

作者： 广东省东莞市长安中学 文合平(wenheping at gmail.com)

一 什么是 PostgreSQL 和 PostGIS

1.1 什么是 PostgreSQL

说起数据库，大家耳熟能详的商业数据库产品当推 Oracle、微软的 SqlServer 和 IBM 的 DB2 等，而开源数据库中则有两大产品 MySQL 和 PostgreSQL。

PostgreSQL 是一种对象-关系型数据库管理系统（ORDBMS），也是目前功能最强大、特性最丰富和最复杂的自由软件数据库系统。它起源于伯克利（BSD）的数据库研究计划，目前是最重要的开源数据库产品开发项目之一， 有着非常广泛的用户。

PostgreSQL 可以说是最富特色的自由数据库管理系统，也有人认为可以是最强大的自由软件数据库管理系统。PostgreSQL 是唯一支持事务、子查询、多版本并行控制系统、数据完整性检查等特性的唯一的一种自由软件的数据库管理系统。能在多平台下---包括 Linux、FreeBSD 和 Windows 等---运行，并且支持多语言的开发。

在两大开源数据库产品的对比中，一般认为 MySQL 速度更快，所以得到更为广泛的使用；而 PostgreSQL 性能更为先进，PostgreSQL 提供很多 MySQL 目前所不支持的特性，比如触发器、视图、存储过程等等，在记录数超千万之后性能表现尤其出色。

当前的最新版本是 PostgreSQL 8.2.3，官方网站是：

www.postgresql.org

1.2 什么是 PostGIS

PostGIS 在对象关系型数据库 PostgreSQL 上增加了存储管理空间数据的能力，相当于 Oracle 的 spatial 部分。PostGIS 最大的特点是符合并且实现了 OpenGIS 的一些规范，是最著名的开源 GIS 数据库。

当前的最新版本是 PostGIS 1.2.1，官方网站是：

www.postgis.org

二 PostgreSQL 和 PostGIS 的安装

2.1 在 windows 下安装。

新版本的 PostgreSQL 在其安装程序中集成了 PostGIS，只需要在安装过程中选中 PostGIS 和 pgsql 项就可以了。

2.2 在 Debian 下安装 PostGIS

```
# apt-get install postgresql-8.1 postgresql-8.1-postgis
```

当前 Etch 中的版本：PostgreSQL 是 8.1.7，而 PostGIS 是 1.1.6，虽然不是最新的版本，但已经足够了。

还需要做的工作是：

1) 创建一个专门用于 PostGIS 的数据库并为专用数据库导入 PostGIS 支持：

```
# su - postgres
$ cd /usr/share/postgresql-8.1-postgis
$ createdb wen1
$ createlang plpgsql wen1
$ psql -d wen1 -f lwpostgis.sql
$ psql -d wen1 -f spatial_ref_sys.sql
```

2) 创建一个专门的用户，并把相应的数据库和数据表的属主设置为该用户：

```
# su - postgres
$ psql
$ create user wen1 password '123456';
$ alter database wen1 owner to wen1;
$ alter table spatial_ref_sys owner to wen1;
$ alter table geometry_columns owner to wen1 ;
```

3) 修改 PostgreSQL 配置文件以便可以远程使用。

修改 PostgreSQL 的配置文件/etc/postgresql/8.1/main/Postgresql.conf，将其中的一句：

```
listen_address='localhost'
```

前面的注释去掉，并把 'localhost' 改为 '*'。

修改 Postgresql 的配置文件/etc/postgresql/8.1/main/pg_hba.conf，在文件后面加一句：

```
host all all 192.168.0.0/24 password
```

这句话的意思是：同网络中 192.168.0.* 的机器可以以密码的形式使用所有的数据库。更具体的参数意义直接看该配置文件中的注释就可以了。

这里一定要配置正确，否则无法在远程连接 PostgreSQL 数据库。

2.3 源码安装 PostgreSQL 和 PostGIS

参阅以前写的老文章。

三 PostGIS 的使用

要使用 PostGIS，需要两方面的预备知识，一是要熟悉基本的 SQL 语法；二是熟悉 PostgreSQL 数据库的基本使用。

(一) 快速入门

我们使用前面创建的数据库 wen2，用户为 wen1，现在创建一个包含三个点的数据库 test1：

首先在 Etch 下以 wen1 登陆，然后打开 wen2 数据库：

```
$ psql -d wen2 ;
```

再在 PostgreSQL 的控制台下输入以下命令：

```
wen2=> create table test1 (myID int4, pt geometry, myName varchar );
wen2=> insert into test1 values (1, 'POINT(0 0)', 'beijing' );
wen2=> insert into test1 values (2, 'POINT(31.5 60.87)', 'shanghai' );
wen2=> insert into test1 values (3, 'POINT(10.77 85.902)', 'tianjin' );
```

这样我们利用 PostGIS 创建了一个包含三个点的 GIS 数据表。

为了能在 QGIS 中打开这一图层，我们还需要为数据表创建一个主键：

```
wen2=> alter table test1 add primary key (myID);
```

可以看到 PostGIS 的使用其实就相当于使用一个经过扩展的 SQL 语法，上述语句熟悉 SQL 语法的人一看都很熟悉，都是普通的 SQL 语句，不同的只不过是增加了 PostGIS 特殊的 geometry 数据类型。

你可以再试试这些 SQL 语句：

```
select * from test1;
```

```
select myID,AsText(pt) from test1;  
select Distance(pt, 'POINT(0 0)') from test1;
```

(二) PostGIS 的 Geometry 数据类型

Geometry 可以说是 PostGIS 最重要的一个概念，是“几何体”的意思，由于 PostGIS 很好地遵守 OGC 的“Simple Feature for Specification for SQL”规范，目前支持的几何体类型包含其实例有：

```
POINT(1 1)  
MULTIPOINT(1 1, 3 4, -1 3)  
LINESTRING(1 1, 2 2, 3 4)  
POLYGON((0 0, 0 1, 1 1, 1 0, 0 0))  
MULTIPOLYGON((0 0, 0 1, 1 1, 1 0, 0 0), (5 5, 5 6, 6 6, 6 5, 5 5))  
MULTILINESTRING((1 1, 2 2, 3 4),(2 2, 3 3, 4 5))
```

而 geometry 具体表现又有两种形式，一种叫做 WKT(Well Known Text)形式，如上面的例子。或者使用如下 SQL 语句浏览：

```
select AsText(pt) from test1;
```

另一种叫做“Canonical Form”形式，看上去是一串古怪的数字，其实是一种增强的十六进制编码，使用如下 SQL 语句就可以浏览了：

```
select pt from test1;
```

(三) 读写 PostGIS 数据

建设好 PostGIS 数据库之后，我们现在需要进行读写 GIS 数据了，包括把 GIS 写入到 PostGIS 数据库中以及在应用程序中使用 PostGIS 数据库的数据。读写 PostGIS 目前主要有以下四种方式：

3.1 使用 psql 语言

Psql 语言是 PostgreSQL 内嵌的一个命令行工具，其语法基本上和标准的 SQL 语法是一致的，可以使用 Psql 工具，结合标准 SQL 语法和一些 PostGIS 的扩展对 PostGIS 数据库进行读写操作。

具体例子如上面“快速入门”中的例子。

这种方式功能强大，但全部需要手工操作，烦琐且繁重。

3.2 使用一些小工具

有两个很有用的小的转换工具，一是 shp2pg; 一是 ogr2ogr

3.2.1 shp2pgsql 和 pgsql2shp

shp2pgsql 和 pgsql2shp 是 PostGIS 自身携带的一对在 Shape 文件和 PostGIS 数据库之间进行转换的工具，在 Debian 中安装好 PostGIS 之后，这两个程序已经位于可执行文件的搜索路径上，因此可以在系统中任何位置使用。

假如当前目录下有完整的 china.shp 文件(应该有三个同名不同扩展名的文件)，需要把其中数据输入到上述的 wen2 数据库中的数据表 china_postgis 中，具体操作为(操作用户为 wen1)：

```
$ shp2pgsql china china_postgis > tmp.sql
```

```
$ psql -d wen2 -f roads.sql
```

这个工具很好用，缺点在于使用范围有限，只针对 Shape 文件格式。

3.2.2 ogr2ogr

PostGIS 本身的 shp2pg 工具只把 shape 文件转换到 PostGIS 数据库中，那如何把其他的 GIS 数据加入呢？比如说 MapInfo 的 mid 格式，ESRI 的 e00 格式呢？这就要使用 ogr 这个工具了。

Ogr 目前是 gdal 的一个组成部分，那什么是 gdal 呢？其官方主页(<http://www.gdal.org>)上如此介绍：

GDAL is a translator library for raster geospatial data formats that is released under an X/MIT style Open Source license by the Open Source Geospatial Foundation. As a library, it presents a single abstract data model to the calling application for all supported formats. It also comes with a variety of useful commandline utilities for data translation and processing.

简单地说，Gdal 是一个各种 Gis 数据格式的转换软件库，ogr 则是转换矢量 GIS 数据的软件库。

目前 ogr 能够支持的数据格式包括：

Arc/Info Binary Coverage、DWG、ESRI Personal GeoDatabase、ArcSDE、ESRI Shapefile、GML、GRASS、Mapinfo File、Microstation DGN、ODBC、Oracle Spatial 和 PostgreSQL 等。应该说，这就基本包括了我们平常用到的所有矢量型 GIS 文件格式了。

Gdal 支持的栅格数据格式参阅 http://www.gdal.org/formats_list.html

下面我们从源码编译安装 gdal---因为我在 Debian 下使用 apt 方式安装的 gdal 并不支持 postgresql 数据库，其实更为简便的方法也许是去 gdal.org 下载一个 Fwtools 的工具包，可以直接运行不用繁琐的编译，既有 Windows 的，也有 Linux 下用的，只不过这个包有些大。

1) 下载最新的 gdal-1.3.2

2) 解压

3) `./configure --prefix=/opt/gdal --with-pg=/opt/pg/bin/pg_config --without-ogdi`

这里我习惯把软件安装在/opt 目录下。

--with-pg 参数很重要，使 ogr 可以支持 PostGIS，后面的参数是我的 PostGIS 安装的地方。

我在这里编译很多次失败，经过仔细查找，发现问题出在 ogdi 上，所以暂时我只好使它不支持 ogdi。

4) make

5) make install

下面是使用过程，假设现在我要把一个 rai.mid 文件导入到数据库 data1 中：

```
$ ogr2ogr -f Postgresql PG:dbname=data1 rai.mid
```

注意，f 参数后面跟的是导入的数据类型，最后那个文件才是要导出的文件。PG 后面说明的是数据库的名字，需要是已经创建好的数据库，而数据表则由程序自动创建。还要特别注意权限问题，以上命令我是由 postgres 用户执行的。

还有一个要注意的是主键问题。一般情况下，你使用 shp2pg 或者 ogr2ogr 向 postgis 中增加了 Gis 数据后，在 GIS 客户端添加 postgis 图层的时候，常常会因为数据表没有定义主键而出错，解决的办法是，在服务器上使用 psql 或者其他 sql 工具，使用命令：

```
alter table table_name add primary key (column_name);
```

另外我们发现 ogr 竟然是不支持最常见的 GIS 格式 e00 格式，好象是因为版权方面的原因吧。解决的方法就是去 sf.net 查找相关工具。我在上面就找到一个 e002pg 工具，支持将

文件导入到 postGIS 数据库中。

3.3 在其他 GIS 软件中读写 PostGIS 数据

比如在 QGIS 中，能够打开 PostGIS 图层，还有 SPIT 插件可以把 Shape 文件输入到 PostGIS 数据库中。

其他 GIS 软件如 uDig，Grass 等，甚至连 ArcInfo 都支持或部分支持读写 PostGIS 数据。

3.4 利用接口在应用程序中读写 PostGIS 数据

广大的开源 GIS 程序员几乎为每一种程序设计语言设计好了读写 PostGIS 的接口，如利用 PostgreSQL 的 JDBC 库，可以使用 Java 语言在程序中读写 PostGIS 数据；利用 libpq 库，可以使用 C 语言读写 PostGIS 数据。

(四) 使用 PostGIS 函数

4.1 PostGIS 函数

要能熟练使用 PostGIS，掌握 PostGIS 的所有函数是关键。通过掌握这些 PostGIS 函数，我们能够高效、稳定地处理各种地理数据。由于 PostGIS 的函数设计时都遵守 OpenGIS 规范，我们也可以把这些 PostGIS 叫做 OpenGIS 函数。

4.2 PostGIS 函数的分类

PostGIS 函数大致可以分为以下四类

1) 字段处理函数

这类函数当前一共有 3 个，分别是：

AddGeometryColumn(var1,var2,var3,var4,var5,var6): 为已有的数据表增加一个地理几何数据字段。Var1 代表数据表的模式(schema)的名字，一般是 public，也可以省略，则使用当前的缺省模式；var2 是已有的数据表的名字；var3 是新的地理数据字段的名字；var4 是 SRID 值，不确定的话就取-1 吧；var5 是地理数据的类型，可以是 POINT 等；var6 是指该几何数据是二维还是三维数据。

前面的 SQL 语句 create table test1 (myID int4, pt geometry, myName varchar)更规范的写法为：

```
create table test1 (myID int4, myName varchar );
select AddGeometryColumn('test1','pt',-1,'GEOMETRY',2);
```

DropGeometryColumn 函数显然是删除一个地理数据字段的；

SetSRID 函数显然是设置 SRID 值的。

2) 几何关系函数

这类函数目前共有 10 个，分别是：

Distance Equals Disjoint Intersects Touches Crosses Within Overlaps Contains Relate

3) 几何分析函数

这类函数目前共有 12 个，分别是：Centroid Area Length PointOnSurface Boundary Buffer ConvexHull Intersection SymDifference Difference GeomUnion MemGeomUnion

4) 读写函数

这类函数很多，主要是用于在各种数据类型之间的转换，尤其是在于 Geometry 数据类型与其他如字符型等数据类型之间的转换，函数名如 AsText、GeomFromText 等，其作用是显然的。

4.3 PostGIS 函数使用示例。

假设在我们的 wen2 数据库中，有两张表，一张为 guojia，是从“国家基础地理数据”网站下载的国界数据表，数据类型为 LINE；二为 shengqu_polygon，也从同一个网站下载，地理数据类型为多边形。

1) 查询我国边境线的长度：

```
wen1=> select sum(length(the_geom)) as lenth from guojie;
```

2) 查询我国面积最大的省区名字：

```
wen1=> select name area(the_geom) as myarea  
       from shengqu_polygon  
       order by myarea DESC  
       LIMIT 1;
```

(五) 使用 PostGIS 扩展函数

除了上述遵循 OpenGIS 的函数之外，PostGIS 还自行扩展了一些当前 OpenGIS 规范之外的函数，主要包括以下几类：

5.1 管理类函数

扩展的管理类函数主要包括一些软件版本查询函数，如 postgis_version()、postgis_geos_version()、postgis_proj_version() 函数等，分别查询当前的 PostGIS 的版本及其使用的 Geos 和 Proj 库的版本。

5.2 数据类型的输入输出函数

除了 OpenGIS 定义的地理数据类型之外，PostGIS 还对数据类型进行了扩展，这种扩展主要是两方面的扩展，一是把二维的数据向三维和四维扩展；二就是在 WKT 和 WKB 数据类型基础上扩展出 EWKT 和 EWKB 数据类型。

PostGIS 提供了在这些地理数据类型和常用数据类型如字符型、浮点型数据之间进行转换的函数。

5.3 量算函数

如 length3d 函数是对 length2d 函数的扩展。

5.4 几何操作函数

如 addBBox(geometry) 函数给所给的几何体加上一个边框。

如 simplify(geometry,tolerance) 函数可以对折线和多边形利用 Douglas-Peucker 算法进行一些节点进行删除，从而使表现的图形更简单而清晰，在网络传输数据时具有更高的效率。

5.5 操作符号

5.6 其他扩展函数

(六) 建立 PostGIS 索引

当数据库的记录增大的时候，如果没有建立索引的话，操作的效率就显著下降。PostGIS 建议当记录数超过几千的时候就应该建立索引，而 GIS 数据库一般都是海量数据，所以对 PostGIS 而言，索引就非常重要。