian Solbach, Brian Spendolini and others)

Do not hesitate to contact us at:
eric.grancher@cern.ch.
manuel.martin.marquez@cern.ch,
sebastien.masson@cern.ch,

`https://indico.cern.ch/event/789142/`

home.cern