

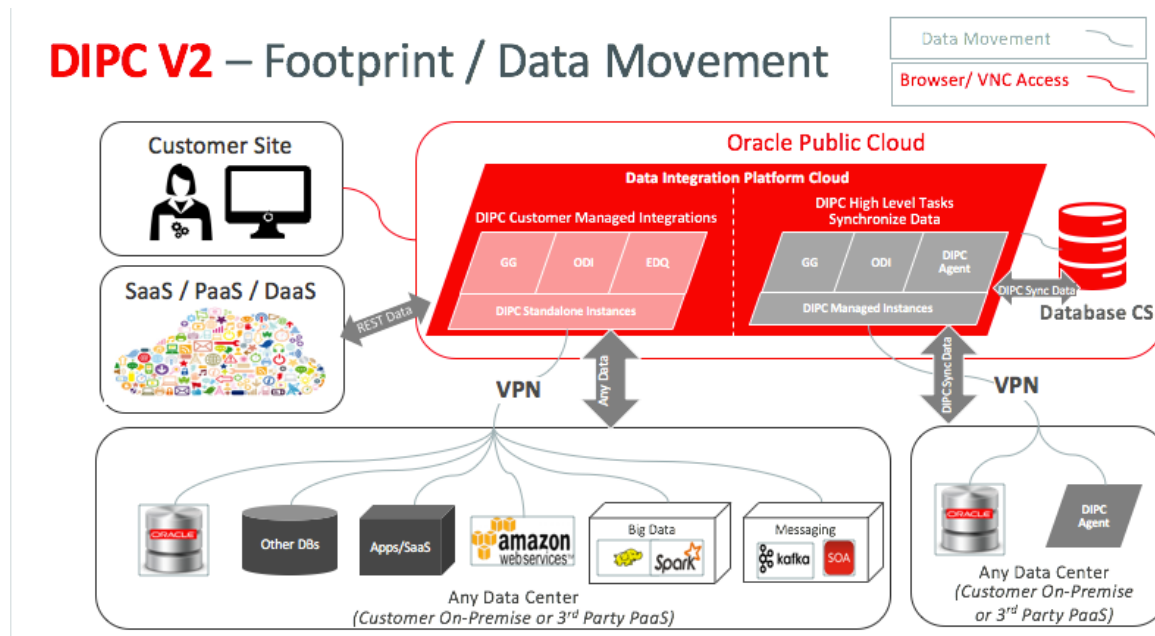


## Data Integration Platform Cloud Hands On Lab



## Architecture Overview

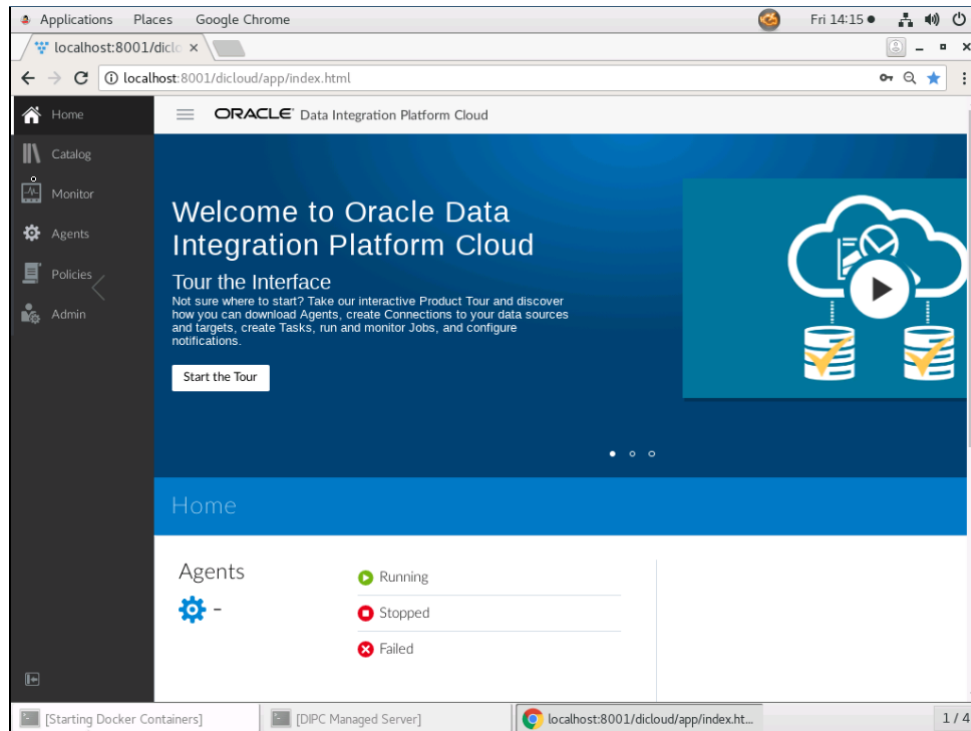
The rapid adoption of enterprise cloud-based solutions brings with it a new set of challenges. Data Integration becomes one of the greatest challenges of any enterprise cloud-based solution. Join this hands-on lab to have firsthand experience of the power and simplicity of Oracle Data Integration Platform Cloud. See how DIPC helps with data lake automation, enabling an intuitive instantiation and copy of data into a data lake, in an effort to help reduce some of the existing data engineer/ data scientist friction through a new Data Lake Builder task! You can quickly create a comprehensive, end-to-end repeatable data pipeline to your data lake.





## Task 1: Create Object Storage Classic Connection

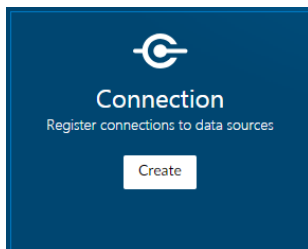
1. Log into Data Integration Platform Cloud
  - a. From your local laptop/desktop start Chrome (you can also continue to use Chrome on the Ravello instance)
  - b. Go to your <Ravello Instance hostname>:8001/dicloud/login.html
  - c. Login with weblogicwelcome1The following page should appear –



2. Click Home
3. Click Home and click on the right arrow in the carousel



4. Select Create under Connection





- a. Using your own Object Storage Classic account fill in the required Parameters

**Connection**

**General Information**

Name \* DL\_OBJECT\_STORAGE

Identifier \* DL\_OBJECT\_STORAGE

Description Storage for Data Lake

Agent DIPC.localdomain:7005

Type \* Oracle Object Storage Classic

**Connection Settings**

Domain \* Storage-...

Service URL \* https://storage-...us.storage.or

Container \* DIPC\_Demo

Username \* j...@oracle.com

Password \* .....

Test Connection Cancel Save

- b. Test the Connection and Save it
  - c. Then click on Save
5. The Connection now appears in the Catalog

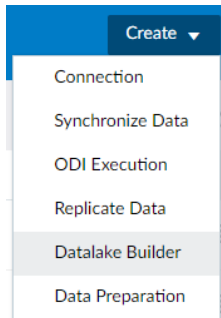
Category	Name	Type	Last Updated	Popularity
	DL_OBJECT_STORAGE	Oracle Object Storage Classic	8/7/2018, 1:03:11 PM	

**Note:** This Connection could also have been created from the Data Lake Builder Task



## Task 2: Create Data Lake Builder Task

1. Click on the Create button in the Catalog and select Datalake Builder



2. The Add Data to Datalake Task screen appears

3. Enter the following information
  - a. **Name:** Load Customer Data to Data Lake
  - b. **Description:** Load Customer data from Oracle into Object Storage Data Lake
  - c. Under Source Configuration:
    - **Connection:** select 'Sync Source' (refer to Synchronize Data Lab to create it in case it doesn't exist)
    - **Schema:** DIPC\_SRC



- Click on select next to Data Entity, pick SRC\_CUSTOMER then click Select

Select Data Entity

Connection

Sync Source

Schema

DIPC\_SRC

Available Data Entities

SRC\_

SRC\_AGE\_GROUP

SRC\_CITY

SRC\_CUSTOMER

SRC\_ORDERS

SRC\_ORDER\_LINES

SRC\_PRODUCT

SRC\_REGION

SRC\_SALES\_PERSON

Cancel

Select

## d. Under Target Configuration

- Connection: select 'DL\_OBJECT\_STORAGE'
- Data Entity: customer\_data
- Type: Parquet
- File Path: leave it blank

## Target Configuration ?

Connection \*

DL\_OBJECT\_STORAGE

+

Data Entity \*

customer\_data

?

Type \*

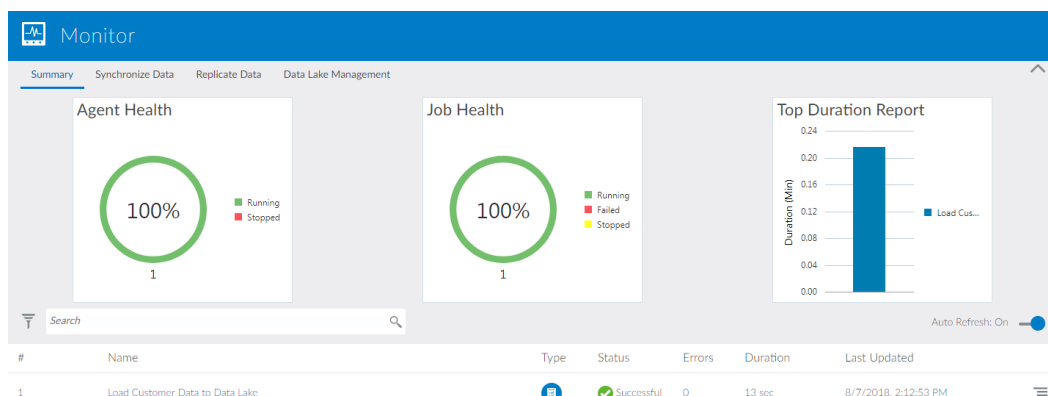
Parquet

▼

File Path

Directory/File

4. Click Save and Execute
5. Review Job Execution





Load Customer Data to Data Lake

Summary History Auto Refresh: Off

✓ Load Customer Data to Data Lake  
Type: Data Lake Add Data Task  
ID: 3762f459-30f7-4bc1-b2d5-b1851b05e0cf  
Last Updated: 8/7/2018, 2:12:53 PM

Errors	0	Start	8/7/2018, 2:12:40 PM
Processed	35	End	8/7/2018, 2:12:53 PM
Bytes Read	2893	Duration	13 sec
Bytes Written	2893	Created By	weblogic

Expand All | Collapse All

Action Name	Status	Processed	Errors	Duration	Bytes Read	Bytes Written
DATALAKE_Process_Data Lake Add Data Task0	✓ Successful	35	-	-	2893	2893

- Use the Object Storage Classic Service Console or REST APIs to list the files stored in the target Container

```
< X-Container-Meta-Policy-Georeplication: container
< X-Last-Modified-Timestamp: 1532602497.46756
< Content-Type: text/plain;charset=utf-8
< Content-Length: 42
< Server: Oracle-Storage-Cloud-Service
DL_CUSTOMER.parquet
customer_data.parquet
* Connection #0 to host storage-fbda41478ab043c49ffdb0e33636f5cb.us.storage.oraclecloud.com left intact
```

- The target Data Entity will also be listed in the Catalog

Catalog Data Entities Create

Search

Category	Name	Type	Last Updated	Popularity
	customer_data	File	8/7/2018, 2:12:52 PM	

customer\_data

Summary History

Name	customer_data
Identifier	CUSTOMER_DATA
Resource Name	customer_data
Description	
Type	File
Namespace	
Contact	
Popularity	
Tags	Add Tags...

## Summary

You have now successfully completed the Hands on Lab, and have successfully loaded data into a data lake using Oracle's Data Integration Platform Cloud.