DSpace Release 1.5.x Documentation for Older Database Upgrades

<?xml version="1.0" encoding="utf-8"?>
<html>
This document covers my experience with upgrading DSpace@MITs Postgres instance during the 1.4.2 to 1.5. More speciically, thanks to deal's uncovered by Larry Stone we discovered that Postgres 7.3 or earlier used raw Triggers to implement system Constrains in Postgres and that if you've been maintaing an instance of DSpace that originated on postgres 7.3 or earlier that dealing with database table drops in the 1.5 upgrade process will leave behind broken triggers that have random identifiers and cannot be easily removed with serious review of the database schema.

I've worked a more general solution to the above problem that works by (1) performing a small number of SQL statements that will remove unnecessary tables and columns from the database (2) separating the database data in a backup from from its schema and (3) restoring that data to a fresh instance of the appropriate schema for DSpace 1.4.2 (that has been altered to remove any data inserts. This script and schema are included here. However, it may require tailoring to meet the requirements of your system.

Caution, this script can bring much harm to a database that contains many local customizations, you should seriously consult with your system admin and DSpace admin (as well as the community) before performing any of the following.

Drop any tables that are not represented in the DSpace 1.4.2 Schema

1. these tables, sequences and columns are no longer used at this time.
   psql -U dspace -d dspace.mit.edu -c 'DROP TABLE "action";'
   psql -U dspace -d dspace.mit.edu -c 'DROP TABLE "statistics";'
   psql -U dspace -d dspace.mit.edu -c 'DROP SEQUENCE statistics_seq;'
   psql -U dspace -d dspace.mit.edu -c 'ALTER TABLE item DROP COLUMN withdrawal_date;'

1. This table contains dulicate keys which require dropping, the content of this table is temporary

   a. Many of the existing records are stale. The table will be restored by the new schema
      psql -U dspace -d dspace.mit.edu -c 'DROP TABLE registrationdata;'
      psql -U dspace -d dspace.mit.edu -c 'DROP SEQUENCE registrationdata_seq;'

Dump the existing Database data to a backup file

1. Dump the database using the following settings
2. -a : data only
3. -D : use inserts with colmn names to correct for column ordering
4. -F c : Compressed tar
5. the rest is self explanitory
   pg_dump -a -D -F c -U dspace dspace.mit.edu -f dump.sql.custom

Drop and recreate the database using the altered 1.4.2 schema

1. drop and initialize the new database
   dropdb -U dspace dspace.mit.edu || error_exit "Failed to drop old db! Aborting";
   createdb -E UTF8 -U dspace dspace.mit.edu || error_exit "Failed to create new db! Aborting";
   psql -U dspace -d dspace.mit.edu < database_schema.sql

Enable dspace as a super user

1. Enable dspace to be superuser to disable triggers on inserts
   psql -U postgres -d dspace.mit.edu -c 'ALTER ROLE dspace WITH SUPERUSER'

Restore the data

1. restore the data to the new schema
   pg_restore -v --disable-triggers -U dspace -d dspace.mit.edu -e -a dump.sql.custom

Disable the Superuser

1. Disable dspace as superuser
   psql -U postgres -d dspace.mit.edu -c 'ALTER ROLE dspace WITH NOSUPERUSER'

Required files to perform the above work

Database_schema_1_4_2_altered_sql.mht

Repair-dspace-1_4_2.sh

</html>