PostGres V8.2.x Comparison to V7.4
NWS/Office of Hydrologic Development
Hydrologic Software Engineering Branch
2/4/2009

This document contains assorted notes and lessons learned on PostGres V8.2.x, especially as it compares to V7.4. It is intended for OHD/HSEB use.

Date Values Restrictions

On 3/28/2007, NWRFC reported a problem with bad date values in the rrevise and ratedat fields in the RiverStat table. Both of these fields are type DATE. They unloaded the database with the bad DATE values and attempted to load it into a new database on a V8.2 server. The load of the RiverStat table failed with an error.

This is an incompatibility between V7.4 and V8.2 of postgres. V7.4 allows records with the bad DATE value to be inserted into the database. V7.4 also allows the database containing the bad DATE values to be dumped and restored without error. V8.2 does not allow records with the bad DATE value to be inserted. Attempting to restore the V7.4 dump file into a database on a V8.2 server will result in the records with the bad DATE values not inserted. Since these records are part of the RiverStat table which maintains a foreign key relationship with other tables, the database restore will fail due to foreign key constraint violations.

In all cases, the bad DATE value shows up as a DATE of

5881610-07-12

The bad DATE value appears in both the ratedat and rrevise fields of the RiverStat table.

The SQL statements

UPDATE RiverStat SET ratedat = NULL where ratedat = '5881610-07-12';

UPDATE RiverStat SET rrevise = NULL where rrevise = '5881610-07-12';

will change all bad values of ratedat and rrevise to NULL. Doing this before the dump of the database will allow the records in the RiverStat table to be restored.

Null Check Syntax

In V7.4 in ecpg, a WHERE clause which checks for a NULL as … = '' (i.e. empty single quotes)

will not work in V8.2. This code will compile properly but will fail with a runtime error when executed. It and must be replaced with … IS NULL

Creating or Migrating Databases on the 8.2.x Server

One difference between the 7.4.8 server and the new 8.2.x server is the default encoding type for databases. The 7.4.8 server had encoding type as "SQL_ASCII" as its default. The 8.2.x server uses "UTF-8" as its default type. For an explanation of these encoding types, see Section 21.2.2 of the postgres 8.2.x documentation.

The default encoding type can be changed in the postgresql.conf file or it can be defined when postgres is initialized using the initdb command. At OHD, the default value of "UTF-8" is used. At AWIPS sites, the encoding type is "SQL_ASCII".

At OHD, a consequence of the above change in default encoding is that dumping some databases on the 7.4.8 server and then attempting to restore them on the 8.2.x server MAY RESULT IN LOSS OF DATA RECORDS. Records in tables such as the Descrip and Observer tables having special characters such as "/" and "#" will not be handled properly when an insert is attempted on the 8.2.x server. This problem will result in all records being lost in these tables.

To get around this problem, the user must explicitly set the encoding type to "SQL_ASCII" using the "-E" option in the CREATEDB statement when creating the database on the 8.2.x server. For example

createdb -U postgres -E SQL_ASCII hd_ob83fwr

will create a database which can be loaded using the "psql" command as has been done previously without loss of data.

This problem does not occur at AWIPS sites which use "SQL_ASCII" as the default encoding. At AWIPS sites, databases created using the "createdb" command do not need to include the -E" option.

PostgreSQL Version 8.2.x New Features

1) Copying records from a table based on a SELECT query. This feature was available in Informix.

To copy the lid and obsdate fields for records in the hourlypp table with
obsdate > 2007-06-21 to a file named "file.unl":

psql db_name
\copy (select lid, obsdate from hourlypp where obsdate > '2007-06-21') to file.unl

To copy the lid and ts fields for all records in the hourlypp table to a file named "file.unl":

psql db_name
\copy hourlypp (lid, obsdate) to file.unl

2) By default, Version 8.2.x treats the "\" character as an escape character. Version 7.4.x treated "\" as a literal. The TextProduct deletion portion of the db_purge application in some cases pulls in "\" as part of an SQL delete. This causes a warning message to be printed in the postgres log. To prevent this warning and to force Version 8.2.x to treat a "\" as a literal, the following statement was added to set_hydro_env:

>   export   PGOPTIONS='-c standard_conforming_strings=on'

See Section 17.12.1 of the postgres documentation for an explanation of the "standard_conforming_strings" parameter.

<div align="center">ecpg</div>

Version 8.2.6 includes ecpg version 4.2.1

Upgrading from V7.4 to V8.2.x caused the following problems to surface when attempting the buildall:

1) Following functions are not in ecpglib.h file

>   ECPGis_informix_null   ECPGset_informix_null

However, the original Informix names of these functions

>   risnull   rsetnull

are available. The code was changed to use these names and the link succeeded. See Section 31.9.2 of the documentation. Note that the risnull/rsetnull functions call the postgres functions ECPGis_noind_null/ECPGset_noind_null.

2) Initially, OHD installed version 8.2.4 of postgres. The file ecpg.h was found to have the typedef statement for the dtime_t datatype missing. This caused compilation of our generated code to fail. The following steps were taken to correct the problem:

>   - problem was corrected by postgres people – they sent us updated code
>   – Jim Williams created new COTS tarball
>   - Randy Brown installed new COTS package on lx5

<div align="center">Version 8.2.4 vs 8.2.5 vs 8.2.6</div>

When OHD first began building software for OB8.3 in August 2007, the latest postgres version was 8.2.4. This version was found to have a serious bug in the ecpg code related to the dtime_t type which prevented our software from compiling. The postgres people sent us a new version with a fix. This new version was installed and found to work properly. Subsequently, postgres released a newer version (8.2.5) which contained this fix along with other bug fixes. This newer version was initially installed at OHD on Sep 19, 2007.

Version 8.2.6 was released in early January 2008. This version of postgres contained security patches and some ecpg parser changes. The SwEG has decided to implement version 8.2.6 with the AWIPS OB8.3 software build.

Accessing the postgres Server for OB9.0 at OHD

In OHD in August 2008, a virtual machine named genessee was created to host the postgres server. The server is running Version 8.2.6 of postgres.

Users needing to access the server should setup their environment as follows:

export PGHOST=genessee
export PGUSER=pguser
export PGPORT=5432

Executables such as psql and ecpg are located in /usr/bin. Make sure that your PATH env variable is set up to access this directory.

Tables Defined With oids

NWRFC reports that using forms through the pgaccess utility does not work with postgres 8.2.5. They discovered that pgaccess requires that the tables be created with oids (Object Identifier Types). By default, version 8.2.x creates tables without oids.

Version 7.4.x of postgres creates tables WITH OIDS by default. Version 8.0.x changed this behavior to create tables WITHOUT OIDS by default. Oids are unsigned 4-byte integers. Each record in a table is assigned an oid if the table is created with the "WITH OIDS" option. This adds 4 bytes per record to the size of the tables. Generally, oids are of limited value and this is why the default behavior was changed.

The default behavior of creating tables with or without oids is controlled by the system configuration file

/data/db/pgdata/postgresql.conf

The file contains the line

#default_with_oids = off

which controls this behavior.

The CREATE TABLE statement has "WITH OIDS" and "WITHOUT OIDS" options allowing individual tables in a database to be created with or without oids.

----
ecpg current_timestamp Usage

In V7.4, the timezone associated with the current_timestamp function was GMT. In V8.2, it changed to the local time zone. The env variable PGTZ must be set to "GMT" to get the previous behavior for current_timestamp.

Dump/Restore for Transition

A dump/restore of databases is needed to transition from V7.4.8 to V8.2.6.

Referring to the following web page:
http://developer.postgresql.org/pgdocs/postgres/release.html

For 8.2.6, it states:
E.1.1. Migration to version 8.2.6
A dump/restore is not required for those running 8.2.X.

For 8.2, it states:
E.6.2. Migration to version 8.2
A dump/restore using pg_dump is required for those wishing to migrate data from any previous release.