DSpace Oracle to PostgreSQL migration

The following procedure is only one example of an Oracle to PostgreSQL migration done while upgrading from DSpace v6.3 to v7.2. It was kindly donated by the team at Université Laval. We encourage other institutions performing this migration to help enhance this guide, based on your own experiences. DSpace Oracle Support is deprecated, and will be removed by mid 2023. See the announcement at https://github.com/DSpace/DSpace/issues/8214. Therefore, this guide may be useful to any institutions looking to migrate to PostgreSQL before Oracle support ends.

Install / Setup Ora2Pg

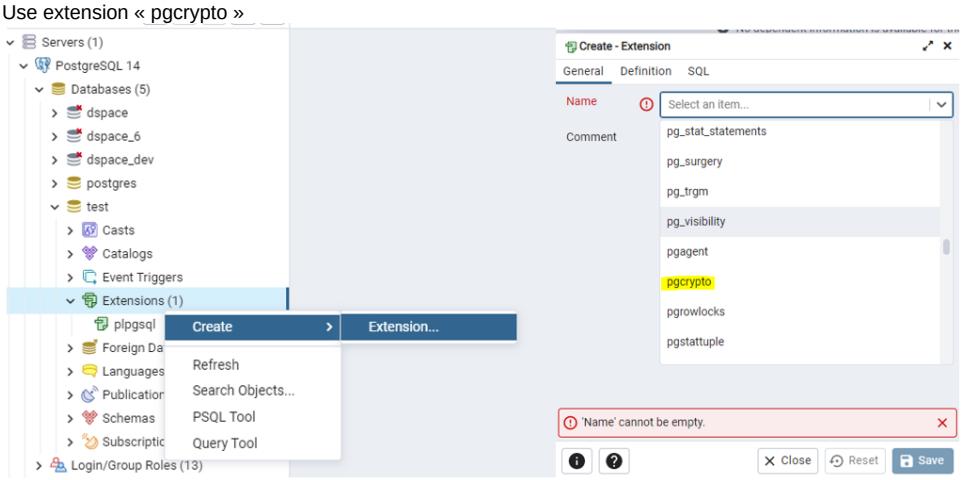
This migration was made using Ora2pg: https://ora2pg.darold.net/documentation.html

1. Install PERL: http://strawberryperl.com/
2. Download Ora2Pg: https://sourceforge.net/projects/ora2pg/
    - Unzip and access the file from the command line
3. Install Ora2Pg :
   perl Makefile.PL
   gmake && gmake install

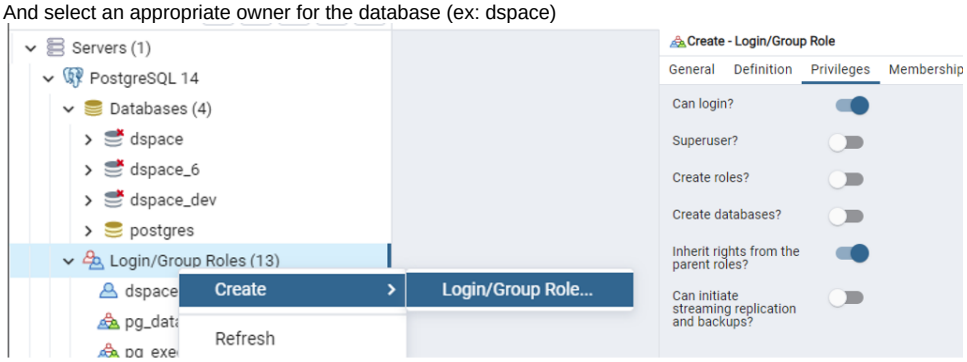
4. ora2pg files will be installed in Strawberry directory
    - C:\Strawberry\perl\site\bin
    - C:\ora2pg
      Note : Do not use another path for Ora2Pg, otherwise configuration will not be made correctly in Strawberry Perl.
5. Rename file « C:\ora2pg\ora2pg_conf.dist » en « ora2pg.conf »
    a. Modify variable ORACLE_HOME to point to Oracle client
    b. Modify string connexion to Oracle DB :
        i. User/Password
        ii. ORACLE_DSN, ORACLE_USER, ORACLE_PWD
    c. TYPE = TABLE,SEQUENCE,INSERT
    d. UUID_FUNCTION = gen_random_uuid
    e. STOP_ON_ERROR = 0
    f. FKEY_DEFERRABLE = 1
    g. DEFER_FKEY = 1
    h. MODIFY_TYPE
    bitstream:size_bytes:bigint,checksum_history:check_id:bigint,checksum_history:checksum_expected:varchar,checksum_history:checksum_calculated:varchar,checksum_history:result:varchar,checksum_results:result_code:varchar,checksum_results:result_description:varchar,harvested_collection:oai_source:varchar,harvested_collection:oai_set_id:varchar,harvested_collection:harvest_message:varchar,harvested_collection:metadata_config_id:varchar,harvested_collection:harvest_start_time:timestamp with time zone,harvested_collection:last_harvested:timestamp with time zone,harvested_item:last_harvested:timestamp with time zone,harvested_item:oai_id:varchar,item:last_modified:timestamp with time zone,metadatavalue:text_lang:varchar(24),most_recent_checksum:expected_checksum:varchar,most_recent_checksum:current_checksum:varchar,most_recent_checksum:checksum_algorithm:varchar,most_recent_checksum:result:varchar,resourcepolicy:start_date:date,resourcepolicy:end_date:date,webapp:url:varchar,webapp:isui:integer
    i. REPLACE_AS_BOOLEAN NUMBER:1 CHAR:1
    j. DATA_TYPE NUMBER(*\,0):integer,VARCHAR2(2000):text
    k. SCHEMA DSPACE

Data migration procedure

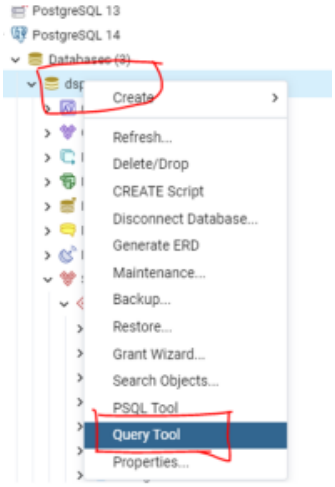
1. Open command line and move to following directory:
   C:\Strawberry\perl\site\bin.
2. Execute command line ora2pg to launch migration script from Oracle to SQL files for PostgresSQL objects creation.
3. Create a database on PostgreSQL server
    a. Use extension « pgcrypto »

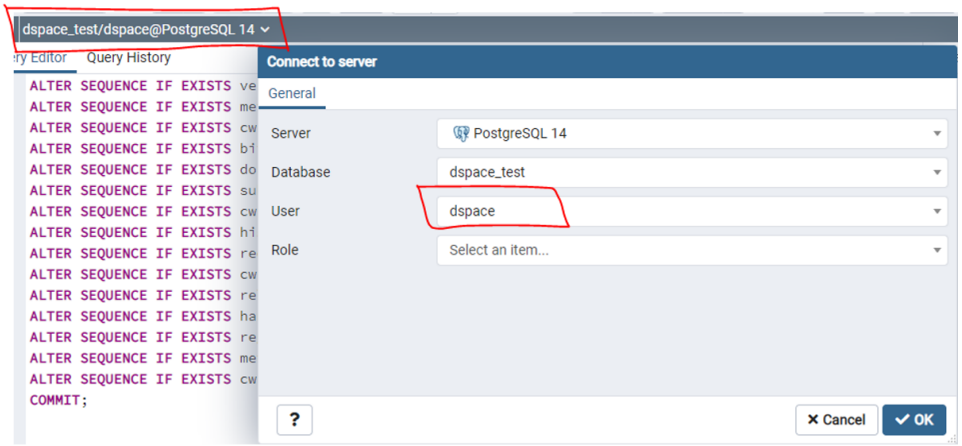


b. And select an appropriate owner for the database (ex: dspace)



4. Execute obtained SQL scripts as ***« dspace »*** using the above mentionned command in point 2.
5. For a SQL script that is too large, it is necessary to use the following commands in admin mode:
   a. SET PGCLIENTENCODING=utf-8
   b. psql -h localhost -p 5432 -U dspace -d dspace -f INSERT_output.sql 1>out.txt 2>err.txt





Warning: You must have previously defined a password for the user “dspace”. When clicking on OK, you will be prompt for this password.

6. Replace content of the « schema_version » table, doing a « truncate » and using the import function in the PostgreSQL admin tool. The file to import is csv. You will have to use “comma” as delimiter and select that there is a Header in the file. In the 2nd tab, use NULL as value for empty strings.
7. Run scripts SQL_FUNCTIONS.sql and SQL_SEQUENCES.sql.
   Note: These 2 file are a local one.
8. Run script 0__DS-2701_xml_workflow_migration.sql.
9. Run script 0__DS-2701_data_workflow_migration.sql.
10. Run script SQL_TABLES.sql.
    Note: This file is a local one.
11. (Optional), backup the database and run the migration program to DSpace 7.2 as documented on DSpace 7 Documentation: https://wiki.lyrasis.org/display/DSDOC7x/Upgrading+DSpace.