

# PostgreSQL 6.4.2 手册

---

# PostgreSQL 6.4.2 手册

涵盖正式发布的 v6.4

PostgreSQL 版权所有 (C) 1998, Postgres 全球开发组。

---

# 目录

|                             |      |
|-----------------------------|------|
| 摘要 .....                    | xiii |
| I. 教程 .....                 | 1    |
| 1. 引言 .....                 | 3    |
| 1.1. 什么是Postgres? .....     | 3    |
| 1.2. Postgres 简史 .....      | 3    |
| 1.3. 关于本发行版 .....           | 5    |
| 1.4. 资源 .....               | 5    |
| 1.5. 术语 .....               | 6    |
| 1.6. 记法 .....               | 6    |
| 1.7. Y2K 声明 .....           | 7    |
| 1.8. 版权与商标 .....            | 7    |
| 2. 体系结构 .....               | 9    |
| 2.1. Postgres 体系结构概念 .....  | 9    |
| 3. 从头开始 .....               | 10   |
| 3.1. 设置你的环境 .....           | 10   |
| 3.2. 启动交互式监控器 (psql) .....  | 10   |
| 3.3. 管理数据库 .....            | 11   |
| 4. 查询语言 .....               | 13   |
| 4.1. 交互式监视器 .....           | 13   |
| 4.2. 概念 .....               | 13   |
| 4.3. 创建新类 .....             | 13   |
| 4.4. 用实例填充类 .....           | 14   |
| 4.5. 查询类 .....              | 14   |
| 4.6. 重定向 SELECT 查询 .....    | 15   |
| 4.7. 类之间的连接 .....           | 15   |
| 4.8. 更新 .....               | 16   |
| 4.9. 删除 .....               | 16   |
| 4.10. 使用聚合函数 .....          | 16   |
| 5. Postgres SQL 的高级特性 ..... | 17   |
| 5.1. 继承 .....               | 17   |
| 5.2. 非原子值 .....             | 18   |
| 5.3. 时间旅行 .....             | 19   |
| 5.4. 更多高级特性 .....           | 20   |
| II. 用户指南 .....              | 21   |
| 6. 设置你的环境 .....             | 25   |
| 7. 管理数据库 .....              | 26   |
| 7.1. 数据库创建 .....            | 26   |
| 7.2. 备选数据库位置 .....          | 26   |
| 7.3. 访问数据库 .....            | 27   |
| 7.4. 删除数据库 .....            | 28   |
| 8. 数据类型 .....               | 30   |
| 8.1. 数字类型 .....             | 31   |
| 8.2. 货币类型 .....             | 32   |
| 8.3. 字符类型 .....             | 32   |
| 8.4. 日期/时间类型 .....          | 33   |
| 8.5. 布尔类型 .....             | 37   |
| 8.6. 几何类型 .....             | 37   |
| 8.7. IPv4 网络与主机地址 .....     | 39   |
| 9. 操作符 .....                | 41   |
| 9.1. 词法优先级 .....            | 41   |
| 9.2. 通用操作符 .....            | 42   |

|                           |     |
|---------------------------|-----|
| 9.3. 数值操作符 .....          | 42  |
| 9.4. 几何操作符 .....          | 43  |
| 9.5. 时间间隔操作符 .....        | 44  |
| 9.6. IP V4 操作符 .....      | 44  |
| 9.7. IP V4 操作符 .....      | 45  |
| 10. 函数 .....              | 46  |
| 10.1. 数学函数 .....          | 46  |
| 10.2. 字符串函数 .....         | 46  |
| 10.3. 日期/时间函数 .....       | 47  |
| 10.4. 几何函数 .....          | 48  |
| 10.5. IP V4 函数 .....      | 50  |
| 11. 类型转换 .....            | 51  |
| 11.1. 概述 .....            | 51  |
| 11.2. 操作符 .....           | 52  |
| 11.3. 函数 .....            | 54  |
| 11.4. 查询目标 .....          | 56  |
| 11.5. UNION 查询 .....      | 57  |
| 12. 索引和键 .....            | 59  |
| 13. 数组 .....              | 61  |
| 14. 继承 .....              | 63  |
| 15. 查询语言 .....            | 65  |
| 15.1. 概念 .....            | 65  |
| 15.2. 创建新类 .....          | 65  |
| 15.3. 用实例填充类 .....        | 65  |
| 15.4. 查询类 .....           | 66  |
| 15.5. 重定向 SELECT 查询 ..... | 66  |
| 15.6. 类之间的连接 .....        | 67  |
| 15.7. 更新 .....            | 68  |
| 15.8. 删除 .....            | 68  |
| 15.9. 使用聚合函数 .....        | 68  |
| 16. 磁盘存储 .....            | 69  |
| 17. psql .....            | 70  |
| 17.1. 分页到屏幕 .....         | 70  |
| 18. pgaccess .....        | 71  |
| 19. SQL 命令 .....          | 72  |
| ABORT .....               | 73  |
| ALTER TABLE .....         | 74  |
| ALTER USER .....          | 77  |
| BEGIN WORK .....          | 79  |
| CLOSE .....               | 80  |
| CLUSTER .....             | 81  |
| COMMIT .....              | 83  |
| COPY .....                | 84  |
| CREATE AGGREGATE .....    | 88  |
| CREATE DATABASE .....     | 91  |
| CREATE FUNCTION .....     | 93  |
| CREATE INDEX .....        | 96  |
| CREATE LANGUAGE .....     | 99  |
| CREATE OPERATOR .....     | 103 |
| CREATE RULE .....         | 107 |
| CREATE SEQUENCE .....     | 110 |
| CREATE TABLE .....        | 114 |
| CREATE TABLE AS .....     | 128 |

|                      |     |
|----------------------|-----|
| CREATE TRIGGER ..... | 129 |
| CREATE TYPE .....    | 131 |
| CREATE USER .....    | 135 |
| CREATE VIEW .....    | 138 |
| DECLARE .....        | 140 |
| DELETE .....         | 143 |
| DROP AGGREGATE ..... | 145 |
| DROP DATABASE .....  | 147 |
| DROP FUNCTION .....  | 148 |
| DROP INDEX .....     | 150 |
| DROP LANGUAGE .....  | 151 |
| DROP OPERATOR .....  | 153 |
| DROP RULE .....      | 155 |
| DROP SEQUENCE .....  | 156 |
| DROP TABLE .....     | 157 |
| DROP TRIGGER .....   | 159 |
| DROP TYPE .....      | 160 |
| DROP USER .....      | 162 |
| DROP VIEW .....      | 163 |
| EXPLAIN .....        | 165 |
| FETCH .....          | 167 |
| GRANT .....          | 170 |
| INSERT .....         | 174 |
| LISTEN .....         | 176 |
| LOAD .....           | 178 |
| LOCK .....           | 180 |
| MOVE .....           | 182 |
| NOTIFY .....         | 184 |
| RESET .....          | 186 |
| REVOKE .....         | 187 |
| ROLLBACK .....       | 190 |
| SELECT .....         | 191 |
| SELECT INTO .....    | 197 |
| SET .....            | 198 |
| SHOW .....           | 203 |
| UNLISTEN .....       | 205 |
| UPDATE .....         | 207 |
| VACUUM .....         | 209 |
| 20. 工具应用程序 .....     | 211 |
| createdb .....       | 212 |
| createuser .....     | 214 |
| destroydb .....      | 216 |
| destroyuser .....    | 218 |
| initdb .....         | 220 |
| initlocation .....   | 223 |
| pg_dump .....        | 225 |
| pg_dumpall .....     | 228 |
| psql .....           | 231 |
| III. 管理员指南 .....     | 239 |
| 21. 引言 .....         | 242 |
| 21.1. 资源 .....       | 242 |
| 21.2. 术语 .....       | 243 |
| 21.3. 记法 .....       | 243 |
| 21.4. Y2K 声明 .....   | 244 |

|                                   |     |
|-----------------------------------|-----|
| 21.5. 版权与商标 .....                 | 244 |
| 22. 平台移植 .....                    | 246 |
| 22.1. 当前支持的平台 .....               | 246 |
| 22.2. 未支持的平台 .....                | 248 |
| 23. 配置选项 .....                    | 249 |
| 23.1. 配置参数 (configure) .....      | 249 |
| 23.2. 构建参数 (make) .....           | 249 |
| 23.3. 区域支持 .....                  | 251 |
| 23.4. Kerberos 认证 .....           | 252 |
| 24. 安装 .....                      | 254 |
| 24.1. 运行 Postgres 的要求 .....       | 254 |
| 24.2. 安装过程 .....                  | 254 |
| 24.3. 玩转 Postgres .....           | 265 |
| 24.4. 下一步 .....                   | 266 |
| 24.5. 移植说明 .....                  | 266 |
| 25. 运行时环境 .....                   | 268 |
| 25.1. 在 Unix 上使用 Postgres .....   | 268 |
| 26. 系统布局 .....                    | 269 |
| 27. 使用 pg_options .....           | 271 |
| 28. 启动 postmaster .....           | 276 |
| 29. 添加和删除用户 .....                 | 277 |
| 30. 磁盘管理 .....                    | 278 |
| 30.1. 备选位置 .....                  | 278 |
| 31. 排障 .....                      | 279 |
| 32. 管理数据库 .....                   | 280 |
| 32.1. 创建数据库 .....                 | 280 |
| 32.2. 访问数据库 .....                 | 280 |
| 32.3. 删除数据库 .....                 | 281 |
| 33. 数据库恢复 .....                   | 282 |
| 34. 回归测试 .....                    | 283 |
| 34.1. 回归测试环境 .....                | 283 |
| 34.2. 目录布局 .....                  | 283 |
| 34.3. 回归测试过程 .....                | 284 |
| 34.4. 回归分析 .....                  | 285 |
| 35. 发行说明 .....                    | 287 |
| 35.1. 版本 6.4 .....                | 287 |
| 35.2. 版本 6.3.2 .....              | 291 |
| 35.3. 版本 6.3.1 .....              | 291 |
| 35.4. 版本 6.3 .....                | 292 |
| 35.5. 版本 6.2.1 .....              | 296 |
| 35.6. 版本 6.2 .....                | 297 |
| 35.7. 版本 6.1.1 .....              | 300 |
| 35.8. 版本 6.1 .....                | 300 |
| 35.9. 版本 6.0 .....                | 302 |
| 35.10. 版本 1.09 .....              | 305 |
| 35.11. 版本 1.02 .....              | 305 |
| 35.12. 版本 1.01 .....              | 306 |
| 35.13. 版本 1.0 .....               | 309 |
| 35.14. Postgres95 Beta 0.03 ..... | 310 |
| 35.15. Postgres95 Beta 0.02 ..... | 312 |
| 35.16. Postgres95 Beta 0.01 ..... | 312 |
| 35.17. 计时结果 .....                 | 312 |
| IV. 程序员指南 .....                   | 314 |
| 36. 引言 .....                      | 317 |

|                                           |     |
|-------------------------------------------|-----|
| 36.1. 资源 .....                            | 317 |
| 36.2. 术语 .....                            | 318 |
| 36.3. 记法 .....                            | 318 |
| 36.4. Y2K 声明 .....                        | 319 |
| 36.5. 版权与商标 .....                         | 319 |
| 37. 体系结构 .....                            | 321 |
| 37.1. Postgres 体系结构概念 .....               | 321 |
| 38. 扩展 SQL: 概览 .....                      | 324 |
| 38.1. 可扩展性如何运作 .....                      | 324 |
| 38.2. Postgres 类型系统 .....                 | 324 |
| 38.3. 关于 Postgres 系统目录 .....              | 324 |
| 39. 扩展 SQL: 函数 .....                      | 327 |
| 39.1. 查询语言 (SQL) 函数 .....                 | 327 |
| 39.2. 编程语言函数 .....                        | 329 |
| 40. 扩展 SQL: 类型 .....                      | 333 |
| 40.1. 用户定义的类型 .....                       | 333 |
| 41. 扩展 SQL: 操作符 .....                     | 335 |
| 42. 扩展 SQL: 聚合 .....                      | 336 |
| 43. Postgres 规则系统 .....                   | 338 |
| 43.1. 什么是查询树? .....                       | 338 |
| 43.2. 视图和规则系统 .....                       | 339 |
| 43.3. INSERT、UPDATE 和 DELETE 上的规则 .....   | 348 |
| 43.4. 规则和权限 .....                         | 358 |
| 43.5. 规则与触发器 .....                        | 359 |
| 44. 索引扩展接口 .....                          | 362 |
| 45. GiST Indices .....                    | 367 |
| 46. 链接动态装载的函数 .....                       | 369 |
| 46.1. ULTRIX .....                        | 370 |
| 46.2. DEC OSF/1 .....                     | 370 |
| 46.3. SunOS 4.x、Solaris 2.x 和 HP-UX ..... | 370 |
| 47. 触发器 .....                             | 372 |
| 47.1. 触发器创建 .....                         | 372 |
| 47.2. 与触发器管理器的交互 .....                    | 373 |
| 47.3. 数据更改的可见性 .....                      | 374 |
| 47.4. 示例 .....                            | 374 |
| 48. 服务器编程接口 .....                         | 378 |
| 48.1. 接口函数 .....                          | 378 |
| 48.2. 接口支持函数 .....                        | 386 |
| 48.3. 内存管理 .....                          | 399 |
| 48.4. 数据更改的可见性 .....                      | 400 |
| 48.5. 示例 .....                            | 400 |
| 49. 过程语言 .....                            | 403 |
| 49.1. 安装过程语言 .....                        | 403 |
| 49.2. PL/pgSQL .....                      | 404 |
| 49.3. PL/Tcl .....                        | 412 |
| V. 接口 .....                               | 418 |
| 50. 函数 .....                              | 421 |
| 51. Large Objects .....                   | 422 |
| 51.1. 历史注记 .....                          | 422 |
| 51.2. Inversion 大对象 .....                 | 422 |
| 51.3. 大对象接口 .....                         | 422 |
| 51.4. 内建已注册函数 .....                       | 423 |
| 51.5. 从 LIBPQ 访问大对象 .....                 | 424 |
| 51.6. 示例程序 .....                          | 424 |

|                                      |     |
|--------------------------------------|-----|
| 52. ecpg - C 中的嵌入式 SQL .....         | 429 |
| 52.1. 为什么用嵌入式 SQL? .....             | 429 |
| 52.2. 概念 .....                       | 429 |
| 52.3. 如何使用 ecpg .....                | 429 |
| 52.4. 限制 .....                       | 432 |
| 52.5. 从其他 RDBMS 软件包移植 .....          | 432 |
| 52.6. 安装 .....                       | 432 |
| 52.7. 给开发者 .....                     | 432 |
| 53. libpq .....                      | 437 |
| 53.1. 数据库连接函数 .....                  | 437 |
| 53.2. 查询执行函数 .....                   | 440 |
| 53.3. 异步查询处理 .....                   | 443 |
| 53.4. 快速路径 .....                     | 445 |
| 53.5. 异步通知 .....                     | 446 |
| 53.6. 与 COPY 命令相关的函数 .....           | 446 |
| 53.7. libpq 跟踪函数 .....               | 448 |
| 53.8. libpq 控制函数 .....               | 448 |
| 53.9. 用户认证函数 .....                   | 449 |
| 53.10. 环境变量 .....                    | 449 |
| 53.11. 注意事项 .....                    | 450 |
| 53.12. 示例程序 .....                    | 450 |
| 54. pgtcl .....                      | 459 |
| 54.1. 命令 .....                       | 459 |
| 54.2. 示例 .....                       | 459 |
| 54.3. pgtcl 命令参考信息 .....             | 460 |
| 55. ODBC 接口 .....                    | 479 |
| 55.1. 背景 .....                       | 479 |
| 55.2. Windows应用 .....                | 479 |
| 55.3. Unix 安装 .....                  | 480 |
| 55.4. 配置文件 .....                     | 483 |
| 55.5. ApplixWare .....               | 484 |
| 56. JDBC 接口 .....                    | 491 |
| 56.1. 构建 JDBC 接口 .....               | 491 |
| 56.2. 为 JDBC 准备数据库 .....             | 491 |
| 56.3. 使用驱动 .....                     | 492 |
| 56.4. 导入 JDBC .....                  | 492 |
| 56.5. 载入驱动 .....                     | 492 |
| 56.6. 连接到数据库 .....                   | 493 |
| 56.7. 发出查询并处理结果 .....                | 493 |
| 56.8. 执行更新 .....                     | 494 |
| 56.9. 关闭连接 .....                     | 494 |
| 56.10. 使用大对象 .....                   | 494 |
| 56.11. Postgres 对 JDBC API 的扩展 ..... | 495 |
| 56.12. 延伸阅读 .....                    | 531 |
| VI. 开发者指南 .....                      | 532 |
| 57. 架构 .....                         | 534 |
| 57.1. Postgres 体系结构概念 .....          | 534 |
| 58. 数据库系统中的遗传查询优化 .....              | 537 |
| 58.1. 将查询处理看成是一个复杂的优化问题 .....        | 537 |
| 58.2. 遗传算法 (GA) .....                | 537 |
| 58.3. Postgres 中的遗传查询优化 (GEQO) ..... | 538 |
| 58.4. Postgres GEQO 的未来实现任务 .....    | 538 |
| 59. 前端/后端协议 .....                    | 540 |

|                                |     |
|--------------------------------|-----|
| 59.1. Overview .....           | 540 |
| 59.2. Protocol .....           | 540 |
| 59.3. Message Data Types ..... | 545 |
| 59.4. 消息格式 .....               | 545 |
| 60. Postgres 信号 .....          | 553 |
| 61. gcc 默认优化 .....             | 555 |
| 62. 后端接口 .....                 | 556 |
| 62.1. BKI 文件格式 .....           | 556 |
| 62.2. 通用命令 .....               | 556 |
| 62.3. 宏命令 .....                | 557 |
| 62.4. 调试命令 .....               | 557 |
| 62.5. 示例 .....                 | 558 |
| 63. 页文件 .....                  | 559 |
| 63.1. 页结构 .....                | 559 |
| 63.2. 文件 .....                 | 559 |
| 63.3. 缺陷 .....                 | 560 |
| VII. 附录 .....                  | 561 |
| A. 文档 .....                    | 563 |
| A.1. 文档路线图 .....               | 563 |
| A.2. 文档源码 .....                | 563 |
| A.3. 文档项目 .....                | 564 |
| A.4. 构建文档 .....                | 565 |
| A.5. v6.4 的硬拷贝生成 .....         | 566 |
| A.6. 工具集 .....                 | 567 |
| A.7. 备用工具集 .....               | 571 |
| B. 联系方式 .....                  | 572 |
| SQL 参考文献 .....                 | 573 |

---

## 插图清单

|                              |     |
|------------------------------|-----|
| 2.1. 如何建立一个连接 .....          | 9   |
| 26.1. Postgres 文件布局 .....    | 269 |
| 37.1. 如何建立连接 .....           | 322 |
| 38.1. Postgres 的主要系统目录 ..... | 325 |
| 57.1. 连接是如何建立的 .....         | 535 |

---

## 表格清单

|                               |     |
|-------------------------------|-----|
| 8.1. 数据类型 .....               | 30  |
| 8.2. 常量 .....                 | 31  |
| 8.3. 数值 .....                 | 31  |
| 8.4. 货币 .....                 | 32  |
| 8.5. 字符 .....                 | 32  |
| 8.6. 特殊字符 .....               | 33  |
| 8.7. 日期/时间 .....              | 33  |
| 8.8. 范围 .....                 | 33  |
| 8.9. 样式 .....                 | 34  |
| 8.10. 顺序 .....                | 34  |
| 8.11. 常量 .....                | 35  |
| 8.12. 布尔值 .....               | 37  |
| 8.13. 几何 .....                | 37  |
| 8.14. IPV4 .....              | 40  |
| 8.15. Postgres IP 类型示例 .....  | 40  |
| 9.1. 操作符排序（优先级递减） .....       | 41  |
| 9.2. Operators .....          | 42  |
| 9.3. Operators .....          | 42  |
| 9.4. Operators .....          | 43  |
| 9.5. Operators .....          | 44  |
| 9.6. Operators .....          | 44  |
| 9.7. Operators .....          | 45  |
| 10.1. 数学函数 .....              | 46  |
| 10.2. SQL92 字符串函数 .....       | 46  |
| 10.3. 字符串函数 .....             | 46  |
| 10.4. 日期/时间函数 .....           | 47  |
| 10.5. 几何函数 .....              | 48  |
| 10.6. 几何类型转换函数 .....          | 49  |
| 10.7. 几何升级函数 .....            | 49  |
| 10.8. Postgres IP V4 函数 ..... | 50  |
| 19.1. 二进制复制文件的内容 .....        | 86  |
| 22.1. 支持的平台 .....             | 246 |
| 22.2. 不兼容 .....               | 248 |
| 23.1. Kerberos .....          | 253 |
| 38.1. 目录 .....                | 324 |
| 44.1. 索引 .....                | 362 |
| 44.2. B-tree .....            | 363 |
| 44.3. pg_amproc .....         | 365 |
| 54.1. pgctl 命令 .....          | 459 |
| 60.1. 信号 .....                | 553 |
| 63.1. Page Layout .....       | 559 |
| A.1. Postgres 文档产品 .....      | 563 |

---

## 范例清单

|                          |     |
|--------------------------|-----|
| 19.1. 循环重写规则组合的例子。 ..... | 108 |
|--------------------------|-----|

---

# 摘要

Postgres 最初在加州大学伯克利分校计算机科学系开发，首创了许多对象关系概念，这些概念目前正在一些商业数据库中变得可用。它提供 SQL92/SQL3 语言支持、事务完整性和类型可扩展性。PostgreSQL 是这套原始伯克利代码的公有领域开源后继版本。

---

# 部分 I. 教程

面向新用户的介绍。

---

# 目录

|                              |    |
|------------------------------|----|
| 1. 引言 .....                  | 3  |
| 1.1. 什么是Postgres? .....      | 3  |
| 1.2. Postgres 简史 .....       | 3  |
| 1.2.1. 伯克利 Postgres 项目 ..... | 3  |
| 1.2.2. Postgres95 .....      | 4  |
| 1.2.3. PostgreSQL .....      | 4  |
| 1.3. 关于本发行版 .....            | 5  |
| 1.4. 资源 .....                | 5  |
| 1.5. 术语 .....                | 6  |
| 1.6. 记法 .....                | 6  |
| 1.7. Y2K 声明 .....            | 7  |
| 1.8. 版权与商标 .....             | 7  |
| 2. 体系结构 .....                | 9  |
| 2.1. Postgres 体系结构概念 .....   | 9  |
| 3. 从头开始 .....                | 10 |
| 3.1. 设置你的环境 .....            | 10 |
| 3.2. 启动交互式监控器 (psql) .....   | 10 |
| 3.3. 管理数据库 .....             | 11 |
| 3.3.1. 创建数据库 .....           | 11 |
| 3.3.2. 访问数据库 .....           | 11 |
| 3.3.3. 删除数据库 .....           | 12 |
| 4. 查询语言 .....                | 13 |
| 4.1. 交互式监视器 .....            | 13 |
| 4.2. 概念 .....                | 13 |
| 4.3. 创建新类 .....              | 13 |
| 4.4. 用实例填充类 .....            | 14 |
| 4.5. 查询类 .....               | 14 |
| 4.6. 重定向 SELECT 查询 .....     | 15 |
| 4.7. 类之间的连接 .....            | 15 |
| 4.8. 更新 .....                | 16 |
| 4.9. 删除 .....                | 16 |
| 4.10. 使用聚合函数 .....           | 16 |
| 5. Postgres SQL 的高级特性 .....  | 17 |
| 5.1. 继承 .....                | 17 |
| 5.2. 非原子值 .....              | 18 |
| 5.2.1. 数组 .....              | 18 |
| 5.3. 时间旅行 .....              | 19 |
| 5.4. 更多高级特性 .....            | 20 |

---

# 第 1 章 引言

本文档是最初在加州大学伯克利分校开发的 PostgreSQL<sup>1</sup> 数据库管理系统的用户手册。PostgreSQL 基于 Postgres release 4.2<sup>2</sup>。Postgres 项目由 Michael Stonebraker 教授领导，得到了国防高级研究计划局（DARPA）、陆军研究办公室（ARO）、国家科学基金会（NSF）以及 ESL 公司的赞助。

## 1.1. 什么是Postgres?

传统的关系数据库管理系统（DBMSs）支持的数据模型由一组命名的关系组成，关系中包含特定类型的属性。在当前的商业系统中，可用的类型包括浮点数、整数、字符串、货币和日期。人们普遍认识到，这个模型对于未来的数据处理应用是不够的。关系模型之所以能够成功地取代以前的模型，部分原因在于它的"斯巴达式的简洁"。然而，如前所述，这种简洁性常使某些应用的实现变得非常困难。Postgres通过以下面这种方式纳入下列四个附加基本概念，提供了可观的额外能力，使用户可以轻松地扩展系统：

- 类
- 继承
- 类型
- 函数

其他特性提供了额外的能力和灵活性：

- 约束
- 触发器
- 规则
- 事务完整性

这些特性使Postgres属于被称为对象-关系的数据库类别。注意，这不同于那些被称为面向对象的数据库，后者通常不太适合支持传统的关系数据库语言。因此，尽管 Postgres有一些面向对象的特性，它仍然坚定地站在关系数据库的世界中。事实上，一些商业数据库最近纳入了Postgres开创的特性。

## 1.2. Postgres 简史

### 1.2.1. 伯克利 Postgres 项目

Postgres DBMS 的实现始于 1986 年。该系统的最初概念见于The Design of Postgres，而最初数据模型的定义见于The Postgres Data Model。当时规则系统的设计见于The Design of the Postgres Rules System。存储管理器的设计动机和体系结构详见The Postgres Storage System。

Postgres 此后经历了多个主要版本。第一个"演示版"（demoware）系统于 1987 年开始运行，并在 1988 年的 ACM-SIGMOD 大会上进行了展示。我们在The Implementation of Postgres中描述了版本 1，并于 1989 年 6 月将其发布给少数外部用户。针对第一代规则系统所受到的批评（A Commentary on the Postgres Rules System），规则系统被重新设计（On Rules, Procedures, Caching and Views in Database Systems），版本 2 于 1990 年 6 月随新规则系统一同发布。

---

<sup>1</sup><http://postgresql.org/>

<sup>2</sup><http://s2k-ftp.CS.Berkeley.EDU:8000/postgres/postgres.html>

版本 3 出现于 1991 年，增加了对多种存储管理器、改进的查询执行器和重写的重写规则系统的支持。此后的版本大部分都集中在可移植性和可靠性上。

Postgres 已被用于实现许多不同的研究和生产应用。这些应用包括金融数据分析系统、喷气发动机性能监测软件包、小行星跟踪数据库、医学信息数据库以及多个地理信息系统。Postgres 还被多所大学用作教学工具。最后，Illustra Information Technologies<sup>3</sup>（后来并入 Informix<sup>4</sup>）接手了这套代码并将其商业化。Postgres 在 1992 年末成为 Sequoia 2000<sup>5</sup> 科学计算项目的主要数据管理器。此外，1993 年期间外部用户社区的规模几乎翻了一番。人们越来越清楚地意识到，对原型代码的维护和支持工作占用了大量本应用于数据库研究的时间。为了减轻这一支持负担，该项目在版本 4.2 时正式结束。

## 1.2.2. Postgres95

1994 年，Andrew Yu<sup>6</sup> 和 Jolly Chen<sup>7</sup> 为 Postgres 添加了一个 SQL 语言解释器，随后代码被发布到 Web 上自寻出路。Postgres95 是这份原始伯克利代码的公有领域、开源后裔。

Postgres95 是 Postgres（版本 4.2）最后一个官方版本的衍生品。代码现在完全采用 ANSI C，并且代码体积缩减了 25%。有许多内部更改提升了性能和代码可维护性。Postgres95 v1.0.x 在 Wisconsin Benchmark 上比 v4.2 快大约 30-50%。除了错误修复之外，这些就是主要的增强：

- 查询语言 Postquel 已被 SQL（在服务器中实现）取代。我们尚不支持子查询（可以用用户定义的 SQL 函数来模拟）。聚合函数已被重新实现。我们还增加了对“GROUP BY”的支持。libpq 接口仍然可用于 C 程序。
- 除了 monitor 程序之外，我们还提供了一个新程序（psql），它支持 GNU readline。
- 我们添加了一个新的前端库 libpgtcl，它支持基于 Tcl 的客户端。一个示例 shell pgctlsh 提供了新的 Tcl 命令，用于让 tcl 程序与 Postgres95 服务器交互。
- 大对象接口经过了全面改造。我们保留反转大对象作为存储大对象的唯一机制。（不要把它与已被移除的反转文件系统混淆。）
- 实例级规则系统已被移除。规则仍然可以作为重写规则使用。
- 源代码中附带了一份介绍常规 SQL 特性以及我们自有特性的简短教程。
- 构建使用的是 GNU make（而不是 BSD make）。此外，Postgres95 可以用未经修改的 gcc 编译（双精度浮点数的数据对齐问题已修复）。

## 1.2.3. PostgreSQL

到 1996 年，很明显“Postgres95”这个名字经不起时间的考验。我们选择了新名字 PostgreSQL，以体现原始 Postgres 与后来具备 SQL 能力的版本之间的关系。同时，版本号被重置为从 6.0 开始，使编号回到最初由 Postgres 项目开启的序列中。

Postgres95 v1.0.x 版本的开发重点是稳定后端代码。而在 PostgreSQL 的 v6.x 系列中，重点已经从识别和理解后端中存在的问题转移到增强特性和能力，尽管各方面的工作仍在继续。

主要增强包括：

<sup>3</sup><http://www.illustra.com/>

<sup>4</sup><http://www.informix.com/>

<sup>5</sup>[http://www.sdsc.edu/0/Parts\\_Collabs/S2K/s2k\\_home.html](http://www.sdsc.edu/0/Parts_Collabs/S2K/s2k_home.html)

<sup>6</sup><mailto:ayu@informix.com>

<sup>7</sup><http://http.cs.berkeley.edu/~jolly/>

- 实现了重要的后端特性，包括子查询、默认值、约束和触发器。
- 增加了额外的符合 SQL92 的语言特性，包括主键、带引号的标识符、字面串类型强制转换、类型转换以及二进制和十六进制整数输入。
- 内置类型得到了改进，包括新的大范围日期/时间类型和额外的几何类型支持。
- 后端代码的整体速度提升了大约 20-40%，而且自 v6.0 发布以来后端启动时间缩短了 80%。

## 1.3. 关于本发行版

PostgreSQL 是免费提供的。本手册描述的是 PostgreSQL 的 6.4 版。

我们把 Postgres 用作以 PostgreSQL 名称发行的版本的简称。

当前支持的机器列表请查阅管理员指南。一般来说，Postgres 可以移植到任何具有完整 libc 库支持的 Unix/Posix 兼容系统。

## 1.4. 资源

本手册集被组织为若干部分：

### 教程

面向新用户的介绍。不涉及高级特性。

### 用户指南

面向用户的一般信息，包括可用的命令和数据类型。

### 程序员指南

面向应用程序员的高级信息。主题包括类型和函数可扩展性、库接口以及应用设计问题。

### 管理员指南

安装和管理信息。受支持的机器列表。

### 开发者指南

面向 Postgres 开发者的信息。此书面向为 Postgres 项目做贡献的人；应用开发的信息见程序员指南。目前包含在程序员指南中。

### 参考手册

关于命令语法的详细参考信息。目前包含在用户指南中。

除本手册集之外，还有其他资源可以帮助你安装和使用 Postgres：

### man 手册页

man 手册页包含关于命令语法的一般信息。

### FAQ

常见问题（FAQ）文档既阐述一般性问题，也阐述一些平台特定的问题。

## README

某些附带的软件包提供了 README 文件。

## 网站

Postgres<sup>8</sup> 网站上可能有发行包中未包含的信息。那里有一个 mhonarc 邮件列表流量存档，是许多主题的丰富资源。

## 邮件列表

Postgres Questions<sup>9</sup> 邮件列表是一个让用户问题得到解答的好地方。还有其他可用的邮件列表；详情请参阅该网页。

## 你自己！

Postgres 是一项开源产品。因此，它的持续支持依赖于用户社区。当你开始使用 Postgres 时，你会依赖他人的帮助——无论是通过文档还是通过邮件列表。请考虑回馈你的知识。如果你学到了文档中没有的东西，把它写出来并贡献回去。如果你给代码添加了特性，也把它们贡献出去。即使经验不多的人也能为文档提供更正和小改动，而这是一个很好的起点。Postgres Documentation<sup>10</sup> 邮件列表就是开始的地方。

## 1.5. 术语

在后面的文档中，站点（site）可以解释为安装 Postgres 的主机。由于在单台主机上可以安装多套 Postgres 数据库，这个术语更确切地指任何一套特定的已安装的 Postgres 二进制文件和数据库。

Postgres 超级用户是名为 *postgres* 的用户，他拥有 Postgres 的二进制文件和数据库文件。作为数据库超级用户，所有的保护机制都可以被绕过，任何数据都可以任意访问。此外，Postgres 超级用户还被允许执行一些通常不向所有用户开放的支持程序。注意，Postgres 超级用户不是 Unix 超级用户（后者将称为 *root*）。出于安全原因，超级用户应具有非零的用户标识符（UID）。

数据库管理员（DBA）负责安装 Postgres 并为站点建立实施安全策略的机制。DBA 可以按下述方法添加新用户，并维护一组供 *createdb* 使用的模板数据库。

*postmaster* 是为发往 Postgres 系统的请求充当中转站的进程。前端应用连接到 *postmaster*，由它跟踪系统错误和后端进程之间的通信。*postmaster* 可以接受若干命令行参数来调整其行为。不过，只有当你打算运行多个站点或非默认站点时才需要提供参数。

Postgres 后端（实际的可执行程序 *postgres*）可以由 Postgres 超级用户直接在用户 shell 中执行（以数据库名作为参数）。但这样做会绕过与 *postmaster*/站点关联的共享缓冲池和锁表，因此在多用户站点中不建议这样做。

## 1.6. 记法

文件名前面的 “...” 或 `/usr/local/pgsql/` 用来表示 Postgres 超级用户主目录的路径。

在命令概要中，方括号 “[” 和 “]” 表示可选的短语或关键字。花括号 (“{” 和 “}”) 中含有竖线 (“|”) 时表示你必须选择其中一个。

<sup>8</sup> [postgresql.org](http://postgresql.org)

<sup>9</sup> <mailto:questions@postgresql.org>

<sup>10</sup> <mailto:docs@postgresql.org>

在示例中，圆括号（“(”和“)”）用于给布尔表达式分组。“|”是布尔运算符 OR。

示例将展示从不同账户和程序执行的命令。从 root 账户执行的命令前带有“>”。从 Postgres 超级用户账户执行的命令前带有“%”，而从无特权用户账户执行的命令前带有“\$”。SQL 命令前带有“=>”提示符，也可能没有前导提示符，视上下文而定。

### 注意

在撰写本文时（Postgres v6.4），整个文档集中标记命令的记法并不完全一致。如发现问题请报告给文档邮件列表<sup>11</sup>。

## 1.7. Y2K 声明

### 作者

由 Thomas Lockhart<sup>12</sup> 写于 1998-10-22。

PostgreSQL 全球开发队伍将 Postgres 软件代码树作为公共服务提供，不对它的行为或性能作任何保证，也不承担任何责任。不过，在撰写本文时：

- 本声明的作者自 1996 年 11 月起就是 Postgres 支持团队的一名志愿者，他不知道 Postgres 代码库中存在任何与 2000 年 1 月 1 日前后的时间过渡（Y2K）相关的问题。
- 本声明的作者没有听说在回归测试中或在 Postgres 近期或当前版本的其他实际使用中发现过任何 Y2K 问题的报告。考虑到已安装的用户基础以及用户在支持邮件列表上的积极参与，如果存在问题，我们本应有所耳闻。
- 就作者所知，Postgres 对用两位数字年份指定的日期所做的假定记录在当前用户指南<sup>13</sup>关于数据类型的章节中。对于两位年份，关键的过渡年份是 1970 而不是 2000；例如“70-01-01”被解释为“1970-01-01”，而“69-01-01”被解释为“2069-01-01”。
- 底层操作系统中与获取“当前时间”相关的任何 Y2K 问题都可能传播为 Postgres 中表面的 Y2K 问题。

关于 Y2K 问题的进一步讨论，特别是与开源、免费软件相关的讨论，请参阅 The Gnu Project<sup>14</sup> 和 The Perl Institute<sup>15</sup>。

## 1.8. 版权与商标

PostgreSQL 版权所有 (C) 1996-8，归 PostgreSQL 全球开发组所有，并按照 Berkeley 许可的条款分发。

Postgres95 版权所有 (C) 1994-5，归加利福尼亚大学董事会所有。特此授权，可为任何目的使用、复制、修改和分发本软件及其文档，无需付费，也无需书面协议，但条件是所有副本中都必须包含上述版权声明、本段以及后续两段内容。

<sup>11</sup><mailto:docs@postgresql.org>

<sup>12</sup><mailto:lockhart@alumni.caltech.edu>

<sup>13</sup><http://www.postgresql.org/docs/user/datatype.htm>

<sup>14</sup><http://www.gnu.org/software/year2000.html>

<sup>15</sup><http://language.perl.com/news/y2k.html>

在任何情况下，加利福尼亚大学都不对任何一方承担因使用本软件及其文档而产生的任何直接、间接、特殊、附带或后果性损害赔偿责任，包括利润损失，即使加利福尼亚大学已被告知此类损害发生的可能性。

加利福尼亚大学明确不作任何担保，包括但不限于适销性和特定用途适用性的默示担保。本协议项下提供的软件按"原样"提供，加利福尼亚大学没有义务提供维护、支持、更新、增强或修改。

UNIX 是 X/Open, Ltd. 的商标。Sun4、SPARC、SunOS 和 Solaris 是 Sun Microsystems, Inc. 的商标。DEC、DECstation、Alpha AXP 和 ULTRIX 是 Digital Equipment Corp. 的商标。PA-RISC 和 HP-UX 是 Hewlett-Packard Co. 的商标。OSF/1 是 Open Software Foundation 的商标。

---

## 第 2 章 体系结构

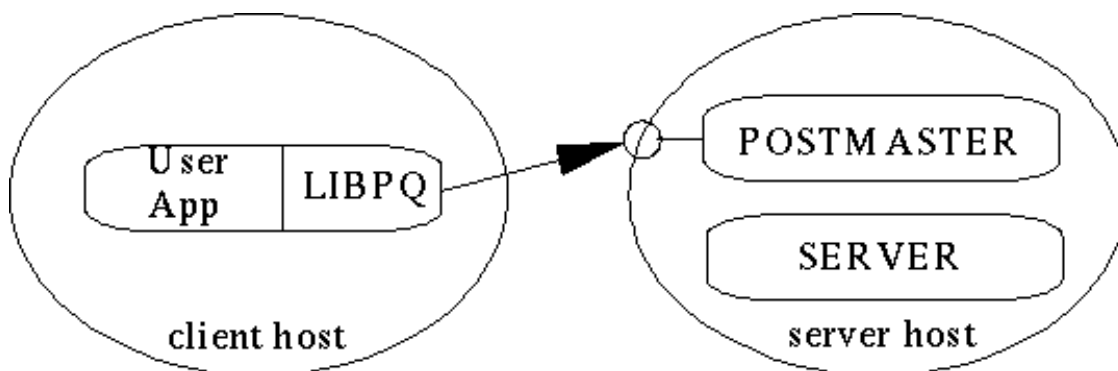
### 2.1. Postgres 体系结构概念

在我们开始之前，你应当理解基本的 Postgres 系统体系结构。理解 Postgres 的各个部分如何交互将使下一章更容易理解。用数据库术语来说，Postgres 使用一种简单的 "每用户一进程" 客户端/服务器模型。一个 Postgres 会话由下列相互协作的 UNIX 进程（程序）组成：

- 一个监管守护进程（postmaster），
- 用户的前端应用（例如 psql 程序），以及
- 一个或多个后端数据库服务器（postgres 进程本身）。

一个 postmaster 管理单个主机上一个给定的数据库集合。这样的数据库集合称为一个安装或站点。希望访问安装内某个给定数据库的前端应用对库进行调用。库通过网络把用户请求发送给 postmaster（图 2.1），后者转而启动一个新的后端服务器进程

图 2.1. 如何建立一个连接



并把前端进程连接到新的服务器。从那时起，前端进程和后端服务器进行通信而无需 postmaster 干预。因此，postmaster 总是在运行，等待请求，而前端和后端进程则来来去去。

libpq 库允许单个前端建立到多个后端进程的连接。但是，前端应用仍然是单线程进程。目前 libpq 不支持多线程的前端/后端连接。这种体系结构的一个含义是 postmaster 和后端总是在同一台机器（数据库服务器）上运行，而前端应用可以运行在任何地方。你应当牢记这一点，因为在客户机上可以访问的文件在数据库服务器机器上可能无法访问（或者只能用不同的文件名访问）。

你还应当知道，postmaster 和 postgres 服务器以 Postgres "超级用户" 的用户 ID 运行。注意 Postgres 超级用户并不非得是特殊用户（例如名为 "postgres" 的用户）。而且，Postgres 超级用户绝对不应当是 UNIX 超级用户（"root"）！任何情况下，与某个数据库相关的所有文件都应当属于这个 Postgres 超级用户。

---

## 第 3 章 从头开始

新用户如何开始使用 Postgres。

使用 Postgres 所需的一些步骤可以由任何 Postgres 用户执行，而另一些必须由站点数据库管理员完成。站点管理员是安装软件、创建数据库目录并启动 `postmaster` 进程的人。这个人不一定是 UNIX 超级用户（“root”）或计算机系统管理员；任何人都可以在没有任何特殊账户或权限的情况下安装和使用 Postgres。

如果你要自己安装 Postgres，请参阅管理员指南中的安装说明，安装完成后再回到本指南。

在本手册中，任何以字符 “%” 开头的示例都是应当在 UNIX shell 提示符下键入的命令。以字符 “\*” 开头的示例是 Postgres 查询语言 Postgres SQL 中的命令。

### 3.1. 设置你的环境

本节讨论如何设置你自己的环境，以便使用前端应用。我们假定 Postgres 已经成功安装并启动；关于如何安装 Postgres，请参阅管理员指南和安装说明。

Postgres 是一个客户端/服务器应用。作为用户，你只需要访问安装中的客户端部分（客户端应用的一个例子是交互式监控器 `psql`）。为简单起见，我们假定 Postgres 安装在目录 `/usr/local/pgsql` 中。因此，凡是看到目录 `/usr/local/pgsql` 之处，你都应该把它替换为 Postgres 实际安装的目录名。所有 Postgres 命令都安装在目录 `/usr/local/pgsql/bin` 中。因此，你应当把这个目录加入你的 shell 命令路径。如果你使用 Berkeley C shell 的变体（如 `csh` 或 `tcsh`），可以在你家目录的 `.login` 文件中加入

```
% set path = ( /usr/local/pgsql/bin path )
```

如果你使用 Bourne shell 的变体（如 `sh`、`ksh` 或 `bash`），则应当把

```
% PATH=/usr/local/pgsql/bin PATH
% export PATH
```

加入你家目录的 `.profile` 文件。从现在起，我们假定你已把 Postgres 的 `bin` 目录加入了路径。此外，本文档会频繁提到“设置 shell 变量”或“设置环境变量”。如果你没有完全理解上一段关于修改搜索路径的内容，应当先查阅描述你所用的 shell 的 UNIX 手册页，然后再继续。

如果你的站点管理员没有按默认方式完成设置，你可能还需要做一些额外工作。例如，如果数据库服务器所在的机器是远程机器，你就需要把 `PGHOST` 环境变量设置为数据库服务器机器的名称。环境变量 `PGPORT` 也可能需要设置。归根结底就是：如果你试着启动某个应用程序，而它抱怨无法连接到 `postmaster`，你应当立即咨询站点管理员，以确保你的环境已正确设置。

### 3.2. 启动交互式监控器（psql）

假定你的站点管理员已正确启动 `postmaster` 进程并授权你使用数据库，你（作为用户）就可以开始启动应用了。如前所述，你应当把 `/usr/local/pgsql/bin` 加入你的 shell 搜索路径。在大多数情况下，这就是你在准备方面需要做的全部工作。

从 Postgres v6.3 起，支持两种不同风格的连接。站点管理员或者选择允许 TCP/IP 网络连接，或者把数据库访问限制为仅本地（同机）套接字连接。如果你在连接数据库时遇到问题，这些选择就变得重要了。

如果你从某个 Postgres 命令（如 `psql` 或 `createdb`）得到下面的错误消息：

```
% psql template1
Connection to database 'postgres' failed.
connectDB() failed: Is the postmaster running and accepting
connections
    at 'UNIX Socket' on port '5432'?
```

or

```
% psql -h localhost template1
Connection to database 'postgres' failed.
connectDB() failed: Is the postmaster running and accepting TCP/IP
(with -i) connections at 'localhost' on port '5432'?
```

通常是因为 (1) postmaster 没有在运行，或者 (2) 你试图连接到错误的服务器主机。如果你得到下面的错误消息：

```
FATAL 1:Feb 17 23:19:55:process userid (2360) != database owner (268)
```

这表示站点管理员以错误的用户身份启动了 postmaster。让他以 Postgres 超级用户身份重新启动它。

## 3.3. 管理数据库

既然 Postgres 已经启动并运行，我们可以创建一些数据库来做实验。这里我们描述管理数据库的基本命令。

大多数 Postgres 应用假定，如果没有指定数据库名，它就与你的计算机账户名相同。

如果你的数据库管理员在设置你的账户时没有给它创建数据库的权限，那么她应当已经告诉你你的数据库名是什么。如果是这种情况，你可以跳过关于创建和删除数据库的小节。

### 3.3.1. 创建数据库

比如说你想创建一个名为 mydb 的数据库。你可以用下面的命令完成：

```
% createdb mydb
```

如果你没有创建数据库所需的权限，你会看到：

```
% createdb mydb
WARN:user "your username" is not allowed to create/destroy databases
createdb: database creation failed on mydb.
```

Postgres 允许你在一个给定站点上创建任意数量的数据库，并且你会自动成为你所创建数据库的数据库管理员。数据库名必须以字母开头，长度限制为 32 个字符。并非每个用户都有成为数据库管理员的授权。如果 Postgres 拒绝为你创建数据库，那么站点管理员需要授予你创建数据库的权限。如果发生这种情况，请咨询你的站点管理员。

### 3.3.2. 访问数据库

构建好数据库后，你可以通过以下方式访问它：

- 运行 Postgres 终端监控器程序（如 psql），它允许你交互式地输入、编辑并执行 SQL 命令。

- 用 C 语言编写使用 LIBPQ 子例程库的程序。这允许你从 C 提交 SQL 命令，并把答案和状态消息返回给你的程序。这个接口在 libpq 中进一步讨论。

你可能想启动 psql 来试一下本手册中的示例。为 mydb 数据库启动它只需键入命令：

```
% psql mydb
```

你会看到下面的欢迎消息：

```
Welcome to the POSTGRES interactive sql monitor:
Please read the file COPYRIGHT for copyright terms of POSTGRES

type \? for help on slash commands
type \q to quit
type \g or terminate with semicolon to execute query
You are currently connected to the database: template1

mydb=>
```

这个提示符表示终端监控器正在等待你的输入，你可以把 SQL 查询键入由终端监控器维护的工作区中。psql 程序会响应以反斜杠字符 “\” 开头的转义码。例如，键入下面的命令可以得到各种 Postgres SQL 命令语法的帮助：

```
mydb=> \h
```

把查询输入工作区后，你可以通过键入下面的命令把工作区的内容传递给 Postgres 服务器：

```
mydb=> \g
```

这会告诉服务器处理该查询。如果你用分号终止查询，则不需要 “\g”。psql 会自动处理以分号终止的查询。要从文件（比如 myFile）读取查询而不是交互式输入，键入：

```
mydb=> \i fileName
```

要退出 psql 并返回 UNIX，键入

```
mydb=> \q
```

这样 psql 就会退出并返回到命令 shell。（要了解更多转义码，在监控器提示符下键入 \h。）在 SQL 查询中可以自由使用空白（即空格、制表符和换行符）。单行注释用 “--” 表示。从双划线到行尾的所有内容都会被忽略。多行注释和行内注释用 “/\* ... \*/” 表示。

### 3.3.3. 删除数据库

如果你是数据库 mydb 的数据库管理员，你可以用下面的 UNIX 命令删除它：

```
% destroydb mydb
```

这个操作会从物理上删除与该数据库关联的所有 UNIX 文件，而且无法撤销，因此务必三思而后行。

---

## 第 4 章 查询语言

Postgres 查询语言是下一代标准草案 SQL3 的一个变种。它有许多扩展，例如可扩展的类型系统、继承、函数和产生式规则。这些都是从最初的 Postgres 查询语言 PostQuel 继承下来的特性。本节概述如何使用 Postgres SQL 执行简单的操作。本手册只是让你对我们这种风格的 SQL 有个概念，绝不是 SQL 的完整教程。关于 SQL 已有许多著作，包括 [MELT93] 和 [DATE97]。你应当知道，某些语言特性并不属于 ANSI 标准。

### 4.1. 交互式监视器

在后面的例子中，我们假设你已按上一小节的说明创建了 mydb 数据库并启动了 psql。本手册中的例子也可以在 `/usr/local/pgsql/src/tutorial/` 中找到。如何使用它们请参考该目录下的 README 文件。要开始教程，请执行以下操作：

```
% cd /usr/local/pgsql/src/tutorial
% psql -s mydb
Welcome to the POSTGRESQL interactive sql monitor:
  Please read the file COPYRIGHT for copyright terms of POSTGRESQL

  type \? for help on slash commands
  type \q to quit
  type \g or terminate with semicolon to execute query
  You are currently connected to the database: postgres

mydb=> \i basics.sql
```

`\i` 命令从指定的文件中读入查询。`-s` 选项让你进入单步模式，它会在把查询发送给后端之前暂停。本节使用的查询在文件 `basics.sql` 中。

psql 有多种用于显示系统信息的 `\d` 命令。更多细节请查阅这些命令；要查看列表，请在 psql 提示符下输入 `\?`。

### 4.2. 概念

Postgres 中的基本概念是类，它是一组有名字的对象实例。每个实例都有相同的命名属性集合，每个属性都有特定的类型。此外，每个实例都有一个永久的对象标识符（OID），它在整个安装范围内唯一。由于 SQL 语法说的是表，我们将把表和类作为同义词混用。同样，SQL 的行是实例，SQL 的列就是属性。如前所述，类被组织成数据库，而由单个 postmaster 进程管理的一组数据库构成一个安装或站点。

### 4.3. 创建新类

你可以通过指定类名以及所有属性的名称和类型来创建一个新类：

```
CREATE TABLE weather (
    city          varchar(80),
    temp_lo       int,           -- low temperature
    temp_hi       int,           -- high temperature
    prcp          real,         -- precipitation
    date          date
);
```

注意关键字和标识符都不区分大小写；标识符可以用双引号括起来以变得区分大小写，这是 SQL92 所允许的。Postgres 的 SQL 支持常见的 SQL 类型 `int`、`float`、`real`、`smallint`、`char(N)`、`varchar(N)`、`date`、`time` 和 `timestamp`，还有其他一些通用类型以及一套丰富的几何类型。我们后面会看到，Postgres 可以用任意数量的用户自定义数据类型来定制。因此，类型名不是语法上的关键字，除非为了支持 SQL92 标准中的特殊情形所必需。到目前为止，Postgres 的 `create` 命令看上去与传统关系系统中建表的命令完全一样。但我们马上就会看到，类还有一些性质是对关系模型的扩展。

## 4.4. 用实例填充类

`insert` 语句用于向类中填充实例：

```
INSERT INTO weather
VALUES ('San Francisco', 46, 50, 0.25, '11/27/1994')
```

你也可以用 `copy` 命令从平面 (ASCII) 文件装载大量数据。

## 4.5. 查询类

`weather` 类可以用普通的关系选择和投影查询来查询。这由 SQL 的 `select` 语句完成。该语句分为目标列表（列出要返回的属性的部分）和限定条件（指定任何限制的部分）两部分。例如，要检索 `weather` 的所有行，输入：

```
SELECT * FROM WEATHER;
```

输出应该是：

```
+-----+-----+-----+-----+-----+
|city      | temp_lo | temp_hi | prcp  | date      |
+-----+-----+-----+-----+-----+
|San Francisco | 46      | 50      | 0.25  | 11-27-1994 |
+-----+-----+-----+-----+-----+
|San Francisco | 43      | 57      | 0      | 11-29-1994 |
+-----+-----+-----+-----+-----+
|Hayward      | 37      | 54      |        | 11-29-1994 |
+-----+-----+-----+-----+-----+
```

你可以在目标列表中指定任意表达式。例如：

```
SELECT city, (temp_hi+temp_lo)/2 AS temp_avg, date FROM weather;
```

查询的限定条件中允许使用任意布尔运算符 (`and`、`or` 和 `not`)。例如，

```
SELECT * FROM weather
WHERE city = 'San Francisco'
AND prcp > 0.0;
```

```
+-----+-----+-----+-----+-----+
|city      | temp_lo | temp_hi | prcp  | date      |
+-----+-----+-----+-----+-----+
|San Francisco | 46      | 50      | 0.25  | 11-27-1994 |
+-----+-----+-----+-----+-----+
```

最后提一句，你可以指定选择的结果按排序顺序返回，或者去掉重复的实例。

```
SELECT DISTINCT city
FROM weather
ORDER BY city;
```

## 4.6. 重定向 SELECT 查询

任何 select 查询都可以重定向到一个新类：

```
SELECT * INTO TABLE temp FROM weather;
```

这构成一条隐式的 create 命令，用 select into 命令目标列表中指定的属性名和类型创建新类 temp。当然，之后我们可以像对其他类一样对结果类执行任何操作。

## 4.7. 类之间的连接

到目前为止，我们的查询一次只访问一个类。查询可以同时访问多个类，也可以用让类的多个实例同时被处理的方式访问同一个类。一次访问同一个或多个类的多个实例的查询称为连接查询。举个例子，假设我们希望找出所有位于其他记录温度范围之内的记录。实际上，我们需要把每个 EMP 实例的 temp\_lo 和 temp\_hi 属性与所有其他 EMP 实例的 temp\_lo 和 temp\_hi 属性进行比较。

### 注意

这只是一个概念模型。实际的连接可能以更高效的方式执行，但这对用户是不可见的。

我们可以用下面的查询做到：

```
SELECT W1.city, W1.temp_lo AS low, W1.temp_hi AS high,
       W2.city, W2.temp_lo AS low, W2.temp_hi AS high
FROM weather W1, weather W2
WHERE W1.temp_lo < W2.temp_lo
AND W1.temp_hi > W2.temp_hi;
```

| city          | low | high | city          | low | high |
|---------------|-----|------|---------------|-----|------|
| San Francisco | 43  | 57   | San Francisco | 46  | 50   |
| San Francisco | 37  | 54   | San Francisco | 46  | 50   |

### 注意

这种连接的语义是：限定条件是针对查询中所指各类的笛卡尔积定义的一个真值表达式。对于笛卡尔积中使限定条件为真的那些实例，Postgres 计算并返回目标列表中指定的值。Postgres 的 SQL 不对此类表达式中的重复值赋予任何意义。这意味着 Postgres 有时会多次重复计算同一目标列表；当布尔表达式用 "or" 连接时经常出现这种情况。要去掉这样的重复，必须使用 select distinct 语句。

在这个例子中，W1 和 W2 都是类 weather 中某个实例的替身，二者都遍历该类的所有实例。（用大多数数据库系统的术语说，W1 和 W2 被称为范围变量。）查询可以包含任意数量的类名和替身。

## 4.8. 更新

你可以用 update 命令更新现有的实例。假设你发现从 11 月 28 日起所有温度读数都偏了 2 度，你可以这样更新数据：

```
UPDATE weather
    SET temp_hi = temp_hi - 2,  temp_lo = temp_lo - 2
    WHERE date > '11/28/1994';
```

## 4.9. 删除

删除是用 delete 命令执行的：

```
DELETE FROM weather WHERE city = 'Hayward';
```

所有属于 Hayward 的天气记录都被删除了。对如下形式的查询应当保持警惕

```
DELETE FROM classname;
```

没有限定条件时，delete 会直接删除给定类中的所有实例，使其变空。系统在此之前不会请求确认。

## 4.10. 使用聚合函数

与大多数其他查询语言一样，PostgreSQL 支持聚合函数。当前实现的 Postgres 聚合函数有一些限制。具体来说，虽然有一些聚合可以计算一组实例的 count（计数）、sum（求和）、avg（平均值）、max（最大值）和 min（最小值），但聚合只能出现在查询的目标列表中，不能直接出现在限定条件（where 子句）里。举个例子，

```
SELECT max(temp_lo) FROM weather;
```

是允许的，而

```
SELECT city FROM weather WHERE temp_lo = max(temp_lo);
```

则不行。不过，正如常见的情况，可以把查询改写成另一种形式来达到想要的结果，这里用的是子选择：

```
SELECT city FROM weather WHERE temp_lo = (SELECT max(temp_lo) FROM
weather);
```

聚合还可以带 group by 子句：

```
SELECT city, max(temp_lo)
    FROM weather
    GROUP BY city;
```

---

## 第 5 章 Postgres SQL 的高级特性

在介绍了使用 Postgres SQL 访问数据的基础知识之后，我们现在来讨论 Postgres 区别于传统数据库管理器的那些特性。这些特性包括继承、时间旅行以及非原子数据值（数组值和集合值属性）。本节的示例也可以在教程目录的 `advance.sql` 中找到。（关于如何使用它，请参阅第 4 章。）

### 5.1. 继承

让我们创建两个类。`capitals` 类包含也是城市的州首府。自然，`capitals` 类应当从 `cities` 继承。

```
CREATE TABLE cities (
    name          text,
    population     float,
    altitude       int          -- (in ft)
);

CREATE TABLE capitals (
    state          char2
) INHERITS (cities);
```

在这种情况下，`capitals` 的一个实例会从它的父类 `cities` 继承全部属性（`name`、`population` 和 `altitude`）。属性 `name` 的类型是 `text`，这是 Postgres 原生的一种变长 ASCII 字符串类型。属性 `population` 的类型是 `float`，一种 Postgres 原生的双精度浮点数类型。州首府有一个额外的属性 `state`，用来表示所在的州。在 Postgres 中，一个类可以从零个或多个其他类继承，而一个查询既可以引用一个类的所有实例，也可以引用一个类及其所有后代的所有实例。

#### 注意

继承层次结构是一个有向无环图。

例如，下面的查询会找出所有海拔在 500 英尺或更高的城市：

```
SELECT name, altitude
FROM cities
WHERE altitude > 500;
```

```
+-----+-----+
|name      | altitude |
+-----+-----+
|Las Vegas | 2174     |
+-----+-----+
|Mariposa  | 1953     |
+-----+-----+
```

另一方面，要找出所有海拔超过 500 英尺的城市名称（包括州首府），查询是：

```
SELECT c.name, c.altitude
FROM cities* c
WHERE c.altitude > 500;
```

它返回：

```

+-----+-----+
|name      | altitude |
+-----+-----+
|Las Vegas | 2174      |
+-----+-----+
|Mariposa  | 1953      |
+-----+-----+
|Madison   | 845       |
+-----+-----+

```

这里，cities 后面的 “\*” 表示该查询应当在 cities 以及继承层次中位于 cities 之下的所有类上运行。我们已经讨论过的许多命令（select、update 和 delete）以及 alter 等其他命令都支持这种 “\*” 记法。

## 5.2. 非原子值

关系模型的信条之一是关系的属性是原子的。Postgres 没有这个限制；属性本身可以包含可以从查询语言访问的子值。例如，你可以创建由基本类型构成的数组属性。

### 5.2.1. 数组

Postgres 允许把一个实例的属性定义为定长或变长的多维数组。可以创建任何基本类型或用户定义类型的数组。为了说明它们的用法，我们首先创建一个带有基本类型数组的类。

```

CREATE TABLE SAL_EMP (
    name          text,
    pay_by_quarter int4[],
    schedule       text[][]
);

```

上面的查询会创建一个名为 SAL\_EMP 的类，它带有一个 text 字符串（name）、一个 int4 一维数组（pay\_by\_quarter，表示雇员按季度的薪水）和一个 text 二维数组（schedule，表示雇员的每周日程）。现在我们做一些 INSERTS；注意，向数组追加内容时，我们把值放在花括号内并用逗号分隔。如果你了解 C，这很像初始化结构的语法。

```

INSERT INTO SAL_EMP
VALUES ('Bill',
       '{10000, 10000, 10000, 10000}',
       '{{"meeting", "lunch"}, {}}');

INSERT INTO SAL_EMP
VALUES ('Carol',
       '{20000, 25000, 25000, 25000}',
       '{{"talk", "consult"}, {"meeting"}}');

```

默认情况下，Postgres 对数组使用“从 1 开始”的编号约定——即一个有 n 个元素的数组从 array[1] 开始到 array[n] 结束。现在，我们可以在 SAL\_EMP 上运行一些查询。首先，我们演示如何一次访问数组的一个元素。下面的查询检索薪水在第二季度发生变化的雇员的姓名：

```

SELECT name
FROM SAL_EMP
WHERE SAL_EMP.pay_by_quarter[1] <>
      SAL_EMP.pay_by_quarter[2];

```

```
+-----+
|name   |
+-----+
|Carol  |
+-----+
```

下面的查询检索所有雇员的第三季度薪水：

```
SELECT SAL_EMP.pay_by_quarter[3] FROM SAL_EMP;
```

```
+-----+
|pay_by_quarter |
+-----+
|10000          |
+-----+
|25000          |
+-----+
```

我们还可以访问数组的任意切片或子数组。下面的查询检索 Bill 的日程中一周头两天的第一项。

```
SELECT SAL_EMP.schedule[1:2][1:1]
       FROM SAL_EMP
       WHERE SAL_EMP.name = 'Bill';
```

```
+-----+
|schedule          |
+-----+
|{{"meeting"},{""}} |
+-----+
```

## 5.3. 时间旅行

从 Postgres v6.2 起，不再支持时间旅行。原因有几个：性能影响、存储大小，以及一个会在短时间内增长到无限大小的 `pg_time` 文件。

触发器等新特性允许在需要时模拟时间旅行的行为，而在不需要时（对大多数用户来说，大多数时候都不需要）不产生开销。更多信息见 `contrib` 目录中的示例。

### 时间旅行已废弃

本节余下的文字只是暂时保留，直到能够用实现相同目的的新技术的语境重写为止。  
有人志愿吗？ -- thomas 1998-01-12

Postgres 支持时间旅行的概念。这个特性允许用户运行历史查询。例如，要查找 Mariposa 市当前的人口，可以查询：

```
SELECT * FROM cities WHERE name = 'Mariposa';
```

```
+-----+-----+-----+
|name      | population | altitude |
+-----+-----+-----+
```

```
|Mariposa | 1320          | 1953          |
+-----+-----+-----+
```

Postgres 会自动找到在当前时间有效的 Mariposa 记录版本。还可以给出一个时间范围。例如要查看 Mariposa 过去和现在的人口，可以查询：

```
SELECT name, population
       FROM cities['epoch', 'now']
       WHERE name = 'Mariposa';
```

其中"epoch"表示系统时钟的起点。

### 注意

在 UNIX 系统上，这总是 1970 年 1 月 1 日 GMT 午夜。

如果你已经执行了目前为止的所有示例，那么上面的查询返回：

```
+-----+-----+
|name      | population |
+-----+-----+
|Mariposa  | 1200       |
+-----+-----+
|Mariposa  | 1320       |
+-----+-----+
```

时间范围的默认起点是系统可表示的最早时间，默认终点是当前时间；因此，上面的时间范围可以缩写为 ``[,]"

## 5.4. 更多高级特性

Postgres 还有许多特性是这个面向 SQL 新用户的入门教程中没有涉及的。这些特性在用户指南和程序员指南中有更详细的讨论。

---

## 部分 II. 用户指南

面向用户的信息。

---

# 目录

|                             |    |
|-----------------------------|----|
| 6. 设置你的环境 .....             | 25 |
| 7. 管理数据库 .....              | 26 |
| 7.1. 数据库创建 .....            | 26 |
| 7.2. 备选数据库位置 .....          | 26 |
| 7.3. 访问数据库 .....            | 27 |
| 7.3.1. 数据库权限 .....          | 28 |
| 7.3.2. 表权限 .....            | 28 |
| 7.4. 删除数据库 .....            | 28 |
| 8. 数据类型 .....               | 30 |
| 8.1. 数字类型 .....             | 31 |
| 8.2. 货币类型 .....             | 32 |
| 8.3. 字符类型 .....             | 32 |
| 8.4. 日期/时间类型 .....          | 33 |
| 8.4.1. 日期/时间样式 .....        | 34 |
| 8.4.2. 时区 .....             | 34 |
| 8.4.3. 日期/时间输入 .....        | 35 |
| 8.4.4. datetime .....       | 35 |
| 8.4.5. timespan .....       | 36 |
| 8.4.6. abstime .....        | 36 |
| 8.4.7. reltime .....        | 36 |
| 8.4.8. timestamp .....      | 37 |
| 8.4.9. interval .....       | 37 |
| 8.4.10. tinterval .....     | 37 |
| 8.5. 布尔类型 .....             | 37 |
| 8.6. 几何类型 .....             | 37 |
| 8.6.1. 点 .....              | 38 |
| 8.6.2. 线段 .....             | 38 |
| 8.6.3. 方框 .....             | 38 |
| 8.6.4. 路径 .....             | 38 |
| 8.6.5. 多边形 .....            | 39 |
| 8.6.6. 圆 .....              | 39 |
| 8.7. IPv4 网络与主机地址 .....     | 39 |
| 8.7.1. 用于 IP 网络的 inet ..... | 40 |
| 8.7.2. 用于 IP 网络的 inet ..... | 40 |
| 9. 操作符 .....                | 41 |
| 9.1. 词法优先级 .....            | 41 |
| 9.2. 通用操作符 .....            | 42 |
| 9.3. 数值操作符 .....            | 42 |
| 9.4. 几何操作符 .....            | 43 |
| 9.5. 时间间隔操作符 .....          | 44 |
| 9.6. IP V4 操作符 .....        | 44 |
| 9.7. IP V4 操作符 .....        | 45 |
| 10. 函数 .....                | 46 |
| 10.1. 数学函数 .....            | 46 |
| 10.2. 字符串函数 .....           | 46 |
| 10.3. 日期/时间函数 .....         | 47 |
| 10.4. 几何函数 .....            | 48 |
| 10.5. IP V4 函数 .....        | 50 |
| 11. 类型转换 .....              | 51 |
| 11.1. 概述 .....              | 51 |
| 11.1.1. 指南 .....            | 52 |

|                           |     |
|---------------------------|-----|
| 11.2. 操作符 .....           | 52  |
| 11.2.1. 转换过程 .....        | 52  |
| 11.2.2. 示例 .....          | 53  |
| 11.3. 函数 .....            | 54  |
| 11.3.1. 示例 .....          | 55  |
| 11.4. 查询目标 .....          | 56  |
| 11.4.1. 示例 .....          | 56  |
| 11.5. UNION 查询 .....      | 57  |
| 11.5.1. 示例 .....          | 57  |
| 12. 索引和键 .....            | 59  |
| 13. 数组 .....              | 61  |
| 14. 继承 .....              | 63  |
| 15. 查询语言 .....            | 65  |
| 15.1. 概念 .....            | 65  |
| 15.2. 创建新类 .....          | 65  |
| 15.3. 用实例填充类 .....        | 65  |
| 15.4. 查询类 .....           | 66  |
| 15.5. 重定向 SELECT 查询 ..... | 66  |
| 15.6. 类之间的连接 .....        | 67  |
| 15.7. 更新 .....            | 68  |
| 15.8. 删除 .....            | 68  |
| 15.9. 使用聚合函数 .....        | 68  |
| 16. 磁盘存储 .....            | 69  |
| 17. psql .....            | 70  |
| 17.1. 分页到屏幕 .....         | 70  |
| 18. pgaccess .....        | 71  |
| 19. SQL 命令 .....          | 72  |
| ABORT .....               | 73  |
| ALTER TABLE .....         | 74  |
| ALTER USER .....          | 77  |
| BEGIN WORK .....          | 79  |
| CLOSE .....               | 80  |
| CLUSTER .....             | 81  |
| COMMIT .....              | 83  |
| COPY .....                | 84  |
| CREATE AGGREGATE .....    | 88  |
| CREATE DATABASE .....     | 91  |
| CREATE FUNCTION .....     | 93  |
| CREATE INDEX .....        | 96  |
| CREATE LANGUAGE .....     | 99  |
| CREATE OPERATOR .....     | 103 |
| CREATE RULE .....         | 107 |
| CREATE SEQUENCE .....     | 110 |
| CREATE TABLE .....        | 114 |
| CREATE TABLE AS .....     | 128 |
| CREATE TRIGGER .....      | 129 |
| CREATE TYPE .....         | 131 |
| CREATE USER .....         | 135 |
| CREATE VIEW .....         | 138 |
| DECLARE .....             | 140 |
| DELETE .....              | 143 |
| DROP AGGREGATE .....      | 145 |
| DROP DATABASE .....       | 147 |

|                     |     |
|---------------------|-----|
| DROP FUNCTION ..... | 148 |
| DROP INDEX .....    | 150 |
| DROP LANGUAGE ..... | 151 |
| DROP OPERATOR ..... | 153 |
| DROP RULE .....     | 155 |
| DROP SEQUENCE ..... | 156 |
| DROP TABLE .....    | 157 |
| DROP TRIGGER .....  | 159 |
| DROP TYPE .....     | 160 |
| DROP USER .....     | 162 |
| DROP VIEW .....     | 163 |
| EXPLAIN .....       | 165 |
| FETCH .....         | 167 |
| GRANT .....         | 170 |
| INSERT .....        | 174 |
| LISTEN .....        | 176 |
| LOAD .....          | 178 |
| LOCK .....          | 180 |
| MOVE .....          | 182 |
| NOTIFY .....        | 184 |
| RESET .....         | 186 |
| REVOKE .....        | 187 |
| ROLLBACK .....      | 190 |
| SELECT .....        | 191 |
| SELECT INTO .....   | 197 |
| SET .....           | 198 |
| SHOW .....          | 203 |
| UNLISTEN .....      | 205 |
| UPDATE .....        | 207 |
| VACUUM .....        | 209 |
| 20. 工具应用程序 .....    | 211 |
| createdb .....      | 212 |
| createuser .....    | 214 |
| destroydb .....     | 216 |
| destroyuser .....   | 218 |
| initdb .....        | 220 |
| initlocation .....  | 223 |
| pg_dump .....       | 225 |
| pg_dumpall .....    | 228 |
| psql .....          | 231 |

---

## 第 6 章 设置你的环境

本节讨论如何设置你自己的环境，使你能够使用前端应用程序。我们假设 Postgres 已经被成功安装并启动；关于如何安装 Postgres，请参阅管理员指南和安装说明。

Postgres 是一个客户/服务器应用程序。作为用户，你只需要访问安装的客户端部分（客户端应用程序的一个例子是交互式监视程序 `psql`）。为简单起见，我们假设 Postgres 已经安装在目录 `/usr/local/pgsql` 中。因此，无论在哪里看到目录 `/usr/local/pgsql`，你都应该用 Postgres 实际安装所在的目录名来替换它。所有 Postgres 命令都安装在目录 `/usr/local/pgsql/bin` 中。因此，你应当把这个目录加到你的 shell 命令路径中。如果你使用 Berkeley C shell 的一个变体，例如 `csh` 或 `tcsh`，你可以把

```
set path = ( /usr/local/pgsql/bin path )
```

加到你主目录中的 `.login` 文件里。如果你使用 Bourne shell 的一个变体，例如 `sh`、`ksh` 或 `bash`，那么你可以把

```
PATH=/usr/local/pgsql/bin PATH
export PATH
```

加到你主目录中的 `.profile` 文件里。从现在起，我们将假设你已经把 Postgres 的 `bin` 目录加到了你的路径中。此外，我们在本文档中将频繁提到“设置一个 shell 变量”或“设置一个环境变量”。如果你没有完全理解上一段关于修改搜索路径的内容，在继续之前你应当查阅描述你的 shell 的 Unix 手册页。

如果你的站点管理员没有按默认方式完成设置，你可能还有一些工作要做。例如，如果数据库服务器机器是一台远程机器，你就需要把 `PGHOST` 环境变量设置为数据库服务器机器的名字。环境变量 `PGPORT` 也可能必须设置。底线是：如果你试图启动一个应用程序而它抱怨无法连接到 `postmaster`，你应当立即咨询你的站点管理员，确保你的环境已正确设置。

---

## 第 7 章 管理数据库

### 注意

这一节目前基本上是教程的一份简单改写的拷贝。需要扩充。- thomas 1998-01-12

虽然站点管理员负责 Postgres 安装的整体管理，但安装中的某些数据库可以由另一个人管理，这个人被指定为数据库管理员。这种职责分配发生在创建数据库时。一个用户可以被授予创建数据库和/或创建新用户的显式权限。被同时授予这两种权限的用户可以在 Postgres 内执行大多数管理任务，但默认不会拥有与站点管理员相同的操作系统权限。

数据库管理员指南更详细地讨论这些主题。

### 7.1. 数据库创建

数据库由从 Postgres 内部发出的 `create database` 命令创建。`createdb` 是一个命令行实用程序，用来在 Postgres 之外提供同样的功能。

无论哪种方法，要成功执行都必须有 Postgres 后端在运行，而且发出命令的用户必须是 Postgres 超级用户，或者已被超级用户授予了数据库创建权限。

要从命令行创建一个名为“mydb”的新数据库，输入

```
% createdb mydb
```

要在 psql 内做同样的事，输入

```
* CREATE DATABASE mydb;
```

如果你没有创建数据库所需的权限，你会看到下列消息：

```
% createdb mydb
WARN:user "your username" is not allowed to create/destroy databases
createdb: database creation failed on mydb.
```

Postgres 允许你在给定站点创建任意数量的数据库，并且你会自动成为你刚创建的数据库的数据库管理员。数据库名的第一个字符必须是字母，并且长度限制为 32 个字符。

### 7.2. 备选数据库位置

可以在安装的默认位置之外的位置创建数据库。请记住，所有数据库访问实际上都是通过数据库后端进行的，因此所指定的任何位置都必须是后端可以访问的。

可以指定一个绝对路径名或者一个环境变量作为位置。任何指定备选位置的环境变量都必须在后端启动之前已经定义。关于预先配置的备选数据库位置，请咨询站点管理员。

## 注意

环境变量风格的指定方式更可取，因为它让站点管理员在管理磁盘存储时有更大的灵活性。

出于安全和完整性原因，所指定的任何路径或环境变量都会被追加一些额外的路径字段。

备选数据库位置必须通过运行 `initlocation` 来准备。

要在 `/alt/postgres/data` 中创建一个数据存储区，请确保 `/alt/postgres` 已经存在。在命令行输入

```
% initlocation /alt/postgres/data
Creating Postgres database system directory /alt/postgres/data

Creating Postgres database system directory /alt/postgres/data/base
```

要使用环境变量 `PGDATA2` 做同样的事，输入

```
% initlocation $PGDATA2
Creating Postgres database system directory /alt/postgres/data

Creating Postgres database system directory /alt/postgres/data/base
```

要从命令行在备选存储区 `/alt/postgres/data` 中创建数据库，输入

```
% createdb -D /alt/postgres/data mydb
```

或者

```
% createdb -D PGDATA2 mydb
```

要在 `psql` 内做同样的事，输入

```
* CREATE DATABASE mydb WITH LOCATION = 'PGDATA2';
```

如果你没有创建数据库所需的权限，你会看到下列消息：

```
% createdb mydb
WARN:user "your username" is not allowed to create/destroy databases
createdb: database creation failed on mydb.
```

如果指定的位置不存在，或者数据库后端没有权限访问它或写入其下的目录，你会看到下列消息：

```
% createdb -D /alt/postgres/data mydb
ERROR:  Unable to create database directory /alt/postgres/data/base/
mydb
createdb: database creation failed on mydb.
```

## 7.3. 访问数据库

一旦构造好数据库，就可以通过以下方式访问它：

- 运行 Postgres 终端监视程序（例如 psql），它允许你交互式地输入、编辑和执行 SQL 命令。
- 使用 LIBPQ 子程序库编写 C 程序。这允许你从 C 中提交 SQL 命令，并把答案和状态消息返回给你的程序。这一接口在第 ?? 节中有进一步讨论。

你也许想启动 psql，来试试本手册中的例子。可以通过输入以下命令为 mydb 数据库激活它：

```
% psql mydb
```

你会看到下面的欢迎消息：

```
Welcome to the POSTGRESQL interactive sql monitor:
Please read the file COPYRIGHT for copyright terms of POSTGRESQL

type \? for help on slash commands
type \q to quit
type \g or terminate with semicolon to execute query
You are currently connected to the database: template1
```

```
mydb=>
```

这个提示表明终端监视程序正在监听你，你可以把 SQL 查询输入到由终端监视程序维护的工作区中。psql 程序响应以反斜杠字符 “\” 开头的转义码。例如，你可以获得各种 Postgres SQL 命令的语法帮助：

```
mydb=> \h
```

一旦在工作区中输入完查询，你可以通过输入下列内容把工作区的内容传递给 Postgres 服务器：

```
mydb=> \g
```

这告诉服务器处理该查询。如果你用分号终止查询，“\g”就不是必需的。psql 会自动处理以分号终止的查询。要从文件（例如 myFile）中读取查询而不是交互式地输入，输入：

```
mydb=> \i fileName
```

要退出 psql 并返回 UNIX，输入

```
mydb=> \q
```

然后 psql 将退出并把你返回到命令 shell。（关于更多转义码，在监视程序提示符下输入 \h。）空白（即空格、制表符和换行符）可以在 SQL 查询中自由使用。单行注释用 “--” 表示。连字符之后直到行尾的所有内容都被忽略。多行注释以及行内注释用 “/\* ... \*/” 表示

## 7.3.1. 数据库权限

## 7.3.2. 表权限

TBD

## 7.4. 删除数据库

如果你是数据库 mydb 的数据库管理员，你可以使用下面的 UNIX 命令销毁它：

```
% destroydb mydb
```

这个动作会物理地删除与该数据库关联的所有 UNIX 文件，而且无法撤销，因此只应在深思熟虑之后才进行。

# 第 8 章 数据类型

描述 Postgres 中可用的内建数据类型。

Postgres 有一套丰富的本机数据类型可供用户使用。用户可以用在别处描述的 `define type` 命令向 Postgres 添加新类型。

在数据类型的语境下，下面几节将讨论 SQL 标准兼容性、移植问题和使用方法。某些 Postgres 类型直接对应于 SQL92 兼容类型。另一些情况下，由 SQL92 语法定义的数据类型被直接映射为 Postgres 本机类型。许多内建类型都有显而易见的外部格式。但也有一些类型要么为 Postgres 所独有（例如开放和闭合路径），要么有多种可选格式（例如日期和时间类型）。

表 8.1. Postgres 数据类型

| Postgres 类型 | SQL92 或 SQL3 类型          | 描述                 |
|-------------|--------------------------|--------------------|
| bool        | boolean                  | 逻辑布尔值（真/假）         |
| box         |                          | 二维平面上的矩形框          |
| char(n)     | character(n)             | 定长字符串              |
| cidr        |                          | IPv4 网络或主机地址       |
| circle      |                          | 二维平面上的圆            |
| date        | date                     | 不含时间的日历日期          |
| float4/8    | float(p)                 | 精度为 p 的浮点数         |
| float8      | real, double precision   | 双精度浮点数             |
| inet        |                          | IPv4 网络或主机地址       |
| int2        | smallint                 | 有符号二字节整数           |
| int4        | int, integer             | 有符号 4 字节整数         |
| int4        | decimal(p,s)             | p <= 9、s = 0 的精确数值 |
| int4        | numeric(p,s)             | p == 9、s = 0 的精确数值 |
| int8        |                          | 有符号 8 字节整数         |
| line        |                          | 二维平面上的无限直线         |
| lseg        |                          | 二维平面上的线段           |
| money       | decimal(9,2)             | 美式货币               |
| path        |                          | 二维平面上的开放和闭合几何路径    |
| point       |                          | 二维平面上的几何点          |
| polygon     |                          | 二维平面上的封闭几何路径       |
| serial      |                          | 用于索引和交叉引用的唯一 id    |
| time        | time                     | 一天中的时间             |
| timespan    | interval                 | 通用时间段              |
| timestamp   | timestamp with time zone | 日期/时间              |
| varchar(n)  | character varying(n)     | 变长字符串              |

注意

`cidr` 和 `inet` 类型设计为可处理任何 IP 类型，但当前实现只处理 `ipv4`。这里所有关于 `ipv4` 的内容在未来的版本中都将适用于 `ipv6`。

表 8.2. Postgres 函数常量

| Postgres 函数      | SQL92 常量          | 描述         |
|------------------|-------------------|------------|
| getpgusername()  | current_user      | 当前会话中的用户名  |
| date('now')      | current_date      | 当前事务的日期    |
| time('now')      | current_time      | 当前事务的时间    |
| timestamp('now') | current_timestamp | 当前事务的日期和时间 |

Postgres 的特性处于 ORDBMS 开发的前沿。除了符合 SQL3 之外，还支持 SQL92 的很大一部分。尽管我们力求 SQL92 兼容，但该标准有一些方面考虑欠周，不应延续到后续标准中。Postgres 不会花大力气去符合这些特性；不过它们往往只适用于很少使用或冷僻的场景，典型用户不太可能遇到。

大多数对应于基本类型（如整数和浮点数）的输入和输出函数会做一些错误检查。为了提高执行速度，某些运算符和函数（例如加法和乘法）不做运行时错误检查。例如在某些系统上，某些数据类型的数值运算符可能静默下溢或上溢。

注意，有些输入和输出函数是不可逆的。也就是说，输出函数的结果与原始输入相比可能损失精度。

### 注意

从 Berkeley 收到的最初 Postgres v4.2 代码会把所有双精度浮点结果四舍五入到六位输出。从 v6.1 开始，浮点数允许保留该类型的大部分固有精度（双精度通常 15 位，4 字节浮点通常 6 位）。底层为浮点字段的其他类型（如几何类型）具有类似的精度。

## 8.1. 数字类型

数值类型包括二字节和四字节整数以及四字节和八字节浮点数。

表 8.3. Postgres 数值类型

| 数值类型   | 存储尺寸 | 描述       | 范围                        |
|--------|------|----------|---------------------------|
| float4 | 4字节  | 可变精度     | 6 位十进制小数                  |
| float8 | 8字节  | 可变精度     | 15 位十进制小数                 |
| int2   | 2字节  | 定点数      | -32768 到 +32767           |
| int4   | 4字节  | 定点数的常用选择 | -2147483648 到 +2147483647 |
| int8   | 8字节  | 极大范围定点数  | +/- > 18 位十进制小数           |
| serial | 4字节  | 标识符或交叉引用 | 0 到 +2147483647           |

数值类型有一整套相应的算术运算符和函数。更多信息请参阅数值操作符和数学函数。

`serial` 类型是由 Postgres 用其他现有构件构造的一种特殊类型。它通常用于为表条目创建唯一标识符。在当前的实现中，指定

```
CREATE TABLE tablename (colname SERIAL);
```

等效于指定：

```
CREATE SEQUENCE tablename_colname_seq;
CREATE TABLE tablename
    (colname INT4 DEFAULT nextval('tablename_colname_seq');
CREATE UNIQUE INDEX tablename_colname_key on tablename (colname);
```

### 小心

为 serial 类型创建的隐式序列在表被删除时不会被自动删除。因此，按顺序执行下面的命令很可能会失败：

```
CREATE TABLE tablename (colname SERIAL);
DROP TABLE tablename;
CREATE TABLE tablename (colname SERIAL);
```

该序列会一直留在数据库中，直到用 DROP SEQUENCE 显式删除。

精确数值 decimal 和 numeric 拥有完整实现的语法，但目前（Postgres v6.4）只支持较小的精度和/或取值范围。int8 类型并非在所有平台上都可用，因为它依赖编译器对此的支持。

## 8.2. 货币类型

money 类型支持以固定小数点表示的美式货币。如果 Postgres 编译时定义了 USE\_LOCALE 则 money 类型应使用为 locale(7) 定义的货币惯例。

表 8.4. Postgres 货币类型

| 货币类型  | 存储尺寸 | 描述   | 范围                          |
|-------|------|------|-----------------------------|
| money | 4 字节 | 固定精度 | -21474836.48 到 +21474836.47 |

numeric 最终应取代 money 类型。它拥有完整实现的语法，但目前（Postgres v6.4）只支持较小的精度和/或取值范围，尚不能充分替代 money 类型。

## 8.3. 字符类型

SQL92 定义了两种主要的字符类型：char 和 varchar。Postgres 支持这些类型，此外还有更通用的 text 类型；与 varchar 不同，它不要求为字段大小声明上限。

表 8.5. Postgres 字符类型

| 字符类型       | 存储尺寸     | 建议       | 描述       |
|------------|----------|----------|----------|
| char       | 1 字节     | SQL92 兼容 | 单字符      |
| char(n)    | (4+n) 字节 | SQL92 兼容 | 定长，空白填充  |
| text       | (4+x) 字节 | 最佳选择     | 变长       |
| varchar(n) | (4+n) 字节 | SQL92 兼容 | 有长度限制的变长 |

还有另一种定长字符类型。name 类型只有一个用途，即给 Postgres 提供一个用于内部名称的特殊类型。它并非供普通用户使用。其长度目前定义为 32 个字符，但应使用 NAMEDATALEN 引用。这在编译时设定，在未来的版本中可能改变。

表 8.6. Postgres 特殊字符类型

| 字符类型 | 存储尺寸 | 描述        |
|------|------|-----------|
| name | 32字节 | 32 字符内部类型 |

## 8.4. 日期/时间类型

日期时间度量有两种基本类型：绝对时钟时间和相对时间区间。

两种量都应体现连续性和平滑性，就像时间本身一样。Postgres 提供两种主要的面向用户的日期时间类型 `datetime` 和 `timespan`，以及相关的 SQL92 类型 `timestamp`、`interval`、`date` 和 `time`。

在未来的版本中，`datetime` 和 `timespan` 很可能与 SQL92 类型 `timestamp`、`interval` 合并。另外还有一些其他日期时间类型可用，大多出于历史原因。

表 8.7. Postgres 日期/时间类型

| 日期/时间类型                | 存储尺寸 | 建议        | 描述                        |
|------------------------|------|-----------|---------------------------|
| <code>abstime</code>   | 4字节  | 原始日期和时间   | 范围有限                      |
| <code>date</code>      | 4字节  | SQL92 类型  | 范围广                       |
| <code>datetime</code>  | 8字节  | 最佳通用日期和时间 | 范围广、精度高                   |
| <code>interval</code>  | 12字节 | SQL92 类型  | 等价于 <code>timespan</code> |
| <code>reltime</code>   | 4字节  | 原始时间区间    | 范围有限、精度低                  |
| <code>time</code>      | 4字节  | SQL92 类型  | 范围广                       |
| <code>timespan</code>  | 12字节 | 最佳通用时间区间  | 范围广、精度高                   |
| <code>timestamp</code> | 4字节  | SQL92 类型  | 范围有限                      |

`timestamp` 目前与 `datetime` 分开实现，不过它们共用输入和输出例程。

表 8.8. Postgres 日期/时间范围

| 日期/时间类型                | 最早           | 最晚          | 精度       |
|------------------------|--------------|-------------|----------|
| <code>abstime</code>   | 1901-12-14   | 2038-01-19  | 1秒       |
| <code>date</code>      | 4713 BC      | 32767 AD    | 1日       |
| <code>datetime</code>  | 4713 BC      | 1465001 AD  | 1微秒到14位  |
| <code>interval</code>  | -1780000000年 | 1780000000年 | 1微秒      |
| <code>reltime</code>   | -68年         | +68年        | 1秒       |
| <code>time</code>      | 00:00:00.00  | 23:59:59.99 | 1微秒      |
| <code>timespan</code>  | -1780000000年 | 1780000000年 | 1微秒（14位） |
| <code>timestamp</code> | 1901-12-14   | 2038-01-19  | 1秒       |

Postgres 力求在典型用法上与 SQL92 定义兼容。SQL92 标准对日期和时间类型及能力有一种奇怪的组合。两个明显的问题是：

- 尽管 `date` 类型没有关联的时区，`time` 类型却可以或确实有关联的时区。
- 默认时区被指定为相对 GMT/UTC 的固定整数偏移。

然而，现实世界中的时区若不同时关联日期和时间就可能没有意义，因为偏移量在一年中可能随夏令时边界而变化。

为解决这些困难，Postgres 只把时区关联到同时包含日期和时间的日期时间类型，而对只包含假定使用本地时间。此外，时区支持源自底层操作系统的时区能力，因此可以处理夏令时和其他预期行为。

在未来的版本中，日期时间类型的数量将会减少：当前 `datetime` 的实现将变成 `timestamp`，`timespan` 变成 `interval`，而 `abstime` 和 `reltime`（可能）被弃用，转而使用 `timestamp` 和 `interval`。SQL92 标准中日期时间定义的那些更晦涩的特性不太可能被追求。

### 8.4.1. 日期/时间样式

输出格式可以设为四种风格之一：ISO-8601、SQL (Ingres)、传统 Postgres 和德式。

**表 8.9. Postgres 日期样式**

| 样式说明     | 描述          | 示例                           |
|----------|-------------|------------------------------|
| ISO      | ISO-8601 标准 | 1997-12-17 07:37:16-08       |
| SQL      | 传统样式        | 12/17/1997 07:37:16.00 PST   |
| Postgres | 原始样式        | Wed Dec 17 07:37:16 1997 PST |
| German   | 地区样式        | 17.12.1997 07:37:16.00 PST   |

SQL 风格有欧式和非欧式（美式）两种变体，它们决定月跟在日后还是日跟在月前。

**表 8.10. Postgres 日期顺序惯例**

| 样式说明        | 描述   | 示例                         |
|-------------|------|----------------------------|
| European    | 地区惯例 | 17/12/1997 15:37:16.00 MET |
| NonEuropean | 地区惯例 | 12/17/1997 07:37:16.00 PST |
| US          | 地区惯例 | 12/17/1997 07:37:16.00 PST |

有几种方法可以影响日期/时间类型的外观：

- `postmaster` 启动时后端直接使用的 `PGDATESTYLE` 环境变量。
- 会话启动时前端 `libpq` 使用的 `PGDATESTYLE` 环境变量。
- `SET DateStyle SQL` 命令。

对于 Postgres v6.4（及更早版本），默认的日期/时间样式是“非欧式传统 Postgres”。在未来的版本中，默认值可能变成 ISO-8601，它会减轻日期指定的歧义和 Y2K 排序问题。

### 8.4.2. 时区

Postgres 从底层操作系统获得时区支持。所有日期和时间在内部都以协调世界时（UTC，也称为格林尼治标准时间 GMT）存储。时间在发送给客户端前端之前会被转换为数据库服务器上的本地时间，因此默认情况下采用服务器时区。

有几种方法可以影响时区行为：

- `postmaster` 启动时后端直接使用 `TZ` 环境变量作为默认时区。
- 在客户端设置的 `PGTZ` 环境变量由 `libpq` 用于在连接时向后端发送时区信息。
- SQL 命令 `SET TIME ZONE` 设置会话的时区。

如果指定了无效的时区，时区会变成 GMT（至少在大多数系统上如此）。

### 8.4.3. 日期/时间输入

通用日期时间的输入可采用多种风格，包括 ISO 兼容、SQL 兼容、传统 Postgres 以及日期时间的其他排列组合。在解释可能有歧义的情况下（许多传统日期指定风格很可能如此），Postgres 使用一个风格设置来解决歧义。

大多数日期时间类型共用数据输入代码。对这些类型，输入可以采用多种多样的风格。

对于数字日期表示，欧式与美式惯例可能不同，正确的解释要在输入数据之前使用 `set datestyle` 命令来获得。注意，风格设置并不妨碍在输入中使用各种风格；它主要用于确定输出风格和解决歧义。

提供了特殊值 `'current'`、`'infinity'` 和 `'-infinity'`。`'infinity'` 指定一个晚于任何其他有效时间的时间，而 `'-infinity'` 指定一个早于任何其他有效时间的时间。`'current'` 表示每当该值出现在计算中时就以当前时间替换。字符串 `'now'`、`'today'`、`'yesterday'`、`'tomorrow'` 和 `'epoch'` 可用于指定时间值。`'now'` 表示当前事务时间，它与 `'current'` 的区别在于当前时间会被立即替换进去。`'epoch'` 表示 Jan 1 00:00:00 1970 GMT。

**表 8.11. Postgres 特殊日期/时间常量**

| 常量                     | 描述                                |
|------------------------|-----------------------------------|
| <code>current</code>   | 当前事务时间，延迟求值                       |
| <code>epoch</code>     | 1970-01-01 00:00:00+00（Unix系统时间0） |
| <code>infinity</code>  | 晚于其他有效时间                          |
| <code>-infinity</code> | 早于其他有效时间                          |
| <code>invalid</code>   | 非法输入                              |
| <code>now</code>       | 当前事务时间                            |
| <code>today</code>     | 今天午夜                              |
| <code>tomorrow</code>  | 明天午夜                              |
| <code>yesterday</code> | 昨天午夜                              |

### 8.4.4. datetime

通用日期时间的输入可采用多种风格，包括 ISO 兼容、SQL 兼容、传统 Postgres（见“绝对时间”一节）以及日期时间的其他排列组合。输出风格可以是 ISO 兼容、SQL 兼容或传统 Postgres，默认设置为与 Postgres v6.0 兼容。

`datetime` 用下列语法指定：

```
Year-Month-Day [ Hour : Minute : Second ] [AD,BC] [ Timezone ]
YearMonthDay [ Hour : Minute : Second ] [AD,BC] [ Timezone ]
Month Day [ Hour : Minute : Second ] Year [AD,BC] [ Timezone ]
```

其中

```
Year (年) 是 4013 BC, ..., 非常大
Month (月) 是 Jan, Feb, ..., Dec 或 1, 2, ..., 12
Day (日) 是 1, 2, ..., 31
Hour (时) 是 00, 02, ..., 23
Minute (分) 是 00, 01, ..., 59
Second (秒) 是 00, 01, ..., 59 (闰秒为 60)
Timezone (时区) 是 3 个字符或相对 GMT 的 ISO 偏移
```

有效日期从公元前 4013 年 11 月 13 日 00:00:00 GMT 到遥远的未来。时区要么是三个字符（如 "GMT" 或 "PST"），要么是相对于 GMT 的 ISO 兼容偏移（如太平洋标准时间为 "-08"

或 "-08:00")。日期在内部以格林尼治标准时间存储。输入和输出例程会在必要时把时间转换为服务器的本地时区。

## 8.4.5. timespan

通用时间区间的输入可采用多种语法，包括 ISO 兼容、SQL 兼容、传统 Postgres（见“相对时间”一节）以及时间区间的其他排列组合。输出格式可以是 ISO 兼容、SQL 兼容或传统 Postgres，默认设置为与 Postgres 兼容。月和年是“定性”时间区间，与日或小时等其他“定量”时间区间分开存储。在日期运算中，定性时间单位会在相关日期或时间的上下文中实例化。

时间区间用下列语法指定：

```
Quantity Unit [Quantity Unit...] [Direction]
@ Quantity Unit [Direction]
```

其中

Quantity (数量) 是 ..., '-1', '0', '1', '2', ...  
Unit (单位) 是 'second', 'minute', 'hour', 'day', 'week', 'month', 'year', 'decade', 'century', 'millenium', 或这些单位的缩写和复数形式。  
Direction (方向) 是 'ago'。

## 8.4.6. abstime

绝对时间 (abstime) 是一种范围有限 ( $\pm 68$  年)、精度有限 (1 秒) 的日期数据类型。可能更宜使用 datetime，因为它覆盖更大的范围且精度更高。

绝对时间用下列语法指定：

```
Month Day [ Hour : Minute : Second ] Year [ Timezone ]
```

其中

Month (月) 是 Jan, Feb, ..., Dec  
Day (日) 是 1, 2, ..., 31  
Hour (时) 是 01, 02, ..., 24  
Minute (分) 是 00, 01, ..., 59  
Second (秒) 是 00, 01, ..., 59  
Year (年) 是 1901, 1902, ..., 2038

有效日期从 Dec 13 20:45:53 1901 GMT 到 Jan 19 03:14:04 2038 GMT。

### 历史注记

从版本 3.0 起，时间的读写不再使用格林尼治标准时间；输入和输出例程默认使用本地时区。

所有允许用于 datetime 的特殊值也同样允许用于“绝对时间”。

## 8.4.7. reltime

相对时间 reltime 是一种范围有限 ( $\pm 68$  年)、精度有限 (1 秒) 的时间区间数据类型。应优先使用 timespan，因为它覆盖更大的范围、精度更高，而且更重要的是能区分相对单位（月和年）与定量单位（日、小时等）。而 reltime 必须把一个月强制当作恰好 30 天，因此时间运算并不总是按预期工作。例如，把一个 reltime 的 year 加到 abstime 的 today 上，得到的不是从今天起一年的日期，而是从今天起 360 天的日期。

`reltime` 与其他时间区间类型共用输入和输出例程。`timespan` 一节对此有更详细的介绍。

## 8.4.8. timestamp

这是一种目前范围有限的绝对时间，与 `abstime` 数据类型非常相似。它与其他日期时间类型共用通用输入解析器。在未来的版本中，此类型将吸收 `datetime` 类型的能力，并向 SQL92 兼容的方向发展。

`timestamp` 使用与 `datetime` 相同的语法指定。

## 8.4.9. interval

`interval` 是一种 SQL92 数据类型，目前被映射到 `timespan` Postgres 数据类型。

## 8.4.10. tinterval

时间范围指定为：

```
[ 'abstime' 'abstime' ]
```

其中

`abstime` 是绝对时间格式的一个时间。

可以使用诸如 `'current'`、`'infinity'` 和 `'-infinity'` 这样的特殊 `abstime` 值。

## 8.5. 布尔类型

Postgres 支持 `bool` 作为 SQL3 布尔类型。`bool` 只能处于两种状态之一：`'true'`（真）或 `'false'`（假）。第三种状态 `'unknown'`（未知）未实现且 SQL3 也不建议使用；`NULL` 是一个有效的替代。`bool` 可用于任何布尔表达式，而布尔表达式的求值结果总是与此类型兼容。

`bool` 使用 4 字节存储。

**表 8.12. Postgres 布尔类型**

| 状态 | 输出               | 输入                                                                                                                     |
|----|------------------|------------------------------------------------------------------------------------------------------------------------|
| 真  | <code>'t'</code> | <code>TRUE</code> , <code>'t'</code> , <code>'true'</code> , <code>'y'</code> , <code>'yes'</code> , <code>'1'</code>  |
| 假  | <code>'f'</code> | <code>FALSE</code> , <code>'f'</code> , <code>'false'</code> , <code>'n'</code> , <code>'no'</code> , <code>'0'</code> |

## 8.6. 几何类型

几何类型表示二维的空间物体。最基础的类型是点，它是所有其他类型的基础。

**表 8.13. Postgres 几何类型**

| 几何类型               | 存储尺寸     | 表示                             | 描述           |
|--------------------|----------|--------------------------------|--------------|
| <code>point</code> | 16字节     | <code>(x,y)</code>             | 空间中的点        |
| <code>line</code>  | 32字节     | <code>((x1,y1),(x2,y2))</code> | 无限直线         |
| <code>lseg</code>  | 32字节     | <code>((x1,y1),(x2,y2))</code> | 有限线段         |
| <code>box</code>   | 32字节     | <code>((x1,y1),(x2,y2))</code> | 矩形框          |
| <code>path</code>  | 4+32n 字节 | <code>((x1,y1),...)</code>     | 封闭路径（类似于多边形） |

| 几何类型    | 存储尺寸     | 表示            | 描述           |
|---------|----------|---------------|--------------|
| path    | 4+32n 字节 | [(x1,y1),...] | 开放路径         |
| polygon | 4+32n 字节 | ((x1,y1),...) | 多边形（类似于封闭路径） |
| circle  | 24字节     | <(x,y),r>     | 圆（圆心和半径）     |

有一套丰富的函数和运算符可用于执行各种几何操作，如缩放、平移、旋转以及求交。

### 8.6.1. 点

点用下列语法指定：

```
( x , y )
  x , y
```

其中

$x$  是  $x$  轴坐标，以浮点数表示  
 $y$  是  $y$  轴坐标，以浮点数表示

### 8.6.2. 线段

线段 (lseg) 由点对表示。

lseg 用下列语法指定：

```
( ( x1 , y1 ) , ( x2 , y2 ) )
  ( x1 , y1 ) , ( x2 , y2 )
    x1 , y1      , x2 , y2
```

其中

(x1,y1) 和  
 (x2,y2) 是线段的两个端点

### 8.6.3. 方框

方框由作为其两个对角的点对表示。

box 用下列语法指定：

```
( ( x1 , y1 ) , ( x2 , y2 ) )
  ( x1 , y1 ) , ( x2 , y2 )
    x1 , y1      , x2 , y2
```

其中

(x1,y1) 和 (x2,y2) 是方框的对角

方框使用第一种语法输出。输入时角点会被重新排序，先存左下角，后存右上角。  
 也可以输入方框的其他角点，但左下角和右上角由输入确定并存储。

### 8.6.4. 路径

路径由相连的点集表示。路径可以是“open”（开放）的（集合中第一个点和最后一个点不相连），也可以是“closed”（封闭）的（第一个点和最后一个点相连）。函数 `popen(p)` 和 `pclose(p)` 可用于强制路径开放或封闭，而函数 `isopen(p)` 和 `isclosed(p)` 可用于在查询中选择这两种类型。

path 用下列语法指定：

```
( ( x1 , y1 ) , ... , ( xn , yn ) )
[ ( x1 , y1 ) , ... , ( xn , yn ) ]
( x1 , y1 ) , ... , ( xn , yn )
( x1 , y1 , ... , xn , yn )
x1 , y1 , ... , xn , yn
```

其中

(x1,y1),..., (xn,yn) 是第 1 到第 n 个点

前导 "[" 表示开放路径

前导 "(" 表示封闭路径

路径使用第一种语法输出。注意 Postgres v6.1 之前的版本对路径使用另一种格式：单个前导圆括号、一个“closed”（封闭）标志、一个点数的整数计数，然后是点列表和一个闭合圆括号。内建函数 `upgradepath` 用于转换从 v6.1 之前的数据库转储并重新加载的路径。

## 8.6.5. 多边形

多边形由点集表示。多边形或许应被视为等价于封闭路径，但其存储方式不同，并且有自己的一套支持例程。

polygon 用下列语法指定：

```
( ( x1 , y1 ) , ... , ( xn , yn ) )
( x1 , y1 ) , ... , ( xn , yn )
( x1 , y1 , ... , xn , yn )
x1 , y1 , ... , xn , yn
```

其中

(x1,y1),..., (xn,yn) 是第 1 到第 n 个点

多边形使用第一种语法输出。注意 Postgres v6.1 之前的版本对多边形使用另一种格式：单个前导圆括号、x 轴坐标列表、y 轴坐标列表，再接一个闭合圆括号。内建函数 `upgradepoly` 用于转换从 v6.1 之前的数据库转储并重新加载的多边形。

## 8.6.6. 圆

圆由一个圆心和一个半径表示。

circle 用下列语法指定：

```
< ( x , y ) , r >
( ( x , y ) , r )
( x , y ) , r
x , y , r
```

其中

(x,y) 是圆的圆心

r 是圆的半径

圆使用第一种语法输出。

## 8.7. IPv4 网络与主机地址

`cidr` 类型存储以 CIDR 记法指定的网络。`inet` 类型以 CIDR 记法存储主机和网络。

表 8.14. Postgres IPv4 类型

| IPv4 类型 | 存储尺寸 | 描述      | 范围              |
|---------|------|---------|-----------------|
| cidr    | 可变   | CIDR 网络 | 有效的 IPV4 CIDR 块 |
| inet    | 可变   | 网络和主机   | 有效的 IPV4 CIDR 块 |

### 8.7.1. 用于 IP 网络的 inet

cidr 类型保存一个 CIDR 网络。指定网络的格式是 "x.x.x.x/y"，其中 "x.x.x.x" 是网络，"/y" 是网络掩码的位数。如果省略了 "/y" 部分，则按旧的有类别系统的假定来计算，但至少会大到包含输入中写出的所有八位组。下面是一些示例：

表 8.15. Postgres IP 类型示例

| CIDR 输入   | CIDR 显示      |
|-----------|--------------|
| 192.168.1 | 192.168.1/24 |
| 192.168   | 192.168.0/24 |
| 128.1     | 128.1/16     |
| 128       | 128.0/16     |
| 128.1.2   | 128.1.2/24   |
| 10.1.2    | 10.1.2/24    |
| 10.1      | 10.1/16      |
| 10        | 10/8         |

### 8.7.2. 用于 IP 网络的 inet

inet 类型设计为在单个字段中保存关于一台主机的全部信息，包括它所在的 CIDR 式子网。注意，如果你想存储真正的 CIDR 网络，请参见 cidr 类型。inet 类型与 cidr 类型类似，区别在于主机部分的位可以非零。存在可用于提取该字段各组成元素的函数。

此类型的输入格式是 "x.x.x.x/y"，其中 "x.x.x.x" 是一台互联网主机，y 是网络掩码的位数。如果省略了 "/y" 部分，则按 "/32" 对待。输出时，如果 "/y" 部分为 /32，则不会打印它。这样只要省略位数部分，该类型就可以直接用作单纯的主机类型。

# 第 9 章 操作符

描述Postgres中可用的内建操作符。

Postgres 为系统类型提供了大量的内建操作符。这些操作符声明在系统目录 pg\_operator 中。pg\_operator 中的每个项都包含实现该操作符的过程的名称以及输入和输出类型的类 OIDs。

要查看 “||” 字符串连接操作符的所有变体，可以试试

```
SELECT oprleft, oprright, oprresult, oprcode
FROM pg_operator WHERE oprname = '||';

oprleft|oprright|oprresult|opcode
-----+-----+-----+-----
      25|      25|      25|textcat
    1042|    1042|    1042|textcat
    1043|    1043|    1043|textcat
(3 rows)
```

用户可以使用操作符名称来调用操作符，例如：

```
select * from emp where salary < 40000;
```

或者，用户也可以直接调用实现这些操作符的函数。在这种情况下，上面的查询将表达为：

```
select * from emp where int4lt(salary, 40000);
```

psql 有一个命令 (\dd) 用来显示这些操作符。

## 9.1. 词法优先级

操作符有优先级，目前优先级被硬编码在解析器中。大多数操作符具有相同的优先级并且不结合。这可能导致不直观的行为；例如布尔操作符 "<" 和 ">" 的优先级与布尔操作符 "<=" 和 ">=" 不同。

表 9.1. 操作符排序（优先级递减）

| 元素     | 优先级   | 描述                 |
|--------|-------|--------------------|
| UNION  | left  | SQL select 构造      |
| ::     |       | Postgres 类型转换      |
| []     | left  | 数组定界符              |
| .      | left  | 表/列定界符             |
| -      | right | 一元负号               |
| ;      | left  | 语句结束、对数            |
| :      | right | 求幂                 |
|        | left  | 间隔的起点              |
| * /    | left  | 乘法、除法              |
| + -    | left  | 加法、减法              |
| IS     |       | 测试 TRUE、FALSE、NULL |
| ISNULL |       | 测试 NULL            |

| 元素        | 优先级   | 描述       |
|-----------|-------|----------|
| NOTNULL   |       | 测试非 NULL |
| (所有其他操作符) |       | 内建和用户定义  |
| IN        |       | 集合成员     |
| BETWEEN   |       | 包含       |
| LIKE      |       | 字符串模式匹配  |
| < >       |       | 布尔不等     |
| =         | right | 等于       |
| NOT       | right | 非        |
| AND       | left  | 逻辑交      |
| OR        | left  | 逻辑并      |

## 9.2. 通用操作符

这里列出的操作符是为若干种原生数据类型定义的，从数值类型到日期/时间类型都有。

**表 9.2. Postgres 操作符**

| 操作符 | 描述             | 用法                                   |
|-----|----------------|--------------------------------------|
| <   | 小于吗？           | 1 < 2                                |
| <=  | 小于或等于吗？        | 1 <= 2                               |
| <>  | 不相等吗？          | 1 <> 2                               |
| =   | 相等吗？           | 1 = 1                                |
| >   | 大于吗？           | 2 > 1                                |
| >=  | 大于或等于吗？        | 2 >= 1                               |
|     | 连接字符串          | 'Postgre'    'SQL'                   |
| !=  | NOT IN         | 3 != i                               |
| ~~  | LIKE           | 'scrappy,marc,hermit' ~~ '%scrappy%' |
| !~~ | NOT LIKE       | 'bruce' !~~ '%al%'                   |
| ~   | 匹配（正则），区分大小写   | 'thomas' ~ '!.*thomas.*'             |
| ~*  | 匹配（正则），不区分大小写  | 'thomas' ~* '.*Thomas.*'             |
| !~  | 不匹配（正则），区分大小写  | 'thomas' !~ '.*Thomas.*'             |
| !~* | 不匹配（正则），不区分大小写 | 'thomas' !~* '.*vadim.*'             |

## 9.3. 数值操作符

**表 9.3. Postgres 数值操作符**

| 操作符 | 描述       | 用法    |
|-----|----------|-------|
| !   | 阶乘       | 3 !   |
| !!  | 阶乘（左操作符） | !! 3  |
| %   | 取模       | 5 % 4 |
| %   | 截断       | % 4.5 |

| 操作符 | 描述   | 用法        |
|-----|------|-----------|
| *   | 乘法   | 2 * 3     |
| +   | 加法   | 2 + 3     |
| -   | 减法   | 2 - 3     |
| /   | 除法   | 4 / 2     |
| :   | 自然指数 | : 3.0     |
| ;   | 自然对数 | (; 5.0)   |
| @   | 绝对值  | @ -5.0    |
| ^   | 求幂   | 2.0 ^ 3.0 |
| /   | 平方根  | / 25.0    |
| /   | 立方根  | / 27.0    |

## 9.4. 几何操作符

表 9.4. Postgres 几何操作符

| 操作符 | 描述       | 用法                                                |
|-----|----------|---------------------------------------------------|
| +   | 平移       | '((0,0),(1,1))'::box + '(2,0,0)'::point           |
| -   | 平移       | '((0,0),(1,1))'::box - '(2,0,0)'::point           |
| *   | 缩放/旋转    | '((0,0),(1,1))'::box * '(2,0,0)'::point           |
| /   | 缩放/旋转    | '((0,0),(2,2))'::box / '(2,0,0)'::point           |
| #   | 相交       | '((1,-1),(-1,1))' # '((1,1),(-1,-1))'             |
| #   | 多边形中的点数  | # '((1,0),(0,1),(-1,0))'                          |
| ##  | 最近邻点     | '(0,0)'::point ## '((2,0),(0,2))'::lseg           |
| &&  | 重叠吗?     | '((0,0),(1,1))'::box && '((0,0),(2,2))'::box      |
| &<  | 向左重叠吗?   | '((0,0),(1,1))'::box &< '((0,0),(2,2))'::box      |
| &>  | 向右重叠吗?   | '((0,0),(3,3))'::box &> '((0,0),(2,2))'::box      |
| <-> | ...之间的距离 | '((0,0),1)'::circle <-> '((5,0),1)'::circle       |
| <<  | 在其左侧吗?   | '((0,0),1)'::circle << '((5,0),1)'::circle        |
| <^  | 在其下方吗?   | '((0,0),1)'::circle <^ '((0,5),1)'::circle        |
| >>  | 在其右侧吗?   | '((5,0),1)'::circle >> '((0,0),1)'::circle        |
| >^  | 在其上方吗?   | '((0,5),1)'::circle >^ '((0,0),1)'::circle        |
| ?#  | 相交或重叠    | '((-1,0),(1,0))'::lseg ?# '((-2,-2),(2,2))'::box; |

| 操作符 | 描述     | 用法                                                 |
|-----|--------|----------------------------------------------------|
| ?-  | 是水平的吗? | '(1,0)::point ?- '(0,0)::point                     |
| ?   | 是垂直的吗? | '((0,0),(0,1))::lseg ?-  '((0,0),(1,0))::lseg      |
| @-@ | 长度或周长  | @-@ '((0,0),(1,0))::path                           |
| ?   | 是竖直的吗? | '(0,1)::point ?  '(0,0)::point                     |
| ?   | 是平行的吗? | '((-1,0),(1,0))::lseg ?   '((-1,2),(1,2))::lseg    |
| @   | 包含于或位于 | '(1,1)::point @ '((0,0),2)::circle                 |
| @@  | ...的中心 | @@ '((0,0),10)::circle                             |
| ~=  | 等同于    | '((0,0),(1,1))::polygon ~= '((1,1),(0,0))::polygon |

## 9.5. 时间间隔操作符

时间间隔数据类型 `interval` 是最初的日期/时间类型遗留下来的产物，不像更现代的类型那样得到良好的支持。这种类型有若干个操作符。

表 9.5. Postgres 时间间隔操作符

| 操作符 | 描述        | 用法 |
|-----|-----------|----|
| #<  | 间隔小于吗?    |    |
| #<= | 间隔小于或等于吗? |    |
| #<> | 间隔不相等吗?   |    |
| #=  | 间隔相等吗?    |    |
| #>  | 间隔大于吗?    |    |
| #>= | 间隔大于或等于吗? |    |
| <#> | 转换为时间间隔   |    |
| <<  | 间隔小于吗?    |    |
|     | 间隔的起点     |    |
| ~=  | 等同于       |    |
| <?> | 时间在间隔内吗?  |    |

## 9.6. IP V4 操作符

表 9.6. Postgres IP V4 操作符

| 操作符   | 描述    | 用法                                         |
|-------|-------|--------------------------------------------|
| <     | 小于    | '192.168.1.5'::cidr < '192.168.1.6'::cidr  |
| <=    | 小于或等于 | '192.168.1.5'::cidr <= '192.168.1.5'::cidr |
| =     | 等于    | '192.168.1.5'::cidr = '192.168.1.5'::cidr  |
| %gt;= | 大于或等于 | '192.168.1.5'::cidr >= '192.168.1.5'::cidr |

| 操作符  | 描述     | 用法                                            |
|------|--------|-----------------------------------------------|
| %gt; | 大于     | '192.168.1.5'::cidr %gt; '192.168.1.4'::cidr  |
| <>   | 不等于    | '192.168.1.5'::cidr <> '192.168.1.4'::cidr    |
| <<   | 包含于    | '192.168.1.5'::cidr << '192.168.1/24'::cidr   |
| <=<  | 包含于或等于 | '192.168.1/24'::cidr <=< '192.168.1/24'::cidr |
| >>   | 包含     | '192.168.1/24'::cidr >> '192.168.1.5'::cidr   |
| >=>  | 包含或等于  | '192.168.1/24'::cidr >=> '192.168.1/24'::cidr |

## 9.7. IP V4 操作符

表 9.7. PostgresIP V4 操作符

| 操作符   | 描述     | 用法                                            |
|-------|--------|-----------------------------------------------|
| <     | 小于     | '192.168.1.5'::inet < '192.168.1.6'::inet     |
| <=    | 小于或等于  | '192.168.1.5'::inet <= '192.168.1.5'::inet    |
| =     | 等于     | '192.168.1.5'::inet = '192.168.1.5'::inet     |
| %gt;= | 大于或等于  | '192.168.1.5'::inet >= '192.168.1.5'::inet    |
| %gt;  | 大于     | '192.168.1.5'::inet %gt; '192.168.1.4'::inet  |
| <>    | 不等于    | '192.168.1.5'::inet <> '192.168.1.4'::inet    |
| <<    | 包含于    | '192.168.1.5'::inet << '192.168.1/24'::inet   |
| <=<   | 包含于或等于 | '192.168.1/24'::inet <=< '192.168.1/24'::inet |
| >>    | 包含     | '192.168.1/24'::inet >> '192.168.1.5'::inet   |
| >=>   | 包含或等于  | '192.168.1/24'::inet >=> '192.168.1/24'::inet |

# 第 10 章 函数

描述 Postgres 中可用的内置函数。

许多数据类型都有可用于转换到其他相关类型的函数。此外，还有一些类型特定的函数。有些函数也可通过操作符使用，并且可能只作为操作符记录。

## 10.1. 数学函数

表 10.1. 数学函数

| 函数                  | 返回     | 描述       | 例子              |
|---------------------|--------|----------|-----------------|
| dexp(float8)        | float8 | e 的指定幂   | dexp(2.0)       |
| dpow(float8,float8) | float8 | 数的指定幂    | dpow(2.0, 16.0) |
| float(int)          | float8 | 把整数转换为浮点 | float(2)        |
| float4(int)         | float4 | 把整数转换为浮点 | float4(2)       |
| integer(float)      | int    | 把浮点转换为整数 | integer(2.0)    |

## 10.2. 字符串函数

SQL92 用特定语法定义了一些字符串函数。其中一些使用其他 Postgres 函数实现。

表 10.2. SQL92 字符串函数

| 函数                                             | 返回   | 描述        | 例子                            |
|------------------------------------------------|------|-----------|-------------------------------|
| position(text in text)                         | int4 | 指定子串的位置   | position('o' in 'Tom')        |
| substring(text [from int] [for int])           | text | 提取指定子串    | substring('Tom' from 2 for 2) |
| trim([leading trailing both] [text] from text) | text | 移除文本两侧的字符 | trim(both 'x' from 'x Tomx')  |

text、varchar() 和 char() 类型有许多可用的字符串函数。其中一些在内部用于实现上面列出的 SQL92 字符串函数。

表 10.3. 字符串函数

| 函数                  | 返回   | 描述                    | 例子                             |
|---------------------|------|-----------------------|--------------------------------|
| char(text)          | char | 把 text 转换为 char 类型    | char('text string')            |
| char(varchar)       | char | 把 varchar 转换为 char 类型 | char(varchar 'varchar string') |
| initcap(text)       | text | 把每个单词的首字母转换为大写        | initcap('thomas')              |
| lower(text)         | text | 把文本转换为小写              | lower('TOM')                   |
| lpad(text,int,text) | text | 在左侧填充字符串到指定长度         | lpad('hi',4,'??')              |
| ltrim(text,text)    | text | 移除文本左侧的字符             | ltrim('xxxtrim','x')           |

| 函数                      | 返回      | 描述                    | 例子                             |
|-------------------------|---------|-----------------------|--------------------------------|
| position(text,text)     | text    | 提取指定子串                | position('high','ig')          |
| rpad(text,int,text)     | text    | 在右侧填充字符串到指定长度         | rpad('hi',4,'x')               |
| rtrim(text,text)        | text    | 移除文本右侧的字符             | rtrim('trimxxx','x')           |
| substr(text,int[,int])  | text    | 提取指定子串                | substr('hi there',3,5)         |
| text(char)              | text    | 把 char 转换为 text 类型    | text('char string')            |
| text(varchar)           | text    | 把 varchar 转换为 text 类型 | text(varchar 'varchar string') |
| translate(text,from,to) | text    | 转换字符串中的字符             | translate('12345', '1', 'a')   |
| varchar(char)           | varchar | 把 char 转换为 varchar 类型 | varchar('char string')         |
| varchar(text)           | varchar | 把 text 转换为 varchar 类型 | varchar('text string')         |
| upper(text)             | text    | 把文本转换为大写              | upper('tom')                   |

显式为 text 定义的函数大多数也适用于 char() 和 varchar() 参数。

## 10.3. 日期/时间函数

日期/时间函数提供了一套强大的工具，用于操作各种日期/时间类型。

**表 10.4. 日期/时间函数**

| 函数                        | 返回       | 描述           | 例子                                               |
|---------------------------|----------|--------------|--------------------------------------------------|
| abstime(datetime)         | abstime  | 转换为 abstime  | abstime('now'::datetime)                         |
| age(datetime,datetime)    | timespan | 保留月和年的跨度     | age('now','1957-06-13'::datetime)                |
| datetime(abstime)         | datetime | 转换为 datetime | datetime('now'::abstime)                         |
| datetime(date)            | datetime | 转换为 datetime | datetime('today'::date)                          |
| datetime(date,time)       | datetime | 转换为 datetime | datetime('1998-02-24'::datetime, '23:07'::time); |
| date_part(text,datetime)  | float8   | 日期字段的指定部分    | date_part('dow','now'::datetime)                 |
| date_part(text,timespan)  | float8   | 时间字段的指定部分    | date_part('hour','4 hrs 3 mins'::timespan)       |
| date_trunc(text,datetime) | datetime | 按指定单位截断日期    | date_trunc('month','now'::abstime)               |
| isfinite(abstime)         | bool     | 是否为有限时间?     | isfinite('now'::abstime)                         |
| isfinite(datetime)        | bool     | 是否为有限时间?     | isfinite('now'::datetime)                        |
| isfinite(timespan)        | bool     | 是否为有限时间?     | isfinite('4 hrs'::timespan)                      |

| 函数                | 返回       | 描述           | 例子                            |
|-------------------|----------|--------------|-------------------------------|
| reltime(timespan) | reltime  | 转换为 reltime  | reltime('4 hrs':::timespan)   |
| timespan(reltime) | timespan | 转换为 timespan | timespan('4 hours':::reltime) |

对于 `date_part` 和 `date_trunc` 函数，参数可以是 `'year'`、`'month'`、`'day'`、`'hour'`、`'minute'` 和 `'second'`，也可以是更专门的量 `'decade'`、`'century'`、`'millenium'`、`'millisecond'` 和 `'microsecond'`。`date_part` 还允许 `'dow'` 返回星期几，允许 `'epoch'` 返回自 1970 年以来的秒数（用于 `datetime`），或 `'epoch'` 返回总流逝秒数（用于 `timespan`）。

## 10.4. 几何函数

几何类型 `point`、`box`、`lseg`、`line`、`path`、`polygon` 和 `circle` 有一大套原生支持的函数和操作符。

表 10.5. 几何函数

| 函数               | 返回     | 描述         | 例子                                                      |
|------------------|--------|------------|---------------------------------------------------------|
| area(box)        | float8 | 框的面积       | area('((0,0),(1,1))':::box)                             |
| area(circle)     | float8 | 圆的面积       | area('((0,0),2.0)':::circle)                            |
| box(box,box)     | box    | 框转为交集框     | box('((0,0),(1,1))','((0.5,0.5),(2,2))')                |
| center(box)      | point  | 对象的中心      | center('((0,0),(1,2))':::box)                           |
| center(circle)   | point  | 对象的中心      | center('((0,0),2.0)':::circle)                          |
| diameter(circle) | float8 | 圆的直径       | diameter('((0,0),2.0)':::circle)                        |
| height(box)      | float8 | 框的垂直尺寸     | height('((0,0),(1,1))':::box)                           |
| isclosed(path)   | bool   | 是否为闭合路径?   | isclosed('((0,0),(1,1),(2,0))':::path)                  |
| isopen(path)     | bool   | 是否为开放路径?   | isopen('[(0,0),(1,1),(2,0)]':::path)                    |
| length(lseg)     | float8 | 线段的长度      | length('((-1,0),(1,0))':::lseg)                         |
| length(path)     | float8 | 路径的长度      | length('((0,0),(1,1),(2,0))':::path)                    |
| pclose(path)     | path   | 把路径转换为闭合形式 | popen('[(0,0),(1,1),(2,0)]':::path)                     |
| point(lseg,lseg) | point  | 转换为点（交点）   | point('((-1,0),(1,0))':::lseg,'((-2,-2),(2,2))':::lseg) |
| points(path)     | int4   | 路径中的点数     | points('[(0,0),(1,1),(2,0)]':::path)                    |
| popen(path)      | path   | 把路径转换为开放形式 | popen('((0,0),(1,1),(2,0))':::path)                     |

| 函数             | 返回     | 描述     | 例子                             |
|----------------|--------|--------|--------------------------------|
| radius(circle) | float8 | 圆的半径   | radius('((0,0),2.0)')::circle) |
| width(box)     | float8 | 框的水平尺寸 | width('((0,0),(1,1))')::box)   |

表 10.6. 几何类型转换函数

| 函数                   | 返回      | 描述            | 例子                                                      |
|----------------------|---------|---------------|---------------------------------------------------------|
| box(circle)          | box     | 把圆转换为框        | box('((0,0),2.0)')::circle)                             |
| box(point,point)     | box     | 把点转换为框        | box('(0,0)')::point,'(1,1)')::point)                    |
| box(polygon)         | box     | 把多边形转换为框      | box('((0,0),(1,1),(2,0))')::polygon)                    |
| circle(box)          | circle  | 转换为圆          | circle('((0,0),(1,1))')::box)                           |
| circle(point,float8) | circle  | 转换为圆          | circle('(0,0)')::point,2.0)                             |
| lseg(box)            | lseg    | 把对角线转换为 lseg  | lseg('((-1,0),(1,0))')::box)                            |
| lseg(point,point)    | lseg    | 转换为 lseg      | lseg('(-1,0)')::point,'(1,0)')::point)                  |
| path(polygon)        | point   | 转换为路径         | path('((0,0),(1,1),(2,0))')::polygon)                   |
| point(circle)        | point   | 转换为点（中心）      | point('((0,0),2.0)')::circle)                           |
| point(lseg,lseg)     | point   | 转换为点（交点）      | point('((-1,0),(1,0))')::lseg,'((-2,-2),(2,2))')::lseg) |
| point(polygon)       | point   | 多边形的中心        | point('((0,0),(1,1),(2,0))')::polygon)                  |
| polygon(box)         | polygon | 转换为 12 点多边形   | polygon('((0,0),(1,1))')::box)                          |
| polygon(circle)      | polygon | 转换为 12 点多边形   | polygon('((0,0),2.0)')::circle)                         |
| polygon(npts,circle) | polygon | 转换为 npts 点多边形 | polygon(12,'((0,0),2.0)')::circle)                      |
| polygon(path)        | polygon | 转换为多边形        | polygon('((0,0),(1,1),(2,0))')::path)                   |

表 10.7. 几何升级函数

| 函数                   | 返回      | 描述                 | 例子                                          |
|----------------------|---------|--------------------|---------------------------------------------|
| isoldpath(path)      | path    | 测试路径是否为 v6.1 之前的形式 | isoldpath('(1,3,0,0,1,1,2,0)')::path)       |
| revertpoly(polygon)  | polygon | 转换 v6.1 之前的多边形     | revertpoly('((0,0),(1,1),(2,0))')::polygon) |
| upgradepath(path)    | path    | 转换 v6.1 之前的路径      | upgradepath('(1,3,0,0,1,1,2,0)')::path)     |
| upgradepoly(polygon) | polygon | 转换 v6.1 之前的多边形     | upgradepoly('(0,1,2,0,1,0)')::polygon)      |

## 10.5. IP V4 函数

表 10.8. Postgres IP V4 函数

| 函数              | 返回   | 描述          | 例子                                              |
|-----------------|------|-------------|-------------------------------------------------|
| broadcast(cidr) | text | 以文本形式构造广播地址 | broadcast('192.168.1.5/24') ==> '192.168.1.255' |
| broadcast(inet) | text | 以文本形式构造广播地址 | broadcast('192.168.1.5/24') ==> '192.168.1.255' |
| host(inet)      | text | 以文本形式提取主机地址 | host('192.168.1.5/24') ==> '192.168.1.5'        |
| masklen(cidr)   | int4 | 计算网络掩码长度    | masklen('192.168.1.5/24') ==> 24                |
| masklen(inet)   | int4 | 计算网络掩码长度    | masklen('192.168.1.5/24') ==> 24                |
| netmask(inet)   | text | 以文本形式构造网络掩码 | netmask('192.168.1.5/24') ==> '255.255.255.0'   |

---

# 第 11 章 类型转换

SQL 查询可能有意或无意地要求在同一表达式中混用不同的数据类型。Postgres 提供了丰富的机制来求值混合类型表达式。

在很多情况下，用户不需要了解类型转换机制的细节。但 Postgres 所做的隐式转换会影响查询的表面结果，用户或程序员可以用显式类型强制转换来定制这些结果。

本章介绍 Postgres 的类型转换机制和约定。关于特定数据类型及允许的函数和操作符的更多信息，请参阅用户指南和程序员指南中的相关章节。

程序员指南中有关于隐式类型转换和强制转换所用确切算法的更多细节。

## 11.1. 概述

SQL 是一种强类型语言。也就是说，每个数据项都有一个相关联的数据类型，它决定了该数据项的行为和允许的用法。Postgres 拥有一个可扩展的类型系统，比其他 RDBMS 实现更加通用和灵活。因此，Postgres 中大部分类型转换行为应当由一般规则而不是 ad-hoc（特定）启发式规则支配，以使混合类型表达式即使在有用户定义类型时也有意义。

Postgres 的扫描器/解析器只把词法元素解码成五种基本类别：整数、浮点数、字符串、名字和关键字。大多数扩展类型首先被词法化为字符串。SQL 语言定义允许用字符串指定类型名，这一机制被 Postgres 用来让解析器走上正确的路径。例如，查询

```
tgl=> SELECT text 'Origin' AS "Label", point '(0,0)' AS "Value";
Label |Value
-----+-----
Origin|(0,0)
(1 row)
```

中有两个字符串，类型分别为 `text` 和 `point`。如果没有指定类型，则最初会赋予占位类型 `unknown`，并在后面阶段按如下所述解析。

在 Postgres 解析器中有四种基本的 SQL 结构需要专门的类型转换规则：

操作符

Postgres 允许带左一元和右一元（单参数）操作符的表达式，也允许二元（双参数）操作符的表达式。

函数调用

Postgres 类型系统的很大一部分是围绕一套丰富的函数构建的。函数调用有一个或多个参数，对任何具体的查询，这些参数都必须与系统目录中可用的函数匹配。

查询目标

SQL 的 INSERT 语句把查询的结果放到表中。查询中的表达式必须与插入的目标列匹配，必要时还要转换成目标列的类型。

UNION 查询

由于 UNION SELECT 语句的所有查询结果必须出现在一组列中，每个 SELECT 子句的类型必须匹配并转换成统一的集合。

许多一般类型转换规则使用了建立在 Postgres 函数和操作符系统表之上的简单约定。转换规则中还包含一些启发式规则，以便更好地支持 SQL92 标准固有类型（如 `smallint`、`integer` 和 `float`）的约定。

Postgres 解析器使用的约定是：所有类型转换函数都接受一个源类型的参数，并且以目标类型的名称命名。满足这一标准的任何函数都被认为是有效的转换函数，解析器可以把它当作转换函数使用。

这个简单的假设使解析器无需硬编码就能探索类型转换的可能性，让扩展的用户定义类型透明地使用这些相同的特性。

解析器中还提供了一个额外的启发式规则，以便对 SQL 标准类型的正确行为做出更好的猜测。定义了五种类型类别：`boolean`、`string`、`numeric`、`geometric` 和用户自定义。除用户自定义外，每个类别都有一个用于解决候选歧义的“首选类型”。

每个“用户自定义”类型都是自己的“首选类型”，因此只含一个用户定义类型的有歧义表达式（有多个候选解析解的表达式）可以解析出单一的最佳选择，而含多个用户定义类型的表达式仍会有歧义并报错。

候选解只落在一个类型类别内的有歧义表达式很可能得到解析，而候选跨越多个类别的有歧义表达式则很可能报错并请求用户澄清。

### 11.1.1. 指南

所有类型转换规则的设计都考虑了以下几条原则：

- 隐式转换绝不应有令人意外或不可预测的结果。
- 解析器没有先验知识的用户定义类型在类型层级中应该更“高”。在混合类型表达式中，固有类型总是应当转换为用户定义类型（当然，只在需要转换时如此）。
- 用户定义的类型之间没有关联。目前，Postgres 没有关于类型之间关系的信息，只有内置类型的硬编码启发式规则以及基于目录中可用函数的隐式关系。
- 如果查询不需要隐式类型转换，解析器或执行器就不应有额外的开销。也就是说，如果查询构造良好且类型已经匹配，查询就应当继续进行，不必在解析器上花费额外时间，也不必向查询中引入不必要的隐式转换函数。

此外，如果某个查询通常需要为函数做隐式转换，而之后用户用正确的参数类型定义了一个显式函数，解析器就应使用这个新函数，不再用旧函数做隐式转换。

## 11.2. 操作符

### 11.2.1. 转换过程

#### 操作符求值

1. 在 `pg_operator` 系统目录中检查精确匹配。
  - a. (可选的) 如果二元操作符的一个参数是 `unknown`，就假定它与另一个参数的类型相同。
  - b. 交换参数，寻找与一个指向自身为可交换的操作符的精确匹配。如果找到，就在解析树中交换参数并使用该操作符。
2. 寻找最佳匹配。
  - a. (可选的) 列出所有同名的操作符。

- b. 如果列表中只有一个操作符，输入类型可强制转换时就使用它，类型不能强制转换时就报错。
- c. 保留类型显式匹配最多的操作符。如果没有显式匹配则全部保留并进入下一步。如果只剩一个候选，类型可强制转换时就使用它。
- d. 如果有输入参数是"unknown"，把输入候选归类为 boolean、numeric、string、geometric 或用户自定义。如果类别混合，或者用户定义类型多于一个，就报错，因为没有更多线索就无法推断出正确的选择。如果只有一个类别，就把"首选类型"赋给先前为"unknown"的输入列。
- e. 选择类型精确匹配最多且与上一步中每个列类别的 "首选类型"匹配的候选。如果候选仍多于一个，或者一个也没有，就报错。

## 11.2.2. 示例

### 11.2.2.1. 求幂操作符

目录中只定义了一个求幂操作符，它接受 float8 参数。扫描器为这个查询表达式的两个参数都赋予初始类型 int4：

```
tgl=> select 2 ^ 3 AS "Exp";
Exp
---
  8
(1 row)
```

于是解析器对两个操作数都做类型转换，查询等价于

```
tgl=> select float8(2) ^ float8(3) AS "Exp";
Exp
---
  8
(1 row)
```

或

```
tgl=> select 2.0 ^ 3.0 AS "Exp";
Exp
---
  8
(1 row)
```

#### 注意

最后这种形式的开销最小，因为不需要调用任何函数来做隐式类型转换。这对小查询不是问题，但可能影响涉及大表的查询的性能。

### 11.2.2.2. 字符串连接

类字符串的语法既用于处理字符串类型，也用于处理复杂的扩展类型。未指定类型的字符串会与可能的操作符候选匹配。

一个参数未指定类型：

```
tgl=> SELECT text 'abc' || 'def' AS "Text and Unknown";
Text and Unknown
-----
abcdef
(1 row)
```

在这种情况下，解析器会查看是否存在一个两边参数都接受`text`的操作符。既然存在，它就会假定第二个参数应解释为`text`类型。

未指定类型上的串接：

```
tgl=> SELECT 'abc' || 'def' AS "Unspecified";
Unspecified
-----
abcdef
(1 row)
```

这种情况下没有关于使用哪个类型的初始提示，因为查询中没有指定类型。于是解析器查找所有候选操作符，发现所有候选的所有参数都是 `string` 类型。它为此查询选择 `string` 的“首选类型”`text`。

### 注意

如果用户定义了一个新类型并为它定义了“||”操作符来配合它工作，这个查询就会不再按原样成功。解析器此时会有来自两个类别的候选类型，无法决定用哪一个。

#### 11.2.2.3. 阶乘

这个例子展示了一个有趣的结果。传统上阶乘操作符只对整数定义。Postgres 的操作符目录中只有一个阶乘条目，接受整型操作数。如果给定一个非整数的数值参数，Postgres 会尝试把该参数转换成整数来求阶乘。

```
tgl=> select (4.3 !);
?column?
-----
      24
(1 row)
```

### 注意

当然，这会得到数学上可疑的结果，因为原则上非整数的阶乘是没有定义的。不过，数据库的角色不是教数学，而是做数据操作的工具。如果用户选择求浮点数的阶乘，Postgres 会尽量照办。

## 11.3. 函数

## 函数求值

1. 在 pg\_proc 系统目录中检查精确匹配。
2. 寻找最佳匹配。
  - a. 列出所有同名且参数个数相同的函数。
  - b. 如果列表中只有一个函数，输入类型可强制转换时就使用它，类型不能强制转换时就报错。
  - c. 保留类型显式匹配最多的函数。如果没有显式匹配则全部保留并进入下一步。如果只剩一个候选，类型可强制转换时就使用它。
  - d. 如果有输入参数是"unknown"，把输入的候选参数归类为 boolean、numeric、string、geometric 或用户自定义。如果类别混合，或者用户定义类型多于一个，就报错，因为没有更多线索就无法推断出正确的选择。如果只有一个类别，就把"首选类型"赋给先前为"unknown"的输入列。
  - e. 选择类型精确匹配最多且与上一步中每个列类别的 "首选类型"匹配的候选。如果候选仍多于一个，或者一个也没有，就报错。

## 11.3.1. 示例

### 11.3.1.1. 阶乘函数

pg\_proc 目录中只定义了一个阶乘函数。因此下面的查询自动把 int2 参数转换为 int4:

```
tgl=> select int4fac(int2 '4');
int4fac
-----
      24
(1 row)
```

并且实际上被解析器变换为

```
tgl=> select int4fac(int4(int2 '4'));
int4fac
-----
      24
(1 row)
```

### 11.3.1.2. 子串函数

pg\_proc 中声明了两个 substr 函数。但只有一个接受两个参数，类型是 text 和 int4。

如果用未指定类型的字符串常量调用，该类型会直接与唯一的候选函数类型匹配：

```
tgl=> select substr('1234', 3);
substr
-----
      34
(1 row)
```

如果该字符串被声明为 varchar 类型（例如它来自一张表时就是这种情况），则解析器会试图把它强制转换成 text：

```

tgl=> select substr(varchar '1234', 3);
substr
-----
      34
(1 row)

```

它被解析器变换成

```

tgl=> select substr(text(varchar '1234'), 3);
substr
-----
      34
(1 row)

```

### 注意

解析器中有一些启发式规则来优化 `char`、`varchar` 和 `text` 类型之间的关系。在这个例子中，`substr` 直接用 `varchar` 字符串调用，而不是插入显式的转换调用。

而且，如果用 `int4` 调用该函数，解析器会试图把它转换成 `text`：

```

tgl=> select substr(1234, 3);
substr
-----
      34
(1 row)

```

实际执行为

```

tgl=> select substr(text(1234), 3);
substr
-----
      34
(1 row)

```

## 11.4. 查询目标

### 目标求值

1. 检查是否与目标类型精确匹配。
2. 必要时尝试把表达式直接强制转换成目标类型。
3. 如果目标是定长类型（例如声明了长度的 `char` 或 `varchar`），则尝试寻找一个与类型同名、接受两个参数的尺寸调整函数，第一个参数是类型名，第二个是 `integer` 长度。

### 11.4.1. 示例

#### 11.4.1.1. `varchar` 存储

对于声明为 `varchar(4)` 的目标列，下面的查询确保目标的大小正确：

```
tgl=> CREATE TABLE vv (v varchar(4));
CREATE
tgl=> INSERT INTO vv SELECT 'abc' || 'def';
INSERT 392905 1
tgl=> select * from vv;
v
----
abcd
(1 row)
```

## 11.5. UNION 查询

UNION 结构有些不同，它必须把可能不相似的类型匹配起来构成单一结果集。

### UNION 求值

1. 检查所有结果的类型是否一致。
2. 把来自各 UNION 子句的每个结果强制转换成与第一个 SELECT 子句或目标列相同的类型。

### 11.5.1. 示例

#### 11.5.1.1. 未充分指定的类型

```
tgl=> SELECT text 'a' AS "Text" UNION SELECT 'b';
Text
----
a
b
(2 rows)
```

#### 11.5.1.2. 简单 UNION

```
tgl=> SELECT 1.2 AS Float8 UNION SELECT 1;
Float8
-----
1
1.2
(2 rows)
```

#### 11.5.1.3. 转置 UNION

union 的类型被强制匹配 union 中第一个/顶端子句的类型：

```
tgl=> SELECT 1 AS "All integers"
tgl-> UNION SELECT '2.2'::float4
tgl-> UNION SELECT 3.3;
All integers
-----
1
2
3
(3 rows)
```

另一种解析器策略是中选择"最佳"类型，但由于解析器使用了巧妙的递归技术，这更难实现。不过在插入（into）表时会使用"最佳"类型：

```
tgl=> CREATE TABLE ff (f float);
CREATE
tgl=> INSERT INTO ff
tgl-> SELECT 1
tgl-> UNION SELECT '2.2'::float4
tgl-> UNION SELECT 3.3;
INSERT 0 3
tgl=> SELECT f AS "Floating point" from ff;
    Floating point
-----
                1
2.20000004768372
                3.3
(3 rows)
```

---

# 第 12 章 索引和键

Herouth Maoz

## 作者

本文由 Herouth Maoz<sup>1</sup> 撰写

## 编者注

本文最初发表在邮件列表上，是对下面这个问题的回答："PRIMARY KEY 和 UNIQUE 约束有什么区别？"。

Subject: Re: [QUESTIONS] PRIMARY KEY | UNIQUE

下面两者有什么区别：

```
PRIMARY KEY(fields,...) 和  
UNIQUE (fields,...)
```

- 这是一个别名吗？
- 如果 PRIMARY KEY 已经是唯一的，那为什么还会有另一种名为 UNIQUE 的键？

主键是用于标识特定行的字段（或字段组合）。例如，标识一个人的社会保障号码。

单纯的 UNIQUE 字段组合与行的标识无关，它只是一个完整性约束。举个例子：我有一些链接集合。每个集合由一个唯一的编号标识，这个编号就是主键。这个键在关系中使用。

不过，我的应用还要求每个集合有一个唯一的名称。

为什么？这样想要修改集合的人才能识别它。如果你有两个名为 "Life Science" 的集合，就很难知道标记为 24433 的那个才是你需要的，而标记为 29882 的那个不是。

因此，用户通过名称来选择集合。为此我们在数据库中确保名称是唯一的。但是，数据库中没有任何表通过集合名称与 collections 表关联——那样做会非常低效。

此外，尽管是唯一的，集合名称实际上并不定义集合！例如，如果有人决定把集合的名称从 "Life Science" 改为 "Biology"，它仍然是同一个集合，只是名称不同而已。只要名称保持唯一，这就没问题。

因此：

- 主键：
  - 用于标识行并建立到它的关联。
  - 不可能（或很难）更新。
  - 不应允许 NULL。
- 唯一字段（组合）：

---

<sup>1</sup>herouth@oumail.openu.ac.il

- 用作访问行的另一种途径。
- 只要保持唯一就可以更新。
- 可以接受 NULL。

那么，为什么标准 SQL 语法中没有显式定义非唯一键？你必须理解，索引是与实现相关的。SQL 不定义实现，只定义数据库中数据之间的关系。Postgres 确实允许非唯一索引，但用于实施 SQL 键的索引总是唯一的。

因此，你可以按表中列的任意组合查询一个表，即使你在这些列上并没有索引。索引只是每种 RDBMS 向你提供的一种实现上的辅助手段，目的是让常用的查询执行得更加高效。有些 RDBMS 可能会提供额外的措施，例如把键保存在主存中。它们会有一条特殊的命令，例如

```
CREATE MEMSTORE ON <table> COLUMNS <cols>
```

（这不是一条实际存在的命令，只是一个例子）。

事实上，当你创建主键或唯一的字段组合时，SQL 规范中没有任何地方说会创建索引，也没有说通过键检索数据会比顺序扫描更高效！

所以，如果你想用一个非唯一的字段组合作为辅助键，你其实根本不必指定任何东西——直接按那个组合开始检索就行了！但如果你想让检索更加高效，就得求助于你的 RDBMS 供应商提供的手段——无论是索引、我虚构的 MEMSTORE 命令，还是一个聪明的 RDBMS，它会根据你基于特定键组合发送的大量查询，在你不知情的情况下创建索引……（它从经验中学习）。

## 第 13 章 数组

### 注意

这里必须变成一个关于数组行为的章节。有人志愿吗？ - thomas 1998-01-12

Postgres 允许把实例的属性定义为定长或变长的多维数组。  
可以创建任何基本类型或用户定义类型的数组。为了说明它们的用途，我们先创建一个带有基本类型数组的类。

```
CREATE TABLE SAL_EMP (  
    name          text,  
    pay_by_quarter int4[],  
    schedule      text[][]  
);
```

上面的查询将创建一个名为 SAL\_EMP 的类，其中有一个 text 字符串 (name)、一个 int4 的一维数组 (pay\_by\_quarter, 表示员工的季度薪水) 以及一个 text 的二维数组 (schedule, 表示员工的周计划)。现在我们做一些 INSERTS; 注意，向数组追加时，我们把值用花括号括起来并用逗号分隔。如果你了解 C，这很像初始化结构的语法。

```
INSERT INTO SAL_EMP  
VALUES ('Bill',  
    '{10000, 10000, 10000, 10000}',  
    '{{"meeting", "lunch"}, {}}');  
  
INSERT INTO SAL_EMP  
VALUES ('Carol',  
    '{20000, 25000, 25000, 25000}',  
    '{{"talk", "consult"}, {"meeting"}}');
```

默认情况下，Postgres 对数组使用 "one-based" (从 1 开始) 的编号约定——也就是说，一个有 n 个元素的数组从 array[1] 开始，以 array[n] 结束。现在，我们可以在 SAL\_EMP 上运行一些查询。首先，我们展示如何一次访问数组的一个元素。下面的查询检索第二季度薪水发生变化的员工的姓名：

```
SELECT name  
FROM SAL_EMP  
WHERE SAL_EMP.pay_by_quarter[1] <>  
SAL_EMP.pay_by_quarter[2];
```

```
+-----+  
|name  |  
+-----+  
|Carol |  
+-----+
```

这个查询检索所有员工的第三季度薪水：

```
SELECT SAL_EMP.pay_by_quarter[3] FROM SAL_EMP;
```

```
+-----+
|pay_by_quarter |
+-----+
|10000          |
+-----+
|25000          |
+-----+
```

我们也可以访问数组的任意切片或子数组。下面的查询检索 Bill 周计划前两天的第一项。

```
SELECT SAL_EMP.schedule[1:2][1:1]
FROM SAL_EMP
WHERE SAL_EMP.name = 'Bill';
```

```
+-----+
|schedule      |
+-----+
|{"meeting"}, {""} |
+-----+
```

---

## 第 14 章 继承

我们来创建两个类。capitals 类包含各州首府，它们同时也是城市。自然地，capitals 类应当继承自 cities。

```
CREATE TABLE cities (  
    name            text,  
    population      float,  
    altitude        int          -- (in ft)  
);  
  
CREATE TABLE capitals (  
    state           char2  
) INHERITS (cities);
```

在这种情况下，capitals 的一个实例从其父类 cities 继承所有属性（name、population 和 altitude）。属性 name 的类型是 text，这是 Postgres 用于变长 ASCII 字符串的原生类型。属性 population 的类型是 float，这是 Postgres 用于双精度浮点数的原生类型。州首府有一个额外的属性 state，表示它所在的州。在 Postgres 中，一个类可以从零个或多个其他类继承，而查询既可以只引用一个类的所有实例，也可以引用一个类及其所有后代的所有实例。

### 注意

继承层次实际上是一个有向无环图。

例如，下面的查询找出所有位于海拔 500 英尺或以上的城市：

```
SELECT name, altitude  
FROM cities  
WHERE altitude > 500;
```

```
+-----+-----+  
|name      | altitude |  
+-----+-----+  
|Las Vegas | 2174     |  
+-----+-----+  
|Mariposa  | 1953     |  
+-----+-----+
```

另一方面，要找出所有位于海拔 500 英尺以上的城市（包括州首府）的名称，查询是：

```
SELECT c.name, c.altitude  
FROM cities* c  
WHERE c.altitude > 500;
```

它返回：

```
+-----+-----+  
|name      | altitude |  
+-----+-----+  
|Las Vegas | 2174     |  
+-----+-----+  
|Mariposa  | 1953     |
```

```
+-----+-----+
|Madison  | 845    |
+-----+-----+
```

这里 cities 后面的 “\*” 表示查询应在 cities 以及继承层次中位于 cities 之下的所有类上执行。我们已经讨论过的许多命令——select、update 和 delete——都支持这种 “\*” 记法，其他命令如 alter 也是如此。

---

# 第 15 章 查询语言

## 注意

本章必须深入讲解查询语言的每个领域。目前是教程的一份副本。- thomas 1998-01-12

Postgres 查询语言是 SQL3 的一个变体。它有许多扩展，例如可扩展的类型系统、继承、函数和产生式规则。这些是从原始的 Postgres 查询语言 PostQuel 继承来的特性。本节概述如何使用 Postgres SQL 执行简单的操作。本手册只想让你了解我们的 SQL 的风味，绝非完整的 SQL 教程。关于 SQL 已有许多著作。例如，可查阅 Understanding the New SQL 或 A Guide to the SQL Standard。你还应注意 Postgres 的某些特性并不属于 ANSI 标准。

## 15.1. 概念

Postgres 中的基本概念是类 (class)，它是对象实例的一个命名集合。每个实例都具有同一组命名属性，而每个属性都有特定的类型。此外，每个实例都有一个在整个安装中唯一的永久对象标识符 (OID)。由于 SQL 语法使用表的说法，我们将把术语表 (table) 和类 (class) 互换使用。同样，SQL 的行 (row) 就是实例 (instance)，而 SQL 的列 (column) 就是属性 (attribute)。如前所述，类被组织成数据库，而由单个 postmaster 进程管理的一组数据库构成一个安装或站点。

## 15.2. 创建新类

你可以通过指定类名以及所有属性名及其类型来创建一个新类：

```
CREATE TABLE weather (  
    city          varchar(80),  
    temp_lo       int,          -- low temperature  
    temp_hi       int,          -- high temperature  
    prcp          real,         -- precipitation  
    date          date  
);
```

注意关键字是大小写不敏感的，标识符通常也是大小写不敏感的。Postgres 允许 SQL92 的分隔标识符（由双引号包围的标识符）包含大小写混合以及空格、制表符等。

Postgres SQL 支持通常的 SQL 类型 int、float、real、smallint、char(N)、varchar(N)、date、time、timestamp，以及其他通用类型和一组丰富的几何类型。后面会看到，Postgres 可以用任意数量的用户自定义数据类型来定制。因此，类型名不是语法关键字，除非为支持 SQL92 标准中的特殊情况所必需。到目前为止，Postgres 的 create 命令看起来与传统关系系统中创建表所用的命令完全一样。但我们马上就会看到，类具有一些关系模型扩展性质的属性。

## 15.3. 用实例填充类

insert 语句用于向类中填充实例：

```
INSERT INTO weather
```

```
VALUES ('San Francisco', 46, 50, 0.25, '11/27/1994')
```

你还可以使用 `copy` 命令从平面 (ASCII) 文件装载大量数据。这通常更快，因为数据作为单个原子事务直接读入（或写出）目标表。例如：

```
COPY INTO weather FROM '/home/user/weather.txt'
  USING DELIMITERS '|';
```

其中源文件的路径名必须为后端服务器机器可访问，而不仅仅是客户端。

## 15.4. 查询类

`weather` 类可以用正常的关系选择和投影查询来查询。这通过 SQL 的 `select` 语句完成。该语句分为选择列表（列出要返回的属性的部分）和限定条件（指定限制条件的部分）。例如，要检索 `weather` 的所有行，键入：

```
SELECT * FROM WEATHER;
```

输出应为：

```
+-----+-----+-----+-----+-----+
|city      | temp_lo | temp_hi | prcp  | date      |
+-----+-----+-----+-----+-----+
|San Francisco | 46      | 50      | 0.25  | 11-27-1994 |
+-----+-----+-----+-----+-----+
|San Francisco | 43      | 57      | 0      | 11-29-1994 |
+-----+-----+-----+-----+-----+
|Hayward      | 37      | 54      |       | 11-29-1994 |
+-----+-----+-----+-----+-----+
```

你可以在选择列表中指定任意表达式。例如：

```
SELECT city, (temp_hi+temp_lo)/2 AS temp_avg, date FROM weather;
```

任意布尔操作符 (`and`、`or` 和 `not`) 都可用于任何查询的限定条件中。例如，

```
SELECT * FROM weather
  WHERE city = 'San Francisco'
  AND prcp > 0.0;
```

```
+-----+-----+-----+-----+-----+
|city      | temp_lo | temp_hi | prcp  | date      |
+-----+-----+-----+-----+-----+
|San Francisco | 46      | 50      | 0.25  | 11-27-1994 |
+-----+-----+-----+-----+-----+
```

最后要说明的是，你可以指定 `select` 的结果以排序顺序返回，或者去除重复实例。

```
SELECT DISTINCT city
  FROM weather
  ORDER BY city;
```

## 15.5. 重定向 SELECT 查询

任何 `select` 查询都可以被重定向到一个新类

```
SELECT * INTO TABLE temp FROM weather;
```

这构成一个隐式的 `create` 命令，按照 `select into` 命令的选择列表中指定的属性名和类型创建新类 `temp`。当然，随后我们可以对这个结果类执行任何可对其他类执行的操作。

## 15.6. 类之间的连接

到目前为止，我们的查询一次只访问一个类。查询可以一次访问多个类，或者以某种方式访问同一个类，使得该类的多个实例同时被处理。一次访问同一个或不同类的多个实例的查询称为连接查询。例如，假设我们希望找出所有处于其他记录温度范围内的记录。实际上，我们需要把每个 `EMP` 实例的 `temp_lo` 和 `temp_hi` 属性与所有其他 `EMP` 实例的 `temp_lo` 和 `temp_hi` 属性进行比较。

### 注意

这只是一种概念模型。实际的连接可能以更高效的方式执行，但这对用户是不可见的。

我们可以用以下查询做到这一点：

```
SELECT W1.city, W1.temp_lo, W1.temp_hi,
       W2.city, W2.temp_lo, W2.temp_hi
FROM   weather W1, weather W2
WHERE  W1.temp_lo < W2.temp_lo
AND    W1.temp_hi > W2.temp_hi;
```

```
+-----+-----+-----+-----+-----+-----+
--+
|city      | temp_lo | temp_hi | city      | temp_lo | temp_hi |
|          |         |         |          |         |         |
+-----+-----+-----+-----+-----+-----+
--+
|San Francisco | 43      | 57      | San Francisco | 46      | 50      |
|          |         |         |          |         |         |
+-----+-----+-----+-----+-----+-----+
--+
|San Francisco | 37      | 54      | San Francisco | 46      | 50      |
|          |         |         |          |         |         |
+-----+-----+-----+-----+-----+-----+
--+
```

### 注意

这种连接的语义是：限定条件是为查询中指定的各类的笛卡尔积定义的真值表达式。对于笛卡尔积中限定条件为真的那些实例，Postgres 计算并返回选择列表中指定的值。Postgres SQL 不为此类表达式中的重复值赋予任何意义。这意味着 Postgres 有时会重复计算同一选择列表多次；当布尔表达式用 "or" 连接时经常发生这种情况。要去除此类重复，必须使用 `select distinct` 语句。

在本例中，`W1` 和 `W2` 都是类 `weather` 一个实例的代理（surrogate），二者都取遍该类的所有实例。（用大多数数据库系统的术语来说，`W1` 和 `W2` 被称为范围变量。）查询可以包含任意数量的类名和代理。

## 15.7. 更新

你可以使用 `update` 命令更新现有实例。假设你发现从 11 月 28 日起所有温度读数都偏高 2 度，可以这样更新数据：

```
UPDATE weather
    SET temp_hi = temp_hi - 2, temp_lo = temp_lo - 2
    WHERE date > '11/28/1994';
```

## 15.8. 删除

删除使用 `delete` 命令执行：

```
DELETE FROM weather WHERE city = 'Hayward';
```

所有属于 Hayward 的天气记录都被删除。对如下形式的查询应当小心

```
DELETE FROM classname;
```

不带限定条件时，`delete` 会直接删除给定类的所有实例，使其变空。系统在执行此操作前不会请求确认。

## 15.9. 使用聚合函数

与大多数其他查询语言一样，PostgreSQL 支持聚合函数。当前 Postgres 聚合函数的实现有一些限制。具体来说，虽然有计算 `count`、`sum`、`avg`（平均值）、`max`（最大值）和 `min`（最小值）等函数（在一组实例上）的聚合，聚合只能出现在查询的选择列表中，而不能直接出现在限定条件（`where` 子句）中。例如，

```
SELECT max(temp_lo) FROM weather;
```

是允许的，而

```
SELECT city FROM weather WHERE temp_lo = max(temp_lo);
```

则不行。不过，正如常见的情况，查询可以改写以达到预期的结果；这里使用子查询：

```
SELECT city FROM weather WHERE temp_lo = (SELECT max(temp_lo) FROM
weather);
```

聚合还可以带 `group by` 子句：

```
SELECT city, max(temp_lo)
    FROM weather
   GROUP BY city;
```

---

# 第 16 章 磁盘存储

本节尚待编写。FAQ 中有一些相关信息。有人愿意吗？ - thomas 1998-01-11

---

# 第 17 章 `psql`

本节需要从原始的 `psql` 手册页转换而来。欢迎志愿者！

## 提示

要更改结果的分页行为，请设置/取消设置你的 `PAGER` 环境变量。

## 17.1. 分页到屏幕

## 作者

来自 Brett McCormick 1998-04-04 在邮件列表上的文章。

要影响 `psql` 输出的分页行为，请设置或取消设置你的 `PAGER` 环境变量。我总是必须先设置它才会暂停。当然，你必须在启动程序之前完成此操作。

在 `csh/tcsh` 或其他 C shell 中：

```
unsetenv PAGER
```

而在 `sh/bash` 或其他 Bourne shell 中：

```
unset PAGER
```

---

# 第 18 章 pgaccess

本节有待撰写。欢迎志愿者！

---

# 第 19 章 SQL 命令

这里是 Postgres 所支持的 SQL 命令的参考信息。

# ABORT

ABORT — 中止当前事务

## 大纲

ABORT

## 输入

无。

## 输出

ABORT

成功时返回的消息。

NOTICE: UserAbortTransactionBlock and not in in-progress state ABORT

当前没有正在进行的事务时给出。

## 描述

ABORT回卷当前事务并使该事务所做的所有更新被丢弃。  
这条命令的行为与SQL92命令ROLLBACK 完全相同，其存在仅出于历史原因。

## 注解

使用COMMIT语句来成功终止一个事务。

## 用法

```
--To abort all changes
--
ABORT WORK;
```

## 兼容性

## SQL92

这条命令是一个因历史原因而保留的Postgres 扩展。ROLLBACK是等价的SQL92 命令。

# ALTER TABLE

ALTER TABLE — 修改表属性

## 大纲

```
ALTER TABLE table
    [ * ] ADD [ COLUMN ] column type
ALTER TABLE table
    [ * ] RENAME [ COLUMN ] column TO newcolumn
ALTER TABLE table
    RENAME TO newtable
```

## Inputs

*table*

要更改的现有表的名称。

*column*

新列或现有列的名称。

*type*

新列的类型。

*newcolumn*

现有列的新名称。

*newtable*

现有列的新名称。

## Outputs

→ ALTER

列或表重命名后返回的消息。

→ NEW

添加列后返回的消息。

→ ERROR

表或列不可用时返回的消息。

## Description

ALTER TABLE 更改一个现有表的定义。新列及其类型的指定风格和限制与 CREATE TABLE 中相同。RENAME 子句更改表或列的名称而不更改受影响表中包含的任何数据。因此，该命令执行后，表或列仍保持相同的类型和大小。

要更改表的模式，你必须拥有该表。

## Notes

关键字 COLUMN 只是噪声，可以省略。

“[\*]”跟在表名之后表示该语句应作用于该表以及继承层次中位于其下的所有表。PostgreSQL 用户指南中还有关于继承的更多信息。

有关有效参数的进一步说明，请参见 CREATE TABLE。

## Usage

要添加类型为 VARCHAR 的列到表中：

```
ALTER TABLE distributors ADD COLUMN address VARCHAR(30);
```

要重命名一个现有列：

```
ALTER TABLE distributors RENAME COLUMN address TO city;
```

要重命名一个现有的表：

```
ALTER TABLE distributors RENAME TO suppliers;
```

## Compatibility

### SQL92

ALTER TABLE/RENAME 是 Postgres 的语言扩展。

SQL92 为 ALTER TABLE 语句规定了一些 Postgres 尚不直接支持的附加能力：

```
ALTER TABLE table ALTER [ COLUMN ] column
SET DEFAULT default
```

```
ALTER TABLE table ALTER [ COLUMN ] column
ADD [ CONSTRAINT constraint ] table-constraint
```

把指定的默认值或约束放入表中该列的定义。默认值和表约束子句的语法见 CREATE TABLE。如果默认值子句已存在，它将被新定义替换。如果该列上已有任何约束，它们将通过与新约束的布尔 AND 关系保留。

目前，要为现有列设置新的默认约束，必须重建并重新装载该表：

```
CREATE TABLE temp AS SELECT * FROM distributors;
DROP TABLE distributors;
CREATE TABLE distributors (
    did      DECIMAL(3) DEFAULT 1,
    name     VARCHAR(40) NOT NULL,
    city     VARCHAR(30)
);
INSERT INTO distributors SELECT * FROM temp;
DROP TABLE temp;
```

```
ALTER TABLE table
    DROP DEFAULT default
ALTER TABLE table
    DROP CONSTRAINT constraint { RESTRICT | CASCADE }
```

从表的定义中移除由 *default* 指定的默认值或由 *constraint* 指定的规则。如果指定了 RESTRICT，只有没有依赖约束的约束才能被销毁。如果指定了 CASCADE，任何依赖于该约束的约束也会被删除。

目前，要移除现有列上的默认值或约束，必须重建并重新装载该表：

```
CREATE TABLE temp AS SELECT * FROM distributors;
DROP TABLE distributors;
CREATE TABLE distributors AS SELECT * FROM temp;
DROP TABLE temp;
```

```
ALTER TABLE table
    DROP [ COLUMN ] column { RESTRICT | CASCADE }
```

从表中移除一个列。如果指定了 RESTRICT，只有没有依赖对象的列才能被销毁。如果指定了 CASCADE，所有依赖于该列的对象也会被删除。

目前，要移除一个现有列，必须重建并重新装载该表：

```
CREATE TABLE temp AS SELECT did, city FROM distributors;
DROP TABLE distributors;
CREATE TABLE distributors (
    did      DECIMAL(3)  DEFAULT 1,
    name     VARCHAR(40) NOT NULL,
);
INSERT INTO distributors SELECT * FROM temp;
DROP TABLE temp;
```

# ALTER USER

ALTER USER — 修改用户账户信息

## 大纲

```
ALTER USER username
    [ WITH PASSWORD password ]
    [ CREATEDB | NOCREATEDB ]
    [ CREATEUSER | NOCREATEUSER ]
    [ IN GROUP groupname [, ...] ]
    [ VALID UNTIL 'abstime' ]
```

## 输入

各子句的详细描述参见 CREATE USER。

*username*

要更改其详细信息的用户的 Postgres 账户名。

*password*

此帐号要使用的新密码。

*groupname*

该账户将被放入的访问组的名称。

*abstime*

该用户的访问将被终止的日期（以及可选的时间）。

## 输出

→ ALTER USER

更改成功时返回的消息。

→ ERROR: alterUser: user "username" does not exist

如果指定的用户不存在，则返回此错误消息。

## 描述

ALTER USER 用于更改用户 Postgres 账户的属性。请注意，无法通过 alter user 语句更改用户的"usesysid"。而且，只有 Postgres 用户或任何对 "pg\_shadow" 具有读和修改权限的用户才能更改用户口令。

如果省略了 alter user 语句的任何子句，"pg\_shadow" 表中相应的值保持不变。

## 注意

ALTER USER 语句是一种 Postgres 语言扩展。

要创建或删除用户账户，参见 CREATE/DROP USER。

在当前版本（v6.4）中，IN GROUP 子句会被解析但不产生作用。当它被完全实现时，将用于修改 pg\_group 关系。

## 用法

更改用户口令

```
ALTER USER davide WITH PASSWORD hu8jmn3;
```

更改用户的有效期限日期

```
ALTER USER manuel VALID UNTIL 'Jan 31 2030';
```

更改用户的有效期限日期，指定他的授权应在 1998 年 5 月 4 日正午、使用比 UTC 早一小时的时区过期

```
ALTER USER chris VALID UNTIL 'May 4 12:00:00 1998 +1';
```

授予用户创建其他用户和新数据库的能力。

```
ALTER USER miriam CREATEUSER CREATEDB;
```

把一个用户放入两个组

```
ALTER USER miriam IN GROUP sales, payroll;
```

## 兼容性

### SQL92

SQL92 中没有 ALTER USER 语句。该标准把用户的定义留给实现。

# BEGIN WORK

BEGIN WORK — 开始一个事务

## 大纲

```
BEGIN [ WORK | TRANSACTION ]
```

## 输入

无

## 输出

→ BEGIN

这表示一个新的事务已经启动。

→ NOTICE: BeginTransactionBlock and not in default state

这表示已经有一个事务在进行中。当前事务不受影响。

## 描述

**BEGIN** 开启一个用户事务，Postgres 将保证它相对于所有并发执行的事务是可串行化的。Postgres 使用两阶段锁来完成这一任务。如果事务被提交，Postgres 将确保要么所有更新都完成，要么一个都不完成。事务具有标准的 ACID（原子性、一致性、隔离性、持久性）属性。

## 注意

关键字 TRANSACTION 只是 WORK 的一个修饰性替代。两个关键字都不必指定。

关于在事务中锁定表的进一步信息，请参阅 LOCK 语句。

使用 COMMIT 或 ROLLBACK 终止一个事务。

## 用法

开始一个用户事务：

```
BEGIN WORK;
```

## 兼容性

BEGIN 是 Postgres 语言扩展。

## SQL92

SQL92 中没有显式的 BEGIN WORK 命令；事务的开启总是隐式的，它以 COMMIT 或 ROLLBACK 语句终止。

# CLOSE

CLOSE — 关闭一个游标

## 大纲

```
CLOSE cursor
```

## 输入

```
cursor
```

要关闭的打开游标的名称。

## 输出

→ CLOSE

游标成功关闭时返回的消息。

→ NOTICE PerformPortalClose: portal "*cursor*" not found

如果 *cursor* 没有被声明或者已经被关闭，就会给出这条警告。

## 描述

CLOSE 释放与一个打开的游标相关联的资源。游标关闭后，不允许再对其执行任何后续操作。当不再需要游标时，应将其关闭。

当事务被 COMMIT 或 ROLLBACK 终止时，每个打开的游标都会被隐式关闭。

## 注意

Postgres 没有显式的 OPEN 游标语句；游标在声明时即被视为打开。请使用 DECLARE 语句声明游标。

## 用法

关闭游标 *liahona*：

```
CLOSE liahona;
```

## 兼容性

### SQL92

CLOSE 与 SQL92 完全兼容。

# CLUSTER

CLUSTER — 向后端给出存储聚簇建议

## 大纲

```
CLUSTER indexname ON table
```

## 输入

*indexname*

一个索引的名称。

*table*

一个表的名称。

## 输出

→ CLUSTER

聚簇已成功完成。

→ ERROR: relation <tablerepresentation\_number> inherits "invoice"

→ ERROR: Relation x does not exist!

## 描述

CLUSTER指示 Postgres近似地基于 *indexname*所指定的索引，对 *classname* 所指定的类进行聚簇。该索引必须已经定义在 *classname*上。

当一个类被聚簇时，它会基于索引信息被物理地重新排序。聚簇是静态的。换句话说，随着类被更新，这些更改不会被聚簇。系统不会尝试让新实例或更新过的元组保持聚簇。如果需要，可以通过再次发出该命令来手动重新聚簇。

## 注意

表实际上被按索引顺序复制到一个临时表中，然后改回原来的名称。因此，执行聚簇时会丢失所有 GRANT 权限和其他索引。

如果你在表中随机访问单个行，数据在堆表中的实际顺序并不重要。但如果你倾向于比其他数据更常访问某些数据，并且有一个索引把它们聚集在一起，使用 CLUSTER会带来好处。

CLUSTER有帮助的另一个场合是使用索引从一个表中取出多行时。如果你从一个表中请求某个索引值范围，或者一个有多行匹配的单个索引值，CLUSTER

会有帮助，因为一旦索引识别出第一行匹配所在的堆页，所有其他匹配的行很可能已经在同一个堆页上，这样就节省了磁盘访问并加快了查询。

有两种聚簇数据的方法。第一种是使用 `CLUSTER` 命令，它使用你指定的索引的顺序对原表重新排序。在大表上这可能很慢，因为行是按索引顺序从堆中取出的，如果堆表是无序的，各项就散布在随机页面上，因此每移动一行就要检索一个磁盘页。Postgres有缓存，但大表的大部分放不进缓存。

另一种聚簇数据的方法是使用

```
SELECT ... INTO TABLE temp FROM ... ORDER BY ...
```

这在 `ORDER BY` 中使用 Postgres的排序代码来匹配索引，对于无序数据它要快得多。然后你删除旧表，使用 `ALTER TABLE/RENAME` 把 *temp* 改名为旧名称，并重建任何索引。唯一的问题是 `OID` 不会被保留。此后，`CLUSTER` 应该会很快，因为堆数据大部分已经有序，并且使用的是现有的索引。

## 用法

基于其 `salary` 属性聚簇 `employees` 关系

```
CLUSTER emp_ind ON emp
```

## 兼容性

### SQL92

SQL92 中没有 `CLUSTER` 语句。

# COMMIT

COMMIT — 提交当前事务

## 大纲

```
COMMIT [ WORK | TRANSACTION ]
```

## 输入

无

## 输出

→ END

事务成功提交时返回的消息。

→ NOTICE EndTransactionBlock and not inprogress/abort state

如果当前没有正在进行的事务。

## 描述

COMMIT提交当前事务。该事务所做的所有更改都会对其他会话可见，并且如果发生崩溃，也能保证其持久性。

## 注意

关键字 WORK 和 TRANSACTION 只是噪音，可以省略。

使用 ROLLBACK 语句中止一个事务。

## 用法

要使所有更改永久生效：

```
COMMIT WORK;
```

## 兼容性

### SQL92

完全兼容。

# COPY

COPY — 在文件和表之间复制数据

## 大纲

```
COPY [ BINARY ] table [ WITH OIDS ]  
    FROM { 'filename' | stdin }  
    [ USING DELIMITERS 'delimiter' ]  
COPY [ BINARY ] table [ WITH OIDS ]  
    TO { 'filename' | stdout }  
    [ USING DELIMITERS 'delimiter' ]
```

## 输入

### BINARY

更改字段格式化的行为，强制所有数据以二进制对象而不是文本格式存储或读取。

*table*

一个现有表的名称。

### WITH OIDS

为每一行复制内部唯一对象 id (OID)。

*filename*

输入或输出文件的绝对 Unix 路径名。

stdin

指定输入来自管道或终端。

stdout

指定输出到管道或终端。

*delimiter*

分隔输入或输出字段的字符。

## 输出

→ COPY

复制成功完成。

→ ERROR: *error message*

复制因错误消息中所述的原因而失败。

## 描述

`COPY` 在 Postgres 表和标准 Unix 文件之间移动数据。

`COPY` 指示 Postgres 后端直接读取或写入文件。该文件必须对后端直接可见，且名称必须从后端的角度指定。如果指定了 `stdin` 或 `stdout`，数据通过客户端前端流经后端。

## 注解

**BINARY** 关键字强制所有数据以二进制对象而不是文本格式存储/读取。它比普通的复制命令略快，但通常不可移植，且生成的文件略大，尽管这一因素高度依赖于数据本身。默认情况下，文本复制使用制表符（`"\t"`）作为分隔符。也可以用关键字短语 `USING DELIMITERS` 把分隔符改为任何其他单个字符。数据字段中恰好匹配分隔符的字符将被引用。

你必须对其值被 `COPY` 读取的任何表拥有 `select` 访问权限，并对正被 `COPY` 插入值的表拥有 `insert` 或 `update` 访问权限。对于被 `COPY` 读取或写入的任何文件，后端还需要适当的 Unix 权限。

关键字短语 `USING DELIMITERS` 指定一个用于列之间所有分隔符的单个字符。如果在分隔符字符串中指定了多个字符，只使用第一个字符。

### 提示

不要把 `COPY` 与 `psql` 指令 `\copy` 混淆。

## 文件格式

### 文本格式

当 `COPY TO` 不带 `BINARY` 选项使用时，生成的文件中每一行（实例）占一行，各列（属性）由分隔字符分隔。嵌入的分隔字符前面会有一个反斜线字符（`"\"`）。

属性值本身是由与各属性类型关联的输出函数生成的字符串。

类型的输出函数不应尝试生成反斜线字符；这将由 `COPY` 自身处理。

每个实例的实际格式为

```
<attr1><separator><attr2><separator>...<separator><attrn><newline>
```

如果指定了 `WITH OIDS`，`oid` 放在行的开头。

如果 `COPY` 把输出发送到标准输出而不是文件，完成时它将在单独一行上发送一个反斜线（`"\"`）和一个句点（`"."`）后紧跟一个换行符。类似地，如果 `COPY` 从标准输入读取，它将期待一个反斜线（`"\"`）和一个句点（`"."`）后跟一个换行符，作为一行上表示文件结束的前三个字符。不过，如果在找到这个特殊的文件结束模式之前遇到真正的 `EOF`，`COPY` 将终止（随后后端自身也终止）。

反斜线字符还有其他特殊含义。`NULL` 属性输出为 `"\N"`。字面反斜线字符输出为两个连续的反斜线（`"\\"`）。字面制表符表示为一个反斜线加一个制表符。

字面换行符表示为一个反斜线加一个换行符。在加载不是由 Postgres 生成的文本数据时，你需要把反斜线字符（`"\"`）转换为双反斜线（`"\\"`）以确保正确加载。

## 二进制格式

在 `COPY BINARY` 的情况下，文件的前四个字节是文件中的实例数。如果此数为零，`COPY BINARY` 命令将读取直到遇到文件结尾。否则，读满这个数目的实例后将停止读取。文件中的其余数据将被忽略。

文件中每个实例的格式如下。注意必须严格遵循此格式。下表中无符号四字节整数称为 `uint32`。

**表 19.1. 二进制复制文件的内容**

| 在文件的开头                           |               |
|----------------------------------|---------------|
| <code>uint32</code>              | 元组的数目         |
| 对于每个元组                           |               |
| <code>uint32</code>              | 元组数据的总长度      |
| <code>uint32</code>              | oid (如果指定)    |
| <code>uint32</code>              | 空属性的数目        |
| <code>[uint32,...,uint32]</code> | 属性编号 (从 0 计数) |
| -                                | <tuple data>  |

## 二进制数据的对齐

在 Sun-3 上，2 字节属性按两字节边界对齐，所有更大的属性按四字节边界对齐。字符属性按单字节边界对齐。在大多数其他机器上，所有大于 1 字节的属性都按四字节边界对齐。注意变长属性前面有该属性的长度；数组只是数组元素类型的连续流。

## 用法

下面的例子把一个表复制到标准输出，使用竖线 (|) 作为字段分隔符：

```
COPY country TO stdout USING DELIMITERS '|';
```

把数据从一个 Unix 文件复制到表 "country"：

```
COPY country FROM '/usr1/proj/bray/sql/country_data';
```

下面是一段适合从 `stdin` 复制到表中的数据示例（所以最后一行有终止序列）：

```
AF      AFGHANISTAN
AL      ALBANIA
DZ      ALGERIA
...
ZM      ZAMBIA
ZW      ZIMBABWE
\.
```

相同的数据在一台 Linux/i586 机器上以二进制格式输出。数据经过 Unix 工具 `od -c` 过滤后显示。该表有三个字段；第一个是 `char(2)`，第二个是 `text`。所有行的第三个字段都是空值。注意 `char(2)` 字段如何用空字节填充到四字节，`text` 字段前面有它的长度：

```
355  \0  \0  \0 027  \0  \0  \0 001  \0  \0  \0 002  \0  \0  \0
```

```

006  \0  \0  \0  A  F  \0  \0 017  \0  \0  \0  A  F  G  H
      A  N  I  S  T  A  N 023  \0  \0  \0 001  \0  \0  \0 002
      \0  \0  \0 006  \0  \0  \0  A  L  \0  \0  \v  \0  \0  \0  A
      L  B  A  N  I  A 023  \0  \0  \0 001  \0  \0  \0 002  \0
      \0  \0 006  \0  \0  \0  D  Z  \0  \0  \v  \0  \0  \0  A  L
      G  E  R  I  A
...
      \n  \0  \0  \0  Z  A  M  B  I  A 024  \0
      \0  \0 001  \0  \0  \0 002  \0  \0  \0 006  \0  \0  \0  Z  W
      \0  \0  \f  \0  \0  \0  Z  I  M  B  A  B  W  E

```

## 缺陷

`COPY` 在第一个错误处停止操作。这对 `COPY FROM` 不会导致问题，但在 `COPY TO` 中目标关系当然会被部分修改。失败的复制之后应当用 `VACUUM` 查询清理。

由于 Postgres 后端的当前工作目录通常与用户的工作目录不同，复制到文件 "foo"（不带附加路径信息）可能给天真的用户带来意外的结果。在这种情况下，foo 最终会出现在 `$PGDATA/foo` 中。一般而言，指定要复制的文件时应使用后端服务器机器上所见完整路径名。

用作 `COPY` 参数的文件必须位于数据库服务器机器上或可由其访问，方法是在本地磁盘上或网络文件系统上。

当使用从一台机器到另一台机器的 TCP/IP 连接并指定了目标文件时，目标文件将写入后端运行的机器而不是用户的机器。

## 兼容性

### SQL92

在 SQL92 中没有 `COPY` 语句。

# CREATE AGGREGATE

CREATE AGGREGATE — 定义一个新的聚合函数

## 大纲

```
CREATE AGGREGATE name [ AS ]
    ( BASETYPE      = data_type
    [ , SFUNC1      = sfunc1
    [ , STYPE1      = sfunc1_return_type ]
    [ , SFUNC2      = sfunc2
    [ , STYPE2      = sfunc2_return_type ]
    [ , FINALFUNC   = ffunc ]
    [ , INITCOND1   = initial_condition1 ]
    [ , INITCOND2   = initial_condition2 ]
    )
```

## 输入

*name*

要创建的聚合函数的名称。

*data\_type*

此聚合函数所操作的基本数据类型。

*sfunc1*

会为源列的每个非 NULL 字段调用的状态转移函数。它接受一个类型为 *sfunc1\_return\_type* 的变量作为第一个参数，并接受该字段作为第二个参数。

*sfunc1\_return\_type*

第一个转移函数的返回类型。

*sfunc2*

会为源列的每个非 NULL 字段调用的状态转移函数。它接受一个类型为 *sfunc2\_return\_type* 的变量作为唯一参数，并返回同一类型的变量。

*sfunc2\_return\_type*

第二个转移函数的返回类型。

*ffunc*

在遍历完所有输入字段之后调用的最终函数。该函数必须接受两个参数，类型分别为 *sfunc1\_return\_type* 和 *sfunc2\_return\_type*。

*initial\_condition1*

第一个转移函数参数的初始值。

```
initial_condition2
```

第二个转移函数参数的初始值。

## 输出

→ CREATE

命令成功完成时返回的消息。

## 描述

CREATE AGGREGATE 允许用户或程序员通过定义新的聚合函数来扩展 Postgres 的功能。基本发行版中已经提供了一些针对基本类型的聚合函数，例如 `min(int4)` 和 `avg(float8)`。如果定义了新的类型，或者需要一个尚未提供的聚合函数，就可以使用 CREATE AGGREGATE 来提供所需的功能。

一个聚合函数最多可以需要三个函数：两个状态转移函数，*sfunc1* 和 *sfunc2*：

```
sfunc1( internal-state1, next-data_item ) ---> next-internal-state1
sfunc2( internal-state2 ) ---> next-internal-state2
```

以及一个最终计算函数，*ffunc*：

```
ffunc(internal-state1, internal-state2) ---> aggregate-value
```

Postgres 会创建最多两个临时变量（这里称为 *temp1* 和 *temp2*），用来保存用作转移函数参数的中间结果。

这些转移函数必须具有下列性质：

- *sfunc1* 的参数必须是 *temp1*（类型为 *sfunc1\_return\_type*）和 *column value*（类型为 *data\_type*）。返回值必须是 *sfunc1\_return\_type* 类型，并将用作下一次调用 *sfunc1* 时的第一个参数。
- *sfunc2* 的参数和返回值必须是 *temp2*（类型为 *sfunc2\_return\_type*）。
- 最终计算函数的参数必须是 *temp1* 和 *temp2*，而且其返回值必须是 Postgres 的基本类型（不必是为 BASETYPE 指定的 *data\_type*）。
- 当且仅当两个状态转移函数都指定时，才应指定 FINALFUNC。

聚合函数还可以需要一个或两个初始条件，每个转移函数一个。它们在数据库中以 text 类型的字段的形式指定和存储。

## 注意

使用 DROP AGGREGATE 删除聚合函数。

可以指定具有不同状态函数和最终函数组合的聚合函数。例如，count 聚合需要 SFUNC2（一个递增函数）但不需要 SFUNC1 或 FINALFUNC；而 sum 聚合需要 SFUNC1（一个加法函数）但不需要 SFUNC2 或 FINALFUNC；而 avg 聚合则需要上述两个状态函数以及一个 FINALFUNC（一个除法函数）来产生它的结果。无论如何，至少要定义一个状态函数，而且任何 SFUNC2 都必须有一个对应的 INITCOND2。

## 用法

请参阅PostgreSQL 程序员指南中关于聚合函数的章节中关于聚合函数的内容，以获取完整的使用示例。

## 兼容性

### SQL92

`CREATE AGGREGATE` 是 Postgres 的语言扩展。SQL92 中没有`CREATE AGGREGATE`。

# CREATE DATABASE

CREATE DATABASE — 创建一个新数据库

## 大纲

```
CREATE DATABASE name [ WITH LOCATION = 'dbpath' ]
```

## 输入

*name*

要创建的数据库的名称。

*dbpath*

备选位置可以指定为一个后端服务器已知的环境变量（例如 `'PGDATA2'`），或者一个绝对路径名（例如 `'/usr/local/pgsql/data'`）。无论哪种情况，该位置都必须由 `initlocation` 预先配置。

## 输出

→ CREATEDB

命令成功完成时返回的消息。

→ WARN: createdb: database "*name*" already exists.

如果所指定的 *database* 已经存在，就会出现这个错误。

→ ERROR: Unable to create database directory *directory*

创建所需的目录时出了问题；这个操作将需要 `postgres` 用户在指定位置上的权限。

## 描述

CREATE DATABASE 创建一个新的 Postgres 数据库。创建者成为新数据库的管理员。

## 注意

CREATE DATABASE 是一种 Postgres 语言扩展。

使用 DROP DATABASE 删除一个数据库。

## 用法

创建一个新的数据库：

```
olly=>
create database lusiadas;
```

要在备选区域 `~/private_db` 中创建一个新数据库：

```
$
mkdir private_db
$
initlocation ~/private_db
Creating Postgres database system directory /home/olly/
private_db/base
$
psql olly
Welcome to the POSTGRESQL interactive sql monitor:
Please read the file COPYRIGHT for copyright terms of POSTGRESQL

type \? for help on slash commands
type \q to quit
type \g or terminate with semicolon to execute query
You are currently connected to the database: template1

olly=>
create database elsewhere with location = '/home/olly/
private_db';
CREATEDB
```

## 缺陷

使用以绝对路径名指定的备选数据库位置会涉及安全和数据完整性问题，默认情况下，只有后端已知的环境变量才能被指定为备选位置。更多信息见《管理员指南》。

## 兼容性

### SQL92

SQL92 中没有 `CREATE DATABASE` 语句。

标准 SQL 中等价的命令是 `CREATE SCHEMA`。

# CREATE FUNCTION

CREATE FUNCTION — 定义一个新函数

## 大纲

```
CREATE FUNCTION name ( [ ftype [, ...] ] )  
    RETURNS rtype  
    AS path  
    LANGUAGE 'langname'
```

## 输入

*name*

要创建的函数的名称。

*ftype*

函数参数的数据类型。

*rtype*

返回数据类型。

*path*

可以是一个 SQL 查询，也可以是一个对象文件的绝对路径。

*langname*

可以是 'c'、'sql'、'internal' 或 '*plname*'，其中 '*plname*' 是一种已创建过程语言的名  
称。详见CREATE LANGUAGE。

## 输出

→ CREATE

如果命令成功完成，就返回这个结果。

## 描述

CREATE FUNCTION允许一个 Postgres用户向一个数据库注册一个函数。随后，这个用户被视为该函数的拥有者。

## 注意

关于编写函数的更多信息，请参考 PostgreSQL Programmer's Guide 中关于函数的章节。

使用DROP FUNCTION 删除用户定义的函数。

## 用法

创建一个简单的 SQL 函数：

```
CREATE FUNCTION one() RETURNS int4
AS 'SELECT 1 AS RESULT'
LANGUAGE 'sql';
```

```
SELECT one() AS answer;
```

```
answer
-----
1
```

创建一个 C 函数，调用一个用户创建的共享库中的例程。这个特定的例程计算一个校验位，如果函数参数中的校验位正确就返回 TRUE。它用于 CHECK 约束中。

```
CREATE FUNCTION ean_checkdigit(bpchar, bpchar) RETURNS
bool
AS '/usr1/proj/bray/sql/funcs.so' LANGUAGE 'c';

CREATE TABLE product
(
    id                char(8)                PRIMARY KEY,
    eanprefix         char(8)                CHECK (eanprefix ~ '[0-9]{2}-[0-9]{5}'),
    eancode            char(6)                REFERENCES brandname(ean_prefix),
    CONSTRAINT ean CHECK (ean_checkdigit(eanprefix, eancode))
);
```

## 缺陷

C 函数不能返回一组值。

## 兼容性

CREATE FUNCTION 是 Postgres 语言扩展。

## SQL/PSM

### 注意

PSM 表示持久存储模块 (Persistent Stored Modules)。它是一种过程语言，最初希望在 1996 年底之前 PSM 能被批准为官方标准。截至 1998 年中，这还没有发生，但希望 PSM 最终能成为一个标准。

SQL/PSM CREATE FUNCTION 的语法如下：

```
CREATE FUNCTION name
( [ [ IN | OUT | INOUT ] parm type [, ...] ] )
RETURNS rtype
LANGUAGE 'langname'
```

ESPECIFIC *routine*  
*SQL-statement*

# CREATE INDEX

CREATE INDEX — 构建一个二级索引

## 大纲

```
CREATE [ UNIQUE ] INDEX index_name
    ON table [ USING acc_name ]
    ( column [ ops_name] [, ...] )
CREATE [ UNIQUE ] INDEX index_name
    ON table [ USING acc_name ]
    ( func_name( column [, ... ] ) ops_name )
```

## 输入

### UNIQUE

使系统在创建索引时（如果数据已存在）以及每次添加数据时检查表中的重复值。试图插入或更新非重复数据将产生一个错误。

*index\_name*

要创建的索引名称。

*table*

要建立索引的表名。

*acc\_name*

用于索引的访问方法的名称。默认访问方法是 BTREE。Postgres 为二级索引提供了三种访问方法：

#### BTREE

Lehman-Yao 高并发 btrees 的一个实现。

#### RTREE

使用 Guttman 的二次分裂算法实现标准 rtrees。

#### HASH

Litwin 线性散列的一个实现。

*column*

表中一个列的名称。

*ops\_name*

一个关联的操作符类。下面的 select 列出所有 ops\_name：

```
SELECT am.amname AS acc_name,
```

```

        opc.opcname AS ops_name,
        opr.oprname AS ops_comp
FROM pg_am am, pg_amop amop,
     pg_opclass opc, pg_operator opr
WHERE amop.amopid = am.oid AND
      amop.amopclaid = opc.oid AND
      amop.amopopr = opr.oid
ORDER BY acc_name, ops_name, ops_comp

```

*func\_name*

一个用户定义的函数，返回一个可以建立索引的值。

## 输出

→ CREATE

索引成功创建时返回的消息。

→ ERROR: Cannot create index: 'index\_name' already exists.

无法创建索引时出现此错误。

## 描述

CREATE INDEX 在指定的 *table* 上构建一个索引 *index\_name*。

### 提示

索引主要用于提升数据库性能。但使用不当会导致性能下降。

在上面所示的第一种语法中，索引的键字段以列名指定；一个列还可以有一个关联的操作符类。操作符类用于为特定索引指定要使用的操作符。例如，四字节整数上的 `btree` 索引会使用 `int4_ops` 类；此操作符类包含四字节整数的比较函数。默认操作符类是该字段类型的相应操作符类。

在第二种语法中，索引定义在把用户定义的函数 *func\_name* 应用于单个类的一个或多个属性所得到的结果上。这些函数索引可用于基于操作符的快速数据访问，而这些操作符通常需要对基础数据做某种变换才能应用。

## 注解

目前，只有 BTREE 访问方法支持多列索引。最多可指定 7 个键。

使用 DROP INDEX 删除索引。

## 用法

要在表 `films` 的字段 `title` 上创建一个 `btree` 索引：

```

CREATE UNIQUE INDEX title_idx
ON films (title);

```

## 兼容性

### SQL92

CREATE INDEX 是 Postgres 的语言扩展。

SQL92 中没有 CREATE INDEX 命令。

# CREATE LANGUAGE

CREATE LANGUAGE — 为函数定义一种新语言

## 大纲

```
CREATE [ TRUSTED ] PROCEDURAL LANGUAGE 'langname'  
    HANDLER call_handler  
    LANCOMPILER 'comment'
```

## 输入

TRUSTED

TRUSTED指定该语言的调用处理器是安全的；也就是说，它不会为无特权用户提供任何绕过访问限制的功能。如果注册语言时省略这个关键字，则只有拥有 Postgres超级用户权限的用户能用这种语言创建新函数（就像'C'语言那样）。

*langname*

新过程语言的名称。语言名不区分大小写。过程语言不能覆盖 Postgres的内建语言之一。

HANDLER *call\_handler*

*call\_handler*是一个先前已注册函数的名称，该函数将被调用来执行 PL 过程。

*comment*

LANCOMPILER参数是将被插入到新的 `pg_language`项的LANCOMPILER属性中的字符串。目前，Postgres完全不以任何方式使用这个属性。

## 输出

→ CREATE

如果语言成功创建，就返回此消息。

→ ERROR: PL handler function *funcname*() doesn't exist

如果找不到函数 *funcname*()，就返回此错误。

## 描述

使用CREATE LANGUAGE，Postgres用户可以向 Postgres注册一种新语言。随后，就可以用这种新语言定义函数和触发器过程。注册新语言的用户必须拥有 Postgres超级用户权限。

## 编写 PL 处理器

过程语言的调用处理器必须用 'C'之类的编译器语言编写，并作为不带参数、返回 `opaque`类型的函数注册到 Postgres中，这个类型是为未指定或未定义类型准备的占位类型。这可以防止调用处理器被当作函数从查询中直接调用。

不过，当要执行处理器所提供的语言中的某个 PL 函数或触发器过程时，实际调用中必须提供参数。

- 从触发器管理器调用时，唯一的参数是过程的pg\_proc项的对象 ID。来自触发器管理的所有其他信息都可以在全局CurrentTriggerData指针中找到。
- 从函数管理器调用时，参数是过程pg\_proc项的对象 ID、给 PL 函数的参数个数、放在FmgrValues结构中的参数，以及一个指向布尔值的指针，函数通过它告诉调用方返回值是否为 SQL NULL 值。

获取pg\_proc项并分析被调用过程的参数和返回类型，是调用处理器的责任。过程的 CREATE FUNCTION中的 AS 子句可以在pg\_proc表项的prosrc属性中找到。它可以是过程语言自身的源文本（例如 PL/Tcl），可以是一个文件的路径名，也可以是任何能告诉调用处理器具体做什么的东西。

## 注意

使用CREATE FUNCTION创建函数。

使用DROP LANGUAGE删除过程语言。

参阅表pg\_language以获得更多信息：

| Table = pg_language                                                                                                   |      |        |
|-----------------------------------------------------------------------------------------------------------------------|------|--------|
| Field                                                                                                                 | Type | Length |
| lanname                                                                                                               | name | 32     |
| lancompiler                                                                                                           | text | var    |
| <pre> lanname  lancompiler -----+----- internal n/a lisp     /usr/ucb/liszt C        /bin/cc sql      postgres </pre> |      |        |

## 限制

由于过程语言的调用处理器必须以 'C'语言注册到 Postgres中，它继承了 'C'函数的所有能力和限制。

## 缺陷

目前，过程语言的定义一旦创建就不能更改。

## 用法

这是一个用 'C'编写的 PL 处理器模板：

```
#include "executor/spi.h"
```

```

#include "commands/trigger.h"
#include "utils/elog.h"
#include "fmgr.h"          /* for FmgrValues struct */
#include "access/heapam.h"
#include "utils/syscache.h"
#include "catalog/pg_proc.h"
#include "catalog/pg_type.h"

Datum
plsample_call_handler(
    Oid          prooid,
    int          pronargs,
    FmgrValues    *proargs,
    bool          *isNull)
{
    Datum        retval;
    TriggerData   *trigdata;

    if (CurrentTriggerData == NULL) {
        /*
         * Called as a function
         */

        retval = ...
    } else {
        /*
         * Called as a trigger procedure
         */
        trigdata = CurrentTriggerData;
        CurrentTriggerData = NULL;

        retval = ...
    }
    *isNull = false;
    return retval;
}

```

只需在省略号处再加上几千行代码，就能完成这个 PL 调用处理器。关于如何把它编译成可装载模块，见CREATE FUNCTION。

下面的命令随后注册这个示例过程语言：

```

CREATE FUNCTION plsample_call_handler () RETURNS opaque
AS '/usr/local/pgsql/lib/plsample.so'
LANGUAGE 'C';
CREATE PROCEDURAL LANGUAGE 'plsample'
HANDLER plsample_call_handler
LANCOMPILER 'PL/Sample';

```

## 兼容性

CREATE LANGUAGE 是 Postgres 扩展。

## SQL92

SQL92中没有 `CREATE LANGUAGE` 语句。

# CREATE OPERATOR

CREATE OPERATOR — 定义一个新的用户操作符

## 大纲

```
CREATE OPERATOR name (  
    PROCEDURE = func_name  
    [, LEFTARG = type1 ]  
    [, RIGHTARG = type2 ]  
    [, COMMUTATOR = com_op ]  
    [, NEGATOR = neg_op ]  
    [, RESTRICT = res_proc ]  
    [, HASHES ]  
    [, JOIN = join_proc ]  
    [, SORT = sort_op [, ...] ]  
)
```

## 输入

*name*

要定义的操作符。允许的字符见下文。

*func\_name*

用于实现该操作符的函数。

*type1*

操作符左侧的类型（如果有的话）。对于右一元操作符，这个选项要省略。

*type2*

操作符右侧的类型（如果有的话）。对于左一元操作符，这个选项要省略。

*com\_op*

相应的可交换操作符。

*neg\_op*

相应的求反操作符。

*res\_proc*

相应的限制操作符。

HASHES

此操作符可以支持哈希连接算法。

*join\_proc*

支持表连接的过程。

`sort_op`

用于排序的操作符。

## 输出

→ CREATE

操作符创建成功时返回的消息。

## 描述

CREATE OPERATOR 定义一个新的操作符 *name*。定义操作符的用户将成为其所有者。

操作符的 *name* 是一个最多由三十二（32）个字符组成的序列，这些字符可以从下列字符中任意组合：

`+ - * / < > = ~ ! @ # % ^ & | ` ? $ :`

### 注意

操作符名称中不允许使用字母字符。这使 Postgres 能够把 SQL 输入解析成记号而无需在每个记号之间加空格。

操作符 `!="` 在输入时会被映射为 `<>`，因此二者等效。

LEFTARG 和 RIGHTARG 至少必须定义一个。对于二元操作符，两者都应定义。对于右一元操作符，只应定义 LEFTARG；对于左一元操作符，只应定义 RIGHTARG。

此外，过程 *func\_name* 必须此前已用 CREATE FUNCTION 定义，并且必须定义为接受正确数量的参数（一个或两个）。

交换子操作符的存在是为了让 Postgres 可以在需要时反转操作数的顺序。例如，面积小于操作符 `<<<` 会有一个交换子操作符——面积大于 `>>>`。于是，查询优化器就可以自由地把：

```
"0,0,1,1"::box >>> MYBOXES.description
```

转换为

```
MYBOXES.description <<< "0,0,1,1"::box
```

这使执行代码总能使用后一种表示，并在一定程度上简化了查询优化器。

假设存在一个操作符——面积等于 `===`，同时也存在面积不等于 `!=="`。求反器操作符使查询优化器可以把

```
NOT MYBOXES.description === "0,0,1,1"::box
```

转换为

```
MYBOXES.description != "0,0,1,1"::box
```

如果给出了交换子操作符的名称，Postgres 会在目录中搜索它。如果找到了，而它自己还没有交换子，那么该交换子的条目就会被更新，把当前（新）操作符作为其交换子。这对求反器同样适用。

这样做的目的是允许定义两个互为交换子或互为求反器的操作符：第一个操作符应当不带交换子或求反器（视情况而定）先定义；定义第二个操作符时，把第一个指名为交换子或求反器，第一个就会作为副作用被更新。

后两项说明的存在是为了支持查询优化器执行连接。Postgres 总是能通过迭代替换 [WONG76] 求一个连接的值（即处理一个由返回 `boolean` 的操作符分隔两个元组变量的子句）。此外，Postgres 计划实现一种类似 [SHAP86] 的哈希连接算法；但它必须知道这一策略是否适用。例如，哈希连接算法可用于形如：

```
MYBOXES.description === MYBOXES2.description
```

的子句，但不能用于形如：

```
MYBOXES.description <<< MYBOXES2.description.
```

HASHES 标志向查询优化器提供所需的信息，指明哈希连接策略对所讨论的操作符是否可用。

类似地，两个排序操作符向查询优化器指明归并排序是否是可用的连接策略，以及应该用哪些操作符来排序两个操作数类。对于上面的 `===` 子句，优化器必须用操作符 `<<<` 对两个关系都排序。而另一方面，归并排序不能用于子句：

```
MYBOXES.description <<< MYBOXES2.description
```

如果将来发现其他连接策略切实可行，Postgres 会修改优化器和运行时系统来使用它们，并在定义操作符时要求额外的说明。幸运的是，研究界并不经常发明新的连接策略，而用户自定义连接策略带来的额外通用性被认为不值得为之增加复杂度。

说明的最后两项是为了让查询优化器能够估算结果大小。如果限定条件中出现了形如：

```
MYBOXES.description <<< "0,0,1,1"::box
```

的子句，那么 Postgres 可能必须估算 MYBOXES 中满足该子句的实例比例。函数 `res_proc` 必须是一个已注册的函数（即已用 `define function(l)` 定义），它接受一个正确数据类型的参数并返回一个浮点数。查询优化器只需调用该函数，传入参数 `"0,0,1,1"`，再把结果乘以关系的大小，就得到想要的预期实例数。

类似地，当操作符的两个操作数都含有实例变量时，查询优化器必须估算所得连接的大小。函数 `join_proc` 会返回另一个浮点数，把它乘以所涉及两个类的基数，就得到想要的预期结果大小。

函数

```
my_procedure_1 (MYBOXES.description, "0,0,1,1"::box)
```

与操作符

```
MYBOXES.description === "0,0,1,1"::box
```

的区别在于：Postgres 会尝试优化操作符，并可在涉及操作符时决定使用索引来限制搜索空间。但系统不会尝试优化函数，函数是靠蛮力执行的。此外，函数可以有任意数量的参数，而操作符只能有一个或两个。

## 注意

更多信息参见 PostgreSQL 用户指南中关于操作符的章节。参见 `DROP OPERATOR` 以从数据库中删除用户定义的操作符。

## 用法

下面的命令为 BOX 数据类型定义一个新的操作符——面积相等。

```
CREATE OPERATOR == (
    LEFTARG = box,
    RIGHTARG = box,
    PROCEDURE = area_equal_procedure,
    COMMUTATOR = ==,
    NEGATOR = !=,
    RESTRICT = area_restriction_procedure,
    HASHES,
    JOIN = area-join-procedure,
    SORT = <<<, <<<)
```

## 兼容性

`CREATE OPERATOR` 是 Postgres 的扩展。

## SQL92

在 SQL92 中没有 `CREATE OPERATOR` 语句。

# CREATE RULE

CREATE RULE — 定义一条新规则

## 大纲

```
CREATE RULE name
  AS ON event
  TO object [ WHERE condition ]
  DO [ INSTEAD ] [ action | NOTHING ]
```

## 输入

*name*

要创建的规则的名称。

*event*

事件是 `select`、`update`、`delete` 或 `insert` 之一。

*object*

Object 是 `table` 或 `table.column`。

*condition*

任意的 SQL WHERE 子句。只要 SQL 中允许出现实例变量，`new` 或 `current` 就可以出现在实例变量的位置上。

*action*

任意的 SQL 语句。只要 SQL 中允许出现实例变量，`new` 或 `current` 就可以出现在实例变量的位置上。

## 输出

→ CREATE

规则成功创建时返回的消息。

## 描述

一条规则的语义是：当单个实例被访问、更新、插入或删除时，会有一个当前实例（对于检索、更新和删除）和一个新实例（对于更新和追加）。如果 ON 子句中指定的 *event* 和 WHERE 子句中指定的 *condition* 对当前实例为真，该规则的 *action* 部分就会被执行。不过首先，当前实例和/或新实例中字段的值会替换到 `current.attribute-name` 和 `new.attribute-name` 上。

规则的 *action* 部分与触发激活它的用户命令使用相同的命令标识符和事务标识符执行。

## 注解

关于 SQL 规则有一点需要提醒。如果同一个类名或实例变量出现在规则的 *event*、*condition* 和 *action* 部分中，它们都会被视为不同的元组变量。更准确地说，*new* 和 *current* 是这些子句之间仅有的共享元组变量。例如，下面的两条规则具有相同的语义：

```
on update to EMP.salary where EMP.name = "Joe"
do update EMP ( ... ) where ...
```

```
on update to EMP-1.salary where EMP-2.name = "Joe"
do update EMP-3 ( ... ) where ...
```

每条规则都可以带可选的 *INSTEAD* 标记。如果没有该标记，当规则 *condition* 部分中的 *event* 发生时，*action* 会随用户命令一起执行。或者，*action* 部分也可以替代用户命令执行。在后一种情况下，*action* 可以是关键字 *NOTHING*。

在为一个特定的规则应用选择重写规则系统还是实例规则系统时，请记住：在重写系统中，*current* 指的是一个关系和一些限定词，而在实例系统中它指的是一个实例（元组）。

特别需要注意，重写规则系统既不会检测也不会处理循环规则。例如，尽管下面两条规则定义中的每一条都被 Postgres 接受，但 *retrieve* 命令会导致 Postgres 崩溃：

### 例 19.1. 循环重写规则组合的例子。

```
create rule bad_rule_combination_1 is
on select to EMP
do instead select to TOYEMP

create rule bad_rule_combination_2 is
on select to TOYEMP
do instead select to EMP
```

这个从 EMP 中检索的尝试会导致 Postgres 崩溃。

```
select * from EMP
```

要在一个类上定义规则，你必须拥有该类的规则定义权限。使用 *GRANT* 和 *REVOKE* 来更改权限。

## 用法

让 Sam 获得与 Joe 相同的薪水调整：

```
create rule example_1 is
on update EMP.salary where current.name = "Joe"
do update EMP (salary = new.salary)
where EMP.name = "Sam"
```

在 Joe 接受薪水调整时，该事件将变为真，并且 Joe 的当前实例和提议的新实例对执行例程可用。因此，他的新薪水会被代入该规则的 *action* 部分，随后被执行。这样就把 Joe 的薪水传播给了 Sam。

让 Bill 在访问时获得 Joe 的薪水：

```
create rule example_2 is
  on select to EMP.salary
  where current.name = "Bill"
do instead
  select (EMP.salary) from EMP
  where EMP.name = "Joe"
```

拒绝 Joe 访问鞋部门员工的薪水（current\_user 返回当前用户的名称）：

```
create rule example_3 is
  on select to EMP.salary
  where current.dept = "shoe" and current_user = "Joe"
do instead nothing
```

创建一个在玩具部门工作的员工的视图。

```
create TOYEMP(name = char16, salary = int4)

create rule example_4 is
  on select to TOYEMP
do instead
  select (EMP.name, EMP.salary) from EMP
  where EMP.dept = "toy"
```

所有新员工的薪水必须不超过 5,000

```
create rule example_5 is
  on insert to EMP where new.salary > 5000
do update newset salary = 5000
```

## 缺陷

SQL 规则中的对象不能是数组引用，也不能带参数。

除 "oid" 字段外，系统属性不能在规则中的任何地方被引用。这尤其意味着实例的函数（例如 "foo(emp)"，其中 "emp" 是一个类）不能在规则中的任何地方被调用。

规则系统将规则文本和查询计划存储为文本属性。

这意味着如果规则加上其各种内部表示超过大约一页（8KB）的某个值，规则的创建可能会失败。

## 兼容性

CREATE RULE 语句是 Postgres 的语言扩展。

## SQL92

SQL92 中没有 CREATE RULE 语句。

# CREATE SEQUENCE

CREATE SEQUENCE — 创建一个新的序列号发生器

## 大纲

```
CREATE SEQUENCE seqname
    [ INCREMENT increment ]
    [ MINVALUE minvalue ]
    [ MAXVALUE maxvalue ]
    [ START start ]
    [ CACHE cache ]
    [ CYCLE ]
```

## 输入

*seqname*

要创建的序列的名称。

*increment*

INCREMENT *increment* 子句是可选的。正值将生成递增序列，负值生成递减序列。默认值为一（1）。

*minvalue*

可选子句 MINVALUE *minvalue* 决定序列可以生成的最小值。递增和递减序列的默认值分别是 1 和 -2147483647。

*maxvalue*

用可选子句 MAXVALUE *maxvalue* 决定序列的最大值。递增和递减序列的默认值分别是 2147483647 和 -1。

*start*

可选的 START *start* 子句让序列可以从任何地方开始。默认起始值：递增序列为 *minvalue*，递减序列为 *maxvalue*。

*cache*

CACHE *cache* 选项使序列号可以预分配并存储在内存中以便更快访问。最小值为 1（一次只能生成一个值，即不缓存），这也是默认值。

CYCLE

可选的 CYCLE 关键字可以让序列在递增或递减序列分别达到 *maxvalue* 或 *minvalue* 时继续。如果到达了限制，生成的下一个数将是 *minvalue* 或 *maxvalue*，视情况而定。

## 输出

→ CREATE

命令成功时返回的消息。

→ ERROR: amcreate: 'seqname' relation already exists

指定的序列已存在。

→ ERROR: DefineSequence: START value (*start*) can't be > MAXVALUE (*maxvalue*)

指定的起始值超出范围。

→ ERROR: DefineSequence: START value (*start*) can't be < MINVALUE (*minvalue*)

指定的起始值超出范围。

→ ERROR: DefineSequence: MINVALUE (*minvalue*) can't be >= MAXVALUE (*maxvalue*)

如果最小值和最大值不一致。

## 描述

CREATE SEQUENCE will enter a new sequence number generator into the current data base. This involves creating and initialising a new single-row table with the name *seqname*. The generator will be "owned" by the user issuing the command.

序列创建之后，可以用函数 `nextval(seqname)` 从序列获取一个新的数。函数 `currval('seqname')` 可用来确定当前会话中对指定序列最后一次调用 `nextval(seqname)` 所返回的数。函数 `setval('seqname', newvalue)` 可用来设置指定序列的当前值。下一次调用 `nextval(seqname)` 将返回给定的值加上序列增量。

Use a query like

```
SELECT * FROM sequence_name;
```

to get the parameters of a sequence. Aside from fetching the original parameters, you can use

```
SELECT last_value FROM sequence_name;
```

to obtain the last value allocated by any backend. parameters, you can use

使用低级锁来支持对一个发生器的多个同时调用。

### 小心

Unexpected results may be obtained if a cache setting greater than one is used for a sequence object that will be used concurrently by multiple backends. Each backend will allocate "cache" successive sequence values during one access to the sequence object and increase the sequence object's `last_value` accordingly. Then, the next cache-1 uses of `nextval` within that backend simply return the preallocated values without touching the shared object. So, numbers allocated but not used

in the current session will be lost. Furthermore, although multiple backends are guaranteed to allocate distinct sequence values, the values may be generated out of sequence when all the backends are considered. (For example, with a cache setting of 10, backend A might reserve values 1..10 and return nextval=1, then backend B might reserve values 11..20 and return nextval=11 before backend A has generated nextval=2.) Thus, with a cache setting of one it is safe to assume that nextval values are generated sequentially; with a cache setting greater than one you should only assume that the nextval values are all distinct, not that they are generated purely sequentially. Also, last\_value will reflect the latest value reserved by any backend, whether or not it has yet been returned by nextval.

## 注解

要删除序列，请参阅 DROP SEQUENCE 语句。

每个后端用自己的缓存存储已分配的数字。当前会话中已缓存但未使用的数字将会丢失，从而在序列中留下"空洞"。

## 用法

创建一个名为 serial 的递增序列，从 101 开始：

```
CREATE SEQUENCE serial START 101;
```

从该序列中选择下一个数

```
SELECT NEXTVAL ('serial');
```

```
nextval
-----
      114
```

在 INSERT 中使用该序列：

```
INSERT INTO distributors VALUES (NEXTVAL('serial'),'nothing');
```

在 COPY FROM 之后设置序列值：

```
CREATE FUNCTION distributors_id_max() RETURNS INT4
AS 'SELECT max(id) FROM distributors'
LANGUAGE 'sql';
BEGIN;
COPY distributors FROM 'input_file';
SELECT setval('serial', distributors_id_max());
END;
```

## 兼容性

CREATE SEQUENCE 是一种 Postgres 语言扩展。

## SQL92

SQL92 中没有 `CREATE SEQUENCE` 语句。

# CREATE TABLE

CREATE TABLE — Creates a new table

## 大纲

```
CREATE TABLE table (
    column type
    [ DEFAULT value ]
    [, NOT NULL ] [ , UNIQUE ]
    [column_constraint_clause | PRIMARY KEY } [ ... ] ]
    [, ... ]
    [, PRIMARY KEY ( column [, ...] ) ]
    [, CHECK ( condition ) ]
    [, table_constraint_clause ]
) [ INHERITS ( inherited_table [, ...] ) ]
```

## 输入

*table*

要创建的新表的名称。

*column*

列的名称。

*type*

列的类型。这可以包括数组说明符。有关数据类型和数组的更多信息，请参阅 [PostgreSQL 用户指南](#)。

DEFAULT *value*

列的默认值。详见 DEFAULT 子句。

*column\_constraint\_clause*

可选的列约束子句指定一组完整性约束（测试），要使插入或更新操作成功，新条目或更新后的条目必须满足这些约束。每个约束都必须能求值为布尔表达式。尽管 SQL92 要求 *column\_constraint\_clause* 只引用该列，Postgres 却允许在单个列约束中引用多个列。详见列约束子句。

*table\_constraint\_clause*

可选的表 CONSTRAINT 子句指定一组完整性约束，要使插入或更新操作成功，新条目或更新后的条目必须满足这些约束。每个约束都必须能求值为布尔表达式。单个约束中可以引用多个列。一个表只能指定一个 PRIMARY KEY 子句；PRIMARY KEY *column*（表约束）与 PRIMARY KEY（列约束）互斥。详见表约束子句。

INHERITS *inherited\_table*

可选的 INHERITS 子句指定一组表名，该表自动从中继承所有字段。如果某个被继承的字段名出现多于一次，Postgres 会报错。Postgres 会自动让新表继承其在继承层次中上方表上的函数。

## 题外话

函数的继承按照 Common Lisp Object System (CLOS) 的惯例进行。

## 输出

→ CREATE

表创建成功时返回的消息。

→ ERROR

表创建失败时返回的消息。通常还伴随一些描述性文字，例如：

```
amcreate: "table" relation already exists
```

，如果指定的表在数据库中已存在，就会在运行时出现该错误。

→ ERROR: DEFAULT: type mismatched

如果默认值的数据类型与列定义的数据类型不匹配。

## 描述

CREATE TABLE 将向当前数据库中输入一个新表。该表将由发出命令的用户“拥有”。

新表作为堆创建，不含初始数据。一个表最多可以有 1600 列（现实中，由于元组大小必须小于 8192 字节，这个上限还会更低），某些站点还可能把该上限配置得更低。表不能与系统目录表同名。

## DEFAULT 子句

DEFAULT *value*

## 输入

*value*

默认值表达式可能的取值有：

- 字面值
- 用户函数
- niladic 函数（无参函数）

## 输出

## 描述

DEFAULT 子句为一个列指定默认数据值（通过 CREATE TABLE 语句中的列定义）。默认值的数据类型必须与列定义的数据类型匹配。

如果一个 INSERT 操作涉及的列没有指定默认值，且没有为它提供显式数据值，那么将把 NULL 值赋给该列。Default *literal* 表示默认值就是指定的常量值。Default *niladic-function* 或 *user-function* 表示默认值是 INSERT 时该指定函数的值。

niladic 函数有两种类型：

niladic USER

CURRENT\_USER / USER

见 CURRENT\_USER 函数

SESSION\_USER

尚不支持

SYSTEM\_USER

尚不支持

niladic datetime

CURRENT\_DATE

见 CURRENT\_DATE 函数

CURRENT\_TIME

见 CURRENT\_TIME 函数

CURRENT\_TIMESTAMP

见 CURRENT\_TIMESTAMP 函数

在当前版本（v6.4）中，Postgres 会在定义表的时候求值所有默认表达式。因此，诸如 CURRENT\_TIMESTAMP 这类“不可缓存”的函数可能不会产生预期的效果。对于日期/时间类型这一特殊情况，可以用“DEFAULT TEXT 'now'”代替“DEFAULT 'now'”或“DEFAULT CURRENT\_TIMESTAMP”。这会迫使 Postgres 把该常量当作字符串类型，然后在运行时再把该值转换为 timestamp。

## 用法

为列 did 和 number 指定常量作为默认值，并为列 did 指定一个字符串字面值：

```
CREATE TABLE video_sales (
    did      VARCHAR(40) DEFAULT 'luso films',
    number   INTEGER DEFAULT 0,
    total    CASH DEFAULT '$0.0'
);
```

为列 did 指定一个已存在的序列作为默认值，并为列 name 指定一个字面值：

```
CREATE TABLE distributors (
    did      DECIMAL(3)  DEFAULT NEXTVAL('serial'),
    name     VARCHAR(40) DEFAULT 'luso films'
);
```

## 列 CONSTRAINT 子句

```
[ CONSTRAINT name ] { NOT NULL | UNIQUE | PRIMARY KEY |  
CHECK constraint } [, ...]
```

### 输入

*name*

赋予完整性约束的一个任意名称。如果未指定 *name*，它将由表名和列名生成，这应当能保证 *name* 的唯一性。

NOT NULL

该列不允许包含 NULL 值。这等价于列约束 CHECK (*column* NOT NULL)。

UNIQUE

该列必须具有唯一值。在 Postgres 中，这是通过在表上隐式创建一个唯一索引来强制实现的。

PRIMARY KEY

该列是一个主键，这蕴含着唯一性由系统强制，并且其他表可以把该列当作行的唯一标识符来依赖。详见 PRIMARY KEY。

*constraint*

约束的定义。

### 描述

约束是一条命名的规则：一个通过对在基表上执行的 INSERT、UPDATE 或 DELETE 操作的结果加以限制，来帮助定义有效值集合的 SQL 对象。

定义完整性约束有两种方式：稍后讲述的表约束，以及这里讲述的列约束。

列约束是作为列定义的一部分定义的完整性约束，一旦创建，在逻辑上就成为表约束。可用的列约束有：

PRIMARY KEY  
REFERENCES  
UNIQUE  
CHECK  
NOT NULL

### 注意

Postgres 在 6.4 版中尚不支持 REFERENCES 完整性约束。解析器接受 REFERENCES 语法，但忽略该子句。

## NOT NULL 约束

```
[ CONSTRAINT name ] NOT NULL
```

NOT NULL 约束指定一条规则：一列只能包含非空值。

NOT NULL 约束只是列约束，不允许作为表约束。

### 输出

```
status
```

```
→ ERROR: ExecAppend: Fail to add null value in not null attribute
"column".
```

如果试图向一个带 NOT NULL 约束的列插入空值，就会在运行时出现该错误。

### 描述

### 用法

在表 `distributors` 上定义两个 NOT NULL 列约束，其中一个是命名约束：

```
CREATE TABLE distributors (  
    did          DECIMAL(3) CONSTRAINT no_null NOT NULL,  
    name         VARCHAR(40) NOT NULL  
);
```

## UNIQUE 约束

```
[ CONSTRAINT name ] UNIQUE
```

### 输入

```
CONSTRAINT name
```

赋予约束的一个任意标签。

### 输出

```
status
```

```
→ ERROR: Cannot insert a duplicate key into a unique index.
```

如果试图向一个列插入重复值，就会在运行时出现该错误。

### 描述

UNIQUE 约束指定一条规则：表中一个或多个不同列组成的一组只能包含唯一值。

被指定列的列定义不必包含 NOT NULL 约束就可以纳入 UNIQUE 约束。一个不带 NOT NULL 约束的列中有多个空值并不违反 UNIQUE 约束。（这偏离了 SQL92 的定义，但是一个更合乎常理的约定。详见兼容性一节。）

每个 UNIQUE 列约束所命名的列，必须与为该表定义的任何其他 UNIQUE 或 PRIMARY KEY 约束所命名的列集合不同。

### 注意

Postgres 会自动为每个 UNIQUE 约束创建一个唯一索引，以确保数据完整性。详见 CREATE INDEX。

## 用法

为表 distributors 定义一个 UNIQUE 列约束。UNIQUE 列约束只能定义在表的一列上：

```
CREATE TABLE distributors (
    did      DECIMAL(3),
    name     VARCHAR(40) UNIQUE
);
```

它等价于把以下内容指定为表约束：

```
CREATE TABLE distributors (
    did      DECIMAL(3),
    name     VARCHAR(40),
    UNIQUE (name)
);
```

## CHECK 约束

```
[ CONSTRAINT name ] CHECK ( condition [, ...] )
```

### 输入

→ *name*

赋予约束的一个任意名称。

*condition*

任何产生布尔结果的合法条件表达式。

### 输出

*status*

→ ERROR: ExecAppend: rejected due to CHECK constraint "*table\_column*".

如果试图向一个受 CHECK 约束约束的列插入非法值，就会在运行时出现该错误。

## 描述

CHECK 约束对列中允许的值指定一个限制。CHECK 约束也允许作为表约束。

SQL92 的 CHECK 列约束只能定义在表的一列上并只引用该列。Postgres 没有这一限制。

## PRIMARY KEY 约束

```
[ CONSTRAINT name ] PRIMARY KEY
```

## 输入

CONSTRAINT → *name*

约束的一个任意名称。

## 输出

→ ERROR: Cannot insert a duplicate key into a unique index.

如果试图向一个受 PRIMARY KEY 约束约束的列插入重复值，就会在运行时出现该错误。

## 描述

PRIMARY KEY 列约束指定一个表的某一列只能包含唯一（非重复）的非 NULL 值。被指定列的定义不必包含显式的 NOT NULL 约束就可以纳入 PRIMARY KEY 约束。

一个表只能指定一个 PRIMARY KEY。

## 注解

Postgres 会自动创建一个唯一索引来确保数据完整性。（见 CREATE INDEX 语句）

PRIMARY KEY 约束所命名的列集合应当与为同一表定义的任何 UNIQUE 约束所命名的列集合不同，否则会导致等价索引的重复和不必要的额外运行时开销。不过，Postgres 并不特别禁止这样做。

## 表 CONSTRAINT 子句

```
[ CONSTRAINT name ] { PRIMARY KEY | UNIQUE } ( column [, ...] )  
[ CONSTRAINT name ] CHECK ( constraint )
```

## 输入

CONSTRAINT *name*

赋予完整性约束的一个任意名称。

*column* [, ...]

要为其定义唯一索引的列名，对于 PRIMARY KEY 还要加 NOT NULL 约束。

`CHECK ( constraint )`

一个将被作为约束求值的布尔表达式。

## 输出

表约束子句可能的输出与列约束子句相应部分的输出相同。

## 描述

表约束是定义在一个基表的一列或多列上的完整性约束。“表约束”的四种变体是：

UNIQUE  
CHECK  
PRIMARY KEY  
FOREIGN KEY

### 注意

Postgres 在 6.4 版中尚不支持 FOREIGN KEY 完整性约束。解析器理解 FOREIGN KEY 语法，但只打印一条提示，此外便忽略该子句。外键可以用触发器部分模拟（见 CREATE TRIGGER 语句）。

## UNIQUE 约束

`[ CONSTRAINT name ] UNIQUE ( column [, ...] )`

## 输入

`CONSTRAINT name`

赋予完整性约束的一个任意名称。

*column*

表中某一列的名称。

## 输出

*status*

ERROR: Cannot insert a duplicate key into a unique index.

如果试图向一个受 PRIMARY KEY 约束约束的列插入重复值，就会在运行时出现该错误。

## 描述

UNIQUE 约束指定一条规则：表中一个或多个不同列组成的一组只能包含唯一值。UNIQUE 表约束的行为与列约束相同，只是它还能跨越多列。

更多细节见 UNIQUE 列约束一节。

## 用法

为表 distributors 定义一个 UNIQUE 表约束：

```
CREATE TABLE distributors (  
    did          DECIMAL(03),  
    name         VARCHAR(40),  
    UNIQUE(name)  
);
```

## PRIMARY KEY 约束

```
[ CONSTRAINT name ] PRIMARY KEY ( column [, ...] )
```

### 输入

CONSTRAINT *name*

约束的一个任意名称。

*column* [, ...]

要为其定义唯一索引的列名，对于 PRIMARY KEY 还要加 NOT NULL 约束。

### 输出

*status*

ERROR: Cannot insert a duplicate key into a unique index.

如果试图向一个受 PRIMARY KEY 约束约束的列插入重复值，就会在运行时出现该错误。

### 描述

PRIMARY KEY 约束指定一条规则：表中一个或多个不同列组成的一组只能包含唯一（非重复）的非空值。被指定列的列定义不必包含 NOT NULL 约束就可以纳入 PRIMARY KEY 约束。PRIMARY KEY 表约束与相应的列约束类似，只是它还能涵盖多列。

更多信息见 PRIMARY KEY 列约束一节。

## 用法

创建表 films 和表 distributors

```
CREATE TABLE films (  
    code          CHARACTER(5) CONSTRAINT firstkey PRIMARY KEY,  
    title         CHARACTER VARYING(40) NOT NULL,  
    did           DECIMAL(3) NOT NULL,  
    date_prod     DATE,  
    kind          CHAR(10),  
    len           INTERVAL HOUR TO MINUTE  
);
```

```
CREATE TABLE distributors (  
    did          DECIMAL(03) PRIMARY KEY DEFAULT NEXTVAL('serial'),  
    name         VARCHAR(40) NOT NULL CHECK (name <> '')  
);
```

创建一个带二维数组的表

```
CREATE TABLE array (  
    vector INT[][]  
);
```

为表 films 定义一个 UNIQUE 表约束。UNIQUE 表约束可以定义在表的一列或多列上：

```
CREATE TABLE films (  
    code         CHAR(5),  
    title        VARCHAR(40),  
    did          DECIMAL(03),  
    date_prod    DATE,  
    kind         CHAR(10),  
    len          INTERVAL HOUR TO MINUTE,  
    CONSTRAINT production UNIQUE(date_prod)  
);
```

定义一个 CHECK 列约束。

```
CREATE TABLE distributors (  
    did          DECIMAL(3) CHECK (did > 100),  
    name         VARCHAR(40)  
);
```

定义一个 CHECK 表约束：

```
CREATE TABLE distributors (  
    did          DECIMAL(3),  
    name         VARCHAR(40)  
    CONSTRAINT con1 CHECK (did > 100 AND name > '')  
);
```

为表 films 定义一个 PRIMARY KEY 表约束。PRIMARY KEY 表约束可以定义在表的一列或多列上：

```
CREATE TABLE films (  
    code         CHAR(05),  
    title        VARCHAR(40),  
    did          DECIMAL(03),  
    date_prod    DATE,  
    kind         CHAR(10),  
    len          INTERVAL HOUR TO MINUTE,  
    CONSTRAINT code_title PRIMARY KEY(code,title)  
);
```

为表 distributors 定义一个 PRIMARY KEY 列约束。PRIMARY KEY 列约束只能定义在表的一列上（下面两个示例是等价的）

```
CREATE TABLE distributors (
    did      DECIMAL(03),
    name     CHAR VARYING(40),
    PRIMARY KEY(did)
);

CREATE TABLE distributors (
    did      DECIMAL(03) PRIMARY KEY,
    name     VARCHAR(40)
);
```

## 注解

CREATE TABLE/INHERITS 是 Postgres 的语言扩展。

## 兼容性

## SQL92

除普通的 CREATE TABLE 外，SQL92 还定义了 CREATE TEMPORARY TABLE 语句：

```
CREATE [ {GLOBAL | LOCAL} ] TEMPORARY TABLE table (
    column type [DEFAULT value] [CONSTRAINT column_constraint] [,
... ] )
    [CONSTRAINT table_constraint ]
    [ ON COMMIT {DELETE | PRESERVE} ROWS ]
```

对于临时表，CREATE TEMPORARY TABLE 语句命名一个新表，并定义该表的列和约束。

CREATE TEMPORARY TABLE 的可选 ON COMMIT 子句指定每当执行 COMMIT 时是否清空临时表中的行。如果省略 ON COMMIT 子句，则假定使用默认选项 ON COMMIT DELETE ROWS。

创建一个临时表：

```
CREATE TEMPORARY TABLE actors (
    id      DECIMAL(03),
    name    VARCHAR(40),
    CONSTRAINT actor_id CHECK (id < 150)
) ON COMMIT DELETE ROWS
```

临时表目前在 Postgres 中尚不可用。

### 提示

在 Postgres 的当前版本（v6.4）中，要使用临时表，必须用显式命令创建和删除表。

## UNIQUE 子句

SQL92 为 UNIQUE 指定了一些额外的能力：

表约束定义

```
[ CONSTRAINT name ]  
UNIQUE ( column [, ...] )  
[ {INITIALLY DEFERRED | INITIALLY IMMEDIATE} ]  
[ [ NOT ] DEFERRABLE ]
```

列约束定义

```
[ CONSTRAINT name ]  
UNIQUE  
[ {INITIALLY DEFERRED | INITIALLY IMMEDIATE} ]  
[ [ NOT ] DEFERRABLE ]
```

## NOT NULL 子句

SQL92 为 NOT NULL 指定了一些额外的能力：

```
[ CONSTRAINT name ] NOT NULL  
[ {INITIALLY DEFERRED | INITIALLY IMMEDIATE} ]  
[ [ NOT ] DEFERRABLE ]
```

## CONSTRAINT 子句

SQL92 为约束指定了一些额外的能力，并且还定义了断言和域约束。

### 注意

Postgres 尚不支持域和断言。

断言是一种特殊类型的完整性约束，与其他约束共享同一名字空间。不过，断言不一定像约束那样依赖于某个特定的基表，因此 SQL-92 提供了 CREATE ASSERTION 语句作为定义约束的另一种方法：

```
CREATE ASSERTION name CHECK ( condition )
```

域约束由 CREATE DOMAIN 或 ALTER DOMAIN 语句定义：

域约束：

```
[ CONSTRAINT name ]  
CHECK constraint  
[ {INITIALLY DEFERRED | INITIALLY IMMEDIATE} ]  
[ [ NOT ] DEFERRABLE ]
```

表约束定义:

```
[ CONSTRAINT name ]
{ PRIMARY KEY constraint |
  FOREIGN KEY constraint |
  UNIQUE constraint |
  CHECK constraint }
[ { INITIALLY DEFERRED | INITIALLY IMMEDIATE } ]
[ [ NOT ] DEFERRABLE ]
```

列约束定义:

```
[ CONSTRAINT name ]
{ NOT NULL constraint |
  PRIMARY KEY constraint |
  FOREIGN KEY constraint |
  UNIQUE constraint |
  CHECK constraint }
[ { INITIALLY DEFERRED | INITIALLY IMMEDIATE } ]
[ [ NOT ] DEFERRABLE ]
```

一个 **CONSTRAINT** 定义可以按任意顺序包含一个可推迟性属性子句和/或一个初始约束模式子句。

**NOT DEFERRABLE**

表示该约束必须在每条 SQL 语句执行之后就检查是否违反其规则。

**DEFERRABLE**

这控制约束是否可以推迟到事务结束检查。如果使用了 **SET CONSTRAINTS ALL DEFERRED**，或者约束被设置为 **INITIALLY DEFERRED**，这将使外键只在事务结束时才被检查。

每个约束的约束模式总是有一个初始默认值，它在事务开始时为该约束设置。

**INITIALLY IMMEDIATE**

表示从事务开始起，该约束必须在每条 SQL 语句执行之后就检查是否违反其规则。

**INITIALLY DEFERRED**

每条语句之后检查约束。

## CHECK 子句

SQL92 为表约束或列约束中的 **CHECK** 指定了一些额外的能力。

表约束定义:

```
[ CONSTRAINT name ]
CHECK ( VALUE condition )
[ { INITIALLY DEFERRED | INITIALLY IMMEDIATE } ]
[ [ NOT ] DEFERRABLE ]
```

列约束定义:

```
[ CONSTRAINT name ]  
CHECK ( VALUE condition )  
[ {INITIALLY DEFERRED | INITIALLY IMMEDIATE} ]  
[ [ NOT ] DEFERRABLE ]
```

## PRIMARY KEY 子句

SQL92 为 PRIMARY KEY 指定了一些额外的能力：

表约束定义：

```
[ CONSTRAINT name ]  
PRIMARY KEY ( column [, ...] )  
[ {INITIALLY DEFERRED | INITIALLY IMMEDIATE} ]  
[ [ NOT ] DEFERRABLE ]
```

列约束定义：

```
[ CONSTRAINT name ]  
PRIMARY KEY  
[ {INITIALLY DEFERRED | INITIALLY IMMEDIATE} ]  
[ [ NOT ] DEFERRABLE ]
```

# CREATE TABLE AS

CREATE TABLE AS — Creates a new table

## 大纲

```
CREATE TABLE table [ ( column [, ...] ) ] AS select_clause
```

## 输入

*table*

要创建的新表的名称。

*column*

列的名称。可以用逗号分隔的列名列表指定多个列名。

*select\_clause*

一条合法的查询语句。允许的语法见 SELECT。

## 输出

有关可能的输出消息的汇总，请参阅 CREATE TABLE 和 SELECT。

## 描述

CREATE TABLE AS 使得可以从一个现有表的内容创建一个新表。它的功能等价于 SELECT TABLE INTO，只是语法可能更为直观。

# CREATE TRIGGER

CREATE TRIGGER — 创建一个新的触发器

## 大纲

```
CREATE TRIGGER name { BEFORE | AFTER }  
    { event [OR ...] }  
    ON table FOR EACH { ROW | STATEMENT }  
    EXECUTE PROCEDURE funcname ( arguments )
```

## Inputs

→ *name*

现有触发器的名称。

→ *table*

表的名称。

→ *event*

INSERT、DELETE 或 UPDATE 之一。

→ *funcname*

一个用户提供的函数。

## Outputs

→ CREATE

触发器创建成功时返回此消息。

## 描述

CREATE TRIGGER 将把一个新触发器登记到当前数据库中。该触发器将与关系 *relname* 关联，并执行指定的函数 *funcname*。

可以指定触发器在操作试图作用于元组之前（即约束检查和 INSERT、UPDATE 或 DELETE 尝试之前）引发，或在操作尝试之后（例如约束检查完毕且 INSERT、UPDATE 或 DELETE 已完成之后）引发。如果触发器在事件之前引发，它可以跳过对当前元组的操作，或者修改被插入的元组（仅适用于 INSERT 和 UPDATE 操作）。如果触发器在事件之后引发，那么所有更改（包括最后一次插入、更新或删除）对触发器都是"可见"的。

更多信息参见 PostgreSQL 程序员指南中关于 SPI 和触发器的章节。

## 注意

CREATE TRIGGER 是一种 Postgres 语言扩展。

只有关系的所有者才能在该关系上创建触发器。

截至当前版本 (v6.4) , STATEMENT 触发器尚未实现。

关于如何移除触发器, 参见 DROP TRIGGER。

## 用法

在向表 films 追加或更新行之前, 检查指定的发行商代码是否存在于 distributors 表中:

```
CREATE TRIGGER if_dist_exists
    BEFORE INSERT OR UPDATE ON films FOR EACH ROW
    EXECUTE PROCEDURE check_primary_key ('did', 'distributors', 'did')
;
```

在取消一个发行商或更新其代码之前, 移除对表 films 的所有引用:

```
CREATE TRIGGER if_film_exists
    BEFORE DELETE OR UPDATE ON distributors FOR EACH ROW
    EXECUTE PROCEDURE check_foreign_key (1, 'CASCADE', 'did', 'films',
    'did');
```

## 兼容性

### SQL92

SQL92 中没有 CREATE TRIGGER。

上面的第二个例子也可以使用外键 (FOREIGN KEY) 约束来完成, 如下所示:

```
CREATE TABLE distributors (
    did          DECIMAL(3),
    name         VARCHAR(40),
    CONSTRAINT if_film_exists
        FOREIGN KEY(did) REFERENCES films
            ON UPDATE CASCADE ON DELETE CASCADE
);
```

不过, 截至 6.4 版本, Postgres 尚未实现外键。

# CREATE TYPE

CREATE TYPE — 定义一种新的基本数据类型

## 大纲

```
CREATE TYPE typename (
    INPUT          = input_function
    , OUTPUT       = output_function
    , INTERNALLENGTH = (internallength | VARIABLE)
    [ , EXTERNALLENGTH = (externallength | VARIABLE) ]
    [ , ELEMENT      = element ]
    [ , DELIMITER     = delimiter ]
    [ , DEFAULT       = "default" ]
    [ , SEND          = send_function ]
    [ , RECEIVE       = receive_function ]
    [ , PASSEDBYVALUE ]
)
```

## 输入

*typename*

要创建的类型的名称。

INTERNALLENGTH *internallength*

一个字面值，指定新类型的内部长度。

EXTERNALLENGTH *externallength*

一个字面值，指定新类型的外部长度。

INPUT *input\_function*

一个由 CREATE FUNCTION 创建的函数的名称，它把数据从外部形式转换为该类型的内部形式。

OUTPUT *output\_function*

一个由 CREATE FUNCTION 创建的函数的名称，它把数据从内部形式转换为适合显示的形式。

*element*

正在创建的类型是数组；该参数指定数组元素类型。

*delimiter*

数组的分隔符字符。

*default*

为表示"数据不存在"而显示的默认文本。

*send\_function*

一个由 CREATE FUNCTION 创建的函数的名称，它把该类型的数据转换为适合传输到另一台机器的形式。

*receive\_function*

一个由 CREATE FUNCTION 创建的函数的名称，它把该类型的数据从适合从另一台机器传输的形式转换为内部形式。

## 输出

→ CREATE

类型创建成功时返回的消息。

## 描述

CREATE TYPE 允许用户向 Postgres 注册一种新的用户数据类型，供当前数据库使用。定义类型的用户将成为其所有者。*Typename* 是新类型的名称，在该数据库已定义的类型中必须唯一。

CREATE TYPE 要求在定义类型之前（用 create function）先注册两个函数。新基本类型的表示由 *input\_function* 决定，它把该类型的外部表示转换为为该类型定义的操作符和函数可用的内部表示。自然，*output\_function* 执行反向的转换。输入和输出函数都必须声明为接受一个或两个 "opaque" 类型的参数。

新的基本数据类型可以是定长的，这时 *internallength* 是一个正整数；也可以是变长的，这时 Postgres 假定新类型与 Postgres 内置的数据类型 "text" 具有相同的格式。要指明一个类型是变长的，把 *internallength* 设为 VARIABLE。外部表示用 *externallength* 关键字以类似的方式指定。

要指明一个类型是数组并指明它有数组元素，可以用 *element* 关键字指定数组元素的类型。例如，要定义一个 4 字节整数 ("int4") 的数组，指定

```
ELEMENT = int4
```

要指明该类型数组上使用的分隔符，可以把 *delimiter* 设为某个特定的字符。默认分隔符是逗号 (",")。

如果用户希望某个特定的位模式表示"数据不存在"，可以选择性地提供一个默认值。用 DEFAULT 关键字指定默认值。

可选函数 *send\_function* 和 *receive\_function* 用于请求 Postgres 服务的应用程序位于另一台机器上的情形。在这种情况下，Postgres 所在机器为该数据类型使用的格式可能与远程机器使用的格式不同。此时，服务器向客户端发送数据时把数据项转换为标准形式，而在服务器从客户端接收数据时从标准格式转换为机器特定的格式，是恰当的做法。如果不指定这两个函数，就假定该类型的内部格式在所有相关机器体系结构上都是可接受的。例如，单个字符从 Sun-4 传到 DECstation 不需要转换，但许多其他类型需要。

可选标志 PASSEDBYVALUE 表示使用该数据类型的操作符和函数应按值而不是按引用传递参数。注意，内部表示超过 4 字节的类型不能按值传递。

对于新的基本类型，用户可以使用本节描述的相应设施定义操作符、函数和聚合。

## 数组类型

存在两个广义的内置函数 `array_in` 和 `array_out`，用于快速创建变长数组类型。这些函数可用于任何现有 Postgres 类型的数组。

## 大对象类型

"常规"的 Postgres 类型长度只能达到 8192 字节。如果你需要更大的类型，就必须创建大对象类型。这些类型的接口在第 7 节"大对象接口"中有详细讨论。所有大对象类型的长度永远是 VARIABLE。

## 示例

这条命令创建 `box` 数据类型，然后在类定义中使用该类型：

```
CREATE TYPE box (INTERNALLENGTH = 8,
                INPUT = my_procedure_1, OUTPUT = my_procedure_2)

CREATE TABLE myboxes (id INT4, description box)
```

这条命令创建一个带整数元素的变长数组类型。

```
CREATE TYPE int4array
    (INPUT = array_in, OUTPUT = array_out,
     INTERNALLENGTH = VARIABLE, ELEMENT = int4)

CREATE TABLE myarrays (id int4, numbers int4array)
```

这条命令创建一个对象类型并在类定义中使用它。

```
CREATE TYPE bigobj
    (INPUT = lo_filein, OUTPUT = lo_fileout,
     INTERNALLENGTH = VARIABLE)

CREATE TABLE big_objs (id int4, obj bigobj)
```

## 限制

类型名不能以下划线字符 ("\_") 开头，且长度只能是 15 个字符。这是因为 Postgres 会为每个基本类型悄悄创建一个数组类型，其名称由基本类型名前加下划线组成。

## 注意

参见 `DROP TYPE` 以删除现有类型。

另请参阅 `CREATE FUNCTION`、`CREATE OPERATOR` 以及 PostgreSQL 程序员指南中关于大对象的章节。

## 兼容性

## SQL3

CREATE TYPE 是一个 SQL3 语句。

# CREATE USER

CREATE USER — 为一个新用户创建账号信息

## 大纲

```
CREATE USER username
    [ WITH PASSWORD password ]
    [ CREATEDB      | NOCREATEDB ]
    [ CREATEUSER   | NOCREATEUSER ]
    [ IN GROUP     groupname [, ...] ]
    [ VALID UNTIL  'abstime' ]
```

## 输入

→ *username*

用户的名称。

→ *password*

WITH PASSWORD 子句在 "pg\_shadow" 表中设置用户的口令。由于这个原因, "pg\_shadow" 不再可被该 Postgres 实例访问, 而 Postgres 用户的口令被初始设置为 NULL。当 "pg\_shadow" 表中某个用户的口令为 NULL 时, 用户认证按历史上的方式进行 (HBA、PG\_PASSWORD 等)。但是, 如果为某个用户设置了口令, 一个新的认证系统会取代为该 Postgres 实例配置的任何其他认证系统, 而存储在 "pg\_shadow" 表中的口令将被用于认证。关于这个认证系统如何工作的更多细节见 pg\_crypt(3)。如果省略 WITH PASSWORD 子句, 用户的口令将被设置为空字符串, 它等同于上面提到的认证系统中的 NULL 值。

→ CREATEDB/NOCREATEDB

这些子句定义用户创建数据库的能力。如果指定了 CREATEDB, 将被定义的用户将被允许创建他自己的数据库。使用 NOCREATEDB 将拒绝用户创建数据库的能力。如果这个子句被省略, 默认使用 NOCREATEDB。

→ CREATEUSER/NOCREATEUSER

这些子句决定是否允许用户在一个 Postgres 实例中创建新用户。省略这个子句将把该用户的这个属性值设置为 NOCREATEUSER。

→ *groupname*

把该用户作为新成员插入其中的组的名称。

→ *abstime*

VALID UNTIL 子句设置一个绝对时间, 超过该时间后用户的 Postgres 登录不再有效。请注意, 如果某个用户在 "pg\_shadow" 表中没有定义口令, 在用户认证期间不会检查有效截止日期。如果省略这个子句, 该属性在 "pg\_shadow" 中存储为 NULL 值, 登录将永久有效。

## 输出

→ CREATE USER

命令成功完成时返回的消息。

→ ERROR: removeUser: user "username" does not exist

如果没有找到 "username"。

## 描述

CREATE USER 将向一个 Postgres 实例添加一个新用户。

新用户将被赋予这样的 `usesysid`: `'SELECT MAX(usesysid) + 1 FROM pg_shadow'`。这意味着 Postgres 用户的 `usesysid` 不会与他们的操作系统 (OS) 用户 ID 对应。这条规则的例外是 'postgres' 用户, 在 `initdb` 过程中它的 OS 用户 ID 被用作 `usesysid`。如果你仍想让任意给定用户的 OS 用户 ID 与 `usesysid` 匹配, 可以使用 Postgres 发行版附带的 "createuser" 脚本。

## 注意

CREATE USER 语句是一种 Postgres 语言扩展。

使用 DROP USER 或 ALTER USER 语句删除或修改一个用户账号。

进一步的信息请参阅 `pg_shadow` 表。

| Table = pg_shadow |         |        |
|-------------------|---------|--------|
| Field             | Type    | Length |
| username          | name    | 32     |
| usesysid          | int4    | 4      |
| usecreatedb       | bool    | 1      |
| usetrace          | bool    | 1      |
| usesuper          | bool    | 1      |
| usecatupd         | bool    | 1      |
| passwd            | text    | var    |
| valuntil          | abstime | 4      |

## 用法

创建一个没有口令的用户:

```
CREATE USER jonathan
```

创建一个带口令的用户:

```
CREATE USER davide WITH PASSWORD jw8s0F4
```

创建一个带口令的用户, 其账号有效到 2001 年底。注意 2002 年到来一秒之后, 该账号不再有效:

```
CREATE USER miriam WITH PASSWORD jw8s0F4 VALID UNTIL 'Jan 1 2002'
```

创建一个用户可以创建数据库的账号：

```
CREATE USER manuel WITH PASSWORD jw8s0F4 CREATEDB
```

## 兼容性

### SQL92

SQL92 中没有 CREATE USER 语句。

# CREATE VIEW

CREATE VIEW — 构造一个虚拟表

## 大纲

```
CREATE VIEW view
  AS SELECT query
```

## 输入

→ *view*

要创建的视图的名称。

→ *query*

一条 SQL 查询，它将提供视图的列和行。

关于有效参数的更多信息，参见 SELECT 语句。

## 输出

→ CREATE

视图创建成功时返回的消息。

→ WARN amcreate: "view" relation already exists

如果指定的视图已存在于数据库中，就会出现这个错误。

→ NOTICE create: attribute named "column" has an unknown type

如果你不指定类型，创建出的视图就会带有一个未知类型的列。例如，下面的命令会产生一个错误：

```
CREATE VIEW vista AS SELECT 'Hello World'
```

而这条命令则不会：

```
CREATE VIEW vista AS SELECT 'Hello World'::text
```

## 描述

CREATE VIEW 将定义一个表的一个视图。该视图不会被物理物化。具体来说，会自动生成一条查询重写检索规则来支持对视图的检索操作。

## 注意

使用 DROP VIEW 语句删除视图。

## 缺陷

目前视图是只读的。

## 用法

创建一个由所有喜剧（Comedy）影片组成的视图：

```
CREATE VIEW kinds AS
  SELECT *
  FROM films
  WHERE kind = 'Comedy';
```

```
SELECT * FROM kinds;
```

| code  | title                     | did | date_prod  | kind   | len   |
|-------|---------------------------|-----|------------|--------|-------|
| UA502 | Bananas                   | 105 | 1971-07-13 | Comedy | 01:22 |
| C_701 | There's a Girl in my Soup | 107 | 1970-06-11 | Comedy | 01:36 |

## 兼容性

### SQL92

SQL92 为 CREATE VIEW 语句规定了一些额外的能力：

```
CREATE VIEW view [ column [, ...] ]
  AS SELECT expression [AS colname] [, ...]
  FROM table
  [ WHERE condition ]
  [ WITH [ CASCADE | LOCAL ] CHECK OPTION ]
```

#### CHECK OPTION

此选项与可更新视图有关。视图上的所有 INSERT 和 UPDATE 都会被检查，以确保数据满足定义视图的条件。如果不满足，更新将被拒绝。

#### LOCAL

在该视图上检查完整性。

#### CASCADE

在该视图以及任何依赖视图上检查完整性。如果既未指定 CASCADE 也未指定 LOCAL，则假定使用 CASCADE。

# DECLARE

DECLARE — 定义一个用于表访问的游标

## 大纲

```
DECLARE cursor [ BINARY ] [ INSENSITIVE ] [ SCROLL ]  
        CURSOR FOR query  
        [ FOR { READ ONLY | UPDATE [ OF column [, ...] ] } ]
```

## 输入

*cursor*

用于后续 FETCH 操作的游标的名称。

BINARY

使游标以二进制而不是文本格式获取数据。

INSENSITIVE

SQL92 关键字，表示从游标检索的数据应不受其他进程或游标的更新的影响。由于 Postgres 中游标操作发生在事务内，情况总是如此。这个关键字没有任何效果。

SCROLL

SQL92 关键字，表示每次 FETCH 操作可以检索多行。由于 Postgres 总是允许这样做，这个关键字没有任何效果。

*query*

一个提供将由该游标管理的行的 SQL 查询。关于有效参数的更多信息请参阅 SELECT 语句。

READ ONLY

SQL92 关键字，表示游标将以只读模式使用。由于这是 Postgres 中唯一可用的游标访问模式，这个关键字没有任何效果。

UPDATE

SQL92 关键字，表示游标将用于更新表。由于 Postgres 目前不支持游标更新，这个关键字会引发一条信息性错误消息。

*column*

要更新的列。由于 Postgres 目前不支持游标更新，UPDATE 子句会引发一条信息性错误消息。

## 输出

SELECT

SELECT 成功运行时返回的消息。

NOTICE BlankPortalAssignName: portal "*cursor*" already exists

如果游标 "*cursor*" 已被声明，就会发生这个错误。

ERROR: Named portals may only be used in begin/end transaction blocks

如果游标不是在事务块内声明的，就会发生这个错误。

## 描述

DECLARE 允许用户创建游标，可用于从较大的查询中一次检索少量行。游标可以以文本或二进制格式返回数据。

普通游标以文本格式（ASCII 或其他编码方案，取决于 Postgres 后端的构建方式）返回数据。由于数据本来就是以二进制格式存储的，系统必须做一次转换来产生文本格式。此外，文本格式通常比对应的二进制格式更大。信息以文本形式返回后，客户端应用可能仍需要把它转换为二进制格式才能处理。

BINARY 游标以本机二进制表示形式返回数据。因此二进制游标往往稍微快一些，因为它们的转换开销更小。

例如，如果一个查询从整数列返回值一，用默认游标会得到字符串 '1'，而用二进制游标会得到一个等于 control-A ('^A') 的 4 字节值。

### 小心

使用 BINARY 游标应当小心。诸如 psql 这样的用户应用不识别二进制游标，并期望数据以文本格式返回。

不过，字符串表示与体系结构无关，而二进制表示在不同的机器体系结构之间可能不同。因此，如果你的客户端机器和服务端机器使用不同的表示（例如 "big-endian" 对 "little-endian"），你大概不想让数据以二进制格式返回。

### 提示

如果你想以 ASCII 显示数据，以 ASCII 取回它会省去客户端的一些工作。

## 注解

游标只在事务中可用。

Postgres 没有显式的 OPEN *cursor* 语句；游标在声明时即视为打开。

### 注意

在 SQL92 中，游标只在嵌入式应用中可用。ecpg（Postgres 的嵌入式 SQL 预处理器）支持 SQL92 的约定，包括涉及 DECLARE 和 OPEN 语句的那些。

## 用法

声明一个游标：

```
DECLARE liahona CURSOR  
    FOR SELECT * FROM films;
```

## 兼容性

### SQL92

SQL92 只允许在嵌入式 SQL 和模块中使用游标。Postgres 允许交互式地使用游标。SQL92 允许嵌入式或模块游标更新数据库信息。所有 Postgres 游标都是只读的。BINARY 关键字是一种 Postgres 扩展。

# DELETE

DELETE — 删除一个表中的行

## 大纲

```
DELETE FROM table [ WHERE condition ]
```

## 输入

*table*

一个已存在表的名称。

*condition*

这是一个 SQL 选择查询，它返回要删除的行。

关于 WHERE 子句的进一步描述请参考 SELECT 语句。

## 输出

```
→ DELETE count
```

成功删除了项时返回的消息。*count* 是删除的行数。

如果 *count* 为 0，表示没有行被删除。

## 描述

DELETE 从指定表中删除满足 WHERE *condition* 的行。

如果省略 *condition*，则删除表中的所有行。结果是一个有效但为空的表。

要修改一个表，你必须拥有对它的写权限，同时对 *condition* 中读取其值的任何表拥有读权限。

## 用法

删除除音乐剧外的所有电影：

```
DELETE FROM films WHERE kind <> 'Musical';
```

```
SELECT * FROM films;
```

| code  | title                     | id  | date_prod  | kind    | len   |
|-------|---------------------------|-----|------------|---------|-------|
| UA501 | West Side Story           | 105 | 1961-01-03 | Musical | 02:32 |
| TC901 | The King and I            | 109 | 1956-08-11 | Musical | 02:13 |
| WD101 | Bed Knobs and Broomsticks | 111 |            | Musical | 01:57 |

(3 rows)

清空表 `films`:

```
DELETE FROM films;
```

```
SELECT * FROM films;
code|title|did|date_prod|kind|len
----+-----+---+-----+----+---
(0 rows)
```

## 兼容性

### SQL92

SQL92允许定位式 `DELETE` 语句:

```
DELETE FROM table WHERE
    CURRENT OF cursor
```

其中 *cursor* 标识一个已打开的游标。Postgres中的交互式游标是只读的。

# DROP AGGREGATE

DROP AGGREGATE — 删除一个聚集函数的定义

## 大纲

```
DROP AGGREGATE name type
```

## 输入

*name*

一个已存在聚集函数的名称。

*type*

一个已存在聚集函数的类型。（关于数据类型的更多信息请参考 [PostgreSQL User's Guide](#)）。

## 输出

→ DROP

命令成功时返回的消息。

→ WARN RemoveAggregate: aggregate '*name*' for '*type*' does not exist

如果指定的聚集函数在数据库中不存在，就会出现这条消息。

## 描述

DROP AGGREGATE 将删除一个已存在的聚集定义。要执行这条命令，当前用户必须是该聚集函数的拥有者。

## 注意

DROP AGGREGATE 语句是一种 Postgres 语言扩展。

请参考 CREATE AGGREGATE 语句来创建聚集函数。

## 用法

删除类型 `int4` 的 `myavg` 聚集函数：

```
DROP AGGREGATE myavg int4;
```

## 兼容性

## SQL92

SQL92中没有 DROP AGGREGATE 语句。

# DROP DATABASE

DROP DATABASE — 销毁一个现有数据库

## 大纲

```
DROP DATABASE name
```

## 输入

→ *name*

要删除的现有数据库的名称。

## 输出

→ DESTROYDB

命令成功时返回此消息。

→ WARN: destroydb: database "*name*" does not exist.

如果指定的数据库不存在，就会出现此消息。

→ ERROR: destroydb cannot be executed on an open database

如果指定的数据库不存在，就会出现此消息。

## 描述

DROP DATABASE 删除现有数据库的目录项并删除包含数据的目录。它只能由数据库管理员执行（细节见CREATE DATABASE命令）。

## 注解

DROP DATABASE 语句是Postgres 的语言扩展。

### 提示

当连接到目标数据库时无法执行这条查询。通常更可取的做法是改用destroydb脚本。

关于如何创建数据库的信息，请参阅 CREATE DATABASE语句。

## 兼容性

### SQL92

DROP DATABASE 在 SQL92 中没有。

# DROP FUNCTION

DROP FUNCTION — 移除一个用户定义的 C 函数

## 大纲

```
DROP FUNCTION name ( [ type [, ...] ] )
```

## Inputs

*name*

现有函数的名称。

*type*

函数参数的类型。

## Outputs

→ DROP

命令成功完成时返回的消息。

→ WARN RemoveFunction: Function "*name*" ("*types*") does not exist

如果指定的函数在当前数据库中不存在，则给出此消息。

## 描述

DROP FUNCTION 将删除对现有 C 函数的引用。要执行此命令，用户必须是该函数的所有者。必须指定函数的输入参数类型，因为只有具有给定名称和参数类型的函数才会被删除。

## 注意

关于创建聚合函数，参见 CREATE FUNCTION。

## 用法

这条命令删除平方根函数：

```
DROP FUNCTION sqrt(int4);
```

## 缺陷

系统不会检查依赖该函数的类型、操作符或访问方法是否已被先行删除。

## 兼容性

DROP FUNCTION 是一种 Postgres 语言扩展。

## SQL/PSM

SQL/PSM 是一个实现函数可扩展性的提议标准。SQL/PSM 的 DROP FUNCTION 语句语法如下：

```
DROP [ SPECIFIC ] FUNCTION name { RESTRICT | CASCADE }
```

# DROP INDEX

DROP INDEX — 从数据库中移除一个索引

## 大纲

```
DROP INDEX index_name
```

## 输入

```
→ index_name
```

要移除的该索引的名称。

## 输出

```
→ DROP
```

索引被成功删除时返回的消息。

```
→ ERROR: index "index_name" nonexistent
```

如果 *index\_name* 不是数据库中的一个索引，就会出现这条消息。

## 描述

DROP INDEX 从数据库系统中删除一个现有的索引。要执行这个命令，你必须是指定索引的拥有者。

## 注意

DROP INDEX 是一种 Postgres 语言扩展。

关于如何创建索引的信息，请参阅 CREATE INDEX 语句。

## 用法

这条命令将移除 `title_idx` 索引：

```
DROP INDEX title_idx;
```

## 兼容性

### SQL92

SQL92 定义了访问通用关系数据库的命令。索引是一种实现相关的特性，因此 SQL92 语言中没有索引特定的命令或定义。

# DROP LANGUAGE

DROP LANGUAGE — 移除一个用户定义的过程语言

## 大纲

```
DROP PROCEDURAL LANGUAGE 'langname'
```

## 输入

*langname*

现有语言的名称。

## 输出

→ DROP

语言成功删除时返回此消息。

→ ERROR: Language "*langname*" doesn't exist

如果找不到名为 "*langname*" 的语言，就会出现这条消息。

## 描述

DROP PROCEDURAL LANGUAGE 将删除此前注册的名为 '*langname*' 的过程语言的定义。

## 注意

DROP PROCEDURAL LANGUAGE 语句是一种 Postgres 语言扩展。

关于如何创建过程语言的信息，参见 CREATE PROCEDURAL LANGUAGE。

## 缺陷

系统不会检查是否有使用该语言注册的函数或触发器过程仍然存在。  
为了不删除并重建全部函数就能重新启用它们，必须把这些函数在 pg\_proc 中的 prolang 属性调整为为该 PL 重建的 pg\_language 条目的新对象 ID。

## 用法

这条命令删除 PL/Sample 语言：

```
DROP PROCEDURAL LANGUAGE 'plsample'
```

## 兼容性

## SQL92

SQL92 中没有 `DROP PROCEDURAL LANGUAGE`。

# DROP OPERATOR

DROP OPERATOR — 从数据库中移除一个操作符

## 大纲

```
DROP OPERATOR id ( type | NONE [,...] )
```

## 输入

*id*

现有操作符的标识符。

*type*

函数参数的类型。

## 输出

→ DROP

命令成功时返回的消息。

```
→ ERROR: RemoveOperator: binary operator 'id' taking 'type1' and 'type2' does not exist
```

如果指定的二元操作符不存在，就会出现这条消息。

```
→ ERROR: RemoveOperator: left unary operator 'id' taking 'type' does not exist
```

如果指定的左一元操作符不存在，就会出现这条消息。

```
→ ERROR: RemoveOperator: right unary operator 'id' taking 'type' does not exist
```

如果指定的右一元操作符不存在，就会出现这条消息。

## 描述

DROP OPERATOR 语句从数据库中删除一个现有的操作符。要执行此命令，你必须是该操作符的所有者。

左一元或右一元操作符的左类型或右类型可以相应地指定为 NONE。

## 注意

DROP OPERATOR 语句是一种 Postgres 语言扩展。

关于如何创建操作符的信息，参见 CREATE OPERATOR。

删除任何依赖被删操作符的访问方法和操作符类是用户的责任。

## 用法

移除 `int4` 的幂操作符 `a^n`:

```
DROP OPERATOR ^ (int4, int4);
```

移除布尔值的左一元操作符 `!a`:

```
DROP OPERATOR ! (none, bool);
```

移除 `int4` 的右一元阶乘操作符 `a!`:

```
DROP OPERATOR ! (int4, none);
```

## 兼容性

### SQL92

SQL92 中没有 `DROP OPERATOR`。

# DROP RULE

DROP RULE — 从数据库中删除一个已存在的规则

## 大纲

```
DROP RULE name
```

## 输入

*name*

要删除的已存在规则的名称。

## 输出

→ DROP

成功时返回的消息。

→ ERROR: RewriteGetRuleEventRel: rule "*name*" not found

如果指定的规则不存在，就会出现这条消息。

## 描述

DROP RULE 从指定的 Postgres 规则系统中删除一个规则。Postgres 会立即停止强制执行该规则，并从系统目录中清除其定义。

## 注意

DROP RULE 语句是一种 Postgres 语言扩展。

关于如何创建规则的信息，请参考 CREATE RULE。

## 缺陷

规则一旦被删除，对该规则所写入的历史信息的访问可能会随之消失。

## 用法

删除重写规则 *newrule*:

```
DROP RULE newrule;
```

## 兼容性

## SQL92

SQL92 中没有 DROP RULE。

# DROP SEQUENCE

DROP SEQUENCE — 移除一个现有序列

## 大纲

```
DROP SEQUENCE seqname [, ...]
```

## 输入

*seqname*

序列的名称。

## 输出

→ DROP

序列被成功删除时返回的消息。

→ WARN: Relation "*seqname*" does not exist.

指定的序列不存在时出现这条消息。

## 描述

DROP SEQUENCE 从数据库中移除序列号生成器。在当前把序列实现为特殊表的方式下，它的工作方式与 DROP TABLE 语句一样。

## 注意

DROP SEQUENCE 语句是一种 Postgres 语言扩展。

关于如何创建序列的信息，请参阅 CREATE SEQUENCE 语句。

## 用法

要从数据库中移除序列 *serial*：

```
DROP SEQUENCE serial
```

## 兼容性

### SQL92

DROP SEQUENCE 没有出现在 SQL92 中。

# DROP TABLE

DROP TABLE — 从数据库中删除已存在的表

## 大纲

```
DROP TABLE table [, ...]
```

## 输入

→ *table*

要删除的已存在表或视图的名称。

## 输出

→ DROP

命令成功完成时返回的消息。

→ ERROR Relation "*table*" Does Not Exist!

如果指定的表或视图在数据库中不存在。

## 描述

DROP TABLE 从数据库中删除表和视图。只有其所有者才能删除一个表或视图。一个表可以通过使用 DELETE 清空行，而不被删除。

如果被删除的表上有二级索引，会先删除它们。只删除一个二级索引不会影响底层表的内容。

## 注意

关于如何创建或修改表的信息，请参考 CREATE TABLE 和 ALTER TABLE。

## 用法

删除 films 和 distributors 两个表：

```
DROP TABLE films, distributors
```

## 兼容性

### SQL92

SQL92 为 DROP TABLE 规定了一些额外的能力：

```
DROP TABLE table { RESTRICT | CASCADE }
```

## RESTRICT

保证只有没有任何依赖视图或完整性约束的表才能被删除。

## CASCADE

所有引用它的视图或完整性约束也将被删除。

### 提示

目前，要删除一个被引用的视图，必须显式删除它。

# DROP TRIGGER

DROP TRIGGER — 移除一个触发器的定义

## 大纲

```
DROP TRIGGER name ON table
```

## 输入

*name*

一个现有触发器的名称。

*table*

一个表的名称。

## 输出

→ DROP

触发器被成功删除时返回的消息。

→ ERROR: DropTrigger: there is no trigger *name* on relation "*table*"

指定的触发器不存在时出现这条消息。

## 描述

DROP TRIGGER 将移除对一个现有触发器定义的所有引用。要执行这个命令当前用户必须是该触发器的拥有者。

## 注意

DROP TRIGGER 是一种 Postgres 语言扩展。

关于如何创建触发器的信息，请参阅 CREATE TRIGGER。

## 用法

销毁表 `films` 上的 `if_dist_exists` 触发器：

```
DROP TRIGGER if_dist_exists ON films;
```

## 兼容性

### SQL92

DROP TRIGGER 语句没有出现在 SQL92 中。

# DROP TYPE

DROP TYPE — 从系统目录中移除一个用户定义的类型

## 大纲

```
DROP TYPE typename
```

## 输入

*typename*

现有类型的名称。

## 输出

→ DROP

命令成功时返回的消息。

→ ERROR: RemoveType: type '*typename*' does not exist

如果找不到指定的类型，就会出现这条消息。

## 描述

DROP TYPE 将从系统目录中删除一个用户类型。

只有类型的所有者才能删除它。

## 注意

DROP TYPE 语句是一种 Postgres 语言扩展。

关于如何创建类型的信息，参见 CREATE TYPE。

删除任何使用被删类型的操作符、函数、聚合、访问方法、子类型和类是用户的责任。

## 缺陷

如果删除了内置类型，后端的行为将不可预料。

## 用法

要删除 box 类型：

```
DROP TYPE box
```

## 兼容性

## SQL3

DROP TYPE 是一个 SQL3 语句。

# DROP USER

DROP USER — 删除一个用户账户的信息

## 大纲

```
DROP USER username
```

## 输入

*username*

一个已存在用户的名称。

## 输出

→ DROP

成功删除用户时返回的消息。

→ ERROR: removeUser: user "*username*" does not exist.

用户名未找到时出现此消息。

## 描述

DROP USER 从数据库中删除指定的用户，连同该用户拥有的任何数据库。它不会删除该用户在不属于他自己的数据库中拥有的表、视图或触发器。这条语句可以代替 `destroyuser` 脚本使用，无论该用户是如何创建的。

## 注解

DROP USER 是一种 Postgres 语言扩展。

如何创建或修改用户账户的信息请参考 CREATE USER 和 ALTER USER。

## 用法

删除一个用户账户：

```
DROP USER Jonathan;
```

## 兼容性

### SQL92

DROP USER 在 SQL92 中没有定义。

# DROP VIEW

DROP VIEW — 从数据库中删除一个现有视图

## 大纲

```
DROP VIEW view
```

## 输入

```
view
```

现有视图的名称。

## 输出

→ DROP

命令成功时返回的消息。

→ ERROR: RewriteGetRuleEventRel: rule "\_RETview" not found

如果指定的视图在数据库中不存在，就会出现此消息。

## 描述

DROP VIEW从数据库中删除一个现有视图。要执行这条命令，你必须是该视图的所有者。

## 注解

Postgres DROP TABLE语句也会删除视图。

关于如何创建视图的信息，请参阅CREATE VIEW。

## 用法

这条命令将删除名为kinds的视图：

```
DROP VIEW kinds;
```

## 兼容性

## SQL92

SQL92为DROP VIEW规定了一些额外的能力：

```
DROP VIEW view { RESTRICT | CASCADE }
```

## 输入

RESTRICT

确保只有没有依赖视图或完整性约束的视图才能被销毁。

CASCADE

所有引用它的视图和完整性约束也将被删除。

## 注解

### 提示

目前，要从Postgres数据库中删除一个被引用的视图，你必须显式地删除它。

# EXPLAIN

EXPLAIN — 显示语句的执行细节

## 大纲

```
EXPLAIN [ VERBOSE ] query
```

## 输入

VERBOSE

显示详细查询计划的标志。

*query*

任何 *query*。

## 输出

NOTICE: QUERY PLAN: *plan*

来自Postgres后端的显式查询计划。

EXPLAIN

查询计划显示完毕后发出的标志。

## 描述

这条命令输出给定查询的细节。默认的输出是计算得到的查询代价。VERBOSE 在你的屏幕上显示完整的查询计划和代价，并把美化打印的计划发送到 postmaster 日志文件。

## 注解

关于优化器在Postgres中对代价信息的使用，目前只有很少的文档。  
关于查询优化的代价估计的一般信息可以在数据库教科书中找到。更多信息请参阅程序员指南中关于索引和遗传查询优化器的章节。

## 用法

要显示一个简单查询的查询计划：

```
postgres=> explain select * from foo;  
NOTICE:  QUERY PLAN:
```

```
Seq Scan on foo  (cost=0.00 size=0 width=4)
```

EXPLAIN

## 兼容性

## SQL92

SQL92 中没有定义 EXPLAIN 语句。

# FETCH

FETCH — 使用游标获取行

## 大纲

```
FETCH [ selector ] [ count ] { IN | FROM } cursor
FETCH [ RELATIVE ] [ { [ # | ALL | NEXT | PRIOR ] } ] FROM ] cursor
```

## 输入

*selector*

*selector* 定义提取方向。它可以是下列之一：

**FORWARD**

提取下一个（或下一些）行。如果省略 *selector*，这是默认值。

**BACKWARD**

提取前一个（或前一些）行。

**RELATIVE**

为兼容 SQL92 而存在的噪声词。

*count*

*count* 决定提取多少行。它可以是下列之一：

**#**

一个有符号整数，指定要提取多少行。注意，负整数等价于反转 **FORWARD** 和 **BACKWARD** 的方向。

**ALL**

检索所有剩余的行。

**NEXT**

等价于指定数量 1。

**PRIOR**

等价于指定数量 -1。

*cursor*

一个已打开游标的名称。

## 输出

FETCH 返回由指定游标定义的查询的结果。如果查询失败，将返回下列消息：

NOTICE: PerformPortalFetch: portal "*cursor*" not found

如果 *cursor* 此前未被声明。游标必须在事务块内声明。

NOTICE: FETCH/ABSOLUTE not supported, using RELATIVE

Postgres 不支持游标的绝对定位。

ERROR: FETCH/RELATIVE at current position is not supported

SQL92 允许使用语法

```
FETCH RELATIVE 0 FROM cursor
```

反复检索游标"当前位置"上的行。Postgres 目前不支持这一概念；实际上，值零被保留用来表示要检索所有行，等价于指定 ALL 关键字。如果使用了 RELATIVE 关键字，Postgres 假定用户想要的是 SQL92行为，并返回这条错误消息。

## 描述

FETCH 允许用户使用游标检索行。检索的行数由 # 指定。如果游标中剩余的行数少于 #，则只提取那些可用的行。用关键字 ALL 代替数字将检索游标中所有剩余的行。可以沿 FORWARD 和 BACKWARD 两个方向提取行。默认方向是 FORWARD。

### 提示

现在允许把行数指定为负数。负数等价于反转 FORWARD 和 BACKWARD 关键字的方向。例如，FORWARD -1与BACKWARD 1相同。

注意，FORWARD 和 BACKWARD 关键字是 Postgres 扩展。也支持 SQL92语法，即命令的第二种形式。关于兼容性问题的细节见下文。

一旦所有行都被提取，之后的每次提取访问都不再返回行。

Postgres 不支持通过游标更新数据，因为把游标更新映射回基表通常不可行，VIEW 更新也是同样的情况。因此，用户必须发出显式的 UPDATE 命令来替换数据。

游标只能在事务内部使用，因为它们保存的数据跨越用户的多次查询。

## 注意

参阅 MOVE 语句来改变游标位置。参阅 DECLARE 语句来声明游标。参阅 BEGIN WORK、COMMIT WORK、ROLLBACK WORK 语句以获得关于事务的更多信息。

## 用法

```
--set up and use a cursor:
--
BEGIN WORK;
  DECLARE liahona CURSOR
    FOR SELECT * FROM films;

--Fetch first 5 rows in the cursor liahona:
```

```
--
    FETCH FORWARD 5 IN liahona;

    code |title                                |did| date_prod|kind        |len
    -----+-----+-----+-----+-----+-----
    BL101|The Third Man                            |101|1949-12-23|Drama       |01:44
    BL102|The African Queen                        |101|1951-08-11|Romantic    |01:43
    JL201|Une Femme est une Femme                 |102|1961-03-12|Romantic    |01:25
    P_301|Vertigo                                  |103|1958-11-14|Action      |02:08
    P_302|Becket                                  |103|1964-02-03|Drama       |02:28

--Fetch previous row:
--
    FETCH BACKWARD 1 IN liahona;

    code |title                                |did| date_prod|kind        |len
    -----+-----+-----+-----+-----+-----
    P_301|Vertigo                                  |103|1958-11-14|Action      |02:08

-- close the cursor and commit work:
--
    CLOSE liahona;
    COMMIT WORK;
```

## 兼容性

游标的非嵌入式使用是 Postgres 扩展。这里把游标的语法和用法与 SQL92 定义的嵌入式游标形式相比较。

## SQL92

SQL92 允许 FETCH 绝对定位游标，并允许把结果放入显式变量：

```
FETCH ABSOLUTE #
    FROM cursor
    INTO :variable [, ...]
```

### ABSOLUTE

游标应被定位到指定的绝对行号。Postgres 中的所有行号都是相对编号，因此不支持这一能力。

:*variable*

目标主变量。

# GRANT

GRANT — 把访问权限授予一个用户、一个组或所有用户

## 大纲

```
GRANT privilege [, ...]
    ON object [, ...]
    TO { PUBLIC | GROUP group | username }
```

## 输入

*privilege*

可能的权限有：

SELECT

访问指定表/视图的所有列。

INSERT

向指定表的所有列插入数据。

UPDATE

更新指定表的所有列。

DELETE

从指定表中删除行。

RULE

在表/视图上定义规则（见 CREATE RULE 语句）。

ALL

授予所有权限。

*object*

要授予访问权限的对象的名称。可能的对象有：

- table
- view
- sequence
- index

PUBLIC

代表所有用户的简写形式。

GROUP *group*

被授予权限的一个组 *group*。在当前版本中，组必须按下面描述的方式显式创建。

*username*

被授予权限的用户的名称。PUBLIC 是代表所有用户的简写形式。

## 输出

CHANGE

如果成功则返回此消息。

ERROR: ChangeAcl: class "*object*" not found

如果指定的对象不可用，或者无法把权限授予指定的组或用户，则返回此消息。

## 描述

GRANT 允许对象的创建者把特定权限授予所有用户（PUBLIC）或某个用户或组。除非创建者在对象创建之后 GRANT 权限，除创建者之外的用户没有任何访问权限。

一旦用户拥有了对象上的某个权限，他就能够行使该权限。无需向对象的创建者 GRANT 权限，创建者自动持有所有权限，并且也可以删除对象。

## 注意

使用 `psql \z` 命令获取现有对象上权限的更多信息：

```
Database      = lusitania
+-----+-----+
| Relation    | Grant/Revoke Permissions |
+-----+-----+
| mytable     | {"=rw","miriam=arwR","group todos=rw"} |
+-----+-----+
Legend:
      uname=arwR -- privileges granted to a user
      group gname=arwR -- privileges granted to a GROUP
              =arwR -- privileges granted to PUBLIC

              r -- SELECT
              w -- UPDATE/DELETE
              a -- INSERT
              R -- RULE
          arwR -- ALL
```

### 提示

目前，要创建 GROUP，必须手工向表 `pg_group` 中插入数据：

```
INSERT INTO pg_group VALUES ('todos');
CREATE USER miriam IN GROUP todos;
```

参阅 REVOKE 语句来撤销访问权限。

## 用法

```
-- grant insert privilege to all users on table films:
--
GRANT INSERT ON films TO PUBLIC;

-- grant all privileges to user manuel on view kinds:
--
GRANT ALL ON kinds TO manuel;
```

## 兼容性

### SQL92

SQL92 的 GRANT 语法允许为表中的单个列设置权限，也允许设置把相同权限授予他人的权限。

```
GRANT privilege [, ...]
    ON object [ ( column [, ...] ) ] [, ...]
    TO { PUBLIC | username [, ...] }
    [ WITH GRANT OPTION ]
```

各字段与 Postgres 实现中的兼容，并有下列补充：

*privilege* SELECT

SQL92 允许指定额外的权限：

REFERENCES

允许在完整性约束中引用指定表/视图的部分或全部列。

USAGE

允许使用域、字符集、排序规则或转换。如果对象指定的不是表/视图，*privilege* 必须只指定 USAGE。

### 提示

目前，在 Postgres 中要只对少数几个列授予权限，必须创建一个包含所需列的视图，然后把权限授予该视图。

*object*

*object*

SQL92 允许一个额外的无功能关键字：

[ TABLE ] *table*

CHARACTER SET

允许使用指定的字符集。

**COLLATION**

允许使用指定的排序序列。

**TRANSLATION**

允许使用指定的字符集转换。

**DOMAIN**

允许使用指定的域。

**WITH GRANT OPTION**

允许把相同的权限授予他人。

# INSERT

INSERT — 向一个表中插入新行

## 大纲

```
INSERT INTO table [ ( column [, ...] ) ]  
    { VALUES ( expression [, ...] ) | SELECT query }
```

## 输入

*table*

一个已存在表的名称。

*column*

*table* 中一个列的名称。

*expression*

要赋予 *column* 的一个有效表达式或值。

*query*

一个有效的查询。有效参数的进一步描述请参考 SELECT 语句。

## 输出

→ INSERT oid1

只插入了一行时返回的消息。→ *oid* 是被插入行的行标识符。

→ INSERT 0 #

插入了多行时返回的消息。→ # 是插入的行数。

## 描述

INSERT 允许我们向一个表中插入新行。可以一次插入一行，也可以插入一个查询结果的多个行。目标列表中的列可以按任意顺序列出。对于目标列表中没有出现的每个列，将插入默认值；如果列没有声明的默认值，则假定为 NULL。如果每个列的表达式数据类型不正确，将尝试自动类型强制转换。

要向一个表追加数据，你必须拥有该表上的 insert 权限，并且对 WHERE 子句中指定的任何表拥有 select 权限。

## 用法

```
--Insert a single row into table films;  
--(in the second example the column date_prod is omitted
```

```
--therefore will be stored in it a default value of NULL):
--
INSERT INTO films VALUES
    ('UA502', 'Bananas', 105, '1971-07-13', 'Comedy', INTERVAL '82 minute')
;

INSERT INTO films (code, title, did, date_prod, kind)
    VALUES ('T_601', 'Yojimbo', 106, DATE '1961-06-16', 'Drama');

--Insert a single row into table distributors, note that
--only column "name" is specified, to the non specified
--column "did" will be assigned its default value:
--
INSERT INTO distributors (name) VALUES ('British Lion');

--Insert several rows into table films from table tmp:
--
INSERT INTO films
    SELECT * FROM tmp;

--Insert into arrays:
--Create an empty 3x3 gameboard for noughts-and-crosses
--(all of these queries create the same board attribute)
--(Refer to the PostgreSQL User's Guide for further
--information about arrays).

INSERT INTO tictactoe (game, board[1:3][1:3])
    VALUES (1, '{{"","",""}, {}, {"",""}}');
INSERT INTO tictactoe (game, board[3][3])
    VALUES (2, '{}');
INSERT INTO tictactoe (game, board)
    VALUES (3, '{{,,},{,,},{,,}}');
```

## 兼容性

### SQL92

INSERT 语句与 SQL92 完全兼容。*query*子句特性方面可能存在的限制记录在 SELECT 语句中。

# LISTEN

LISTEN — 监听某个通知条件上的通知

## 大纲

```
LISTEN notifysname
```

## 输入

*notifysname*

通知条件的名称。

## 输出

→ LISTEN

注册成功完成时返回的消息。

→ NOTICE Async\_Listen: We are already listening on *notifysname*

如果该后端已经注册了该通知条件。

## 描述

LISTEN 把当前的Postgres后端注册为通知条件 *notifysname*的监听者。

每当命令 `NOTIFY notifysname` 被调用时——无论是被本后端还是被连接到同一数据库的其他后端调用——所有当前正在监听该通知条件的后端都会被通知，而每个后端又将进而通知它所连接的前端应用。更多信息请参阅 `NOTIFY` 的讨论。

可以用 `UNLISTEN` 命令取消一个后端对给定通知条件的注册。此外，后端的监听注册在后端进程退出时会被自动清除。

前端应用检测通知事件所必须使用的方法取决于它使用的 Postgres 应用编程接口。使用基本的 `libpq` 库时，应用像发出普通 SQL 命令一样发出 `LISTEN`，然后必须周期性地调用例程 `PQnotifies` 来查明是否收到了任何通知事件。其他接口（例如 `libpqctl`）提供了处理通知事件的更高层方法；实际上，使用 `libpqctl` 时，应用程序员甚至不必直接发出 `LISTEN` 或 `UNLISTEN`。更多细节请参阅你所使用的库的文档。

`NOTIFY` 的参考页包含关于 `LISTEN` 和 `NOTIFY` 用法的更广泛的讨论。

## 注解

*notifysname* 可以是任何可作为名称使用的有效字符串；它不需要对应任何实际表的名称。如果 *notifysname* 被双引号括起来，它甚至不必是语法上有效的名称，而可以是长度不超过 31 个字符的任何字符串。

在Postgres的一些早期版本中，*notifysname* 不对应任何现有表名时，即使它在语法上是有效的名称，也必须用双引号把它括起来。现在已不再需要这样。

## 用法

```
-- Configure and execute a listen/notify sequence from psql
postgres=> listen virtual;
LISTEN
postgres=> notify virtual;
NOTIFY
ASYNC NOTIFY of 'virtual' from backend pid '11239' received
```

## 兼容性

### SQL92

LISTEN 在 SQL92 中没有。

# LOAD

LOAD — 动态装载一个目标文件

## 大纲

```
LOAD 'filename'
```

## 输入

*filename*

用于动态装载的目标文件。

## 输出

→ LOAD

成功完成时返回的消息。

→ ERROR: LOAD: could not open file '*filename*'

找不到指定的文件时返回的消息。该文件必须对 Postgres 后端可见，并指定适当的完整路径名，才能避免此消息。

## 描述

把一个目标（或".o"）文件装载到 Postgres 后端地址空间中。文件一旦装载，该文件中的所有函数都可以被访问。这个函数用于支持用户定义的类型和函数。

如果文件没有用 LOAD 装载，该文件将在函数第一次被 Postgres 调用时被自动装载。LOAD 也可以用来在目标文件被编辑并重新编译之后重新装载它。目前只支持由 C 语言文件创建的目标。

## 注解

已装载目标文件中的函数不应调用其他通过 LOAD 命令装载的目标文件中的函数。例如，文件 A 中的所有函数应当互相调用、调用标准库或数学库中的函数、或调用 Postgres 本身中的函数。它们不应调用另一个已装载文件 B 中定义的函数。这是因为如果 B 被重新装载，Postgres 装载器无法把 A 中函数的调用重定位到 B 的新地址空间。但如果 B 没有被重新装载，则不会有问题。

目标文件必须编译为包含位置无关代码。例如，在 DECstation 上，编译要装载的目标文件时必须使用 /bin/cc 加上 -G 0 选项。

注意，如果你正在把 Postgres 移植到新平台，LOAD 必须能够工作才能支持 ADT。

## 用法

```
--Load the file /usr/postgres/demo/circle.o
--
LOAD '/usr/postgres/demo/circle.o'
```

## 兼容性

### SQL92

LOAD 在 SQL92 中没有。

# LOCK

LOCK — 在事务内显式锁定一个表

## 大纲

```
LOCK [ TABLE ] table
```

## 输入

*table*

要锁定的现有表的名称。

## 输出

DELETE 0

锁成功时返回的消息。LOCK 实现为一条保证不会删除任何行的 `DELETE FROM table`。

ERROR *table*: Table does not exist.

如果 *table* 不存在则返回此消息。

## 描述

LOCK 在事务内以排他模式锁定一个表。它的经典用例是：你想在一个事务中先选出一些数据，然后更新它们。如果不用 LOCK 语句显式锁定表，表只会在第一次 UPDATE、INSERT 或 DELETE 操作时才被隐式锁定。如果你在 select 之前没有排他锁定该表，别的用户也可能读到被选中的数据，并尝试做他们自己的更新，导致你们双方互相等待对方释放由 select 产生的共享锁以便取得排他锁去做更新，从而形成死锁。

死锁的另一个例子是：一个用户锁定了一张表，另一个用户锁定了第二张表。在双方都保持现有锁的同时，第一个用户试图锁定第二个用户的表，第二个用户试图锁定第一个用户的表。两个用户都死锁等待表变为可用。唯一的解决办法是两个用户以相同的顺序锁定表，使各用户的锁获取和请求不会形成死锁。

### 注意

Postgres 确实会检测死锁，并会回滚事务来解决死锁。通常至少有一个死锁的事务会成功完成。

## 注意

LOCK 是一种 Postgres 语言扩展。

LOCK 只在事务内部起作用。

### Bug

如果被锁定的表被删除，那么即使事务仍在进行中，它也会被自动解锁。

## 用法

```
--Explicit locking to prevent deadlock:
--
BEGIN WORK;
    LOCK films;
    SELECT * FROM films;
    UPDATE films SET len = INTERVAL '100 minute'
        WHERE len = INTERVAL '117 minute';
COMMIT WORK;
```

## 兼容性

### SQL92

SQL92 中没有 `LOCK TABLE`，而是使用 `SET TRANSACTION` 来指定事务的并发级别。

# MOVE

MOVE — 移动游标位置

## 大纲

```
MOVE [ selector ] [ count ]
    { IN | FROM } cursor
FETCH [ RELATIVE ] [ { [ # | ALL | NEXT | PRIOR ] } ] FROM ] cursor
```

## 描述

MOVE允许用户把游标位置移动指定的行数。MOVE的工作方式类似FETCH命令，但它只定位游标而不返回行。

有关语法和用法的细节，请参阅FETCH命令。

## 注意

MOVE是一个Postgres 语言扩展。

有效参数的描述请参阅FETCH。声明游标请参阅DECLARE。有关事务的更多信息，请参阅BEGIN WORK、COMMIT WORK、ROLLBACK WORK语句。

## 用法

```
--建立并使用一个游标：
--
BEGIN WORK;
    DECLARE liahona CURSOR  FOR SELECT * FROM films;

    --跳过前 5 行：
    --
    MOVE FORWARD 5 IN liahona;
MOVE
    --从游标 liahona 中提取第 6 行：
    --
    FETCH 1 IN liahona;
FETCH
code |title |did| date_prod|kind      |len
-----+-----+---+-----+-----+-----
P_303|48 Hrs|103|1982-10-22|Action    | 01:37
(1 row)
    -- 关闭游标 liahona 并提交事务：
    --
    CLOSE liahona;
COMMIT WORK;
```

## 兼容性

## SQL92

SQL92 中没有MOVE语句。作为替代，SQL92允许从游标的绝对位置 FETCH行，从而隐式地把游标移动到正确的地点。

# NOTIFY

NOTIFY — 向所有正在监听某个通知条件的前端和后端发信号

## 大纲

`NOTIFY notifyname`

## 输入

*notifyname*

要发信号的通知条件。

## 输出

NOTIFY

确认 notify 命令已执行。

Notify events

事件被送达正在监听的前端；每个前端应用是否响应以及如何响应取决于它的程序设计。

## 描述

NOTIFY 命令向当前数据库中每个此前已为指定通知条件执行过 `LISTEN notifyname` 的前端应用发送一个通知事件。

传递给前端的通知事件信息包括通知条件名和发出通知的后端进程的 PID。由数据库设计者来定义给定数据库中将要使用的条件名以及每个条件名的含义。

通常，通知条件名与数据库中某张表的名称相同，通知事件本质上意味着“我改动了这张表，看看它有什么新内容”。但 NOTIFY 和 LISTEN 命令并不强制这种关联。例如，数据库设计者可以用几个不同的条件名来表示对单个表不同种类的更改。

NOTIFY 为访问同一 Postgres 数据库的一组进程提供了一种简单的信号或 IPC（进程间通信）机制。通过使用数据库中的表，可以构建从通知者向监听者传递附加数据（而不仅仅是条件名）的更高层机制。

当 NOTIFY 用于表明某张特定表发生变化时，一种很有用的编程技巧是把 NOTIFY 放进由表更新触发的规则中。这样一来，每当表发生变化时就会自动发出通知，应用程序员也不容易忘记这么做。

NOTIFY 在若干重要方面与 SQL 事务交互。首先，如果在事务内部执行了 NOTIFY，通知事件直到且仅当事务提交时才会被送达。这是恰当的，因为如果事务中止，我们希望其中的所有命令都不产生效果，包括 NOTIFY。但如果期望通知事件立即送达，这可能会让人不安。其次，如果一个正在监听的后端在处于事务中时收到一个通知信号，该通知事件要到事务刚完成（提交或中止）之后才会送达其连接的前端。同样，其理由是：如果一个通知在某个后来被中止的事务内被送达，人们会希望以某种方式撤销该通知——但后端一旦把通知发给了前端，就无法“收回”它。因此通知事件只在事务之间送达。由此得出的结论是：使用 NOTIFY 进行实时信号传递的应用应尽量保持事务简短。

NOTIFY 在一个重要方面的行为类似 Unix 信号：如果在短时间内多次发出同一通知名的信号，接收者可能对多次执行的 NOTIFY 只收到一个通知事件。因此，依赖收到通知的数量并不是一个好主意。正确的做法是用 NOTIFY 唤醒需要关注某事的应用，而用一个数据库对象（例如序列）来跟踪发生了什么以及发生了多少次。

发送 NOTIFY 的前端自己同时也在监听同一通知名，这是很常见的情形。在这种情况下，它会像其他所有监听前端一样收到一个通知事件。根据应用逻辑，这可能导致无用功——例如再次读取数据库表，找到的却正是该前端自己刚刚写入的更新。在 Postgres 6.4 及之后的版本中，可以通过检查发出通知的后端进程的 PID（在通知事件消息中提供）是否与自己的后端的 PID（可从 libpq 获得）相同来避免这种额外工作。二者相同时，该通知事件就是自己发的工作回弹回来，可以忽略。（尽管前一段有那样的说法，但这是一种安全的技术。Postgres 会把自身通知与来自其他后端的的通知分开保存，因此忽略自己的通知并不会让你错过来自外部的通知。）

## 注意

*notifyname* 可以是任何作为名称有效的字符串；它不必对应任何实际表的名称。如果 *notifyname* 用双引号括起，它甚至不必是语法上有效的名称，而可以是长度最多 31 个字符的任意字符串。

在 Postgres 的一些早期版本中，当 *notifyname* 不对应任何现有表名时，即使语法上是有效的名称，也必须用双引号括起。这一要求已不再存在。

在 6.4 之前的 Postgres 版本中，通知消息中传递的后端 PID 总是前端自己的后端的 PID。因此在那些早期版本中，无法把自己的通知与其他客户的通知区分开。

## 用法

```
-- Configure and execute a listen/notify sequence from psql
postgres=> listen virtual;
LISTEN
postgres=> notify virtual;
NOTIFY
ASYNC NOTIFY of 'virtual' from backend pid '11239' received
```

## 兼容性

### SQL92

SQL92 中没有 NOTIFY 语句。

# RESET

RESET — 恢复会话的运行时参数为默认值

## 大纲

```
RESET variable
```

## 输入

```
variable
```

有关可用变量的更多信息，请参阅 SET 语句。

## 输出

```
RESET VARIABLE
```

如果 *variable* 被成功重置为默认值，返回此消息。

## 描述

RESET 把变量恢复为默认值。有关允许的取值和默认值，请参阅 SET 命令。RESET 是

```
SET variable = DEFAULT
```

的替代形式。

## 注意

RESET 语句是一个 Postgres 语言扩展。

设置/显示变量的值请参阅 SET/SHOW 语句。

## 用法

```
-- 把 DateStyle 重置为默认值;  
RESET DateStyle;
```

```
-- 把 Geqo 重置为默认值;  
RESET GEQO;
```

## 兼容性

### SQL92

SQL92 中没有 RESET。

# REVOKE

REVOKE — 从一个用户、一个组或所有用户回收访问权限。

## 大纲

```
REVOKE privilege [, ...]
      ON object [, ...]
      FROM { PUBLIC | GROUP group | username }
```

## 输入

*privilege*

可能的权限有：

SELECT

访问一个指定表/视图的所有列的权限。

INSERT

向一个指定表的所有列插入数据的权限。

UPDATE

更新一个指定表的所有列的权限。

DELETE

从一个指定表删除行的权限。

RULE

在表/视图上定义规则的权限。（见 CREATE RULE）。

ALL

回收所有权限。

*object*

要回收访问权限的对象的名称。可能的对象有：

- 表
- 视图
- 序列
- 索引

*group*

要从中回收权限的组的名称。

*username*

要从中回收权限的用户的名称。使用 PUBLIC 关键字指定所有用户。

PUBLIC

为所有用户回收指定的权限。

输出

CHANGE

成功时返回的消息。

ERROR

对象不可用或不可能从一个组或用户回收权限时返回的消息。

描述

REVOKE 允许一个对象的创建者回收先前授予所有用户（通过 PUBLIC）或某个用户或组的权限。

注意

关于现有对象上权限的更多信息，请参阅 psql 的 \z 命令：

```
Database      = lusitania
+-----+-----+
| Relation    | Grant/Revoke Permissions |
+-----+-----+
| mytable     | {"=rw","miriam=arwR","group todos=rw"} |
+-----+-----+
Legend:
    uname=arwR -- privileges granted to a user
    group gname=arwR -- privileges granted to a GROUP
               =arwR -- privileges granted to PUBLIC

    r -- SELECT
    w -- UPDATE/DELETE
    a -- INSERT
    R -- RULE
    arwR -- ALL
```

提示

目前，要创建一个 GROUP，你必须手动向表 pg\_group 中插入数据：

```
INSERT INTO pg_group VALUES ('todos');
CREATE USER miriam IN GROUP todos;
```

用法

```
-- revoke insert privilege from all users on table films:
--
REVOKE INSERT ON films FROM PUBLIC;
```

```
-- revoke all privileges from user manuel on view kinds:
--
REVOKE ALL ON kinds FROM manuel;
```

## 兼容性

### SQL92

SQL92 的 REVOKE 语法有更多回收权限的能力，包括对表中个别列的权限：

```
REVOKE { SELECT | DELETE | USAGE | ALL PRIVILEGES } [, ...]
      ON object
      FROM { PUBLIC | username [, ...] } { RESTRICT | CASCADE }
REVOKE { INSERT | UPDATE | REFERENCES } [, ...] [ ( column [,
... ] ) ]
      ON object
      FROM { PUBLIC | username [, ...] } { RESTRICT | CASCADE }
```

各个字段的细节请参阅 GRANT 命令。

```
REVOKE GRANT OPTION FOR privilege [, ...]
      ON object
      FROM { PUBLIC | username [, ...] } { RESTRICT | CASCADE }
```

回收一个用户把指定权限授予其他人的权力。各个字段的细节请参阅 GRANT 命令。

可能的对象有：

[ TABLE ] 表/视图  
CHARACTER SET 字符集  
COLLATION 排序规则  
TRANSLATION 转换  
DOMAIN 域

如果 user1 把一个权限带 WITH GRANT OPTION 地授予 user2，而 user2 又把它授予 user3，那么 user1 可以使用 CASCADE 关键字级联回收该权限。

如果 user1 把一个权限带 WITH GRANT OPTION 地授予 user2，而 user2 又把它授予 user3，那么当 user1 试图回收该权限时，如果他/她指定了 RESTRICT 关键字，回收将失败。

# ROLLBACK

ROLLBACK — 中止当前的事务

## 大纲

```
ROLLBACK [ WORK ]
```

## 输入

无。

## 输出

ABORT

成功时返回的消息。

NOTICE: UserAbortTransactionBlock and not in in-progress state ABORT

如果当前没有正在进行的事务。

## 描述

ROLLBACK 回滚当前事务并使该事务所做的所有更新被丢弃。

## 注意

关键字 WORK 是噪声词，可以省略。

使用 COMMIT 语句成功地终止一个事务。

## 用法

```
--To abort all changes:
--
ROLLBACK WORK;
```

## 兼容性

### SQL92

完全兼容。

# SELECT

SELECT — 从表或视图中检索行

## 大纲

```
SELECT [ALL|DISTINCT [ON column] ]  
      expression [ AS name ] [, ...]  
      [ INTO [TABLE] new_table ]  
      [ FROM table [alias] [, ...] ]  
      [ WHERE condition ]  
      [ GROUP BY column [, ...] ]  
      [ HAVING condition [, ...] ]  
      [ UNION [ALL] select ]  
      [ ORDER BY column [ ASC | DESC ] [, ...] ]
```

## 输入

*expression*

表的列名或一个表达式。

*name*

用 AS 子句为列或表达式指定另一个名称。*name* 不能用于 WHERE 条件中。但它可以在相关的 ORDER BY 或 GROUP BY 子句中引用。

*new\_table*

如果指定了 INTO TABLE 子句，查询结果将存储在另一个具有指定名称的表中。如果 *new\_table* 不存在，将自动创建它。更多信息参见 SELECT INTO。

### 注意

CREATE TABLE AS 语句也会从一个 select 查询创建一个新表。

*table*

FROM 子句引用的现有表的名称。

*alias*

前面表的替代名称。它用于简洁或消除单个表内连接的歧义。

*condition*

一个结果为真或假的布尔表达式。见 WHERE 子句。

*column*

表的列名。

*select*

一个除 ORDER BY 子句外具有所有特性的 select 语句。

## 输出

Rows

由查询说明产生的完整行集合。

→ *count*

查询返回的行数。

## 描述

SELECT 将取回满足 WHERE 条件的所有行，如果省略 WHERE，则取回一个表的所有行。

GROUP BY 子句允许用户在概念上把一个表划分为若干组。（见 GROUP BY 子句）。

HAVING 子句指定通过从先前指定子句的结果中消除某些组而导出的分组表。（见 HAVING 子句）。

ORDER BY 子句允许用户指定希望行按升序或降序模式操作符排序。（见 ORDER BY 子句）

UNION 子句指定一个由笛卡尔积并连接导出的表。（见 UNION 子句）。

你必须对一个表拥有 SELECT 权限才能读取它的值（见 GRANT/REVOKE 语句）。

## WHERE 子句

可选的 WHERE 条件具有如下一般形式：

*WHERE expr cond\_op expr [ log\_op ... ]*

其中 *cond\_op* 可以是 =、<、<=、>、>=、<> 之一，或 ALL、ANY、IN、LIKE 等条件操作符，而 *log\_op* 可以是 AND、OR、NOT 之一。比较返回 TRUE 或 FALSE，如果表达式求值为 FALSE 则所有实例都被丢弃。

## GROUP BY 子句

GROUP BY 指定通过应用此子句导出的分组表：

*GROUP BY column [, ...]*

## HAVING 子句

可选的 HAVING 条件具有如下一般形式：

*HAVING cond\_expr*

其中 *cond\_expr* 与为 WHERE 子句指定的相同。

HAVING 指定通过从先前指定子句的结果中消除不满足 *cond\_expr* 的组而导出的分组表。

*cond\_expr* 中引用的每一列都应无歧义地引用一个分组列。

## ORDER BY 子句

```
ORDER BY column [ ASC | DESC ] [, ...]
```

*column* 既可以是列名，也可以是序号。

序号指列的（从左到右的）位置。此特性使得可以基于没有合适名称的列定义排序。这永远不是绝对必要的，因为总是可以用 AS 子句为计算列赋予一个名称，例如：

```
SELECT title, date_prod + 1 AS newlen FROM films ORDER BY newlen;
```

ORDER BY 中的列必须出现在 SELECT 子句中。因此下面的语句是非法的：

```
SELECT name FROM distributors ORDER BY code;
```

还可以在 ORDER BY 子句中的每个列名后面添加关键字 DESC（降序）或 ASC（升序）。如果未指定，默认为 ASC。

## UNION 子句

```
table_query UNION [ ALL ]
table_query
    [ ORDER BY column [ ASC | DESC ] [, ...] ]
```

其中 *table\_query* 指定不带 ORDER BY 子句的任何 select 表达式。

UNION 操作符指定一个由笛卡尔积导出的表。表示 UNION 的直接操作数的两个表必须具有相同数量的列，并且对应的列必须是兼容的数据类型。

默认情况下，UNION 的结果不包含任何重复行，除非指定了 ALL 子句。

同一 SELECT 语句中的多个 UNION 操作符从左到右求值。注意 ALL 关键字不是全局性的，只应用于当前这对表结果。

## 用法

把表 films 与表 distributors 连接：

```
SELECT f.title, f.did, d.name, f.date_prod, f.kind
FROM distributors d, films f
WHERE f.did = d.did
```

| title                   | did | name            | date_prod  | kind     |
|-------------------------|-----|-----------------|------------|----------|
| The Third Man           | 101 | British Lion    | 1949-12-23 | Drama    |
| The African Queen       | 101 | British Lion    | 1951-08-11 | Romantic |
| Une Femme est une Femme | 102 | Jean Luc Godard | 1961-03-12 | Romantic |

|                           |                      |                     |
|---------------------------|----------------------|---------------------|
| Vertigo                   | 103 Paramount        | 1958-11-14 Action   |
| Becket                    | 103 Paramount        | 1964-02-03 Drama    |
| 48 Hrs                    | 103 Paramount        | 1982-10-22 Action   |
| War and Peace             | 104 Mosfilm          | 1967-02-12 Drama    |
| West Side Story           | 105 United Artists   | 1961-01-03 Musical  |
| Bananas                   | 105 United Artists   | 1971-07-13 Comedy   |
| Yojimbo                   | 106 Toho             | 1961-06-16 Drama    |
| There's a Girl in my Soup | 107 Columbia         | 1970-06-11 Comedy   |
| Taxi Driver               | 107 Columbia         | 1975-05-15 Action   |
| Absence of Malice         | 107 Columbia         | 1981-11-15 Action   |
| Storia di una donna       | 108 Westward         | 1970-08-15 Romantic |
| The King and I            | 109 20th Century Fox | 1956-08-11 Musical  |
| Das Boot                  | 110 Bavaria Atelier  | 1981-11-11 Drama    |
| Bed Knobs and Broomsticks | 111 Walt Disney      | Musical             |

对所有影片的 len 列求和并按 kind 分组:

```
SELECT kind, SUM(len) AS total FROM films GROUP BY kind;
```

| kind        | total |
|-------------|-------|
| -----+----- |       |
| Action      | 07:34 |
| Comedy      | 02:58 |
| Drama       | 14:28 |
| Musical     | 06:42 |
| Romantic    | 04:38 |

对所有影片的 len 列求和, 按 kind 分组, 并显示少于 5 小时的组总计:

```
SELECT kind, SUM(len) AS total
FROM films
GROUP BY kind
HAVING SUM(len) < INTERVAL '5 hour';
```

| kind        | total |
|-------------|-------|
| -----+----- |       |
| Comedy      | 02:58 |
| Romantic    | 04:38 |

下面两个例子是按照第二列 (name) 的内容对各个结果排序的相同方式:

```
SELECT * FROM distributors ORDER BY name;
SELECT * FROM distributors ORDER BY 2;
```

| did  | name             |
|------|------------------|
| ---- | -----            |
| 109  | 20th Century Fox |
| 110  | Bavaria Atelier  |
| 101  | British Lion     |
| 107  | Columbia         |
| 102  | Jean Luc Godard  |
| 113  | Luso films       |

```

104|Mosfilm
103|Paramount
106|Toho
105|United Artists
111|Walt Disney
112|Warner Bros.
108|Westward

```

这个例子展示如何获得表 `distributors` 和 `actors` 的并集，把结果限制为每个表中以字母 W 开头的行。只使用不重复的行，所以省略了 `ALL` 关键字：

```

--          distributors:          actors:
--          did|name                id|name
--          ---+-----            ---+-----
--          108|Westward            1|Woody Allen
--          111|Walt Disney         2|Warren Beatty
--          112|Warner Bros.        3|Walter Matthau
--          ...                    ...

SELECT distributors.name
   FROM distributors
  WHERE distributors.name LIKE 'W%'
UNION
SELECT actors.name
   FROM actors
  WHERE actors.name LIKE 'W%'

name
-----
Walt Disney
Walter Matthau
Warner Bros.
Warren Beatty
Westward
Woody Allen

```

## 兼容性

## SQL92

### SELECT 子句

在 SQL92 标准中，可选的关键字 `AS` 只是无意义的噪音，可以省略而不影响含义。Postgres 解析器在重命名列时要求这个关键字，因为类型可扩展特性在此上下文中会导致解析歧义。

在 SQL92 标准中，`AS` 子句中指定的新列名可以在 `GROUP BY` 和 `HAVING` 子句中引用。目前在 Postgres 中不允许这样做。

### UNION 子句

SQL92 的 `UNION` 语法允许一个附加的 `CORRESPONDING BY` 子句：

```
table_query UNION [ALL]  
  [CORRESPONDING [BY (column [, ...])]]  
  table_query
```

Postgres 不支持 CORRESPONDING BY 子句。

# SELECT INTO

SELECT INTO — 从一个现有表或视图创建一个新表

## 大纲

```
SELECT [ ALL | DISTINCT ] expression [ AS name ] [, ...]  
    INTO [ TABLE ] new_table [  
    [ FROM table [alias] [, ...] ]  
    [ WHERE condition ]  
    [ GROUP BY column [, ...] ]  
    [ HAVING condition [, ...] ]  
    [ UNION [ ALL ] select]  
    [ ORDER BY column [ ASC | DESC ] [, ...] ]
```

## 输入

所有输入字段在 SELECT 中有详细描述。

## 输出

所有输出字段在 SELECT 中有详细描述。

## 描述

SELECT INTO 从查询的结果创建一个新表。通常，这个查询从现有的表中提取数据，但允许任何 SQL 查询。

### 注意

CREATE TABLE AS 在功能上等价于 SELECT INTO 命令。

# SET

SET — 设置会话的运行时参数

## 大纲

```
SET variable { TO | = } { 'value' | DEFAULT }  
SET TIME ZONE { 'timezone' | LOCAL };
```

## 输入

*variable*

可设置的全局参数。

*value*

参数的新值。

可能的变量和允许的值有：

DateStyle

ISO

使用 ISO 8601 风格的日期和时间

SQL

使用 Oracle/Ingres 风格的日期和时间

Postgres

使用传统的 Postgres 格式

European

数值日期表示使用 dd/mm/yyyy。

NonEuropean

数值日期表示使用 mm/dd/yyyy。

German

数值日期表示使用 dd.mm.yyyy。

US

与 'NonEuropean' 相同

default

恢复默认值 ('US,Postgres')

日期格式的初始化可以通过以下方式完成：

设置 PGDATESTYLE 环境变量。

使用 -oe 参数运行 postmaster，把日期设为 'European' 约定。注意这只影响日期风格的一些组合；例如 ISO 风格不受这个参数影响。

更改 src/backend/utils/init/globals.c 中的变量。

globals.c 中可以更改的变量有：

```
bool EuroDates = false
                true
int   DateStyle = USE_ISO_DATES
                USE_POSTGRES_DATES
                USE_ISO_DATES
                USE_SQL_DATES
                USE_GERMAN_DATES
```

## TIMEZONE

timezone 的可用值取决于你的操作系统。例如在 Linux 上，/usr/lib/zoneinfo 包含时区数据库。

下面是一些有效的 timezone 值：

'PST8PDT'

把时区设为加利福尼亚州

'Portugal'

把时区设为葡萄牙。

'Europe/Rome'

把时区设为意大利。

## DEFAULT

把时区设为你的本地时区（TZ 环境变量的值）。

如果指定了无效的时区，时区会变成 GMT（至少在大多数系统上如此）。

使用 libpq 的前端可以通过设置 PGTZ 环境变量来初始化。

上面所示的第二种语法允许用与 SQL92 SET TIME ZONE 类似的语法设置时区。LOCAL 关键字只是为兼容 SQL92 而提供的 DEFAULT 的另一种形式。

还有一些内部参数或优化参数也可以由 SET 命令指定：

## COST\_HEAP

设置优化器使用的堆扫描默认代价。

*float4*

把堆扫描的代价设为指定的浮点值。

## DEFAULT

把堆扫描的代价设为默认值。

前端可以通过设置 PG\_COST\_HEAP 环境变量来初始化。

## COST\_INDEX

设置优化器使用的索引扫描默认代价。

*float4*

把索引扫描的代价设为指定的浮点值。

## DEFAULT

把索引扫描的代价设为默认值。

前端可以通过设置 PG\_COST\_INDEX 环境变量来初始化。

## GEQO

设置使用遗传优化器算法的阈值。

On

对涉及 8 个或更多表的语句启用遗传优化器算法。

On=#

接受一个整数参数，对查询中涉及 # 个或更多表的语句启用遗传优化器算法。

Off

禁用遗传优化器算法。

## DEFAULT

等价于指定 `SET GEQO='on'`

这个算法默认开启，即对涉及八个或更多表的语句使用 GEQO。（更多信息见程序员指南中的 GEQO 一章。）

前端可以通过设置 PGGEQO 环境变量来初始化。

## R\_PLANS

决定是否允许右手计划求值：

On

启用计划的右手求值。

Off

禁用计划的右手求值。

## DEFAULT

等价于指定 `SET R_PLANS='off'`。

在大关系与小关系连接时可能有用。这个算法默认关闭。GEQO 也不会使用它。

前端可以通过设置 PGRPLANS 环境变量来初始化。

## KSQO

键集查询优化器强制查询优化器优化诸如 Microsoft Access 生成的重复 OR 子句：

On

启用这种优化。

Off

禁用这种优化。

DEFAULT

等价于指定 `SET KSQO='off'`。

在大关系与小关系连接时可能有用。这个算法默认关闭。GEQO 也不会使用它。

前端可以通过设置 PGRPLANS 环境变量来初始化。

## QUERY\_LIMIT

设置查询返回的行数。

Value

查询返回的最大行数。默认允许不限数量的行。

#

把查询返回的最大行数设为 #。

DEFAULT

把查询返回的最大行数设为不限。

默认情况下，查询返回的行数没有限制。

## 输出

→ SET VARIABLE

成功时返回的消息。

→ WARN: Bad value for variable (value)

如果命令未能设置变量。

## 描述

SET 会在会话期间修改变量的配置参数。

当前值可以用 SHOW 获得，值可以用 RESET 恢复为默认值。参数和值不区分大小写。注意值字段总是以字符串形式指定，因此用单引号括起。

SET TIME ZONE 更改会话的默认时区偏移。SQL 会话总是以一个初始的默认时区偏移开始。SET TIME ZONE 语句用于更改当前 SQL 会话的默认时区偏移。

## 注解

`SET variable` 语句是一种 Postgres 语言扩展。

请参阅 `SHOW` 和 `RESET` 来显示或重置当前值。

## 用法

```
--Set the style of date to ISO:
--
SET DATESTYLE TO 'ISO';

--Set GEQO to default:
--
SET GEQO = DEFAULT;

--Turn on right-hand evaluation of plans:
--
SET R_PLANS TO 'on';

--set the timezone for Berkeley, California:
SET TIME ZONE 'PST8PDT';

SELECT CURRENT_TIMESTAMP AS today;

      today
-----
1998-03-31 07:41:21-08

--set the timezone for Italy:
SET TIME ZONE 'Europe/Rome';

SELECT CURRENT_TIMESTAMP AS today;

      today
-----
1998-03-31 17:41:31+02
```

## 兼容性

### SQL92

SQL92 中没有 `SET variable`。SQL92 的 `SET TIME ZONE` 语法稍有不同，只允许用单个整数值指定时区：

```
SET TIME ZONE { interval_value_expression | LOCAL }
```

# SHOW

SHOW — 显示会话的运行时参数

## 大纲

```
SHOW variable
```

## 输入

```
variable
```

可用的变量更多信息请参考 SET。

## 输出

```
→ NOTICE: variable is value SHOW VARIABLE
```

成功时返回的消息。

```
→ NOTICE: Unrecognized variable value
```

如果 → *value* 不存在则返回此消息。

```
NOTICE: Time zone is unknown SHOW VARIABLE
```

如果没有设置 TZ 环境变量。

## 描述

SHOW 将显示会话期间变量的当前配置参数。

会话可以用 SET 语句配置，值可以用 RESET 语句恢复为默认值。参数和值不区分大小写。

## 注解

SHOW 是一种 Postgres 语言扩展。

设置/重置变量的值请参考 SET/RESET。另见 SET TIME ZONE。

## 用法

```
-- show DateStyle;
SHOW DateStyle;
NOTICE:DateStyle is Postgres with US (NonEuropean) conventions

-- show Geco;
SHOW GEQO;
NOTICE:GEQO is ON
```

## 兼容性

## SQL92

SHOW 在 SQL92 中没有定义。

# UNLISTEN

UNLISTEN — 停止监听通知

## 大纲

```
UNLISTEN { notifyname | * }
```

## 输入

*notifyname*

先前注册的通知条件的名称。

\*

清除该后端当前的所有监听注册。

## 输出

→ UNLISTEN

语句已执行的确认。

## 描述

UNLISTEN 用于移除一个现有的NOTIFY注册。UNLISTEN 取消当前 Postgres会话作为通知条件 *notifyname* 监听者的任何现有注册。特殊的条件通配符"\*"取消当前会话的所有监听者注册。

NOTIFY中包含关于LISTEN和 NOTIFY用法的更广泛的讨论。

## 注解

*classname* 不需要是有效的类名，而可以是任何可作为名称使用的、长度不超过 32 个字符的字符串。

后端不会因为你 UNLISTEN 一个你并没有在监听的东西而抱怨。每个后端退出时都会自动执行UNLISTEN \*。

Postgres一些早期版本中，不对应实际表的 *classname*必须用双引号括起来的限制已不存在。

## 用法

```
postgres=> LISTEN virtual;
LISTEN
postgres=> NOTIFY virtual;
NOTIFY
ASYNC NOTIFY of 'virtual' from backend pid '12317' received

postgres=> UNLISTEN virtual;
UNLISTEN
postgres=> NOTIFY virtual;
```

```
NOTIFY
-- notice no NOTIFY event is received
postgres=>
```

## 兼容性

### SQL92

UNLISTEN 在 SQL92 中没有。

# UPDATE

UPDATE — 替换表中各列的值

## 大纲

```
UPDATE table SET column = expression [, ...]
    [ FROM fromlist ]
    [ WHERE condition ]
```

## 输入

*table*

一个现有表的名称。

*column*

*table* 中一个列的名称。

*expression*

赋给列的一个有效表达式或值。

*fromlist*

一个 Postgres 非标准扩展，允许来自其他表的列出现在 WHERE 条件中。

*condition*

WHERE 子句的进一步描述请参阅 SELECT 语句。

## 输出

UPDATE #

成功时返回的消息。# 表示被更新的行数。如果 # 等于 0，则没有行被更新。

## 描述

UPDATE 更改所有满足条件的行中指定列的值。只有要修改的列才需要作为列出现。

数组引用使用与 SELECT 中相同的语法。也就是说，可以用单次查询替换单个数组元素、数组元素的一个范围或整个数组。

要修改表，你必须拥有对它的写权限，以及对 WHERE 条件中提到其值的任何表的读权限。

## 用法

```
--Change word "Drama" with "Dramatic" on column kind:
--
UPDATE films
    SET kind = 'Dramatic'
```

```
WHERE kind = 'Drama';

SELECT * FROM films WHERE kind = 'Dramatic' OR kind = 'Drama';
```

| code  | title         | did | date_prod  | kind     | len   |
|-------|---------------|-----|------------|----------|-------|
| BL101 | The Third Man | 101 | 1949-12-23 | Dramatic | 01:44 |
| P_302 | Becket        | 103 | 1964-02-03 | Dramatic | 02:28 |
| M_401 | War and Peace | 104 | 1967-02-12 | Dramatic | 05:57 |
| T_601 | Yojimbo       | 106 | 1961-06-16 | Dramatic | 01:50 |
| DA101 | Das Boot      | 110 | 1981-11-11 | Dramatic | 02:29 |

## 兼容性

### SQL92

SQL92 为定位式 UPDATE 语句定义了一种不同的语法：

```
UPDATE table SET column = expression [, ...]
WHERE CURRENT OF cursor
```

其中 *cursor* 标识一个打开的游标。

# VACUUM

VACUUM — 清理并分析一个 Postgres 数据库

## 大纲

```
VACUUM [ VERBOSE ] [ ANALYZE ] [ table ]
VACUUM [ VERBOSE ] ANALYZE [ table [ (column [, ...] ) ] ]
```

## 输入

VERBOSE

为每个表打印一份详细的清理活动报告。

ANALYZE

更新优化器用来确定执行查询最有效方式所使用的列统计信息。这些统计信息表示每列数据的 disbursion（离散度）。当存在多条可能的执行路径时，这些信息很有价值。

*table*

要清理的指定表的名称。默认为所有表。

*column*

要分析的指定列的名称。默认为所有列。

## 输出

→ VACUUM

命令已被接受，数据库正在被清理。

NOTICE: --Relation *table*--

*table*的报告头。

NOTICE: Pages 98: Changed 25, Reapped 74, Empty 0, New 0; Tup 1000: Vac 3000, Crash 0, UnUsed 0, MinLen 188, MaxLen 188; Re-using: Free/Avail. Space 586952/586952; EndEmpty/Avail. Pages 0/74. Elapsed 0/0 sec.

对*table*本身的分析。

NOTICE: Index *index*: Pages 28; Tuples 1000: Deleted 3000. Elapsed 0/0 sec.

对目标表上某个索引的分析。

## 描述

VACUUM在 Postgres 中有两个目的：既是回收存储空间的手段，也是为优化器收集信息的手段。

VACUUM打开数据库中的每个类，清除回滚事务留下的记录，并更新系统目录中的统计信息。所维护的统计信息包括所有类中存储的元组数和页数。周期性地运行VACUUM将提高数据库处理用户查询的速度。

## 注意

打开的数据库是 VACUUM 的目标。

我们建议对活跃的生产数据库每晚清理一次，以保持统计信息相对最新。不过 VACUUM 查询可以在任何时候执行。特别是在把一个大类复制进 Postgres 或删除大量记录之后，发出一条 VACUUM 查询可能是个好主意。这会用所有最近变更的结果更新系统目录，使 Postgres 查询优化器在规划用户查询时做出更好的选择。

如果在 VACUUM 命令执行期间服务器崩溃，很可能留下一个悬挂的锁文件。重新运行 VACUUM 命令会导致关于创建锁文件的错误消息。如果你确定 VACUUM 没有在运行，删除你数据库目录中的 pg\_vlock 文件（即 PGDATA/base/dbname/pg\_vlock）。

## 用法

下面是对回归数据库中的一个表运行 VACUUM 的例子：

```
regression=> vacuum verbose analyze onek;
NOTICE:  --Relation onek--
NOTICE:  Pages 98: Changed 25, Reapped 74, Empty 0, New 0;
          Tup 1000: Vac 3000, Crash 0, Unused 0, MinLen 188, MaxLen
          188;
          Re-using: Free/Avail. Space 586952/586952; EndEmpty/Avail.
          Pages 0/74.
          Elapsed 0/0 sec.
NOTICE:  Index onek_stringu1: Pages 28; Tuples 1000: Deleted 3000.
          Elapsed 0/0 sec.
NOTICE:  Index onek_hundred: Pages 12; Tuples 1000: Deleted 3000.
          Elapsed 0/0 sec.
NOTICE:  Index onek_unique2: Pages 19; Tuples 1000: Deleted 3000.
          Elapsed 0/0 sec.
NOTICE:  Index onek_unique1: Pages 17; Tuples 1000: Deleted 3000.
          Elapsed 0/0 sec.
NOTICE:  Rel onek: Pages: 98 --> 25; Tuple(s) moved: 1000. Elapsed 0/
1 sec.
NOTICE:  Index onek_stringu1: Pages 28; Tuples 1000: Deleted 1000.
          Elapsed 0/0 sec.
NOTICE:  Index onek_hundred: Pages 12; Tuples 1000: Deleted 1000.
          Elapsed 0/0 sec.
NOTICE:  Index onek_unique2: Pages 19; Tuples 1000: Deleted 1000.
          Elapsed 0/0 sec.
NOTICE:  Index onek_unique1: Pages 17; Tuples 1000: Deleted 1000.
          Elapsed 0/0 sec.
VACUUM
```

## 兼容性

### SQL92

SQL92 中没有 VACUUM 语句。

---

## 第 20 章 工具应用程序

这里是 Postgres 支持工具的参考信息。

# createdb

createdb — 创建一个新的 Postgres 数据库

## 大纲

```
createdb [ dbname ]  
createdb [ -h host ] [ -p port ]  
        [ -D datadir ]  
        [ -u ] [ dbname ]
```

## 输入

**-h** *host*

指定 postmaster 所在机器的主机名。默认使用本地 Unix 域套接字而不是 IP 连接。

**-p** *port*

指定 postmaster 正在监听连接的 Internet TCP/IP 端口或本地 Unix 域套接字文件扩展名。端口号默认为 5432，或 PGPORT 环境变量的值（如果设置了的话）。

**-u**

使用口令认证。提示输入 *username* 和 *password*。

**-D** *datadir*

指定本次数据库安装的备选数据库位置。这是安装系统表的位置，而不是这个具体数据库的位置，后者可能不同。

*dbname*

指定要创建的数据库的名称。该名称在本安装中的所有 Postgres 数据库中必须唯一。*dbname* 默认为 USER 环境变量的值。

## 输出

createdb 将为新数据库在 PGDATA/*dbname*/ 数据区中创建文件。

Connection to database 'template1' failed. connectDB() failed: Is the postmaster running and accepting connections at 'UNIX Socket' on port '*port*'? createdb: database creation failed on *dbname*.

createdb 无法连接到指定主机和端口上的 postmaster 进程。如果看到这条消息，请确保 postmaster 在正确的主机上运行，并且你指定了正确的端口。如果你的站点使用认证系统，请确保你已获得所需的认证凭据。

Connection to database 'template1' failed. FATAL 1: SetUserId: user '*username*' is not in 'pg\_shadow' createdb: database creation failed on *dbname*.

你在关系 pg\_shadow 中没有有效的条目，将不被允许访问 Postgres。请联系你的 Postgres 管理员。

ERROR: user '*username*' is not allowed to create/destroy databases createdb: database creation failed on *dbname*.

你没有创建新数据库的权限。请联系你的 Postgres 站点管理员。

ERROR: createdb: database '*dbname*' already exists. createdb: database creation failed on *dbname*.

数据库已存在。

createdb: database creation failed on *dbname*.

psql 或后端服务器发生了内部错误。请确保你的站点管理员已正确安装 Postgres 并用 initdb 初始化了站点。

### 注意

createdb 在内部从 psql 运行 CREATE DATABASE, 同时连接到 `template1` 数据库。

## 描述

createdb 创建一个新的 Postgres 数据库。执行这个命令的人成为该数据库的数据库管理员 (DBA), 并且是 Postgres 超级用户之外唯一可以销毁它的人。

createdb 是一个调用 psql 的 shell 脚本。因此, 在执行 createdb 之前, 数据库服务器主机上必须有一个 postmaster 进程在运行。PGOPTION 和 PGREALM 环境变量将被传递给 psql 并按 psql 中的描述处理。

## 用法

要用本地主机 5432 端口上的 postmaster 创建数据库 demo:

```
createdb demo
```

要用主机 eden 上 5000 端口上的 postmaster 创建数据库 demo:

```
createdb -p 5000 -h eden demo
```

# createuser

createuser — 创建一个新的 Postgres 用户

## 大纲

```
createuser [ username ]  
createuser [ -h host ] [ -p port ]  
           [ -i userid ] [ -d | -D ] [ -u | -U ]  
           [ username ]
```

## 输入

**-h** *host*

指定运行 postmaster 的主机名。默认使用本地 unix 域套接字而不是 IP 连接..

**-p** *port*

指定 postmaster 正在监听连接的 Internet TCP/IP 端口或本地 unix 域套接字文件扩展名。端口号默认为 5432，或 PGPORT 环境变量的值（如果已设置）。

**-d**

允许该用户创建数据库。

**-D**

禁止该用户创建数据库。

**-i** *userid*

指定要与此用户关联的数字标识符。这个标识符必须在所有 Postgres用户之间唯一，并且不需要与操作系统的 uid 一致。如果命令行上没有指定标识符，系统会提示你输入，并且它会建议一个与 uid 匹配的标识符。

**-u**

允许该用户创建其他用户。

**-U**

禁止该用户创建其他用户。

*username*

指定要创建的 Postgres 用户的名称。这个名称必须在所有 Postgres用户之间唯一。如果命令行上没有指定名称，系统会提示你输入。

## 输出

createuser将在 pg\_user或pg\_shadow系统表中添加一项。

Connection to database 'template1' failed. connectDB() failed: Is the postmaster running and accepting connections at 'UNIX SOCKET' on port '*port*'? createuser: database access failed.

createuser无法在指定的主机和端口上连接到 postmaster 进程。如果你看到这条消息，请确保 postmaster 在正确的主机上运行，并且你指定了正确的端口。如果你的站点使用认证系统，请确保你已获得所需的认证凭据。

Connection to database 'template1' failed. FATAL 1: SetUserId: user '*username*' is not in 'pg\_shadow' createuser: database access failed.

你在关系 pg\_shadow中没有有效的项，将不被允许访问 Postgres。联系你的 Postgres管理员。

createuser: *username* cannot create users.

你没有创建新用户的权限；联系你的 Postgres站点管理员。

createuser: user "*username*" already exists

要添加的用户已经在 pg\_shadow类中有一项。

database access failed

psql或后端服务器中发生了内部错误。请确保你的站点管理员已正确安装 Postgres并用 initdb初始化了站点。

### 注意

createuser在内部连接到 template1数据库并从 psql运行 CREATE USER。

## 描述

createuser创建一个新的 Postgres用户。只有 pg\_shadow类中usesuper被设置的用户才能创建新的 Postgres用户。就发行版本而言，用户 postgres可以创建用户。

createuser是一个调用 psql的 shell 脚本。因此，在执行 createuser之前，数据库服务器主机上必须有一个 postmaster 进程在运行。PGOPTION 和 PGREALM 环境变量将被传递给 psql，并按 psql中所述处理。一旦被调用，createuser 会提出一系列问题来获取命令行上未指定的参数。必须指定新用户的数据库登录名和数字用户标识符。

### 注意

Postgres用户标识符不需要与该用户的 Unix UID 相同。不过，通常它们会被指定为相同。

出于安全目的，你还必须描述新用户的权限。具体来说，系统会询问新用户是否应能充当 Postgres超级用户、新用户是否可以创建新数据库、以及新用户是否被允许创建其他新用户。

# destroydb

destroydb — 创建一个新的Postgres数据库

## 大纲

```
destroydb [ dbname ]
destroydb [ -h host ] [ -p port ]
          [ -i ] [ dbname ]
```

## 输入

**-h** *host*

指定运行 postmaster 的主机名。默认使用本地 Unix 域套接字而非 IP 连接。

**-p** *port*

指定 postmaster 监听连接所用的 Internet TCP/IP 端口或本地 Unix 域套接字文件扩展名。端口号默认为 5432，或为 PGPORT 环境变量的值（如果设置了该变量）。

**-i**

以交互模式运行。在删除数据库之前提示确认。

*dbname*

指定要删除的数据库的名称。该数据库必须是本安装中现有的 Postgres 数据库之一。*dbname* 默认为 USER 环境变量的值。

## 输出

destroydb 将从现有数据库的 PGDATA/*dbname*/ 数据区中移除文件。

Connection to database 'template1' failed. connectDB() failed: Is the postmaster running and accepting connections at 'UNIX Socket' on port '*port*'? destroydb: database destroy failed on *dbname*.

destroydb 无法连接到指定主机和端口上的 postmaster 进程。如果你看到这条消息，请确认 postmaster 正在正确的主机上运行，并且你指定了正确的端口。如果你的站点使用了认证系统，请确认你已经获得了所需的认证凭据。

Connection to database 'template1' failed. FATAL 1: SetUserId: user '*username*' is not in 'pg\_shadow' destroydb: database destroy failed on *dbname*.

你在 pg\_shadow 关系中没有有效的条目，将不被允许访问 Postgres。请联系你的 Postgres 管理员。

ERROR: user '*username*' is not allowed to create/destroy databases destroydb: database destroy failed on *dbname*.

你没有删除（或创建）数据库的权限。请联系你的 Postgres 站点管理员。

ERROR: destroydb: database '*dbname*' does not exist. destroydb: database destroy failed on *dbname*.

要删除的数据库在 pg\_database 类中没有条目。

ERROR: destroydb: database '*dbname*' is not owned by you. destroydb: database destroy failed on *dbname*.

你不是指定数据库的数据库管理员（DBA）。

destroydb: database destroy failed on *dbname*.

psql 或后端服务器中发生了内部错误。请确保你的站点管理员已正确安装 Postgres 并使用 initdb 初始化了站点。

### 注意

destroydb 内部在连接到 `template1` 数据库后通过 psql 运行 `DESTROY DATABASE`。

## 描述

destroydb 删除一个现有的 Postgres 数据库。执行此命令的人必须是数据库管理员（DBA），或者必须是 Postgres 超级用户。该程序静默运行；不会显示任何确认消息。数据库删除完毕后，将重新出现 Unix shell 提示符。

对该数据库的所有引用都将被移除，包括包含该数据库及其相关文件的目录。

destroydb 是一个调用 psql 的 shell 脚本。因此，在执行 destroydb 之前，数据库服务器主机上必须有一个 postmaster 进程正在运行。PGOPTION 和 PGREALM 环境变量将被传递给 psql，并按照 psql 中的描述进行处理。

## 用法

要使用本地主机上端口为 5432 的 postmaster 删除数据库 demo：

```
destroydb demo
```

要使用主机 eden 上端口为 5000 的 postmaster 删除数据库 demo：

```
destroydb -p 5000 -h eden demo
```

# destroyuser

destroyuser — 销毁一个 Postgres 用户及其关联的数据库

## 大纲

```
destroyuser [ username ]
destroyuser [ -h host ] [ -p port ]
           [ username ]
```

## 输入

**-h** *host*

指定 postmaster 正在运行的机器的主机名。默认使用本地 Unix 域套接字而不是 IP 连接。

**-p** *port*

指定 postmaster 正在监听连接的 Internet TCP/IP 端口或本地 Unix 域套接字文件扩展。端口号默认为 5432，或 PGPORT 环境变量的值（如果设置了）。

*username*

指定要移除的 Postgres 用户的名称。该名称必须存在于 Postgres 安装中。如果命令行上没有指定名称，将会提示你输入一个。

## 输出

destroyuser 将移除 pg\_user 或 pg\_shadow 系统表中的一个项，并且将移除该用户作为管理员（DBA）的所有数据库。

Connection to database 'template1' failed. connectDB() failed: Is the postmaster running and accepting connections at 'UNIX Socket' on port '*port*'? destroyuser: database access failed.

destroyuser 无法连接到指定主机和端口上的 postmaster 进程。如果你看到这条消息，请确保 postmaster 正在正确的主机上运行，并且你指定了正确的端口。如果你的站点使用认证系统，请确保你已经获得了所需的认证凭据。

Connection to database 'template1' failed. FATAL 1: SetUserId: user '*username*' is not in 'pg\_shadow' destroyuser: database access failed.

你在关系 pg\_shadow 中没有有效的项，因而不允许访问 Postgres。请联系你的 Postgres 管理员。

destroyuser: *username* cannot delete users.

你没有删除用户的权限；请联系你的 Postgres 站点管理员。

destroyuser: user "*username*" already exists

要添加的用户已经在 pg\_shadow 类中有一个项。

database access failed

psql 或后端服务器中发生了内部错误。请确保你的站点管理员已经正确安装 Postgres 并已用 initdb 初始化了站点。

destroydb on *dbname* failed - exiting

psql 或后端服务器中发生了内部错误。所指定的数据库可能存在 Unix 权限问题。

delete of user *username* was UNSUCCESSFUL

psql 或后端服务器中发生了内部错误。

### 注意

destroyuser 内部从 psql 运行 DROP USER，同时连接到 `template1` 数据库。

## 描述

destroyuser 移除一个现有的 Postgres 用户以及该用户作为数据库管理员的数据库。只有在 `pg_shadow` 类中设置了 `usesuper` 的用户才能销毁 Postgres 用户。按发行时的状态，用户 `postgres` 可以移除用户。

destroyuser 是一个调用 psql 的 shell 脚本。因此，在执行 destroyuser 之前，数据库服务器主机上必须有一个 `postmaster` 进程在运行。`PGOPTION` 和 `PGREALM` 环境变量将被传递给 psql，并按 psql 中的描述处理。

一旦被调用，destroyuser 会警告你哪些数据库将在过程中被销毁，并允许你在需要时中止对该用户的移除。

# initdb

initdb — 创建一个新的 Postgres 数据库安装

## 大纲

```
initdb [ --pgdata=dbdir | -r dbdir ]
      [ --pglib=libdir | -l libdir ]
      [ --template=template | -t template ]
      [ --username=name | -u name ]
      [ --noclean | -n ] [ --debug | -d ]
```

## 输入

```
--pglib=libdir
-l libdir
PGLIB
```

构成 Postgres 的文件在哪里？除了那些因功能原因必须放在特定目录中的文件外，构成 Postgres 软件的文件都安装在名为 *libdir* 的目录中。initdb 需要的文件中有一个可以在那里找到的例子是 `global1.bki.source`，它包含进入共享目录表的全部信息。

```
--pgdata=dbdir
-r dbdir
PGDATA
```

你希望数据库数据放在 Unix 文件系统的什么位置？顶层目录称为 PGDATA 目录。

```
--username=name
-u name
PGUSER
```

谁将成为这个数据库系统的 Postgres 超级用户？Postgres 超级用户是一个 Unix 用户，他拥有存储数据库系统的所有文件，也拥有访问这些文件的 postmaster 和后端进程。或者就让它默认为你（运行 initdb 的 Unix 用户）。

### 注意

只有 Unix 超级用户（root）才能创建拥有者不同于 Postgres 超级用户的数据库系统。

还有其他一些较少使用的参数：

```
--template=template
-t template
```

替换现有数据库系统中的 `template1` 数据库，其他什么都不动。当你需要用来自更新版本 Postgres 的 initdb 升级你的 `template1` 数据库时，或者当你的 `template1` 数据库因某些系统问题而损坏时，这很有用。通常，`template1` 的内容在数据库系统的整个生命期内保持不变。用 initdb 加 `--template` 选项运行不会破坏任何东西。

`--noclean`  
`-n`

默认情况下，当 `initdb` 判定某个错误使它无法完全创建数据库系统时，它会删除在判定无法完成工作之前可能已创建的所有文件。其中包括它调用的程序留下的任何 `core` 文件。这个选项抑制任何清理，因此对调试有用。

`--debug`  
`-d`

打印引导后端的调试输出。引导后端是 `initdb` 用来创建目录表的程序。这个选项会产生大量的输出。它还会关闭最终的 `vacuum` 步骤。

文件也是 `initdb` 的输入：

`postconfig`

如果出现在 Unix 命令搜索路径（由 `PATH` 环境变量定义）中的话。这是一个为某些命令选项指定默认值的程序。见下文。

`PGLIB/global1.bki.source`

新数据库系统中共享目录表的内容。这个文件是 Postgres 软件的一部分。

`PGLIB/local1_template1.bki.source`

新数据库系统中 `template1` 表的内容。这个文件是 Postgres 软件的一部分。

## 输出

`initdb` 会在 `PGDATA` 数据区域中创建作为完整安装的系统表和框架的文件。

## 描述

`initdb` 创建一个新的 Postgres 数据库系统。数据库系统是由同一个 Unix 用户管理、并由单个 `postmaster` 管理的一组数据库的集合。

创建数据库系统包括：创建存放数据库数据的目录，生成共享目录表（不属于任何特定数据库的表），以及创建 `template1` 数据库。`template1` 数据库是什么？当你创建数据库时，Postgres 是通过复制 `template1` 数据库中的所有内容来完成的。其中包含为内建类型等内容填好的目录表。

在 `initdb` 创建数据库之后，它会运行 `vacuum` 来完成初始化，这会重置一些优化参数。

有三种方式向 `initdb` 传递参数。第一，可以使用 `initdb` 命令选项。第二，可以在调用 `initdb` 之前设置环境变量。第三，可以在 Unix 命令搜索路径中放一个名为 `postconfig` 的程序。`initdb` 会调用该程序，该程序随后把 `initdb` 参数写到它的标准输出流。不过第三种方式并不常见。

命令选项总是覆盖以其他任何方式指定的参数。`postconfig` 返回的值覆盖任何环境变量，但如果你希望使用环境变量的值，你的 `postconfig` 程序可以让它的输出基于环境变量。

`postconfig` 输出的值必须具有格式

```
var1=value1 var2=value2 ...
```

如果它不想提供任何参数，也可以什么都不输出。`var` 的值与相应环境变量名相同。例如，

```
PGDATA=/tmp/postgres_test
```

的效果与以一个名为 PGDATA、值为 /tmp/postgres\_test 的环境变量调用 initdb 相同。

# initlocation

initlocation — 创建一个辅助的 Postgres 数据库存储区

## 大纲

```
initlocation [ --location=altdir | -D altdir ]
             [ --username=name | -u name ]
             [ altdir ]
```

## 输入

```
--location=altdir
-D altdir
altdir
```

你想让备选数据库放在你的 Unix 文件系统的什么地方？顶层目录称为 PGDATA 目录，所以你可能想把你的第一个备选位置指向 PGDATA2。

```
--username=name
-u name
PGUSER
```

谁将成为这个数据库存储区的 Unix 文件系统所有者？Postgres 超级用户是一个 Unix 用户，他拥有存储数据库系统的所有文件，并且拥有访问它们的 postmaster 和后端进程。通常，这就是应该运行 initlocation 的用户，他因此将拥有这些目录和文件。

### 注意

只有 Unix 超级用户才能创建一个由不是 Postgres 超级用户的用户拥有的数据库系统。指定一个不是 Postgres 超级用户的用户可能导致数据库安全和数据完整性问题。更多信息请参阅 PostgreSQL 管理员指南。

## 输出

initlocation 将在指定的位置创建目录。

```
We are initializing the database area with username postgres (uid=500). This user will
own all the files and must also own the server process. Creating Postgres database system
directory altdir Creating Postgres database system directory altdir
```

成功完成。

```
We are initializing the database area with username postgres (uid=500). This user will
own all the files and must also own the server process. Creating Postgres database system
directory /usr/local/src/testlocation mkdir: cannot make directory 'altdir': Permission
denied
```

你没有写入指定目录区域的文件系统权限。

Valid username not given. You must specify the username for the Postgres superuser for the database system you are initializing, either with the `--username` option or by default to the `USER` environment variable.

你指定的用户名不是 Postgres 超级用户。

Can't tell what username to use. You don't have the `USER` environment variable set to your username and didn't specify the `--username` option

指定 `--username` 命令行选项。

## 描述

`initlocation` 创建一个新的 Postgres 辅助数据库存储区。一个辅助存储区包含一棵必需的目录树，并且这些目录上有正确的文件权限。

创建一个数据库存储区包括创建数据库数据可能存放的目录。

`initlocation` 有两类参数。首先，你可以指定一个环境变量（例如 `PGDATA2`）。这个环境变量应当为后端所知，供以后在 `CREATE DATABASE/WITH LOCATION` 或 `createdb -D altdir` 中使用。但是，后端守护进程的环境中必须有这个变量，这才能成功。其次，或许可以指定存储区顶层目录的一个显式绝对路径。但是，第二个选项只有在 Postgres 安装期间显式启用时才可用。它通常被禁用，以减轻安全和数据完整性方面的顾虑。

### 注意

Postgres 会把 `/base/` 追加到指定的路径来创建存储区。

后端要求 `WITH LOCATION` 的任何全大写且不含路径分隔符的参数都是环境变量。

## 用法

要在一个备选位置创建数据库，使用环境变量：

```
% setenv PGDATA2 /opt/postgres/data
```

```
% initlocation PGDATA2
```

```
% createdb -D PGDATA2
```

# pg\_dump

pg\_dump — Extract a Postgres database into a script file

## 大纲

```
pg_dump [ dbname ]
pg_dump [ -h host ] [ -p port ]
      [ -t table ]
      [ -f outputfile ]
      [ -a ] [ -d ] [ -D ] [ -o ] [ -s ] [ -u ] [ -v ] [ -z ]
      [ dbname ]
```

## 输入

pg\_dump 接受下列命令行参数：

*dbname*

指定要抽取的数据库名称。*dbname* 默认为 `USER` 环境变量的值。

`-a`

只转储数据，不转储模式（定义）。

`-d`

将数据转储为正规的插入字符串。

`-D`

将数据转储为带属性名的插入

`-f filename`

指定输出文件。默认为 `stdout`。

`-n`

除非绝对必要，抑制标识符周围的双引号。如果标识符中使用了保留字，这可能导致装载这份转储数据时出问题。

`-o`

为每个表转储对象标识符（OID）。

`-s`

只转储模式（定义），不转储数据。

`-t table`

只转储 *table* 的数据。

`-u`

使用密码认证。提示输入用户名和密码。

-V

指定详细模式

-Z

包含 ACL (grant/revoke 命令) 和表所有权信息。

pg\_dump 还接受下列用于连接参数的命令行参数：

-h *host*

指定运行 postmaster 的主机名。默认使用本地 Unix 域套接字而不是 IP 连接。。

-p *port*

指定 postmaster 正在监听连接的 Internet TCP/IP 端口，或本地 Unix 域套接字文件扩展名。端口号默认为 5432，或 PGPORT 环境变量的值（如果已设置）。

-u

使用密码认证。提示输入 *username* 和 *password*。

## 输出

pg\_dump 会创建一个文件或写入 stdout。

Connection to database 'template1' failed. connectDB() failed: Is the postmaster running and accepting connections at 'UNIX Socket' on port '*port*'?

pg\_dump 无法连接到指定主机和端口上的 postmaster 进程。如果你看到这条消息，请确保 postmaster 运行在正确的主机上，并且你指定了正确的端口。如果你的站点使用认证系统，请确保你已获得所需的认证凭据。

Connection to database '*dbname*' failed. FATAL 1: SetUserId: user '*username*' is not in 'pg\_shadow'

你在关系 pg\_shadow 中没有有效的条目，将不被允许访问 Postgres。请联系你的 Postgres 管理员。

dumpSequence(*table*): SELECT failed

你没有读取该数据库的权限。请联系你的 Postgres 站点管理员。

### 注意

pg\_dump 内部执行 SELECT 语句。如果在运行 pg\_dump 时遇到问题，请确保你能够使用例如 psql 从数据库中选择信息。

## 描述

pg\_dump 是一个把 Postgres 数据库转储为包含查询命令的脚本文件的工具。脚本文件为文本格式，可用于重建数据库，甚至是在其他机器和其他架构上。pg\_dump 会产生重新生成所有用户

定义类型、函数、表、索引、聚合和操作符所需的查询。此外，所有数据都以文本格式复制出来，因此既可以方便地重新复制进去，也可以导入编辑工具。

pg\_dump 可用于转储数据库的内容，以便从一个 Postgres 安装迁移到另一个安装。运行 pg\_dump 之后，应当检查输出脚本文件中是否有任何警告，尤其是考虑到下面列出的限制。

## 注意

pg\_dump 有一些限制。这些限制大多源于难以从系统目录中抽取某些元信息。

部分索引

pg\_dump 不理解部分索引。原因同上；部分索引谓词是以计划形式存储的。

大对象

pg\_dump 不处理大对象。大对象会被忽略，必须手工处理。

## 用法

要转储一个与用户同名的数据库：

```
% pg_dump > db.out
```

要重新装载这个数据库：

```
psql -e database < db.out
```

# pg\_dumpall

pg\_dumpall — 把所有 Postgres 数据库抽取到一个脚本文件

## 大纲

```
pg_dumpall
pg_dumpall [ -h host ] [ -p port ]
           [ -a ] [ -d ] [ -D ] [ -o ] [ -s ] [ -u ] [ -v ] [ -z ]
```

## 输入

pg\_dumpall 接受下列命令行参数：

-a

只转储数据，不转储模式（定义）。

-d

把数据转储为正确的插入字符串。

-D

把数据转储为带属性名的插入

-n

除非绝对必要，抑制标识符两边的双引号。如果有保留字被用作标识符，加载这些转储数据时可能出问题。

-o

转储每个表的对象标识符（OID）。

-s

只转储模式（定义），不转储数据。

-u

使用口令认证。提示输入用户名和口令。

-v

指定详细模式

-z

包含 ACL（grant/revoke 命令）和表所有权信息。

pg\_dumpall 还接受下列用于连接参数的命令行参数：

-h *host*

指定 postmaster 所在机器的主机名。默认使用本地 Unix 域套接字而不是 IP 连接。

`-p port`

指定 postmaster 正在监听连接的 Internet TCP/IP 端口或本地 Unix 域套接字文件扩展名。端口号默认为 5432，或 PGPORT 环境变量的值（如果设置了的话）。

`-u`

使用口令认证。提示输入 *username* 和 *password*。

## 输出

pg\_dumpall 将创建一个文件或写入 stdout。

Connection to database 'template1' failed. connectDB() failed: Is the postmaster running and accepting connections at 'UNIX Socket' on port '*port*'?

pg\_dumpall 无法连接到指定主机和端口上的 postmaster 进程。如果看到这条消息，请确保 postmaster 在正确的主机上运行，并且你指定了正确的端口。如果你的站点使用认证系统，请确保你已获得所需的认证凭据。

Connection to database '*dbname*' failed. FATAL 1: SetUserId: user '*username*' is not in 'pg\_shadow'

你在关系 pg\_shadow 中没有有效的条目，将不被允许访问 Postgres。请联系你的 Postgres 管理员。

dumpSequence(*table*): SELECT failed

你没有读取数据库的权限。请联系你的 Postgres 站点管理员。

### 注意

pg\_dumpall 在内部执行 SELECT 语句。如果运行 pg\_dumpall 遇到问题，请确保你能用例如 psql 从数据库选择信息。

## 描述

pg\_dumpall 是一个把所有 Postgres 数据库转储到一个文件的实用工具。它还转储对所有数据库全局的 pg\_shadow 表。pg\_dumpall 在这个文件中包含在加载之前自动创建每个被转储数据库的相应命令。

pg\_dumpall 接受所有 pg\_dump 选项，但 `-f`、`-t` 和 *dbname* 应当省略。

关于这一功能的更多信息请参阅 pg\_dump。

## 用法

要转储所有数据库：

```
% pg_dumpall -o > db.out
```

**提示**

可以为 pg\_dumpall 使用大多数 pg\_dump 选项。

要重新加载这个数据库：

```
psql -e template1 < db.out
```

**提示**

重新加载时可以使用大多数 psql 选项。

# psql

psql — Postgres 的交互式客户端

## 大纲

```
psql [ dbname ]
psql -A [ -c query ] [ -d dbname ]
      -e [ -f filename ] [ -F separator ] [ -h hostname ] -Hln
      [ -o filename ] [ -p port ] -qsSt [ -T table_options ] -ux [
      dbname ]
```

## 输入

psql 接受许多命令行参数、一整套丰富的元命令，以及 Postgres 支持的完整 SQL 语言。最常见的命令行参数有：

*dbname*

The name of an existing database to access. *dbname* defaults to the value of the `USER` environment variable or, if that's not set, to the Unix account name of the current user.

`-C query`

要运行的单条查询。完成后 psql 将退出。

完整的命令行参数和元命令集合在后续小节中描述。

There are some environment variables which can be used in lieu of command line arguments. Additionally, the Postgres frontend library used by the psql application looks for other optional environment variables to configure, for example, the style of date/time representation and the local time zone. Refer to the chapter on `libpq` in the Programmer's Guide for more details.

You may set any of the following environment variables to avoid specifying command-line options:

`PGHOST`

The DNS host name of the database server. Setting `PGHOST` to a non-zero-length string causes TCP/IP communication to be used, rather than the default local Unix domain sockets.

`PGPORT`

Postgres 服务器正在监听的端口号。默认为 5432。

`PGTTY`

客户端支持库消息的显示目标。不是必需的。

`PGOPTION`

如果指定了 `PGOPTION`，其中包含的选项会在任何命令行选项之前被解析。

## PGREALM

PGREALM only applies if Kerberos authentication is in use. If this environment variable is set, Postgres will attempt authentication with servers for this realm and will use separate ticket files to avoid conflicts with local ticket files. See the PostgreSQL Administrator's Guide for additional information on Kerberos.

## 输出

psql returns 0 to the shell on successful completion of all queries, 1 for errors, 2 for abrupt disconnection from the backend. The default TAB delimiter is used. psql will also return 1 if the connection to a database could not be made for any reason.

## 描述

psql is a character-based front-end to Postgres. It enables you to type in queries interactively, issue them to Postgres, and see the query results.

psql is a Postgres client application. Hence, a postmaster process must be running on the database server host before psql is executed. In addition, the correct parameters to identify the database server, such as the postmaster host name, may need to be specified as described below.

When psql starts, it reads SQL commands from `/etc/psqlrc` and then from `$(HOME)/.psqlrc`. This allows SQL commands like `SET` which can be used to set the date style to be run at the start of every session.

## 连接到数据库

psql attempts to make a connection to the database at the hostname and port number specified on the command line. If the connection could not be made for any reason (e. g. insufficient privileges, postmaster is not running on the server, etc) .IR psql will return an error that says

```
Connection to database failed.
```

The reason for the connection failure is not provided.

## 输入查询

In normal operation, psql provides a prompt with the name of the database that psql is current connected to followed by the string `"=>"`. For example,

```
$ psql testdb
Welcome to the POSTGRESQL interactive sql monitor:
Please read the file COPYRIGHT for copyright terms of POSTGRESQL

type \e? for help on slash commands
type \eq to quit
type \eg or terminate with semicolon to execute query
You are currently connected to the database: testdb

testdb=>
```

At the prompt, the user may type in SQL queries. Unless the `-S` option is set, input lines are sent to the backend when a query-terminating semicolon is reached.

Whenever a query is executed, `psql` also polls for asynchronous notification events generated by `LISTEN` and `NOTIFY`.

`psql` can be used in a pipe sequence, and automatically detects when it is not listening or talking to a real tty.

## 命令行选项

`psql` understands the following command-line options:

`-A`

Turn off fill justification when printing out table elements.

`-C query`

Specifies that `psql` is to execute one query string, *query*, and then exit. This is useful for shell scripts, typically in conjunction with the `-q` option in shell scripts.

`-d dbname`

Specifies the name of the database to connect to. This is equivalent to specifying *dbname* as the last field in the command line.

`-e`

Echo the query sent to the backend

`-f filename`

Use the file *filename* as the source of queries instead of reading queries interactively. This file must be specified for and visible to the client frontend.

`-F separator`

Use *separator* as the field separator. The default is an ASCII vertical bar ("`|`").

`-h hostname`

Specifies the host name of the machine on which the postmaster is running. Without this option, communication is performed using local Unix domain sockets.

`-H`

Turns on HTML 3.0 tabular output.

`-l`

Lists all available databases, then exit. Other non-connection options are ignored.

`-n`

Do not use the readline library for input line editing and command history.

`-o filename`

Put all output into file *filename*. The path must be writable by the client.

`-p port`

Specifies the TCP/IP port or, by omission, the local Unix domain socket file extension on which the postmaster is listening for connections. Defaults to the value of the `PGPORT` environment variable, if set, or to 5432.

`-q`

Specifies that `psql` should do its work quietly. By default, it prints welcome and exit messages and prompts for each query, and prints out the number of rows returned from a query. If this option is used, none of this happens. This is useful with the `-c` option.

`-s`

Run in single-step mode where the user is prompted for each query before it is sent to the backend.

`-S`

Runs in single-line mode where each query is terminated by a newline, instead of a semicolon.

`-t`

Turn off printing of column names. This is useful with the `-c` option in shell scripts.

`-T table_options`

Allows you to specify options to be placed within the `table ...` tag for HTML 3.0 tabular output. For example, `border` will give you tables with borders. This must be used in conjunction with the `-H` option.

`-u`

Asks the user for the user name and password before connecting to the database. If the database does not require password authentication then these are ignored. If the option is not used (and the `PGPASSWORD` environment variable is not set) and the database requires password authentication, then the connection will fail. The user name is ignored anyway.

`-x`

Turns on extended row format mode. When enabled each row will have its column names printed on the left with the column values printed on the right. This is useful for rows which are otherwise too long to fit into one screen line. HTML row output supports this mode also.

You may set environment variables to avoid typing some of the above options. See the section on environment variables below.

## psql 元命令

Anything you enter in `psql` that begins with an unquoted backslash is a `psql` meta-command. Anything else is SQL and simply goes into the current query buffer (and once you have at least one complete query, it gets automatically submitted to the backend). `psql` meta-commands are also called slash commands.

The format of a psql command is the backslash, followed immediately by a command verb, then any arguments. The arguments are separated from the command verb and each other by any number of white space characters.

With single character command verbs, you don't actually need to separate the command verb from the argument with white space, for historical reasons. You should anyway.

The following meta-commands are defined:

`\a`

Toggle field alignment when printing out table elements.

`\Ccaption`

Set the HTML3.0 table caption to `"caption"`.

`\connectdbname [ username ]`

Establish a connection to a new database, using the default `username` if none is specified. The previous connection is closed.

`\copydbname { FROM | TO } filename`

Perform a frontend (client) copy. This is an operation that runs a SQL COPY command, but instead of the backend reading or writing the specified file, and consequently requiring backend access and special user privilege, psql reads or writes the file and routes the data to or from the backend. The default `tab` delimiter is used.

### 提示

此操作不如 SQL 的 COPY 命令高效，因为所有数据都必须通过客户端/服务器的 IP 或套接字连接。对于大量数据，另一种技术可能更可取。

`\d [ table ]`

List tables in the database, or if `table` is specified, list the columns in that table. If table name is specified as an asterisk ( `"*"` ), list all tables and column information for each tables.

`\da`

List all available aggregates.

`\ddobject`

List the description from `pg_description` of the specified object, which can be a table, table.column, type, operator, or aggregate.

### 提示

并非所有对象都在 `pg_description` 中有描述。此元命令可用于快速了解 Postgres 原生特性的说明。

`\df`

List functions.

`\di`

List only indexes.

`\do`

List only operators.

`\ds`

List only sequences.

`\dS`

List system tables and indexes.

`\dt`

List only non-system tables.

`\dT`

List types.

`\e [ filename ]`

Edit the current query buffer or the contents of the file *filename*.

`\E [ filename ]`

Edit the current query buffer or the contents of the file *filename* and execute it upon editor exit.

`\f [ separator ]`

Set the field separator. Default is a single blank space.

`\g [ { filename | command } ]`

Send the current query input buffer to the backend and optionally save the output in *filename* or pipe the output into a separate Unix shell to execute *command*.

`\h [ command ]`

Give syntax help on the specified SQL command. If *command* is not a defined SQL command (or is not documented in psql), or if *command* is not specified, then psql will list all the commands for which syntax help is available. If *command* is an asterisk ( `"*"` ), then give syntax help on all SQL commands.

`\H`

Toggle HTML3 output. This is equivalent to the `-H` command-line option.

`\i filename`

Read queries from the file *filename* into the query input buffer.

`\l`

List all the databases in the server.

`\m`

Toggle the old monitor-like table display, which includes border characters surrounding the table. This is standard SQL output. By default, psql includes only field separators between columns.

`\o [{ filename} | command]`

Save future query results to the file *filename* or pipe future results into a separate Unix shell to execute *command*. If no arguments are specified, send query results to stdout.

`\p`

Print the current query buffer.

`\q`

Quit the psql program.

`\r`

Reset(clear) the query buffer.

`\s [ filename]`

Print or save the command line history to *filename*. If *filename* is omitted, do not save subsequent commands to a history file. This option is only available if psql is configured to use readline.

`\t`

Toggle display of output column name headings and row count footer (defaults to on).

`\T table_options`

Allows you to specify options to be placed within the `table ...` tag for HTML 3.0 tabular output. For example, `border` will give you tables with borders. This must be used in conjunction with the `\H` meta-command.

`\x`

Toggles extended row format mode. When enabled each row will have its column names printed on the left with the column values printed on the right. This is useful for rows which are otherwise too long to fit into one screen line. HTML row output mode supports this flag too.

`\w filename`

Outputs the current query buffer to the file *filename*.

\z

Produces a list of all tables in the database with their appropriate ACLs (grant/revoke permissions) listed.

\! [ *command* ]

Escape to a separate Unix shell or execute the Unix command *command*.

\?

Get help information about the slash ( "\ " ) commands.

---

## 部分 III. 管理员指南

安装与维护信息。

---

# 目录

|                                 |     |
|---------------------------------|-----|
| 21. 引言 .....                    | 242 |
| 21.1. 资源 .....                  | 242 |
| 21.2. 术语 .....                  | 243 |
| 21.3. 记法 .....                  | 243 |
| 21.4. Y2K 声明 .....              | 244 |
| 21.5. 版权与商标 .....               | 244 |
| 22. 平台移植 .....                  | 246 |
| 22.1. 当前支持的平台 .....             | 246 |
| 22.2. 未支持的平台 .....              | 248 |
| 23. 配置选项 .....                  | 249 |
| 23.1. 配置参数 (configure) .....    | 249 |
| 23.2. 构建参数 (make) .....         | 249 |
| 23.3. 区域支持 .....                | 251 |
| 23.3.1. 有什么好处? .....            | 252 |
| 23.3.2. 有什么缺点? .....            | 252 |
| 23.4. Kerberos 认证 .....         | 252 |
| 23.4.1. 可用性 .....               | 252 |
| 23.4.2. 安装 .....                | 252 |
| 23.4.3. 操作 .....                | 253 |
| 24. 安装 .....                    | 254 |
| 24.1. 运行 Postgres 的要求 .....     | 254 |
| 24.2. 安装过程 .....                | 254 |
| 24.3. 玩转 Postgres .....         | 265 |
| 24.4. 下一步 .....                 | 266 |
| 24.5. 移植说明 .....                | 266 |
| 24.5.1. Ultrix4.x .....         | 266 |
| 24.5.2. Linux .....             | 266 |
| 24.5.3. BSD/OS .....            | 267 |
| 24.5.4. NeXT .....              | 267 |
| 25. 运行时环境 .....                 | 268 |
| 25.1. 在 Unix 上使用 Postgres ..... | 268 |
| 26. 系统布局 .....                  | 269 |
| 27. 使用 pg_options .....         | 271 |
| 28. 启动 postmaster .....         | 276 |
| 29. 添加和删除用户 .....               | 277 |
| 30. 磁盘管理 .....                  | 278 |
| 30.1. 备选位置 .....                | 278 |
| 31. 排障 .....                    | 279 |
| 32. 管理数据库 .....                 | 280 |
| 32.1. 创建数据库 .....               | 280 |
| 32.2. 访问数据库 .....               | 280 |
| 32.3. 删除数据库 .....               | 281 |
| 33. 数据库恢复 .....                 | 282 |
| 34. 回归测试 .....                  | 283 |
| 34.1. 回归测试环境 .....              | 283 |
| 34.2. 目录布局 .....                | 283 |
| 34.3. 回归测试过程 .....              | 284 |
| 34.4. 回归分析 .....                | 285 |
| 34.4.1. 比较期望/实际输出 .....         | 285 |
| 34.4.2. 错误消息差异 .....            | 285 |
| 34.4.3. OID 差异 .....            | 285 |

|                                      |     |
|--------------------------------------|-----|
| 34.4.4. 日期和时间差异 .....                | 286 |
| 34.4.5. 浮点差异 .....                   | 286 |
| 34.4.6. 多边形差异 .....                  | 286 |
| 34.4.7. 随机差异 .....                   | 286 |
| 34.4.8. “expected” 文件 .....          | 286 |
| 35. 发行说明 .....                       | 287 |
| 35.1. 版本 6.4 .....                   | 287 |
| 35.1.1. 迁移到版本 6.4 .....              | 287 |
| 35.1.2. 详细变更列表 .....                 | 287 |
| 35.2. 版本 6.3.2 .....                 | 291 |
| 35.2.1. 详细变更列表 .....                 | 291 |
| 35.3. 版本 6.3.1 .....                 | 291 |
| 35.3.1. 详细变更列表 .....                 | 292 |
| 35.4. 版本 6.3 .....                   | 292 |
| 35.4.1. 迁移到版本 6.3 .....              | 293 |
| 35.4.2. 详细变更列表 .....                 | 293 |
| 35.5. 版本 6.2.1 .....                 | 296 |
| 35.5.1. 从版本 6.2 迁移到版本 6.2.1 .....    | 297 |
| 35.5.2. 详细变更列表 .....                 | 297 |
| 35.6. 版本 6.2 .....                   | 297 |
| 35.6.1. 从版本 6.1 迁移到版本 6.2 .....      | 297 |
| 35.6.2. 从版本 1.x 迁移到版本 6.2 .....      | 297 |
| 35.6.3. 详细变更列表 .....                 | 297 |
| 35.7. 版本 6.1.1 .....                 | 300 |
| 35.7.1. 从版本 6.1 迁移到版本 6.1.1 .....    | 300 |
| 35.7.2. 详细变更列表 .....                 | 300 |
| 35.8. 版本 6.1 .....                   | 300 |
| 35.8.1. 迁移到版本 6.1 .....              | 301 |
| 35.8.2. 详细变更列表 .....                 | 301 |
| 35.9. 版本 6.0 .....                   | 302 |
| 35.9.1. 从版本 1.09 迁移到版本 6.0 .....     | 303 |
| 35.9.2. 从 1.09 之前版本迁移到版本 6.0 .....   | 303 |
| 35.9.3. 详细变更列表 .....                 | 303 |
| 35.10. 版本 1.09 .....                 | 305 |
| 35.11. 版本 1.02 .....                 | 305 |
| 35.11.1. 从版本 1.02 迁移到版本 1.02.1 ..... | 305 |
| 35.11.2. 转储/重载程序 .....               | 305 |
| 35.11.3. 详细变更列表 .....                | 306 |
| 35.12. 版本 1.01 .....                 | 306 |
| 35.12.1. 从版本 1.0 迁移到版本 1.01 .....    | 306 |
| 35.12.2. 详细变更列表 .....                | 308 |
| 35.13. 版本 1.0 .....                  | 309 |
| 35.13.1. 详细变更列表 .....                | 309 |
| 35.14. Postgres95 Beta 0.03 .....    | 310 |
| 35.14.1. 详细变更列表 .....                | 310 |
| 35.15. Postgres95 Beta 0.02 .....    | 312 |
| 35.15.1. 详细变更列表 .....                | 312 |
| 35.16. Postgres95 Beta 0.01 .....    | 312 |
| 35.17. 计时结果 .....                    | 312 |
| 35.17.1. 版本 6.4beta .....            | 313 |
| 35.17.2. 版本 6.3 .....                | 313 |
| 35.17.3. 版本 6.1 .....                | 313 |

---

# 第 21 章 引言

本文档是最初在加州大学伯克利分校开发的 PostgreSQL<sup>1</sup> 数据库管理系统的管理员手册。PostgreSQL 基于 Postgres release 4.2<sup>2</sup>。Postgres 项目由 Michael Stonebraker 教授领导，得到了国防高级研究计划局（DARPA）、陆军研究办公室（ARO）、国家科学基金会（NSF）以及 ESL 公司的赞助。

## 21.1. 资源

本手册集被组织为若干部分：

### 教程

面向新用户的介绍。不涉及高级特性。

### 用户指南

面向用户的一般信息，包括可用的命令和数据类型。

### 程序员指南

面向应用程序员的高级信息。主题包括类型和函数可扩展性、库接口以及应用设计问题。

### 管理员指南

安装和管理信息。受支持的机器列表。

### 开发者指南

面向 Postgres 开发者的信息。此书面向为 Postgres 项目做贡献的人；应用开发的信息见程序员指南。目前包含在程序员指南中。

### 参考手册

关于命令语法的详细参考信息。目前包含在用户指南中。

除本手册集之外，还有其他资源可以帮助你安装和使用 Postgres：

### man 手册页

man 手册页包含关于命令语法的一般信息。

### FAQ

常见问题（FAQ）文档既阐述一般性问题，也阐述一些平台特定的问题。

### README

某些附带的软件包提供了 README 文件。

---

<sup>1</sup><http://postgresql.org/>

<sup>2</sup><http://s2k-ftp.CS.Berkeley.EDU:8000/postgres/postgres.html>

## 网站

Postgres<sup>3</sup> 网站上可能有发行包中未包含的信息。那里有一个 mhonarc 邮件列表流量存档，是许多主题的丰富资源。

## 邮件列表

Postgres Questions<sup>4</sup> 邮件列表是一个让用户问题得到解答的好地方。还有其他可用的邮件列表；详情请参阅该网页。

## 你自己！

Postgres 是一项开源产品。因此，它的持续支持依赖于用户社区。当你开始使用 Postgres 时，你会依赖他人的帮助——无论是通过文档还是通过邮件列表。请考虑回馈你的知识。如果你学到了文档中没有的东西，把它写出来并贡献回去。如果你给代码添加了特性，也把它们贡献出去。即使经验不多的人也能为文档提供更正和小改动，而这是一个很好的起点。Postgres Documentation<sup>5</sup> 邮件列表就是开始的地方。

## 21.2. 术语

在后面的文档中，站点（site）可以解释为安装 Postgres 的主机。由于在单台主机上可以安装多套 Postgres 数据库，这个术语更确切地指任何一套特定的已安装的 Postgres 二进制文件和数据库。

Postgres 超级用户是名为 *postgres* 的用户，他拥有 Postgres 的二进制文件和数据库文件。作为数据库超级用户，所有的保护机制都可以被绕过，任何数据都可以任意访问。此外，Postgres 超级用户还被允许执行一些通常不向所有用户开放的支持程序。注意，Postgres 超级用户不是 Unix 超级用户（后者将称为 *root*）。出于安全原因，超级用户应具有非零的用户标识符（UID）。

数据库管理员（DBA）负责安装 Postgres 并为站点建立实施安全策略的机制。DBA 可以按下述方法添加新用户，并维护一组供 *createdb* 使用的模板数据库。

*postmaster* 是为发往 Postgres 系统的请求充当中转站的进程。前端应用连接到 *postmaster*，由它跟踪系统错误和后端进程之间的通信。*postmaster* 可以接受若干命令行参数来调整其行为。不过，只有当你打算运行多个站点或非默认站点时才需要提供参数。

Postgres 后端（实际的可执行程序 *postgres*）可以由 Postgres 超级用户直接在用户 shell 中执行（以数据库名作为参数）。但这样做会绕过与 *postmaster*/站点关联的共享缓冲池和锁表，因此在多用户站点中不建议这样做。

## 21.3. 记法

文件名前面的 “...” 或 */usr/local/pgsql/* 用来表示 Postgres 超级用户主目录的路径。

在命令概要中，方括号 “[” 和 “]” 表示可选的短语或关键字。花括号 (“{” 和 “}”) 中含有竖线 (“|”) 时表示你必须选择其中一个。

在示例中，圆括号 (“(” 和 “)”) 用于给布尔表达式分组。“|” 是布尔运算符 OR。

示例将展示从不同账户和程序执行的命令。从 *root* 账户执行的命令前带有 “>”。从 Postgres 超级用户账户执行的命令前带有 “%”，而从无特权用户账户执行的命令前带有 “\$”。SQL 命令前带有 “=>” 提示符，也可能没有前导提示符，视上下文而定。

<sup>3</sup> [postgresql.org](http://postgresql.org)

<sup>4</sup> <mailto:questions@postgresql.org>

<sup>5</sup> <mailto:docs@postgresql.org>

## 注意

在撰写本文时 (Postgres v6.4)，整个文档集中标记命令的记法并不完全一致。如发现问题请报告给文档邮件列表<sup>6</sup>。

## 21.4. Y2K 声明

### 作者

由 Thomas Lockhart<sup>7</sup> 写于 1998-10-22。

PostgreSQL 全球开发队伍将 Postgres 软件代码树作为公共服务提供，不对它的行为或性能作任何保证，也不承担任何责任。不过，在撰写本文时：

- 本声明的作者自 1996 年 11 月起就是 Postgres 支持团队的一名志愿者，他不知道 Postgres 代码库中存在任何与 2000 年 1 月 1 日前后的时间过渡 (Y2K) 相关的问题。
- 本声明的作者没有听说在回归测试中或在 Postgres 近期或当前版本的其他实际使用中发现过任何 Y2K 问题的报告。考虑到已安装的用户基础以及用户在支持邮件列表上的积极参与，如果存在问题，我们本应有所耳闻。
- 就作者所知，Postgres 对用两位数字年份指定的日期所做的假定记录在当前用户指南<sup>8</sup>关于数据类型的章节中。对于两位年份，关键的过渡年份是 1970 而不是 2000；例如“70-01-01”被解释为“1970-01-01”，而“69-01-01”被解释为“2069-01-01”。
- 底层操作系统中与获取“当前时间”相关的任何 Y2K 问题都可能传播为 Postgres 中表面的 Y2K 问题。

关于 Y2K 问题的进一步讨论，特别是与开源、免费软件相关的讨论，请参阅 The Gnu Project<sup>9</sup> 和 The Perl Institute<sup>10</sup>

## 21.5. 版权与商标

PostgreSQL 版权所有 (C) 1996-8，归 PostgreSQL 全球开发组所有，并按照 Berkeley 许可的条款分发。

Postgres95 版权所有 (C) 1994-5，归加利福尼亚大学董事会所有。特此授权，可为任何目的使用、复制、修改和分发本软件及其文档，无需付费，也无需书面协议，但条件是所有副本中都必须包含上述版权声明、本段以及后续两段内容。

在任何情况下，加利福尼亚大学都不对任何一方承担因使用本软件及其文档而产生的任何直接、间接、特殊、附带或后果性损害赔偿责任，包括利润损失，即使加利福尼亚大学已被告知此类损害发生的可能性。

<sup>6</sup><mailto:docs@postgresql.org>

<sup>7</sup><mailto:lockhart@alumni.caltech.edu>

<sup>8</sup><http://www.postgresql.org/docs/user/datatype.htm>

<sup>9</sup><http://www.gnu.org/software/year2000.html>

<sup>10</sup><http://language.perl.com/news/y2k.html>

加利福尼亚大学明确不作任何担保，包括但不限于适销性和特定用途适用性的默示担保。本协议项下提供的软件按"原样"提供，加利福尼亚大学没有义务提供维护、支持、更新、增强或修改。

UNIX 是 X/Open, Ltd. 的商标。Sun4、SPARC、SunOS 和 Solaris 是 Sun Microsystems, Inc. 的商标。DEC、DECstation、Alpha AXP 和 ULTRIX 是 Digital Equipment Corp. 的商标。PA-RISC 和 HP-UX 是 Hewlett-Packard Co. 的商标。OSF/1 是 Open Software Foundation 的商标。

# 第 22 章 平台移植

本手册描述 Postgres 6.4 版。Postgres 开发者社区已经在许多平台上编译和测试过 Postgres。最新信息请查阅网站<sup>1</sup>。

## 22.1. 当前支持的平台

截至发布时，下列平台已经过测试：

表 22.1. 支持的平台

| OS                 | 处理器     | 版本     | 报告日期       | 备注                                                                 |
|--------------------|---------|--------|------------|--------------------------------------------------------------------|
| AIX 4.2.1          | RS6000  | v6.4   | 1998-10-27 | (Andreas Zeugswetter <sup>2</sup> )                                |
| BSDI               | x86     | v6.4   | 1998-10-25 | (Bruce Momjian <sup>3</sup> )                                      |
| FreeBSD 2.2.x-3.x  | x86     | v6.4   | 1998-10-26 | (Tatsuo Ishii <sup>4</sup> , Marc Fournier <sup>5</sup> )          |
| DGUX 5.4R4.11      | m88k    | v6.3   | 1998-03-01 | v6.4 可能没问题。需要新的维护者。(Brian E Gallew <sup>6</sup> )                  |
| Digital Unix 4.0   | Alpha   | v6.4   | 1998-10-29 | 有些小问题，可通过补丁解决 (Pedro J. Lobo <sup>7</sup> )                        |
| HPUX               | PA-RISC | v6.4   | 1998-10-25 | 9.0x 和 10.20 皆可 (Tom Lane <sup>8</sup> , Stan Brown <sup>9</sup> ) |
| IRIX 6.x           | MIPS    | v6.3   | 1998-03-01 | 5.x 有所不同 (Andrew Martin <sup>10</sup> )                            |
| linux 2.0.x        | Alpha   | v6.3.2 | 1998-04-16 | 基本成功。v6.4 尚需完善。(Ryan Kirkpatrick <sup>11</sup> )                   |
| linux 2.0.x        | x86     | v6.4   | 1998-10-27 | (Thomas Lockhart <sup>12</sup> )                                   |
| linux 2.0.x/glibc2 | x86     | v6.4   | 1998-10-25 | (Oliver Elphick <sup>13</sup> , Taral <sup>14</sup> )              |
| linux 2.0.x        | Sparc   | v6.4   | 1998-10-25 | (Tom Szybist <sup>15</sup> )                                       |

<sup>1</sup><http://www.postgresql.org/docs/admin/ports.htm>

<sup>2</sup><mailto:Andreas.Zeugswetter@telecom.at>

<sup>3</sup><mailto:maillist@candle.pha.pa.us>

<sup>4</sup><mailto:t-ishii@sra.co.jp>

<sup>5</sup><mailto:scrappy@hub.org>

<sup>6</sup><mailto:geek+@cmu.edu>

<sup>7</sup><mailto:pjlobo@euitt.upm.es>

<sup>8</sup><mailto:tgl@sss.pgh.pa.us>

<sup>9</sup><mailto:stanb@awod.com>

<sup>10</sup><mailto:martin@biochemistry.ucl.ac.uk>

<sup>11</sup><mailto:rkirkpat@nag.cs.colorado.edu>

<sup>12</sup><mailto:lockhart@alumni.caltech.edu>

<sup>13</sup><mailto:olly@lfix.co.uk>

<sup>14</sup><mailto:taral@mail.utexas.edu>

<sup>15</sup><mailto:szybist@boxhill.com>

| OS                    | 处理器     | 版本     | 报告日期       | 备注                                                                    |
|-----------------------|---------|--------|------------|-----------------------------------------------------------------------|
| linuxPPC 2.1.24       | PPC603e | v6.4   | 1998-10-26 | Powerbook 2400c(<br>Tatsuo Ishii <sup>16</sup> )                      |
| mklinux DR3           | PPC750  | v6.4   | 1998-09-16 | PowerMac 7600 (<br>Tatsuo Ishii <sup>17</sup> )                       |
| NetBSD/i386 1.3.2     | x86     | v6.4   | 1998-10-25 | (Brook Milligan <sup>18</sup> )                                       |
| NetBSD-current        | NS32532 | v6.4   | 1998-10-27 | (日期/时间运算有小<br>问题 (Jon Buller <sup>19</sup> ))                         |
| NetBSD/sparc 1.<br>3H | Sparc   | v6.4   | 1998-10-27 | (Tom I Helbek<br>kmo <sup>20</sup> )                                  |
| NetBSD 1.3            | VAX     | v6.3   | 1998-03-01 | (Tom I Helbek<br>kmo <sup>21</sup> )                                  |
| SCO UnixWare 2.x      | x86     | v6.3   | 1998-03-01 | 即 UNIVEL (Billy G.<br>Allie <sup>22</sup> )                           |
| SCO UnixWare 7        | x86     | v6.4   | 1998-10-04 | (Billy G. Allie <sup>23</sup> )                                       |
| Solaris               | x86     | v6.4   | 1998-10-28 | (Marc Fournier <sup>24</sup> )                                        |
| Solaris 2.6-2.7       | Sparc   | v6.4   | 1998-10-28 | (Tom Szybist <sup>25</sup> ,<br>Frank Ridder<br>busch <sup>26</sup> ) |
| SunOS 4.1.4           | Sparc   | v6.3   | 1998-03-01 | 已提交补丁 (Tatsuo<br>Ishii <sup>27</sup> )                                |
| SVR4                  | MIPS    | v6.4   | 1998-10-28 | 不支持 64 位整<br>数 (Frank Ridder<br>busch <sup>28</sup> )                 |
| SVR4 4.4              | m88k    | v6.2.1 | 1998-03-01 | 打补丁后确认可<br>用 (Doug Winter<br>burn <sup>29</sup> )                     |
| Windows NT            | x86     | v6.4   | 1998-10-08 | 配合 Cygwin 库<br>大体可用。尚<br>无 DLL (Horak<br>Daniel <sup>30</sup> )       |

v6.3.x 所列的平台应该也能在 v6.4 上工作，但在编制本列表时我们没有收到明确的确认。

<sup>16</sup>mailto:t-ishii@sra.co.jp

<sup>17</sup>mailto:t-ishii@sra.co.jp

<sup>18</sup>mailto:brook@trillium.NMSU.Edu

<sup>19</sup>mailto:jonb@metronet.com

<sup>20</sup>mailto:tih@hamartun.priv.no

<sup>21</sup>mailto:tih@hamartun.priv.no

<sup>22</sup>mailto:Bill.Allie@mug.org

<sup>23</sup>mailto:Bill.Allie@mug.org

<sup>24</sup>mailto:scrappy@hub.org

<sup>25</sup>mailto:szybist@boxhill.com

<sup>26</sup>mailto:ridderbusch.pad@sni.de

<sup>27</sup>mailto:t-ishii@sra.co.jp

<sup>28</sup>mailto:ridderbusch.pad@sni.de

<sup>29</sup>mailto:dlw@seavme.xroads.com

<sup>30</sup>mailto:Dan.Horak@email.cz

**注意**

对于 Windows NT, Postgres 的服务器端移植最近已经完成。最新信息请查阅 Askesis Postgres 主页<sup>31</sup>。你也可以在 Postgres 网站<sup>32</sup>上查找可能的补丁。

## 22.2. 未支持的平台

有几个平台曾被尝试过，并据报道无法在标准发行版本上工作。这里列出的另一些平台则没有提供足够的库支持，无法尝试。

**表 22.2. 可能不兼容的平台**

| OS       | 处理器       | 版本   | 报告日期       | 备注                                                |
|----------|-----------|------|------------|---------------------------------------------------|
| MacOS    | all       | v6.3 | 1998-03-01 | 库不兼容；请使用 ODBC/JDBC                                |
| NetBSD   | arm32     | v6.3 | 1998-03-01 | 尚未可用 (Dave Millen <sup>33</sup> )                 |
| NetBSD   | m68k      | v6.3 | 1998-03-01 | Amiga、HP300、Mac；尚未可用 (Henry Hotz <sup>34</sup> )  |
| NextStep | x86       | v6.x | 1998-03-01 | 仅客户端支持；v1.0.9 加补丁可用 (David Wetzel <sup>35</sup> ) |
| Ultrix   | MIPS,VAX? | v6.x | 1998-03-01 | 近期无报告；已过时？                                        |
| Windows  | x86       | v6.3 | 1998-03-01 | 库不兼容；客户端也许可行；请使用 ODBC/JDBC                        |

注意，使用第三方的 Posix 移植工具和库，前端在 Windows 上的移植显然是可能的。

<sup>31</sup><http://www.asksis.nl/AskesisPostgresIndex.html>

<sup>32</sup><http://postgresql.org>

<sup>33</sup><mailto:dmill@globalnet.co.uk>

<sup>34</sup><mailto:hotz@jpl.nasa.gov>

<sup>35</sup><mailto:dave@turbocat.de>

---

# 第 23 章 配置选项

## 23.1. 配置参数 (configure)

configure 可用的全部参数可以通过键入以下命令获得：

```
$ ./configure --help
```

安装者可能对以下参数感兴趣：

Directory and file names:

|                               |                                                                                       |
|-------------------------------|---------------------------------------------------------------------------------------|
| <code>--prefix=PREFIX</code>  | install architecture-independent files in PREFIX<br>[ <code>/usr/local/pgsql</code> ] |
| <code>--bindir=DIR</code>     | user executables in DIR [ <code>EPREFIX/bin</code> ]                                  |
| <code>--libdir=DIR</code>     | object code libraries in DIR [ <code>EPREFIX/lib</code> ]                             |
| <code>--includedir=DIR</code> | C header files in DIR [ <code>PREFIX/include</code> ]                                 |
| <code>--mandir=DIR</code>     | man documentation in DIR [ <code>PREFIX/man</code> ]                                  |

Features and packages:

|                                     |                                                                    |
|-------------------------------------|--------------------------------------------------------------------|
| <code>--disable-FEATURE</code>      | do not include FEATURE (same as <code>--enable-FEATURE=no</code> ) |
| <code>--enable-FEATURE[=ARG]</code> | include FEATURE [ <code>ARG=yes</code> ]                           |
| <code>--with-PACKAGE[=ARG]</code>   | use PACKAGE [ <code>ARG=yes</code> ]                               |
| <code>--without-PACKAGE</code>      | do not use PACKAGE (same as <code>--with-PACKAGE=no</code> )       |

--enable and --with options recognized:

|                                       |                                                              |
|---------------------------------------|--------------------------------------------------------------|
| <code>--with-template=template</code> | use operating system template file<br>see template directory |
| <code>--with-includes=incdir</code>   | site header files for tk/tcl, etc in DIR                     |
| <code>--with-libs=libdir</code>       | also search for libraries in DIR                             |
| <code>--with-libraries=libdir</code>  | also search for libraries in DIR                             |
| <code>--enable-locale</code>          | enable locale support                                        |
| <code>--enable-recode</code>          | enable cyrillic recode support                               |
| <code>--with-mb=encoding</code>       | enable multi-byte support                                    |
| <code>--with-pgport=portnum</code>    | change default startup port                                  |
| <code>--with-tcl</code>               | use tcl                                                      |
| <code>--with-tclconfig=tcldir</code>  | tclConfig.sh and tkConfig.sh are in DIR                      |
| <code>--with-perl</code>              | use perl                                                     |
| <code>--with-odbc</code>              | build ODBC driver package                                    |
| <code>--with-odbcinst=odbcdir</code>  | change default directory for odbcinst.ini                    |
| <code>--enable-cassert</code>         | enable assertion checks (debugging)                          |
| <code>--with-CC=compiler</code>       | use specific C compiler                                      |
| <code>--with-CXX=compiler</code>      | use specific C++ compiler                                    |

有些系统在构建Postgres的某个特定特性时可能会遇到问题。例如，C++ 编译器损坏的系统可能需要指定`--without-CXX`来劝使构建过程忽略 `libpq++` 的构建。

## 23.2. 构建参数 (make)

许多与安装相关的参数可以在Postgres 安装的构建阶段设置。

在大多数情况下，这些参数应当放在专门用于此目的的文件 `Makefile.custom` 中。默认的发行版不包含这个可选文件，因此你要用自己选择的文本编辑器创建它。升级安装时，只需在构建之前把旧的 `Makefile.custom` 复制到新安装中即可。

```
make [ variable=value [,...] ]
```

可以指定的许多变量中的几个如下：

`POSTGRES DIR`

安装树的顶层。

`BINDIR`

应用程序和实用工具的位置。

`LIBDIR`

目标库的位置，包括共享库。

`HEADERDIR`

头文件的位置。

`ODBCINST`

安装范围的psqlODBC（ODBC）配置文件的位置。

还有一些不那么常用的可选参数。  
下面列出的许多参数适合在进行Postgres服务器代码开发时使用。

`CFLAGS`

为 C 编译器设置标志。应当用 `"+="` 赋值以保留相关的默认参数。

`YFLAGS`

为 yacc/bison 解析器设置标志。`-v` 可以用来帮助诊断构建新解析器时的问题。应当用 `"+="` 赋值以保留相关的默认参数。

`USE_TCL`

启用 Tcl 接口的构建。

`HSTYLE`

用于从头构建文档的 DocBook HTML 样式表。除非你正在根据 `doc/src/sgml/` 中与 DocBook 兼容的 SGML 源文档开发新文档，否则不会用到。

`PSTYLE`

用于从头构建印刷文档的 DocBook 样式表。除非你正在根据 `doc/src/sgml/` 中与 DocBook 兼容的 SGML 源文档开发新文档，否则不会用到。

下面是一个 PentiumPro Linux 系统的 `Makefile.custom` 示例：

```
# Makefile.custom
```

```
# Thomas Lockhart 1998-03-01

POSTGRES_DIR= /opt/postgres/current
CFLAGS+= -m486 # -g -O0
USE_TCL= true
TCL_LIB= -ltcl
X_LIBS= -L/usr/X11/lib
TK_LIB= -ltk

# documentation

HSTYLE= /home/tgl/SGML/db118.d/docbook/html
PSTYLE= /home/tgl/SGML/db118.d/docbook/print
```

## 23.3. 区域支持

### 注意

由 Oleg Bartunov 撰写。关于区域和俄语支持的更多信息，请见Oleg 的网页<sup>1</sup>。

在为俄罗斯莫斯科的一家公司做项目时，我遇到了 postgresql 不支持国家字母表的问题。在寻找可能的变通办法之后，我决定自己开发区域支持。我不是 C 程序员，但在使用 perl (调试) 和 glimpse 工作时已经有一些区域编程的经验。在 Postgres源码树中挖掘了几天之后，我对 src/backend/utils/adt/varlena.c 和 src/backend/main/main.c 做了非常小的修改，就得到了我需要的东西！我当时只做了 LC\_CTYPE和LC\_COLLATE的支持，但后来其他人又添加了LC\_MONETARY。我收到许多人关于这个补丁的消息，因此我决定把它发给开发者，而（让我惊讶的是）它被并入了Postgres 发行版。

人们经常抱怨区域支持对他们不起作用。有几种常见的错误：

- 编译之前没有正确配置 postgresql。必须用 --enable-locale 选项运行 configure 来启用区域支持。启动 postmaster 时没有正确设置环境。必须在运行 postmaster 之前定义环境变量 LC\_CTYPE和LC\_COLLATE，因为后端从环境中获取区域信息。我使用下面的 shell 脚本 (runpostgres)：

```
#!/bin/sh

export LC_CTYPE=koi8-r
export LC_COLLATE=koi8-r
postmaster -B 1024 -S -D/usr/local/pgsql/data/ -o '-Fe'
```

并从 rc.local 中这样运行它：

```
/bin/su - postgres -c "/home/postgres/runpostgres"
```

- 操作系统的区域支持损坏（例如，Linux 下 libc 的区域支持曾多次变化，这引起了很多问题）。最新的 perl 也支持区域，如果区域损坏，perl -v会抱怨类似下面的内容：

```
8:17[mira]:~/WWW/postgres>setenv LC_CTYPE not_exist
8:18[mira]:~/WWW/postgres>perl -v
perl: warning: Setting locale failed.
```

<sup>1</sup><http://www.sai.msu.su/~megeera/postgres/>

```
perl: warning: Please check that your locale settings:
    LC_ALL = (unset),
    LC_CTYPE = "not_exist",
    LANG = (unset)
are supported and installed on your system.
perl: warning: Falling back to the standard locale ("C").
```

- 区域文件的位置错误！可能的位置包括：`/usr/lib/locale` (Linux, Solaris), `/usr/share/locale` (Linux), `/usr/lib/nls/loc` (DUX 4.0). 请查看`man locale`以找到正确的位置。在 Linux 下，我在`/usr/lib/locale`和`/usr/share/locale`之间做了一个符号链接，以确保下一个 `libc` 不会破坏我的区域。

### 23.3.1. 有什么好处？

可以对包含国家字母表字符的字符串使用 `~*` 和 `order by` 操作符。非英语用户肯定需要它。如果你不想使用区域相关的东西，只需取消定义 `USE_LOCALE` 变量。

### 23.3.2. 有什么缺点？

使用区域有一个明显的缺点——速度！所以，只在真正需要时才使用区域。

## 23.4. Kerberos 认证

Kerberos 是一种工业标准的安全认证系统，适合在公共网络上进行分布式计算。

### 23.4.1. 可用性

Kerberos 认证系统并不随 Postgres 一起发行。Kerberos 的各个版本通常以操作系统厂商的可选软件的形式提供。此外，还可以通过 MIT Project Athena<sup>2</sup> 获取源码发行版。

#### 注意

即使你的厂商提供了某个版本，你也可能希望获取 MIT 版本，因为有些厂商的移植版本被故意削弱或弄得与 MIT 版本无法互操作。

位于美国和加拿大之外的用户请注意，Kerberos 中实际加密代码的分发受美国政府出口法规的限制。

关于 Kerberos 的询问应当指向你的厂商或 MIT Project Athena<sup>3</sup>。注意，FAQL（常见问题列表）会定期张贴到 Kerberos 邮件列表<sup>4</sup>（发送邮件订阅<sup>5</sup>）和 USENET 新闻组<sup>6</sup>。

### 23.4.2. 安装

Kerberos 本身的安装在 Kerberos 安装说明中有详细介绍。请确保服务器密钥文件（`srvtab` 或 `keytab`）能被 Postgres 账户以某种方式读取。

<sup>2</sup><ftp://athena-dist.mit.edu>

<sup>3</sup>[info-kerberos@athena.mit.edu](mailto:info-kerberos@athena.mit.edu)

<sup>4</sup><mailto:kerberos@ATHENA.MIT.EDU>

<sup>5</sup><mailto:kerberos-request@ATHENA.MIT.EDU>

<sup>6</sup><news:comp.protocols.kerberos>

Postgres及其客户端可以编译为使用 MIT Kerberos协议的版本 4 或版本 5，方法是把文件 `src/Makefile.global` 中的 `KRBVERS` 变量设置为适当的值。你还可以更改Postgres期望找到相关库、头文件和它自己的服务器密钥文件的位置。

编译完成后，必须把Postgres 注册为一个Kerberos 服务。关于注册服务的更多细节，请参阅 Kerberos 操作说明及相关的手册页。

### 23.4.3. 操作

初次安装之后，Postgres 在各方面都应当像正常的 Kerberos服务一样运行。关于认证使用的细节，请参阅postmaster和 psql的 PostgreSQL 用户指南参考章节。

在Kerberos版本 5 的钩子中，对用户和服务命名做了以下假设：

- 假定用户主体名（aname）的第一个组件中包含实际的 Unix/Postgres用户名。
- 假定Postgres服务有两个组件，即服务名和一个主机名，主机名按版本 4 的方式规范化（即去掉所有域名后缀）。

**表 23.1. Kerberos 参数示例**

| 参数   | 示例                                      |
|------|-----------------------------------------|
| user | frew@S2K.ORG                            |
| user | aoki/HOST=miyu.S2K.Berkeley.EDU@S2K.ORG |
| host | postgres_dbms/ucbvax@S2K.ORG            |

在 MIT 发布版本 5 的正式版本之后的某个时候，将不再支持版本 4。

---

# 第 24 章 安装

Postgres v6.4 的完整安装说明。

在安装 Postgres 之前，你可以访问 [www.postgresql.org](http://www.postgresql.org)<sup>1</sup> 获取最新信息、补丁等。

这些安装说明假定：

- 命令与 Unix 兼容。见下面的注记。
- 除特别说明外均使用默认值。
- 用户 `postgres` 是 Postgres 超级用户。
- 源代码路径为 `/usr/src/pgsql`（其他路径也可以）。
- 运行时路径为 `/usr/local/pgsql`（其他路径也可以）。

这些命令是在 RedHat Linux 4.2 版上使用 `tcsh` shell 测试的。除特别说明外，它们在大多数系统上很可能都能工作。像 `ps` 和 `tar` 这样的命令，应该使用哪些选项在各个平台之间可能差异极大。输入这些命令之前请运用常识。

我们的 Makefile 需要 GNU make（本文档中称“`gmake`”）。它们不能与非 GNU 的 `make` 程序一起工作。如果你的 GNU make 是以“`make`”而不是“`gmake`”的名称安装的，那么你就使用 `make` 命令。这没有问题，但你必须拥有 GNU 形式的 `make` 才能安装成功。

## 24.1. 运行 Postgres 的要求

关于所支持平台的最新信息位于 <http://www.postgresql.org/docs/admin/install.htm><sup>2</sup>。一般来说，大多数拥有现代库的 Unix 兼容平台都应该能够运行 Postgres。

尽管运行 Postgres 所需的最小内存可以低至 8MB，但在一个运行 X-Windows 的相对较快的双处理器系统上，把内存扩展到 96MB 时回归测试的运行时间有明显改善。规律是：内存永远不嫌多。

检查你是否有足够的磁盘空间。`/usr/src/pgsql` 大约需要 30 MB，`/usr/local/pgsql`（不含你的数据库）大约需要 5 MB，一个空数据库需要 1 MB。回归测试期间数据库会临时增长到约 20 MB。发行版 tar 文件还需要约 3 MB。

因此我们建议，在安装和测试期间，`/usr/local` 之下保留远超 20 MB 的空闲空间，在包含数据库的磁盘分区上另外保留 25 MB 空闲空间。删除源文件、tar 文件和回归数据库之后，`/usr/local/pgsql` 需要 2 MB，空数据库需要 1 MB，再加上大约五倍于把数据库数据存成平面文件所需的空间。

要检查磁盘空间，使用

```
$ df -k
```

## 24.2. 安装过程

### Postgres 安装

对于全新安装或从旧版本 Postgres 升级：

---

<sup>1</sup><http://www.postgresql.org>

<sup>2</sup><http://www.postgresql.org/docs/admin/install.htm>

1. 阅读任何最后时刻的信息和针对特定平台的移植说明。本文件末尾有关于 Ultrix4.x、Linux、BSD/OS 和 NeXT 的一些平台特定说明。目录 `/usr/src/pgsql/doc` 中还有其他文件，包括 FAQ-Irix 和 FAQ-Linux。也看看目录 `ftp://ftp.postgresql.org/pub`。如果该目录中有一个名为 `INSTALL` 的文件，那么该文件包含最新的安装信息。

请注意，前面列表中的“已测试”平台只是意味着有人曾经花力气确保 Postgres 发行版能在该平台上不经修改地编译和运行。由于当前的开发者无法访问所有这些平台，其中一些可能因小问题而无法在当前版本中干净地编译并通过回归测试。任何此类已知问题及其解决方案都会公布在 `ftp://ftp.postgresql.org/pub/INSTALL`。

2. (可选的) 如果 Postgres 超级用户账号尚不存在，则创建它（通常使用 `postgres`）。

Postgres 文件的所有者可以是任何无特权的用户账号。它绝不能是 `root`、`bin` 或任何其他拥有特殊访问权限的账号，否则会造成安全风险。

3. 以 Postgres 超级用户账号登录。安装中余下的大多数步骤都将在该账号下进行。
4. 从 Internet 上 Ftp 文件 `ftp://ftp.postgresql.org/pub/postgresql-v6.4.tar.gz`<sup>3</sup>。把它存放在你的主目录中。
5. 某些平台使用 flex。如果你的系统使用 flex，请确保你有一个好的版本。要检查，输入

```
$ flex --version
```

如果找不到 flex 命令，那么你很可能不需要它。如果版本是 2.5.2、2.5.4 或更高，就没有问题。如果是 2.5.3 或低于 2.5.2，你就得升级 flex。可以从 `ftp://prep.ai.mit.edu/pub/gnu/flex-2.5.4.tar.gz` 获取。

如果你需要 flex 却没有它或版本不对，那么当你尝试编译程序时会被告知。如果不确定是否需要，尽可以跳过这一步。如果确实需要，当你尝试编译 Postgres 时会被告知安装/升级 flex。

你可以从 root 账号完成整个 flex 安装，尽管这并非绝对必要。假设你希望安装把文件放在常规的默认位置，输入以下命令：

```
$ su -
$ cd /usr/local/src
ftp prep.ai.mit.edu
ftp> cd /pub/gnu/
ftp> binary
ftp> get flex-2.5.4.tar.gz
ftp> quit
$ gunzip -c flex-2.5.4.tar.gz | tar xvf -
$ cd flex-2.5.4
$ configure --prefix=/usr
$ gmake
$ gmake check
# You must be root when typing the next line:
$ gmake install
$ cd /usr/local/src
$ rm -rf flex-2.5.4
```

这将更新文件 `/usr/man/man1/flex.1`、`/usr/bin/flex`、`/usr/lib/libfl.a`、`/usr/include/FlexLexer.h`，并添加一个指向 flex 的链接 `/usr/bin/flex++`。

<sup>3</sup>`ftp://ftp.postgresql.org/pub/postgresql-v6.4.tar.gz`

6. 如果你不是在升级现有系统，则跳到步骤 9。如果你是在升级现有系统，请备份你的数据库。对于 alpha 和 beta 级的发布，数据库格式容易变化，往往每隔几周就变一次，除了 HACKERS 邮件列表中一句快速评论之外没有任何通知。正式发布总是要求从旧版本进行转储/重载。因此跳过这一步是很糟糕的主意。

### 提示

不要使用 v6.0 的 pg\_dumpall 脚本，否则所有对象都将归 Postgres 超级用户所有。

要转储你相当新的 post-v6.0 数据库安装，输入

```
$ pg_dumpall -z > db.out
```

要在升级 Postgres 之前对你现有的旧数据库使用最新的 pg\_dumpall 脚本，请从新发行版中取出最新版本的 pg\_dumpall:

```
$ cd
$ gunzip -c postgresql-v6.4.tar.gz \
  | tar xvf - src/bin/pg_dump/pg_dumpall
$ chmod a+x src/bin/pg_dump/pg_dumpall
$ src/bin/pg_dump/pg_dumpall -z > db.out
$ rm -rf src
```

如果你想保留对象 id (oid)，则在运行 pg\_dumpall 时使用 -o 选项。但除非你有特殊理由这样做（例如在表中把 OID 用作键），否则不要这样做。

如果 pg\_dumpall 命令似乎花了很长时间而你觉得它可能已经死掉，那么可以从另一个终端多次输入

```
$ ls -l db.out
```

看文件的大小是否在增长。

请注意，如果你是从 Postgres95 v1.09 之前的版本升级，那么你必须备份数据库、安装 Postgres95 v1.09、恢复数据库，然后再次备份。你还应当阅读发行说明，其中应该涵盖任何针对特定发行的问题。

### 小心

你必须确保数据库在备份期间没有被更新。必要时，关闭 postmaster，编辑文件 /usr/local/pgsql/data/pg\_hba.conf 的权限使只有你能访问，然后把 postmaster 重新启动。

7. 如果你是在升级现有系统，则终止 postmaster。输入

```
$ ps -ax | grep postmaster
```

这会列出若干进程的进程号。输入下面这行，其中 *pid* 替换为 postmaster 进程的进程 id。（不要使用进程 "grep postmaster" 的 id。）输入

```
$ kill pid
```

来真正停止该进程。

### 提示

在开机时自动启动 Postgres 的系统上，很可能有一个能完成同样事情的启动文件。例如，在我的 Linux 系统上我可以输入

```
$ /etc/rc.d/init.d/postgres.init stop
```

来停止 Postgres。

8. 如果你是在升级现有系统，则把旧目录移开。如果磁盘空间紧张，你可能得备份并删除这些目录。如果这样做，请把旧数据库保存在 `/usr/local/pgsql/data` 目录树中。至少要保存文件 `/usr/local/pgsql/data/pg_hba.conf`。

输入以下命令：

```
$ su -
$ cd /usr/src
$ mv pgsql pgsql_6_0
$ cd /usr/local
$ mv pgsql pgsql_6_0
$ exit
```

如果你不使用 `/usr/local/pgsql/data` 作为数据目录（检查环境变量 `PGDATA` 是否被设置成了别的值），那么你也应该以同样的方式移动该目录。

9. 创建新的源代码目录和安装目录。你的安装中实际路径可以不同，但在整个过程中必须保持一致。

### 注意

本安装过程中有两处可以让你指定程序、库、文档和其他文件的安装位置。通常在安装的 `make install` 阶段指定这些就足够了。

输入

```
$ su
$ cd /usr/src
$ mkdir pgsql
$ chown postgres:postgres pgsql
$ cd /usr/local
$ mkdir pgsql
$ chown postgres:postgres pgsql
$ exit
```

10. 解压并解开新的源文件。输入

```
$ cd /usr/src/pgsql
$ gunzip -c ~/postgresql-v6.4.tar.gz | tar xvf -
```

11. 为你的系统配置源代码。在这一步你可以为构建过程指定实际的安装路径（见下面的 `--prefix` 选项）。输入

```
$ cd /usr/src/pgsql/src
$ ./configure [ options ]
```

- a. (可选的) 除其他工作外, `configure` 脚本会从 `template` 子目录中提供的文件里选择一个特定系统的"模板"文件。如果它猜不出你的系统该用哪一个, 它会说明并退出。那种情况下你需要弄清楚该用哪一个, 然后再次运行 `configure`, 这次给出 `--with-template=TEMPLATE` 选项来选中正确的文件。

### 请报告问题

如果你的系统没有被 `configure` 自动识别而不得不这样做, 请发电子邮件到 [scrappy@hub.org](mailto:scrappy@hub.org)<sup>4</sup>, 附上程序 `./config.guess` 的输出。指明模板文件应该是什么。

- b. (可选的) 选择配置选项。详情见配置选项。不过, 对于不带多字节字符支持或区域排序支持等额外选项的普通首次安装, 选好安装区域并不加额外选项地运行 `configure` 也许就够了。`configure` 脚本接受许多额外的选项, 如果你不喜欢默认配置可以用它们。要全部看到, 输入

```
./configure --help
```

一些较常用的选项有:

|                                       |                                                                                              |
|---------------------------------------|----------------------------------------------------------------------------------------------|
| <code>--prefix=BASEDIR</code>         | 为 Postgres 配置的安装选择一个不同的基目录。默认为 <code>/usr/local/pgsql</code> 。                               |
| <code>--with-template=TEMPLATE</code> | 使用模板文件 <code>TEMPLATE</code> — 模板文件假定在目录 <code>src/template</code> 中, 请到那里查找合适的值。            |
| <code>--with-tcl</code>               | 构建需要 Tcl/Tk 的接口库和程序, 包括 <code>libpgtcl</code> 、 <code>pgtclsh</code> 和 <code>pgtksh</code> 。 |
| <code>--with-perl</code>              | 构建 Perl 接口库。                                                                                 |
| <code>--with-odbc</code>              | 构建 ODBC 驱动程序包。                                                                               |
| <code>--enable-hba</code>             | 启用基于主机的认证 (默认)                                                                               |
| <code>--disable-hba</code>            | 禁用基于主机的认证                                                                                    |
| <code>--enable-locale</code>          | 启用 <code>USE_LOCALE</code>                                                                   |
| <code>--enable-cassert</code>         | 启用 <code>ASSERT_CHECKING</code>                                                              |
| <code>--with-CC=compiler</code>       | 使用 <code>configure</code> 脚本找不到的某个特定 C 编译器。                                                  |
| <code>--with-CXX=compiler</code>      | 使用 <code>configure</code> 脚本找不到的某个特定 C++ 编译器,                                                |
| <code>--without-CXX</code>            | 或者完全排除 C++ 编译。(目前这只影响 <code>libpq++</code> 。)                                                |

<sup>4</sup><mailto:scrappy@hub.org>

- c. 下面是在 Sparc Solaris 2.5 系统上使用的 configure 脚本，安装基目录指定为 /opt/postgres:

```
$ ./configure --prefix=/opt/postgres \
  --with-template=sparc_solaris-gcc --with-pgport=5432 \
  --enable-hba --disable-locale
```

### 提示

当然，你可以把这三行都写在同一行上。

12. 安装 man 和 HTML 文档。输入

```
$ cd /usr/src/pgsql/doc
$ gmake install
```

文档也有 Postscript 格式。在同一目录中查找以 .ps.gz 结尾的文件。

13. 编译程序。输入

```
$ cd /usr/src/pgsql/src
$ gmake all >& make.log &
$ tail -f make.log
```

但愿最后显示的一行是

```
All of PostgreSQL is successfully made. Ready to install.
```

这时（或者更早，随你愿意）输入 control-C 退出 tail。（如果以后遇到问题，你可以检查文件 make.log 中的警告和错误消息。）

### 注意

你可能会在 make.log 中发现许多警告消息。除非以后遇到问题，这些消息可以放心地忽略。

如果编译器失败并给出找不到 flex 命令的消息，则按前述方法安装 flex。然后回到本目录，输入

```
$ make clean
```

再重新编译。

编译器选项（如优化和调试）可以在命令行上用 COPT 变量指定。例如，输入

```
$ gmake COPT="-g" all >& make.log &
```

会在构建的所有步骤中调用你编译器的 -g 选项。更多细节见 src/Makefile.global.in。

14. 安装程序。输入

```
$ cd /usr/src/pgsql/src
$ gmake install >& make.install.log &
$ tail -f make.install.log
```

最后显示的一行将是

```
gmake[1]: Leaving directory `/usr/src/pgsql/src/man'
```

这时（或者更早，随你愿意）输入 control-C 退出 tail。

15. 如有必要，告诉你的系统如何找到新的共享库。你可以做下面之一，最好是第一种：

- a. （可选的）以 root 身份编辑文件 `/etc/ld.so.conf`。向该文件加入一行

```
/usr/local/pgsql/lib
```

然后运行命令 `/sbin/ldconfig`。

- b. （可选的）在 bash shell 中，输入

```
export LD_LIBRARY_PATH=/usr/local/pgsql/lib
```

- c. （可选的）在 csh shell 中，输入

```
setenv LD_LIBRARY_PATH /usr/local/pgsql/lib
```

请注意，上述命令在不同的操作系统之间可能差异极大。请查阅平台特定说明，例如 Ultrix4.x 和非 ELF Linux 的那些。

如果你创建数据库时收到消息

```
pg_id: can't load library 'libpq.so'
```

那么上面的步骤就是必要的。照做之后，再试着创建数据库。

16. （可选的）如果你配置时用了 `--with-perl` 选项，请检查安装日志看 Perl 模块是否真的被安装了。如果你听从了我们的建议让 Postgres 文件归一个无特权用户所有，那么 Perl 模块会因为对 Perl 库目录没有写权限而未被安装。你可以现在或以后补装：通过 `su` 变成拥有 Perl 库的用户（通常是 root），然后执行

```
$ cd /usr/src/pgsql/src/interfaces/perl5
$ gmake install
```

17. 如果尚未做过，则为使用 Postgres 准备账号 `postgres`。任何要使用 Postgres 的账号都必须同样地准备。

有多种方法可以影响 Postgres 服务器的运行时环境。更多信息参见管理员指南。

### 注意

下面的说明针对 bash/sh shell。对其他 shell 请相应调整。

- a. 把下面几行加入你的登录环境：shell 的 `~/.bash_profile`：

```
PATH=$PATH:/usr/local/pgsql/bin
MANPATH=$MANPATH:/usr/local/pgsql/man
PGLIB=/usr/local/pgsql/lib
PGDATA=/usr/local/pgsql/data
export PATH MANPATH PGLIB PGDATA
```

- b. 如果用户的区域排序方案与标准 C 区域不同，一些回归测试可能失败。

如果你用 `--enable-locale` 选项配置并编译了 Postgres，那么在启动 `postmaster` 之前，把区域环境设置为 C（或取消所有 `LC_*` 变量），为此把这几行额外加入你的登录环境：

```
LC_COLLATE=C
LC_CTYPE=C
LC_COLLATE=C
export LC_COLLATE LC_CTYPE LC_COLLATE
```

- c. 在继续余下步骤之前，确保你已经定义了这些变量。最简单的做法是输入：

```
$ source ~/.bash_profile
```

18. 以 Postgres 超级用户账号（通常是 `postgres`）创建数据库安装。不要以 `root` 身份做下面这件事！那会是一个重大的安全漏洞。输入

```
$ initdb
```

19. 设置访问数据库系统的权限。编辑文件 `/usr/local/pgsql/data/pg_hba.conf` 来完成。说明包含在该文件中。（如果你的数据库不在默认位置，即如果 `PGDATA` 被设置为指向别处，那么该文件的位置也会相应改变。）完成后应把该文件重新设为只读。如果你是从 v6.0 或更晚版本升级，可以把旧数据库中的 `pg_hba.conf` 文件复制到新数据库的上面，而不必从头重做该文件。

20. 简要地测试一下后端能否启动并运行：从命令行运行它。

- a. 在后台启动 `postmaster` 守护进程，输入

```
$ cd
$ postmaster -i
```

- b. 创建一个数据库，输入

```
$ createdb
```

- c. 连接到新数据库：

```
$ psql
```

- d. 运行一个示例查询：

```
postgres=> SELECT datetime 'now';
```

- e. 退出 `psql`：

```
postgres=> \q
```

- f. 删除测试数据库（除非你以后还想用它做其他测试）：

```
$ destroydb
```

21. 以 Postgres 超级用户账号（通常是 postgres）在后台运行 postmaster。不要从 root 账号运行 postmaster！

通常，你会想修改你的计算机，使其在每次启动时自动启动 postmaster。这不是必需的；Postgres 服务器可以在没有 root 干预的情况下成功地以非特权账号运行。

下面是一些关于如何做到这一点的建议，由各位用户贡献。

无论你怎么做，postmaster 都必须由 Postgres 超级用户（postgres?）而不是 root 运行。这就是下面所有示例都先切换用户（su）到 postgres 的原因。这些命令还考虑到了 PATH 和 PGDATA 等环境变量可能没有被正确设置的事实。示例如下。请格外谨慎地使用它们。

- 如果你是从无特权账号安装且没有 root 权限，那么启动 postmaster 并把它送到后台：

```
$ cd
$ nohup postmaster > regress.log 2>&1 &
```

- 在 NetBSD 上编辑文件 rc.local，或在 SPARC Solaris 2.5.1 上编辑文件 rc2.d，使其包含下面这一行：

```
su postgres -c "/usr/local/pgsql/bin/postmaster -S -D /usr/local/pgsql/data"
```

- 在 FreeBSD 2.2-RELEASE 中编辑 /usr/local/etc/rc.d/pgsql.sh 使其包含下面的行，并对它执行 chmod 755 和 chown root:bin。

```
#!/bin/sh
[ -x /usr/local/pgsql/bin/postmaster ] && {
    su -l pgsql -c 'exec /usr/local/pgsql/bin/postmaster
        -D/usr/local/pgsql/data
        -S -o -F > /usr/local/pgsql/errlog' &
    echo -n ' pgsql'
}
```

你可以像上面那样放置换行。只要表达式未结束，shell 就足够聪明，会继续解析越过行尾的内容。exec 在 postmaster 进程之下节省了一层 shell，因此父进程是 init。

- 在 RedHat Linux 中添加一个文件 /etc/rc.d/init.d/postgres.init，它基于 contrib/linux/ 中的示例。然后从 /etc/rc.d/rc5.d/S98postgres.init 做一个指向该文件的软链接。
- 在 RedHat Linux 中编辑文件 /etc/inittab，把下面的内容作为单独一行加入：

```
pg:2345:respawn:/bin/su - postgres -c
    "/usr/local/pgsql/bin/postmaster -D/usr/local/pgsql/data
    >> /usr/local/pgsql/server.log 2>&1 </dev/null"
```

（这个例子的作者说该例会在 postmaster 死掉时复活它，但他不知道是否有其他副作用。）

22. 运行回归测试。文件 /usr/src/pgsql/src/test/regress/README 有运行和解释回归测试的详细说明。这里给出一个简短的版本：

- 输入

```
$ cd /usr/src/pgsql/src/test/regress
$ gmake clean
$ gmake all runtest
```

如果这是你第一次运行测试，不需要输入 `gmake clean`。

你应该在屏幕上（同时也写入文件 `./regress.out`）看到一系列语句，说明哪些测试通过、哪些测试失败。请注意，某些测试在某些平台上"失败"可能是正常的。只要测试的实际输出与期望输出之间有任何差异，脚本就说该测试失败了。因此，由于你的系统与回归测试参考平台之间错误消息措辞的细微差别、浮点舍入的小差异等，测试可能"失败"。这类"失败"并不表示 Postgres 有问题。文件 `./regression.diffs` 包含你机器上实际测试输出与"期望"输出（即参考系统产生的输出）之间的文本差异。你应当仔细检查列出的每一处差异，看它是否看起来是一个重大问题。

例如，

- 对于 i686/Linux-ELF 平台，没有测试失败，因为它是 v6.4 回归测试的参考平台。
- 对于 SPARC/Linux-ELF 平台，使用 Postgres v6.2 的 970525 beta 版时下列测试"失败"：float8 和 geometry 因浮点数的微小精度差异而"失败"。select\_views 产生大量不同的输出，但差异源于微小的浮点差异。

即使某个测试结果明确表明是真正的失败，它也可能是一个不影响你的局部问题。一个例子是：如果你的机器和 C 编译器不提供 64 位整数数据类型（或者虽然提供但 `configure` 没有发现），int8 测试会失败并产生明显错误的输出。除非你需要存储 64 位整数，否则这不需要担心。

结论？如果你确实看到了失败，试着理解这些差异的性质，然后判断这些差异是否会影响你对 Postgres 的预期用途。回归测试是一个有用的工具，但要发挥用处可能需要一些研究。

运行回归测试之后，输入

```
$ destroydb regression
$ cd /usr/src/pgsql/src/test/regress
$ gmake clean
```

以回收测试占用的磁盘空间。（做这件事之前，你可能想把 `regression.diffs` 文件保存到别处。）

23. 如果你还没有这样做，现在是修改你的计算机进行定期维护的好时机。下面的事情应当定期进行：

### 最小备份过程

1. 运行 SQL 命令 `VACUUM`。这会清理你的数据库。
2. 备份你的系统。（你或许应该保留最近几次的备份。）最好在此期间没有其他人使用系统。

理想情况下，上述任务应该由一个 shell 脚本完成，由 cron 每夜或每周运行。可以看看 crontab 的手册页作为入门。（如果你这样做了，请把你的 shell 脚本用电子邮件发给我们一份。我们也在自己的系统上这样做。）

24. 如果你是在升级现有系统，则重新装入你的旧数据库。输入

```
$ cd
$ psql -e template1 < db.out
```

如果你的 pre-v6.2 数据库使用了 path 或 polygon 几何数据类型，那么你需要升级包含这些类型的列。为此，输入（在 psql 内）

```
UPDATE FirstTable SET PathCol = UpgradePath(PathCol);
UPDATE SecondTable SET PathCol = UpgradePath(PathCol);
...
VACUUM;
```

UpgradePath() 检查一个 path 值是否与旧语法一致，不会更新未通过该检查的列。UpgradePoly() 无法验证一个 polygon 是否确实来自旧语法，但提供了 RevertPoly() 来撤销一次误用的升级。

25. 如果你是新用户，可以像下面描述的那样玩一玩 Postgres。

26. 清理你留下的东西。输入

```
$ rm -rf /usr/src/pgsql_6_0
$ rm -rf /usr/local/pgsql_6_0
# Also delete old database directory tree if it is not in
# /usr/local/pgsql_6_0/data
$ rm ~/postgresql-v6.2.1.tar.gz
```

27. 你可能想打印文档。如果你有 Postscript 打印机，或者你的机器已经用打印过滤器设置为可以接受 Postscript 文件，那么要打印用户指南只需输入

```
$ cd /usr/local/pgsql/doc
$ gunzip user.ps.tz | lpr
```

下面是如果你的系统上有 Ghostscript 并且向激光打印机打印时的做法。

```
$ alias gshp='gs -sDEVICE=laserjet -r300 -dNOPAUSE'
$ export GS_LIB=/usr/share/ghostscript:/usr/share/ghostscript/
fonts
$ gunzip user.ps.gz
$ gshp -sOUTPUTFILE=user.hp user.ps
$ gzip user.ps
$ lpr -l -s -r manpage.hp
```

28. Postgres 团队希望让 Postgres 在所有受支持的平台上都能工作。因此我们请你告诉我们你在你的系统上有没有让 Postgres 跑起来。请发邮件到 [pgsql-ports@postgresql.org](mailto:pgsql-ports@postgresql.org)<sup>5</sup> 告诉我们以下内容：

- Postgres 的版本（v6.4、6.3.2、beta 981014 等）。
- 你的操作系统（例如 RedHat v5.1 Linux v2.0.34）。

<sup>5</sup><mailto:pgsql-ports@postgresql.org>

- 你的硬件（SPARC、i486 等）。
  - 你是否干净地编译、安装并运行了回归测试？如果没有，你改了哪些源代码（即你应用的补丁、你做的修改等）、哪些测试失败了等。编译时出现许多警告是正常的。这些不需要报告。
29. 现在按需要创建、访问和操纵数据库。编写访问数据库服务器的客户端程序。换句话说，尽情享受！

## 24.3. 玩转 Postgres

在 Postgres 安装完成、数据库系统创建好、postmaster 守护进程在运行、回归测试通过之后，你会想看看 Postgres 做点什么。这很容易。调用 Postgres 的交互界面 psql：

```
% psql template1
```

（psql 必须打开一个特定的数据库，但此时唯一存在的是 template1 数据库，它总是存在的。我们连接到它只是为了创建另一个数据库然后切换过去。）

psql 的响应是：

```
Welcome to the POSTGRES interactive sql monitor:
Please read the file COPYRIGHT for copyright terms of POSTGRES

type \? for help on slash commands
type \q to quit
type \g or terminate with semicolon to execute query
You are currently connected to the database: template1
```

```
template1=>
```

创建数据库 foo：

```
template1=> create database foo;
CREATEDB
```

（养成写上 SQL 分号的习惯。在看到分号或 "\g" 之前 psql 不会执行任何东西，而且分号是分隔多条语句所必需的。）

现在连接到新数据库：

```
template1=> \c foo
connecting to new database: foo
```

（"斜杠"命令不是 SQL，所以不用分号。用 \? 可以看到所有斜杠命令。）

再创建一个表：

```
foo=> create table bar (i int4, c char(16));
CREATE
```

然后查看新表：

```
foo=> \d bar

Table      = bar
```

| +-----+ |  | +-----+   |  |
|---------|--|-----------|--|
|         |  |           |  |
| Length  |  |           |  |
| Field   |  | Type      |  |
| +-----+ |  | +-----+   |  |
| i       |  | int4      |  |
| 4       |  |           |  |
| c       |  | (bp) char |  |
| 16      |  |           |  |
| +-----+ |  | +-----+   |  |
| +-----+ |  | +-----+   |  |

依此类推。你明白意思了。

## 24.4. 下一步

有问题？Bug？反馈？首先阅读目录 `/usr/src/pgsql/doc/` 中的文件。该目录中的 FAQ 可能特别有用。

如果 `Postgres` 在你的计算机上编译失败，那么填写文件 `/usr/src/pgsql/doc/bug.template` 中的表单，并邮寄到表单顶部指示的地址。

在网站 <http://www.postgresql.org> 上查看各种支持邮件列表的更多信息。

## 24.5. 移植说明

### 注意

在源码发行版的 `doc/` 目录中查找任何平台特定的 FAQ。对于某些移植，下面的说明可能已经过时。

### 24.5.1. Ultrix4.x

### 注意

最近没有关于 Ultrix 使用 Postgres 的报告。

你需要安装 `libdl-1.1` 包，因为 Ultrix 4.x 没有动态加载器。它可从 `s2k-ftp.CS.Berkeley.EDU:pub/personal/andrew/libdl-1.1.tar.Z` 获取

### 24.5.2. Linux

#### 24.5.2.1. Linux ELF

Thomas G. Lockhart

回归测试参考机器是一台运行 `linux-2.0.30/libc-5.3.12/RedHat-4.2` 的双处理器 i686。linux-elf 移植可以干净地安装。更多细节见 Linux FAQ。

### 24.5.2.2. Linux a.out

对于非 ELF 的 Linux, 必须获取 GNU dld 库并安装到系统上。它为 Postgres 移植启用动态链接加载能力。dld 库可以从 [sunsite linux](#) 发行版获取。当前名称是 dld-3.2.5。Jalon Q. Zimmerman<sup>6</sup>

### 24.5.3. BSD/OS

对于 BSD/OS 2.0 和 2.01, 你需要获取 GNU dld 库。

### 24.5.4. NeXT

v1.09 的 NeXT 移植由 Tom R. Hageman<sup>7</sup> 提供。它需要一个 SysV IPC 模拟库以及用于共享库和信号量功能的头文件。Tom 恰好销售这样的产品, 可以联系他获取信息。他还表示将向公众提供 Postgres 的 NEXTSTEP 二进制发行版。联系 [Info@RnA.nl](mailto:Info@RnA.nl) 获取信息。

我们没有关于 NeXT 成功安装的最新报告 (截至 v6.2.1)。不过, 即使后端不受支持, 客户端库也应该能工作。

---

<sup>6</sup>[sneaker@powergrid.electriciti.com](mailto:sneaker@powergrid.electriciti.com)

<sup>7</sup>[mailto:tom@basil.icce.rug.nl](mailto:mailto:tom@basil.icce.rug.nl)

---

# 第 25 章 运行时环境

本章概述 Postgres 与操作系统的交互。

## 25.1. 在 Unix 上使用 Postgres

所有直接从 Unix shell 执行的 Postgres 命令都位于目录 “.../bin” 中。把这个目录包含在你的搜索路径中会让执行这些命令更容易。

每个站点都存在一组系统目录。其中包括一个类 (pg\_user)，它为每个有效的 Postgres 用户包含一个实例。该实例指定了一组 Postgres 权限，例如充当 Postgres 超级用户的能力、创建/删除数据库的能力以及更新系统目录的能力。在把适当的实例安装到这个类中之前，Unix 用户无法用 Postgres 做任何事。有关系统目录的更多信息可以通过在相应的类上运行查询获得。

## 第 26 章 系统布局

图 26.1. Postgres 文件布局

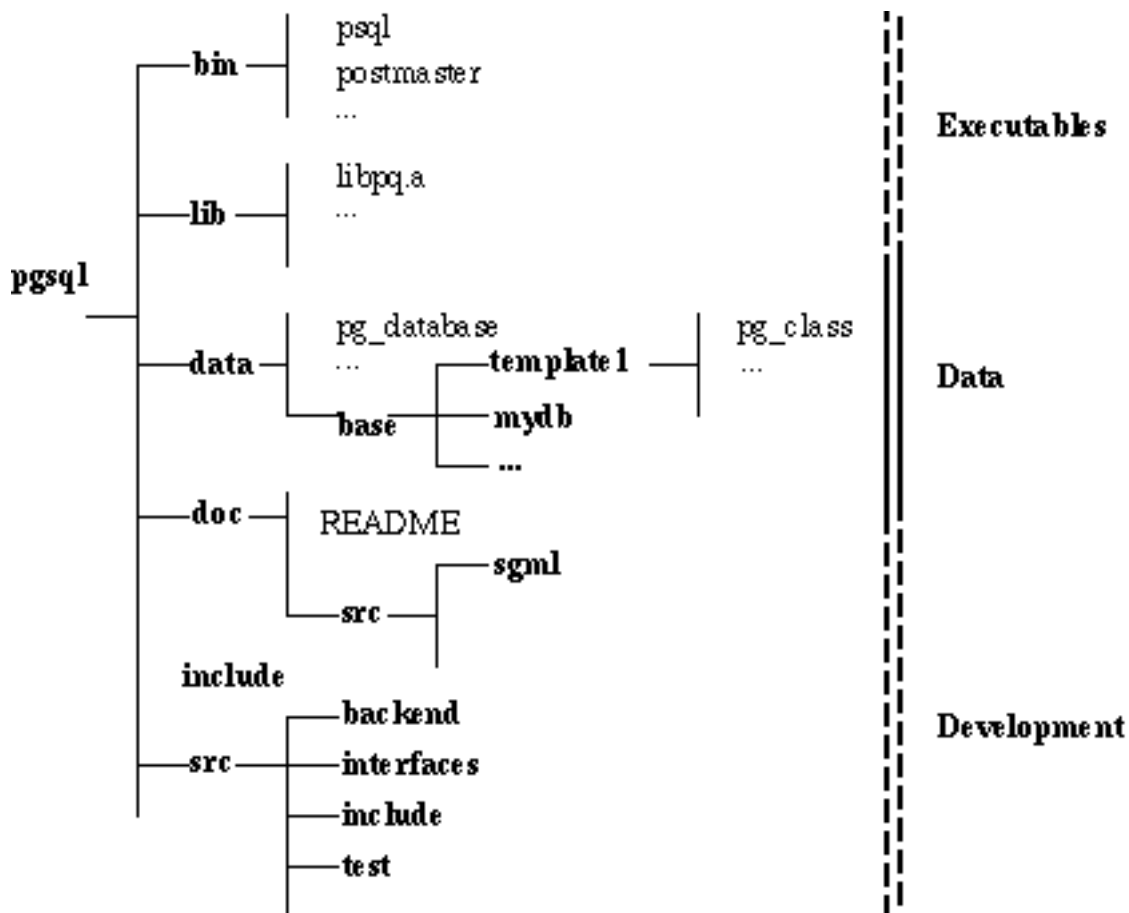


图 26.1 展示以默认方式安装时 Postgres 发行版的布局。为简单起见，我们假设 Postgres 安装在目录 `/usr/local/pgsql`。因此，凡是你看到目录 `/usr/local/pgsql` 之处，都应替换为 Postgres 实际安装所在目录的名称。所有 Postgres 命令都安装在目录 `/usr/local/pgsql/bin` 中。因此，你应当把该目录加入你的 shell 命令路径。如果你使用 Berkeley C shell 的变体，例如 `csh` 或 `tcsh`，可以在你主目录的 `.login` 文件中加入

```
set path = ( /usr/local/pgsql/bin path )
```

如果你使用 Bourne shell 的变体，例如 `sh`、`ksh` 或 `bash`，则可以在你主目录的 `.profile` 文件中加入

```
PATH=/usr/local/pgsql/bin PATH
export PATH
```

从现在起，我们假设你已把 Postgres 的 `bin` 目录加入了你的路径。此外，本文档会频繁提到“设置 shell 变量”或“设置环境变量”。如果你没有完全理解上一段关于修改搜索路径的内容，在继续之前应当查阅描述你所用的 shell 的 UNIX 手册页。

如果你没有按默认方式设置，可能还有更多工作要做。例如，如果数据库服务器机器是一台远程机器，你需要把 `PGHOST` 环境变量设置为数据库服务器机器的

名称。环境变量 `PGPORT` 可能也需要设置。底线是：如果你试图启动一个应用程序而它抱怨无法连接到 `postmaster`，就必须回过头确认你的环境已正确设置。

---

# 第 27 章 使用 pg\_options

Massimo Dal Zotto

## 注意

由 Massimo Dal Zotto<sup>1</sup> 供稿

可选文件 `data/pg_options` 包含后端使用的运行时选项，用于控制跟踪消息和其他后端可调参数。这个文件的有趣之处在于，后端收到 `SIGHUP` 信号时会重新读取它，因此无需重启 `Postgres` 就可以即时更改运行时选项。该文件中指定的选项可以是跟踪包（`backend/utils/misc/trace.c`）使用的调试标志，也可以是后端用来控制其行为的数值参数。新的选项和参数必须在 `backend/utils/misc/trace.c` 和 `backend/include/utils/trace.h` 中定义。

例如，假设我们想为文件 `foo.c` 中的代码添加条件跟踪消息和一个可调的数值参数。我们只需把常量 `TRACE_FOO` 和 `OPT_FOO_PARAM` 加入 `backend/include/utils/trace.h`：

```
/* file trace.h */
enum pg_option_enum {
    ...
    TRACE_FOO,      /* trace foo functions */
    OPT_FOO_PARAM,   /* foo tunable parameter */

    NUM_PG_OPTIONS      /* must be the last item of enum */
};
```

并在 `backend/utils/misc/trace.c` 中加入对应的一行：

```
/* file trace.c */
static char *opt_names[] = {
    ...
    "foo",              /* trace foo functions */
    "fooparam"          /* foo tunable parameter */
};
```

两个文件中的选项必须以完全相同的顺序指定。现在在 `foo` 源文件中就可以这样引用新的标志：

```
/* file foo.c */
#include "trace.h"
#define foo_param pg_options[OPT_FOO_PARAM]

int
foo_function(int x, int y)
{
    TPRINTF(TRACE_FOO, "entering foo_function, foo_param=%d", foo_param);
    if (foo_param > 10) {
        do_more_foo(x, y);
    }
}
```

---

<sup>1</sup>[mailto:dz@cs.unitn.it](mailto:mailto:dz@cs.unitn.it)

```
}
```

使用私有跟踪标志的现有文件只需加上以下代码即可改造：

```
#include "trace.h"
/* int my_own_flag = 0; -- removed */
#define my_own_flag pg_options[OPT_MY_OWN_FLAG]
```

所有 `pg_options` 在后端启动时都被初始化为零。如果需要不同的默认值，必须在 `Postgres Main` 开头添加一些初始化代码。现在我们可以向 `data/pg_options` 文件写入值来设置 `foo_param` 并启用 `foo` 跟踪：

```
# file pg_options
...
foo=1
fooparam=17
```

新选项会在每个新后端启动时被读取。要让更改对所有正在运行的后端生效，需要向 `postmaster` 发送 `SIGHUP`。该信号会自动被发送到所有后端。也可以直接向某个特定后端发送 `SIGHUP`，只对它激活更改。

`pg_options` 也可以通过 `Postgres` 的 `-T` 开关来指定：

```
postgres options -T "verbose=2,query,hostlookup="
```

用于打印错误和调试消息的函数现在可以使用 `syslog(2)` 设施。打印到 `stdout` 或 `stderr` 的消息都带有一个时间戳前缀，其中还包含后端 `pid`：

```
#timestamp      #pid      #message
980127.17:52:14.173 [29271] StartTransactionCommand
980127.17:52:14.174 [29271] ProcessUtility: drop table t;
980127.17:52:14.186 [29271] SIIncNumEntries: table is 70% full
980127.17:52:14.186 [29286] Async_NotifyHandler
980127.17:52:14.186 [29286] Waking up sleeping backend process
980127.19:52:14.292 [29286] Async_NotifyFrontEnd
980127.19:52:14.413 [29286] Async_NotifyFrontEnd done
980127.19:52:14.466 [29286] Async_NotifyHandler done
```

这种格式提高了日志的可读性，让人能准确了解哪个后端在什么时间做了什么。它也让编写简单的 `awk` 或 `perl` 脚本来监控日志、检测数据库错误或问题、或计算事务时间统计变得更容易。

打印到 `syslog` 的消息使用日志设施 `LOG_LOCAL0`。是否使用 `syslog` 可以用 `syslog pg_option` 控制。遗憾的是，许多函数直接调用 `printf()` 把消息打印到 `stdout` 或 `stderr`，这样的输出既无法重定向到 `syslog`，也无法加上时间戳。最好把所有对 `printf` 的调用都替换为 `PRINTF` 宏，并把输出到 `stderr` 的调用改为使用 `EPRINTF`，这样我们就能以统一的方式控制所有输出。

新的 `pg_options` 机制比定义新的后端选项开关更方便，原因如下：

- 我们不必为想要控制的每样东西定义不同的开关。  
所有选项都定义为存储在数据目录中一个外部文件里的关键字。
- 我们不必为了更改某个选项的设置而重启 `Postgres`。通常后端选项是指定给 `postmaster` 并在每个后端启动时传递给它的。现在它们从文件中读取。
- 我们可以在后端运行时即时更改选项。  
这样我们就可以只在问题出现时才激活调试消息来研究问题。  
我们还可以为可调参数尝试不同的值。

The format of the `pg_options` file is as follows:

```
# comment
option=integer_value # set value for option
option                # set option = 1
option+               # set option = 1
option-               # set option = 0
```

Note that *keyword* can also be an abbreviation of the option name defined in `backend/`  
`utils/misc/trace.c`.

目前定义在 `backend/utils/misc/trace.c` 中的选项如下：

**all**

全局跟踪标志。允许的值有：

0

单独启用各条跟踪消息

1

启用所有跟踪消息

-1

禁用所有跟踪消息

**verbose**

详细程度标志。允许的值有：

0

不打印消息。这是默认值。

1

打印信息消息。

2

打印更多信息消息。

**query**

查询跟踪标志。允许的值有：

0

不打印查询。

1

以一行打印压缩形式的查询。

4

打印完整查询。

plan

打印查询计划。

parse

打印解析器输出。

rewritten

打印重写后的查询。

parserstats

打印解析器统计信息。

plannerstats

打印规划器统计信息。

executorstats

打印执行器统计信息。

shortlocks

目前未用，但将来启用特性时需要。

locks

跟踪锁。

userlocks

跟踪用户锁。

spinlocks

跟踪自旋锁。

notify

跟踪通知函数。

malloc

目前未用。

palloc

目前未用。

lock\_debug\_oidmin

被锁跟踪的最小关系 oid。

lock\_debug\_relid

被锁跟踪的关系 oid（非零时）。

lock\_read\_priority

目前未用。

deadlock\_timeout

死锁检查计时器。

syslog

syslog 标志。允许的值有：

0

消息输出到 stdout/stderr。

1

消息输出到 stdout/stderr 和 syslog。

2

消息只输出到 syslog。

hostlookup

在 ps\_status 中启用主机名查找。

showportnumber

在 ps\_status 中显示端口号。

notifyunlock

notify 之后解锁 pg\_listener。

notifyhack

从 pg\_listener 中移除重复元组。

例如我的 pg\_options 文件包含下列值：

```
verbose=2
query
hostlookup
showportnumber
```

一些使用私有变量和选项开关的现有代码已被改为使用 pg\_options 特性，主要是在 postgres.c 中。最好把所有现有代码都改成这种方式，这样我们就可以去掉 Postgres 命令行上的许多开关，并拥有更多可调选项，而选项值集中放在一个地方。

---

## 第 28 章 启动 postmaster

除非 `postmaster` 进程在运行，否则数据库什么也做不了。作为站点管理员，在启动 `postmaster` 之前，有一些事情你应当记住。这些内容在本手册的安装和配置章节中讨论。不过，如果 Postgres 是严格按照安装说明安装的，下面的简单命令就是你启动 `postmaster` 所需的全部：

```
% postmaster
```

`postmaster` 偶尔会打印出一些在排障时往往有帮助的消息。如果你希望查看来自 `postmaster` 的调试消息，可以用 `-d` 选项启动它并把输出重定向到日志文件：

```
% postmaster -d >& pm.log &
```

如果你不希望看到这些消息，可以输入

```
% postmaster -S
```

这样 `postmaster` 就是 "Silent"（静默）的。注意最后一个例子的末尾没有 `and` 符（`&`）。

---

## 第 29 章 添加和删除用户

`createuser` 使特定的用户能够访问 Postgres。`destroyuser` 移除用户并阻止他们访问 Postgres。注意，这些命令只影响与 Postgres 相关的用户；它们对用户底层操作系统方面的其他权限或地位没有影响。

---

# 第 30 章 磁盘管理

## 30.1. 备选位置

可以在安装的默认位置之外的位置创建数据库。请记住，所有数据库访问实际上都是通过数据库后端进行的，因此所指定的任何位置都必须是后端可以访问的。

可以指定一个绝对路径名或者一个环境变量作为位置。注意，出于安全和完整性原因，所有这样指定的路径和环境变量都会被追加一些额外的路径字段。

### 注意

环境变量风格的指定方式更可取，因为它让站点管理员在管理磁盘存储时有更大的灵活性。

请记住，数据库的创建实际上是由数据库后端执行的。因此，任何指定备选位置的环境变量都必须在后端启动之前已经定义。要定义一个指向 `/home/postgres/data` 的备选位置 `PGDATA2`，输入

```
% setenv PGDATA2 /home/postgres/data
```

通常，你会想在 Postgres 超级用户的 `.profile` 或 `.cshrc` 初始化文件中定义这个变量，以确保它在系统启动时被定义。

要在 `/home/postgres/data` 中创建一个数据存储区，请确保 `/home/postgres` 已经存在并且可写。然后在命令行输入

```
% initlocation $PGDATA2
Creating Postgres database system directory /home/postgres/data

Creating Postgres database system directory /home/postgres/data/base
```

要测试新位置，输入下列命令创建一个数据库 `test`

```
% createdb -D PGDATA2 test
% destroydb test
```

---

## 第 31 章 排障

假设你的站点管理员已经正确地启动了 `postmaster` 进程并且授权你使用数据库，你（作为用户）就可以开始启动应用程序了。如前所述，你应当把 `/usr/local/pgsql/bin` 加到你的 `shell` 搜索路径中。在大多数情况下，这就是你在准备方面要做的全部。

如果你从一个 Postgres 命令（例如 `psql` 或 `createdb`）得到下列错误消息：

```
connectDB() failed: Is the postmaster running at 'localhost' on port
'5432'?
```

这通常要么是因为 `postmaster` 没有在运行，要么是你试图连接到了错误的服务器主机。如果你得到下列错误消息：

```
FATAL 1:Feb 17 23:19:55:process userid (2360) != database owner (268)
```

则意味着站点管理员以错误的用户启动了 `postmaster`。让他以 Postgres 超级用户身份重启它。

---

## 第 32 章 管理数据库

现在 Postgres 已经启动并运行，我们可以创建一些数据库来做实验。这里我们描述管理数据库的基本命令。

### 32.1. 创建数据库

假设你想创建一个名为 mydb 的数据库。你可以使用下面的命令完成：

```
% createdb mydb
```

Postgres 允许你在给定站点创建任意数量的数据库，并且你会自动成为你刚创建的数据库的数据库管理员。数据库名的第一个字符必须是字母，并且长度限制为 16 个字符。并非每个用户都有权成为数据库管理员。如果 Postgres 拒绝为你创建数据库，那么站点管理员就需要授予你创建数据库的权限。如果发生这种情况，请咨询你的站点管理员。

### 32.2. 访问数据库

一旦构造好数据库，就可以通过以下方式访问它：

- 运行 Postgres 终端监视程序（psql），它允许你交互式地输入、编辑和执行 SQL 命令。
- 使用 libpq 子程序库编写 C 程序。这允许你从 C 中提交 SQL 命令，并把答案和状态消息返回给你的程序。这一接口在 PostgreSQL 程序员指南中有进一步讨论。

你也许想启动 psql，来试试本手册中的例子。可以通过输入以下命令为 mydb 数据库激活它：

```
% psql mydb
```

你会看到下面的欢迎消息：

```
Welcome to the Postgres interactive sql monitor:
```

```
type \? for help on slash commands
type \q to quit
type \g or terminate with semicolon to execute query
You are currently connected to the database: mydb
```

```
mydb=>
```

这个提示表明终端监视程序正在监听你，你可以把 SQL 查询输入到由终端监视程序维护的工作区中。psql 程序响应以反斜杠字符 “\” 开头的转义码。例如，你可以通过输入下列内容获得各种 Postgres SQL 命令的语法帮助：

```
mydb=> \h
```

一旦在工作区中输入完查询，你可以通过输入下列内容把工作区的内容传递给 Postgres 服务器：

```
mydb=> \g
```

这告诉服务器处理该查询。如果你用分号终止查询，backslash-g 就不是必需的。psql 会自动处理以分号终止的查询。要从文件（例如 myFile）中读取查询而不是交互式地输入，输入：

```
mydb=> \i fileName
```

要退出 psql 并返回 UNIX, 输入

```
mydb=> \q
```

然后 psql 将退出并把你返回到命令 shell。(关于更多转义码, 在监视程序提示符下输入 backslash-h。)空白(即空格、制表符和换行符)可以在 SQL 查询中自由使用。单行注释用 "--" 表示。连字符之后直到行尾的所有内容都被忽略。多行注释以及行内注释用 "/\* ... \*/" 表示

## 32.3. 删除数据库

如果你是数据库 mydb 的数据库管理员, 你可以使用下面的 UNIX 命令销毁它:

```
% destroydb mydb
```

这个动作会物理地删除与该数据库关联的所有 UNIX 文件, 而且无法撤销, 因此只应在深思熟虑之后才进行。

---

# 第 33 章 数据库恢复

本节尚待编写。有人愿意吗？

---

# 第 34 章 回归测试

回归测试的说明与分析。

PostgreSQL 回归测试是对由 Jolly Chen 和 Andrew Yu 开发的、嵌入在 PostgreSQL 中的 SQL 实现的一套全面测试。它既测试标准 SQL 操作，也测试 PostgreSQL 的扩展能力。

这些测试最近由 Marc Fournier 和 Thomas Lockhart 修订，现在被打包为功能单元，这应当使它们更易于运行、更易于解读。从 PostgreSQL v6.1 起，回归测试对每个正式发行版都是最新的。

某些安装正确且功能完备的 PostgreSQL 安装可能会因浮点表示和时区支持的特有问题而在其中一些回归测试上失败。这些测试目前是用一种简单的 "diff" 算法来评估的，对细微的系统差异很敏感。对于表面上失败的测试，检查差异可能会发现这些差异并不重要。

下面的回归测试说明假定以下条件（除非另有说明）：

- 命令是 Unix 兼容的。见下面的说明。
- 除非另有说明，均使用默认值。
- 用户 postgres 是 Postgres 超级用户。
- 源代码路径是 /usr/src/pgsql（其他路径也是可能的）。
- 运行时路径是 /usr/local/pgsql（其他路径也是可能的）。

## 34.1. 回归测试环境

回归测试由 `make` 命令调用，该命令把一个 C 程序编译成当前目录中的一个共享库。当前目录中还会创建本地化的 shell 脚本。输出文件模板被整理成 `./expected/*.out` 文件。本地化会把源文件中的宏替换为绝对路径名和用户名。

通常，回归测试应当以 `pg_superuser` 运行，因为 `'src/test/regress'` 目录及其子目录为 `pg_superuser` 所有。如果你以其他用户运行回归测试，那么 `'src/test/regress'` 目录树应当可被该用户写入。

`postmaster` 应当在系统时区设置为加利福尼亚州伯克利的情况下调用。回归测试脚本会自动完成这项设置。但是，它确实要求机器支持 PST8PDT 时区。

要验证你的机器确实支持这一点，请键入：

```
setenv TZ PST8PDT
date
```

上面的 "date" 命令应当以 PST8PDT 时区返回当前系统时间。如果 PST8PDT 数据库不可用，你的系统可能返回的是 GMT 时间。如果 PST8PDT 时区不可用，你可以显式设置时区规则：

```
setenv PGTZ PST8PDT7,M04.01.0,M10.05.03
```

## 34.2. 目录布局

### 注意

这里以后应当变成上一节中的一个表。

```
input/ .... .source files that are converted using 'make all' into
           some of the .sql files in the 'sql' subdirectory
```

```

output/ ... .source files that are converted using 'make all' into
           .out files in the 'expected' subdirectory

sql/ ..... .sql files used to perform the regression tests

expected/ . .out files that represent what we *expect* the results
to
           look like

results/ .. .out files that represent what the results *actually*
look
           like. Also used as temporary storage for table copy
testing.

```

## 34.3. 回归测试过程

这些命令是在使用 bash shell 的 RedHat Linux 4.2 版上测试的。除非另有说明，它们在大多数系统上可能都能工作。像 `ps` 和 `tar` 这样的命令在各个平台上应使用的选项差异极大。键入这些命令之前请运用常识。

### Postgres 回归测试配置

对于全新安装或从先前版本的 Postgres 升级：

1. 构建回归测试。键入

```

cd /usr/src/pgsql/src/test/regress
gmake all

```

2. (可选的) 如果你之前调用过回归测试，用以下命令清理工作目录：

```

cd /usr/src/pgsql/src/test/regress
make clean

```

3. 文件 `/usr/src/pgsql/src/test/regress/README` 包含关于运行和解读回归测试的详细说明。这里给出一个简短的版本：

如果 `postmaster` 还没有运行，在一个可用的窗口中键入以下命令启动 `postmaster`

```
postmaster
```

或者键入以下命令让 `postmaster` 守护进程在后台运行

```

cd
nohup postmaster > regress.log 2>&1 &

```

从你的 Postgres 超级用户账户（通常是 `postgres` 账户）运行 `postmaster`。

### 注意

不要从 `root` 账户运行 `postmaster`。

#### 4. 运行回归测试。键入

```
cd /usr/src/pgsql/src/test/regress
gmake runtest
```

如果这是你第一次运行测试，不需要键入 "gmake clean"。

5. 你应当在屏幕上（同时也写入文件 `./regress.out`）看到一系列说明哪些测试通过、哪些测试失败的语句。请注意，某些测试"失败"可能是正常的。对于失败的测试，用 `diff` 比较 `./results` 和 `./expected` 目录中的文件。如果 `float8` 失败了，键入类似下面的命令：

```
cd /usr/src/pgsql/src/test/regress
diff -w expected/float8.out results
```

#### 6. 运行测试后，键入

```
destroydb regression
cd /usr/src/pgsql/src/test/regress
gmake clean
```

## 34.4. 回归分析

"失败"的测试可能是由于错误消息、数学库或输出格式的细微差异而失败。这类"失败"并不表示 Postgres 有问题。

对于 i686/Linux-ELF 平台，没有测试失败，因为它是 v6.2.1 回归测试的参考平台。

对于 SPARC/Linux-ELF 平台，使用 Postgres v6.2 的 970525 beta 版时下列测试"失败"：`float8` 和 `geometry` 因浮点数的微小精度差异而"失败"。`select_views` 产生大量不同的输出，但差异源于微小的浮点差异。

结论是什么？如果你确实看到失败，请设法理解这些差异的本质，然后判断这些差异是否会影响你对 Postgres 的预期使用。不过请记住，这很可能是迄今为止最坚实的 Postgres 发行版，它合并了来自 v6.1 的许多缺陷修复，而且先前版本的 Postgres 已经成功使用了一段时间。

### 34.4.1. 比较期望/实际输出

结果位于 `./results` 目录中的文件里。这些结果可以用 'diff' 与 `./expected` 目录中的结果进行比较。这些文件可能不会完全一致。下面几段试图解释这些差异。

### 34.4.2. 错误消息差异

某些回归测试包含故意构造的非法输入值。错误消息可能来自 Postgres 代码，也可能来自宿主平台的系统例程。在后一种情况下，消息会因平台而异，但应反映相近的信息。这类消息差异会导致回归测试显示为"失败"，但可以通过检查确认其有效性。

### 34.4.3. OID 差异

PostgreSQL OID（对象标识符）出现在 'regress.out' 中的若干地方。OID 是唯一的 32 位整数，由 PostgreSQL 后端在每次插入或更新表行时生成。如果你在一个非全新的数据库上运行回归测试，或者多次运行它，报告的 OID 将具有不同的值。'misc.out' 中的下列 SQL 语句表现出了这种行为：`QUERY: SELECT user_relns() AS user_relns ORDER BY user_relns; 'a,523676'` 行就是由一个 OID 组成的。

### 34.4.4. 日期和时间差异

在许多受支持的平台上，你可以让 PostgreSQL 认为自己运行在与加利福尼亚州伯克利相同的时区中。详见关于如何运行回归测试的那一节。如果你不把时区环境显式设置为 PST8PDT，那么大多数日期和时间结果将反映你的本地时区，并且无法通过回归测试。

有些系统似乎不接受显式设置本地时区规则的推荐语法。某些使用公共域时区包的系统在 1970 年前的 PDT 时间上表现出小问题，把它们表示为 PST。

### 34.4.5. 浮点差异

某些测试涉及从表列计算 64 位 (float8) 数。已经观察到涉及 float8 列数学函数的结果存在差异。这些差异出现在同一平台上使用不同操作系统的情况，即 Intel/86 上的 BSDI 和 SOLARIS；以及同一操作系统用于不同平台的情况，即 SPARC 和 Intel/86 上的 SOLARIS。需要人工目测比较来确定这些差异的真实意义，它们通常出现在小数点右侧第 10 位。某些系统发出 pow() 和 exp() 错误的方式与当前 Postgres 代码所预期的机制不同。

### 34.4.6. 多边形差异

若干测试涉及对加利福尼亚州奥克兰/伯克利街道地图地理数据的操作。

地图数据被表示为多边形，其顶点用一对 float8 数（十进制纬度和经度）表示。首先创建一些表并装入地理数据，然后创建一些使用多边形相交操作符 (##) 连接两个表的视图，再对视图做一次 select。在比较不同平台的结果时，差异出现在小数点右侧第 2 或第 3 位。出现这些问题的 SQL 语句是：

```
QUERY: SELECT * from street;
QUERY: SELECT * from iexit;
```

### 34.4.7. 随机差异

random.out 中至少有一个测试用例旨在产生随机结果。这会使 random 回归测试失败。键入

```
diff results/random.out expected/random.out
```

因此应当只产生一行或几行差异，但不同体系结构上的其他浮点差异可能造成多得多的差异。见下面的发行说明。

### 34.4.8. “expected” 文件

./expected/\*.out 文件改编自 Jolly Chen 等人提供的原始整体式 expected.input 文件。在各种开发机器上生成的这些文件的新版本经过仔细 (?) 检查后被替换了进来。许多开发机器运行的是 i386 硬件上的某种 Unix 变体 (FreeBSD、Linux 等)。原始的 expected.input 文件是在一台 SPARC Solaris 2.4 系统上使用 postgres5-1.02a5.tar.gz 源树创建的。它与一台 i386 Solaris 2.4 系统上创建的文件比较后，差异仅在于浮点多边形小数点右侧第 3 位。（见下面）原始的 sample.regress.out 文件来自 Jolly Chen 构建的 postgres-1.01 发行版，这里收录仅供参考。它可能是在一台 DEC ALPHA 机器上创建的，因为 postgres-1.01 发行版中的 Makefile.global 有 PORTNAME=alpha。

---

# 第 35 章 发行说明

## 35.1. 版本 6.4

此版本有许多新特性和改进。感谢我们的开发者和维护者，自上一个版本以来系统的几乎每个方面都得到了关注。以下是简要而不完整的总结：

- 得益于 Jan Wieck 在重写规则系统中的大量新代码，视图和规则现在可以工作了。他还为程序员指南写了相关章节。
- Jan 还贡献了第二个过程语言 PL/pgSQL，与他在上个版本贡献的最初的 PL/pgTCL 过程语言相伴。
- 我们有 Tatsuo Ishii 提供的可选多字节字符集支持，以补充现有的区域设置支持。
- 客户端/服务器通信得到了清理，感谢 Tom Lane，对异步消息和中断的支持更好了。
- 解析器现在执行自动类型强制，以把参数匹配到可用的操作符和函数，并把列和表达式与目标列匹配。这使用了支持 Postgres 类型可扩展特性的通用机制。用户指南有涵盖此主题的新章节。
- 新增了三种数据类型。其中两种，inet 和 cidr，支持各种形式的 IP 网络、子网和机器寻址。某些平台上现在有 8 字节整数类型可用。详情见用户指南中的数据类型章节。第四种类型 serial 现在被解析器支持，作为 int4 类型、序列和唯一索引的混合体。
- 新增了更多 SQL92 兼容的语法特性，包括 INSERT DEFAULT VALUES
- 自动配置和安装系统得到了一些关注，应比以往在更多平台上更健壮。

### 35.1.1. 迁移到版本 6.4

希望从任何以前的 Postgres 版本迁移数据的用户，需要使用 pg\_dump 或 pg\_dumpall 进行转储/恢复。

### 35.1.2. 详细变更列表

#### 缺陷修复

-----

修复 PQsetdb/PQfinish 中的小内存泄漏 (Bryan)  
移除 char2-16 数据类型，使用 char/varchar (Darren)  
Pqfn 现在处理 NOTICE 消息 (Anders)  
降低多后端时自旋锁的忙等开销 (dg)  
卡死自旋锁检测 (dg)  
修复 "ISO 风格" timespan 的解码和编码 (Thomas)  
修复事务回滚后删除表的问题 (Vadim)  
更改错误消息并移除无效的更新消息 (Vadim)  
修复 COPY 数组检查的问题  
修复 SELECT 1 UNION SELECT NULL  
修复大对象调用中的缓冲区泄漏 (Pascal)  
把 owner 从 oid 类型改为 int4 (Bruce)  
修复 Oracle 兼容函数 btrim()、ltrim() 和 rtrim() 中的缺陷  
修复共享失效缓存溢出 (Massimo)  
防止失败的 COPY 中的文件描述符泄漏 (Bruce)  
修复 libpgtcl 的 pg\_select 中的内存泄漏 (Constantin)  
修复超过 8 字符的用户名/口令问题 (Tom)  
修复后端处理异步 NOTIFY 的问题 (Tom)  
修复许多错误的系统表条目 (Tom)

## 增强

-----

升级 `ecpg` 和 `ecpglib`, 参见 `src/interfaces/ecpg/ChangeLog` (Michael)

在 `EXPLAIN` 中显示使用的索引 (Zeugswetter)

`EXPLAIN` 调用规则系统并显示被重写查询的计划 (Jan)

通过 `configure` 使许多数据类型和函数感知多字节 (Tatsuo)

新 `configure --with-mb` 选项 (Tatsuo)

新 `initdb --pgencoding` 选项 (Tatsuo)

新 `createdb -E` 多字节选项 (Tatsuo)

`Select version()`; 现在返回 PostgreSQL 版本 (Jeroen)

`libpq` 现在允许异步客户端 (Tom)

允许从客户端取消后端查询 (Tom)

`psql` 现在可用 `Control-C` 取消查询 (Tom)

`libpq` 用户不必发送哑查询即可获取 `NOTIFY` 消息 (Tom)

`NOTIFY` 现在发送发送者的 `PID`, 可以判断是否是自己的 (Tom)

`PGresult` 结构现在包含相关的错误消息 (如果有) (Tom)

定义 `date_part()` 的 `"tz_hour"` 和 `"tz_minute"` 参数 (Thomas)

添加 `varchar` 和 `bpchar` 之间转换的例程 (Thomas)

添加允许把 `varchar` 和 `bpchar` 调整到目标列大小的例程 (Thomas)

添加位标志以支持数据检索中的时区小时和分钟 (Thomas)

允许有效浮点数的更多变体 (如 `".1"`、`"1e6"`) (Thomas)

修复带前导空格的一元负号解析 (Thomas)

按 SQL92 规范实现 `TIMEZONE_HOUR`、`TIMEZONE_MINUTE` (Thomas)

检查并正确忽略 `FOREIGN KEY` 列约束 (Thomas)

按 SQL92 规范定义 `USER` 为 `CURRENT_USER` 的同义词 (Thomas)

启用 `HAVING` 子句但其他地方尚无修复。

使 `"char"` 类型成为 `"char(1)"` 的同义词 (实际实现为 `bpchar`) (Thomas)

为 `DEFAULT` 子句处理保存指定的字符串类型 (Thomas)

强制涉及不同数据类型的操作 (Thomas)

允许对不同类型列的某些索引使用 (Thomas)

添加自动类型转换能力 (Thomas)

清理大对象, 使文件在打开时被截断 (Peter)

`Readline` 清理 (Tom)

允许 `psql \f \` 用空格作为分隔符 (Bruce)

把 `pg_attribute.atttypmod` 传递给前端以获取列字段长度 (Tom、Bruce)

`/contrib` 中的 `Msql` 兼容库 (Aldrin)

移除 `ORDER/GROUP BY` 子句标识符必须包含在目标列表中的要求 (David)

转换列以匹配 `UNION` 子句中的列 (Thomas)

移除 `fork()/exec()` 只做 `fork()` (Bruce)

`Jdbc` 清理 (Peter)

在 `ps` 命令行上显示后端状态 (只在某些平台上有效) (Bruce)

`Pg_hba.conf` 的 `database` 字段现在有 `sameuser` 选项

使 `lo_unlink` 接受 `oid` 参数而不是 `int4`

为无法处理我们宏的编译器新增 `DISABLE_COMPLEX_MACRO` (Bruce)

`Libpgtcl` 现在把 `NOTIFY` 作为 `Tcl` 事件处理, 不必发送哑查询 (Tom)

`libpgtcl` 清理 (Tom)

向 `libpgtcl` 的 `pg_result` 命令添加 `-error` 选项 (Tom)

新区域设置补丁, 参见 `docs/README/locale` (Oleg)

修复 `pg_dump` 使 `CONSTRAINT` 和 `CHECK` 语法正确 (ccb)

新 `contrib/lo` 代码用于移除孤立大对象 (Peter)

多字节特性的新 `psql` 命令 `"SET CLIENT_ENCODING TO 'encoding'"`, 参见 `/doc/README.mb` (Tatsuo)

`contrib/noupdate` 代码用于撤销列上的更新权限

libpq 现在可以在 win32 上编译 (Magnus)  
 在 libpq 中添加 PQsetdbLogin()  
 新 8 字节整数类型, 由 configure 检查操作系统支持 (Thomas)  
 更好地支持带引号的表/列名 (Thomas)  
 在 pg\_dump 中用双引号包围表和列名 (Thomas)  
 PQreset() 现在可用于口令 (Tom)  
 处理 GROUP BY 目标列表列号越界的情形 (David)  
 允许子查询中的 UNION  
 向 \d? 命令添加屏幕自动调整大小 (Bruce)  
 用 UNION 在一个查询中显示所有 \d? 结果 (Bruce)  
 添加 \d? 字段搜索特性 (Bruce)  
 Pg\_dump 发出更少的 \connect 请求 (Tom)  
 使 pg\_dump -z 标志更好地工作并在手册页中记录 (Tom)  
 添加对子查询和 union 完全支持的 HAVING 子句 (Stephan)  
 contrib/fulltextindex 中的全文索引例程 (Maarten)  
 事务 id 现在存储在共享内存中 (Vadim)  
 发出 COPY 命令时的新 PGCLIENTENCODING (Tatsuo)  
 支持 SQL92 语法 "SET NAMES" (Tatsuo)  
 支持 LATIN2-5 (Tatsuo)  
 添加 UNICODE 回归测试用例 (Tatsuo)  
 锁管理器清理, LLL 的新锁定模式 (Vadim)  
 允许 OR 子句使用索引 (Bruce)  
 允许 "SELECT NULL ORDER BY 1;"  
 Explain VERBOSE 打印计划, 现在还把计划美化打印到  
 postmaster 日志文件 (Bruce)  
 向 \d 命令添加索引显示 (Bruce)  
 允许对函数 GROUP BY (David)  
 大对象的新 pg\_class.relkind (Bruce)  
 把 libpq NOTICE 消息发送到其他位置的新方法 (Tom)  
 psql 的新 \w 写入命令 (Bruce)  
 新 /contrib/findoidjoins 扫描 oid 列以发现连接关系 (Bruce)  
 检查含常量的限制子句的有效索引时  
 允许考虑二进制兼容的索引 (Thomas)  
 /contrib/isbn\_issn 中的新 ISBN/ISSN 代码  
 允许 NOT LIKE、IN、NOT IN、BETWEEN 和 NOT BETWEEN 约束 (Thomas)  
 新重写系统修复了规则和视图的许多问题 (Jan)  
 \* 关系上的规则可用  
 \* insert/update/delete 上的事件限定可用  
 \* 新的 OLD 变量引用 CURRENT, CURRENT 未来将被移除  
 \* 更新规则可在规则限定/动作中引用 NEW 和 OLD  
 \* 视图上的 insert/update/delete 规则可用  
 \* 现在支持多个规则动作, 用括号包围  
 \* 普通用户可以在有 RULE 权限的表上创建视图/规则  
 \* 规则和视图继承创建者的权限  
 \* 没有列级别的规则  
 \* 没有 UPDATE NEW/OLD 规则  
 \* 新 pg\_tables、pg\_indexes、pg\_rules 和 pg\_views 系统视图  
 \* SELECT 规则只有单个动作  
 \* 彻底的重写改造, 也许在 6.5  
 \* 处理子查询  
 \* 处理视图上的聚合  
 \* 处理 insert into select from view 可用  
 系统索引现在是多键的 (Bruce)  
 移除 Oidint2、oidint4 和 oidname 类型 (Bruce)

更多系统表查找使用系统缓存 (Bruce)  
 backend/pl 中的新后端编程语言 PL/pgSQL (Jan)  
 新 SERIAL 数据类型, 自动创建序列/索引 (Thomas)  
 无需重新编译即可启用断言检查 (Massimo)  
 用户锁增强 (Massimo)  
 设置序列值的新 setval() 命令 (Massimo)  
 启动时若无 postmaster 运行则自动移除 unix 套接字文件 (Massimo)  
 条件跟踪包 (Massimo)  
 新 UNLISTEN 命令 (Massimo)  
 psql 和 libpq 现在可用 win32.mak 在 Windows 下编译 (Magnus)  
 Lo\_read 不再存储尾随 NULL (Bruce)  
 标识符现在内部截断为 31 字符 (Bruce)  
 Createuser 选项现在可在命令行上使用  
 添加 64 位整数支持代码, configure 已测试, int8 类型 (Thomas)  
 防止失败 COPY 的文件描述符泄漏 (Bruce)  
 新 pg\_upgrade 命令 (Bruce)  
 更新的 /contrib 目录 (Massimo)  
 新 CREATE TABLE DEFAULT VALUES 语句可用 (Thomas)  
 新 INSERT INTO TABLE DEFAULT VALUES 语句可用 (Thomas)  
 新 DECLARE 和 FETCH 特性 (Thomas)  
 libpq 的内部结构现在不导出 (Tom)  
 允许多达 8 键的索引 (Bruce)  
 移除 ARCHIVE 关键字, 它不再使用 (Thomas)  
 pg\_dump -n 标志以抑制标识符周围的引号  
 禁用视图的系统列 (Jan)  
 用于网络地址的新 INET 和 CIDR 类型 (TomH、Paul)  
 psql 输出中不再有双引号  
 pg\_dump 现在转储视图 (Terry)  
 新 SET QUERY\_LIMIT (Tatsuo、Jan)

#### 源代码树变更

-----  
 /contrib 清理 (Jun)  
 内联一些每行都调用的小函数 (Bruce)  
 Alpha/linux 修复  
 HP/UX 清理 (Tom)  
 多字节回归测试 (Soonmyung.)  
 从 configure 移除 --disabled 选项  
 定义 PGDOC 默认使用 POSTGRES DIR  
 使回归可选  
 移除 pgindent 的额外花括号代码 (Bruce)  
 添加 bsd 共享库支持 (Bruce)  
 新 --without-CXX 支持 configure 选项 (Brook)  
 新 FAQ\_CVS  
 更新 tools/backend 中的后端流程图 (Bruce)  
 把 atttypmid 从 int16 改为 int32 (Bruce、Tom)  
 为没有 Getrusage() 的平台提供修复 (Tom)  
 把 PQconnectdb、PGUSER、PGPASSWORD 添加到 libpq 手册页  
 NS32K 平台修复 (Phil Nelson、John Buller)  
 SCO 7/UnixWare 2.x 修复 (Billy 等)  
 Sparc/Solaris 2.5 修复 (Ryan)  
 Pgbuiltin.3 已过时, 移到 doc 文件 (Thomas)  
 更多的文档 (Thomas)  
 Nextstep 支持 (Jacek)

Aix 支持 (David)  
pginterface 手册页 (Bruce)  
所有共享库都有版本号  
把所有操作系统特定的共享库定义合并到一个文件  
更智能的 TCL/TK 配置检查 (Billy)  
更智能的 perl 配置 (Brook)  
未找到 install 脚本时 configure 使用随附的 install-sh (Tom)  
用于共享库配置的新 Makefile.shlib (Tom)

## 35.2. 版本 6.3.2

这是 6.3.x 的缺陷修复版本。有关 6.3 版本系列新特性更完整的摘要，请参阅 6.3 版本的发行说明。

摘要：

- 修复某些平台（包括 Linux）的自动配置支持，该问题是 6.3.1 版本无意中引入的。
- 正确处理 BETWEEN 和 LIKE 子句左侧的函数调用。

对于运行 6.3 或 6.3.1 的用户，不需要进行转储/恢复。只需做一次 make distclean、make 和 make install。最后一步应在 postmaster 未运行时执行。你应重新链接任何使用 PostgreSQL 库的自定义应用程序。

对于从 6.3 之前安装的升级，请参阅 6.3 版本的安装和迁移说明。

### 35.2.1. 详细变更列表

变更

-----

对 tcl/tk 的配置检测改进 (Brook Milligan、Alvin)  
手册页改进 (Bruce)  
BETWEEN 和 LIKE 修复 (Thomas)  
修复 pg\_dump 所用 psql \connect 的问题 (Oliver Elphick)  
新 odbc 驱动  
pgaccess 0.86 版  
移除 qsort，现使用 libc 版本并清理 (Jeroen)  
修复检测到的缓冲区越界 (Maurice Gittens)  
修复 libpgtcl 中的缓冲区越界 (Randy Kunkee)  
修复带 DISTINCT 或 ORDER BY 的 UNION (Bruce)  
gettimeofday configure 检查 (Doug Winterburn)  
修复 "indexes not used" 缺陷 (Vadim)  
文档补充 (Thomas)  
修复后端内存泄漏 (Bruce)  
libreadline 清理 (Erwan MAS)  
移除 DISTDIR (Bruce)  
Makefile 依赖清理 (Jeroen van Vianen)  
ASSERT 修复 (Bruce)

## 35.3. 版本 6.3.1

摘要：

- 对多字节字符集的额外支持。

- 修复混合端序客户端和服务器的字节顺序。
- 对允许的 SQL 语法做小的更新。
- 改进安装时配置的自动检测。
- 

对于运行 6.3 的用户，不需要进行转储/恢复。只需做一次 `make distclean`、`make` 和 `make install`。最后一步应在 `postmaster` 未运行时执行。你应重新链接任何使用 PostgreSQL 库的自定义应用程序。

对于从 6.3 之前安装的升级，请参阅 6.3 版本的安装和迁移说明。

### 35.3.1. 详细变更列表

#### 变更

```
-----
ecpg 清理/修复, 现为 1.1 版 (Michael Meskes)
pg_user 清理 (Bruce)
pg_dump 和 tclsh 的大对象修复 (alvin)
修复多个相邻下划线的 LIKE
修复内建函数的重定义 (Thomas)
ultrix4 清理
升级到 pg_access 0.83
更新 CLUSTER 手册页
多字节字符集支持, 参见 doc/README.mb (Tatsuo)
configure --with-pgport 修复
pg_ident 修复
修复后端通信的大端问题 (Kataoka)
修复 SUBSTR() 和 substring() (Jan)
若干 jdbc 修复 (Peter)
libpgtcl 改进, 见 libptcl/README (Randy Kunkee)
修复 "Datasize = 0" 错误 (Vadim)
防止 \do 折行 (Bruce)
移除重复的俄文字符集条目
Sunos4 清理
允许 LOCK 和 SELECT INTO 中使用可选的 TABLE 关键字 (Thomas)
CREATE SEQUENCE 选项允许负整数 (Thomas)
添加 "PASSWORD" 作为允许的列标识符 (Thomas)
添加对 UNION 目标字段的检查 (Bruce)
修复 Alpha 移植 (Dwayne Bailey)
修复包含引号的 text 数组 (Doug Gibson)
Solaris 编译修复 (Albert Chin-A-Young)
更好地识别 tcl 与 tk 库和头文件 (Bruce)
```

## 35.4. 版本 6.3

此版本有许多新特性和改进。以下是简要而不完整的总结：

- 许多新的 SQL 特性，包括完整的 SQL92 子查询能力（只缺目标列表子查询，其余一切都有了）。
- 支持用客户端环境变量指定时区和日期样式。
- 客户端/服务器连接的套接字接口。现在这是默认，所以你可能需要用“-i”标志启动 `postmaster`。
- 更好的口令认证机制。默认表权限已经更改。
- 旧式“时间旅行”已被移除。性能得到改进。

**注意**

Bruce Momjian 写了以下说明来介绍新版本。

我想提一些 6.3 的普遍问题。这里只是一句话说不完的大项。仍需查阅详细变更列表。

首先，我们现在有了子查询。既然有了它们，我想指出：没有子查询的 SQL 是一种非常受限的语言。子查询是一项重大特性，你应当检查自己的代码，看看哪些地方子查询能为你的查询提供更好的解决方案。我想你会发现，子查询的用途比你想象的更多。Vadim 用子查询让我们登上了 SQL 的大舞台，而且是功能完备的子查询。子查询唯一不能做的，是用在目标列表中。

其次，6.3 默认使用 Unix 域套接字而不是 TCP/IP。要启用来自其他机器的连接，必须使用新的 `postmaster -i` 选项，当然还要编辑 `pg_hba.conf`。此外，由于这个原因，`pg_hba.conf` 的格式已经更改。

第三，`char()` 字段现在允许比 `varchar()` 或 `text` 更快的访问。具体来说，`text` 和 `varchar()` 在访问该类型第一列之后的任何列时有性能损失。`char()` 过去也有这种访问损失，但现在没有了。这可能提示你重新设计一些表，特别是如果你有定义为 `varchar()` 或 `text` 的短字符列。这项和其他变更使 6.3 比以前的版本更快。

我们现在可以独立于任何 Unix 文件定义口令。有新的 SQL `USER` 命令。更多信息见 `pg_hba.conf` 手册页。还有一个新表 `pg_shadow`，用于存储用户信息和用户口令，默认只有 `postgres` 超级用户可以 `SELECT`。`pg_user` 现在是 `pg_shadow` 的视图，可被 `PUBLIC SELECT`。你的应用程序应继续使用 `pg_user` 而无需更改。

用户创建的表现默认不再对 `PUBLIC` 授予 `SELECT` 权限。这样做是因为 ANSI 标准的要求。当然，你可以在表创建之后随意 `GRANT` 你想要的任何权限。系统表仍然可被 `PUBLIC SELECT`。

我们还有真正的死锁检测代码。不再有六十秒超时。而且新的锁定代码更好地实现了 FIFO，因此在重度使用期间资源饥饿应该会更少。

以前的版本中有很多关于文档不足的抱怨。Thomas 在此版本的许多新手册上投入了大量精力。请查看 `doc/` 目录。

出于性能原因，时间旅行已被移除，但可以用触发器实现（参见 `pgsql/contrib/spi/README`）。请查看新的 `\d` 命令以了解类型、操作符等。此外，视图现在有自己的权限，不再基于底层表，因此必须单独设置视图的权限。查看 `/pgsql/interfaces` 了解与 PostgreSQL 交互的一些新方式。

这是第一个真正需要为现有用户做解释的版本。在许多方面，这是必要的，因为新版本移除了许多限制，人们以前使用的变通方法不再需要。

## 35.4.1. 迁移到版本 6.3

希望从任何以前的 PostgreSQL 版本迁移数据的用户，需要使用 `pg_dump` 或 `pg_dumpall` 进行转储/恢复。

## 35.4.2. 详细变更列表

缺陷修复

修复被 `MOVE` 实现破坏的二进制游标 (Vadim)  
 修复 `tcl` 库崩溃 (Jan)  
 修复数组处理的问题，来自 Gerhard Hintermayer

修复 acl 错误并移除重复的 pqtrace (Bruce)  
 修复 psql \e 对空文件的问题 (Bruce)  
 修复 varchar() 字段上的 textcat (Bruce)  
 修复 DBT Sendproc (Zeugswetter Andres)  
 修复 vacuum analyze 语法问题 (Bruce)  
 修复国际标识符的问题 (Tatsuo)  
 修复继承表上的聚合 (Bruce)  
 修复 substr() 的越界数据问题  
 修复 select 1=1 or 2=2、select 1=1 and 2=2 和 select sum(2+2) (Bruce)  
 修复 notty 输出以显示状态结果。-q 选项仍会关闭它 (Bruce)  
 修复 count(\*)、带视图和多表的聚合以及 sum(3) (Bruce)  
 修复 cluster (Bruce)  
 修复 PQtrace 多次启动/停止的问题 (Bruce)  
 修复多种锁定问题，例如较新的锁等待者  
 先于较旧的等待者获得锁、读锁持有者在有写者  
 等待锁时不共享锁、以及等待的写者没有  
 优先于等待的读者 (Bruce)  
 修复从外部文件执行查询时 psql 的崩溃 (James)  
 修复多列 order by 且第一列含 NULL 值的问题 (Jeroen)  
 为 float8 和 int4 使用正确的哈希表支持函数 (Thomas)  
 重新启用 CREATE OPERATOR 语句中的 JOIN= 选项 (Thomas)  
 更改布尔操作符的优先级以符合预期行为 (Thomas)  
 对过大整数生成 elog(ERROR) (Bruce)  
 允许在约束子句中使用多参数函数 (Thomas)  
 检查布尔输入字面量 'true'、'false'、'yes'、'no'、'1'、'0'，  
 无法识别时抛出 elog(ERROR) (Thomas)  
 大型对象的重要修复  
 修复 GROUP BY 显示重复项的问题 (Vadim)  
 修复 MergeJoin 中的索引扫描 (Vadim)

## 增强

-----  
 带 EXISTS、IN、ALL、ANY 关键字的子查询 (Vadim、Bruce、Thomas)  
 新用户手册 (Thomas 等)  
 通过内联一些频繁调用的函数提速  
 真正的死锁检测，不再依赖超时 (Bruce)  
 添加 SQL92 "常量" CURRENT\_DATE、CURRENT\_TIME、CURRENT\_TIMESTAMP、  
 CURRENT\_USER (Thomas)  
 修改约束语法以符合 SQL92 (Thomas)  
 用索引实现 SQL92 PRIMARY KEY 和 UNIQUE 子句 (Thomas)  
 识别 FOREIGN KEY 的 SQL92 语法。抛出 elog 通知 (Thomas)  
 允许 NOT NULL UNIQUE 约束子句 (以前各自单独允许) (Thomas)  
 允许对非常量使用 Postgres 风格的转换 ("::") (Thomas)  
 添加对 SQL3 TRUE 和 FALSE 布尔常量的支持 (Thomas)  
 支持 IS TRUE/IS FALSE/IS NOT TRUE/IS NOT FALSE 的 SQL92 语法 (Thomas)  
 允许更短的布尔字面量字符串 (如 "t"、"tr"、"tru") (Thomas)  
 允许 SQL92 分隔标识符 (Thomas)  
 实现 SQL92 二进制和十六进制字符串解码 (b'10' 和 x'1F') (Thomas)  
 支持字面字符串类型强制转换的 SQL92 语法  
 (如 "DATETIME 'now'") (Thomas)  
 添加 int2、int4 和 OID 类型与 text 之间的转换 (Thomas)  
 构建索引时使用共享锁 (Vadim)  
 事务块内的用户查询完成后释放为其分配的内存  
 此功能在 <= 6.2.1 中被关闭 (Vadim)

新 SQL 语句 CREATE PROCEDURAL LANGUAGE (Jan)  
 新 Postgres 过程语言 (PL) 后端接口 (Jan)  
 把 pg\_dump -H 选项改名为 -h (Bruce)  
 添加 Java 对口令和欧洲日期的支持 (Peter)  
 为 LIKE 和 ~、!~ 操作使用索引 (Bruce)  
 添加 datetime 和 timespan 的哈希函数 (Thomas)  
 移除时间旅行 (Vadim、Bruce)  
 为 \d 和 \z 添加分页并修复 \i (Bruce)  
 向后端和前端库添加 Unix 域套接字支持 (Goran)  
 实现 CREATE DATABASE/WITH LOCATION 和 initlocation 工具 (Thomas)  
 允许更多 SQL92 和/或 Postgres 保留字作为列标识符 (Thomas)  
 增强对 SQL92 SET TIME ZONE... 的支持 (Thomas)  
 SET/SHOW/RESET TIME ZONE 使用 TZ 后端环境变量 (Thomas)  
 实现 SET keyword = DEFAULT 和 SET TIME ZONE DEFAULT (Thomas)  
 启用使用 TZ 环境变量的 SET TIME ZONE (Thomas)  
 把 PGDATESTYLE 环境变量添加到前端和后端初始化 (Thomas)  
 添加 PGTZ、PGCOSTHEAP、PGCOSTINDEX、PGRPLANS、PGGEQO  
 前端库初始化环境变量 (Thomas)  
 回归测试时区用 "setenv PGTZ PST8PDT" 自动设置 (Thomas)  
 添加 pg\_description 表以存放表、列、操作符、类型等的信息  
 聚合的信息 (Bruce)  
 把系统表/索引名 16 字符的限制增加到 32 字符 (Bruce)  
 重命名系统索引 (Bruce)  
 向 SET DATESTYLE 添加 'GERMAN' 选项 (Thomas)  
 定义带 "hh:mm:ss" 字段的 "ISO 风格" timespan 输出格式 (Thomas)  
 允许 delta time 的小数值 (如 '2.5 days') (Thomas)  
 更仔细地验证 delta time 的数字输入 (Thomas)  
 实现把年积日作为 date\_part() 的可能输入 (Thomas)  
 定义 timespan\_finite() 和 text\_timespan() 函数 (Thomas)  
 移除归档内容 (Bruce)  
 允许有独立于系统口令文件的 pg\_password 认证数据库 (Todd)  
 转储 ACL、GRANT、REVOKE 权限 (Matt)  
 定义 text、varchar 和 bpchar 字符串长度函数 (Thomas)  
 修复 Query 对继承的处理以及代价计算 (Bruce)  
 实现 CREATE TABLE/AS SELECT (SELECT/INTO 的替代) (Thomas)  
 允许约束中的 NOT、IS NULL、IS NOT NULL (Thomas)  
 实现 SELECT 的 UNION (Bruce)  
 向 INSERT 添加 UNION、GROUP、DISTINCT (Bruce)  
 varchar() 在磁盘上只存储必要的字节 (Bruce)  
 修复 BLOB 的问题 (Peter)  
 JDBC 超大补丁.....变更列表见 README\_6.3 (Peter)  
 从 PQconnectdb() 移除未使用的 "option"  
 新 LOCK 命令和描述死锁的锁手册页 (Bruce)  
 添加新 psql \da、\dd、\df、\do、\ds 和 \dT 命令 (Bruce)  
 增强 psql \z 以显示序列 (Bruce)  
 在 psql \d 表中显示 NOT NULL 和 DEFAULT (Bruce)  
 新 psql .psqlrc 文件启动 (Andrew)  
 修改 contrib/linux 中的示例启动脚本以显示 syslog (Thomas)  
 contrib/ip\_and\_mac 中 IP 和 MAC 地址的新类型 (TomH)  
 contrib/unixdate 中与日期/时间类型的 Unix 系统时间转换 (Thomas)  
 contrib 内容的更新 (Massimo)  
 向 DBD::Pg 添加 Unix 套接字支持 (Goran)  
 新 python 接口 (PyGreSQL 2.0) (D'Arcy)  
 新的前端/后端协议带有版本号和网络字节序 (Phil)

pg\_hba.conf 的安全特性得到增强和记录, 大量清理 (Phil)  
 CHAR() 现在比 VARCHAR() 或 TEXT 访问更快  
 ecpg 嵌入式 SQL 预处理器  
 降低系统列的开销 (Vadmin)  
 移除 pg\_time 表 (Vadim)  
 添加 pg\_type 属性以标识需要长度的类型 (bpchar、varchar)  
 COPY 命令失败时报告出错的行  
 允许对视图设置的权限独立于底层表。  
 为安全起见, 酌情对视图使用 GRANT/REVOKE (Jan)  
 表现在没有默认的 GRANT SELECT TO PUBLIC。你必须  
 显式授予此类权限。  
 清理教程示例 (Darren)

#### 源码树变更

-----  
 添加新的 html 开发工具和 /tools/backend 中的流程图  
 修复 SCO 编译  
 Stratus 计算机移植 (Robert Gillies)  
 为 BSD44\_derived & i386\_solaris 添加 shlib 支持  
 使 configure 更加自动化 (Brook)  
 添加检查回归测试结果的脚本  
 将解析器函数拆分为更小的文件并分组 (Bruce)  
 把 heap\_create 改名为 heap\_create\_and\_catalog, 把 heap\_creatr  
 改名为 heap\_create() (Bruce)  
 Sparc/Linux 的锁定补丁 (TomS)  
 移除 PORTNAME 并重组移植特定内容 (Marc)  
 添加优化器 README 文件 (Bruce)  
 移除优化器中的部分递归并清理那里的代码 (Bruce)  
 修复 NetBSD 锁定 (Henry)  
 修复 libptcl 的制作 (Tatsuo)  
 AIX 补丁 (Darren)  
 把 IS TRUE、IS FALSE、... 更改为使用 "=" 的表达式而不是  
 对 isttrue() 或 isfalse() 的函数调用以允许优化 (Thomas)  
 各种 NetBSD/Sparc 相关修复 (TomH)  
 Alpha linux 锁定 (Travis、Ryan)  
 把 elog(WARN) 改为 elog(ERROR) (Bruce)  
 FreeBSD 的 FAQ (Marc)  
 将 PostODBC 源码树纳入标准发行版 (Marc)  
 HP/UX 10 与 9 的小补丁 (Stan)  
 新 pg\_attribute.atttypmod 用于 varchar 长度等类型特定信息 (Bruce)  
 UnixWare 补丁 (Billy)  
 自旋锁 asm 的新 i386 'lock' (Billy)  
 移除对多路复用后端的支持  
 开始 OpenBSD 移植  
 开始 AUX 移植  
 开始 Cygnus 移植  
 在回归测试套件中添加字符串函数 (Thomas)  
 扩展一些以前截断为 16 字符的函数名 (Thomas)  
 移除不需要的 malloc() 调用并用 palloc() 替换 (Bruce)

## 35.5. 版本 6.2.1

6.2.1 是 6.2 之上的缺陷修复和易用性版本。

摘要:

- 按照 SQL92, 允许字符串跨行。
- 包含在表更新时插入用户名的示例触发器函数。

这是 6.2 之上的次要缺陷修复版本。从 6.2 之前的系统升级需要完整的转储/重载。请参阅 6.2 版本的发行说明了解操作说明。

## 35.5.1. 从版本 6.2 迁移到版本 6.2.1

这是次要缺陷修复版本。从 6.2 版本不需要转储/重载, 但从 6.2 之前的任何版本需要。

从 6.2 版本升级时, 如果你选择转储/重载, 你会发现 `avg(money)` 现在计算正确。所有其他缺陷修复在更新可执行文件后生效。

避免转储/重载的另一种方法是在 `psql` 中使用以下 SQL 命令更新现有的系统表:

```
update pg_aggregate set aggfinalfn = 'cash_div_flt8'
where aggname = 'avg' and aggbasetype = 790;
```

这需要对每个现有数据库执行, 包括 `template1`。

## 35.5.2. 详细变更列表

此发行版中的变更

-----  
允许 `TIME` 和 `TYPE` 作为列名 (Thomas)  
允许更大范围的 `true/false` 布尔值 (Thomas)  
支持 `"now"` 和 `"current"` 的输出 (Thomas)  
正确处理 `INSERT NULL` 与 `DEFAULT` (Vadim)  
修复缓冲区管理器中关系引用计数的问题 (Vadim)  
像 `ANSI` 一样允许字符串跨行 (Thomas)  
修复带 `ORDER BY` 的反向游标 (Vadim)  
修复 `avg(cash)` 计算 (Thomas)  
修复 `ORDER/GROUP BY` 中重复指定列的问题 (Vadim)  
记录返回受影响行数的新 `libpq` 函数 `PQcmdTuples` (Bruce)  
用于在 `INSERT/UPDATE` 时插入用户名的触发器函数 (Brook Milligan)

## 35.6. 版本 6.2

希望从以前的 PostgreSQL 版本迁移数据的用户, 需要进行转储/恢复。

### 35.6.1. 从版本 6.1 迁移到版本 6.2

此迁移需要完整转储 6.1 数据库并在 6.2 中恢复。

注意, 应使用 6.2 的 `pg_dump` 和 `pg_dumpall` 工具来转储 6.1 数据库。

### 35.6.2. 从版本 1.x 迁移到版本 6.2

从更早的 1.\* 版本迁移的用户应先升级到 1.09, 因为 `COPY` 输出格式自 1.02 版本起得到了改进。

### 35.6.3. 详细变更列表

缺陷修复

-----  
 修复 pg\_dump 对继承、序列、归档表的问题 (Bruce)  
 修复因移位、无符号和 Solaris 的错误原型  
 导致的溢出编译错误 (Diab Jerius)  
 修复几何线段算术中的缺陷 (错误的交点计算) (Thomas)  
 检查几何对象在端点处的相交以避免舍入问题 (Thomas)  
 捕获无效的删除尝试 (Vadim)  
 把时间函数名改得更一致 (Michael Reifenberg)  
 检查除零 (Michael Reifenberg)  
 修复一个非常老的缺陷: 命令本身可以看见  
 由该命令更改/插入的元组 (因此会导致对已更新  
 元组的多次更新等) (Vadim)  
 修复 SELECT null, 'fail' FROM pg\_am (Patrick)  
 现在允许 SELECT NULL as EMPTY\_FIELD (Patrick)  
 从 contrib/pginterface 中移除不需要的信号代码  
 修复 OR (where x <> 1 or x isnull 不返回 x NULL 的元组) (Vadim)  
 修复 time\_cmp 函数 (Vadim)  
 修复 WHERE 子句中对首参数为非属性的函数的处理 (Vadim)  
 修复条目顺序与目标列表顺序不同时的 GROUP BY (Vadim)  
 修复 pg\_dump 对无 sfunci 聚合的处理 (Vadim)

## 增强

-----  
 默认遗传优化器 GEQO 参数现在是 8 (Bruce)  
 允许在目标列表中使用带聚合的参数 (Vadim)  
 添加 JDBC 驱动作为接口 (Adrian & Peter)  
 pg\_password 工具  
 返回 INSERT/UPDATE/DELETE 等插入/影响的行数 (Vadim)  
 用 CREATE TRIGGER 实现触发器 (SQL3) (Vadim)  
 SPI (服务器编程接口) 允许在 C 函数内执行查询 (Vadim)  
 实现 NOT NULL (SQL92) (Robson Paniago de Miranda)  
 纳入用于字符串处理、外连接和联合的保留字 (Thomas)  
 用独占状态实现扩展注释 ("/\* ... \*/") (Thomas)  
 添加 "/\*" 单行注释 (Bruce)  
 移除对操作符名称中字符的一些限制 (Thomas)  
 实现表的 DEFAULT 和 CONSTRAINT (SQL92) (Vadim & Thomas)  
 添加文本连接操作符和函数 (SQL92) (Thomas)  
 支持 WITH TIME ZONE 语法 (SQL92) (Thomas)  
 支持 INTERVAL unit TO unit 语法 (SQL92) (Thomas)  
 定义类型 DOUBLE PRECISION、INTERVAL、CHARACTER  
 和 CHARACTER VARYING (SQL92) (Thomas)  
 定义类型 FLOAT(p) 和基础的 DECIMAL(p,s)、NUMERIC(p,s) (SQL92) (Thomas)  
 定义 EXTRACT()、POSITION()、SUBSTRING() 和 TRIM() (SQL92) (Thomas)  
 定义 CURRENT\_DATE、CURRENT\_TIME、CURRENT\_TIMESTAMP (SQL92) (Thomas)  
 为 UNION、HAVING、INNER 和 OUTER JOIN 添加语法和警告 (SQL92) (Thomas)  
 添加更多保留字, 主要为符合 SQL92 (Thomas)  
 允许 timespan/retime 类型的 hh:mm:ss 时间输入 (Thomas)  
 为 lseg、path、polygon 添加 center() 例程 (Thomas)  
 为 circle-polygon、polygon-polygon 添加 distance() 例程 (Thomas)  
 用轴交叉算法显式检查多边形内包含的点和多边形 (Thomas)  
 添加 circle-box 转换例程 (Thomas)  
 合并不同几何数据类型的冲突操作符 (Thomas)  
 将距离操作符 "<==>" 替换为 "<->" (Thomas)  
 把"上方"操作符 "!^" 替换为 ">^", "下方"操作符 "!!" 替换为 "<^" (Thomas)

添加两端修剪文本、子串和字符串位置的例程 (Thomas)  
添加转换例程 circle(box) 和 poly(circle) (Thomas)  
允许内部排序存储在内存中而不是文件中 (Bruce & Vadim)  
允许内部相同的类型上的函数和操作符成功执行 (Bruce)  
在性能分析后加速后端启动 (Bruce)  
为性能内联频繁调用的函数 (Bruce)  
减少 open() 调用 (Bruce)  
psql: 为 \h 和 \? 添加 PAGER, \C 修复  
修复无 tty 时 psql 分页器的问题 (Bruce)  
新 entab 工具 (Bruce)  
用于引用完整性的通用触发器函数 (Vadim)  
用于时间旅行的通用触发器函数 (Vadim)  
用于 AUTOINCREMENT/IDENTITY 特性的一般触发器函数 (Vadim)  
MOVE 实现 (Vadim)

#### 源码树变更

-----  
HP-UX 10 补丁 (Vladimir Turin)  
添加 SCO 支持 (Daniel Harris)  
MkLinux 补丁 (Tatsuo Ishii)  
把几何 box 术语从 "length" 改为 "width" (Thomas)  
废弃几何代码中临时不存储的斜率字段 (Thomas)  
从 INSTALL 移除重启说明 (Bruce)  
先在 /usr/ucb 中查找 install (Bruce)  
修复 c++ 复制示例代码 (Thomas)  
在 psql 手册页添加 -o (Bruce)  
防止 relname 未分配字符串长度被复制进数据库 (Bruce)  
清理 NAMEDATALEN 的使用 (Bruce)  
修复输出中超过 15 字符的 pg\_proc 名 (Bruce)  
添加 strNcpy() 函数 (Bruce)  
移除一些不必要的 (void) 转换 (Bruce)  
新的 interfaces 目录 (Marc)  
用对 fd.c 函数的调用替换 fopen() 调用 (Bruce)  
尽可能把函数设为 static (Bruce)  
把未使用的函数包在 #ifdef NOT\_USED 中 (Bruce)  
移除时间戳支持中对 difftime() 的调用以修复 SunOS (Bruce & Thomas)  
针对 Digital Unix 的变更  
pg\_dumpall 的可移植性修复 (Bruce)  
把 pg\_attribute.attnvals 改名为 attdisbursion (Bruce)  
"intro/unix" 手册页现为 "pgintro" (Bruce)  
"built-in" 手册页现为 "pgbuiltin" (Bruce)  
"drop" 手册页现为 "drop\_table" (Bruce)  
添加 "create\_trigger"、"drop\_trigger" 手册页 (Thomas)  
添加约束回归测试 (Vadim & Thomas)  
添加注释语法回归测试 (Thomas)  
添加 PGINDENT 和支持程序 (Bruce)  
对所有 \*.c 和 \*.h 文件运行 PGINDENT 的大规模提交 (Bruce)  
文件移到 /src/tools 目录 (Bruce)  
SPI 和触发器编程指南 (Vadim & D'Arcy)

## 35.7. 版本 6.1.1

### 35.7.1. 从版本 6.1 迁移到版本 6.1.1

这是次要缺陷修复版本。从 6.1 版本不需要转储/重载，但从 6.1 之前的任何版本需要。更多细节请参阅 6.1 的发行说明。

### 35.7.2. 详细变更列表

此发行版中的变更

-----  
 修复带选项的 SET (Thomas)  
 允许 pg\_dump/pg\_dumpall 保留所有表/对象的所有权 (Bruce)  
 新 psql \connect 选项允许只更改用户名而不更改数据库  
 修复 initdb --debug 选项 (Yoshihiko Ichikawa)  
 lexttest 清理 (Bruce)  
 哈希修复 (Vadim)  
 修复日期/时间的月份边界算术 (Thomas)  
 修复某些移植的时区夏令时处理 (Thomas、Bruce、Tatsuo)  
 timestamp 彻底改造为使用标准函数 (Thomas)  
 日期/时间例程的其他代码清理 (Thomas)  
 psql 的 \d 现在不区分大小写 (Bruce)  
 psql 的反斜杠命令现在可以带尾部分号 (Bruce)  
 修复使用 \g 时 psql 的内存泄漏 (Bruce)  
 与服务器通信的字节序处理重大修复 (Thomas、Tatsuo)  
 Solaris 汇编器和头文件的修复 (Yoshihiko Ichikawa)  
 允许用户名中使用下划线 (Bruce)  
 pg\_dumpall 现在返回正确状态，可移植性修复 (Bruce)

## 35.8. 版本 6.1

回归测试已经为 PostgreSQL 6.1 版本做了适配和大量修改。

新增了三种数据类型 (datetime、timespan 和 circle)，已加入到 PostgreSQL 本地类型集中。点、框、路径和多边形在各数据类型间的输出格式已统一。misc.out 中的多边形输出只相对原始回归输出做了抽查。

PostgreSQL 6.1 引入了使用遗传算法的新替代优化器。当查询包含多个限定条件或多个表时（优化器可选择求值顺序），这些算法会给查询结果的顺序引入随机行为。已有几个回归测试被修改为显式排序结果，因此对优化器的选择不敏感。少数回归测试针对本质上无序的数据类型（如点和时间间隔），涉及这些类型的测试显式地用 set geqo to 'off' 和 reset geqo 括起来。

数组说明符（原子值周围的花括号）的解释似乎在最初的回归测试生成之后的某个时候发生了变化。当前的 ./expected/\*.out 文件反映了这一新解释，它可能并不正确！

float8 回归测试至少在某些平台上失败。这是由于 pow() 和 exp() 实现的差异以及上溢/下溢条件所用的信号机制不同。

random 测试中的"随机"结果本应导致 random 测试"失败"，因为回归测试是用简单的 diff 评估的。不过，在我的测试机（Linux/gcc/i686）上，"random" 似乎并不会产生随机结果。

## 35.8.1. 迁移到版本 6.1

此迁移需要完整转储 6.0 数据库并在 6.1 中恢复。

从更早的 1.\* 版本迁移的用户应先升级到 1.09, 因为 COPY 输出格式自 1.02 版本起得到了改进。

## 35.8.2. 详细变更列表

### 缺陷修复

-----  
 库例程中的包长度检查  
 锁管理器优先级补丁  
 检查 float8 的下溢/上溢 (Bruce)  
 多表连接修复 (Vadim)  
 SIGPIPE 崩溃修复 (Darren)  
 大对象修复 (Sven)  
 允许 btree 索引处理 NULL (Vadim)  
 时区修复 (D'Arcy)  
 无行时 select SUM(x) 可返回 NULL (Thomas)  
 内部优化器、执行器缺陷修复 (Vadim)  
 修复 < 或 <= 内层循环无行的问题 (Vadim)  
 防止连接索引子句重新交换 (Vadim)  
 修复多表的连接子句 (Vadim)  
 修复数组的哈希和哈希连接 (Vadim)  
 修复 abstime 类型的 btree (Vadim)  
 大对象修复 (Raymond)  
 修复哈希索引中的缓冲区泄漏 (Vadim)  
 修复 rtree 以用于内层扫描 (Vadim)  
 修复 gist 以用于内层扫描并清理 (Vadim、Andrea)  
 避免不必要的本地缓冲区分配 (Vadim、Massimo)  
 修复事务中止时本地缓冲区泄漏 (Vadim)  
 修复文件管理器内存泄漏并清理 (Vadim、Massimo)  
 修复存储管理器内存泄漏 (Vadim)  
 修复 btree 重复项处理 (Vadim)  
 修复 vacuum 导致已删行复活的问题 (Vadim)  
 修复 SELECT varchar()/char() INTO TABLE 生成零长度字段的问题 (Bruce)  
 用 Purify 修复许多 psql、pg\_dump 和 libpq 内存泄漏 (Igor)

### 增强

-----  
 属性优化统计 (Bruce)  
 快得多的新 btree 批量装载代码 (Paul)  
 批量装载代码加入 BTREE UNIQUE (Vadim)  
 新锁调试代码 (Massimo)  
 libpg++ 的大规模变更 (Leo)  
 新 GEQO 优化器加速多表优化 (Martin)  
 对唯一键非唯一插入的新 WARN 消息 (Marc)  
 update x=-3 (无空格) 现在有效 (Bruce)  
 移除标识符的大小写敏感处理 (Bruce、Thomas、Dan)  
 调试后端现在美化打印树 (Darren)  
 新 Oracle 字符函数 (Edmund)  
 新明文口令函数 (Dan)  
 "no such class or insufficient privilege" 拆分为不同的消息 (Dan)

新 ANSI timestamp 函数 (Dan)  
 新 ANSI Time 和 Date 类型 (Thomas)  
 在后端中移动大数据块 (Martin)  
 多列 btree 索引 (Vadim)  
 新 SET var TO value 命令 (Martin)  
 读取时更新事务状态 (Dan)  
 字符类型的新区域设置 (Oleg)  
 新 SEQUENCE 序列号生成器 (Vadim)  
 现在可以使用 GROUP BY 函数 (Vadim)  
 重新组织回归测试 (Thomas、Marc)  
 新优化器操作权重 (Vadim)  
 新 psql \z 授予/许可选项 (Marc)  
 新 MONEY 数据类型 (D'Arcy、Thomas)  
 tcp 套接字通信速度改进 (Vadim)  
 VACUUM 的属性统计及特定列新选项 (Vadim)  
 许多几何类型改进 (Thomas、Keith)  
 附加回归测试 (Thomas)  
 新 datestyle 变量 (Thomas、Vadim、Martin)  
 更多用于排序的比较操作符 (Thomas)  
 新转换函数 (Thomas)  
 新更紧凑的 btree 格式 (Vadim)  
 允许 pg\_dumpall 保留数据库所有权 (Bruce)  
 新 SET GEQO=# 和 R\_PLANS 变量 (Vadim)  
 旧 (!GEQO) 优化器可使用右侧计划 (Vadim)  
 SQL 解析器类型检查改进 (Bruce)  
 新 SET、SHOW、RESET 命令 (Thomas、Vadim)  
 新 \connect database USER 选项  
 新 destroydb -i 选项 (Igor)  
 新 \dt 和 \di psql 命令 (Darren)  
 SELECT "\n" 现在转义换行 (A. Duursma)  
 从旧格式转换的新几何转换函数 (Thomas)

#### 源代码树变更

-----  
 新配置脚本 (Marc)  
 添加 readline 配置选项 (Marc)  
 移除操作系统特定的配置选项 (Marc)  
 新的操作系统特定模板文件 (Marc)  
 不再需要编辑 Makefile.global (Marc)  
 重新安排头文件 (Marc)  
 nextstep 补丁 (Gregor Hoffleit)  
 移除 WIN32 专用代码 (Bruce)  
 移除 postmaster -e 选项, 现在只有 postgres -e 选项 (Bruce)  
 合并前后端中重复的库代码 (Martin)  
 现在可与 eBones、国际 Kerberos 一起工作 (Jun)  
 更多共享库支持  
 c++ 头文件清理 (Bruce)  
 对有缺陷的 flex 发出警告 (Bruce)  
 DG/UX、Ultrix、IRIX、AIX 可移植性修复

## 35.9. 版本 6.0

希望从以前的 PostgreSQL 版本迁移数据的用户, 需要进行转储/恢复。

## 35.9.1. 从版本 1.09 迁移到版本 6.0

此迁移需要完整转储 1.09 数据库并在 6.0 中恢复。

## 35.9.2. 从 1.09 之前版本迁移到版本 6.0

从更早的 1.\* 版本迁移的用户应先升级到 1.09, 因为 COPY 输出格式自 1.02 版本起得到了改进。

## 35.9.3. 详细变更列表

### 缺陷修复

-----  
 ALTER TABLE 缺陷—运行中的 postgres 进程需要重新读取表定义  
 允许对一个表或整个数据库运行 vacuum (Bruce)  
 数组修复  
 修复数组内存写的越界 (Kurt)  
 修复难以捉摸的 btree 范围/非范围缺陷 (Dan)  
 修复某些类型 (如 time 和 date) 上的哈希索引  
 修复 pg\_log 大小爆炸的问题  
 修复 lo\_export() 的权限 (Bruce)  
 修复未初始化的内存读取 (Kurt)  
 修复 ALTER TABLE ... char(3) 缺陷 (Bruce)  
 修复了几个小的内存泄漏  
 修复 EXPLAIN 的选项处理并更改了 full\_path 选项名  
 修复组 acl 权限的输出  
 内存泄漏 (用 Purify 等工具查找并消灭) (Kurt)  
 规则系统的小改进  
 NOTIFY 修复  
 新的运行检查断言  
 彻底改造 parser/analyze 代码以正确报告错误并提高速度  
 Pg\_dump -d 现在正确处理 NULL (Bruce)  
 防止 SELECT NULL 使服务器崩溃 (Bruce)  
 INSERT ... SELECT 列不匹配时正确报告错误  
 插入列名不正确时正确报告错误  
 psql \g filename 现在可用 (Bruce)  
 修复 psql 一行多语句多输出的问题  
 移除重复的系统 oid  
 SELECT \* INTO TABLE . GROUP/ORDER BY 在表已存在时给出 unlink 错误 (Bruce)  
 几项对使后端崩溃查询的修复  
 插入字符串中起始引号的错误 (Bruce)  
 提交空查询现在返回空状态, 而不只是 " " 查询 (Bruce)

### 增强

-----  
 添加 EXPLAIN 手册页 (Bruce)  
 添加 UNIQUE 索引能力 (Dan)  
 添加主机名/用户级访问控制, 而不只是主机名和用户  
 为 <> 添加 != 同义词 (Bruce)  
 允许 "select oid,\* from table"  
 允许 BY、ORDER BY 按编号或按非别名的 table.column 指定列 (Bruce)  
 允许从前端 COPY (Bryan)  
 允许 GROUP BY 使用别名列名 (Bruce)  
 允许真正的压缩, 而不只是同一页上的重用 (Vadim)

允许安装配置选项自动添加所有本地用户 (Bryan)  
 允许 libpq 区分文本值 '' 和 null (Bruce)  
 允许有 createdb 权限的非 postgres 用户 destroydb  
 允许限制谁可以创建 C 函数 (Bryan)  
 允许限制谁可以做后端 COPY (Bryan)  
 可以收缩表、pg\_time 和 pg\_log (Vadim & Erich)  
 更改调试级别 2 为只打印查询, 更改调试标题布局 (Bruce)  
 把十进制常量的默认表示从 float4 改为 float8 (Bruce)  
 postmaster 启动时现在设置欧洲日期格式  
 找不到精确大小写时执行小写函数名  
 聚合/GROUP 处理的修复, 允许 'select sum(func(x),sum(x+y) from z'  
 Gist 现在包含在发行版中 (Marc)  
 本地用户的 Ident 认证 (Bryan)  
 实现 BETWEEN 限定符 (Bruce)  
 实现 IN 限定符 (Bruce)  
 libpq 有 PQgetisnull() (Bruce)  
 libpq++ 改进  
 initdb 的新选项 (Bryan)  
 pg\_dump 允许转储 oid (Bruce)  
 Pg\_dump 为提高速度在表装载后创建索引 (Bruce)  
 Pg\_dumpall 转储所有数据库和用户表  
 Pginterface 对 NULL 值的补充 (Bruce)  
 防止以 root 运行 postmaster  
 psql \h 和 \? 现在可读 (Bruce)  
 psql 允许行内任意位置的反斜杠和分号 (Bruce)  
 psql 更改查询中或引号中行的命令提示符 (Bruce)  
 psql char(3) 现在在 \d 输出中显示为 (bp)char (Bruce)  
 psql 返回码现在更准确 (Bryan?)  
 psql 更新的帮助语法 (Bruce)  
 重新检查并修复 vacuum (Vadim)  
 减小回归 diff 的大小, 移除时区名差异 (Bruce)  
 移除编译器参数以支持二进制发行 (Bryan)  
 反转 HBA 掩码的含义 (Bryan)  
 本地用户的安全认证 (Bryan)  
 加快 vacuum (Vadim)  
 Vacuum 现在有 VERBOSE 选项 (Bruce)

#### 源代码树变更

-----  
 所有函数现在都有与调用比较的原型  
 允许从 Makefile.global 轻松禁用断言 (Bruce)  
 把代码中使用的 oid 常量改为 #define 名称  
 解耦 sparc 和 solaris 定义 (Kurt)  
 Gcc -Wall 干净编译, 警告只来自无法修复的构造  
 主要的头文件重组/精简 (Marc)  
 编译失败时 make 现在停止 (Bryan)  
 Makefile 重构 (Bryan、Marc)  
 把 bsd\_2\_1 合并到 bsd (Bruce)  
 移除 Monitor 程序  
 从 Postgres95 改名为 PostgreSQL  
 新 config.h 文件 (Marc、Bryan)  
 PG\_VERSION 现在设为 6.0 并被 postmaster 使用  
 可移植性补充, 包括 Ultrix、DG/UX、AIX 和 Solaris  
 减少 #define 的数量, 集中化 #define

移除系统表中的重复 OID (Dan)  
 移除重复的系统目录信息或报告不匹配 (Dan)  
 移除了许多操作系统特定的 #define  
 重构目标文件的生成/位置 (Bryan、Marc)  
 重构移植特定文件的位置 (Bryan、Marc)  
 纠正未使用/未初始化的变量

## 35.10. 版本 1.09

抱歉，我们没有记录 1.02 到 1.09 的变更。6.0 中列出的一些变更实际上包含在 1.02.1 到 1.09 版本中。

## 35.11. 版本 1.02

### 35.11.1. 从版本 1.02 迁移到版本 1.02.1

这是 1.02.1 的新迁移文件。它包括 'copy' 变更和一个把旧 ASCII 文件转换为新格式脚本报本。

#### 注意

以下说明是为了帮助希望把数据库从 Postgres95 1.01 和 1.02 迁移到 Postgres95 1.02.1 的用户。

如果你全新开始使用 Postgres95 1.02.1 而不需要迁移旧数据库，则无需再往下读。

为了把较旧的 Postgres95 1.01 或 1.02 版本数据库升级到 1.02.1 版本，需要以下步骤：

1. 启动新的 1.02.1 postmaster
2. 将 1.02.1 新增的内建函数和操作符添加到 1.01 或 1.02 数据库中。方法是用新的 1.02.1 服务器运行你自己的 1.01 或 1.02 数据库，并应用本文件末尾附带的查询。这可以通过 psql 轻松完成。如果你的 1.01 或 1.02 数据库名为 "testdb"，并且你已把命令从本文件末尾剪出并保存到 addfunc.sql 中：

```
% psql testdb -f addfunc.sql
```

从 1.02 升级的用户在执行文件中的最后两条语句时会得到警告，因为它们已经存在于 1.02 中。这无需担心。

### 35.11.2. 转储/重载程序

如果你试图重新装载由旧版本生成的 pg\_dump 或文本模式 'copy tablename to stdout' 输出，需要先将 ASCII 文件运行附带的 sed 脚本，再将其装入数据库。旧格式用 '.' 作为数据结束标记，而现在的结束标记是 '\.'。此外，空字符串现在装载为 '' 而非 NULL。完整细节见 copy 手册页。

```
sed 's/^\.$/\\./g' <in_file >out_file
```

如果你装载的是较旧的二进制 copy 或非 stdout copy，则没有数据结束字符，因此无需转换。

```
-- following lines added by agc to reflect the case-insensitive
-- regexp searching for varchar (in 1.02), and bpchar (in 1.02.1)
```

```
create operator ~* (leftarg = bpchar, rightarg = text, procedure =
    texticregexeq);
create operator !~* (leftarg = bpchar, rightarg = text, procedure =
    texticregexne);
create operator ~* (leftarg = varchar, rightarg = text, procedure =
    texticregexeq);
create operator !~* (leftarg = varchar, rightarg = text, procedure =
    texticregexne);
```

### 35.11.3. 详细变更列表

#### 源代码维护与开发

- \* 全球志愿者团队
- \* 源代码树现在在 ftp.ki.net 的 CVS 中

#### 增强

- \* psql (及底层 libpq 库) 现在有更多输出格式化选项, 包括 HTML
- \* pg\_dump 现在可以输出模式和/或数据, 并有许多修复以提高完整性。
- \* 管理 shell 脚本中用 psql 代替 monitor。monitor 将在下一个发行版中废弃。
- \* 日期/时间函数得到增强
- \* NULL 插入/更新/比较修复/增强
- \* TCL/TK 库和 shell 修复为同时支持 tcl7.4/tk4.0 和 tcl7.5/tk4.1

#### 缺陷修复 (多到几乎无法列举)

- \* 索引
- \* 存储管理
- \* 解引用前检查 NULL 指针
- \* Makefile 修复

#### 新移植

- \* 新增 SolarisX86 移植
- \* 新增 BSD/OS 2.1 移植
- \* 新增 DG/UX 移植

## 35.12. 版本 1.01

### 35.12.1. 从版本 1.0 迁移到版本 1.01

以下说明是为了帮助希望把数据库从 postgres95 1.0 迁移到 postgres95 1.01 的用户。

如果你全新开始使用 postgres95 1.01 而不需要迁移旧数据库, 则无需继续阅读。

为了在用 postgres95 1.0 版本创建的数据库上运行 postgres95 1.01 版本, 需要以下步骤:

1. 把 src/Makefile.global 中 NAMEDATALEN 的定义改为 16, 把 OIDNAMELEN 的定义改为 20。
2. 决定是否要使用基于主机的认证。
  - a. 如果要用, 你必须在顶级数据目录 (通常是 \$PGDATA 的值) 中创建名为 "pg\_hba" 的文件。src/libpq/pg\_hba 给出了示例语法。

- b. 如果不想使用基于主机的认证，可以注释掉 `src/Makefile.global` 中的这一行：

```
HBA = 1
```

注意，基于主机的认证默认开启，如果你不执行上面的 A 或 B 步骤，开箱即用的 1.01 将不允许你连接到 1.0 数据库。

3. 编译并安装 1.01，但不要执行 `initdb` 步骤。
4. 在做任何其他事情之前，终止你的 1.0 `postmaster`，并备份现有的 `$PGDATA` 目录。
5. 把 `PGDATA` 环境变量指向你的 1.0 数据库，但把路径设置为使用 1.01 二进制文件。
6. 把文件 `$PGDATA/PG_VERSION` 从 5.0 改为 5.1
7. 启动新的 1.01 `postmaster`
8. 将 1.01 新增的内建函数和操作符添加到 1.0 数据库中。方法是用新的 1.01 服务器运行你自己的 1.0 数据库，并应用随附保存在文件 `1.0_to_1.01.sql` 中的查询。这可以通过 `psql` 轻松完成。如果你的 1.0 数据库名为 `"testdb"`：

```
% psql testdb -f 1.0_to_1.01.sql
```

然后执行以下命令（从此处剪切并粘贴）：

```
-- add builtin functions that are new to 1.01

create function int4eqoid (int4, oid) returns bool as 'foo'
language 'internal';
create function oideqint4 (oid, int4) returns bool as 'foo'
language 'internal';
create function char2icregexeq (char2, text) returns bool as 'foo'
language 'internal';
create function char2icregexne (char2, text) returns bool as 'foo'
language 'internal';
create function char4icregexeq (char4, text) returns bool as 'foo'
language 'internal';
create function char4icregexne (char4, text) returns bool as 'foo'
language 'internal';
create function char8icregexeq (char8, text) returns bool as 'foo'
language 'internal';
create function char8icregexne (char8, text) returns bool as 'foo'
language 'internal';
create function char16icregexeq (char16, text) returns bool as
'foo'
language 'internal';
create function char16icregexne (char16, text) returns bool as
'foo'
language 'internal';
create function texticregexeq (text, text) returns bool as 'foo'
language 'internal';
create function texticregexne (text, text) returns bool as 'foo'
language 'internal';

-- add builtin functions that are new to 1.01
```

```

create operator = (leftarg = int4, rightarg = oid, procedure =
    int4eqoid);
create operator = (leftarg = oid, rightarg = int4, procedure =
    oideqint4);
create operator ~* (leftarg = char2, rightarg = text, procedure =
    char2icregexeq);
create operator !~* (leftarg = char2, rightarg = text, procedure =
    char2icregexne);
create operator ~* (leftarg = char4, rightarg = text, procedure =
    char4icregexeq);
create operator !~* (leftarg = char4, rightarg = text, procedure =
    char4icregexne);
create operator ~* (leftarg = char8, rightarg = text, procedure =
    char8icregexeq);
create operator !~* (leftarg = char8, rightarg = text, procedure =
    char8icregexne);
create operator ~* (leftarg = char16, rightarg = text, procedure =
    char16icregexeq);
create operator !~* (leftarg = char16, rightarg = text, procedure =
    char16icregexne);
create operator ~* (leftarg = text, rightarg = text, procedure =
    texticregexeq);
create operator !~* (leftarg = text, rightarg = text, procedure =
    texticregexne);

```

## 35.12.2. 详细变更列表

不兼容之处：

- \* 只要用户按照 `MIGRATION_from_1.0_to_1.01` 文件中概述的步骤操作，1.01 与 1.0 数据库向后兼容。

如果未采取这些步骤，1.01 与 1.0 数据库不兼容。

增强：

- \* 为 `libpq` 添加 `PQdisplayTuples()` 并更改 `monitor` 和 `psql` 使用它
- \* 新增 `NeXT` 移植（需要 `SysVIPC` 实现）
- \* 新增 `CAST .. AS` 语法
- \* 添加了 `ASC` 和 `DESC` 关键字
- \* 为 `CREATE FUNCTION` 添加 `'internal'` 作为可能的语言

内部函数是已静态链接

到 `postgres` 后端中的 `C` 函数。

- \* 为系统标识符（表名、

属性名等）添加了新类型 `"name"`。它取代了旧的 `char16` 类型。

`name` 的长度由 `src/Makefile.global` 中的 `NAMEDATALEN` `#define` 设置

- \* 一本描述查询语言的易读参考手册。

- \* 添加了基于主机的访问控制。配置文件 (`$PGDATA/pg_hba`)

用于保存配置数据。如果不要基于主机的访问控制，

请注释掉 `src/Makefile.global` 中的 `HBA=1`。

- \* 更改正则表达式处理为统一使用 Henry Spencer 的正则代码，

与平台无关。正则代码包含在发行版中

- \* 添加了用于不区分大小写正则表达式的函数和操作符。

操作符为 `~*` 和 `!~*`。

- \* `pg_dump` 使用 `COPY` 而非 `SELECT` 循环以获得更好性能

**缺陷修复:**

- \* 修复了一个优化器缺陷, 当在 WHERE 子句的比较中使用函数调用时会导致核心转储
- \* 将所有 getuid 的使用改为 geteuid, 以使用有效 uid
- \* psql 使用 -c 时出错现在返回非零状态
- \* 应用了公开补丁 1-14

## 35.13. 版本 1.0

### 35.13.1. 详细变更列表

**版权变更:**

- \* Postgres 1.0 的版权已放宽为可自由修改并可用于任何目的。  
请阅读 COPYRIGHT 文件。感谢 Michael Stonebraker 教授使之成为可能。

**不兼容之处:**

- \* 日期格式必须为 MM-DD-YYYY (如果使用欧洲风格, 则为 DD-MM-YYYY)。这遵循 SQL-92 规范。
- \* "delimiters" 现在是关键字

**增强:**

- \* 添加了 SQL LIKE 语法
- \* copy 命令现在接受可选的 USING DELIMITER 说明。

分隔符可以是任何单字符串。

- \* 添加了 IRIX 5.3 移植。

感谢 Paul Walmsley 等人。

- \* 更新 pg\_dump 以配合新 libpq 工作
- \* psql 中添加了 \d

感谢 Keith Parks

- \* 对于使用 POSIX 正则的体系结构, 正则性能通过缓存预编译模式得到了改进。

感谢 Alistair Crooks

- \* 新版本的 libpq++

感谢 William Wanders

**缺陷修复:**

- \* 可以在 createuser 脚本中指定任意用户 id
- \* psql 中连接其他数据库的 \c 现在可以工作了。
- \* 修复了 float4inc() 的错误 pg\_proc 条目
- \* 设置了 usecreatedb 字段的用户现在可以创建数据库, 而无需是 usesuper
- \* 当条目不再有任何权限时移除访问控制条目
- \* 修复了不可移植的 datetimes 实现
- \* 在 src/backend/Makefile 中添加 kerberos 标志
- \* libpq 现在支持 kerberos
- \* 用户手册中的排印错误已更正。
- \* 多索引的 btree 从未工作过, 现在尝试使用时会明确告知它们不可用

## 35.14. Postgres95 Beta 0.03

### 35.14.1. 详细变更列表

不兼容变更：

- \* BETA-0.3 与用以前版本创建的数据库不兼容（由于系统目录更改和索引结构更改）。
- \* 双引号（"）作为字符串字面量的引号字符已被废弃；你需要将它们转换为单引号（'）。
- \* 聚合的名称（如 `int4sum`）已按照 SQL 标准重命名（如 `sum`）。
- \* `CHANGE ACL` 语法被 `GRANT/REVOKE` 语法取代。
- \* 浮点字面量（如 `3.14`）现在是 `float4` 类型（而不是以前版本的 `float8`）；如果你依赖它是 `float8` 类型，可能需要进行类型转换。如果忽略类型转换而把浮点字面量赋给 `float8` 类型的字段，可能会得到错误的存储值！
- \* `LIBPQ` 经过彻底翻新，前端应用程序可以连接多个后端
- \* `pg_user` 中的 `usesysid` 字段已从 `int2` 改为 `int4`，以允许更宽范围的 Unix 用户 `id`。
- \* `netbsd/freebsd/bsd` 操作系统移植已合并为单一的 `BSD44_derived` 移植。（感谢 Alistair Crooks）

SQL 标准符合性（以下详述使 `postgres95`

更符合 SQL-92 标准的更改）：

- \* 以下 SQL 类型现在是内建的：`smallint`、`int(eger)`、`float`、`real`、`char(N)`、`varchar(N)`、`date` 和 `time`。

以下是现有 `postgres` 类型的别名：

`smallint` -> `int2`

`integer`, `int` -> `int4`

`float`, `real` -> `float4`

`char(N)` 和 `varchar(N)` 实现为截断的 `text` 类型。此外，`char(N)` 会做空白填充。

- \* 用单引号（'）引用字符串字面量；''（除 \' 外）也可用于在字符串中插入单引号
- \* 使用 SQL 标准聚合名（`MAX`、`MIN`、`AVG`、`SUM`、`COUNT`）（此外，聚合现在可以重载，即你可以定义接受用户自定义类型的 `MAX` 聚合。）
- \* 移除 `CHANGE ACL`。添加 `GRANT/REVOKE` 语法。
  - 可以使用 "GROUP" 关键字把权限授予组。

例如：

```
GRANT SELECT ON foobar TO GROUP my_group;
```

关键字 'PUBLIC' 也受支持，表示所有用户。

每次只能对一个用户或组授予或撤销权限。

不支持 "WITH GRANT OPTION"。只有类所有者可以更改访问控制

- 默认访问控制是给用户只读访问权限。

你必须显式授予用户插入/更新权限。要更改这一点, 请修改  
`src/backend/utils/acl.h`  
 中定义 `ACL_WORLD_DEFAULT`  
 的那一行

#### 缺陷修复:

- \* 空表上的聚合不被运行的缺陷已修复。现在, 在空表上运行的聚合将返回聚合的初始条件。因此, 空表的 `COUNT` 现在会正确返回 0。空表的 `MAX/MIN` 将返回值为 `NULL` 的元组。
- \* 允许在 `monitor` 中使用 `\;`;
- \* `LISTEN/NOTIFY` 异步通知机制现在可以工作
- \* 规则动作体中的 `NOTIFY` 现在可以工作
- \* 哈希索引可以工作了, 访问方法总体上性能应该更好。创建大型 `btree` 索引应该快得多。(感谢 Paul Aoki)

#### 其他更改与增强:

- \* 添加了用于解释查询执行计划的 `EXPLAIN` 语句 (例如 `"EXPLAIN SELECT * FROM EMP"` 会打印出该查询的执行计划)。
- \* `WARN` 和 `NOTICE` 消息不再带时间戳。要打开错误消息的时间戳, 请取消注释 `src/backend/utils/elog.h` 中的这一行:  

```
/* define ELOG_TIMESTAMPS */
```
- \* 访问控制被违反时, 会给出消息  
`"Either no such class or insufficient privilege"`  
 。这与找不到类时返回的消息相同。这可以阻止无权限的用户猜测受权限保护的类的存在。
- \* 还做了一些用户不可见的系统目录更改。

#### libpgtcl 更改:

- \* `"pg_result"` tcl 命令添加了 `-oid` 选项。  
`pg_result -oid` 返回最后插入元组的 `oid`。如果最后的命令不是 `INSERT`, 则 `pg_result -oid` 返回 ""。
- \* 大对象接口以 `pg_lo*` tcl 命令的形式提供:  
`pg_lo_open`, `pg_lo_close`, `pg_lo_creat`, etc.

#### 可移植性增强与新移植:

- \* `flex/lex` 的问题已经解决。现在, 在任何平台上都应该可以用 `flex` 代替 `lex`。我们不再根据你所用的平台假设你使用的词法分析器。
- \* 现在支持 `Linux-ELF` 移植。已测试多种配置: 已知以下配置可以工作:  
`kernel 1.2.10`, `gcc 2.6.3`, `libc 4.7.2`, `flex 2.5.2`, `bison 1.24`  
 所有内容均为 `ELF` 格式,

#### 新实用程序:

- \* `ipcclean` 已加入发行版  
 通常不需要运行 `ipcclean`, 如果你的后端崩溃并留下共享内存段, `ipcclean` 会为你清理它们。

#### 新文档:

- \* 用户手册已经修订并添加了 `libpq` 文档。

## 35.15. Postgres95 Beta 0.02

### 35.15.1. 详细变更列表

不兼容变更：

- \* 创建数据库的 SQL 语句是 'CREATE DATABASE' 而不是 'CREATEDB'。
- 类似地，删除数据库是 'DROP DATABASE' 而不是 'DESTROYDB'。
- 但可执行文件名 'createdb' 和 'destroydb' 保持不变。

新工具：

- \* pgperl - Postgres95 的 Perl (4.036) 接口
- \* pg\_dump - 把 postgres 数据库转储为包含查询命令的脚本文件的工具。脚本文件为 ASCII 格式，可用于重建数据库，甚至在其他机器和其他体系结构上。（也适合把 Postgres 4.2 数据库转换为 Postgres95 数据库。）

以下移植已并入 postgres95-beta-0.02：

- \* Alistair Crooks 的 NetBSD 移植
- \* Mike Tung 的 AIX 移植
- \* Jon Forrest 的 Windows NT 移植（内容更多但尚未完成）
- \* Brian Gallew 的 Linux ELF 移植

postgres95-beta-0.02 中修复了以下缺陷：

- \* COPY OUT 中换行未转义以及首属性为 '.' 时 COPY OUT 的问题
- \* createuser 中无法键入回车以使用默认用户 id
- \* 大表上的 SELECT DISTINCT 崩溃
- \* Linux 安装问题
- \* monitor 不允许使用 'localhost' 作为 PGHOST
- \* psql 执行 \c 或 \l 时核心转储
- \* src/bin/pgtclsh/Makefile 中缺少 "pgtclsh" 目标
- \* libpgtcl 有硬编码的默认端口号
- \* SELECT DISTINCT INTO TABLE 挂起
- \* CREATE TYPE 不接受 'variable' 作为 internallength
- \* 在 SELECT 中使用多个聚合时结果错误

## 35.16. Postgres95 Beta 0.01

初始发行版。

## 35.17. 计时结果

这些计时结果来自使用以下命令运行回归测试

```
% cd src/test/regress
% make all
% time make runtest
```

Linux 2.0.27 下的计时在每次运行之间似乎有大约 5% 的波动，这大概是多任务系统调度的不确定性所致。

## 35.17.1. 版本 6.4beta

本版本的耗时与以前版本不能直接比较，因为纳入了一些额外的回归测试。不过总体而言，v6.4 应比前一版本略快（谢谢 Bruce!）。

| 时间    | 系统                                                                       |
|-------|--------------------------------------------------------------------------|
| 02:26 | Dual Pentium Pro 180, 96MB, UW-SCSI, Linux 2.0.30, gcc 2.7.2.1 -O2 -m486 |

## 35.17.2. 版本 6.3

本版本的耗时与以前版本不能直接比较，因为纳入了一些额外的回归测试并移除了一些涉及时间旅行的过时测试。不过总体而言，v6.3 比以前的版本快得多（谢谢 Bruce!）。

| 时间    | 系统                                                                       |
|-------|--------------------------------------------------------------------------|
| 02:30 | Dual Pentium Pro 180, 96MB, UW-SCSI, Linux 2.0.30, gcc 2.7.2.1 -O2 -m486 |
| 04:12 | Dual Pentium Pro 180, 96MB, EIDE, Linux 2.0.30, gcc 2.7.2.1 -O2 -m486    |

## 35.17.3. 版本 6.1

| 时间    | 系统                                                             |
|-------|----------------------------------------------------------------|
| 06:12 | Pentium Pro 180, 32MB, EIDE, Linux 2.0.30, gcc 2.7.2 -O2 -m486 |
| 12:06 | P-100, 48MB, Linux 2.0.29, gcc                                 |
| 39:58 | Sparc IPC 32MB, Solaris 2.5, gcc 2.7.2.1 -O -g                 |

---

## 部分 IV. 程序员指南

关于扩展 Postgres 的信息。

---

# 目录

|                                           |     |
|-------------------------------------------|-----|
| 36. 引言 .....                              | 317 |
| 36.1. 资源 .....                            | 317 |
| 36.2. 术语 .....                            | 318 |
| 36.3. 记法 .....                            | 318 |
| 36.4. Y2K 声明 .....                        | 319 |
| 36.5. 版权与商标 .....                         | 319 |
| 37. 体系结构 .....                            | 321 |
| 37.1. Postgres 体系结构概念 .....               | 321 |
| 38. 扩展 SQL: 概览 .....                      | 324 |
| 38.1. 可扩展性如何运作 .....                      | 324 |
| 38.2. Postgres 类型系统 .....                 | 324 |
| 38.3. 关于 Postgres 系统目录 .....              | 324 |
| 39. 扩展 SQL: 函数 .....                      | 327 |
| 39.1. 查询语言 (SQL) 函数 .....                 | 327 |
| 39.1.1. 基本类型上的 SQL 函数 .....               | 327 |
| 39.1.2. 复合类型上的 SQL 函数 .....               | 327 |
| 39.2. 编程语言函数 .....                        | 329 |
| 39.2.1. 基本类型上的编程语言函数 .....                | 329 |
| 39.2.2. 复合类型上的编程语言函数 .....                | 331 |
| 39.2.3. 注意事项 .....                        | 332 |
| 40. 扩展 SQL: 类型 .....                      | 333 |
| 40.1. 用户定义的类型 .....                       | 333 |
| 40.1.1. 用户定义类型所需的函数 .....                 | 333 |
| 40.1.2. 大对象 .....                         | 334 |
| 41. 扩展 SQL: 操作符 .....                     | 335 |
| 42. 扩展 SQL: 聚合 .....                      | 336 |
| 43. Postgres 规则系统 .....                   | 338 |
| 43.1. 什么是查询树? .....                       | 338 |
| 43.1.1. 查询树的组成部分 .....                    | 338 |
| 43.2. 视图和规则系统 .....                       | 339 |
| 43.2.1. Postgres 中视图的实现 .....             | 339 |
| 43.2.2. SELECT 规则如何工作 .....               | 340 |
| 43.2.3. 非 SELECT 语句中的视图规则 .....           | 346 |
| 43.2.4. Postgres 中视图的能力 .....             | 347 |
| 43.2.5. 实现上的副作用 .....                     | 347 |
| 43.3. INSERT、UPDATE 和 DELETE 上的规则 .....   | 348 |
| 43.3.1. 与视图规则的区别 .....                    | 348 |
| 43.3.2. 这些规则如何工作 .....                    | 348 |
| 43.3.3. 与视图的协作 .....                      | 352 |
| 43.4. 规则和权限 .....                         | 358 |
| 43.5. 规则与触发器 .....                        | 359 |
| 44. 索引扩展接口 .....                          | 362 |
| 45. GiST Indices .....                    | 367 |
| 46. 链接动态装载的函数 .....                       | 369 |
| 46.1. ULTRIX .....                        | 370 |
| 46.2. DEC OSF/1 .....                     | 370 |
| 46.3. SunOS 4.x、Solaris 2.x 和 HP-UX ..... | 370 |
| 47. 触发器 .....                             | 372 |
| 47.1. 触发器创建 .....                         | 372 |
| 47.2. 与触发器管理器的交互 .....                    | 373 |
| 47.3. 数据更改的可见性 .....                      | 374 |

|                      |     |
|----------------------|-----|
| 47.4. 示例 .....       | 374 |
| 48. 服务器编程接口 .....    | 378 |
| 48.1. 接口函数 .....     | 378 |
| 48.2. 接口支持函数 .....   | 386 |
| 48.3. 内存管理 .....     | 399 |
| 48.4. 数据更改的可见性 ..... | 400 |
| 48.5. 示例 .....       | 400 |
| 49. 过程语言 .....       | 403 |
| 49.1. 安装过程语言 .....   | 403 |
| 49.2. PL/pgSQL ..... | 404 |
| 49.2.1. 概述 .....     | 404 |
| 49.2.2. 描述 .....     | 404 |
| 49.2.3. 示例 .....     | 410 |
| 49.3. PL/Tcl .....   | 412 |
| 49.3.1. 概述 .....     | 412 |
| 49.3.2. 描述 .....     | 412 |

---

# 第 36 章 引言

本文档是 PostgreSQL<sup>1</sup> 数据库管理系统的程序员手册，该系统最初是在加州大学伯克利分校开发的。PostgreSQL 基于 Postgres release 4.2<sup>2</sup>。由 Michael Stonebraker 教授领导的 Postgres 项目曾获得国防高级研究计划局（DARPA）、陆军研究办公室（ARO）、国家科学基金会（NSF）以及 ESL 公司的资助。

本手册的第一部分解释 Postgres 的可扩展性方式，并描述用户如何通过添加用户定义的类型、操作符、聚合以及查询语言和编程语言函数来扩展 Postgres。在讨论 Postgres 规则系统之后，我们讨论触发器和 SPI 接口。本手册最后详细描述各种语言的编程接口和支持库。

我们假定读者熟悉 UNIX 和 C 编程。

## 36.1. 资源

本手册集被组织为若干部分：

### 教程

面向新用户的介绍。不涉及高级特性。

### 用户指南

面向用户的一般信息，包括可用的命令和数据类型。

### 程序员指南

面向应用程序员的高级信息。主题包括类型和函数可扩展性、库接口以及应用设计问题。

### 管理员指南

安装和管理信息。受支持的机器列表。

### 开发者指南

面向 Postgres 开发者的信息。此书面向为 Postgres 项目做贡献的人；应用开发的信息见程序员指南。目前包含在程序员指南中。

### 参考手册

关于命令语法的详细参考信息。目前包含在用户指南中。

除本手册集之外，还有其他资源可以帮助你安装和使用 Postgres：

### man 手册页

man 手册页包含关于命令语法的一般信息。

### FAQ

常见问题（FAQ）文档既阐述一般性问题，也阐述一些平台特定的问题。

---

<sup>1</sup><http://postgresql.org/>

<sup>2</sup><http://s2k-ftp.CS.Berkeley.EDU:8000/postgres/postgres.html>

## README

某些附带的软件包提供了 README 文件。

## 网站

Postgres<sup>3</sup> 网站上可能有发行包中未包含的信息。那里有一个 mhonarc 邮件列表流量存档，是许多主题的丰富资源。

## 邮件列表

Postgres Questions<sup>4</sup> 邮件列表是一个让用户问题得到解答的好地方。还有其他可用的邮件列表；详情请参阅该网页。

## 你自己！

Postgres 是一项开源产品。因此，它的持续支持依赖于用户社区。当你开始使用 Postgres 时，你会依赖他人的帮助——无论是通过文档还是通过邮件列表。请考虑回馈你的知识。如果你学到了文档中没有的东西，把它写出来并贡献回去。如果你给代码添加了特性，也把它们贡献出去。即使经验不多的人也能为文档提供更正和小改动，而这是一个很好的起点。Postgres Documentation<sup>5</sup> 邮件列表就是开始的地方。

## 36.2. 术语

在后面的文档中，站点（site）可以解释为安装 Postgres 的主机。由于在单台主机上可以安装多套 Postgres 数据库，这个术语更确切地指任何一套特定的已安装的 Postgres 二进制文件和数据库。

Postgres 超级用户是名为 *postgres* 的用户，他拥有 Postgres 的二进制文件和数据库文件。作为数据库超级用户，所有的保护机制都可以被绕过，任何数据都可以任意访问。此外，Postgres 超级用户还被允许执行一些通常不向所有用户开放的支持程序。注意，Postgres 超级用户不是 Unix 超级用户（后者将称为 *root*）。出于安全原因，超级用户应具有非零的用户标识符（UID）。

数据库管理员（DBA）负责安装 Postgres 并为站点建立实施安全策略的机制。DBA 可以按下述方法添加新用户，并维护一组供 *createdb* 使用的模板数据库。

*postmaster* 是为发往 Postgres 系统的请求充当中转站的进程。前端应用连接到 *postmaster*，由它跟踪系统错误和后端进程之间的通信。*postmaster* 可以接受若干命令行参数来调整其行为。不过，只有当你打算运行多个站点或非默认站点时才需要提供参数。

Postgres 后端（实际的可执行程序 *postgres*）可以由 Postgres 超级用户直接在用户 shell 中执行（以数据库名作为参数）。但这样做会绕过与 *postmaster*/站点关联的共享缓冲池和锁表，因此在多用户站点中不建议这样做。

## 36.3. 记法

文件名前面的 “...” 或 */usr/local/pgsql/* 用来表示 Postgres 超级用户主目录的路径。

在命令概要中，方括号 “[” 和 “]” 表示可选的短语或关键字。花括号 (“{” 和 “}”) 中含有竖线 (“|”) 时表示你必须选择其中一个。

<sup>3</sup> [postgresql.org](http://postgresql.org)

<sup>4</sup> <mailto:questions@postgresql.org>

<sup>5</sup> <mailto:docs@postgresql.org>

在示例中，圆括号（“(”和“)”）用于给布尔表达式分组。“|”是布尔运算符 OR。

示例将展示从不同账户和程序执行的命令。从 root 账户执行的命令前带有“>”。从 Postgres 超级用户账户执行的命令前带有“%”，而从无特权用户账户执行的命令前带有“\$”。SQL 命令前带有“=>”提示符，也可能没有前导提示符，视上下文而定。

### 注意

在撰写本文时（Postgres v6.4），整个文档集中标记命令的记法并不完全一致。如发现问题请报告给文档邮件列表<sup>6</sup>。

## 36.4. Y2K 声明

### 作者

由 Thomas Lockhart<sup>7</sup> 写于 1998-10-22。

PostgreSQL 全球开发队伍将 Postgres 软件代码树作为公共服务提供，不对它的行为或性能作任何保证，也不承担任何责任。不过，在撰写本文时：

- 本声明的作者自 1996 年 11 月起就是 Postgres 支持团队的一名志愿者，他不知道 Postgres 代码库中存在任何与 2000 年 1 月 1 日前后的时间过渡（Y2K）相关的问题。
- 本声明的作者没有听说在回归测试中或在 Postgres 近期或当前版本的其他实际使用中发现过任何 Y2K 问题的报告。考虑到已安装的用户基础以及用户在支持邮件列表上的积极参与，如果存在问题，我们本应有所耳闻。
- 就作者所知，Postgres 对用两位数字年份指定的日期所做的假定记录在当前用户指南<sup>8</sup>关于数据类型的章节中。对于两位年份，关键的过渡年份是 1970 而不是 2000；例如“70-01-01”被解释为“1970-01-01”，而“69-01-01”被解释为“2069-01-01”。
- 底层操作系统中与获取“当前时间”相关的任何 Y2K 问题都可能传播为 Postgres 中表面的 Y2K 问题。

关于 Y2K 问题的进一步讨论，特别是与开源、免费软件相关的讨论，请参阅 The Gnu Project<sup>9</sup> 和 The Perl Institute<sup>10</sup>

## 36.5. 版权与商标

PostgreSQL 版权所有 (C) 1996-8，归 PostgreSQL 全球开发组所有，并按照 Berkeley 许可的条款分发。

Postgres95 版权所有 (C) 1994-5，归加利福尼亚大学董事会所有。特此授权，可为任何目的使用、复制、修改和分发本软件及其文档，无需付费，也无需书面协议，但条件是所有副本中都必须包含上述版权声明、本段以及后续两段内容。

<sup>6</sup><mailto:docs@postgresql.org>

<sup>7</sup><mailto:lockhart@alumni.caltech.edu>

<sup>8</sup><http://www.postgresql.org/docs/user/datatype.htm>

<sup>9</sup><http://www.gnu.org/software/year2000.html>

<sup>10</sup><http://language.perl.com/news/y2k.html>

在任何情况下，加利福尼亚大学都不对任何一方承担因使用本软件及其文档而产生的任何直接、间接、特殊、附带或后果性损害赔偿责任，包括利润损失，即使加利福尼亚大学已被告知此类损害发生的可能性。

加利福尼亚大学明确不作任何担保，包括但不限于适销性和特定用途适用性的默示担保。本协议项下提供的软件按"原样"提供，加利福尼亚大学没有义务提供维护、支持、更新、增强或修改。

UNIX 是 X/Open, Ltd. 的商标。Sun4、SPARC、SunOS 和 Solaris 是 Sun Microsystems, Inc. 的商标。DEC、DECstation、Alpha AXP 和 ULTRIX 是 Digital Equipment Corp. 的商标。PA-RISC 和 HP-UX 是 Hewlett-Packard Co. 的商标。OSF/1 是 Open Software Foundation 的商标。

---

# 第 37 章 体系结构

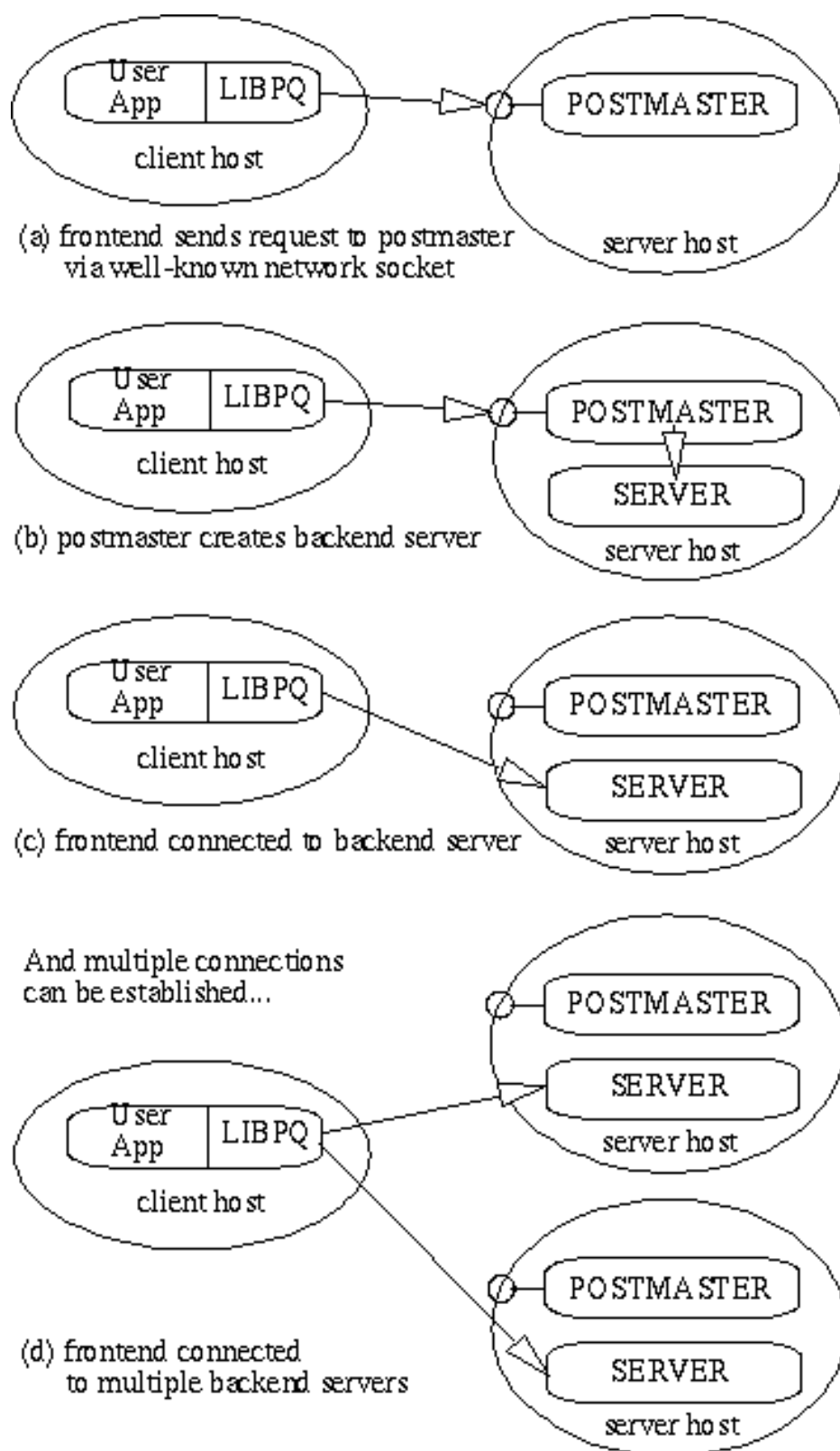
## 37.1. Postgres 体系结构概念

在继续之前，你应当理解 Postgres 系统的基本体系结构。理解 Postgres 各部分的交互方式会让下一章更容易理解。用数据库术语来说，Postgres 使用简单的"每用户一进程"客户端/服务器模型。一个 Postgres 会话由下列协作的 Unix 进程（程序）组成：

- 一个管理守护进程（postmaster），
- 用户的前端应用（例如 psql 程序），以及
- 一个或多个后端数据库服务器（postgres 进程本身）。

单个 postmaster 管理单个主机上给定的数据库集合。这样的数据库集合称为一个安装或站点。希望访问安装中某个数据库的前端应用调用该库。该库通过网络把用户请求发送给 postmaster (图 37.1(a))，后者再启动一个新的后端服务器进程 (图 37.1(b))

图 37.1. 如何建立连接



并把前端进程连接到新服务器（图 37.1(c)）。从那时起，前端进程和后端服务器之间的通信不再需要 `postmaster` 干预。因此，`postmaster` 总是在运行并等待请求，而前端和后端进程则来来去去。`libpq` 库允许单个前端对后端进程建立多个连接。但前端应用仍然是一个单线程进程。`libpq` 目前不支持多线程的前端/后端连接。这一体系结构的一个含义是 `postmaster` 和后端总是运行在同一台机器（数据库服务器）上，而前端应用可以运行在任何地方。你应当牢记这一点，因为在客户端机器上可以访问的文件在数据库服务器机器上可能无法访问（或者只能用不同的文件名访问）。你还应当知道 `postmaster` 和 `postgres` 服务器以 `Postgres "superuser."`（超级用户）的 `user-id` 运行。注意 `Postgres` 超级用户不必是任何特定的用户（例如名为 `"postgres"` 的用户），虽然许多系统就是这样安装的。而且，`Postgres` 超级用户绝对不应该是 `Unix` 超级用户 `"root"`！无论如何，与数据库相关的所有文件都应属于这个 `Postgres` 超级用户。

---

## 第 38 章 扩展 SQL：概览

在后面的各节中，我们将讨论如何通过添加下列内容来扩展 Postgres 的 SQL 查询语言：

- 函数
- 类型
- 操作符
- 聚集

### 38.1. 可扩展性如何运作

Postgres 之所以可扩展，是因为它的操作是目录驱动的。如果你熟悉标准的关系系统，就知道它们把关于数据库、表、列等的信息存储在通常所谓的系统目录中。

（有些系统称之为数据字典。）这些目录在用户看来是像其他类一样的类，但 DBMS 把它的内部簿记存储在其中。Postgres 与标准关系系统之间的一个关键区别是，Postgres 在其目录中存储的信息多得多——不仅有关于表和列的信息，还有关于其类型、函数、访问方法等的信息。

用户可以修改这些类，而由于 Postgres 把它的内部操作建立在这些类之上，这意味着 Postgres 可以由用户扩展。相比之下，传统数据库系统只能通过更改 DBMS 内部的硬编码过程或装载由 DBMS 厂商专门编写的模块来扩展。

Postgres 与大多数其他数据管理器的另一个不同之处在于，服务器可以通过动态装载把用户编写的代码纳入自身。也就是说，

用户可以指定一个实现新类型或函数的目标代码文件（例如编译好的 .o 文件或共享库），Postgres 会按需装载它。用 SQL 编写的代码加入服务器更是轻而易举。这种“即时”修改操作的能力使 Postgres 特别适合新应用和存储结构的快速原型开发。

### 38.2. Postgres 类型系统

Postgres 的类型系统可以从几个方面细分。类型分为基本类型和复合类型。基本类型是像 int4 这样用 C 之类的语言实现的类型。它们大体对应于通常所说的“抽象数据类型”；Postgres 只能通过用户提供的方法操作这类类型，并且只在用户描述它们的程度上理解这类类型的行为。

复合类型在用户创建类时创建。EMP 是复合类型的一个例子。

Postgres 只以一种方式存储这些类型（在存储类的所有实例的文件内），但用户可以从查询语言“窥视”这些类型的属性，并通过（例如）在属性上定义索引来优化它们的检索。Postgres 的基本类型进一步分为内置类型和用户定义类型。内置类型（如 int4）是编译进系统的那些。用户定义类型是用户以下面要描述的方式创建的类型。

### 38.3. 关于 Postgres 系统目录

介绍了基本的可扩展性概念之后，我们现在可以看看目录实际上是如何组织的。

现在可以跳过本节，但后面的某些章节如果没有这里给出的信息将无法理解，所以请给这一页做个标记以便以后参考。所有系统目录的名称都以 pg\_ 开头。下面的类包含可能对最终用户有用的信息。（还有很多其他系统目录，但很少需要直接查询它们。）

**表 38.1. Postgres 系统目录**

| 目录名          | 描述  |
|--------------|-----|
| pg_database  | 数据库 |
| pg_class     | 类   |
| pg_attribute | 类属性 |

| 目录名          | 描述             |
|--------------|----------------|
| pg_index     | 辅助索引           |
| pg_proc      | 过程（C 和 SQL 都是） |
| pg_type      | 类型（基本和复合都是）    |
| pg_operator  | 操作符            |
| pg_aggregate | 聚集和聚集函数        |
| pg_am        | 访问方法           |
| pg_amop      | 访问方法操作符        |
| pg_amproc    | 访问方法支持函数       |
| pg_opclass   | 访问方法操作符类       |

图 38.1. Postgres 的主要系统目录

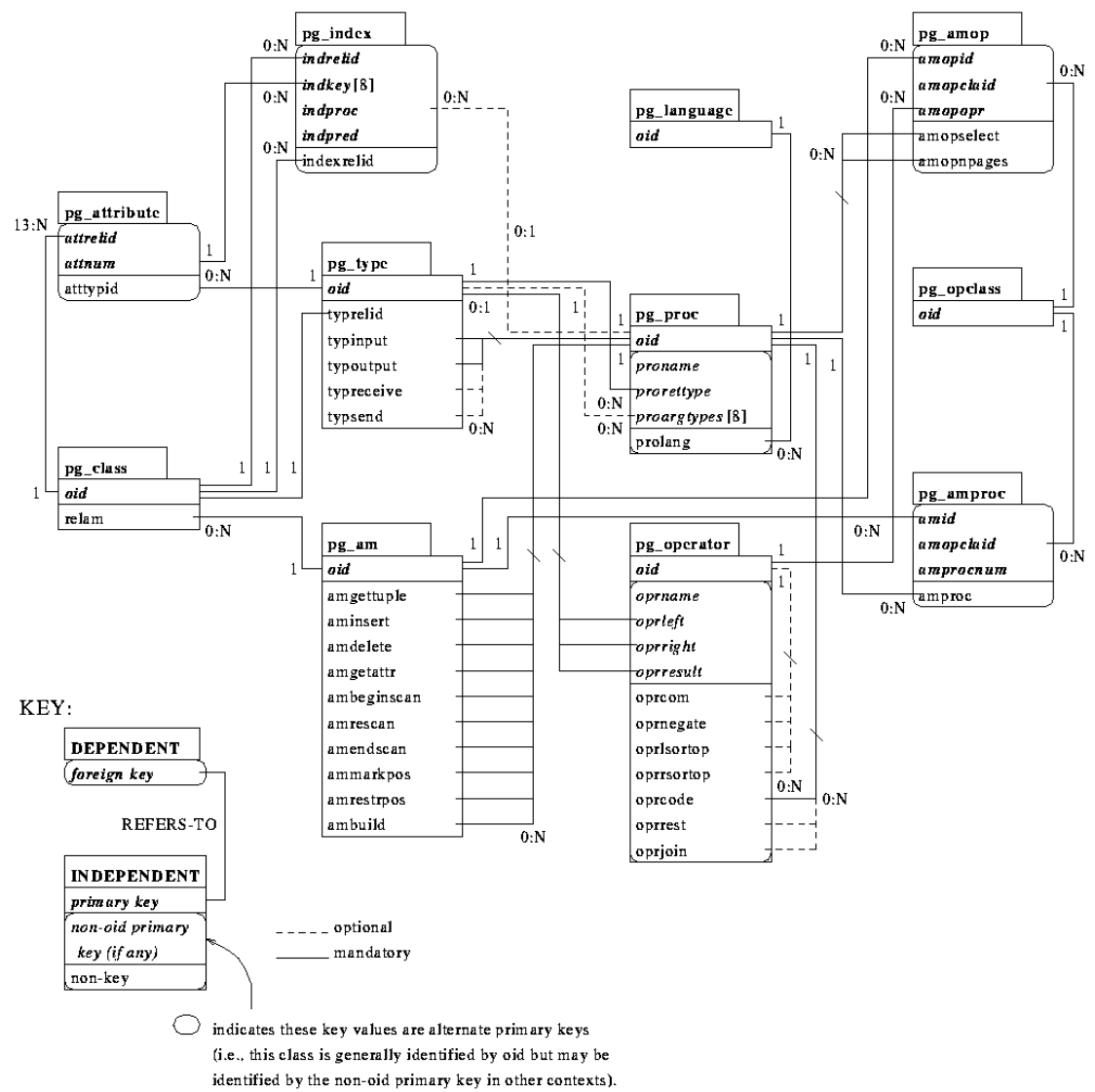


Figure 3. The major POSTGRES system catalogs.

参考手册对这些目录及其属性给出了更详细的解释。不过，图 38.1 展示了系统目录中的主要实体及其关系。（不引用其他实体的属性没有显示，除非它们是主键的一部分。）在你真正开始查看目录的内容并了解它们如何相互关联之前，这张图或多或少难以理解。就目前而言，从这张图中要了解的要如下：

- 在后面的几节中，我们将给出系统目录上的各种连接查询，它们显示我们扩展系统所需的信息。看这张图应该会让其中一些连接查询（它们通常是三路或四路连接）更容易理解，因为你将能看到查询中使用的属性在其他类中构成外键。
- 许多不同的特性（类、属性、函数、类型、访问方法等）在这个模式中紧密集成。一条简单的 `create` 命令就可能修改其中许多目录。
- 类型和过程是这个模式的核心。

### 注意

我们或多或少可以互换地使用过程和函数这两个词。

- 几乎每个目录都包含对这两个类之一或两者的实例的引用。例如，Postgres 经常使用类型签名（例如函数和操作符的签名）来标识其他目录的唯一实例。
- 有许多属性和关系的含义是显而易见的，但也有许多（特别是那些与访问方法有关的）并非如此。`pg_am`、`pg_amop`、`pg_amproc`、`pg_operator` 和 `pg_opclass` 之间的关系特别难以理解，我们将在讨论了基本扩展之后（在把类型和操作符接口到索引的一节中）深入描述它们。

---

## 第 39 章 扩展 SQL：函数

事实证明，定义新类型的一部分工作就是定义描述其行为的函数。因此，虽然可以不定义新类型就定义新函数，反过来却不成立。所以我们在描述如何向 Postgres 添加新类型之前先描述如何添加新函数。Postgres SQL 提供两类函数：查询语言函数（用 SQL 编写的函数）和编程语言函数（用 C 等编译型编程语言编写的函数）。每类函数都可以接受基本类型、复合类型或它们的某种组合作为参数。此外，两类函数都可以返回基本类型或复合类型。定义 SQL 函数更容易，所以我们从它开始。本节的示例也可以在 `funcs.sql` 和 `C-code/funcs.c` 中找到。

### 39.1. 查询语言（SQL）函数

#### 39.1.1. 基本类型上的 SQL 函数

最简单的 SQL 函数没有参数，只是返回一个基本类型，例如 `int4`：

```
CREATE FUNCTION one() RETURNS int4
AS 'SELECT 1 as RESULT' LANGUAGE 'sql';
```

```
SELECT one() AS answer;
```

```
+-----+
|answer |
+-----+
|1      |
+-----+
```

注意我们为函数定义了一个目标列表（名为 `RESULT`），但调用该函数的查询的目标列表覆盖了函数的目标列表。因此结果被标记为 `answer` 而不是 `one`。

定义以基本类型为参数的 SQL 函数几乎同样容易。在下面的示例中，注意我们如何在函数内把参数引用为 `$1` 和 `$2`。

```
CREATE FUNCTION add_em(int4, int4) RETURNS int4
AS 'SELECT $1 + $2;' LANGUAGE 'sql';
```

```
SELECT add_em(1, 2) AS answer;
```

```
+-----+
|answer |
+-----+
|3      |
+-----+
```

#### 39.1.2. 复合类型上的 SQL 函数

指定带复合类型（如 `EMP`）参数的函数时，我们不仅要指明想要哪个参数（如上面用 `$1` 和 `$2`），还要指明该参数的属性。例如，取函数 `double_salary`，它计算你的薪资翻倍后的值。

```
CREATE FUNCTION double_salary(EMP) RETURNS int4
AS 'SELECT $1.salary * 2 AS salary;' LANGUAGE 'sql';
```

```
SELECT name, double_salary(EMP) AS dream
FROM EMP
WHERE EMP.dept = 'toy';
```

```
+-----+-----+
|name | dream |
+-----+-----+
|Sam  | 2400  |
+-----+-----+
```

注意语法 `$1.salary` 的使用。在进入返回复合类型的函数这一主题之前，我们必须先介绍投影属性的函数记法。简单的解释是：`attribute(class)` 和 `class.attribute` 这两种记法通常可以互换使用。

```
--
-- this is the same as:
-- SELECT EMP.name AS youngster FROM EMP WHERE EMP.age < 30
--
SELECT name(EMP) AS youngster
FROM EMP
WHERE age(EMP) < 30;
```

```
+-----+
|youngster |
+-----+
|Sam       |
+-----+
```

然而如下文所见，情况并不总是如此。当我们想使用返回单个实例的函数时，这种函数记法很重要。我们通过在函数内逐属性地组装整个实例来实现。这是一个返回单个 EMP 实例的函数示例：

```
CREATE FUNCTION new_emp() RETURNS EMP
AS 'SELECT \'None\'::text AS name,
    1000 AS salary,
    25 AS age,
    \'none\'::text AS dept;'
LANGUAGE 'sql';
```

在本例中，我们为每个属性都指定了常量值，但这些常量可以换成任何计算或表达式。定义这样的函数可能有些棘手。较重要的一些注意事项如下：

- 查询中的目标列表顺序必须与定义该复合类型的 CREATE TABLE 语句（或执行 `.*` 查询时）中属性出现的顺序完全相同。
- 你必须（使用 `::`）非常小心地对表达式进行类型转换，否则会看到如下错误：

```
WARN::function declared to return type EMP does not
retrieve (EMP.*)
```

- 调用返回实例的函数时，我们无法检索整个实例。我们必须要么从实例中投影出一个属性，要么把整个实例传递给另一个函数。

```
SELECT name(new_emp()) AS nobody;
```

```
+-----+
```

```
|nobody |
+-----+
|None   |
+-----+
```

- 一般而言，我们必须使用函数语法投影函数返回值的属性，原因就是解析器还不理解另一种（点）语法与函数调用组合使用做投影。

```
SELECT new_emp().name AS nobody;
WARN:parser: syntax error at or near "."
```

SQL 查询语言中任何命令的集合都可以打包到一起并定义为函数。这些命令既可以是更新（即 insert、update 和 delete），也可以是 select 查询。但最后一条命令必须是一个 select，它返回函数返回类型所指定的内容。

```
CREATE FUNCTION clean_EMP () RETURNS int4
AS 'DELETE FROM EMP WHERE EMP.salary <= 0;
SELECT 1 AS ignore_this'
LANGUAGE 'sql';
```

```
SELECT clean_EMP();
```

```
+---+
|x |
+---+
|1 |
+---+
```

## 39.2. 编程语言函数

### 39.2.1. 基本类型上的编程语言函数

在内部，Postgres 把基本类型视为一个“内存块”。你针对某个类型定义的用户自定义函数转而定义了 Postgres 能对它进行操作的方式。也就是说，Postgres 只负责在磁盘上存储和检索数据，而使用你的用户自定义函数来输入、处理和输出数据。基本类型可以有以下三种内部格式之一：

- 传值，定长
- 传引用，定长
- 传引用，变长

传值类型的长度只能是 1、2 或 4 字节（即使你的计算机支持其他大小的传值类型）。Postgres 本身只按传值方式传递整数类型。你应当小心地定义类型，使其在所有体系结构上具有相同的大小（以字节计）。例如，long 类型是危险的，因为它在某些机器上是 4 字节而在另一些机器上是 8 字节；而 int 类型在大多数 UNIX 机器上是 4 字节（但在大多数个人电脑上不是）。UNIX 机器上 int4 类型的一个合理实现可以是：

```
/* 4-byte integer, passed by value */
typedef int int4;
```

另一方面，任何大小的定长类型都可以按传引用方式传递。例如，下面是一个 Postgres 类型的实现示例：

```
/* 16-byte structure, passed by reference */
typedef struct {
    char data[16];
} char16;
```

在 Postgres 函数中传入和传出这类类型时只能使用指向它们的指针。最后，所有变长类型也必须以传引用方式传递。所有变长类型都必须以一个恰好 4 字节的长度字段开始，而该类型中要存储的所有数据都位于紧随该长度字段之后的内存中。长度字段是结构的总长度（即它包含长度字段本身的大小）。我们可以这样定义 text 类型：

```
typedef struct {
    int4 length;
    char data[1];
} text;
```

显然，这里所示的 data 字段不足以容纳所有可能的字符串——在 C 中无法声明这样的结构。操作变长类型时，我们必须小心分配正确数量的内存并初始化长度字段。例如，如果我们要在 text 结构中存储 40 字节，可以使用如下代码片段：

```
#include "postgres.h"
...
char buffer[40]; /* our source data */
...
text *destination = (text *) palloc(VARHDRSZ + 40);
destination->length = VARHDRSZ + 40;
memmove(destination->data, buffer, 40);
...
```

既然我们已经讲完了基本类型的所有可能结构，我们可以展示一些真实函数的示例。设 funcs.c 内容如下：

```
#include <string.h>
#include "postgres.h"
int
add_one(int arg)
{
    return(arg + 1);
}
text *
concat_text(text *arg1, text *arg2)
{
    int32 new_text_size = VARSIZE(arg1) + VARSIZE(arg2) -
VARHDRSZ;
    text *new_text = (text *) palloc(new_text_size);

    memset((void *) new_text, 0, new_text_size);
    VARSIZE(new_text) = new_text_size;
    strncpy(VARDATA(new_text), VARDATA(arg1), VARSIZE(arg1)-
VARHDRSZ);
    strncat(VARDATA(new_text), VARDATA(arg2), VARSIZE(arg2)-
VARHDRSZ);
    return (new_text);
}
text *
copytext(text *t)
```

```
{
    /*
     * VARSIZE is the total size of the struct in bytes.
     */
    text *new_t = (text *) palloc(VARSIZE(t));
    memset(new_t, 0, VARSIZE(t));
    VARSIZE(new_t) = VARSIZE(t);
    /*
     * VARDATA is a pointer to the data region of the struct.
     */
    memcpy((void *) VARDATA(new_t), /* destination */
           (void *) VARDATA(t),    /* source */
           VARSIZE(t)-VARHDRSZ);   /* how many bytes */
    return(new_t);
}
```

在 OSF/1 上我们会输入:

```
CREATE FUNCTION add_one(int4) RETURNS int4
AS 'PGROOT/tutorial/obj/funcs.so' LANGUAGE 'c';

CREATE FUNCTION concat_text(text, text) RETURNS text
AS 'PGROOT/tutorial/obj/funcs.so' LANGUAGE 'c';

CREATE FUNCTION copytext(text) RETURNS text
AS 'PGROOT/tutorial/obj/funcs.so' LANGUAGE 'c';
```

在其他系统上, 我们可能必须让文件名以 .sl 结尾 (表示它是共享库)。

## 39.2.2. 复合类型上的编程语言函数

复合类型没有像 C 结构那样的固定布局。复合类型的实例可能包含空字段。此外, 属于某个继承层次的复合类型可能具有与同一继承层次其他成员不同的字段。因此, Postgres 提供了一个从 C 访问复合类型字段的过程式接口。当 Postgres 处理一组实例时, 每个实例都会作为一个 TUPLE 类型的不透明结构传入你的函数。假设我们想编写一个函数来回答查询

```
* SELECT name, c_overpaid(EMP, 1500) AS overpaid
FROM EMP
WHERE name = 'Bill' or name = 'Sam';
```

在上面的查询中, 我们可以把 c\_overpaid 定义为:

```
#include "postgres.h"
#include "libpq-fe.h" /* for TUPLE */
bool
c_overpaid(TUPLE t, /* the current instance of EMP */
           int4 limit)
{
    bool isnull = false;
    int4 salary;
    salary = (int4) GetAttributeByName(t, "salary", &isnull);
    if (isnull)
        return (false);
    return(salary > limit);
}
```

`GetAttributeByName` 是 Postgres 的系统函数，返回当前实例中的属性。它有三个参数：传给函数的 `TUPLE` 类型参数、所需属性的名称，以及一个描述该属性是否为空的返回参数。`GetAttributeByName` 会正确地对齐数据，因此你可以把它的返回值转换为期望的类型。例如，如果你有一个 `name` 类型的属性 `name`，`GetAttributeByName` 调用将形如：

```
char *str;
...
str = (char *) GetAttributeByName(t, "name", &isnull)
```

下面的查询让 Postgres 知道 `c_overpaid` 函数：

```
* CREATE FUNCTION c_overpaid(EMP, int4) RETURNS bool
  AS 'PGROOT/tutorial/obj/funcs.so' LANGUAGE 'c';
```

虽然有办法在 C 函数内构造新实例或修改现有实例，但这些方法过于复杂，本手册不予讨论。

### 39.2.3. 注意事项

现在我们转向编写编程语言函数这一更困难的任务。

请注意：本手册的这一部分不会使你成为程序员。在尝试为 Postgres 编写 C 函数之前，你必须对 C（包括指针和 `malloc` 内存管理器的使用）有良好的理解。虽然也许可以把用 C 以外语言编写的函数装载到 Postgres，但这通常很困难（即使在可能的時候），因为其他语言（如 FORTRAN 和 Pascal）往往不遵循与 C 相同的“调用约定”。也就是说，其他语言在函数之间传递参数和返回值的方式不同。因此，我们将假定你的编程语言函数是用 C 编写的。构建 C 函数的基本规则如下：

- Postgres 的大多数头（include）文件应当已安装在 `PGROOT/include`（见图 2）。你应当总是在 `cc` 命令行中包含

```
-I$PGROOT/include
```

。有时你可能发现需要服务器源码本身中的头文件（即需要一个我们疏忽没有安装到 `include` 中的文件）。那些情况下，你可能需要添加一个或多个

```
-I$PGROOT/src/backend
-I$PGROOT/src/backend/include
-I$PGROOT/src/backend/port/<PORTNAME>
-I$PGROOT/src/backend/obj
```

（其中 `<PORTNAME>` 是移植的名称，例如 `alpha` 或 `sparc`）。

- 分配内存时，使用 Postgres 的例程 `palloc` 和 `pfree`，而不是相应的 C 库例程 `malloc` 和 `free`。`palloc` 分配的内存会在每个事务结束时自动释放，从而防止内存泄漏。
- 总是使用 `memset` 或 `bzero` 把结构体的字节清零。有几个例程（如哈希访问方法、哈希连接和排序算法）会计算结构体中原始位的函数。即使你初始化了结构体的所有字段，结构体中仍可能存在若干包含垃圾值的对齐填充字节（结构体中的空洞）。
- Postgres 的大多数内部类型都在 `postgres.h` 中声明，因此始终也包含该文件是个好主意。包含 `postgres.h` 时还会顺带为你包含 `elog.h` 和 `palloc.h`。
- 编译和装载你的目标代码使其能被动态装载到 Postgres 总是需要特殊的标志。关于如何针对你的特定操作系统进行操作，参见附录 A 的详细说明。

---

# 第 40 章 扩展 SQL：类型

如前所述，Postgres中有两类类型：基本类型（用编程语言定义）和组合类型（实例）。本节中直到索引接口之前的示例可以在 `complex.sql`和`complex.c`中找到。组合类型的示例在`funcs.sql`中。

## 40.1. 用户定义的类型

### 40.1.1. 用户定义类型所需的函数

用户定义的类型必须始终具有输入和输出函数。这些函数决定该类型在字符串中如何表示（供用户输入和输出给用户）以及在内存中如何组织。输入函数接受一个以空字符结尾的字符串作为输入，并返回该类型的内部（内存中）表示。输出函数接受该类型的内部表示，并返回一个以空字符结尾的字符串。假设我们想定义一个表示复数的 `complex` 类型。很自然，我们会选择用下面的C结构在内存中表示复数：

```
typedef struct Complex {
    double      x;
    double      y;
} Complex;
```

并选择形如 (x,y) 的字符串作为外部字符串表示。这些函数通常不难编写，输出函数尤其如此。不过，有几点需要记住：

- 在定义你的外部（字符串）表示时，请记住：你最终必须为该表示编写一个完整而健壮的解析器作为输入函数！

```
Complex *
complex_in(char *str)
{
    double x, y;
    Complex *result;
    if (sscanf(str, " ( %lf , %lf )", &x, &y) != 2)
    {
        elog(WARN, "complex_in: error in parsing");
        return NULL;
    }
    result = (Complex *)palloc(sizeof(Complex));
    result->x = x;
    result->y = y;
    return (result);
}
```

输出函数可以简单地写成：

```
char *
complex_out(Complex *complex)
{
    char *result;
    if (complex == NULL)
        return(NULL);
```

```

        result = (char *) palloc(60);
        sprintf(result, "(%g,%g)", complex->x, complex->y);
    }
    return(result);
}

```

- 你应当尽量让输入和输出函数互为逆函数。否则，在你需要把数据转储到文件中再读回来时（比如读到另一台计算机上某人的数据库中），就会出现严重的问题。当涉及浮点数时，这是一个特别常见的问题。

要定义complex类型，我们需要在创建类型之前先创建 complex\_in 和 complex\_out 这两个用户定义的函数：

```

CREATE FUNCTION complex_in(opaque)
    RETURNS complex
    AS 'PGROOT/tutorial/obj/complex.so'
    LANGUAGE 'c';

CREATE FUNCTION complex_out(opaque)
    RETURNS opaque
    AS 'PGROOT/tutorial/obj/complex.so'
    LANGUAGE 'c';

CREATE TYPE complex (
    internallength = 16,
    input = complex_in,
    output = complex_out
);

```

如前面所讨论的，Postgres完全支持基本类型的数组。此外，Postgres也支持用户定义类型的数组。当你定义一个类型时，Postgres会自动提供对该类型数组的支持。由于历史原因，数组类型的名称与用户定义类型相同，只是在前面加上下划线字符\_。组合类型不需要在其上定义任何函数，因为系统已经知道它们内部是什么样子。

## 40.1.2. 大对象

到此为止讨论的类型都是"小"对象——也就是说，它们的大小小于 8KB。

### 注意

1024 长字 == 8192 字节。实际上，类型必须显著小于 8192 字节，因为Postgres的元组和页开销也必须放进这个 8KB 的限制内。实际能放下的值取决于机器体系结构。

如果你需要一个更大的类型，比如用于文档检索系统或存储位图，你将需要使用Postgres的大对象接口。

---

## 第 41 章 扩展 SQL：操作符

Postgres 支持左一元、右一元和二元操作符。操作符可以重载，即可用不同的参数数量和类型重复使用。如果出现歧义情况而系统无法确定该用哪个操作符，它会返回一个错误，你可能需要对左操作数和/或右操作数进行类型转换，帮助系统理解你想用哪个操作符。要创建一个把两个复数相加的操作符，可以按下面的方法做。首先我们需要创建一个对新类型做加法的函数，然后就可以用该函数创建操作符。

```
CREATE FUNCTION complex_add(complex, complex)
    RETURNS complex
    AS '$PWD/obj/complex.so'
    LANGUAGE 'c';

CREATE OPERATOR + (
    leftarg = complex,
    rightarg = complex,
    procedure = complex_add,
    commutator = +
);
```

我们在这里展示了如何创建二元操作符。要创建一元操作符，只需省略 leftarg（左一元）或 rightarg（右一元）之一即可。只要我们给系统足够的类型信息，它就能自动确定该用哪些操作符。

```
SELECT (a + b) AS c FROM test_complex;
```

```
+-----+
| c      |
+-----+
| (5.2,6.05) |
+-----+
| (133.42,144.95) |
+-----+
```

---

## 第 42 章 扩展 SQL: 聚合

Postgres 中的聚合是按照状态转换函数来表示的。也就是说，聚合可以按照每当处理一个实例就被修改的状态来定义。有些状态函数在计算新状态时会查看实例中的特定值（即 `create aggregate` 语法中的 `sfunc1`），而另一些只跟踪自己的内部状态（`sfunc2`）。如果我们定义一个只使用 `sfunc1` 的聚合，定义的就是一个对每个实例的属性值计算运行函数的聚合。"Sum" 就是这类聚合的一个例子。"Sum" 从零开始，总是把当前实例的值加到它的运行总和上。我们将使用内置在 Postgres 中的 `int4pl` 来执行这个加法。

```
CREATE AGGREGATE complex_sum (  
    sfunc1 = complex_add,  
    basetype = complex,  
    stype1 = complex,  
    initcond1 = '(0,0)'  
);  
  
SELECT complex_sum(a) FROM test_complex;
```

```
+-----+  
|complex_sum |  
+-----+  
| (34,53.9)  |  
+-----+
```

如果我们只定义 `sfunc2`，指定的是一个对每个实例的属性值无关的运行函数计算的聚合。"Count" 是这类聚合最常见的例子。"Count" 从零开始，对每个实例向其运行总数加一，忽略实例的值。这里我们使用内置的 `int4inc` 例程为我们完成这项工作。该例程把输入加一。

```
CREATE AGGREGATE my_count (  
    sfunc2 = int4inc, -- add one  
    basetype = int4,  
    stype2 = int4,  
    initcond2 = '0'  
);  
  
SELECT my_count(*) as emp_count from EMP;
```

```
+-----+  
|emp_count |  
+-----+  
| 5         |  
+-----+
```

"Average"（平均值）是一个既需要计算运行总和的函数又需要计算运行计数的函数的聚合例子。处理完所有实例后，聚合的最终答案是运行总和除以运行计数。我们使用前面用过的 `int4pl` 和 `int4inc` 例程，以及 Postgres 的整数除法例程 `int4div` 来计算总和除以计数。

```
CREATE AGGREGATE my_average (  
    sfunc1 = int4pl,      -- sum  
    basetype = int4,  
    stype1 = int4,  
    sfunc2 = int4inc,    -- count  
    stype2 = int4,  
    finalfunc = int4div, -- division
```

```
    initcond1 = '0',  
    initcond2 = '0'  
);
```

```
SELECT my_average(salary) as emp_average FROM EMP;
```

```
+-----+  
|emp_average |  
+-----+  
|1640        |  
+-----+
```

---

## 第 43 章 Postgres 规则系统

产生式规则系统在概念上很简单，但在实际使用时会涉及许多微妙之处。其中一些要点，以及 Postgres 规则系统的理论基础，可以在 [Stonebraker et al, ACM, 1990] 中找到。

有些其他数据库系统定义了主动型数据库规则。它们通常是存储过程和触发器，在 Postgres 中以函数和触发器的形式实现。

查询重写规则系统（下文简称“规则系统”）与存储过程和触发器完全不同。它会修改查询，使其将规则纳入考虑，然后把修改后的查询交给查询优化器执行。它非常强大，可用于查询语言过程、视图和版本等许多用途。这一规则系统的能力在 [Ong and Goh, 1990] 以及 [Stonebraker et al, ACM, 1990] 中有讨论。

### 43.1. 什么是查询树？

要理解规则系统如何工作，必须先知道它在何时被调用，以及它的输入和输出是什么。

规则系统位于查询解析器和优化器之间。它接收解析器的输出，即一棵查询树，以及来自 `pg_rewrite` 目录的重写规则；这些规则本身也是带有一些附加信息的查询树。它会生成零棵或多棵查询树作为结果。因此，它的输入和输出始终都是解析器本身也可能产生的东西，所以它看到的任何内容基本上都可以表示为一个 SQL 语句。

那么什么是查询树？它是 SQL 语句的一种内部表示，其中构成语句的各个部分被分别存储。如果以调试级别 4 启动 Postgres 后端，并在交互式后端界面中输入查询，就可以看到这些查询树。`pg_rewrite` 系统目录中的规则动作同样以查询树的形式存储。它们的格式与调试输出不同，但包含的却是完全相同的信息。

阅读查询树需要一些经验，我刚开始在规则系统上工作时也经历了一段艰难的时光。我记得我曾站在咖啡机前，看到目标列表里有一个杯子，`rangetable` 里有水和咖啡粉，而所有的按钮出现在一个条件表达式中。由于查询树的 SQL 表示已足以理解规则系统，本文不会讲解如何阅读它们。学习阅读查询树也许会有帮助，而且在后续的描述中需要用到这些命名约定。

#### 43.1.1. 查询树的组成部分

在阅读本文中查询树的 SQL 表示时，需要能够辨认语句在查询树结构中被拆分成了哪些部分。查询树的组成部分是

命令类型

这是一个简单的值，说明了是哪一种命令（SELECT、INSERT、UPDATE、DELETE）产生了解析树。

范围表

范围表是查询中使用的关系列表。在 SELECT 语句中，它们就是 FROM 关键字后面给出的关系。

每个范围表项标识一个表或视图，并说明在查询的其他部分以哪个名称来称呼它。在查询树中，范围表项是按编号而不是按名称引用的，因此即使像 SQL 语句中那样出现重名，这里也无关紧要。这种情况可能出现在规则的范围表被合并之后。本文中的示例不会涉及这种情形。

## 结果关系

这是范围表中的一个索引，用来标识查询结果应写入哪个关系。

SELECT 查询通常没有结果关系。SELECT INTO 这一特殊情况与 CREATE TABLE、INSERT ... SELECT 序列几乎相同，这里不再单独讨论。

在 INSERT、UPDATE 和 DELETE 查询中，结果关系就是要让修改生效的表（或视图！）。

## 目标列表

目标列表是定义查询结果的表达式列表。对于 SELECT，这些表达式构成查询的最终输出。它们对应于关键字 SELECT 和 FROM 之间的表达式。（\* 只是某个关系全部属性名的缩写。）

DELETE 查询不需要目标列表，因为它们不产生任何结果。实际上优化器会向空目标列表加入一个特殊项，但这发生在规则系统之后，将在后面讨论。对规则系统而言，目标列表是空的。

在 INSERT 查询中，目标列表描述的是将要进入结果关系的新行。结果关系中缺失的列将由优化器补上常量空值表达式。它由 VALUES 子句中的表达式，或 INSERT ... SELECT 里 SELECT 子句中的表达式组成。

在 UPDATE 查询中，目标列表描述要替换旧行的新行。这里，优化器会通过插入把旧行的值放到新行中的表达式来补上缺失列，并且也会像 DELETE 那样加上那个特殊项。它由查询中 SET attribute = expression 部分的表达式组成。

目标列表中的每一项都包含一个表达式，它可以是常量值、指向范围表中某个关系属性的变量、一个参数，或者由函数调用、常量、变量、操作符等构成的表达式树。

## 条件

查询的条件是一个表达式，与目标列表项中的表达式很相似。该表达式的结果值是布尔值，用来说明是否应对最终结果行执行相应操作（INSERT、UPDATE、DELETE 或 SELECT）。它就是 SQL 语句的 WHERE 子句。

## 其他

查询树的其他部分，如 ORDER BY 子句，这里不作关注。规则系统在应用规则时会替换其中的项，但这与规则系统的基本原理关系不大。GROUP BY 在出现在视图定义中时是一种特殊情况，仍有待文档说明。

## 43.2. 视图和规则系统

### 43.2.1. Postgres 中视图的实现

在 Postgres 中，视图通过规则系统实现。实际上，以下命令：

```
CREATE VIEW myview AS SELECT * FROM mytab;
```

与下面这两条命令基本没有区别：

```
CREATE TABLE myview (same attribute list as for mytab);
CREATE RULE "_RETmyview" AS ON SELECT TO myview DO INSTEAD
    SELECT * FROM mytab;
```

因为这正是 `CREATE VIEW` 命令在内部所做的事情。这会带来一些副作用。其中之一是，在 Postgres 系统目录中，视图的信息与表的信息完全相同。因此，对查询解析器而言，表和视图完全没有区别。它们是同一种东西——关系。这是目前最重要的一点。

## 43.2.2. SELECT 规则如何工作

ON SELECT 规则会在所有查询上作为最后一步应用，即使给出的命令是 `INSERT`、`UPDATE` 或 `DELETE` 也一样。它们与其他规则在语义上不同，因为它们是就地修改解析树，而不是创建新的解析树。因此我们先讨论 SELECT 规则。

目前，只能有一个动作，而且它必须是一个带有 `INSTEAD` 的 `SELECT` 动作。之所以有这个限制，是为了让规则足够安全，从而能够向普通用户开放；它也把 ON SELECT 规则限制为真正的视图规则。

本文的示例是两个进行一些计算的连接视图，以及另外一些依次使用它们的视图。最初两个视图中的一个会在后面通过为 `INSERT`、`UPDATE` 和 `DELETE` 操作添加规则来定制，从而最终得到一个在行为上像真正的表、但又带有某些特殊功能的视图。作为入门示例，这并不算简单，因此会让理解变得更难一些。但与其使用许多可能令人混淆的不同示例，不如用一个示例逐步覆盖这里讨论的全部要点。

用于在这些示例上操作的数据库名为 `al_bundy`。你很快就会明白为什么取这个名字。它还需要安装过程语言 `PL/pgSQL`，因为我们需要一个返回两个整数值中较小者的 `min()` 函数。我们这样创建它：

```
CREATE FUNCTION min(integer, integer) RETURNS integer AS 'BEGIN
    IF $1 < $2 THEN
        RETURN $1;
    END IF;
    RETURN $2;
END;' LANGUAGE 'plpgsql';
```

在前两节对规则系统的描述（descriptons）中，我们需要用到如下真实表：

```
CREATE TABLE shoe_data (
    shoename    char(10),      -- primary key
    sh_avail    integer,       -- available # of pairs
    slcolor     char(10),      -- preferred shoelace color
    slminlen    float,         -- minimum shoelace length
    slmaxlen    float,         -- maximum shoelace length
    slunit      char(8)        -- length unit
);

CREATE TABLE shoelace_data (
    sl_name     char(10),      -- primary key
    sl_avail    integer,       -- available # of pairs
    sl_color    char(10),      -- shoelace color
    sl_len      float,         -- shoelace length
    sl_unit     char(8)        -- length unit
);

CREATE TABLE unit (
    un_name     char(8),        -- the primary key
    un_fact     float          -- factor to transform to cm
);
```

我想我们大多数人都穿鞋，都能看出这确实是很用的数据。当然，世上有一些不需要鞋带的鞋，但这并不能让 AI 的日子好过一点，所以我们忽略这一点。

这些视图的创建方式如下：

```
CREATE VIEW shoe AS
    SELECT sh.shoename,
           sh.sh_avail,
           sh.slcolor,
           sh.slminlen,
           sh.slminlen * un.un_fact AS slminlen_cm,
           sh.slmaxlen,
           sh.slmaxlen * un.un_fact AS slmaxlen_cm,
           sh.slunit
    FROM shoe_data sh, unit un
    WHERE sh.slunit = un.un_name;

CREATE VIEW shoelace AS
    SELECT s.sl_name,
           s.sl_avail,
           s.sl_color,
           s.sl_len,
           s.sl_unit,
           s.sl_len * u.un_fact AS sl_len_cm
    FROM shoelace_data s, unit u
    WHERE s.sl_unit = u.un_name;

CREATE VIEW shoe_ready AS
    SELECT rsh.shoename,
           rsh.sh_avail,
           rsl.sl_name,
           rsl.sl_avail,
           min(rsh.sh_avail, rsl.sl_avail) AS total_avail
    FROM shoe rsh, shoelace rsl
    WHERE rsl.sl_color = rsh.slcolor
           AND rsl.sl_len_cm >= rsh.slminlen_cm
           AND rsl.sl_len_cm <= rsh.slmaxlen_cm;
```

创建 `shoelace` 视图（这是我们这里最简单的一个例子）的 `CREATE VIEW` 命令会创建一个关系 `shoelace`，并在 `pg_rewrite` 中创建一项，说明只要查询的 `rangetable` 中引用了关系 `shoelace`，就必须应用一条重写规则。该规则没有规则条件（这在非 `SELECT` 规则中讨论，因为 `SELECT` 规则目前不能有规则条件），并且它是 `INSTEAD`。注意，规则条件与查询条件不是一回事！这里我们的规则动作带有一个查询条件。

规则动作本身是一棵查询树，它是视图创建命令中 `SELECT` 语句的一个精确副本。

### 注意

你在 `pg_rewrite` 项中看到的、用于 `NEW` 和 `OLD` 的两个额外范围表项（在打印出来的查询树中，它们出于历史原因被命名为 `*NEW*` 和 `*CURRENT*`），与 `SELECT` 规则无关。

现在我们填充 `unit`、`shoe_data` 和 `shoelace_data`，然后 AI 在他的一生中第一次输入 `SELECT`：

```

al_bundy=> INSERT INTO unit VALUES ('cm', 1.0);
al_bundy=> INSERT INTO unit VALUES ('m', 100.0);
al_bundy=> INSERT INTO unit VALUES ('inch', 2.54);
al_bundy=>
al_bundy=> INSERT INTO shoe_data VALUES
al_bundy->      ('sh1', 2, 'black', 70.0, 90.0, 'cm');
al_bundy=> INSERT INTO shoe_data VALUES
al_bundy->      ('sh2', 0, 'black', 30.0, 40.0, 'inch');
al_bundy=> INSERT INTO shoe_data VALUES
al_bundy->      ('sh3', 4, 'brown', 50.0, 65.0, 'cm');
al_bundy=> INSERT INTO shoe_data VALUES
al_bundy->      ('sh4', 3, 'brown', 40.0, 50.0, 'inch');
al_bundy=>
al_bundy=> INSERT INTO shoelace_data VALUES
al_bundy->      ('sl1', 5, 'black', 80.0, 'cm');
al_bundy=> INSERT INTO shoelace_data VALUES
al_bundy->      ('sl2', 6, 'black', 100.0, 'cm');
al_bundy=> INSERT INTO shoelace_data VALUES
al_bundy->      ('sl3', 0, 'black', 35.0 , 'inch');
al_bundy=> INSERT INTO shoelace_data VALUES
al_bundy->      ('sl4', 8, 'black', 40.0 , 'inch');
al_bundy=> INSERT INTO shoelace_data VALUES
al_bundy->      ('sl5', 4, 'brown', 1.0 , 'm');
al_bundy=> INSERT INTO shoelace_data VALUES
al_bundy->      ('sl6', 0, 'brown', 0.9 , 'm');
al_bundy=> INSERT INTO shoelace_data VALUES
al_bundy->      ('sl7', 7, 'brown', 60 , 'cm');
al_bundy=> INSERT INTO shoelace_data VALUES
al_bundy->      ('sl8', 1, 'brown', 40 , 'inch');
al_bundy=>
al_bundy=> SELECT * FROM shoelace;
sl_name  |sl_avail|sl_color  |sl_len|sl_unit |sl_len_cm
-----+-----+-----+-----+-----+-----
sl1      |      5|black     |  80|cm      |      80
sl2      |      6|black     | 100|cm      |     100
sl7      |      7|brown     |  60|cm      |      60
sl3      |      0|black     |  35|inch    |     88.9
sl4      |      8|black     |  40|inch    |    101.6
sl8      |      1|brown     |  40|inch    |    101.6
sl5      |      4|brown     |   1|m       |     100
sl6      |      0|brown     | 0.9|m       |      90
(8 rows)

```

这是 `al` 在这些视图上能执行的最简单的 `SELECT`，因此我们借此机会说明视图规则的基础。`'SELECT * FROM shoelace'` 由解析器解释后，会生成如下 `parsetree`：

```

SELECT shoelace.sl_name, shoelace.sl_avail,
       shoelace.sl_color, shoelace.sl_len,
       shoelace.sl_unit, shoelace.sl_len_cm
FROM shoelace shoelace;

```

然后它会被交给规则系统。规则系统遍历 `rangetable`，检查 `pg_rewrite` 中是否有关系对应的规则。在处理 `shoelace` 的 `rangetable` 项时（到目前为止只有这一个），它会找到带有如下 `parsetree` 的规则 `'_RETshoelace'`：

```
SELECT s.sl_name, s.sl_avail,
       s.sl_color, s.sl_len, s.sl_unit,
       float8mul(s.sl_len, u.un_fact) AS sl_len_cm
FROM shoelace *OLD*, shoelace *NEW*,
     shoelace_data s, unit u
WHERE bpchareq(s.sl_unit, u.un_name);
```

注意，解析器把计算和限定条件都改写成了对相应函数的调用。但这实际上并没有改变任何东西。重写的第一步是合并两个 rangetable。得到的 parsetree 如下：

```
SELECT shoelace.sl_name, shoelace.sl_avail,
       shoelace.sl_color, shoelace.sl_len,
       shoelace.sl_unit, shoelace.sl_len_cm
FROM shoelace shoelace, shoelace *OLD*,
     shoelace *NEW*, shoelace_data s,
     unit u;
```

第 2 步把规则动作中的限定条件加到该 parsetree 上，得到：

```
SELECT shoelace.sl_name, shoelace.sl_avail,
       shoelace.sl_color, shoelace.sl_len,
       shoelace.sl_unit, shoelace.sl_len_cm
FROM shoelace shoelace, shoelace *OLD*,
     shoelace *NEW*, shoelace_data s,
     unit u
WHERE bpchareq(s.sl_unit, u.un_name);
```

第 3 步把 parsetree 中所有引用当前正在处理的 rangetable 项（即 shoelace 的那一项）的变量，替换为规则动作 targetlist 中的相应表达式。最终得到的查询是：

```
SELECT s.sl_name, s.sl_avail,
       s.sl_color, s.sl_len,
       s.sl_unit, float8mul(s.sl_len, u.un_fact) AS sl_len_cm
FROM shoelace shoelace, shoelace *OLD*,
     shoelace *NEW*, shoelace_data s,
     unit u
WHERE bpchareq(s.sl_unit, u.un_name);
```

把它转换回人类用户会输入的真实 SQL 语句，读作：

```
SELECT s.sl_name, s.sl_avail,
       s.sl_color, s.sl_len,
       s.sl_unit, s.sl_len * u.un_fact AS sl_len_cm
FROM shoelace_data s, unit u
WHERE s.sl_unit = u.un_name;
```

这就是应用的第一条规则。在此过程中，rangetable 已经变大了。于是规则系统继续检查各个 rangetable 项。下一个是第 2 号 (shoelace \*OLD\*)。关系 shoelace 有一条规则，但这个 rangetable 项没有被 parsetree 中的任何变量引用，因此被忽略。由于其余的 rangetable 项要么在 pg\_rewrite 中没有规则、要么未被引用，检查到达 rangetable 末尾。重写至此完成，上面的结果就是交给优化器的最终结果。优化器会忽略那些未被 parsetree 中的变量引用的额外 rangetable 项，优化器/优化器产生的计划，将与 AI 直接输入上面的 SELECT 查询（而不是对视图的选择）完全相同。

现在我们让 Al 面对这样一个问题：布鲁斯兄弟（Blues Brothers）来到他的店里，想买一些新鞋；而且正如布鲁斯兄弟的本色那样，他们要穿一模一样的鞋。他们还想立刻就穿上，所以还需要鞋带。

Al 需要知道，商店里目前哪些鞋子有颜色和尺码都匹配的鞋带，并且完全匹配的总双数大于等于二。我们教会（theach）他怎么做，于是他询问他的数据库：

```
al_bundy=> SELECT * FROM shoe_ready WHERE total_avail >= 2;
shoename  |sh_avail|sl_name    |sl_avail|total_avail
-----+-----+-----+-----+-----
sh1        |        2|sl1        |        5|        2
sh3        |        4|sl7        |        7|        4
(2 rows)
```

Al 是一位鞋类行家，所以他知道只有 sh1 型号的鞋才合适（鞋带 sl7 是棕色的，而需要棕色鞋带的鞋是布鲁斯兄弟绝不会穿的鞋）。

这一次解析器的输出是如下 parsetree：

```
SELECT shoe_ready.shoename, shoe_ready.sh_avail,
       shoe_ready.sl_name, shoe_ready.sl_avail,
       shoe_ready.total_avail
FROM shoe_ready shoe_ready
WHERE int4ge(shoe_ready.total_avail, 2);
```

首先应用的是用于 shoe\_ready 关系的规则，它得到如下 parsetree：

```
SELECT rsh.shoename, rsh.sh_avail,
       rsl.sl_name, rsl.sl_avail,
       min(rsh.sh_avail, rsl.sl_avail) AS total_avail
FROM shoe_ready shoe_ready, shoe_ready *OLD*,
     shoe_ready *NEW*, shoe rsh,
     shoelace rsl
WHERE int4ge(min(rsh.sh_avail, rsl.sl_avail), 2)
      AND (bpchareq(rsl.sl_color, rsh.slcolor)
           AND float8ge(rsl.sl_len_cm, rsh.slminlen_cm)
           AND float8le(rsl.sl_len_cm, rsh.slmaxlen_cm)
      );
```

实际上，限定条件中的 AND 子句会是 AND 类型的操作符节点，各带一个左表达式和一个右表达式。但那样只会让本已不易读的东西更难读，而且还有更多规则要应用。所以我只是把它们放进一些括号里，按添加的顺序把它们分组成逻辑单元，然后继续处理关系 shoe 的规则，因为它是下一个被引用且带有规则的 rangetable 项。应用该规则的结果是：

```
SELECT sh.shoename, sh.sh_avail,
       rsl.sl_name, rsl.sl_avail,
       min(sh.sh_avail, rsl.sl_avail) AS total_avail,
FROM shoe_ready shoe_ready, shoe_ready *OLD*,
     shoe_ready *NEW*, shoe rsh,
     shoelace rsl, shoe *OLD*,
     shoe *NEW*, shoe_data sh,
     unit un
WHERE (int4ge(min(sh.sh_avail, rsl.sl_avail), 2)
      AND (bpchareq(rsl.sl_color, sh.slcolor)
           AND float8ge(rsl.sl_len_cm,
```

```

float8mul(sh.slminlen, un.un_fact))
AND float8le(rsl.sl_len_cm,
float8mul(sh.slmaxlen, un.un_fact))
)
)
AND bpchareq(sh.slunit, un.un_name);

```

最后我们应用那条已经十分熟悉的 shoelace 规则（这一次是在一个稍微复杂一点的 parsetree 上），得到：

```

SELECT sh.shoename, sh.sh_avail,
s.sl_name, s.sl_avail,
min(sh.sh_avail, s.sl_avail) AS total_avail
FROM shoe_ready shoe_ready, shoe_ready *OLD*,
shoe_ready *NEW*, shoe rsh,
shoelace rsl, shoe *OLD*,
shoe *NEW*, shoe_data sh,
unit un, shoelace *OLD*,
shoelace *NEW*, shoelace_data s,
unit u
WHERE ( (int4ge(min(sh.sh_avail, s.sl_avail), 2)
AND (bpchareq(s.sl_color, sh.slcolor)
AND float8ge(float8mul(s.sl_len, u.un_fact),
float8mul(sh.slminlen, un.un_fact))
AND float8le(float8mul(s.sl_len, u.un_fact),
float8mul(sh.slmaxlen, un.un_fact))
)
)
AND bpchareq(sh.slunit, un.un_name)
)
AND bpchareq(s.sl_unit, u.un_name);

```

我们再把它化简为与规则系统最终输出等价的真实 SQL 语句：

```

SELECT sh.shoename, sh.sh_avail,
s.sl_name, s.sl_avail,
min(sh.sh_avail, s.sl_avail) AS total_avail
FROM shoe_data sh, shoelace_data s, unit u, unit un
WHERE min(sh.sh_avail, s.sl_avail) >= 2
AND s.sl_color = sh.slcolor
AND s.sl_len * u.un_fact >= sh.slminlen * un.un_fact
AND s.sl_len * u.un_fact <= sh.slmaxlen * un.un_fact
AND sh.sl_unit = un.un_name
AND s.sl_unit = u.un_name;

```

规则的递归处理把一个对视图的 SELECT 重写成了一个 parsetree，它恰好等价于在没有视图的情况下 AI 必须输入的内容。

## 注意

规则系统目前对视图规则没有递归终止机制（对其他规则才有）。这并不会造成太大问题，因为要把处理推入无限循环（使后端不断膨胀，直至达到内存上限）的唯一方式，是先创建一些表，然后手工用 CREATE RULE 把视图规则设置成一个从另一个选择、而另一个又从这个选择的形式。如果使用 CREATE VIEW，这种

情况绝不会发生，因为在执行第一个 CREATE VIEW 时，第二个关系尚不存在，因此第一个视图无法从第二个视图选择。

### 43.2.3. 非 SELECT 语句中的视图规则

上面关于视图规则的描述没有触及 parsetree 的两个细节：commandtype 和 resultrelation。事实上，视图规则并不需要这些信息（informations）。

SELECT 的 parsetree 与其他任何命令的 parsetree 之间只有少数差别。显然，它们的 commandtype 不同，而且这一次 resultrelation 会指向结果应写入的那个 rangetable 项。除此之外，其余部分完全相同。因此，假设有两个表 t1 和 t2，都具有属性 a 和 b，那么下面两条语句的解析树：

```
SELECT t2.b FROM t1, t2 WHERE t1.a = t2.a;
```

```
UPDATE t1 SET b = t2.b WHERE t1.a = t2.a;
```

几乎是一样的。

- 范围表中包含表 t1 和 t2 的项。
- 目标列表中都包含一个变量，该变量指向表 t2 的范围表项中的属性 b。
- 条件表达式会比较两个范围中的属性 a 是否相等。

结果是，这两个 parsetree 都会产生相似的执行计划。它们都是对这两个表的连接。对于 UPDATE，优化器会把 t1 中缺失的列补入 targetlist，最终的 parsetree 会读作：

```
UPDATE t1 SET a = t1.a, b = t2.b WHERE t1.a = t2.a;
```

因而，执行器在这个连接上运行时会产生与下面语句完全相同的结果集：

```
SELECT t1.a, t2.b FROM t1, t2 WHERE t1.a = t2.a;
```

但在 UPDATE 中有个小问题。执行器并不关心它所做的连接的结果将被用于什么。它只是生成一个行结果集。一个是 SELECT 命令、另一个是 UPDATE 命令这一差别，是在执行器的调用者中处理的。调用者仍然知道（通过查看解析树）这是一个 UPDATE，也知道这个结果应写入表 t1。但那 666 行里的哪一行应当被新行替换？所执行的计划是一个带限定条件的连接，可能以未知顺序产生 0 到 666 行之间的任意数量的行。

为了解决这个问题，UPDATE 和 DELETE 语句的 targetlist 中会额外加入一项：当前元组 ID（ctid）。这是一个带特殊功能的系统属性。它包含该行所在的块号以及在块中的位置。已知表之后，利用 ctid 只需取出一个数据块，就可以在一个包含数百万行、大小为 1.5GB 的表中找到特定的那一行。把 ctid 加入 targetlist 后，最终结果集可以定义为：

```
SELECT t1.a, t2.b, t1.ctid FROM t1, t2 WHERE t1.a = t2.a;
```

现在还要引入 Postgres 的另一个细节。目前，表行不会被覆盖，这也是 ABORT TRANSACTION 之所以很快的原因。在 UPDATE 中，新结果行会被插入表中（去掉 ctid 之后），而 ctid 所指向的行，其元组头中的 cmax 和 xmax 项会被设置为当前命令计数器和当前事务 ID。于是旧行被隐藏起来，事务提交后，清理器（vacuum）最终就可以真正移除它。

知道了那一切之后，我们就可以用完全相同的方式把视图规则应用到任何命令上。没有区别。

## 43.2.4. Postgres 中视图的能力

上文演示了规则系统如何把视图定义整合进原始 `parsetree`。在第二个示例中，从一个视图发出的简单 `SELECT` 最终生成了一个四表连接的 `parsetree`（`unit` 以不同名称使用了两次）。

### 43.2.4.1. 好处

用规则系统实现视图的好处在于：优化器能够在单个 `parsetree` 中同时看到哪些表必须被扫描、这些表之间的关系、来自视图的限制条件，以及原始查询自身的条件。即使原始查询本身已经是对若干视图的连接，情况也仍然如此。现在，优化器必须决定执行查询的最佳路径。优化器掌握的信息越多，这个决定就能做得越好。Postgres实现的规则系统能够保证，到此为止，关于该查询的全部可用信息都已经集中在这里。

### 43.2.4.2. 隐忧

曾经有很长一段时间，Postgres 的规则系统被认为是坏的。不推荐使用规则，唯一能工作的部分是视图规则。而即使是这些视图规则也会带来问题，因为规则系统无法在 `SELECT` 之外的语句上正确应用它们（例如，一个使用了视图数据的 `UPDATE` 无法工作）。

在那段时间里，开发继续推进，解析器和优化器添加了许多特性。规则系统与它们的能力越来越脱节，着手修复也越来越难。于是，没有人去做。

到了 6.4，有人锁上门、深吸一口气，把这该死的东西彻底翻修了一遍。出来的就是本文所描述的这种能力的规则系统。但仍有一些构造没有得到处理，还有一些因为 Postgres 查询优化器目前不支持而失败。

- 带聚合列的视图有严重的问题。限定条件中的聚合表达式必须放在子查询中使用。目前无法对两个各带一个聚合列的视图做连接，并在限定条件中比较这两个聚合值。眼下可以把这些聚合表达式放进带有适当参数的函数中，再在视图定义里使用它们。
- 目前不支持 `union` 的视图。把一个简单的 `SELECT` 重写成 `union` 很容易。但如果该视图是一个做更新的连接的一部分，就有点困难了。
- 不支持视图定义中的 `ORDER BY` 子句。
- 不支持视图定义中的 `DISTINCT`。

没有好的理由说明，优化器为什么不应该处理那些解析器由于 SQL 语法限制而永远无法产生的 `parsetree` 构造。作者希望这些条目将来会消失。

## 43.2.5. 实现上的副作用

用上面描述的规则系统来实现视图，有一个有趣的副作用。下面这个操作看起来不起作用：

```
al_bundy=> INSERT INTO shoe (shoename, sh_avail, slcolor)
al_bundy->      VALUES ('sh5', 0, 'black');
INSERT 20128 1
al_bundy=> SELECT shoename, sh_avail, slcolor FROM shoe_data;
shoename |sh_avail|slcolor
-----+-----+-----
sh1      |      2|black
sh3      |      4|brown
sh2      |      0|black
sh4      |      3|brown
```

(4 rows)

有趣的是，INSERT 的返回码给了我们一个对象 ID，并告诉我们已经插入了 1 行。但它并没有出现在 shoe\_data 中。查看数据库目录可以发现，视图关系 shoe 的数据库文件现在似乎有了一个数据块。而这确实是真的。

我们还可以执行一条 DELETE，如果它不带限定条件，它会告诉我们已经删除了一些行，而下次 vacuum 运行将把该文件重置为零大小。

这种行为的原因在于，INSERT 的 parsetree 没有在任何变量中引用 shoe 关系。targetlist 只包含常量值。因此没有规则可应用，它原封不动地进入执行，行被插入。DELETE 也是如此。

要改变这种情况，我们可以定义规则来修改非 SELECT 查询的行为。这是下一节的主题。

## 43.3. INSERT、UPDATE 和 DELETE 上的规则

### 43.3.1. 与视图规则的区别

定义在 ON INSERT、UPDATE 和 DELETE 上的规则与前一节描述的视图规则完全不同。首先，它们的 CREATE RULE 命令允许更多：

- 它们可以没有动作。
- 它们可以有多个动作。
- INSTEAD 关键字是可选的。
- 伪关系 NEW 和 OLD 可以派上用场。
- 它们可以有规则条件。

第二，它们不是就地修改解析树，而是创建零棵或多棵新解析树，并且可能丢弃原始查询树。

### 43.3.2. 这些规则如何工作

记住以下语法：

```
CREATE RULE rule_name AS ON event
    TO object [WHERE rule_qualification]
    DO [INSTEAD] [action | (actions) | NOTHING];
```

在后文中，“更新规则”是指定义在 ON INSERT、UPDATE 或 DELETE 上的规则。

当解析树的结果关系和命令类型分别等于 CREATE RULE 命令中给出的对象和事件时，规则系统就会应用更新规则。对于更新规则，规则系统会创建一个解析树列表，初始时该列表为空。动作可以有零个（NOTHING 关键字）、一个或多个。为简化说明，我们来看只有一个动作的规则。这个规则可以有条件，也可以没有条件；它可以是 INSTEAD，也可以不是。

什么是规则条件？它是一种限制，用来说明何时执行规则动作、何时不执行。这个条件只能引用 NEW 和/或 OLD 伪关系，它们基本上就是作为对象给出的那个关系，只是带有特殊含义。

因此，对于只有一个动作的规则，有以下四种情况会产生相应的解析树。

- 没有条件，也不是 INSTEAD：

- 规则动作的解析树，其中附加了原始解析树的条件。
- 没有条件，但带有 INSTEAD：
  - 规则动作的解析树，其中附加了原始解析树的条件。
- 给出了条件，不是 INSTEAD：
  - 规则动作的解析树，其中附加了规则条件和原始解析树的条件。
- 给出了条件，带有 INSTEAD：
  - 规则动作的解析树，其中附加了规则条件和原始解析树的条件。
  - 原始解析树，其中附加了规则条件取反后的条件。

最后，如果规则不是 INSTEAD，就把未经更改的原始解析树加入列表。由于只有带条件的 INSTEAD 规则已经加入了原始解析树，因此，对于只有一个动作的规则，最终最多会得到两棵输出解析树。

由规则动作生成的解析树会再次送入重写系统，随后可能还会有更多规则被应用，从而产生更多或更少的解析树。因此，规则动作中的解析树必须具有不同的命令类型，或者具有不同的结果关系。否则，这种递归过程就会陷入循环。目前内置的递归上限是 10 次迭代。如果经过 10 次迭代之后仍有更新规则要应用，规则系统就会认为出现了多个规则定义之间的循环，并中止事务。

pg\_rewrite 系统目录中动作里的解析树只是模板。由于它们可以引用 NEW 和 OLD 的范围表项，因此在使用前必须进行一些替换。对任何 NEW 的引用，都会先在原始查询的目标列表中查找相应项。如果找到了，就用该项的表达式替换该引用。否则，NEW 就与 OLD 含义相同。任何对 OLD 的引用，都会被替换为对作为 resultrelation 的那个 rangetable 项的引用。

### 43.3.2.1. 第一个规则：逐步分析

我们想跟踪 shoelace\_data 关系中 sl\_avail 列的变化。因此，我们建立一个日志表和一条规则，使其在 shoelace\_data 上执行 UPDATE 时，每次都为我们写一些日志项。

```
CREATE TABLE shoelace_log (
    sl_name      char(10),      -- shoelace changed
    sl_avail     integer,       -- new available value
    log_who      name,         -- who did it
    log_when     datetime       -- when
);

CREATE RULE log_shoelace AS ON UPDATE TO shoelace_data
WHERE NEW.sl_avail != OLD.sl_avail
DO INSERT INTO shoelace_log VALUES (
    NEW.sl_name,
    NEW.sl_avail,
    getpgusername(),
    'now'::text
);
```

一个有趣的细节是规则 INSERT 动作中把 'now' 转换为 text 类型。如果不这样做，解析器在 CREATE RULE 时就会看到 shoelace\_log 中的目标类型是 datetime，并试图用它构造一个常量——而且会成功。于是一个常量 datetime 值会被存储在规则动作中，所有日志项的时间都

将是 CREATE RULE 语句的时间。这并不完全是我们想要的。这个转换使解析器改而构造出一个 datetime('now'::text)，它会在规则执行时才被求值。

现在 AI 执行了：

```
al_bundy=> UPDATE shoelace_data SET sl_avail = 6

al_bundy->      WHERE sl_name = 'sl7';
```

然后我们查看日志表：

```
al_bundy=> SELECT * FROM shoelace_log;
sl_name      |sl_avail|log_who|log_when
-----+-----+-----+-----
sl7          |      6|Al     |Tue Oct 20 16:14:45 1998 MET DST
(1 row)
```

这正是我们预期的结果。后台发生的事情如下。解析器创建了 parsetree（这一次突出显示的是原始 parsetree 中的部分，因为对更新规则而言，操作的基础是规则动作）：

```
UPDATE shoelace_data SET sl_avail = 6
      FROM shoelace_data shoelace_data
      WHERE bpchareq(shoelace_data.sl_name, 'sl7');
```

此时存在一条 ON UPDATE 规则 'log\_shoelace'，它带有如下规则条件表达式：

```
int4ne(NEW.sl_avail, OLD.sl_avail)
```

以及一个动作：

```
INSERT INTO shoelace_log SELECT
      *NEW*.sl_name, *NEW*.sl_avail,
      getpgusername(), datetime('now'::text)
FROM shoelace_data *NEW*, shoelace_data *OLD*,
      shoelace_log shoelace_log;
```

不要相信 pg\_rules 系统视图的输出。它专门处理了 INSERT 中只有对 NEW 和 OLD 的引用的情形，输出 INSERT 的 VALUES 格式。实际上，INSERT ... VALUES 和 INSERT ... SELECT 在 parsetree 层面没有区别。它们都有 rangetable、targetlist，可能还有限定条件等。优化器稍后会决定为该 parsetree 创建哪种类型的执行计划：result、seqscan、indexscan、join 或其他。如果 parsetree 中不再有对 rangetable 项的引用（leftin），它就会变成一个 result 执行计划（即 INSERT ... VALUES 版本）。上面的规则动作确实（truely）两种结果都可能出现。

该规则是一条带条件的非 INSTEAD 规则，因此规则系统必须返回两个 parsetree：修改后的规则动作，以及原始 parsetree。第一步中，原始查询的 rangetable 会并入规则动作的 parsetree，得到：

```
INSERT INTO shoelace_log SELECT
      *NEW*.sl_name, *NEW*.sl_avai,
      getpgusername(), datetime('now'::text)
FROM shoelace_data shoelace_data, shoelace_data *NEW*,
      shoelace_data *OLD*, shoelace_log shoelace_log;
```

第 2 步将规则条件加进去，因此结果集被限制为 sl\_avail 发生变化的行：

```
INSERT INTO shoelace_log SELECT
```

```

        *NEW*.sl_name, *NEW*.sl_avai,
        getpgusername(), datetime('now'::text)
    FROM shoelace_data shoelace_data, shoelace_data *NEW*,
        shoelace_data *OLD*, shoelace_log shoelace_log
    WHERE int4ne(*NEW*.sl_avail, *OLD*.sl_avail);

```

第 3 步把原始 parsetree 的条件加进去，把结果集进一步限制为只有那些原始 parsetree 会触及的行：

```

INSERT INTO shoelace_log SELECT
    *NEW*.sl_name, *NEW*.sl_avai,
    getpgusername(), datetime('now'::text)
    FROM shoelace_data shoelace_data, shoelace_data *NEW*,
        shoelace_data *OLD*, shoelace_log shoelace_log
    WHERE int4ne(*NEW*.sl_avail, *OLD*.sl_avail)
        AND bpchareq(shoelace_data.sl_name, 'sl7');

```

第 4 步把 NEW 引用替换为原始 parsetree 中的 targetlist 项，或者用结果关系中相应的变量引用来替换：

```

INSERT INTO shoelace_log SELECT
    shoelace_data.sl_name, 6,
    getpgusername(), datetime('now'::text)
    FROM shoelace_data shoelace_data, shoelace_data *NEW*,
        shoelace_data *OLD*, shoelace_log shoelace_log
    WHERE int4ne(6, *OLD*.sl_avail)
        AND bpchareq(shoelace_data.sl_name, 'sl7');

```

第 5 步把 OLD 引用替换为 resultrelation 引用：

```

INSERT INTO shoelace_log SELECT
    shoelace_data.sl_name, 6,
    getpgusername(), datetime('now'::text)
    FROM shoelace_data shoelace_data, shoelace_data *NEW*,
        shoelace_data *OLD*, shoelace_log shoelace_log
    WHERE int4ne(6, shoelace_data.sl_avail)
        AND bpchareq(shoelace_data.sl_name, 'sl7');

```

至此就完成了。化简到极致之后，规则系统的输出是一个包含两个 parsetree 的列表，它们等同于以下语句：

```

INSERT INTO shoelace_log SELECT
    shoelace_data.sl_name, 6,
    getpgusername(), 'now'
    FROM shoelace_data
    WHERE 6 != shoelace_data.sl_avail
        AND shoelace_data.sl_name = 'sl7';

UPDATE shoelace_data SET sl_avail = 6
    WHERE sl_name = 'sl7';

```

它们会按这个顺序执行，而这正是该规则所定义的行为。上述替换 (substitutions) 以及附加的条件能够保证：如果原始查询是下面这样，

```

UPDATE shoelace_data SET sl_color = 'green'
    WHERE sl_name = 'sl7';

```

就不会写入任何日志项，因为这一次原始 `parsetree` 不包含 `sl_avail` 的 `targetlist` 项，`NEW.sl_avail` 会被 `shoelace_data.sl_avail` 替换，于是规则产生的额外查询是：

```
INSERT INTO shoelace_log SELECT
    shoelace_data.sl_name, shoelace_data.sl_avail,
    getpgusername(), 'now'
FROM shoelace_data
WHERE shoelace_data.sl_avail != shoelace_data.sl_avail
AND shoelace_data.sl_name = 'sl7';
```

而该条件永远不可能为真。由于 `INSERT ... SELECT` 与 `INSERT ... VALUES` 在 `parsetree` 层面没有区别，如果原始查询修改多行，这种机制同样能够正常工作。因此，如果 AI 发出如下命令：

```
UPDATE shoelace_data SET sl_avail = 0
WHERE sl_color = 'black';
```

实际上会更新四行（`sl1`、`sl2`、`sl3` 和 `sl4`）。但 `sl3` 本来就已经是 `sl_avail = 0`。这一次，原始 `parsetree` 的条件不同，因此规则会产生额外的 `parsetree`：

```
INSERT INTO shoelace_log SELECT
    shoelace_data.sl_name, 0,
    getpgusername(), 'now'
FROM shoelace_data
WHERE 0 != shoelace_data.sl_avail
AND shoelace_data.sl_color = 'black';
```

这棵 `parsetree` 必然会插入三条新的日志项。这完全正确。

原始 `parsetree` 最后执行至关重要。Postgres 的“traffic cop”（交通警察）会在两个 `parsetree` 的执行之间做一次命令计数器递增，使第二个能看到第一个所做的更改。如果先执行 `UPDATE`，那么所有行都已经被设为零，记日志的 `INSERT` 就找不到任何满足 `0 != shoelace_data.sl_avail` 的行。

### 43.3.3. 与视图的协作

为了防止有人像前面提到的那样在视图关系上 `INSERT`、`UPDATE` 和 `DELETE` 不可见数据，一种简单的办法是让那些解析树直接被丢弃。我们创建如下规则：

```
CREATE RULE shoe_ins_protect AS ON INSERT TO shoe
DO INSTEAD NOTHING;
CREATE RULE shoe_upd_protect AS ON UPDATE TO shoe
DO INSTEAD NOTHING;
CREATE RULE shoe_del_protect AS ON DELETE TO shoe
DO INSTEAD NOTHING;
```

如果现在 AI 尝试对视图关系 `shoe` 执行这些操作中的任何一种，规则系统就会应用这些规则。由于这些规则没有动作而且是 `INSTEAD`，生成的解析树列表将为空；整个查询也就什么都不会做，因为规则系统处理完后已经没有任何内容可供优化或执行。

#### 注意

这一事实可能会让前端应用程序感到困惑，因为数据库上绝对什么都没有发生，因而后端不会为该查询返回任何东西。在 `libpq` 中连 `PGRES_EMPTY_QUERY` 之类的东西都得不到。在 `psql` 中，什么都不会发生。这一点将来可能会改变。

另一种更完善的做法，是创建一些规则，把解析树重写成在真实表上执行正确操作的解析树。要在 shoelace 视图上做到这一点，我们创建下列规则：

```
CREATE RULE shoelace_ins AS ON INSERT TO shoelace
DO INSTEAD
INSERT INTO shoelace_data VALUES (
    NEW.sl_name,
    NEW.sl_avail,
    NEW.sl_color,
    NEW.sl_len,
    NEW.sl_unit);

CREATE RULE shoelace_upd AS ON UPDATE TO shoelace
DO INSTEAD
UPDATE shoelace_data SET
    sl_name = NEW.sl_name,
    sl_avail = NEW.sl_avail,
    sl_color = NEW.sl_color,
    sl_len = NEW.sl_len,
    sl_unit = NEW.sl_unit
WHERE sl_name = OLD.sl_name;

CREATE RULE shoelace_del AS ON DELETE TO shoelace
DO INSTEAD
DELETE FROM shoelace_data
WHERE sl_name = OLD.sl_name;
```

现在有一批鞋带到货进入 AI 的商店，随附一份很长的部件清单。AI 不太擅长计算，因此我们不想让他手工更新 shoelace 视图。于是我们建立两个小表：一个让他可以插入部件清单中的项目，另一个则采用一个特殊技巧。创建命令如下：

```
CREATE TABLE shoelace_arrive (
    arr_name    char(10),
    arr_quant   integer
);

CREATE TABLE shoelace_ok (
    ok_name     char(10),
    ok_quant    integer
);

CREATE RULE shoelace_ok_ins AS ON INSERT TO shoelace_ok
DO INSTEAD
UPDATE shoelace SET
    sl_avail = sl_avail + NEW.ok_quant
WHERE sl_name = NEW.ok_name;
```

现在 AI 可以坐下来随便干点什么，直到

```
al_bundy=> SELECT * FROM shoelace_arrive;
arr_name |arr_quant
-----+-----
sl3      |         10
sl6      |         20
sl8      |         20
```

(3 rows)

与部件清单上的内容完全一致为止。我们快速查看一下当前数据，

```
al_bundy=> SELECT * FROM shoelace ORDER BY sl_name;
sl_name  |sl_avail|sl_color  |sl_len|sl_unit |sl_len_cm
-----+-----+-----+-----+-----+-----
sl1      |      5|black    |  80|cm     |    80
sl2      |      6|black    | 100|cm     |   100
sl7      |      6|brown    |  60|cm     |    60
sl3      |      0|black    |  35|inch   |   88.9
sl4      |      8|black    |  40|inch   |  101.6
sl8      |      1|brown    |  40|inch   |  101.6
sl5      |      4|brown    |   1|m     |   100
sl6      |      0|brown    | 0.9|m     |    90
(8 rows)
```

把到货的鞋带入库：

```
al_bundy=> INSERT INTO shoelace_ok SELECT * FROM shoelace_arrive;
```

然后检查结果：

```
al_bundy=> SELECT * FROM shoelace ORDER BY sl_name;
sl_name  |sl_avail|sl_color  |sl_len|sl_unit |sl_len_cm
-----+-----+-----+-----+-----+-----
sl1      |      5|black    |  80|cm     |    80
sl2      |      6|black    | 100|cm     |   100
sl7      |      6|brown    |  60|cm     |    60
sl4      |      8|black    |  40|inch   |  101.6
sl3      |     10|black    |  35|inch   |   88.9
sl8      |     21|brown    |  40|inch   |  101.6
sl5      |      4|brown    |   1|m     |   100
sl6      |     20|brown    | 0.9|m     |    90
(8 rows)
```

```
al_bundy=> SELECT * FROM shoelace_log;
sl_name  |sl_avail|log_who|log_when
-----+-----+-----+-----
sl7      |      6|A1     |Tue Oct 20 19:14:45 1998 MET DST
sl3      |     10|A1     |Tue Oct 20 19:25:16 1998 MET DST
sl6      |     20|A1     |Tue Oct 20 19:25:16 1998 MET DST
sl8      |     21|A1     |Tue Oct 20 19:25:16 1998 MET DST
(4 rows)
```

从一条 INSERT ... SELECT 到这些结果，中间要经历相当长的一段过程。对它的描述将是本文中的最后一个（但不是最后一个例子 :-))。首先是解析器的输出：

```
INSERT INTO shoelace_ok SELECT
    shoelace_arrive.arr_name, shoelace_arrive.arr_quant
FROM shoelace_arrive shoelace_arrive, shoelace_ok shoelace_ok;
```

现在应用第一条规则 'shoelace\_ok\_ins'，它会把这一输出转换成：

```
UPDATE shoelace SET
    sl_avail = int4pl(shoelace.sl_avail, shoelace_arrive.arr_quant)
```

```

FROM shoelace_arrive shoelace_arrive, shoelace_ok shoelace_ok,
     shoelace_ok *OLD*, shoelace_ok *NEW*,
     shoelace shoelace
WHERE bpchareq(shoelace.sl_name, shoelace_arrive.arr_name);

```

同时会丢弃针对 shoelace\_ok 的原始 INSERT。这个重写后的查询会再次交给规则系统，随后应用的规则 'shoelace\_upd' 产生：

```

UPDATE shoelace_data SET
    sl_name = shoelace.sl_name,
    sl_avail = int4pl(shoelace.sl_avail, shoelace_arrive.arr_quant)
,
    sl_color = shoelace.sl_color,
    sl_len = shoelace.sl_len,
    sl_unit = shoelace.sl_unit
FROM shoelace_arrive shoelace_arrive, shoelace_ok shoelace_ok,
     shoelace_ok *OLD*, shoelace_ok *NEW*,
     shoelace shoelace, shoelace *OLD*,
     shoelace *NEW*, shoelace_data shoelace_data
WHERE bpchareq(shoelace.sl_name, shoelace_arrive.arr_name)
     AND bpchareq(shoelace_data.sl_name, shoelace.sl_name);

```

这同样是一条 INSTEAD 规则，因此前一个解析树会被丢弃。注意，这个查询仍然使用视图 shoelace。但规则系统在这个循环中尚未完成，因此它会继续在其上应用规则 'RETshoelace'，于是得到：

```

UPDATE shoelace_data SET
    sl_name = s.sl_name,
    sl_avail = int4pl(s.sl_avail, shoelace_arrive.arr_quant),
    sl_color = s.sl_color,
    sl_len = s.sl_len,
    sl_unit = s.sl_unit
FROM shoelace_arrive shoelace_arrive, shoelace_ok shoelace_ok,
     shoelace_ok *OLD*, shoelace_ok *NEW*,
     shoelace shoelace, shoelace *OLD*,
     shoelace *NEW*, shoelace_data shoelace_data,
     shoelace *OLD*, shoelace *NEW*,
     shoelace_data s, unit u
WHERE bpchareq(s.sl_name, shoelace_arrive.arr_name)
     AND bpchareq(shoelace_data.sl_name, s.sl_name);

```

又有一条更新规则被应用，于是车轮继续转动，我们进入了第 3 轮重写。这一次应用的是规则 'log\_shoelace'，这产生了额外的解析树：

```

INSERT INTO shoelace_log SELECT
    s.sl_name,
    int4pl(s.sl_avail, shoelace_arrive.arr_quant),
    getpguser(),
    datetime('now'::text)
FROM shoelace_arrive shoelace_arrive, shoelace_ok shoelace_ok,
     shoelace_ok *OLD*, shoelace_ok *NEW*,
     shoelace shoelace, shoelace *OLD*,
     shoelace *NEW*, shoelace_data shoelace_data,
     shoelace *OLD*, shoelace *NEW*,
     shoelace_data s, unit u,

```

```

        shoelace_data *OLD*, shoelace_data *NEW*
        shoelace_log shoelace_log
WHERE bpchareq(s.sl_name, shoelace_arrive.arr_name)
      AND bpchareq(shoelace_data.sl_name, s.sl_name);
      AND int4ne(int4pl(s.sl_avail, shoelace_arrive.arr_quant), s.sl_
avail);

```

到这里，规则系统已经没有更多规则可用，返回生成的解析树。于是我们最终得到两棵解析树，它们等同于以下 SQL 语句：

```

INSERT INTO shoelace_log SELECT
    s.sl_name,
    s.sl_avail + shoelace_arrive.arr_quant,
    getpgusername(),
    'now'
FROM shoelace_arrive shoelace_arrive, shoelace_data shoelace_data,
     shoelace_data s
WHERE s.sl_name = shoelace_arrive.arr_name
     AND shoelace_data.sl_name = s.sl_name
     AND s.sl_avail + shoelace_arrive.arr_quant != s.sl_avail;

UPDATE shoelace_data SET
    sl_avail = shoelace_data.sl_avail + shoelace_arrive.arr_quant
FROM shoelace_arrive shoelace_arrive, shoelace_data shoelace_data,
     shoelace_data s
WHERE s.sl_name = shoelace_arrive.sl_name
     AND shoelace_data.sl_name = s.sl_name;

```

结果是：来自一个关系的数据被插入到另一个关系中，这个插入被改写为对第三个关系的更新，再被改写为对第四个关系的更新外加在第五个关系中记录该更新，最后整个过程被化简为两个查询。

这里有个稍显难看的小细节。观察这两个查询就会发现，`shoelace_data` 关系在范围表中出现了两次，而实际上完全可以缩减成一次。优化器不会处理这一点，因此规则系统为 INSERT 输出的执行计划会是：

```

Nested Loop
-> Merge Join
    -> Seq Scan
        -> Sort
            -> Seq Scan on s
    -> Seq Scan
        -> Sort
            -> Seq Scan on shoelace_arrive
-> Seq Scan on shoelace_data

```

而省略那个额外的范围表项则会得到：

```

Merge Join
-> Seq Scan
    -> Sort
        -> Seq Scan on s
-> Seq Scan
    -> Sort
        -> Seq Scan on shoelace_arrive

```

后者完全会在日志关系中产生相同的项。因此，规则系统导致了对 `shoelace_data` 关系的一次完全不必要的额外扫描。而在 `UPDATE` 中，同样的冗余扫描还会再发生一次。不过，能让这一切总体上工作起来，已经是一项相当艰巨的工作。

最后来演示一下 Postgres 规则系统及其威力。有一位漂亮的金发女郎卖鞋带。而 AI 永远不会意识到的是，她不仅漂亮，还很聪明——甚至有点太聪明了。于是，AI 不时会订购一些完全卖不出去的鞋带。这一次他订购了 1000 双品红色（magenta）鞋带，而且由于另一种鞋带暂时缺货、而他又已承诺购买一些，他又为粉色（pink）的鞋带做好了数据库准备：

```
al_bundy=> INSERT INTO shoelace VALUES
al_bundy->      ('sl9', 0, 'pink', 35.0, 'inch', 0.0);
al_bundy=> INSERT INTO shoelace VALUES
al_bundy->      ('sl10', 1000, 'magenta', 40.0, 'inch', 0.0);
```

由于这种事经常发生，我们必须不时地找出那些绝对不适合任何鞋子的鞋带项。我们可以每次都有一条复杂的语句来完成，也可以为此建立一个视图。这个视图是：

```
CREATE VIEW shoelace_obsolete AS
SELECT * FROM shoelace WHERE NOT EXISTS
      (SELECT shoename FROM shoe WHERE slcolor = sl_color);
```

它的输出是：

```
al_bundy=> SELECT * FROM shoelace_obsolete;
sl_name    |sl_avail|sl_color  |sl_len|sl_unit |sl_len_cm
-----+-----+-----+-----+-----+-----
sl9         |      0|pink      |  35|inch   |   88.9
sl10        |    1000|magenta   |  40|inch   |  101.6
```

那 1000 双品红色鞋带，我们得先向 AI 讨了债才能把它们扔掉，不过那是另一个问题了。粉色的条目我们要删掉。为了给 Postgres 增加一点难度，我们不直接删除，而是再创建一个视图：

```
CREATE VIEW shoelace_candelelete AS
SELECT * FROM shoelace_obsolete WHERE sl_avail = 0;
```

然后这样执行：

```
DELETE FROM shoelace WHERE EXISTS
      (SELECT * FROM shoelace_candelelete
       WHERE sl_name = shoelace.sl_name);
```

Voila:

```
al_bundy=> SELECT * FROM shoelace;
sl_name    |sl_avail|sl_color  |sl_len|sl_unit |sl_len_cm
-----+-----+-----+-----+-----+-----
sl1         |      5|black     |  80|cm      |   80
sl2         |      6|black     | 100|cm      |  100
sl7         |      6|brown     |  60|cm      |   60
sl4         |      8|black     |  40|inch    |  101.6
sl3         |     10|black     |  35|inch    |   88.9
sl8         |     21|brown     |  40|inch    |  101.6
sl10        |    1000|magenta   |  40|inch    |  101.6
sl5         |      4|brown     |   1|m       |   100
sl6         |     20|brown     |  0.9|m     |    90
(9 rows)
```

作用于一个视图上的 DELETE，其子查询条件总共使用了四个嵌套/连接的视图，其中一个视图自身又带有一个包含视图的子查询条件，并且还用到计算得到的视图列，最终仍会被重写成单棵解析树，从真正的表中删除所请求的数据。

我想在现实世界中，大概只有很少的场景会需要这样的构造。但知道它确实能工作，总归让人安心。

### 实情是

在写这一章的过程中，我又发现了一个 bug。不过修完那个 bug 之后，这一切居然能工作，倒是让我有点惊讶。

## 43.4. 规则和权限

由于 Postgres 规则系统会重写查询，因此可能访问原始查询中并未使用的其他表或视图。使用更新规则时，这甚至可能包括对表的写访问。

重写规则没有单独的所有者。关系（表或视图）的所有者会自动成为为其定义的重写规则的所有者。Postgres 规则系统改变了默认访问控制系统的行为。由于规则而使用的关系，会在重写期间根据其上定义了该规则的关系所有者的权限进行检查。这意味着，用户只需要对其查询中明确命名的表/视图具有所需的权限。

例如，某个用户有一份电话号码列表，其中一部分是私人的，另一部分对办公室秘书有用。该用户可以这样构造：

```
CREATE TABLE phone_data (person text, phone text, private bool);
CREATE VIEW phone_number AS
    SELECT person, phone FROM phone_data WHERE NOT private;
GRANT SELECT ON phone_number TO secretary;
```

除了该用户本人（以及数据库超级用户）之外，没有人可以访问 phone\_data 表。但由于 GRANT 的存在，秘书可以对 phone\_number 视图执行 SELECT。规则系统会把对 phone\_number 的 SELECT 重写成对 phone\_data 的 SELECT，并附加只选取 private 为假的项的条件。由于该用户是 phone\_number 的所有者，对 phone\_data 的读访问会按照该用户的权限进行检查，于是查询被判定为允许。对 phone\_number 本身的访问检查仍会执行，因此除秘书之外没有人能使用它。

权限是逐条规则检查的。因此，目前秘书是唯一能看到公开电话号码的人。但秘书还可以再建立一个视图，并把该视图授予公众访问。这样，任何人都可以通过秘书的视图看到 phone\_number 中的数据。秘书做不到的是创建一个直接访问 phone\_data 的视图（其实可以创建，但它不会起作用，因为每次访问都会在权限检查时中止事务）。而且，一旦用户发现秘书开放了其 phone\_number 视图，用户就可以 REVOKE 秘书的访问权限。那样一来，对秘书视图的任何访问都会立即失败。

有人可能会认为这种逐条规则的检查是一个安全漏洞，但实际上并不是。如果它不这样工作，秘书也可以建立一个与 phone\_number 具有相同列的表，每天把数据复制进去。那样一来，这些就是秘书自己的数据，他同样可以把访问权限授予任何人。GRANT 的含义本来就是“我信任你”。如果某个你信任的人做了上面的事，那么该重新考虑这份信任，然后使用 REVOKE 了。

这种机制对更新规则同样有效。在上一节的示例中，AI 数据库中那些表的所有者可以把 shoelace 视图上的 SELECT、INSERT、UPDATE 和 DELETE 权限授予 al。但对 shoelace\_log 只授予

SELECT 权限。写日志记录的规则动作仍会被成功执行，AI 也能看到日志记录。但他不能伪造记录，也不能操纵或删除已有记录。

### 警告

GRANT ALL 目前会包含 RULE 权限。这意味着被授权的用户可以先删除规则、做出更改、再把规则装回去。我认为这一点应该尽快改变。

## 43.5. 规则与触发器

许多能够用触发器完成的事情，同样也可以用 Postgres 规则系统实现。目前不能用规则实现的内容是一些种类的约束。你可以放置一条带条件的规则：当某列的值没有出现在另一张表中时，把查询重写成 NOTHING。但那样做会悄无声息地丢弃数据，这并不是好主意。如果需要检查值的有效性，并在值无效时生成错误消息，目前就必须使用触发器。

另一方面，在视图上因 INSERT 而触发的触发器可以做到与规则相同的事情：把数据放到别处，并抑制向视图本身的插入。但在 UPDATE 或 DELETE 上它就无能为力了，因为视图关系中没有被扫描的真实数据，触发器因而永远不会被调用。这时只有规则才能解决问题。

对于两者都能实现的事情，哪一种更好取决于数据库的使用方式。触发器会对每个受影响的行触发一次，而规则会修改解析树或生成额外的解析树。因此，如果一条语句会影响很多行，那么发出一条额外查询的规则往往会比触发器更快，因为触发器必须为每一行调用一次，并反复执行其操作。

例如：有两个表：

```
CREATE TABLE computer (
    hostname      text      -- indexed
    manufacturer  text      -- indexed
);

CREATE TABLE software (
    software      text,      -- indexed
    hostname      text      -- indexed
);
```

两个表都有数千行，并且 hostname 上的索引是唯一索引。hostname 列包含计算机的完整域名。规则或触发器应实现这样一个约束：从 software 中删除引用了被删除主机的行。由于从 computer 中删除的每一行都会调用触发器，它可以在一个准备并保存好的计划中使用语句：

```
DELETE FROM software WHERE hostname = $1;
```

并通过参数传入 hostname。规则则会写成：

```
CREATE RULE computer_del AS ON DELETE TO computer
DO DELETE FROM software WHERE hostname = OLD.hostname;
```

现在我们来看不同类型的删除。对于下面这种情况：

```
DELETE FROM computer WHERE hostname = 'mypc.local.net';
```

computer 表会通过索引扫描（很快），由触发器发出的查询也会是一次索引扫描（同样很快）。规则产生的额外查询则是：

```
DELETE FROM software WHERE computer.hostname = 'mypc.local.net'
AND software.hostname = computer.hostname;
```

由于已经建立了合适的索引，优化器将创建一个计划：

```
Nestloop
-> Index Scan using comp_hostidx on computer
-> Index Scan using soft_hostidx on software
```

因此，触发器实现和规则实现之间的速度差异不会太大。在下一个删除场景中，我们想去掉全部 2000 台 hostname 以 'old' 开头的计算机。有两种查询可以做到这件事。其中一种是：

```
DELETE FROM computer WHERE hostname >= 'old'
AND hostname < 'ole'
```

规则查询的计划将是：

```
Hash Join
-> Seq Scan on software
-> Hash
-> Index Scan using comp_hostidx on computer
```

另一种可能的查询是：

```
DELETE FROM computer WHERE hostname ~ '^old';
```

其执行计划为：

```
Nestloop
-> Index Scan using comp_hostidx on computer
-> Index Scan using soft_hostidx on software
```

这说明，当多个条件表达式通过 AND 组合在一起时，优化器并没有意识到，对 computer 上 hostname 的限制同样也可以用于 software 上的索引扫描，而在该查询的正则表达式版本里它却能做到。触发器会对需要删除的 2000 台旧计算机中的每一台各调用一次，这意味着对 computer 做一次索引扫描，并对 software 做 2000 次索引扫描。规则实现则用两条走索引的查询完成这件事。不过，在顺序扫描这种情况下，规则是否仍然更快，还取决于 software 表的总体大小。即使所有用于查找的索引块很快都会进入缓存，通过 SPI 管理器执行 2000 条查询仍然要耗费一些时间。

我们要看的最后一个查询是：

```
DELETE FROM computer WHERE manufacurer = 'bim';
```

同样，这也可能导致从 computer 中删除很多行。因此触发器又会向执行器触发很多条查询。但规则的计划又将是在两个索引扫描上的嵌套循环，只不过使用了 computer 上的另一个索引：

```
Nestloop
-> Index Scan using comp_manufidx on computer
-> Index Scan using soft_hostidx on software
```

它来自规则的查询：

```
DELETE FROM software WHERE computer.manufacurer = 'bim'
AND software.hostname = computer.hostname;
```

在上述任何一种情况下，规则系统产生的额外查询都或多或少与该查询影响的行数无关。

另一种情况是 UPDATE 中是否执行动作取决于某个属性是否发生变化。在 Postgres 6.4 版中，规则事件的属性说明被禁用了（它最迟会在 6.5 恢复，也许更早——敬请期待）。因此，目前创建像 `shoelace_log` 例子中那样规则的唯一方式是使用规则条件。这会导致一条额外的查询总是被执行，即使所关注的属性根本不可能变化，因为它没有出现在初始查询的目标列表中。

等这一功能重新启用后，它将成为规则相对触发器的又一个优势。在这种情况下，触发器的优化按定义必然失败，

因为“其动作只在某个特定属性被更新时执行”这一事实隐藏在它的功能之中。

触发器的定义只允许在行级指定，因此每当某一行被触及，触发器就必须被调用来做决定。

而规则系统会通过查看目标列表知道这一点，并在属性未被涉及时完全抑制那条额外的查询。

所以，无论带不带条件，规则只会在可能有事可做时才执行其扫描。

只有当规则动作导致了规模很大且条件很差的连接（优化器无能为力的情形）时，规则才会明显慢于触发器。规则是一把大锤。不加小心地使用大锤可能造成巨大的破坏。但用对了手法，它们能把任何钉子一击入帽。

# 第 44 章 索引扩展接口

到目前为止所描述的过程使我们能够定义新类型、新函数以及新操作符。然而，我们还不能在一个新类型或其操作符上定义一个二级索引（例如 B-tree、R-tree 或哈希访问方法）。

回头看图 38.1。右半部分显示了我们必须修改的目录，以便告诉 Postgres 如何在索引中使用用户定义的类型和/或用户定义的操作符（即 `pg_am`，`pg_amop`，`pg_amproc` 和 `pg_opclass`）。遗憾的是，没有一条简单的命令可以完成这件事。我们将通过一个贯穿始终的示例来演示如何修改这些目录：为 B-tree 访问方法定义一个新的操作符类，以便按绝对值升序排序整数。

`pg_am` 类为每个用户定义的访问方法保存一个实例。堆访问方法的支持内置于 Postgres 中，但每个其他访问方法都在这里描述。其模式是

表 44.1. 索引模式

| 属性                                               | 描述                                                                   |
|--------------------------------------------------|----------------------------------------------------------------------|
| <code>amname</code>                              | 访问方法的名称                                                              |
| <code>amowner</code>                             | 属主在 <code>pg_user</code> 中实例的对象 id                                   |
| <code>amkind</code>                              | 当前未使用，但保留为占位符并置为 'o'                                                 |
| <code>amstrategies</code>                        | 该访问方法的策略数目（见下文）                                                      |
| <code>amsupport</code>                           | 该访问方法的支持例程数目（见下文）                                                    |
| <code>amgettupl</code> <code>aminsert</code> ... | 访问方法接口例程的过程标识符。例如，打开、关闭访问方法以及从中获取实例的 <code>regproc id</code> 就出现在这里。 |

`pg_am` 中该实例的对象 ID 被用作许多其他类中的外键。你不需要向这个类添加新实例；你所关心的只是你想扩展的访问方法实例的对象 ID：

```
SELECT oid FROM pg_am WHERE amname = 'btree';

+----+
|oid |
+----+
| 403 |
+----+
```

`amstrategies` 属性的存在是为了标准化跨数据类型的比较。例如，B-tree 对键施加了严格的从小到大的顺序。由于 Postgres 允许用户定义操作符，Postgres 不能仅凭操作符名称（例如 ">" 或 "<") 就判断它是哪一类比较。事实上，某些访问方法根本不施加任何排序。例如，R-tree 表达一种矩形包含关系，而哈希数据结构只表达基于哈希函数值的按位相似性。Postgres 需要某种一致的方式，来接受你查询中的条件、查看操作符，然后判断是否存在可用的索引。这蕴含着 Postgres 需要知道，例如，"<=" 和 ">" 操作符划分一个 B-tree。Postgres 使用策略来表达操作符与它们可用于扫描索引的方式之间的这些关系。

定义一组新的策略超出了本讨论的范围，但我们将解释 B-tree 策略如何工作，因为你要添加一个新的操作符类就需要了解它。在 `pg_am` 类中，`amstrategies` 属性是该访问方法所定义的策略数目。对 B-tree 来说，这个数是 5。这些策略对应于

表 44.2. B-tree 策略

| 操作   | 索引号 |
|------|-----|
| 小于   | 1   |
| 小于等于 | 2   |
| 等于   | 3   |
| 大于等于 | 4   |
| 大于   | 5   |

这个思路是：你需要把与上述比较对应的过程添加到 `pg_amop` 关系中（见下文）。访问方法代码可以使用这些策略号（不管数据类型是什么）来弄清如何划分 B-tree、计算选择性等等。现在还不用担心添加过程的细节；只需理解：对 `int2`、`int4`、`oid`，以及 B-tree 可以操作的每个其他数据类型，都必须存在一组这样的过程。有时，仅靠策略信息不足以让系统弄清如何使用索引。某些访问方法还需要其他支持例程才能工作。例如，B-tree 访问方法必须能够比较两个键，并判断其中一个是大于、等于还是小于另一个。类似地，R-tree 访问方法必须能够计算矩形的交集、并集和大小。这些操作并不对应 SQL 查询中的用户条件；它们是访问方法内部使用的管理例程。

为了在所有 PostgreSQL 访问方法之间一致地管理多样的支持例程，`pg_am` 包含一个名为 `amsupport` 的属性。该属性记录一个访问方法所使用的支持例程数目。对 B-tree 来说，这个数是一——接受两个键并根据第一个键小于、等于还是大于第二个键而返回 -1、0 或 +1 的例程。

### 注意

严格地说，该例程可以返回一个负数 (< 0)、0 或一个非零正数 (> 0)。

`pg_am` 中的 `amstrategies` 项只是所论访问方法定义的策略数目。小于、小于等于等的过程并不出现在 `pg_am` 中。类似地，`amsupport` 只是该访问方法所需的支持例程数目。实际的例程列在别处。

下一个我们关心的类是 `pg_opclass`。这个类的存在只是为了把一个名称与一个 `oid` 相关联。在 `pg_amop` 中，每个 B-tree 操作符类都有上文的第 1 到第 5 号一组过程。一些现有的操作符类是 `int2_ops`、`int4_ops` 和 `oid_ops`。你需要把带有你的操作符类名（例如 `complex_abs_ops`）的一个实例添加到 `pg_opclass`。这一实例的 `oid` 是其他类中的外键。

```
INSERT INTO pg_opclass (opcname) VALUES ('complex_abs_ops');
```

```
SELECT oid, opcname
FROM pg_opclass
WHERE opcname = 'complex_abs_ops';
```

```
+-----+-----+
|oid    | opcname      |
+-----+-----+
|17314  | int4_abs_ops |
+-----+-----+
```

注意，你的 `pg_opclass` 实例的 `oid` 会不同！在本讨论中凡出现 17314 之处，你都应当用你得到的值替换。

现在我们有了一个访问方法和一个操作符类。我们还需要一组操作符；定义操作符的过程已在本手册前面讨论过。对于 Btree 上的 `complex_abs_ops` 操作符类，我们需要的操作符是：

```
absolute value less-than
absolute value less-than-or-equal
absolute value equal
absolute value greater-than-or-equal
absolute value greater-than
```

假设实现所定义函数的代码存储在文件 `PGROOT/src/tutorial/complex.c`

代码的一部分看起来像这样：（注意，在余下的示例中我们只展示相等操作符。其他四个操作符非常类似。细节请参阅 `complex.c` 或 `complex.sql`。）

```
#define Mag(c) ((c)->x*(c)->x + (c)->y*(c)->y)

bool
complex_abs_eq(Complex *a, Complex *b)
{
    double amag = Mag(a), bmag = Mag(b);
    return (amag==bmag);
}
```

下面有几点重要的事情正在发生。

首先，注意这里正在为 `int4` 定义小于、小于等于、等于、大于等于和大于操作符。所有这些操作符都已以 `<`、`<=`、`=`、`>=` 和 `>` 的名称为 `int4` 定义过。当然，新的操作符行为不同。为了确保 `Postgres` 使用这些新操作符而不是旧操作符，它们的命名必须与旧操作符不同。这是一个要点：你可以在 `Postgres` 中重载操作符，但仅当该操作符尚未为这些参数类型定义时才行。也就是说，如果你已经为 `(int4, int4)` 定义了 `<`，就不能再定义一次。`Postgres` 在你定义操作符时不检查这一点，所以要小心。为了避免这个问题，将为这些操作符使用奇怪的名字。如果你弄错了，访问方法在你尝试做扫描时很可能会崩溃。

另一个要点是所有操作符函数都返回布尔值。访问方法依赖这一事实。（另一方面，支持函数返回的是特定访问方法所期望的类型——在本例中就是一个有符号整数。）文件中的最后一个例程是我们在讨论 `pg_am` 类的 `amsupport` 属性时提到的“支持例程”。我们稍后会用到它。现在先忽略它。

```
CREATE FUNCTION complex_abs_eq(complex, complex)
    RETURNS bool
    AS 'PGROOT/tutorial/obj/complex.so'
    LANGUAGE 'c';
```

现在定义使用它们的操作符。如前所述，操作符名在所有取两个 `int4` 操作数的操作符中必须唯一。为了查看下面列出的操作符名是否已被占用，我们可以在 `pg_operator` 上做一次查询：

```
/*
 * this query uses the regular expression operator (~)
 * to find three-character operator names that end in
 * the character &
 */
SELECT *
FROM pg_operator
WHERE oprname ~ '^...&$'::text;
```

以查看你想要的名字是否已被用于你想要的类型。这里的重要内容是过程（即上面定义的 `C` 函数）以及限制和连接选择性函数。你应当直接使用下面所用的那些——注意，小于、等于和大于情形有不同的此类函数。必须提供这些函数，

否则访问方法在试图使用该操作符时将会崩溃。你应当照抄 restrict 和 join 的名称，但使用你在上一步中定义的过程名。

```
CREATE OPERATOR = (  
    leftarg = complex, rightarg = complex,  
    procedure = complex_abs_eq,  
    restrict = eqsel, join = eqjoinsel  
)
```

注意，这里定义了对应于小于、小于等于、等于、大于和大于等于的五个操作符。

我们快完成了。最后要做的事情是更新 pg\_amop 关系。为此，我们需要下列属性：

表 44.3. pg\_amproc 模式

| 属性                     | 描述                                                            |
|------------------------|---------------------------------------------------------------|
| amopid                 | pg_am 中 B-tree 实例的 oid (== 403, 见上文)                          |
| amopclaid              | int4_abs_ops 的 pg_opclass 实例的 oid (== 你得到的那个替代 17314 的值, 见上文) |
| amopopr                | 该操作符类的各操作符的 oid (我们马上就会取得)                                    |
| amopselect, amopnpages | 代价函数                                                          |

代价函数供查询优化器用来决定在一次扫描中是否使用某个索引。幸运的是，这些函数已经存在。我们将使用的两个函数是 btreesel (它估计 B-tree 的选择性) 和 btreeenpage (它估计一次搜索将在树中触及的页面数)。

所以我们需要刚才定义的这些操作符的 oid。我们将查找所有取两个 int4 类型操作数的操作符的名称，并挑出我们的那些：

```
SELECT o.oid AS opoid, o.oprname  
    INTO TABLE complex_ops_tmp  
    FROM pg_operator o, pg_type t  
    WHERE o.oprleft = t.oid and o.oprright = t.oid  
        and t.typname = 'complex';
```

```
+-----+-----+  
|oid   |oprname|  
+-----+-----+  
|17321 |<      |  
+-----+-----+  
|17322 |<=     |  
+-----+-----+  
|17323 |=     |  
+-----+-----+  
|17324 |>=     |  
+-----+-----+  
|17325 |>      |  
+-----+-----+
```

(同样，你的某些 oid 号几乎肯定会有所不同。) 我们感兴趣的操作符是 oid 从 17321 到 17325 的那些。你得到的值很可能不同，你应当用它们替换下文中的值。我们可以查看操作符名称并挑出我们刚添加的那些。

现在我们准备好用新操作符类来更新 pg\_amop 了。在整个讨论中最重要的一点是：在 pg\_amop 中，操作符是按从小等于到大等于排序的。我们添加所需的实例：

```

INSERT INTO pg_amop (amopid, amopclaid,
                    amopopr, amopstrategy,
                    amopselect, amopnpages)
SELECT am.oid, opcl.oid, c.opoid, 3,
       'btreesel'::regproc, 'btree'::regproc
FROM pg_am am, pg_opclass opcl, complex_ops_tmp c
WHERE amname = 'btree'
      and opcname = 'complex_abs_ops'
      and c.oprname = '=';

```

注意顺序: "小于"是 1, "小于等于"是 2, "等于"是 3, "大于等于"是 4, "大于"是 5。

最后一步(终于!)是注册此前在讨论 `pg_am` 时描述过的"支持例程"。该支持例程的 `oid` 存储在 `pg_amproc` 类中, 以访问方法 `oid` 和操作符类 `oid` 为键。首先, 我们需要在 Postgres 中注册该函数(回想一下, 我们把实现该例程的 C 代码放在了实现操作符例程的那个文件的底部):

```

CREATE FUNCTION int4_abs_cmp(int4, int4)
RETURNS int4
AS 'PGROOT/tutorial/obj/complex.so'
LANGUAGE 'c';

```

```

SELECT oid, proname FROM pg_proc
WHERE proname = 'int4_abs_cmp';

```

```

+-----+-----+
|oid    | proname    |
+-----+-----+
|17328  | int4_abs_cmp |
+-----+-----+

```

(同样, 你的 `oid` 号很可能会有所不同, 你应当用你看到的值替换下文中的值。)回想 B-tree 实例的 `oid` 是 403 而 `int4_abs_ops` 的是 17314, 我们可以像下面这样添加新实例:

```

INSERT INTO pg_amproc (amid, amopclaid, amproc, amprocnum)
VALUES ('403'::oid, -- btree oid
       '17314'::oid, -- pg_opclass tuple
       '17328'::oid, -- new pg_proc oid
       '1'::int2);

```

# 第 45 章 GiST Indices

Gene Selkov

关于 GiST 的信息在 <http://GiST.CS.Berkeley.EDU:8000/gist/> 关于不同索引和排序方案的更多内容在 <http://s2k-ftp.CS.Berkeley.EDU:8000/personal/jmh/> 另外在伯克利数据库站点还有一些有意思的阅读材料：<http://epoch.cs.berkeley.edu:8000/>。

## 作者

这段摘录来自 Eugene Selkov Jr.<sup>1</sup> 发送的一封电子邮件，其中包含关于 GiST 的有用信息。希望我们将来能学到更多并更新这些信息。 - thomas 1998-03-01

好吧，我不能说完全理解了其中的原理，但我（几乎）成功地把 GiST 示例移植到了 linux。GiST 访问方法已经在 postgres 树中（src/backend/access/gist）。

伯克利的示例<sup>2</sup>附带这些方法的概述，并演示了二维矩形、多边形、整数区间和文本的空间索引机制（另见伯克利的 GiST<sup>3</sup>）。在矩形示例中，使用 GiST 索引本应看到性能提升；它对我来说确实有效，但我没有规模足够大的矩形集合来验证这一点。其他示例也都有效，只有多边形除外：执行

```
test=> create index pix on polytmp
test-> using gist (p:box gist_poly_ops) with (islossy);
ERROR:  cannot open pix
```

(PostgreSQL 6.3

Sun Feb 1 14:57:30 EST 1998)

我没搞懂这条错误消息的含义；它似乎是我们更应该去问开发者的问题（另见下面的注 4）。我这里想建议的是，你们当中的 linux 用户（linux==gcc?）去取上面提到的原始源码并应用我的补丁（见附件），然后告诉我们你们的感受。我觉得很酷，但周围有这么多人，我不想把它扣着不放。

关于这些源码的几点说明：

1. 我没能用上原始的（HPUX）Makefile，于是把古老的 postgres95 教程里的 Makefile 改了改来干活。我试着让它通用一些，但我写 Makefile 的水平很差——只是做了些照猫画虎的活儿。抱歉，不过我想它现在比原来的 makefile 稍微可移植一点了。

2. 我直接在 pgsqll/src 之下构建了示例源码（只是把 tar 文件解压在那里）。前面提到的 Makefile 假定它位于 pgsqll/src 下一级（在我们的例子中即 pgsqll/src/pggist）。

3. 我对 \*.c 文件的改动都只是关于 #include、函数原型和类型转换。除此之外，我只是扔掉了一堆未用的变量，并加了几个括号来取悦 gcc。希望我没有搞砸太多：)

4. polyproc.sql 中有一条注释：

```
-- -- there's a memory leak in rtree poly_ops!!
-- -- create index pix2 on polytmp using rtree (p poly_ops);
```

<sup>1</sup><mailto:selkovjr@mcs.anl.gov>

<sup>2</sup><ftp://s2k-ftp.cs.berkeley.edu/pub/gist/pggist/pggist.tgz>

<sup>3</sup><http://gist.cs.berkeley.edu:8000/gist/>

收到！！我以为它可能与几个版本之前的 Postgres 有关，于是试着执行了该查询。我的系统疯了，大约十分钟后我不得不干掉了 postmaster。

我会继续研究一阵子 GiST，但我也希望看到更多 R 树用法的示例。

## 第 46 章 链接动态装载的函数

在你创建并注册一个用户定义的函数之后，你的工作基本上就完成了。然而，Postgres必须装载实现你的函数的目标代码（例如一个 .o 文件或一个共享库）。如前所述，Postgres会在运行时按需装载你的代码。为了让你的代码能够被动态装载，你可能必须以一种特殊的方式对它进行编译和链接编辑。本节简要描述在你把用户定义的函数装载进运行中的Postgres服务器之前，需要怎样进行编译和链接编辑。注意，这一过程从 4.2 版开始已经改变。

### 提示

旧的Postgres动态装载机制要求编写动态装载器的人深入了解可执行文件格式、可执行指令在内存中的放置和对齐等方面的知识。那样的装载器往往又慢又有缺陷。从 4.2 版开始，Postgres的动态装载机制已被重写为使用操作系统提供的动态装载机制。这种做法通常比我们以前的动态装载机制更快、更可靠也更可移植。原因在于，几乎所有现代版本的 UNIX 都使用动态装载机制来实现共享库，因此必然提供快速可靠的机制。另一方面，目标文件在被装载进 Postgres之前必须先做一点后处理。我们希望速度和可靠性上的大幅提升能弥补便利性上的轻微下降。

如果你有具体问题，应该预期去阅读（反复阅读）C 编译器 `cc(1)` 和链接编辑器 `ld(1)` 的手册页。此外，`PGROOT/src/regress` 目录中的回归测试套件包含这一过程的若干可用示例。如果你照抄这些测试的做法，就不会有任何问题。下面将使用下列术语：

- 动态装载（Dynamic loading）是Postgres对一个目标文件所做的事情。目标文件被复制进运行中的Postgres服务器，文件中的函数和变量变得可供Postgres进程中的函数使用。Postgres使用操作系统提供的动态装载机制来完成这件事。
- 装载和链接编辑（Loading and link editing）是你对一个目标文件所做的事情，目的是产生另一种目标文件（例如一个可执行程序或一个共享库）。你使用链接编辑程序 `ld(1)` 来完成这件事。

下列一般性限制和说明也适用于下面的讨论：

- 传给 `create function` 命令的路径必须是绝对路径（即以 `"/"` 开头），并且指向运行Postgres服务器的机器上可见的目录。

### 提示

相对路径实际上也能工作，但它是相对于数据库所在目录的（前端应用通常看不到这个目录）。显然，让路径相对于用户启动前端应用的目录是没有意义的，因为服务器可能运行在一台完全不同的机器上！

- Postgres用户必须能够遍历传给 `create function` 命令的路径，并且能够读取该目标文件。这是因为Postgres服务器以Postgres用户身份运行，而不是以启动前端进程的用户身份运行。（让文件或上层目录对“postgres”用户不可读和/或不可执行是一个极其常见的错误。）
- 目标文件内定义的符号名彼此不能冲突，也不能与Postgres中定义的符号冲突。
- GNU C 编译器通常不提供使用操作系统动态装载器接口所需的特殊选项。在这种情况下，必须使用操作系统自带的 C 编译器。

## 46.1. ULTRIX

在 ULTRIX 下构建动态装载的目标文件非常容易。ULTRIX 没有任何共享库机制，因此对动态装载器接口没有任何限制。另一方面，我们不得不自己（重）写一个不可移植的动态装载器，也无法使用真正的共享库。在 ULTRIX 下，唯一的限制是每个目标文件都必须用选项 `-G 0` 生成。（注意那是数字“0”而不是字母“O”。）例如，

```
# simple ULTRIX example
% cc -G 0 -c foo.c
```

会产生一个名为 `foo.o` 的目标文件，随后可以把它动态装载进Postgres。不需要再做额外的装载或链接编辑。

## 46.2. DEC OSF/1

在 DEC OSF/1 下，你可以取任意简单的目标文件，用带正确选项的 `ld` 命令处理它来产生一个共享目标文件。相应的命令看起来像这样：

```
# simple DEC OSF/1 example
% cc -c foo.c
% ld -shared -expect_unresolved '*' -o foo.so foo.o
```

得到的共享目标文件随后就可以装载进Postgres。在向 `create function` 命令指定目标文件名时，必须给出共享目标文件的名字（以 `.so` 结尾），而不是简单的目标文件。

### 提示

实际上，只要是一个共享目标文件，Postgres并不关心你把它命名成什么。如果你更喜欢用扩展名 `.o` 来命名共享目标文件，这对Postgres也没问题，只要确保把正确的文件名给了 `create function` 命令。换句话说，你只需保持一致。不过，从实用的角度看，我们不鼓励这种做法，因为你无疑会搞混哪些文件已经被做成了共享目标文件、哪些没有。例如，如果目标文件和共享目标文件都以 `.o` 结尾，就很难编写 Makefile 来自动完成链接编辑！

如果你指定的文件不是共享目标文件，后端将会挂起！

## 46.3. SunOS 4.x、Solaris 2.x 和 HP-UX

在 SunOS 4.x、Solaris 2.x 和 HP-UX 下，必须用特殊的编译器标志编译源文件来创建简单的目标文件，并且必须产生一个共享库。HP-UX 所需的步骤如下。HP-UX C 编译器的 `+z` 标志产生所谓的“位置无关代码”（PIC），`+u` 标志移除 PA-RISC 体系结构通常会强制施加的某些对齐限制。必须使用 HP-UX 链接编辑器并带 `-b` 选项把目标文件变成共享库。这听起来复杂，其实非常简单，因为相应的命令就是：

```
# simple HP-UX example
% cc +z +u -c foo.c
% ld -b -o foo.sl foo.o
```

与上一小节提到的 `.so` 文件一样，必须告诉 `create function` 命令哪个是要装载的正确文件（即必须给出共享库即 `.sl` 文件的位置）。在 SunOS 4.x 下，命令看起来像这样：

```
# simple SunOS 4.x example
% cc -PIC -c foo.c
% ld -dc -dp -Bdynamic -o foo.so foo.o
```

在 Solaris 2.x 下等价的命令行是：

```
# simple Solaris 2.x example
% cc -K PIC -c foo.c
% ld -G -Bdynamic -o foo.so foo.o
```

或者

```
# simple Solaris 2.x example
% gcc -fPIC -c foo.c
% ld -G -Bdynamic -o foo.so foo.o
```

链接共享库时，你可能必须在 ld 命令行上指定一些额外的共享库（通常是系统库，例如 C 库和数学库）。

---

## 第 47 章 触发器

Postgres 有多种客户端接口，例如 Perl、Tcl、Python 和 C，还有两种过程语言（PL）。也可以把 C 函数作为触发器动作调用。注意，当前版本不支持语句（STATEMENT）级触发器事件。目前可以把 BEFORE 或 AFTER 指定在一个元组的 INSERT、DELETE 或 UPDATE 上作为触发器事件。

### 47.1. 触发器创建

一旦触发器事件发生，触发器管理器（由执行器调用）就会初始化全局结构 `TriggerData` \*`CurrentTriggerData`（下文描述），并调用触发器函数来处理该事件。

必须先创建触发器函数，然后才能创建触发器，它声明为不接受参数并返回 opaque 的函数。

创建触发器的语法如下：

```
CREATE TRIGGER <trigger name> <BEFORE|AFTER> <INSERT|DELETE|UPDATE>
ON <relation name> FOR EACH <ROW|STATEMENT>
EXECUTE PROCEDURE <procedure name> (<function args>);
```

触发器的名称在你将来需要删除该触发器时使用。它被用作 DROP TRIGGER 命令的一个参数。

下一个词决定函数是在事件之前还是之后被调用。

命令的下一个元素决定该函数将在哪些事件上触发。可以用 OR 分隔指定多个事件。

关系名决定该事件作用于哪个表。

FOR EACH 语句决定触发器是针对每个受影响的行触发，还是在整条语句完成之前（或之后）触发。

过程名是被调用的 C 函数。

参数通过 `CurrentTriggerData` 结构传给函数。向函数传递参数的目的，是为了让需求相似的不同触发器能够调用同一个函数。

另外，函数也可以用于触发不同的关系（这类函数被称为“通用触发器函数”）。

作为同时使用上述两个特性的例子，可以有一个通用函数，它接受两个字段名作为参数，把当前用户写入其中一个字段，把当前时间戳写入另一个字段。这样就可以在 INSERT 事件上编写触发器，例如自动跟踪某个事务表中记录的创建。如果用在 UPDATE 事件上，它还可以作为一个“最近更新”函数使用。

触发器函数向调用它的执行器返回 `HeapTuple`。对于在 INSERT、DELETE 或 UPDATE 操作之后触发的触发器，该返回值会被忽略，但它允许 BEFORE 触发器：- 返回 NULL，以跳过对当前元组的操作（这样该元组就不会被插入/更新/删除）；- 返回一个指向另一个元组的指针（仅 INSERT 和 UPDATE），该元组将被插入（如果是 UPDATE，则作为被更新元组的新版本），以取代原来的元组。

注意，CREATE TRIGGER 处理器不执行任何初始化。这一点将来会改变。另外，如果为同一关系上的同一事件定义了多个触发器，触发的顺序是不可预测的。这一点将来也可能会改变。

如果触发器函数执行 SQL 查询（使用 SPI），那么这些查询可能会再次触发触发器。这就是所谓的级联触发器。对级联层数没有明确的限制。

如果一个触发器由 INSERT 触发，而又向同一个关系插入了一个新元组，那么该触发器会再次被触发。目前，对于这类情形没有提供任何同步（等）机制，但这一点将来可能改变。目前，回归测试中的函数 funny\_dup17() 使用了一些技巧来停止对自身的递归（级联）.....

## 47.2. 与触发器管理器的交互

如上所述，当函数被触发器管理器调用时，结构 `TriggerData *CurrentTriggerData` 非 NULL 且已初始化。因此，最好在开始时检查 `CurrentTriggerData` 是否为 NULL，并在取回信息之后立即把它设为 NULL，以防止并非来自触发器管理器的对触发器函数的调用。

`struct TriggerData` 定义在 `src/include/commands/trigger.h` 中：

```
typedef struct TriggerData
{
    TriggerEvent tg_event;
    Relation tg_relation;
    HeapTuple tg_trigtuple;
    HeapTuple tg_newtuple;
    Trigger *tg_trigger;
} TriggerData;
```

`tg_event`

描述引发该函数调用的事件。你可以使用下列宏来检查 `tg_event`：

```
TRIGGER_FIRED_BEFORE(event) 如果触发器在之前 (BEFORE) 触发则返回 TRUE;
TRIGGER_FIRED_AFTER(event) 如果触发器在之后 (AFTER) 触发则返回 TRUE;
TRIGGER_FIRED_FOR_ROW(event) 如果触发器针对行 (ROW) 级事件触发则返回
TRUE;
TRIGGER_FIRED_FOR_STATEMENT(event) 如果触发器针对语句 (STATEMENT) 级事件
触发则返回 TRUE;
TRIGGER_FIRED_BY_INSERT(event) 如果触发器由 INSERT 引发则返回 TRUE;
TRIGGER_FIRED_BY_DELETE(event) 如果触发器由 DELETE 引发则返回 TRUE;
TRIGGER_FIRED_BY_UPDATE(event) 如果触发器由 UPDATE 引发则返回 TRUE。
```

`tg_relation`

是一个指针，指向描述被触发关系的结构体。关于这个结构体的细节请参看 `src/include/utils/rel.h`。其中最令人感兴趣的是 `tg_relation->rd_att`（关系元组的描述符）和 `tg_relation->rd_rel->relname`（关系名。它不是 `char*`，而是 `NameData`。如果你需要名称的一个副本，可以使用 `SPI_getrelname(tg_relation)` 来获得 `char*`）。

`tg_trigtuple`

是一个指针，指向引发该触发器的元组。这就是正在被插入（INSERT）、删除（DELETE）或更新（UPDATE）的元组。如果是 INSERT/DELETE，那么当你不想（INSERT 的情形下）用另一个元组替换该元组或跳过该操作时，就应当把它返回给执行器。

`tg_newtuple`

是一个指针，指向元组的新版本（如果 UPDATE）；如果是 INSERT 或 DELETE 则为 NULL。当事件是 UPDATE 而你不想用另一个元组替换该元组或跳过该操作时，就应当把它返回给执行器。

`tg_trigger`

是一个指针，指向 Trigger 结构，它定义在 src/include/utils/rel.h 中：

```
typedef struct Trigger
{
    char *tgname;
    Oid tgfoid;
    func_ptr tgfunc;
    int16 tgtype;
    int16 tgnargs;
    int16 tgattr[8];
    char **tgargs;
} Trigger;
```

tgname 是触发器的名称，tgnargs 是 tgargs 中参数的数量，tgargs 是一个指针数组，指向 CREATE TRIGGER 语句中指定的参数。其他成员仅供内部使用。

## 47.3. 数据更改的可见性

Postgres 的数据更改可见性规则：在一次查询执行期间，该查询自身（通过 SQL 函数、SPI 函数、触发器）所做的数据更改对查询扫描是不可见的。例如，在查询

```
INSERT INTO a SELECT * FROM a
```

中，被插入的元组对 SELECT 扫描不可见。其效果是在数据库表内部复制该表自身（当然要受唯一索引规则的约束），而不会递归。

但请记住 SPI 文档中关于可见性的这段说明：

查询 Q 所做的更改，对所有在查询 Q 之后启动的查询都可见，无论这些查询是在 Q 内部（即 Q 执行期间）启动，还是在 Q 完成之后启动。

这对触发器同样成立，因此，虽然正在被插入的元组 (tg\_trigtuple) 对 BEFORE 触发器中的查询不可见，但这个元组（刚刚插入的）对 AFTER 触发器中的查询可见，也对在此之后触发的 BEFORE/AFTER 触发器中的查询可见！

## 47.4. 示例

更多更复杂的示例见 src/test/regress/regress.c 和 contrib/spi。

下面是一个非常简单的触发器用法示例。函数 trigf 报告被触发关系 ttest 中的元组数，并且当查询试图向 x 插入 NULL 时跳过该操作（也就是说，它起到一个 NOT NULL 约束的作用，但不会中止事务）：

```
#include "executor/spi.h" /* this is what you need to work with SPI */
#include "commands/trigger.h" /* --- and triggers */

HeapTuple trigf(void);

HeapTuple
trigf()
{
    TupleDesc tupdesc;
    HeapTuple rettup;
```

```
char *when;
bool checknull = false;
bool isnull;
int ret, i;

if (!CurrentTriggerData)
    elog(WARN, "trigf: triggers are not initialized");

/* tuple to return to Executor */
if (TRIGGER_FIRED_BY_UPDATE(CurrentTriggerData->tg_event))
    rettupple = CurrentTriggerData->tg_newtuple;
else
    rettupple = CurrentTriggerData->tg_trigtuple;

/* check for NULLs ? */
if (!TRIGGER_FIRED_BY_DELETE(CurrentTriggerData->tg_event) &&
    TRIGGER_FIRED_BEFORE(CurrentTriggerData->tg_event))
    checknull = true;

if (TRIGGER_FIRED_BEFORE(CurrentTriggerData->tg_event))
    when = "before";
else
    when = "after ";

tupdesc = CurrentTriggerData->tg_relation->rd_att;
CurrentTriggerData = NULL;

/* Connect to SPI manager */
if ((ret = SPI_connect()) < 0)
    elog(WARN, "trigf (fired %s): SPI_connect returned %d", when, ret);

/* Get number of tuples in relation */
ret = SPI_exec("select count(*) from ttest", 0);

if (ret < 0)
    elog(WARN, "trigf (fired %s): SPI_exec returned %d", when, ret);

i = SPI_getbinval(SPI_tuptable->vals[0], SPI_tuptable->tupdesc, 1,
    &isnull);

elog (NOTICE, "trigf (fired %s): there are %d tuples in ttest", when,
i);

SPI_finish();

if (checknull)
{
    i = SPI_getbinval(rettuple, tupdesc, 1, &isnull);
    if (isnull)
        rettupple = NULL;
}

return (rettupple);
}
```

现在, 编译并: create table ttest (x int4); create function trigf () returns opaque as '...path\_to\_so' language 'c';

```
vac=> create trigger tbefore before insert or update or delete on
      ttest
      for each row execute procedure trigf();
CREATE
vac=> create trigger tafter after insert or update or delete on ttest
      for each row execute procedure trigf();
CREATE
vac=> insert into ttest values (null);
NOTICE:trigf (fired before): there are 0 tuples in ttest
INSERT 0 0
```

-- Insertion skipped and AFTER trigger is not fired

```
vac=> select * from ttest;
x
-
(0 rows)
```

```
vac=> insert into ttest values (1);
NOTICE:trigf (fired before): there are 0 tuples in ttest
NOTICE:trigf (fired after ): there are 1 tuples in ttest
      ^^^^^^^^
```

remember what we said about visibility.

```
INSERT 167793 1
vac=> select * from ttest;
x
-
1
(1 row)
```

```
vac=> insert into ttest select x * 2 from ttest;
NOTICE:trigf (fired before): there are 1 tuples in ttest
NOTICE:trigf (fired after ): there are 2 tuples in ttest
      ^^^^^^^^
```

remember what we said about visibility.

```
INSERT 167794 1
vac=> select * from ttest;
x
-
1
2
(2 rows)
```

```
vac=> update ttest set x = null where x = 2;
NOTICE:trigf (fired before): there are 2 tuples in ttest
UPDATE 0
vac=> update ttest set x = 4 where x = 2;
NOTICE:trigf (fired before): there are 2 tuples in ttest
NOTICE:trigf (fired after ): there are 2 tuples in ttest
UPDATE 1
vac=> select * from ttest;
```

```
x
-
1
4
(2 rows)
```

```
vac=> delete from ttest;
NOTICE:trigf (fired before): there are 2 tuples in ttest
NOTICE:trigf (fired after ): there are 1 tuples in ttest
NOTICE:trigf (fired before): there are 1 tuples in ttest
NOTICE:trigf (fired after ): there are 0 tuples in ttest
                        ^^^^^^^^
```

remember what we said about visibility.

```
DELETE 2
vac=> select * from ttest;
x
-
(0 rows)
```

---

# 第 48 章 服务器编程接口

Vadim Mikheev

服务器编程接口 (SPI) 使用户能够在用户定义的 C 函数中运行 SQL 查询。可用的过程语言 (PL) 提供了访问这些能力的另一种方法。

实际上, SPI 只是一组简化对解析器、规划器、优化器和执行器访问的本地接口函数。SPI 还做一部分内存管理工作。

为避免误解, 我们用函数指 SPI 接口函数, 而用过程指使用 SPI 的用户定义 C 函数。

SPI 过程总是由某个 (上层) 执行器调用, 而 SPI 管理器使用执行器来运行你的查询。执行器在运行来自你的过程的查询时, 还可能调用其他过程。

注意, 如果在执行来自过程的查询期间事务被中止, 控制将不会返回到你的过程。相反, 所有工作都会被回滚, 服务器会等待客户端的下一条命令。这一点在未来的版本中会被改变。

其他限制是无法执行 BEGIN、END 和 ABORT (事务控制语句) 以及游标操作。这一点将来也会被改变。

成功时, SPI 函数返回非负结果 (或者通过整型返回值返回, 或者如下文所述存放在全局变量 SPI\_result 中)。出错时, 将返回负值或 NULL。

## 48.1. 接口函数

# SPI\_connect

SPI\_connect — 将你的过程连接到 SPI 管理器。

## 大纲

```
int SPI_connect(void)
```

### 输入

无

### 输出

int

返回状态

→ SPI\_OK\_CONNECT

已连接时

→ SPI\_ERROR\_CONNECT

未连接时

## 描述

SPI\_connect 打开到 Postgres 后端的连接。如果需要执行查询，就应调用此函数。有些 SPI 辅助函数可以在未连接的过程中调用。

如果 SPI\_connect 是从已连接的过程中调用的，就可能得到 → SPI\_ERROR\_CONNECT 错误——例如，你从一个已连接的过程直接调用另一个过程。实际上，虽然子过程将能够使用 SPI，但在子过程返回之后（如果子过程调用了 SPI\_finish），你的父过程将无法继续使用 SPI。这是一种不好的做法。

## 用法

XXX thomas 1997-12-24

## 算法

SPI\_connect 执行以下操作：

- 

初始化用于查询执行和内存管理的 SPI 内部结构。

# SPI\_finish

SPI\_finish — 将你的过程与 SPI 管理器断开连接。

## 大纲

```
SPI_finish(void)
```

### 输入

无

### 输出

int

- SPI\_OK\_FINISH 已正确断开连接时
- SPI\_ERROR\_UNCONNECTED 从未连接的过程中调用时

## 描述

SPI\_finish 关闭到 Postgres 后端的现有连接。在完成通过 SPI 管理器所做的操作之后，就应调用此函数。

如果在当前没有有效连接的情况下调用 SPI\_finish，可能会得到错误返回 → SPI\_ERROR\_UNCONNECTED。这并不是什么根本性的问题；它只表示 SPI 管理器没有做任何事情。

## 用法

已连接的过程必须以调用 SPI\_finish 作为最后一步，否则可能得到不可预测的结果！注意，如果你中止了事务（通过 elog(ERROR)），则可以安全地跳过对 SPI\_finish 的调用。

## 算法

SPI\_finish 执行以下操作：

- 

将你的过程与 SPI 管理器断开连接，并释放该过程自 SPI\_connect 以来通过 palloc 所做的全部内存分配。这些分配不能再使用了！参见内存管理。

## SPI\_exec

SPI\_exec — 创建一个执行计划（解析器+规划器+优化器）并执行查询。

### 大纲

```
SPI_exec(query, tcount)
```

#### 输入

```
char *query
```

包含查询计划的字符串

```
int tcount
```

要返回的最大元组数

#### 输出

```
int
```

- SPI\_OK\_EXEC 已正确断开连接时
- SPI\_ERROR\_UNCONNECTED 从未连接的过程中调用时
- SPI\_ERROR\_ARGUMENT query 为 NULL 或 tcount < 0 时。
- SPI\_ERROR\_UNCONNECTED 过程未连接时。
- SPI\_ERROR\_COPY 执行 COPY TO/FROM stdin 时。
- SPI\_ERROR\_CURSOR 执行 DECLARE/CLOSE CURSOR、FETCH 时。
- SPI\_ERROR\_TRANSACTION 执行 BEGIN/ABORT/END 时。
- SPI\_ERROR\_OPUNKNOWN 查询类型未知时（这不应发生）。

如果查询执行成功，则会返回下列（非负）值之一：

- SPI\_OKUTILITY 执行了某个工具命令（例如 CREATE TABLE ...）时
- SPI\_OK\_SELECT 执行了 SELECT（但不是 SELECT ... INTO!）时
- SPI\_OK\_SELINTO 执行了 SELECT ... INTO 时
- SPI\_OK\_INSERT 执行了 INSERT（或 INSERT ... SELECT）时
- SPI\_OK\_DELETE 执行了 DELETE 时
- SPI\_OK\_UPDATE 执行了 UPDATE 时

### 描述

SPI\_exec 创建一个执行计划（解析器+规划器+优化器），并针对 tcount 个元组执行该查询。

### 用法

此函数只应从已连接的过程中调用。如果 tcount 为零，则会对查询扫描返回的所有元组执行该查询。使用 tcount > 0 可以限制该查询将要执行的元组数。例如，

```
SPI_exec ("insert into table select * from table", 5);
```

最多只允许向表中插入 5 个元组。如果查询执行成功，则会返回一个非负值。

### 注意

你可以在一个字符串中传递多条查询，查询字符串也可能被 `RULE` 改写。`SPI_exec` 返回最后执行的那条查询的结果。

（最后一条）查询实际执行所处理的元组数，会通过全局变量 `SPI_processed` 返回（返回值不是 `→ SPI_OK_UTILITY` 时）。如果返回了 `→ SPI_OK_SELECT` 且 `SPI_processed > 0`，则可以使用全局指针 `SPItupleTable *SPI_tuptable` 访问选出的元组：另外注意，`SPI_finish` 会释放所有 `SPItupleTable` 并使它们不可再用！（参见内存管理）。

`SPI_exec` 可能返回下列（负）值之一：

- `SPI_ERROR_ARGUMENT` `query` 为 `NULL` 或 `tcount < 0` 时。
- `SPI_ERROR_UNCONNECTED` 过程未连接时。
- `SPI_ERROR_COPY` 执行 `COPY TO/FROM stdin` 时。
- `SPI_ERROR_CURSOR` 执行 `DECLARE/CLOSE CURSOR`、`FETCH` 时。
- `SPI_ERROR_TRANSACTION` 执行 `BEGIN/ABORT/END` 时。
- `SPI_ERROR_OPUNKNOWN` 查询类型未知时（这不应发生）。

## 算法

`SPI_exec` 执行以下操作：

•

将你的过程与 `SPI` 管理器断开连接，并释放该过程自 `SPI_connect` 以来通过 `palloc` 所做的全部内存分配。这些分配不能再使用了！参见内存管理。

## SPI\_prepare

SPI\_prepare — 将你的过程连接到 SPI 管理器。

### 大纲

```
SPI_prepare(query, nargs, argtypes)
```

#### 输入

*query*

查询字符串

*nargs*

输入参数的数量 (\$1 ... \$nargs -- 与 SQL 函数中相同)

*argtypes*

指向输入参数的类型 OID 列表的指针

#### 输出

void \*

指向执行计划（解析器+规划器+优化器）的指针

### 描述

SPI\_prepare 创建并返回一个执行计划（解析器+规划器+优化器），但不执行该查询。只应从已连接的过程中调用。

### 用法

nargs 是参数的数量 (\$1 ... \$nargs -- 与 SQL 函数中相同)，只有当查询中没有任何 \$1 时，nargs 才可以为 0。

预备执行计划的执行有时会快得多，因此当同一条查询要执行很多次时，此特性可能会很有用。

SPI\_prepare 返回的计划只能在当前这次过程调用中使用，因为 SPI\_finish 会释放为计划分配的内存。参见 SPI\_saveplan。

成功时会返回一个非空指针。否则，将得到 NULL 计划。无论成功还是出错，SPI\_result 都会被设成与 SPI\_exec 返回值类似的值；但如果 query 为 NULL，或 nargs < 0，或 nargs > 0 且 argtypes 为 NULL，则会将其设为 → SPI\_ERROR\_ARGUMENT。

# SPI\_saveplan

SPI\_saveplan — 保存传入的计划

## 大纲

```
SPI_saveplan(plan)
```

## 输入

```
void *query
```

传入的计划

## 输出

```
void *
```

执行计划的位置。不成功时为 NULL。

SPI\_result

→ SPI\_ERROR\_ARGUMENT *plan* 为 NULL 时

→ SPI\_ERROR\_UNCONNECTED 过程未连接时

## 描述

SPI\_saveplan 把由 SPI\_prepare 准备的计划存储在安全内存中，使其不会被 SPI\_finish 或事务管理器释放。

在当前版本的 Postgres 中，还无法把预备计划存储在系统目录中并在执行时从那里取出。这一点将在未来的版本中实现。作为替代，可以在当前会话中该过程的后续调用里复用预备计划。使用 SPI\_execp 执行这个保存的计划。

## 用法

SPI\_saveplan 把传入的计划（由 SPI\_prepare 准备）保存在受保护的内存中，使其不会被 SPI\_finish 和事务管理器释放，并返回指向该已保存计划的指针。可以把返回的指针保存在局部变量中。无论是准备计划时，还是在 SPI\_execp 中使用已准备好的计划时（见下文），都请始终检查该指针是否为 NULL。

### 注意

如果预备计划所引用的对象（关系、函数等）在会话期间被删除（由你的后端或另一个进程删除），那么对该计划执行 SPI\_execp 的结果将是不可预知的。

## SPI\_execp

SPI\_execp — 执行来自 SPI\_saveplan 的计划

### 大纲

```
SPI_execp(plan,  
          values,  
          nulls,  
          tcount)
```

### 输入

`void *plan`

执行计划

`Datum *values`

实际参数值

`char *nulls`

描述哪些参数取 NULL 的数组

'n' 表示允许 NULL

' ' 表示不允许 NULL

`int tcount`

该计划将要执行的元组数

### 输出

`int`

返回与 SPI\_exec 相同的值，以及

→ SPI\_ERROR\_ARGUMENT`plan` 为 NULL 或 `tcount < 0` 时

→ SPI\_ERROR\_PARAM`values` 为 NULL，而 `plan` 在准备时带有参数时。

`SPI_tuptable`

成功时，其初始化方式与 SPI\_exec 中相同

`SPI_processed`

成功时，其初始化方式与 SPI\_exec 中相同

### 描述

SPI\_execp 把由 SPI\_prepare 准备的计划存储在安全内存中，使其不会被 SPI\_finish 或事务管理器释放。

在当前版本的 Postgres 中，还无法把预备计划存储在系统目录中并在执行时从那里取出。这一点将在未来的版本中实现。作为权宜之计，可以在当前会话中该过程的后续调用里复用预备计划。使用 `SPI_execp` 执行这个保存的计划。

## 用法

如果 `nulls` 为 `NULL`，则 `SPI_execp` 假定所有值（如果有的话）都非 `NULL`。

### 注意

如果预备计划所引用的对象（关系、函数等）在会话期间被删除（由你的后端或另一个进程删除），那么对该计划执行 `SPI_execp` 的结果将是不可预知的。

## 48.2. 接口支持函数

下面介绍的所有函数既可由已连接的过程使用，也可由未连接的过程使用。

## SPI\_copytuple

SPI\_copytuple — 在上层执行器上下文中创建元组的副本

### 大纲

```
SPI_copytuple(tuple)
```

### 输入

HeapTuple *tuple*

要复制的输入元组

### 输出

HeapTuple

复制得到的元组

→ 非 NULL *tuple* 不为 NULL 且复制成功时

→ NULL 仅当 *tuple* 为 NULL 时

### 描述

SPI\_copytuple 在上层执行器上下文中创建元组的一个副本。参见内存管理一节。

### 用法

TBD

# SPI\_modifytuple

SPI\_modifytuple — 修改关系的元组

## 大纲

```
SPI_modifytuple(rel, tuple , nattrs
, attnum , Values , Nulls)
```

## 输入

Relation *rel*

HeapTuple *tuple*

要修改的输入元组

int *nattrs*

*attnum* 中属性编号的数量

int \* *attnum*

要被修改的属性的编号组成的数组

Datum \* *Values*

指定属性的新值

char \* *Nulls*

哪些属性为 NULL（如果有的话）

## 输出

HeapTuple

带修改内容的新元组

→ 非 NULL *tuple* 不为 NULL 且修改成功时

→ NULL 仅当 *tuple* 为 NULL 时

SPI\_result

→ SPI\_ERROR\_ARGUMENT *rel* 为 NULL, 或 *tuple* 为 NULL, 或 *natts* ≤ 0, 或 *attnum* 为 NULL, 或 *Values* 为 NULL 时。

→ SPI\_ERROR\_NOATTRIBUTE *attnum* 中存在无效的属性编号时 (*attnum* ≤ 0, 或大于元组中的属性数量)

## 描述

SPI\_modifytuple 在上层执行器上下文中修改元组。参见内存管理一节。

## 用法

成功时，返回指向新元组的指针。新元组在上层执行器上下文中分配（参见内存管理）。传入的元组不会被修改。

# SPI\_fnumber

SPI\_fnumber — 查找指定属性的属性编号

## 大纲

```
SPI_fnumber(tupdesc, fname)
```

## 输入

TupleDesc *tupdesc*

输入元组描述

char \* *fname*

字段名

## 输出

int

属性编号

属性的有效基于 1 的索引编号

→ SPI\_ERROR\_NOATTRIBUTE 找不到指定名称的属性时

## 描述

SPI\_fnumber 返回 *fname* 中指定名称的属性的属性编号。

## 用法

属性编号从 1 开始。

## SPI\_fname

SPI\_fname — 查找指定属性的属性名称

### 大纲

```
SPI_fname(tupdesc, fname)
```

### 输入

TupleDesc *tupdesc*

输入元组描述

char \* *fnumber*

属性编号

### 输出

char \*

属性名称

*fnumber* 超出范围时为 NULL

出错时 SPI\_result 被设置为  $\rightarrow$  SPI\_ERROR\_NOATTRIBUTE

### 描述

SPI\_fname 返回指定属性的属性名称。

### 用法

属性编号从 1 开始。

### 算法

返回一个新分配的属性名称副本。

# SPI\_getvalue

SPI\_getvalue — 返回指定属性的字符串值

## 大纲

```
SPI_getvalue(tuple, tupdesc, fnumber)
```

## 输入

HeapTuple *tuple*

要检查的输入元组

TupleDesc *tupdesc*

输入元组描述

int *fnumber*

属性编号

## 输出

char \*

属性值，以下情况返回 NULL：

属性为 NULL

*fnumber* 超出范围（SPI\_result 被设置为  $\rightarrow$  SPI\_ERROR\_NOATTRIBUTE）

没有可用的输出函数（SPI\_result 被设置为  $\rightarrow$  SPI\_ERROR\_NOOUTFUNC）

## 描述

SPI\_getvalue 返回指定属性值的外部（字符串）表示形式。

## 用法

属性编号从 1 开始。

## 算法

根据该值的需要分配内存。

# SPI\_getbinval

SPI\_getbinval — 返回指定属性的二进制值

## 大纲

```
SPI_getbinval(tuple, tupdesc, fnumber, isnull)
```

## 输入

HeapTuple *tuple*

要检查的输入元组

TupleDesc *tupdesc*

输入元组描述

int *fnumber*

属性编号

## 输出

Datum

属性的二元值

bool \* *isnull*

属性值是否为空的标志

SPI\_result

→ SPI\_ERROR\_NOATTRIBUTE

## 描述

SPI\_getbinval 返回指定属性的二进制值。

## 用法

属性编号从 1 开始。

## 算法

不为该二进制值分配新的空间。

# SPI\_gettype

SPI\_gettype — 返回指定属性的类型名称

## 大纲

```
SPI_gettype(tupdesc, fnumber)
```

### 输入

TupleDesc *tupdesc*

输入元组描述

int *fnumber*

属性编号

### 输出

char \*

指定属性编号的类型名称

SPI\_result

→ SPI\_ERROR\_NOATTRIBUTE

## 描述

SPI\_gettype 返回指定属性类型名称的一个副本。

## 用法

属性编号从 1 开始。

## 算法

不为该二进制值分配新的空间。

# SPI\_gettypeid

SPI\_gettypeid — 返回指定属性的类型 OID

## 大纲

```
SPI_gettypeid(tupdesc, fnumber)
```

## 输入

TupleDesc *tupdesc*

输入元组描述

int *fnumber*

属性编号

## 输出

OID

指定属性编号的类型 OID

SPI\_result

→ SPI\_ERROR\_NOATTRIBUTE

## 描述

SPI\_gettypeid 返回指定属性的类型 OID。

## 用法

属性编号从 1 开始。

## 算法

TBD

# SPI\_getrelname

SPI\_getrelname — 返回指定关系的名称

## 大纲

```
SPI_getrelname(rel)
```

## 输入

Relation *rel*

输入关系

## 输出

char \*

指定关系的名称

## 描述

SPI\_getrelname 返回指定关系的名称。

## 用法

TBD

## 算法

把该关系名称复制到新的存储中。

# SPI\_palloc

SPI\_palloc — 在上层执行器上下文中分配内存

## 大纲

```
SPI_palloc(size)
```

## 输入

Size *size*

要分配的存储空间的字节大小

## 输出

void \*

指定大小的新存储空间

## 描述

SPI\_palloc 在上层执行器上下文中分配内存。参见内存管理一节。

## 用法

TBD

# SPI\_realloc

SPI\_realloc — 在上层执行器上下文中重新分配内存

## 大纲

```
SPI_realloc(pointer, size)
```

## 输入

`void *pointer`

指向现有存储的指针

`Size size`

要分配的存储空间的字节大小

## 输出

`void *`

指定大小的新存储空间，内容从现有区域复制而来

## 描述

`SPI_realloc` 在上层执行器上下文中重新分配内存。参见内存管理一节。

## 用法

TBD

## SPI\_pfree

SPI\_pfree — 从上层执行器上下文中释放内存

### 大纲

```
SPI_pfree(pointer)
```

#### 输入

```
void * pointer
```

指向现有存储的指针

#### 输出

无

### 描述

SPI\_pfree 在上层执行器上下文中释放内存。参见内存管理一节。

### 用法

TBD

## 48.3. 内存管理

服务器在内存上下文中分配内存，使得在一个上下文中做出的分配可以通过销毁该上下文来释放，而不影响在其他上下文中做出的分配。所有分配（通过 `palloc` 等）都是在被选为当前上下文的那个上下文中进行的。如果试图释放（或重新分配）不是在当前上下文中分配的内存，将会得到不可预知的结果。

内存上下文的创建和切换属于 SPI 管理器内存管理的内容。

SPI 过程涉及两个内存上下文：上层执行器内存上下文和过程内存上下文（如果已连接）。

在过程连接到 SPI 管理器之前，当前内存上下文是上层执行器上下文，因此该过程自身在连接到 SPI 之前通过 `palloc/repalloc` 或由 SPI 辅助函数所做的所有分配都位于这个上下文中。

调用 `SPI_connect` 之后，当前上下文是过程的私有上下文。通过 `palloc/repalloc` 或由 SPI 辅助函数所做的所有分配（`SPI_cpytuple`、`SPI_modifytuple`、`SPI_palloc` 和 `SPI_repalloc` 除外）都位于这个上下文中。

当过程（通过 `SPI_finish`）与 SPI 管理器断开连接时，当前上下文会恢复为上层执行器上下文，而在该过程内存上下文中分配的所有内存都会被释放，不能再使用！

如果想把某些东西返回给上层执行器，就必须在上层上下文中为它分配内存！

SPI 无法自动释放上层执行器上下文中的分配！

查询执行期间分配的内存会在该查询完成时被 SPI 自动释放！

## 48.4. 数据更改的可见性

Postgres 的数据更改可见性规则：在一次查询执行期间，该查询自身（通过 SQL 函数、SPI 函数、触发器）所做的数据更改对查询扫描是不可见的。例如，在查询 `INSERT INTO a SELECT * FROM a` 中，被插入的元组对 `SELECT` 的扫描不可见。其效果是在数据库表内部复制该表自身（当然要受唯一索引规则的约束），而不会递归。

查询 Q 所做的更改，对所有在查询 Q 之后启动的查询都可见，无论这些查询是在 Q 内部（即 Q 执行期间）启动，还是在 Q 完成之后启动。

## 48.5. 示例

这个 SPI 用法的示例演示了可见性规则。在 `src/test/regress/regress.c` 和 `contrib/spi` 中还有更复杂的示例。

这是一个非常简单的 SPI 用法示例。过程 `execq` 的第一个参数接受一条 SQL 查询，第二个参数接受 `tcount`，它使用 `SPI_exec` 执行该查询，并返回该查询所处理的元组数：

```
#include "executor/spi.h" /* this is what you need to work with SPI */

int execq(text *sql, int cnt);

int
execq(text *sql, int cnt)
{
    int ret;
    int proc = 0;

    SPI_connect();

    ret = SPI_exec(textout(sql), cnt);

    proc = SPI_processed;
    /*
     * If this is SELECT and some tuple(s) fetched -
     * returns tuples to the caller via elog (NOTICE).
     */
    if (ret == SPI_OK_SELECT && SPI_processed > 0)
    {
        TupleDesc tupdesc = SPI_tuptable->tupdesc;
        SPITupleTable *tuptable = SPI_tuptable;
        char buf[8192];
        int i;

        for (ret = 0; ret < proc; ret++)
        {
            HeapTuple tuple = tuptable->vals[ret];

            for (i = 1, buf[0] = 0; i <= tupdesc->natts; i++)
                sprintf(buf + strlen(buf), " %s",
                    SPI_getvalue(tuple, tupdesc, i),
                    (i == tupdesc->natts) ? " " : " |");
            elog (NOTICE, "EXECQ: %s", buf);
        }
    }
}
```

```

    }
}

SPI_finish();

return (proc);
}

```

现在，编译并创建该函数：

```
create function execq (text, int4) returns int4 as '...path_to_so'
language 'c';
```

```
vac=> select execq('create table a (x int4)', 0);
execq
-----
      0
(1 row)
```

```
vac=> insert into a values (execq('insert into a values (0)',0));
INSERT 167631 1
vac=> select execq('select * from a',0);
NOTICE:EXECQ:  0 <<< 由 execq 插入
```

```
NOTICE:EXECQ:  1 <<< 由 execq 返回并被外层 INSERT 插入的值
```

```
execq
-----
      2
(1 row)
```

```
vac=> select execq('insert into a select x + 2 from a',1);
execq
-----
      1
(1 row)
```

```
vac=> select execq('select * from a', 10);
NOTICE:EXECQ:  0
```

```
NOTICE:EXECQ:  1
```

```
NOTICE:EXECQ:  2 <<< 0 + 2, 只插入了一个元组 — 正如所指定的那样
```

```
execq
-----
      3          <<< 10 只是最大值, 3 才是实际的元组数
(1 row)
```

```
vac=> delete from a;
DELETE 3
vac=> insert into a values (execq('select * from a', 0) + 1);
INSERT 167712 1
vac=> select * from a;
```

```

x
-
1          <<< a 中没有元组 (0) + 1
(1 row)

vac=> insert into a values (execq('select * from a', 0) + 1);
NOTICE:EXECQ: 0
INSERT 167713 1
vac=> select * from a;
x
-
1
2          <<< a 中原先只有一个元组 + 1
(2 rows)

-- 这演示了数据更改可见性规则:

vac=> insert into a select execq('select * from a', 0) * x from a;
NOTICE:EXECQ: 1
NOTICE:EXECQ: 2
NOTICE:EXECQ: 1
NOTICE:EXECQ: 2
NOTICE:EXECQ: 2
INSERT 0 2
vac=> select * from a;
x
-
1
2
2          <<< 2 个元组 * 1 (第一个元组中的 x)
6          <<< 3 个元组 (刚插入的 2 + 1) * 2 (第二个元组中的 x)
(4 rows)          ^^^^^^^^
                    execq() 在不同调用中可见的元组

```

---

# 第 49 章 过程语言

从 6.3 版发布开始, Postgres 支持定义过程语言。对于用过程语言定义的函数或触发器过程, 数据库并不内置关于如何解释函数源文本的知识。相反, 调用会被传递给一个了解该语言细节的调用处理器。

调用处理器本身是一个特殊的编程语言函数, 被编译进共享对象并按需装载。

## 49.1. 安装过程语言

### 过程语言的安装

在数据库中安装过程语言需要三个步骤。

1. 该语言调用处理器的共享对象必须编译并安装。默认情况下, PL/pgSQL 的调用处理器会被构建并安装到数据库库目录中。如果配置了 Tcl/Tk 支持, PL/Tcl 的调用处理器也会被构建并安装到同一位置。

编写新过程语言 (PL) 的调用处理器超出了本手册的范围。

2. 处理器必须用命令

```
CREATE FUNCTION handler_function_name () RETURNS OPAQUE AS
    'path-to-shared-object' LANGUAGE 'C';
```

声明。特殊的返回类型 OPAQUE 告诉数据库, 这个函数不返回任何已定义的基本类型或组合类型, 不能直接在 SQL 语句中使用。

3. PL 必须用命令

```
CREATE [ TRUSTED ] PROCEDURAL LANGUAGE 'language-name'
    HANDLER handler_function_name
    LANCOMPILER 'description';
```

声明。可选关键字 TRUSTED 指明没有超级用户权限的普通数据库用户能否使用这种语言创建函数和触发器过程。由于 PL 函数在数据库后端内部执行, 它只应用于不会访问数据库后端内部结构或文件系统的语言。PL/pgSQL 和 PL/Tcl 语言已知是可信的。

### 示例

1. 下面的命令告诉数据库到哪里查找 PL/pgSQL 语言调用处理器函数的共享对象:

```
CREATE FUNCTION plpgsql_call_handler () RETURNS OPAQUE AS
    '/usr/local/pgsql/lib/plpgsql.so' LANGUAGE 'C';
```

2. 而命令

```
CREATE TRUSTED PROCEDURAL LANGUAGE 'plpgsql'
    HANDLER plpgsql_call_handler
    LANCOMPILER 'PL/pgSQL';
```

则定义对于语言属性为 'plpgsql' 的函数和触发器过程，将调用先前声明的调用处理器函数。

PL 调用处理器函数有一个与普通 C 语言函数不同的特殊调用接口。传给调用处理器的参数之一是要执行的函数在 `pg_proc` 表条目中的对象 ID。调用处理器会检查各种系统目录来分析函数的调用参数及其返回数据类型。函数主体的源文本位于 `pg_proc` 的 `prosrc` 属性中。因此，与 C 语言函数不同，PL 函数可以像 SQL 语言函数一样被重载。只要调用参数不同，就可以有多个同名的不同 PL 函数。

在 `template1` 数据库中定义的过程语言会自动在随后创建的所有数据库中定义。因此数据库管理员可以决定默认提供哪些语言。

## 49.2. PL/pgSQL

PL/pgSQL 是一个可装载的 Postgres 数据库系统过程语言。

这个软件包最初由 Jan Wieck 编写。

### 49.2.1. 概述

PL/pgSQL 的设计目标是创建一种可装载的过程语言，它

- 可以用于创建函数和触发器过程，
- 为 SQL 语言增加控制结构，
- 可以执行复杂计算，
- 继承所有用户定义的类型、函数和操作符，
- 可以定义为被服务器信任，
- 易于使用。

PL/pgSQL 调用处理器在后端第一次调用函数时解析函数的源文本并产生一棵内部二进制指令树。产生的字节码在调用处理器中通过函数的对象 ID 来标识。这确保了通过 `DROP/CREATE` 序列更改函数无需建立新的数据库连接即可生效。

对于函数中使用的所有表达式和 SQL 语句，PL/pgSQL 字节码解释器都会使用 SPI 管理器的 `SPI_prepare()` 和 `SPI_saveplan()` 函数创建预备执行计划。这是在 PL/pgSQL 函数中第一次处理各条语句时完成的。因此，包含许多需要执行计划的语句的条件代码函数，只会预备和保存那些在数据库连接的整个生命周期内真正用到的计划。

除了用户定义类型的输入/输出转换和计算函数之外，C 语言函数中能定义的任何事情都可以用 PL/pgSQL 完成。可以创建复杂的条件计算函数，之后用它们定义操作符或在函数索引中使用。

### 49.2.2. 描述

#### 49.2.2.1. PL/pgSQL 的结构

PL/pgSQL 语言不区分大小写。所有关键字和标识符都可以混合使用大写和小写。

PL/pgSQL 是一种面向块的语言。块的定义是

```

[<<label>>]
[DECLARE
    declarations]
BEGIN
    statements
END;

```

块的语句部分可以有任意数量的子块。子块可以用来对语句块外部隐藏变量。块的声明部分中声明的变量在每次进入该块时都会被初始化为它们的默认值，而不是每次函数调用只初始化一次。

重要的是不要误解 PL/pgSQL 中用于分组语句的 BEGIN/END 与用于事务控制的数据库命令之间的区别。函数和触发器过程不能启动或提交事务，而且 Postgres 没有嵌套事务。

### 49.2.2.2. 注释

PL/pgSQL 中有两种注释。双横线 '--' 开始一个延伸到行尾的注释。'/\*' 开始一个延伸到下一个 '\*/' 出现处的块注释。块注释不能嵌套，但双横线注释可以放在块注释内，双横线也可以隐藏块注释定界符 '/\*' 和 '\*/'。

### 49.2.2.3. 声明

块或其子块中使用的所有变量、行和记录都必须在块的声明区中声明，唯一例外是在整数范围上迭代的 FOR 循环的循环变量。传给 PL/pgSQL 函数的参数会用通常的标识符 \$n 自动声明。声明使用如下语法：

```
name [ CONSTANT ] type [ NOT NULL ] [ DEFAULT | := value ];
```

声明一个指定基本类型的变量。如果变量声明为 CONSTANT，其值不能更改。如果指定了 NOT NULL，赋 NULL 值将导致运行时错误。由于所有变量的默认值都是 SQL NULL 值，所有声明为 NOT NULL 的变量还必须指定默认值。

默认值在每次调用函数时求值。因此把 'now' 赋给一个 *datetime* 类型的变量，会使该变量拥有实际调用函数时的时间，而不是函数预编译为其字节码时的时间。

```
name class %ROWTYPE;
```

声明一个具有给定类的结构的行。类必须是数据库中已存在的表名或视图名。行的字段用点表示法访问。函数的参数可以是复合类型（完整的表行）。在这种情况下，相应的标识符 \$n 将是一个行类型，但必须用下面描述的 ALIAS 命令为它起别名。行中只能访问表行的用户属性，不能访问 Oid 或其他系统属性（因此该行可以来自视图，而视图行没有用的系统属性）。

行类型的字段继承表中 char() 等数据类型的字段大小或精度。

```
name RECORD;
```

记录类似于行类型，但没有预定义的结构。它们用于选择和 FOR 循环中，保存 SELECT 操作返回的一条实际数据库行。同一个记录可以用于不同的选择。当记录中没有实际行时访问记录或试图给记录字段赋值都会导致运行时错误。

触发器中的 NEW 和 OLD 行会作为记录传给过程。这是必要的，因为在 Postgres 中同一个触发器过程可以处理不同表的触发器事件。

```
name ALIAS FOR $n;
```

为了让代码更易读，可以为函数的位置参数定义别名。

对于作为参数传给函数的复合类型，这种别名是必需的。SQL 函数中的点表示法 `$1.salary` 在 PL/pgSQL 中不允许使用。

```
RENAME oldname TO newname;
```

更改变量、记录或行的名称。当触发器过程内部需要用另一个名称引用 NEW 或 OLD 时，这很有用。

#### 49.2.2.4. 数据类型

变量的类型可以是数据库中任何已存在的基本类型。上面声明区中的 *type* 定义为：

- `Postgres-basetype`
- `variable%TYPE`
- `class.field%TYPE`

*variable* 是在同一函数中先前声明、在此处可见的变量的名称。

*class* 是已存在的表或视图的名称，*field* 是其中某个属性的名称。

使用 `class.field%TYPE` 会使 PL/pgSQL 在后端生命周期内第一次调用函数时查找该属性的定义。假设有一个带 `char(20)` 属性的表和一些在局部变量中处理其内容的 PL/pgSQL 函数。现在有人认为 `char(20)` 不够用，转储表、删除它、把这个属性重新定义为 `char(40)` 再重建表并恢复数据。哈——他忘了那些函数。其中的计算会把值截断为 20 个字符。但如果它们是用 `class.field%TYPE` 声明定义的，就会自动处理大小变化，包括新表模式把该属性定义为 `text` 类型的情况。

#### 49.2.2.5. 表达式

PL/pgSQL 语句中使用的所有表达式都由后端的执行器处理。看起来包含常量的表达式实际上可能需要运行时求值（例如 `datetime` 类型的 `'now'`），因此 PL/pgSQL 解析器不可能识别除 NULL 关键字之外的真常量值。所有表达式都在内部通过执行查询

```
SELECT expression
```

用 SPI 管理器求值。在表达式中，变量标识符的出现被替换为参数，变量的实际值通过参数数组传给执行器。PL/pgSQL 函数中使用的所有表达式都只预备和保存一次。

Postgres 主解析器所做的类型检查对常量值的解释有一些副作用。具体来说，下面两个函数的行为有所不同：

```
CREATE FUNCTION logfunc1 (text) RETURNS datetime AS '
DECLARE
    logtxt ALIAS FOR $1;
BEGIN
    INSERT INTO logtable VALUES (logtxt, 'now');
    RETURN 'now';
END;
' LANGUAGE 'plpgsql';
```

and

```
CREATE FUNCTION logfunc2 (text) RETURNS datetime AS '
DECLARE
    logtxt ALIAS FOR $1;
    curtime datetime;
BEGIN
    curtime := 'now';
    INSERT INTO logtable VALUES (logtxt, curtime);
    RETURN curtime;
END;
' LANGUAGE 'plpgsql';
```

在 logfunc1() 的情况中, Postgres 主解析器在为 INSERT 预备计划时知道字符串 'now' 应该被解释为 datetime, 因为 logtable 的目标字段是那种类型。这样, 它此时就会把它变成一个常量, 而这个常量值随后在后端整个生命周期内的所有 logfunc1() 调用中都被使用。不用说, 这并不是程序员想要的。

在 logfunc2() 的例子中, Postgres 主解析器不知道 'now' 应该变成什么类型, 因此它返回一个包含字符串 'now' 的 text 数据类型。在赋值给局部变量 curtime 时, PL/pgSQL 解释器会调用 text\_out() 和 datetime\_in() 函数把这个字符串转换为 datetime 类型。

Postgres 主解析器所做的这种类型检查是在 PL/pgSQL 几乎完工之后实现的。这是 6.3 与 6.4 之间的一个差异, 影响所有使用 SPI 管理器预备计划特性的函数。目前在 PL/pgSQL 中, 按上述方式使用局部变量是让这些值被正确解释的唯一方法。

如果在表达式或语句中使用记录字段,  
在同一个表达式的多次调用之间字段的数据类型不应改变。  
在编写处理多个表的事件的触发器过程时, 请记住这一点。

## 49.2.2.6. 语句

PL/pgSQL 解析器按下面的说明不理解的任何内容都会被放进一个查询并发给数据库引擎执行。这种查询不应返回任何数据。

赋值

给变量或行/记录字段赋一个值的写法是

```
identifier := expression;
```

如果表达式的结果数据类型与变量的数据类型不匹配, 或者变量有已知的大小/精度 (如 char(20)), 结果值将由 PL/pgSQL 字节码解释器使用结果类型的输出函数和变量类型的输入函数隐式转换。注意, 这有可能导致类型输入函数在运行时产生错误。

把一个完整的选择结果赋给记录或行可以用

```
SELECT expressions INTO target FROM ...;
```

完成。*target* 可以是记录、行变量或逗号分隔的变量和记录/行字段列表。

如果用行或变量列表作为目标, 选出的值必须与目标的结构完全匹配, 否则会发生运行时错误。FROM 关键字后面可以跟 SELECT 语句允许的任何有效的限定、分组、排序等。

有一个名为 FOUND 的 bool 类型特殊变量，可以在 SELECT INTO 之后立即用来检查赋值是否成功。

```
SELECT * INTO myrec FROM EMP WHERE empname = myname;
IF NOT FOUND THEN
    RAISE EXCEPTION 'employee % not found', myname;
END IF;
```

如果选择返回多行，只有第一行会被移入目标字段。其余的都被悄悄丢弃。

调用另一个函数

**Postgres** 数据库中定义的所有函数都返回一个值。因此，调用函数的常规方式是执行 SELECT 查询或做一次赋值（导致一次 PL/pgSQL 内部的 SELECT）。但有时人们并不关心函数的结果。

```
PERFORM query
```

通过 SPI 管理器执行 'SELECT query' 并丢弃结果。局部变量等标识符仍会被替换为参数。

从函数返回

```
RETURN expression
```

函数终止，*expression* 的值将返回给上层执行器。函数的返回值不能未定义。如果控制流到达函数顶层块的末尾仍未遇到 RETURN 语句，将发生运行时错误。

表达式的结果会按赋值一节所述自动转换成函数的返回类型。

中止与消息

如上面的示例所示，有一条 RAISE 语句可以把消息抛进 Postgres 的 elog 机制。

```
RAISE level 'format' [, identifier [...]];
```

在格式串内部，“%”用作后续逗号分隔标识符的占位符。可能的级别有 DEBUG（在生产运行的数据库中被悄悄抑制）、NOTICE（写入数据库日志并转发给客户端应用）和 EXCEPTION（写入数据库日志并中止事务）。

条件

```
IF expression THEN
    statements
[ELSE
    statements]
END IF;
```

*expression* 必须返回一个至少可以转换成布尔类型的值。

循环

循环有多种类型。

```
[<<label>>]
```

```
LOOP
    statements
END LOOP;
```

无条件循环，必须由 EXIT 语句显式终止。可选的标签可以被嵌套循环的 EXIT 语句用来指定应终止哪一层嵌套。

```
[<<label>>]
WHILE expression LOOP
    statements
END LOOP;
```

条件循环，只要 *expression* 的求值为真就一直执行。

```
[<<label>>]
FOR name IN [ REVERSE ] expression .. expression LOOP
    statements
END LOOP;
```

在一个整数取值范围上迭代的循环。变量 *name* 自动创建为 integer 类型并且只存在于循环内部。给出范围下界和上界的两个表达式只在进入循环时求值一次。迭代步长总是 1。

```
[<<label>>]
FOR record | row IN select_clause LOOP
    statements
END LOOP;
```

记录或行会被依次赋予 *select* 子句产生的所有行，并对每一行执行这些语句。如果循环被 EXIT 语句终止，最后赋值的行在循环之后仍然可以访问。

```
EXIT [ label ] [ WHEN expression ];
```

如果没有给出 *label*，则终止最内层的循环，接下来执行 END LOOP 之后的语句。如果给出了 *label*，它必须是当前层或嵌套循环块上层的标签。然后命名的循环或块被终止，控制流继续到对应循环/块的 END 之后的语句。

### 49.2.2.7. 触发器过程

PL/pgSQL 可以用来定义触发器过程。它们像平常一样用 CREATE FUNCTION 命令创建一个没有参数、返回类型为 OPAQUE 的函数。

用作触发器过程的函数有一些 Postgres 特有的细节。

首先它们在顶层块的声明区中自动创建了一些特殊变量。它们是

NEW

数据类型 RECORD；在行级触发器的 INSERT/UPDATE 操作中保存新数据库行的变量。

OLD

数据类型 RECORD；在行级触发器的 UPDATE/DELETE 操作中保存旧数据库行的变量。

TG\_NAME

数据类型 name; 包含实际触发的触发器名称的变量。

TG\_WHEN

数据类型 text; 按触发器的定义为 'BEFORE' 或 'AFTER' 的字符串。

TG\_LEVEL

数据类型 text; 按触发器的定义为 'ROW' 或 'STATEMENT' 的字符串。

TG\_OP

数据类型 text; 取值为 'INSERT'、'UPDATE' 或 'DELETE' 的字符串, 指明触发器是为哪个操作实际触发的。

TG\_RELID

数据类型 oid; 导致触发器调用的表的对象 ID。

TG\_RELNAME

数据类型 name; 导致触发器调用的表的名称。

TG\_NARGS

数据类型 integer; CREATE TRIGGER 语句中传给触发器过程的参数个数。

TG\_ARGV[]

数据类型 text 数组; 来自 CREATE TRIGGER 语句的参数。索引从 0 计数并可以用表达式给出。无效索引 (< 0 或 >= tg\_nargs) 得到 NULL 值。

其次它们必须返回 NULL, 或者返回一个恰好具有触发该触发器的表的结构的记录/行。AFTER 触发器总是可以返回 NULL 值而无任何影响。BEFORE 触发器返回 NULL 时会通知触发器管理器跳过这一实际行的操作。否则, 返回的记录/行会替换操作中被插入/更新的行。可以直接在 NEW 中替换单个值并返回它, 或者构造一个完整的新记录/行返回。

### 49.2.2.8. 异常

Postgres 没有一个很聪明的异常处理模型。每当解析器、规划器/优化器或执行器判定一条语句无法继续处理时, 整个事务都会被中止, 系统会跳回主循环去获取客户端应用的下一条查询。

可以挂入错误机制来注意到发生了这种情况。但目前无法判断到底是什么导致了中止(输入/输出转换错误、浮点错误、解析错误)。而且此时数据库后端可能处于不一致状态, 返回上层执行器或发出更多命令可能会损坏整个数据库。即便可以, 此时事务已中止的信息已经发送给了客户端应用, 恢复操作没有任何意义。

因此, PL/pgSQL 目前在函数或触发器过程执行期间遇到中止时, 唯一能做的就是写一些额外的 DEBUG 级别日志消息, 说明发生在哪个函数的什么地方(行号和语句类型)。

### 49.2.3. 示例

这里只给出少量函数来演示编写 PL/pgSQL 函数有多么容易。更复杂的例子可以查看 PL/pgSQL 的回归测试。

在 PL/pgSQL 中编写函数的一个痛苦细节是单引号的处理。CREATE FUNCTION 时函数的源文本必须是字面字符串。字面字符串中的单引号必须加倍或用反斜线引用。我们仍在寻找优雅的替代方案。在此之前，应当像下面的例子那样把单引号加倍。这个问题在 Postgres 未来版本中的任何解决方案都将是向上兼容的。

### 49.2.3.1. 一些简单的 PL/pgSQL 函数

下面两个 PL/pgSQL 函数与 C 语言函数讨论中的对应函数等价。

```
CREATE FUNCTION add_one (int4) RETURNS int4 AS '
BEGIN
    RETURN $1 + 1;
END;
' LANGUAGE 'plpgsql';

CREATE FUNCTION concat_text (text, text) RETURNS text AS '
BEGIN
    RETURN $1 || $2;
END;
' LANGUAGE 'plpgsql';
```

### 49.2.3.2. 复合类型上的 PL/pgSQL 函数

这同样是 C 函数一节中示例的 PL/pgSQL 等价实现。

```
CREATE FUNCTION c_overpaid (EMP, int4) RETURNS bool AS '
DECLARE
    emprec ALIAS FOR $1;
    sallim ALIAS FOR $2;
BEGIN
    IF emprec.salary ISNULL THEN
        RETURN 'f';
    END IF;
    RETURN emprec.salary > sallim;
END;
' LANGUAGE 'plpgsql';
```

### 49.2.3.3. PL/pgSQL 触发器过程

这个触发器确保每当表中插入或更新一行时，当前的用户名和时间都会被盖印到该行上。它还确保给出了雇员的名字，并且工资是正值。

```
CREATE TABLE emp (
    empname text,
    salary int4,
    last_date datetime,
    last_user name);

CREATE FUNCTION emp_stamp () RETURNS OPAQUE AS
BEGIN
    -- Check that empname and salary are given
```

```

        IF NEW.empname ISNULL THEN
            RAISE EXCEPTION 'empname cannot be NULL value';
        END IF;
        IF NEW.salary ISNULL THEN
            RAISE EXCEPTION '% cannot have NULL salary', NEW.
empname;
        END IF;

        -- Who works for us when she must pay for?
        IF NEW.salary < 0 THEN
            RAISE EXCEPTION '% cannot have a negative salary',
NEW.empname;
        END IF;

        -- Remember who changed the payroll when
        NEW.last_date := 'now';
        NEW.last_user := getpgusername();
        RETURN NEW;
    END;
' LANGUAGE 'plpgsql';

CREATE TRIGGER emp_stamp BEFORE INSERT OR UPDATE ON emp
FOR EACH ROW EXECUTE PROCEDURE emp_stamp();

```

## 49.3. PL/Tcl

PL/Tcl 是一个可装载的 Postgres 数据库系统过程语言，它使 Tcl 语言可以用来创建函数和触发器过程。

这个软件包最初由 Jan Wieck 编写。

### 49.3.1. 概述

PL/Tcl 提供了函数编写者在 C 语言中所具备的大部分能力，但有一些限制。

好的限制是所有内容都在一个安全的 Tcl 解释器中执行。除了安全 Tcl 的有限命令集之外，只有少数命令可用于经 SPI 访问数据库以及通过 `elog()` 发出消息。没有办法像 C 语言那样访问数据库后端的内部，也无法获得 Postgres 用户 ID 权限下的 OS 级访问。因此，可以允许任何非特权数据库用户使用这一语言。

另一项内在的限制是，Tcl 过程不能用来为新数据类型创建输入/输出函数。

如果在安装过程的配置步骤中指定了 Tcl/Tk 支持，PL/Tcl 调用处理器的共享对象就会自动构建并安装到 Postgres 库目录中。

### 49.3.2. 描述

#### 49.3.2.1. Postgres 函数与 Tcl 过程名

在 Postgres 中，只要参数个数或类型不同，同一个函数名就可以用于不同的函数。这与 Tcl 过程名会冲突。为了让 PL/Tcl 提供同样的灵活性，内部 Tcl 过程名中把过程 `pg_proc` 行的对象 ID 作为名称的一部分。因此，同一个 Postgres 函数的不同参数类型版本对 Tcl 来说也是不同的。

### 49.3.2.2. 在 PL/Tcl 中定义函数

要用 PL/Tcl 语言创建函数，使用已知的语法

```
CREATE FUNCTION funcname (argument-types) RETURNS returntype AS '
    # PL/Tcl function body
    ' LANGUAGE 'pltcl';
```

在查询中调用这一函数时，参数会以变量 \$1 ... \$n 的形式传给 Tcl 过程体。因此一个返回两个 int4 值中较大者的小小 max 函数可以这样创建：

```
CREATE FUNCTION tcl_max (int4, int4) RETURNS int4 AS '
    if {$1 > $2} {return $1}
    return $2
    ' LANGUAGE 'pltcl';
```

复合类型的参数会以 Tcl 数组的形式传给过程。数组中的元素名是复合类型的属性名。如果实际行中的某个属性为 NULL 值，它不会出现在数组中！下面是一个用 PL/Tcl 定义 overpaid\_2 函数（见较早的 Postgres 文档）的例子

```
CREATE FUNCTION overpaid_2 (EMP) RETURNS bool AS '
    if {200000.0 < $1(salary)} {
        return "t"
    }
    if {$1(age) < 30 && 100000.0 < $1(salary)} {
        return "t"
    }
    return "f"
    ' LANGUAGE 'pltcl';
```

### 49.3.2.3. PL/Tcl 中的全局数据

有时（尤其是使用后文描述的 SPI 函数时），需要在对过程的两次调用之间保存一些全局状态数据。同一后端中执行的所有 PL/Tcl 过程共享同一个安全 Tcl 解释器。为帮助保护 PL/Tcl 过程免受副作用影响，每个过程都可以通过 upvar 命令访问一个数组。这个变量的全局名是过程的内部名，局部名是 GD。

### 49.3.2.4. PL/Tcl 中的触发器过程

在 Postgres 中，触发器过程定义为没有参数、返回类型为 opaque 的函数。在 PL/Tcl 语言中也是如此。

来自触发器管理器的信息会通过下列变量传给过程体：

*\$TG\_name*

来自 CREATE TRIGGER 语句的触发器名称。

*\$TG\_relid*

导致触发器过程被调用的表的对象 ID。

*\$TG\_relatts*

表字段名组成的 Tcl 列表，前面带有一个空列表元素。因此用 lsearch Tcl 命令在列表中查找元素名时，返回的正数编号与字段在 pg\_attribute 系统目录中的编号相同（都从 1 开始）。

*\$TG\_when*

按触发器调用的事件为 BEFORE 或 AFTER 的字符串。

*\$TG\_level*

按触发器调用的事件为 ROW 或 STATEMENT 的字符串。

*\$TG\_op*

按触发器调用的事件为 INSERT、UPDATE 或 DELETE 的字符串。

*\$NEW*

在 INSERT/UPDATE 操作中包含新表行值的数组，DELETE 时空。

*\$OLD*

在 UPDATE/DELETE 操作中包含旧表行值的数组，INSERT 时空。

*\$GD*

上面描述的全局状态数据数组。

*\$args*

按 CREATE TRIGGER 语句给出的传给过程的参数组成的 Tcl 列表。这些参数在过程体中也可以用 \$1 ... \$n 访问。

触发器过程的返回值是字符串 OK 或 SKIP 之一，或者是 'array get' Tcl 命令返回的列表。如果返回值是 OK，触发本次触发的正常操作（INSERT/UPDATE/DELETE）将会进行。显然，SKIP 告诉触发器管理器悄悄地抑制该操作。'array get' 的列表告诉 PL/Tcl 向触发器管理器返回一个修改过的行，用它代替 \$NEW 中给出的行进行插入（仅限 INSERT/UPDATE）。不用说，所有这些只有在触发器为 BEFORE 且 FOR EACH ROW 时才有意义。

下面是一个小示例触发器过程，它强制表中的一个整数值跟踪对该行执行的更新次数。对于新插入的行，该值初始化为 0，然后每次更新操作时递增：

```
CREATE FUNCTION trigfunc_modcount() RETURNS OPAQUE AS '
    switch $TG_op {
        INSERT {
            set NEW($1) 0
        }
        UPDATE {
            set NEW($1) $OLD($1)
            incr NEW($1)
        }
        default {
            return OK
        }
    }
    return [array get NEW]
```

```
' LANGUAGE 'pltcl';

CREATE TABLE mytab (num int4, modcnt int4, desc text);

CREATE TRIGGER trig_mytab_modcount BEFORE INSERT OR UPDATE ON
mytab
    FOR EACH ROW EXECUTE PROCEDURE trigfunc_modcount('modcnt');
```

### 49.3.2.5. 从 PL/Tcl 访问数据库

在 PL/Tcl 过程体中可以使用下列命令来访问数据库：

`elog level msg`

发出一条日志消息。可能的级别有 NOTICE、WARN、ERROR、FATAL、DEBUG 和 NOIND，与 `elog()` C 函数的相同。

`quote string`

复制所有单引号和反斜杠字符。在给 `spi_exec` 或 `spi_prepare` 的查询字符串中使用变量时应该使用它（不用于 `spi_execp` 的值列表）。想一想这样的查询字符串：

```
"SELECT '$val' AS ret"
```

其中 Tcl 变量 `val` 实际包含 "doesn't"。这会得到最终的查询字符串

```
"SELECT 'doesn't' AS ret"
```

它会在 `spi_exec` 或 `spi_prepare` 期间导致解析错误。它应该包含

```
"SELECT 'doesn''t' AS ret"
```

并且必须写成

```
"SELECT '[ quote $val ]' AS ret"
```

`spi_exec ?-count n? ?-array name? query?loop-body?`

为查询调用解析器/规划器/优化器/执行器。可选的 `-count` 值告诉 `spi_exec` 查询要处理的最大行数。

如果查询是 SELECT 语句并且给出了可选的循环体（一组 Tcl 命令，就像在 `foreach` 语句中那样），则对选出的每一行求值，并且在 `continue/break` 上的行为与预期一致。选出字段的值被放入以列名命名的变量中。因此

```
spi_exec "SELECT count(*) AS cnt FROM pg_proc"
```

会把变量 `$cnt` 设置为 `pg_proc` 系统目录中的行数。如果给出了 `-array` 选项，列值将存入名为 `'name'` 的关联数组并以列名为索引，而不是各个单独的变量。

```
spi_exec -array C "SELECT * FROM pg_class" {
```

```

        elog DEBUG "have table %C(relname)"
    }

```

会对 `pg_class` 的每一行打印一条 `DEBUG` 日志消息。`spi_exec` 的返回值是查询影响的行数，与全局变量 `SPI_processed` 中的值相同。

#### `spi_prepare query typelist`

预备并保存一个查询计划供以后执行。它与 C 层的 `SPI_prepare` 稍有不同：计划会被自动复制到顶层内存上下文。因此，目前没有办法预备计划而不保存它。

如果查询引用了参数，类型名必须以 Tcl 列表的形式给出。`spi_prepare` 的返回值是一个查询 ID，用于后续对 `spi_execp` 的调用。示例见 `spi_execp`。

#### `spi_exec ?-count n? ?-array name? ?-nulls str? query ?valuelist? ?loop-body?`

执行 `spi_prepare` 预备的计划并做变量替换。可选的 `-count` 值告诉 `spi_execp` 查询要处理的最大行数。

`-nulls` 的可选值是由空格和 'n' 字符组成的字符串，告诉 `spi_execp` 哪些值是 `NULL`。如果给出，它的长度必须恰好等于值的个数。

`queryid` 是 `spi_prepare` 调用返回的 ID。

如果 `spi_prepare` 给出了类型列表，则必须在查询之后给 `spi_execp` 一个长度恰好相同的 Tcl 值列表。如果 `spi_prepare` 的类型列表为空，则必须省略这个参数。

如果查询是 `SELECT` 语句，循环体和选出字段的变量的处理方式与 `spi_exec` 一节所述相同。

下面是一个使用预备计划的 PL/Tcl 函数示例：

```

CREATE FUNCTION tl_count(int4, int4) RETURNS int4 AS '
    if {[ info exists GD(plan) ]} {
        # prepare the saved plan on the first call
        set GD(plan) [ spi_prepare \
            "SELECT count(*) AS cnt FROM tl WHERE num >= \
$1 AND num <= \\\$2" \
            int4 ]
    }
    spi_execp -count 1 $GD(plan) [ list $1 $2 ]
    return $cnt
' LANGUAGE 'pltcl';

```

注意，Tcl 应该看到的每个反斜杠在创建函数的查询中都必须加倍，因为主解析器在 `CREATE FUNCTION` 时也会处理反斜杠。在给 `spi_prepare` 的查询字符串内部应该使用美元符号标记参数位置，以免 `$1` 被第一次函数调用时给出的值替换。

#### 模块和 unknown 命令

PL/Tcl 对常用的东西提供了特殊支持。它识别两个魔法表 `pltcl_modules` 和 `pltcl_modfuncs`。如果它们存在，'unknown' 模块会在解释器创建后立即装载。每当调用一个未知的 Tcl 过程时，会请求 `unknown` 过程检查该过程是否定义在某个模块中。如果为真，就按需装载该模块。要启用这一行为，PL/Tcl 调用处理器必须在编译时设置 `-DPLTCL_UNKNOWN_SUPPORT`。

在 PL/Tcl 源码的 modules 子目录中有维护这些表的支持脚本，其中包括必须最先安装的 unknown 模块的源码。

---

# 部分 V. 接口

用户与程序员接口。

---

# 目录

|                                |     |
|--------------------------------|-----|
| 50. 函数 .....                   | 421 |
| 51. Large Objects .....        | 422 |
| 51.1. 历史注记 .....               | 422 |
| 51.2. Inversion 大对象 .....      | 422 |
| 51.3. 大对象接口 .....              | 422 |
| 51.3.1. 创建一个对象 .....           | 422 |
| 51.3.2. 导入一个大对象 .....          | 422 |
| 51.3.3. 导出一个大对象 .....          | 423 |
| 51.3.4. 打开一个现有的大对象 .....       | 423 |
| 51.3.5. 向大对象写入数据 .....         | 423 |
| 51.3.6. 在大对象上定位 .....          | 423 |
| 51.3.7. 关闭一个大对象描述符 .....       | 423 |
| 51.4. 内建已注册函数 .....            | 423 |
| 51.5. 从 LIBPQ 访问大对象 .....      | 424 |
| 51.6. 示例程序 .....               | 424 |
| 52. ecpg - C 中的嵌入式 SQL .....   | 429 |
| 52.1. 为什么用嵌入式 SQL? .....       | 429 |
| 52.2. 概念 .....                 | 429 |
| 52.3. 如何使用 ecpg .....          | 429 |
| 52.3.1. 预处理器 .....             | 429 |
| 52.3.2. 库 .....                | 430 |
| 52.3.3. 错误处理 .....             | 430 |
| 52.4. 限制 .....                 | 432 |
| 52.5. 从其他 RDBMS 软件包移植 .....    | 432 |
| 52.6. 安装 .....                 | 432 |
| 52.7. 给开发者 .....               | 432 |
| 52.7.1. 待办列表 .....             | 432 |
| 52.7.2. The Preprocessor ..... | 433 |
| 52.7.3. 一个完整的例子 .....          | 435 |
| 52.7.4. 库 .....                | 436 |
| 53. libpq .....                | 437 |
| 53.1. 数据库连接函数 .....            | 437 |
| 53.2. 查询执行函数 .....             | 440 |
| 53.3. 异步查询处理 .....             | 443 |
| 53.4. 快速路径 .....               | 445 |
| 53.5. 异步通知 .....               | 446 |
| 53.6. 与 COPY 命令相关的函数 .....     | 446 |
| 53.7. libpq 跟踪函数 .....         | 448 |
| 53.8. libpq 控制函数 .....         | 448 |
| 53.9. 用户认证函数 .....             | 449 |
| 53.10. 环境变量 .....              | 449 |
| 53.11. 注意事项 .....              | 450 |
| 53.12. 示例程序 .....              | 450 |
| 53.12.1. 示例程序 1 .....          | 450 |
| 53.12.2. 示例程序 2 .....          | 452 |
| 53.12.3. 示例程序 3 .....          | 455 |
| 54. pgsql .....                | 459 |
| 54.1. 命令 .....                 | 459 |
| 54.2. 示例 .....                 | 459 |
| 54.3. pgsql 命令参考信息 .....       | 460 |
| 55. ODBC 接口 .....              | 479 |

---

|                                      |     |
|--------------------------------------|-----|
| 55.1. 背景 .....                       | 479 |
| 55.2. Windows应用 .....                | 479 |
| 55.2.1. 编写应用 .....                   | 479 |
| 55.3. Unix 安装 .....                  | 480 |
| 55.3.1. 构建驱动 .....                   | 480 |
| 55.4. 配置文件 .....                     | 483 |
| 55.5. ApplixWare .....               | 484 |
| 55.5.1. 配置 .....                     | 484 |
| 55.5.2. 常见问题 .....                   | 485 |
| 55.5.3. 调试ApplixWare ODBC 连接 .....   | 485 |
| 55.5.4. 运行ApplixWare演示 .....         | 486 |
| 55.5.5. 有用的宏 .....                   | 487 |
| 55.5.6. 常见问题 .....                   | 488 |
| 55.5.7. 调试ApplixWare ODBC 连接 .....   | 488 |
| 55.5.8. 运行ApplixWare演示 .....         | 489 |
| 55.5.9. 有用的宏 .....                   | 490 |
| 55.5.10. 支持的平台 .....                 | 490 |
| 56. JDBC 接口 .....                    | 491 |
| 56.1. 构建 JDBC 接口 .....               | 491 |
| 56.1.1. 编译驱动 .....                   | 491 |
| 56.1.2. 安装驱动 .....                   | 491 |
| 56.2. 为 JDBC 准备数据库 .....             | 491 |
| 56.3. 使用驱动 .....                     | 492 |
| 56.4. 导入 JDBC .....                  | 492 |
| 56.5. 载入驱动 .....                     | 492 |
| 56.6. 连接到数据库 .....                   | 493 |
| 56.7. 发出查询并处理结果 .....                | 493 |
| 56.7.1. 使用 Statement 接口 .....        | 493 |
| 56.7.2. 使用 ResultSet 接口 .....        | 494 |
| 56.8. 执行更新 .....                     | 494 |
| 56.9. 关闭连接 .....                     | 494 |
| 56.10. 使用大对象 .....                   | 494 |
| 56.11. Postgres 对 JDBC API 的扩展 ..... | 495 |
| 56.12. 延伸阅读 .....                    | 531 |

---

# 第 50 章 函数

用户可调用函数的参考信息。

## 注意

本节有待撰写。欢迎志愿者！

---

# 第 51 章 Large Objects

在 Postgres 中，数据值存储在元组中，而单个元组不能跨数据页。由于数据页的大小是 8192 字节，数据值大小的上限相对较低。为了支持更大的原子值的存储，Postgres 提供了大对象接口。该接口以面向文件的方式访问被声明为大类型的用户数据。本节描述 Postgres 大对象数据的实现以及编程和查询语言接口。

## 51.1. 历史注记

最初，Postgres 4.2 支持三种标准的大对象实现：作为 Postgres 之外的文件、作为 Postgres 管理的 UNIX 文件、以及作为存储在 Postgres 数据库内部的数据。这给用户带来了相当大的困扰。因此，在 PostgreSQL 中我们只支持把大对象作为存储在 Postgres 数据库内部的数据。尽管访问起来更慢，它提供了更严格的数据完整性。由于历史原因，这种存储方案被称为 Inversion 大对象。（本节中我们将把 Inversion 和大对象作为同义词混用，指同一事物。）

## 51.2. Inversion 大对象

Inversion 大对象的实现把大对象拆分成“块”并把块存储在数据库的元组中。一个 B-tree 索引保证在做随机访问读写时能快速搜索到正确的块号。

## 51.3. 大对象接口

下面描述 Postgres 提供的访问大对象的设施，既包括后端中作为用户自定义函数一部分的用法，也包括前端中作为使用该接口的应用一部分的用法。（对于熟悉 Postgres 4.2 的用户来说，PostgreSQL 有一套新函数，提供了更连贯的接口。该接口对动态装载的 C 函数与对 XXX LOST TEXT? WHAT SHOULD GO HERE?? 是一样的。Postgres 的大对象接口模仿 UNIX 文件系统接口设计，有与 `open(2)`、`read(2)`、`write(2)`、`lseek(2)` 等对应的函数。用户函数调用这些例程来只从大对象中检索感兴趣的数据。例如，假设存在一个存储面孔照片的名为 `mugshot` 的大对象类型，那么可以在 `mugshot` 数据上声明一个名为 `beard` 的函数。`Beard` 可以查看照片的下三分之一部分，并确定那里出现的胡须颜色（如果有的话）。整个大对象的值不需要被 `beard` 函数缓冲，甚至不需要被检查。大对象可以从动态装载的 C 函数或链接了该库的数据库客户端程序访问。Postgres 提供了一组支持打开、读取、写入、关闭和在大对象上定位的例程。

### 51.3.1. 创建一个大对象

例程

```
Oid lo_creat(PGconn *conn, int mode)
```

创建一个新的大对象。`mode` 是一个位掩码，描述新大对象的若干不同属性。这里列出的符号常量定义在 `PGROOT/src/backend/libpq/libpq-fs.h` 中。访问类型（读、写或读写）由把 `INV_READ` 和 `INV_WRITE` 位“或”在一起来控制。如果大对象应当被归档——也就是说，如果它的历史版本应当被定期移到一个特殊的归档关系中——那么应当设置 `INV_ARCHIVE` 位。掩码的低十六位是大对象应驻留的存储管理器编号。对于伯克利以外的站点，这些位应当始终为零。下面的命令创建一个（Inversion）大对象：

```
inv_oid = lo_creat(INV_READ|INV_WRITE|INV_ARCHIVE);
```

### 51.3.2. 导入一个大对象

要把一个 UNIX 文件导入为大对象，调用

```
Oid lo_import(PGconn *conn, text *filename)
```

filename 参数指定要导入为大对象的文件的 UNIX 路径名。

### 51.3.3. 导出一个大对象

要把一个大对象导出到 UNIX 文件，调用

```
int lo_export(PGconn *conn, Oid loobjId, text *filename)
```

loobjId 参数指定要导出的大对象的 Oid，filename 参数指定文件的 UNIX 路径名。

### 51.3.4. 打开一个现有的大对象

要打开一个现有的大对象，调用

```
int lo_open(PGconn *conn, Oid loobjId, int mode, ...)
```

loobjId 参数指定要打开的大对象的 Oid。mode 位控制对象是以读 (INV\_READ)、写还是读写方式打开。大对象在创建之前无法打开。lo\_open 返回一个大对象描述符，供稍后在 lo\_read、lo\_write、lo\_lseek、lo\_tell 和 lo\_close 中使用。

### 51.3.5. 向大对象写入数据

例程

```
int lo_write(PGconn *conn, int fd, char *buf, int len)
```

把 len 字节从 buf 写入大对象 fd。fd 参数必须是先前某个 lo\_open 返回的。返回值是实际写入的字节数。发生错误时，返回值为负。

### 51.3.6. 在大对象上定位

要更改大对象上当前的读或写位置，调用

```
int lo_lseek(PGconn *conn, int fd, int offset, int whence)
```

此例程把 fd 所描述大对象的当前位置指针移动到 offset 指定的新位置。whence 的有效值是 SEEK\_SET、SEEK\_CUR 和 SEEK\_END。

### 51.3.7. 关闭一个大对象描述符

可以通过调用下面的函数来关闭一个大对象：

```
int lo_close(PGconn *conn, int fd)
```

其中 fd 是由 lo\_open 返回的大对象描述符。成功时 lo\_close 返回零。发生错误时，返回值为负。

## 51.4. 内建已注册函数

有两个内置的已注册函数 lo\_import 和 lo\_export，它们便于在 SQL 查询中使用。下面是一个用法的例子：

```
CREATE TABLE image (
```

```
        name          text,
        raster        oid
    );

INSERT INTO image (name, raster)
VALUES ('beautiful image', lo_import('/etc/motd'));

SELECT lo_export(image.raster, "/tmp/motd") from image
WHERE name = 'beautiful image';
```

## 51.5. 从 LIBPQ 访问大对象

下面是一个示例程序，展示了 LIBPQ 中大对象接口的用法。程序的一部分被注释掉了，但为了读者受益仍保留在源码中。该程序可以在 `../src/test/examples` 中找到。使用 LIBPQ 中大对象接口的前端应用应当包含头文件 `libpq/libpq-fs.h` 并链接 `libpq` 库。

## 51.6. 示例程序

```
/*-----
 *
 * testlo.c--
 *   test using large objects with libpq
 *
 * Copyright (c) 1994, Regents of the University of
California
 *
 *
 * IDENTIFICATION
 *   /usr/local/devel/pglite/cvs/src/doc/manual.me,v 1.16
1995/09/01 23:55:00 jolly Exp
 *
 *-----
----
 */
#include <stdio.h>
#include "libpq-fe.h"
#include "libpq/libpq-fs.h"

#define BUFSIZE          1024

/*
 * importFile *      import file "in_filename" into database as
large object "lobjOid"
 *
 */
Oid importFile(PGconn *conn, char *filename)
{
    Oid lobjId;
    int lobj_fd;
    char buf[BUFSIZE];
    int nbytes, tmp;
    int fd;
```

```

/*
 * open the file to be read in
 */
fd = open(filename, O_RDONLY, 0666);
if (fd < 0) { /* error */
    fprintf(stderr, "can't open unix file %s\n", filename);
}

/*
 * create the large object
 */
lobjId = lo_creat(conn, INV_READ|INV_WRITE);
if (lobjId == 0) {
    fprintf(stderr, "can't create large object\n");
}

lobj_fd = lo_open(conn, lobjId, INV_WRITE);
/*
 * read in from the Unix file and write to the inversion
file
 */
while ((nbytes = read(fd, buf, BUFSIZE)) > 0) {
    tmp = lo_write(conn, lobj_fd, buf, nbytes);
    if (tmp < nbytes) {
        fprintf(stderr, "error while reading large object
\n");
    }
}

(void) close(fd);
(void) lo_close(conn, lobj_fd);

return lobjId;
}

void pickout(PGconn *conn, Oid lobjId, int start, int len)
{
    int lobj_fd;
    char* buf;
    int nbytes;
    int nread;

    lobj_fd = lo_open(conn, lobjId, INV_READ);
    if (lobj_fd < 0) {
        fprintf(stderr, "can't open large object %d\n",
            lobjId);
    }

    lo_lseek(conn, lobj_fd, start, SEEK_SET);
    buf = malloc(len+1);

    nread = 0;
    while (len - nread > 0) {

```

```
        nbytes = lo_read(conn, lobj_fd, buf, len - nread);
        buf[nbytes] = ' ';
        fprintf(stderr, ">>> %s", buf);
        nread += nbytes;
    }
    fprintf(stderr, "\n");
    lo_close(conn, lobj_fd);
}

void overwrite(PGconn *conn, Oid lobjId, int start, int len)
{
    int lobj_fd;
    char* buf;
    int nbytes;
    int nwritten;
    int i;

    lobj_fd = lo_open(conn, lobjId, INV_READ);
    if (lobj_fd < 0) {
        fprintf(stderr, "can't open large object %d\n",
            lobjId);
    }

    lo_lseek(conn, lobj_fd, start, SEEK_SET);
    buf = malloc(len+1);

    for (i=0; i<len; i++)
        buf[i] = 'X';
    buf[i] = ' ';

    nwritten = 0;
    while (len - nwritten > 0) {
        nbytes = lo_write(conn, lobj_fd, buf + nwritten, len -
nwritten);
        nwritten += nbytes;
    }
    fprintf(stderr, "\n");
    lo_close(conn, lobj_fd);
}

/*
 * exportFile *      export large object "lobjOid" to file
"out_filename"
 *
 */
void exportFile(PGconn *conn, Oid lobjId, char *filename)
{
    int lobj_fd;
    char buf[BUFSIZE];
    int nbytes, tmp;
    int fd;

    /*
     * create an inversion "object"

```

```
        */
        lobj_fd = lo_open(conn, lobjId, INV_READ);
        if (lobj_fd < 0) {
            fprintf(stderr, "can't open large object %d\n",
                    lobjId);
        }

        /*
         * open the file to be written to
         */
        fd = open(filename, O_CREAT|O_WRONLY, 0666);
        if (fd < 0) { /* error */
            fprintf(stderr, "can't open unix file %s\n",
                    filename);
        }

        /*
         * read in from the Unix file and write to the inversion
file
        */
0) {
        while ((nbytes = lo_read(conn, lobj_fd, buf, BUFSIZE)) >

            tmp = write(fd, buf, nbytes);
            if (tmp < nbytes) {
                fprintf(stderr, "error while writing %s\n",
                        filename);
            }
        }

        (void) lo_close(conn, lobj_fd);
        (void) close(fd);

        return;
    }

void
exit_nicely(PGconn* conn)
{
    PQfinish(conn);
    exit(1);
}

int
main(int argc, char **argv)
{
    char *in_filename, *out_filename;
    char *database;
    Oid lobjOid;
    PGconn *conn;
    PGresult *res;

    if (argc != 4) {
        fprintf(stderr, "Usage: %s database_name in_filename
out_filename\n",
```

```
        argv[0]);
    exit(1);
}

database = argv[1];
in_filename = argv[2];
out_filename = argv[3];

/*
 * set up the connection
 */
conn = PQsetdb(NULL, NULL, NULL, NULL, database);

/* check to see that the backend connection was
successfully made */
if (PQstatus(conn) == CONNECTION_BAD) {
    fprintf(stderr, "Connection to database '%s' failed.\n",
database);
    fprintf(stderr, "%s", PQerrorMessage(conn));
    exit_nicely(conn);
}

res = PQexec(conn, "begin");
PQclear(res);

printf("importing file %s\n", in_filename);
/* lobjOid = importFile(conn, in_filename); */
lobjOid = lo_import(conn, in_filename);
/*
printf("as large object %d.\n", lobjOid);

printf("picking out bytes 1000-2000 of the large object
\n");
pickout(conn, lobjOid, 1000, 1000);

printf("overwriting bytes 1000-2000 of the large object
with X's\n");
overwrite(conn, lobjOid, 1000, 1000);
*/

printf("exporting large object to file %s\n", out_
filename);
/* exportFile(conn, lobjOid, out_filename); */
lo_export(conn, lobjOid, out_filename);

res = PQexec(conn, "end");
PQclear(res);
PQfinish(conn);
exit(0);
}
```

---

# 第 52 章 ecpg - C 中的嵌入式 SQL

Linux Tolke  
Michael Meskes

版权 © 1996-1997 Linus Tolke  
版权 © 1998 Michael Meskes

本文介绍 Postgres 的一个 C 中的嵌入式 SQL 软件包。它由 Linus Tolke<sup>1</sup> 和 Michael Meskes<sup>2</sup> 编写。

## 注意

授予复制和使用的许可，方式与你被允许复制和使用 PostgreSQL 其余部分的方式相同。

## 52.1. 为什么用嵌入式 SQL?

与处理 SQL 查询的其他方式相比，嵌入式 SQL 有一些小的优势。它负责了在 C 程序的变量之间移动信息的全部繁琐工作。许多 RDBMS 软件包都支持这种嵌入式语言。

有一个 ANSI 标准描述了这种嵌入式语言应当如何工作。ecpg 的设计尽可能满足这个标准。因此可以把为其他 RDBMS 软件包编写的嵌入式 SQL 程序移植到 Postgres，从而弘扬自由软件的精神。

## 52.2. 概念

你用 C 加上一些特殊的 SQL 内容编写程序。要声明可以在 SQL 语句中使用的变量，需要把它们放在特殊的 declare 区中。SQL 查询要使用特殊的语法。

编译之前，先用嵌入式 SQL C 预处理器处理文件，它把你使用的 SQL 语句转换成以变量为参数的函数调用。既传递用作 SQL 语句输入的变量，也传递将包含结果的变量。

然后你编译并在链接时与一个包含所用函数的特殊库链接。这些函数（实际上主要是一个函数）从参数中获取信息，用普通接口（libpq）执行 SQL 查询，并把结果放回专门用于输出的参数中。

然后你运行程序，当控制到达 SQL 语句时，该 SQL 语句就会针对数据库执行，你可以接着使用结果。

## 52.3. 如何使用 ecpg

本节介绍如何使用 ecpg 工具。

### 52.3.1. 预处理器

预处理器叫做 ecpg。安装后它位于 Postgres 的 bin/ 目录中。

---

<sup>1</sup>mailto:linus@epact.se

<sup>2</sup>mailto:meskes@debian.org

## 52.3.2. 库

ecpg 库叫做 `libecpg.a` 或 `libecpg.so`。此外，该库用 `libpq` 库与 Postgres 服务器通信，因此你必须用 `-lecpg -lpq` 链接你的程序。

库中有一些"隐藏的"方法，但它们有时可能非常有用。

- `ECPGdebug(int on, FILE *stream)` 在第一个参数非零时打开调试日志。调试日志输出到 `stream`。大多数 SQL 语句会记录其参数和结果。

其中最重要的一个是 `ECPGdo`，除 `EXEC SQL COMMIT`、`EXEC SQL ROLLBACK`、`EXEC SQL CONNECT` 之外的所有 SQL 语句都会调用它，它会同时记录展开后的字符串（即插入了所有输入变量的字符串）以及 Postgres 服务器返回的结果。在排查 SQL 语句中的错误时这会非常有用。

- `ECPGstatus()` 这个方法在已连接到数据库时返回 `TRUE`，否则返回 `FALSE`。

## 52.3.3. 错误处理

为了能够检测来自 Postgres 服务器的错误，你可以在文件的 include 区中包含像下面这样的一行

```
exec sql include sqlca;
```

这会定义一个结构以及一个名为 `sqlca` 的变量，如下：

```
struct sqlca {
    int sqlcode;
    struct {
        int sqlerrml;
        char sqlerrmc[1000];
    } sqlerrm;
} sqlca;
```

如果上一条 SQL 语句发生了错误，则 `sqlca.sqlcode` 将非零。如果 `sqlca.sqlcode` 小于 0，这是某种严重的错误，例如数据库定义与所给的查询不匹配。如果大于 0，则是普通错误，例如表中没有所请求的行。

`sqlca.sqlerrm.sqlerrmc` 将包含一个描述错误的字符串。该字符串以 "line 23." 结尾，其中的行号是源文件中的行号（实际上是预处理器生成的文件中的行号，但我希望能把它修正为输入文件中的行号。）

可能发生的错误列表：

-1, Unsupported type %s on line %d.

通常不发生。这表示预处理器生成了库不认识的内容。  
也许你运行的预处理器和库的版本不兼容。

-1, Too many arguments line %d.

预处理器搞砸了，生成了一些不正确的代码。

-1, Too few arguments line %d.

预处理器搞砸了，生成了一些不正确的代码。

-1, Error starting transaction line %d.

Postgres 向我们表示无法打开连接。

-1, Postgres error: %s line %d.

某种 Postgres 错误。消息中包含来自 Postgres 后端的错误消息。

1, Data not found line %d.

这是一个"正常"错误，告诉你所查询的内容找不到，或者我们已经遍历完了游标。

-1, Too many matches line %d.

这表示查询返回了多行。你执行的 `SELECT` 很可能不唯一。

-1, Not correctly formatted int type: %s line %d.

这表示宿主变量的类型是 `int`，而 Postgres 数据库中的字段是另一种类型，包含不能被解释为 `int` 的值。库用 `strtol` 做这个转换。

-1, Not correctly formatted unsigned type: %s line %d.

这表示宿主变量的类型是 `unsigned int`，而 Postgres 数据库中的字段是另一种类型，包含不能被解释为 `unsigned int` 的值。库用 `strtoul` 做这个转换。

-1, Not correctly formatted floating point type: %s line %d.

这表示宿主变量的类型是 `float`，而 Postgres 数据库中的字段是另一种类型，包含不能被解释为 `float` 的值。库用 `strtod` 做这个转换。

-1, Too few arguments line %d.

这表示 Postgres 返回的记录比我们拥有的匹配变量多。也许你在 `INTO :var1,:var2` 列表中漏掉了几个宿主变量。

-1, Too many arguments line %d.

这表示 Postgres 返回的记录比我们拥有的宿主变量少。也许你在 `INTO :var1,:var2` 列表中放了太多宿主变量。

-1, Empty query line %d.

Postgres 返回了 `PGRES_EMPTY_QUERY`。

-1, Error: %s line %d.

这表示 Postgres 返回了 `PGRES_NONFATAL_ERROR`、`PGRES_FATAL_ERROR` 或 `PGRES_BAD_RESPONSE` 错误之一。是哪一种以及原因会在消息中说明。

-1, Postgres error line %d.

Postgres 返回了库不知道如何处理的内容。这很可能是因为 Postgres 的版本与 `ecpg` 库的版本不匹配。

-1, Error committing line %d.

`COMMIT` 期间出错。`EXEC SQL COMMIT` 被翻译为 Postgres 中的一个 `end` 操作，而那正是无法执行的操作。

-1, Error rolling back line %d.

ROLLBACK 期间出错。EXEC SQL ROLLBACK 被翻译为 Postgres 中的一个 abort 操作，而那正是无法执行的操作。

-1, ECPGconnect: could not open database %s.

连接数据库失败。

## 52.4. 限制

永远不会被包含的内容及其原因，或者这个概念做不到的事情。

Oracle 的单任务可能性

AIX 3 上的 Oracle 7.0 版使用共享内存段上由 OS 支持的锁，允许应用设计者以所谓单任务的方式链接应用。不是每个应用进程启动一个客户端进程，而是数据库部分和应用部分运行在同一个进程中。在 oracle 的后续版本中这已不再受支持。

这需要彻底重新设计 Postgres 的访问模型，而这种努力配不上获得的性能。

## 52.5. 从其他 RDBMS 软件包移植

待由了解各种 RDBMS 软件包并实际移植过东西的人来写.....

## 52.6. 安装

从 0.5 版开始，ecpg 与 Postgres 一起发行。因此作为安装的一部分，你的预编译器、库和头文件默认应该已被编译和安装。

## 52.7. 给开发者

本节面向想要开发 ecpg 接口的人。它描述各部分如何工作。这里的定位是让本节包含给想深入了解的人的内容，而如何使用一节应该足以回答所有常规问题。所以，在研究 ecpg 的内部之前先读这一节。如果你对它究竟如何工作不感兴趣，请跳过本节。

### 52.7.1. 待办列表

这个版本的预处理器有一些缺陷：

并非只限于字符串

ecpg 所使用的 PQ 接口（尤其是 PQexec 函数）依赖请求是以字符串形式构建的。在某些情况下，比如数据中包含空字符时，这会是一个严重的问题。

错误码

不同的错误应当有不同的错误编号，而不是全都用 -1。

库函数

to\_date et al.

## 记录

记录或结构必须在 declare 区中定义。

## 缺失的语句

下列语句迄今尚未实现：

exec sql type

exec sql prepare

exec sql allocate

exec sql free

exec sql whenever sqlwarning

SQLSTATE

message 'no data found'

exec sql insert select from 语句中"no data"的错误消息应该是 100。

sqlwarn[6]

如果 SET DESCRIPTOR 语句中指定的 PRECISION 或 SCALE 值将被忽略, sqlwarn[6] 应为 'W'。

## 脚本转换

建立数据库需要一些包含表定义和其他配置参数的脚本。  
如果你有一套旧数据库的这些脚本, 你会希望直接应用它们得到一个以相同方式工作的 Postgres 数据库。

这种功能可以用一些转换脚本来实现, 但速度永远无法以这种方式达成。要做到这一点, 你需要对数据库的构造和使用有比脚本所能实现的更深入的理解。

## 52.7.2. The Preprocessor

写入输出的前四行是：两行注释和两行与库接口所必需的包含行。

随后预处理器只单遍工作, 边读输入文件边写输出。  
通常它只是把所有内容原样回显到输出而不再细看。

当遇到 EXEC SQL 语句时, 它会介入并根据其类型进行改写。EXEC SQL 语句可以是下列之一：

### 声明区

声明节以

```
exec sql begin declare section;
```

开始, 以

```
exec sql end declare section;
```

结束。节中只允许变量声明。

在这个节中声明的每个变量还会连同其相应的类型一起按名字登记到一个变量列表中。

声明也会被回显到文件中, 使该变量同时成为一个普通的 C 变量。

特殊类型 VARCHAR 和 VARCHAR2 会为每个变量转换成一个命名的结构。像这样的声明:

```
VARCHAR var[180];
```

被转换成

```
struct varchar_var { int len; char arr[180]; } var;
```

包含语句

include 语句的形式是:

```
exec sql include filename;
```

它被转换成

```
#include <filename.h>
```

连接语句

connect 语句的形式是:

```
exec sql connect 'database';
```

该语句被转换成

```
ECPGconnect("database");
```

打开游标语句

open cursor 语句的形式是:

```
exec sql open cursor;
```

它被忽略, 不复制到输出中。

提交语句

commit 语句的形式是

```
exec sql commit;
```

在输出中被翻译为

```
ECPGcommit(__LINE__);
```

### 回滚语句

rollback 语句的形式是

```
exec sql rollback;
```

在输出中被翻译为

```
ECPGrollback(__LINE__);
```

### 其他语句

其他 SQL 语句是其他以 `exec sql` 开始、以 `;` 结束的语句。中间的一切都被当作 SQL 语句并解析变量替换。

当符号以冒号 (`:`) 开头时发生变量替换。

此时会在先前于声明节中声明的变量里找到具有该名字的变量，并且根据 SQL 语句是否把它当作输入或输出变量，把指向变量的指针写入输出以允许函数访问。

对于 SQL 请求中的每个变量，函数都会得到另外五个参数：

以特殊符号表示的类型

指向值的指针

变量的大小（如果是 `varchar`）

数组中的元素个数（用于数组抓取）

到数组中下一个元素的偏移量（用于数组抓取）

由于数组抓取尚未实现，最后两个参数其实并不重要。它们或许本可以省略。

## 52.7.3. 一个完整的例子

下面是一个完整的例子，描述预处理器的输出：

```
exec sql begin declare section;
int index;
int result;
exec sql end declare section;
...
    exec sql select res into :result from mytable where index = :
index;
```

被翻译成：

```
/* These two include files are added by the preprocessor */
#include <ecpgtype.h>
#include <ecpglib.h>
/* exec sql begin declare section */

    int index;
    int result;
/* exec sql end declare section */
```

```
...
    ECPGdo( __LINE__, "select res from mytable where index = ;;",
            ECPGt_int, &index, 0, 0, sizeof(int),
            ECPGt_EOIT,
            ECPGt_int, &result, 0, 0, sizeof(int),
            ECPGt_EORT );
```

(本手册中的缩进是为了可读性而加的，并不是预处理器能做的。)

## 52.7.4. 库

库中最重要的函数是 `ECPGdo` 函数。它接受数量可变的参数。但愿我们不会碰到对 `varchar` 函数可接受的变量数量有限制的机器。这很容易累积到 50 个左右的参数。

参数有：

一个行号

这是原语句的行号，只用于错误消息。

一个字符串

这是要发出的 SQL 请求。该请求被输入变量修改，即那些编译时未知但要输入请求的变量。在变量应出现的位置，字符串中包含 “;”。

输入变量

如关于预处理器的一节所述，每个输入变量得到五个参数。

`ECPGt_EOIT`

一个枚举，表示不再有输入变量。

输出变量

如关于预处理器的一节所述，每个输出变量得到五个参数。这些变量由函数填充。

`ECPGt_EORT`

一个枚举，表示不再有变量。

所有 SQL 语句都在一个事务中执行，除非你发出提交事务。其机制是第一条事务或提交/回滚之后的第一条语句总是会开始一个事务。

待完成：描述其他条目的条目。

---

## 第 53 章 libpq

libpq 是 Postgres 的 C 应用程序接口。libpq 是一组库例程，客户端程序通过它们可以向 Postgres 后端服务器传送查询并接收这些查询的结果。libpq 也是其他几个 Postgres 应用接口的底层引擎，包括 libpq++ (C++)、libpgtcl (Tcl)、perl5 和 ecpg。因此，即使你使用的是上述某个软件包，libpq 行为的某些方面对你来说也很重要。本节末尾包含三个简短程序，展示如何编写使用 libpq 的程序。下列目录中还有几个完整的 libpq 应用程序示例：

```
../src/test/regress
../src/test/examples
../src/bin/psql
```

使用 libpq 的前端程序必须包含头文件 libpq-fe.h，并且必须与 libpq 库链接。

### 53.1. 数据库连接函数

下面的例程用于建立与 Postgres 后端服务器的连接。一个应用程序可以同时打开多个后端连接（原因之一是需要访问多个数据库）。每个连接由一个 PGconn 对象表示，该对象通过 PQconnectdb() 或 PQsetdbLogin() 获得。注意，除非内存太少以至于连 PGconn 对象都无法分配，这些函数总是返回一个非空的对象指针。在通过连接对象发送查询之前，应当调用 PQstatus 函数检查连接是否成功建立。

- PQsetdbLogin 与后端建立一个新连接。

```
PGconn *PQsetdbLogin(const char *pghost,
                    const char *pgport,
                    const char *pgoptions,
                    const char *pgtty,
                    const char *dbName,
                    const char *login,
                    const char *pwd)
```

如果任何参数为 NULL，则检查相应的环境变量（见“环境变量”一节）。如果环境变量也未设置，则使用硬编码的默认值。返回值是指向一个抽象 struct 的指针，该结构表示到后端的连接。

- PQsetdb 与后端建立一个新连接。

```
PGconn *PQsetdb(char *pghost,
                char *pgport,
                char *pgoptions,
                char *pgtty,
                char *dbName)
```

这是一个宏，以空指针作为 login 和 pwd 参数调用 PQsetdbLogin()。它主要是为了与旧程序保持向后兼容而保留。

- PQconnectdb 与后端建立一个新连接。

```
PGconn *PQconnectdb(const char *conninfo)
```

此例程使用从一个字符串中取得的参数打开一个新的数据库连接。与 PQsetdbLogin() 不同，该例程的参数集可以在不改变函数签名的情况下扩展，

因此新的应用程序编程鼓励使用此例程。传入的字符串可以为空以使用全部默认参数，也可以包含一个或多个用空白分隔的参数设置。每个参数设置的形式为 `keyword = value`。（要写入空值或包含空格的值，请用单引号包围，例如 `keyword = 'a value'`。值中的单引号必须写作 `\'`。等号两边的空格是可选的。）目前可识别的参数关键字有：

- `host` -- 要连接的主机。如果指定了非零长度的字符串，则使用 TCP/IP 通信。没有主机名时，libpq 将使用本地 Unix 域套接字连接。
  - `port` -- 服务器主机上要连接的端口号，或 Unix 域连接的套接字文件名扩展。
  - `dbname` -- 数据库名。
  - `user` -- 用于认证的用户名。
  - `password` -- 当后端要求密码认证时使用的密码。
  - `authtype` -- 授权类型。（已不再使用，因为现在由后端自行选择如何对用户进行认证。为了向后兼容，libpq 仍接受并忽略此关键字。）
  - `options` -- 要发送给后端的跟踪/调试选项。
  - `tty` -- 用于后端可选调试输出的文件或 tty。
- 与 PQsetdbLogin 一样，PQconnectdb 对未指定的选项使用环境变量或内建默认值。
- `PQconnndefaults` 返回默认的连接选项。

```
PQconninfoOption *PQconnndefaults(void)
```

```
struct PQconninfoOption
{
    char    *keyword;    /* The keyword of the option */
    char    *envvar;     /* Fallback environment variable
name */
    char    *compiled;   /* Fallback compiled in default
value */
    char    *val;        /* Option's value */
    char    *label;      /* Label for field in connect
dialog */
    char    *dispchar;   /* Character to display for this
field
                        in a connect dialog. Values
are:
value as is           ""          Display entered
hide value            "*"         Password field -
don't                 "D"         Debug options -
                        create a field by default */
    int     dispsize;    /* Field size in characters for
dialog */
};
```

返回连接选项结构的地址。它可用来确定 PQconnectdb 的所有可用选项及其当前默认值。返回指向一个 PQconninfoOption struct 数组，该数组以一个 keyword 指针为 NULL 的条目结

尾。注意，默认值（"val" 字段）依赖于环境变量和其他上下文。调用者必须将连接选项数据视为只读。

- `PQfinish` 关闭与后端的连接，同时释放 `PGconn` 对象使用的内存。

```
void PQfinish(PGconn *conn)
```

注意，即使与后端的连接尝试失败（由 `PQstatus` 指示），应用也应调用 `PQfinish` 释放 `PGconn` 对象使用的内存。调用过 `PQfinish` 之后，不应再使用该 `PGconn` 指针。

- `PQreset` 重置与后端的通信端口。

```
void PQreset(PGconn *conn)
```

此函数将关闭与后端的连接，并尝试使用先前使用的全部相同参数与同一个 `postmaster` 重新建立新连接。如果工作中的连接丢失，这可用于错误恢复。

`libpq` 应用程序员应注意维护 `PGconn` 的抽象。请使用下面的访问函数获取 `PGconn` 的内容。避免直接引用 `PGconn` 结构的字段，因为它们将来可能改变。（从 `Postgres 6.4` 版开始，`libpq-fe.h` 中甚至不再提供 `struct PGconn` 的定义。如果你有直接访问 `PGconn` 字段的旧代码，可以通过同时包含 `libpq-int.h` 继续使用它，但我们建议你尽快修改这些代码。）

- `PQdb` 返回连接的数据库名。

```
char *PQdb(PGconn *conn)
```

`PQdb` 及其后面几个函数返回在建立连接时确定的值。这些值在 `PGconn` 对象的整个生命周期内保持不变。

- `PQuser` 返回连接的用户名。

```
char *PQuser(PGconn *conn)
```

- `PQpass` 返回连接的密码。

```
char *PQpass(PGconn *conn)
```

- `PQhost` 返回连接的服务器主机名。

```
char *PQhost(PGconn *conn)
```

- `PQport` 返回连接的端口。

```
char *PQport(PGconn *conn)
```

- `PQtty` 返回连接的调试 `tty`。

```
char *PQtty(PGconn *conn)
```

- `PQoptions` 返回连接中使用的后端选项。

```
char *PQoptions(PGconn *conn)
```

- `PQstatus` 返回连接的状态。状态可以是 `CONNECTION_OK` 或 `CONNECTION_BAD`。

```
ConnStatusType *PQstatus(PGconn *conn)
```

失败的连接尝试由状态 `CONNECTION_BAD` 表示。通常，`OK` 状态会一直保持到调用 `PQfinish` 为止，但通信故障可能导致状态提前变为 `CONNECTION_BAD`。此时应用可以尝试调用 `PQreset` 来恢复。

- `PQerrorMessage` 返回连接上最近一次操作所产生的错误消息。

```
char *PQerrorMessage(PGconn* conn);
```

几乎所有 libpq 函数在失败时都会设置 `PQerrorMessage`。注意，按照 libpq 的惯例，非空的 `PQerrorMessage` 会包含一个末尾换行符。

- `PQbackendPID` 返回处理此连接的后端服务器的进程 ID。

```
int PQbackendPID(PGconn *conn);
```

后端 PID 可用于调试，也可用于与 `NOTIFY` 消息（其中包含发出通知的后端的 PID）进行比较。注意，该 PID 属于数据库服务器主机上执行的进程，而不是本地主机上的进程！

## 53.2. 查询执行函数

与数据库服务器的连接成功建立后，此处描述的函数用于执行 SQL 查询和命令。

- `PQexec` 向 Postgres 提交一个查询并等待结果。

```
PGresult *PQexec(PGconn *conn,
                 const char *query);
```

返回一个 `PGresult` 指针，也可能返回 `NULL` 指针。除内存耗尽或无法将查询发送到后端等严重错误外，一般会返回非 `NULL` 指针。如果返回了 `NULL`，应将其视同 `PGRES_FATAL_ERROR` 结果。使用 `PQerrorMessage` 获取该错误的更多信息。

`PGresult` 结构封装了后端返回的查询结果。libpq 应用程序员应注意维护 `PGresult` 的抽象。请使用下面的访问函数获取 `PGresult` 的内容。避免直接引用 `PGresult` 结构的字段，因为它们将来可能改变。（从 Postgres 6.4 版开始，libpq-fe.h 中甚至不再提供 `struct PGresult` 的定义。如果你有直接访问 `PGresult` 字段的旧代码，可以通过同时包含 `libpq-int.h` 继续使用它，但我们建议你尽快修改这些代码。）

- `PQresultStatus` 返回查询的结果状态。`PQresultStatus` 可以返回下列值之一：

```
PGRES_EMPTY_QUERY,
PGRES_COMMAND_OK,      /* the query was a command returning no data
*/
PGRES_TUPLES_OK,       /* the query successfully returned tuples */
PGRES_COPY_OUT,        /* Copy Out (from server) data transfer
started */
PGRES_COPY_IN,         /* Copy In (to server) data transfer started
*/
PGRES_BAD_RESPONSE,    /* an unexpected response was received */
PGRES_NONFATAL_ERROR,
PGRES_FATAL_ERROR
```

如果结果状态为 `PGRES_TUPLES_OK`，就可以使用下面描述的例程检索查询返回的元组。注意，碰巧检索到零个元组的 `SELECT` 仍然显示 `PGRES_TUPLES_OK`。`PGRES_COMMAND_OK` 用于永远不会返回元组的命令。

- `PQresultErrorMessage` 返回与查询相关联的错误消息；如果没有错误则返回空字符串。

```
const char *PQresultErrorMessage(PGresult *res);
```

在 `PQexec` 或 `PQgetResult` 调用之后紧接着，（连接上的）`PQerrorMessage` 会返回与（结果上的）`PQresultErrorMessage` 相同的字符串。但 `PGresult` 会保留其错误消息直到被销毁，而连接的错误消息会随后续操作而变化。想知道与某个特定 `PGresult` 相关联的状态时使用 `PQresultErrorMessage`；想知道连接上最新操作的状态时使用 `PQerrorMessage`。

- `PQntuples` 返回查询结果中的元组（实例）数。

```
int PQntuples(PGresult *res);
```

- `PQnfields` 返回查询结果中每个元组的字段（属性）数。

```
int PQnfields(PGresult *res);
```

- `PQbinaryTuples` 如果 `PGresult` 包含二进制元组数据则返回 1，包含 ASCII 数据则返回 0。

```
int PQbinaryTuples(PGresult *res);
```

目前，只有从 BINARY 游标提取数据的查询才能返回二进制元组数据。

- `PQfname` 返回与给定字段编号相关联的字段（属性）名。字段编号从 0 开始。

```
char *PQfname(PGresult *res,  
              int field_index);
```

- `PQfnumber` 返回与给定字段名相关联的字段（属性）编号。

```
int PQfnumber(PGresult *res,  
              char* field_name);
```

如果给定的名称不匹配任何字段，返回 -1。

- `PQftype` 返回与给定字段编号相关联的字段类型。返回的整数是该类型的内部编码。字段编号从 0 开始。

```
Oid PQftype(PGresult *res,  
            int field_num);
```

- `PQfsize` 返回与给定字段编号相关联的字段的大小，以字节计。字段编号从 0 开始。

```
int PQfsize(PGresult *res,  
            int field_index);
```

`PQfsize` 返回数据库元组中为该字段分配的空间，换句话说，即服务器对该数据类型的二进制表示的大小。如果字段是变长的则返回 -1。

- `PQfmod` 返回与给定字段编号相关联的字段类型专属修饰数据。字段编号从 0 开始。

```
int PQfmod(PGresult *res,  
           int field_index);
```

- `PQgetvalue` 返回 `PGresult` 中某个元组的单个字段（属性）值。元组和字段编号都从 0 开始。

```
char* PQgetvalue(PGresult *res,  
                 int tup_num,  
                 int field_num);
```

对大多数查询而言，PQgetvalue 返回的是属性值的以空字符结尾的 ASCII 字符串表示。但如果 PQbinaryTuples() 为 TRUE，PQgetvalue 返回的就是该类型以后端服务器内部格式表示的二进制表示（若字段为变长则不含大小字）。此时由程序员负责把数据转换成正确的 C 类型。PQgetvalue 返回的指针指向属于 PGresult 结构的存储空间。不应修改它；如果需要在 PGresult 结构的生命周期结束后继续使用该值，就必须显式地将它复制到其他存储空间。

- PQgetlength 返回字段（属性）的长度，以字节计。元组和字段编号从 0 开始。

```
int PQgetlength(PGresult *res,
                int tup_num,
                int field_num);
```

这是该特定数据值的实际数据长度，即 PQgetvalue 所指对象的大小。注意，对于以 ASCII 表示的值，此大小与 PQsize 报告的二进制大小几乎没有关系。

- PQgetisnull 测试一个字段是否为 NULL 条目。元组和字段编号从 0 开始。

```
int PQgetisnull(PGresult *res,
                int tup_num,
                int field_num);
```

如果字段包含 NULL，此函数返回 1；包含非 NULL 值则返回 0。（注意，对于 NULL 字段，PQgetvalue 返回的是空字符串而不是空指针。）

- PQcmdStatus 返回生成该 PGresult 的 SQL 命令的命令状态字符串。

```
char *PQcmdStatus(PGresult *res);
```

- PQcmdTuples 返回受 SQL 命令影响的行数。

```
const char *PQcmdTuples(PGresult *res);
```

如果生成该 PGresult 的 SQL 命令是 INSERT、UPDATE 或 DELETE，此函数返回一个包含受影响行数的字符串。如果是其他命令，返回空字符串。

- PQoidStatus 如果 SQL 命令是一个 INSERT，返回一个含有所插入元组对象 id 的字符串。否则返回空字符串。

```
char* PQoidStatus(PGresult *res);
```

- PQprint 打印出所有的元组以及（可选的）属性名到指定的输出流。

```
void PQprint(FILE* fout,          /* output stream */
             PGresult* res,
             PQprintOpt* po);

struct _PQprintOpt
{
    pqbool header;          /* print output field headings
and row count */
    pqbool align;           /* fill align the fields */
    pqbool standard;        /* old brain dead format */
    pqbool html3;           /* output html tables */
    pqbool expanded;        /* expand tables */
    pqbool pager;           /* use pager for output if
needed */
}
```

```
char    *fieldSep;    /* field separator */
char    *tableOpt;    /* insert to HTML <table ...>
*/
char    *caption;     /* HTML <caption> */
char    **fieldName; /* null terminated array of
replacement field names */
};
```

此函数旨在取代现已过时的 PQprintTuples()。psql 程序使用 PQprint() 显示查询结果。

- PQprintTuples 打印出所有的元组以及（可选的）属性名到指定的输出流。

```
void PQprintTuples(PGresult* res,
FILE* fout,          /* output stream */
int printAttName, /* print attribute names or
not*/
int terseOutput, /* delimiter bars or not?*/
int width);        /* width of column, variable
width if 0*/
```

- PQdisplayTuples 打印出所有的元组以及（可选的）属性名到指定的输出流。

```
void PQdisplayTuples(PGresult* res,
FILE* fout,          /* output stream */
int fillAlign,        /* space fill to align
columns */
const char *fieldSep, /* field separator */
int printHeader,      /* display headers? */
int quiet);           /* suppress print of row
count at end */
```

PQdisplayTuples() 本意是取代 PQprintTuples(), 而它又被 PQprint() 取代。

- PQclear 释放与 PGresult 关联的存储。每个查询结果在不再需要时都应通过 PQclear 释放。

```
void PQclear(PGresult *res);
```

PGresult 对象需要保留多久就可以保留多久；发出新查询时它不会消失，即使关闭连接也不会。要清除它，必须调用 PQclear。不这样做将导致前端应用内存泄漏。

- PQmakeEmptyPGresult 以给定的状态构造一个空的 PGresult 对象。

```
PGresult* PQmakeEmptyPGresult(PGconn *conn, ExecStatusType status);
```

这是 libpq 分配并初始化空 PGresult 对象的内部例程。导出它是因为某些应用发现自己生成结果对象（特别是带有错误状态的对象）很有用。如果 conn 非 NULL 且 status 指示一个错误，连接的当前 errorMessage 会被复制到 PGresult。注意，最终也应对该对象调用 PQclear，就像对待 libpq 本身返回的 PGresult 一样。

## 53.3. 异步查询处理

PQexec 函数足以在简单的同步应用中提交查询。但它有几个主要缺陷：

- PQexec 会等待查询完成。应用可能有其他工作要做（例如维护一个用户界面），这种情况下它不会想要阻塞等待响应。

- 由于控制权埋在 PQexec 内部，前端很难在需要时决定尝试取消正在进行的查询。（这可以在信号处理器中完成，但别无他法。）
- PQexec 只能返回一个 PGresult 结构。如果提交的查询串包含多条 SQL 命令，除了最后一个 PGresult 之外都会被 PQexec 丢弃。

不喜欢这些限制的应用可以改用构成 PQexec 的底层函数：PQsendQuery 和 PQgetResult。

- PQsendQuery 向 Postgres 提交一个查询而不等待结果。查询成功发出返回 TRUE，否则返回 FALSE（此时用 PQerrorMessage 获取失败的更多信息）。

```
int PQsendQuery(PGconn *conn,
               const char *query);
```

成功调用 PQsendQuery 后，调用一次或多次 PQgetResult 获取查询结果。在 PQgetResult 返回 NULL（表示查询已完成）之前，不得（在同一连接上）再次调用 PQsendQuery。

- PQgetResult 等待先前一次 PQsendQuery 的下一个结果并将其返回。查询完成且不再有结果时返回 NULL。

```
PGresult *PQgetResult(PGconn *conn);
```

必须反复调用 PQgetResult 直到它返回 NULL，表示查询已完成。（如果在没有活跃查询时调用，PQgetResult 会立即返回 NULL。）对 PQgetResult 返回的每个非空结果，应使用前文所述的同一批 PGresult 访问函数处理。用完每个结果对象后别忘了用 PQclear 释放。注意，只有当查询处于活跃状态且必要的响应数据尚未被 PQconsumeInput 读取时，PQgetResult 才会阻塞。

使用 PQsendQuery 和 PQgetResult 解决了 PQexec 的一个问题：如果查询字符串包含多条 SQL 命令，这些命令的结果可以分别获取。（顺便说一句，这也支持一种简单的重叠处理：前端可以处理某条查询的结果，同时后端继续处理同一查询串中后面的查询。）但调用 PQgetResult 仍会使前端阻塞，直到后端完成下一条 SQL 命令。恰当使用另外三个函数可以避免这一点：

- PQconsumeInput 如果后端有可用的输入，消费它。

```
int PQconsumeInput(PGconn *conn);
```

PQconsumeInput 正常返回 1 表示"no error"（没有错误）；发生时返回 0（此时 PQerrorMessage 会被设置）。注意，返回结果并不说明是否实际读取了输入数据。调用 PQconsumeInput 后，应用可以检查 PQisBusy 和/或 PQnotifies，看它们的状态是否发生了变化。即使应用尚未准备好处理结果或通知，也可以调用 PQconsumeInput。该例程会读取可用数据并保存到缓冲区中，从而使 select(2) 的可读就绪指示消失。因此应用可以用 PQconsumeInput 立即清除 select 的就绪条件，随后再从从容地检查结果。

- PQisBusy 如果查询仍在忙碌则返回 TRUE，也就是说，PQgetResult 会阻塞等待输入。返回 FALSE 表示可以放心调用 PQgetResult 而不会阻塞。

```
int PQisBusy(PGconn *conn);
```

PQisBusy 本身不会尝试从后端读取数据；因此必须先调用 PQconsumeInput，否则忙碌状态永远不会结束。

- PQsocket 获取后端连接套接字的文件描述符号。有效的描述符将  $\geq 0$ ；结果为 -1 表示当前没有打开的后端连接。

```
int PQsocket(PGconn *conn);
```

应当用 `PQsocket` 获取后端套接字描述符，为执行 `select(2)` 做准备。这使应用可以同时等待后端响应或其他条件。如果 `select(2)` 的结果表明可以从后端套接字读取数据，就应调用 `PQconsumeInput` 读取数据；随后即可用 `PQisBusy`、`PQgetResult` 和/或 `PQnotifies` 处理响应。

使用这些函数的典型前端会有一个主循环，用 `select(2)` 等待它必须响应的所有条件。其中一个条件是后端有可读的输入；用 `select` 的话说，就是由 `PQsocket` 标识的文件描述符上有可读数据。主循环检测到输入就绪时，应调用 `PQconsumeInput` 读取输入。然后可以调用 `PQisBusy`，若 `PQisBusy` 返回 `FALSE`，接着调用 `PQgetResult`。还可以调用 `PQnotifies` 检测 `NOTIFY` 消息（见下面的“异步通知”）。示例程序一节给出了一个例子。

使用 `PQsendQuery/PQgetResult` 的前端还可以尝试取消一个仍随后端处理中的查询。

- `PQrequestCancel` 请求 Postgres 放弃对当前查询的处理。

```
int PQrequestCancel(PGconn *conn);
```

取消请求成功发出则返回 `TRUE`，否则返回 `FALSE`。（若为 `FALSE`，`PQerrorMessage` 会说明原因。）但成功发出并不能保证请求会生效。无论 `PQrequestCancel` 的返回值如何，应用都必须继续用 `PQgetResult` 按正常顺序读取结果。如果取消生效，当前查询会提前终止并返回一个错误结果。如果取消失败（例如因为后端已完成该查询的处理），则根本不会有可见的结果。

注意，如果当前查询是事务的一部分，取消将中止整个事务。

`PQrequestCancel` 可以在信号处理器中安全地调用。因此，如果取消的决定可以在信号处理器中做出，也可以将它与普通的 `PQexec` 配合使用。例如，`psql` 从 `SIGINT` 信号处理器中调用 `PQrequestCancel`，从而允许交互式地取消它通过 `PQexec` 发出的查询。注意，如果连接当前未打开，`PQrequestCancel` 将不起作用或后端当前没有在处理查询。

## 53.4. 快速路径

Postgres 提供一个快速路径接口来向后端发送函数调用。这是通向系统内部的一个天窗，可能成为潜在的安全漏洞。大多数用户不需要此特性。

- `PQfn` 通过快速路径接口请求执行一个后端函数。

```
PGresult* PQfn(PGconn* conn,
               int fnid,
               int *result_buf,
               int *result_len,
               int result_is_int,
               PQArgBlock *args,
               int nargs);
```

`fnid` 参数是要执行的函数的对象标识符。`result_buf` 是存放返回值的缓冲区。调用者必须已分配足够的空间来存储返回值（这里不做检查！）。实际结果长度将返回到 `result_len` 所指向的整数中。如果预期结果是 4 字节整数，将 `result_is_int` 设为 1；否则设为 0。（将 `result_is_int` 设为 1 会告诉 `libpq` 在必要时做字节序交换，使值以适合客户端机器的 `int` 值送达。当 `result_is_int` 为 0 时，后端发送的字节串原样返回。）`args` 和 `nargs` 指定要传给函数的参数。

```
typedef struct {
    int len;
    int isint;
    union {
        int *ptr;
```

```
        int integer;
    } u;
} PQArgBlock;
```

PQfn 总是返回有效的 PGresult\*。使用结果前应检查 resultStatus。不再需要结果时，调用者负责用 PQclear 释放 PGresult。

## 53.5. 异步通知

Postgres 通过 LISTEN 和 NOTIFY 命令支持异步通知。后端用 LISTEN 命令注册它对某个特定通知条件的兴趣（并可以用 UNLISTEN 命令停止监听）。当任何后端执行带该条件名的 NOTIFY 命令时，所有监听该条件的后端都会被异步通知。通知者不会向监听者传递任何附加信息。因此，通常需要通信的任何实际数据都通过一个数据库关系传递。条件名通常与相关的关系同名，但并不要求必须有相关的关系。

libpq 应用把 LISTEN 和 UNLISTEN 命令作为普通 SQL 查询提交。随后，可以通过调用 PQnotifies() 检测 NOTIFY 消息的到达。

- PQnotifies 从后端发来的尚未处理的通知消息列表中返回下一条通知。没有待处理的通知时返回 NULL。一条通知被 PQnotifies 返回后即视为已处理，并会从通知列表中移除。

```
PGnotify* PQnotifies(PGconn *conn);

typedef struct pgNotify
{
    char            relname[NAMEDATALEN];        /* name of relation
                                                    * containing data
    */
    int             be_pid;                      /* process id of
    backend */
} PGnotify;
```

处理完 PQnotifies 返回的 PGnotify 对象后，务必用 free() 释放它，以避免内存泄漏。注意：在 Postgres 6.4 及之后的版本中，be\_pid 是发出通知的后端的 PID，而在较早的版本中它总是你自己的后端的 PID。

第二个示例程序演示了异步通知的使用。

PQnotifies() 并不真正读取后端数据；它只是返回先前被其他 libpq 函数吸收的消息。在 libpq 的早期版本中，确保及时收到 NOTIFY 消息的唯一方法是不断地提交查询（哪怕是空查询），然后在每次 PQexec() 之后检查 PQnotifies()。这种方法虽然仍然有效，但由于浪费处理能力已被弃用。在没有有用查询可提交时检查 NOTIFY 消息的更好办法是调用 PQconsumeInput()，然后检查 PQnotifies()。可以用 select(2) 等待后端数据到达，这样除非有事可做就不会消耗 CPU。注意，无论查询用的是 PQsendQuery/PQgetResult 还是普通的旧式 PQexec，这样做都没有问题。但应记住在每次 PQgetResult 或 PQexec 之后检查 PQnotifies()，看查询处理期间是否有通知到来。

## 53.6. 与 COPY 命令相关的函数

Postgres 中的 COPY 命令可以选择从 libpq 所用的网络连接读取或向其写入。因此，需要能直接访问此网络连接的函数，应用才能利用这一能力。

只有在从 PQexec 或 PQgetResult 得到 PGRES\_COPY\_OUT 或 PGRES\_COPY\_IN 结果对象之后，才应执行这些函数。

- `PQgetline` 把一行由后端服务器传来的以换行符结尾的字符读入大小为 `length` 的缓冲区字符串。

```
int PQgetline(PGconn *conn,
              char *string,
              int length)
```

与 `fgets(3)` 一样，此例程最多把 `length-1` 个字符复制到 `string` 中；但与 `gets(3)` 相似，它会把结尾的换行符转换为空字符。`PQgetline` 在 EOF 处返回 EOF，整行已读完返回 0，缓冲区已满但结尾换行符尚未读到。注意，应用必须检查新的一行是否由两个字符 `"\."` 组成，这表示后端服务器已发送完 `copy` 命令的结果。如果应用可能收到超过 `length-1` 个字符的行，就需要小心确保正确识别 `"\."` 行（例如，不要把长数据行的结尾误当作终结行）。../src/bin/psql/psql.c 中的代码包含正确处理 `copy` 协议的例程。

- `PQgetlineAsync` 把一行由后端服务器传来的以换行符结尾的字符读入缓冲区而不阻塞。

```
int PQgetlineAsync(PGconn *conn,
                  char *buffer,
                  int bufsize)
```

此例程类似于 `PQgetline`，但可供必须异步（即不阻塞地）读取 `COPY` 数据的应用使用。发出 `COPY` 命令并得到 `PGRES_COPY_OUT` 响应后，应用应交替调用 `PQconsumeInput` 和 `PQgetlineAsync`，直到检测到数据结束信号。与 `PQgetline` 不同，此例程自己负责检测数据结束。每次调用时，如果 `libpq` 的输入缓冲区中已有完整的一行以换行符结尾的数据行，或者到来的数据行太长放不进调用者提供的缓冲区，`PQgetlineAsync` 就会返回数据。否则，在该行剩余部分到达之前不返回数据。如果已识别到 `copy` 数据结束标记，此例程返回 -1；没有可用数据返回 0；返回正数时该数给出所返回数据的字节数。返回 -1 时，调用者接下来必须调用 `PQendcopy`，然后回到正常处理。返回的数据不会跨过换行符。可能时一次返回整行；但如果调用者提供的缓冲区太小装不下后端发来的一行，就会返回部分数据行。这可以通过检查最后返回的字节是否为 `'\n'` 来检测。返回的字符串不以空字符结尾。（如果想加上结尾空字符，务必让传入的 `bufsize` 比实际可用空间小一。）

- `PQputline` 向后台服务器发送一个以空字符结尾的字符串。成功返回 0，无法发送字符串返回 EOF。

```
int PQputline(PGconn *conn,
              char *string);
```

注意，应用必须在最后一行显式发送两个字符 `"\."`，以告知后端它已发送完数据。

- `PQputnbytes` 向后台服务器发送一个不以空字符结尾的字符串。成功返回 0，无法发送字符串返回 EOF。

```
int PQputnbytes(PGconn *conn,
                const char *buffer,
                int nbytes);
```

它与 `PQputline` 完全一样，只是由于要发送的字节数是直接指定的，数据缓冲区不需要以空字符结尾。

- `PQendcopy` 与后端同步。此函数等待后端完成 `copy`。它应当在用 `PQputline` 向后端发送完最后一个字符串之后，或用 `PQgetline` 从后端收到最后一个字符串之后发出。必须发出此函数，否则后端可能与前端 `"out of sync"`（失去同步）。从此函数返回后，后端就准备好接收下一条查询了。成功完成时返回值为 0，否则非零。

```
int PQendcopy(PGconn *conn);
```

举一个例子：

```
PQexec(conn, "create table foo (a int4, b char16, d float8)");
PQexec(conn, "copy foo from stdin");
PQputline(conn, "3<TAB>hello world<TAB>4.5\n");
PQputline(conn, "4<TAB>goodbye world<TAB>7.11\n");
...
PQputline(conn, "\\.\n");
PQendcopy(conn);
```

使用 `PQgetResult` 时，应用对 `PGRES_COPY_OUT` 结果的响应应是反复执行 `PQgetline`，看到终结行后再执行 `PQendcopy`。然后应回到 `PQgetResult` 循环，直到 `PQgetResult` 返回 `NULL`。类似地，`PGRES_COPY_IN` 结果由一系列 `PQputline` 调用加 `PQendcopy` 处理，然后回到 `PQgetResult` 循环。这样的安排可确保嵌入在一系列 SQL 命令中的 `copy in` 或 `copy out` 命令被正确执行。较老的应用可能通过 `PQexec` 提交 `copy in` 或 `copy out`，并认为 `PQendcopy` 之后事务就完成了。只有当 `copy in/out` 是查询字符串中唯一的 SQL 命令时，这样做才是正确的。

## 53.7. libpq 跟踪函数

- `PQtrace` 把前端/后端通信的跟踪打开到调试文件流。

```
void PQtrace(PGconn *conn
             FILE *debug_port)
```

- `PQuntrace` 关闭由 `PQtrace` 启动的跟踪。

```
void PQuntrace(PGconn *conn)
```

## 53.8. libpq 控制函数

- `PQsetNoticeProcessor` 控制 libpq 产生的通知和警告消息的报告方式。

```
void PQsetNoticeProcessor (PGconn * conn,
                           void (*noticeProcessor) (void * arg, const char * message),
                           void * arg)
```

默认情况下，libpq 把来自后端的“notice”消息以及它自己产生的少量错误消息打印到 `stderr`。可以通过提供一个对这些消息另做处理的回调函数来改变这一行为。

回调函数会收到错误消息的文本（包括末尾换行符），外加一个与传给 `PQsetNoticeProcessor` 的指针相同的 `void` 指针。（需要时可以用该指针访问应用特定的状态。）默认的通知处理器只是

```
static void
defaultNoticeProcessor(void * arg, const char * message)
{
    fprintf(stderr, "%s", message);
}
```

要使用特定的通知处理器，只需在创建新的 `PGconn` 对象之后立即调用 `PQsetNoticeProcessor`。

## 53.9. 用户认证函数

前端/后端认证过程由 `PQconnectdb` 处理，无需任何进一步的干预。认证方法现在完全由 DBA 决定（见 `pga_hba.conf(5)`）。下面的例程已不再起任何作用，不应再使用。

- `fe_getauthname` 返回一个指向静态空间的指针，其中包含用户经过认证所用的名称。强烈建议应用用此例程代替对 `getenv(3)` 或 `getpwuid(3)` 的调用，因为经过认证的用户名完全有可能与 `USER` 环境变量的值或用户在 `/etc/passwd` 中的条目不同。

```
char *fe_getauthname(char* errorMessage)
```

- `fe_setauthsvc` 指定 `libpq` 应使用认证服务 `name` 而不是其编译期默认值。此值通常取自命令行开关。

```
void fe_setauthsvc(char *name,  
                  char* errorMessage)
```

认证尝试的任何错误消息都会返回到 `errorMessage` 参数中。

## 53.10. 环境变量

下面的环境变量可用于选择默认的连接参数值；当调用代码没有直接指定值时，`PQconnectdb` 或 `PQsetdbLogin` 将使用这些值。它们有助于避免把数据库名硬编码进简单的应用程序。

- `PGHOST` 设置默认的服务器名。如果指定了非零长度的字符串，则使用 TCP/IP 通信。没有主机名时，`libpq` 将使用本地 Unix 域套接字连接。
- `PGPORT` 设置与 Postgres 后端通信所用的默认端口或本地 Unix 域套接字文件扩展名。
- `PGDATABASE` 设置默认的 Postgres 数据库名。
- `PGUSER` 设置连接数据库和认证所用的用户名。
- `PGPASSWORD` 设置当后端要求密码认证时使用的密码。
- `PGREALM` 设置与 Postgres 一起使用的 Kerberos realm（当它与本地 realm 不同时）。如果设置了 `PGREALM`，Postgres 应用将尝试对该 realm 中的服务器进行认证，并使用单独的 ticket 文件以避免与本地 ticket 文件冲突。仅当后端选择了 Kerberos 认证时才会使用此环境变量。
- `PGOPTIONS` 设置 Postgres 后端的附加运行时选项。
- `PGTTY` 设置显示来自后端服务器的调试消息的文件或 tty。

下面的环境变量可用于为每个 Postgres 会话指定用户级的默认行为：

- `PGDATESTYLE` 设置日期/时间表示的默认风格。
- `PGTZ` 设置默认时区。

下面的环境变量可用于为每个 Postgres 会话指定默认的内部行为：

- `PGGEQO` 设置遗传优化器的默认模式。
- `PGRPLANS` 设置优化器中允许或禁用右平面（right-sided plans）的默认模式。

- PG\_COST\_HEAP 设置优化器进行堆搜索的默认代价。
- PG\_COST\_INDEX 设置优化器进行索引搜索的默认代价。
- PG\_QUERY\_LIMIT 设置查询返回的最大行数。

关于这些环境变量的正确取值，请参阅 `SET SQL` 命令。

## 53.11. 注意事项

查询缓冲区长 8192 字节，超过该长度的查询将被拒绝。

## 53.12. 示例程序

### 53.12.1. 示例程序 1

```
/*
 * testlibpq.c Test the C version of Libpq, the Postgres frontend
 * library.
 *
 */
#include <stdio.h>
#include "libpq-fe.h"

void
exit_nicely(PGconn *conn)
{
    PQfinish(conn);
    exit(1);
}

main()
{
    char        *pgghost,
                *pgport,
                *pgoptions,
                *pgtty;
    char        *dbName;
    int          nFields;
    int          i,
                j;

    /* FILE *debug; */

    PGconn      *conn;
    PGresult     *res;

    /*
     * begin, by setting the parameters for a backend connection if
     the
     * parameters are null, then the system will try to use reasonable

```

```
    * defaults by looking up environment variables or, failing that,
    * using hardwired constants
    */
    pghost = NULL;                /* host name of the backend server */
    pgport = NULL;                /* port of the backend server */
    pgoptions = NULL;             /* special options to start up the
backend                                server */

    pgtty = NULL;                 /* debugging tty for the backend
server */
    dbName = "template1";

    /* make a connection to the database */
    conn = PQsetdb(pghost, pgport, pgoptions, pgtty, dbName);

    /*
     * check to see that the backend connection was successfully made
     */
    if (PQstatus(conn) == CONNECTION_BAD)
    {
        fprintf(stderr, "Connection to database '%s' failed.\n", dbName);
        fprintf(stderr, "%s", PQerrorMessage(conn));
        exit_nicely(conn);
    }

    /* debug = fopen("/tmp/trace.out", "w"); */
    /* PQtrace(conn, debug); */

    /* start a transaction block */
    res = PQexec(conn, "BEGIN");
    if (PQresultStatus(res) != PGRES_COMMAND_OK)
    {
        fprintf(stderr, "BEGIN command failed\n");
        PQclear(res);
        exit_nicely(conn);
    }

    /*
     * should PQclear PGresult whenever it is no longer needed to
avoid
     * memory leaks
     */
    PQclear(res);

    /*
     * fetch instances from the pg_database, the system catalog of
     * databases
     */
    res = PQexec(conn, "DECLARE mycursor CURSOR FOR select * from pg_
database");
    if (PQresultStatus(res) != PGRES_COMMAND_OK)
    {
        fprintf(stderr, "DECLARE CURSOR command failed\n");
    }
}
```

```
        PQclear(res);
        exit_nicely(conn);
    }
    PQclear(res);
    res = PQexec(conn, "FETCH ALL in mycursor");
    if (PQresultStatus(res) != PGRES_TUPLES_OK)
    {
        fprintf(stderr, "FETCH ALL command didn't return tuples
properly\n");
        PQclear(res);
        exit_nicely(conn);
    }

    /* first, print out the attribute names */
    nFields = PQnfields(res);
    for (i = 0; i < nFields; i++)
        printf("%-15s", PQfname(res, i));
    printf("\n\n");

    /* next, print out the instances */
    for (i = 0; i < PQntuples(res); i++)
    {
        for (j = 0; j < nFields; j++)
            printf("%-15s", PQgetvalue(res, i, j));
        printf("\n");
    }
    PQclear(res);

    /* close the cursor */
    res = PQexec(conn, "CLOSE mycursor");
    PQclear(res);

    /* commit the transaction */
    res = PQexec(conn, "COMMIT");
    PQclear(res);

    /* close the connection to the database and cleanup */
    PQfinish(conn);

    /* fclose(debug); */
}
```

## 53.12.2. 示例程序 2

```
/*
 * testlibpq2.c Test of the asynchronous notification interface
 *
 * populate a database with the following:
 *
 * CREATE TABLE TBL1 (i int4);
 *
 * CREATE TABLE TBL2 (i int4);
 *
```

```
* CREATE RULE r1 AS ON INSERT TO TBL1 DO [INSERT INTO TBL2 values
* (new.i); NOTIFY TBL2];
*
* Then start up this program After the program has begun, do
*
* INSERT INTO TBL1 values (10);
*
*
*
*/
#include <stdio.h>
#include "libpq-fe.h"

void
exit_nicely(PGconn *conn)
{
    PQfinish(conn);
    exit(1);
}

main()
{
    char        *pgghost,
               *pgport,
               *pgoptions,
               *pgtty;
    char        *dbName;
    int         nFields;
    int         i,
               j;

    PGconn      *conn;
    PGresult     *res;
    PGnotify     *notify;

    /*
     * begin, by setting the parameters for a backend connection if
the
     * parameters are null, then the system will try to use reasonable
     * defaults by looking up environment variables or, failing that,
     * using hardwired constants
     */
    pgghost = NULL;           /* host name of the backend server */
    pgport = NULL;            /* port of the backend server */
    pgoptions = NULL;         /* special options to start up the
backend
                               * server */
    pgtty = NULL;             /* debugging tty for the backend
server */
    dbName = getenv("USER");  /* change this to the name of your
test
                               * database */

    /* make a connection to the database */
```

```
conn = PQsetdb(pghost, pgport, pgoptions, pgtty, dbName);

/*
 * check to see that the backend connection was successfully made
 */
if (PQstatus(conn) == CONNECTION_BAD)
{
    fprintf(stderr, "Connection to database '%s' failed.\n", dbName);
    fprintf(stderr, "%s", PQerrorMessage(conn));
    exit_nicely(conn);
}

res = PQexec(conn, "LISTEN TBL2");
if (PQresultStatus(res) != PGRES_COMMAND_OK)
{
    fprintf(stderr, "LISTEN command failed\n");
    PQclear(res);
    exit_nicely(conn);
}

/*
 * should PQclear PGresult whenever it is no longer needed to
avoid
 * memory leaks
 */
PQclear(res);

while (1)
{
    /*
     * wait a little bit between checks; waiting with select()
     * would be more efficient.
     */
    sleep(1);
    /* collect any asynchronous backend messages */
    PQconsumeInput(conn);
    /* check for asynchronous notify messages */
    while ((notify = PQnotifies(conn)) != NULL)
    {
        fprintf(stderr,
            "ASYNC NOTIFY of '%s' from backend pid '%d' received\n",
            notify->relname, notify->be_pid);
        free(notify);
    }
}

/* close the connection to the database and cleanup */
PQfinish(conn);
}
```

### 53.12.3. 示例程序 3

```
/*
 * testlibpq3.c Test the C version of Libpq, the Postgres frontend
 * library. tests the binary cursor interface
 *
 *
 *
 * populate a database by doing the following:
 *
 * CREATE TABLE test1 (i int4, d float4, p polygon);
 *
 * INSERT INTO test1 values (1, 3.567, '(3.0, 4.0, 1.0,
 * 2.0)::polygon);
 *
 * INSERT INTO test1 values (2, 89.05, '(4.0, 3.0, 2.0,
 * 1.0)::polygon);
 *
 * the expected output is:
 *
 * tuple 0: got i = (4 bytes) 1, d = (4 bytes) 3.567000, p = (4
 * bytes) 2 points   boundingbox = (hi=3.000000/4.000000, lo =
 * 1.000000,2.000000) tuple 1: got i = (4 bytes) 2, d = (4 bytes)
 * 89.050003, p = (4 bytes) 2 points   boundingbox =
 * (hi=4.000000/3.000000, lo = 2.000000,1.000000)
 *
 */
#include <stdio.h>
#include "libpq-fe.h"
#include "utils/geo-decls.h"    /* for the POLYGON type */

void
exit_nicely(PGconn *conn)
{
    PQfinish(conn);
    exit(1);
}

main()
{
    char        *pghost,
                *pgport,
                *pgoptions,
                *pgtty;

    char        *dbName;
    int         nFields;
    int         i,
                j;
    int         i_fnum,
                d_fnum,
                p_fnum;
    PGconn      *conn;
```

```
PGresult    *res;

/*
 * begin, by setting the parameters for a backend connection if
the
 * parameters are null, then the system will try to use reasonable
 * defaults by looking up environment variables or, failing that,
 * using hardwired constants
 */
pghost = NULL;                /* host name of the backend server */
pgport = NULL;                /* port of the backend server */
pgoptions = NULL;             /* special options to start up the
backend
                                * server */
pgtty = NULL;                 /* debugging tty for the backend
server */

dbName = getenv("USER");      /* change this to the name of your
test
                                * database */

/* make a connection to the database */
conn = PQsetdb(pghost, pgport, pgoptions, pgtty, dbName);

/*
 * check to see that the backend connection was successfully made
 */
if (PQstatus(conn) == CONNECTION_BAD)
{
    fprintf(stderr, "Connection to database '%s' failed.\n", dbName);
    fprintf(stderr, "%s", PQerrorMessage(conn));
    exit_nicely(conn);
}

/* start a transaction block */
res = PQexec(conn, "BEGIN");
if (PQresultStatus(res) != PGRES_COMMAND_OK)
{
    fprintf(stderr, "BEGIN command failed\n");
    PQclear(res);
    exit_nicely(conn);
}

/*
 * should PQclear PGresult whenever it is no longer needed to
avoid
 * memory leaks
 */
PQclear(res);

/*
 * fetch instances from the pg_database, the system catalog of
 * databases
```

```
    */
    res = PQexec(conn, "DECLARE mycursor BINARY CURSOR FOR select *
from test1");
    if (PQresultStatus(res) != PGRES_COMMAND_OK)
    {
        fprintf(stderr, "DECLARE CURSOR command failed\n");
        PQclear(res);
        exit_nicely(conn);
    }
    PQclear(res);

    res = PQexec(conn, "FETCH ALL in mycursor");
    if (PQresultStatus(res) != PGRES_TUPLES_OK)
    {
        fprintf(stderr, "FETCH ALL command didn't return tuples
properly\n");
        PQclear(res);
        exit_nicely(conn);
    }

    i_fnum = PQfnumber(res, "i");
    d_fnum = PQfnumber(res, "d");
    p_fnum = PQfnumber(res, "p");

    for (i = 0; i < 3; i++)
    {
        printf("type[%d] = %d, size[%d] = %d\n",
            i, PQftype(res, i),
            i, PQfsize(res, i));
    }
    for (i = 0; i < PQntuples(res); i++)
    {
        int          *ival;
        float         *dval;
        int           plen;
        POLYGON       *pval;

        /* we hard-wire this to the 3 fields we know about */
        ival = (int *) PQgetvalue(res, i, i_fnum);
        dval = (float *) PQgetvalue(res, i, d_fnum);
        plen = PQgetlength(res, i, p_fnum);

        /*
         * plen doesn't include the length field so need to
         * increment by VARHDRSZ
         */
        pval = (POLYGON *) malloc(plen + VARHDRSZ);
        pval->size = plen;
        memmove((char *) &pval->npts, PQgetvalue(res, i, p_fnum),
plen);
        printf("tuple %d: got\n", i);
        printf(" i = (%d bytes) %d,\n",
            PQgetlength(res, i, i_fnum), *ival);
        printf(" d = (%d bytes) %f,\n",
```

```
        PQgetlength(res, i, d_fnum), *dval);
    printf(" p = (%d bytes) %d points \tboundingbox = (hi=%f/%f, lo =
%f,%f)\n",
        PQgetlength(res, i, d_fnum),
        pval->npts,
        pval->boundingbox.xh,
        pval->boundingbox.yh,
        pval->boundingbox.xl,
        pval->boundingbox.yl);
}
PQclear(res);

/* close the cursor */
res = PQexec(conn, "CLOSE mycursor");
PQclear(res);

/* commit the transaction */
res = PQexec(conn, "COMMIT");
PQclear(res);

/* close the connection to the database and cleanup */
PQfinish(conn);
}
```

---

# 第 54 章 pgctl

pgctl 是一个 tcl 包，供前端程序与 Postgres 后端交互。它使 tcl 脚本可以使用 libpq 的大部分功能。

本包最初由 Jolly Chen 编写。

## 54.1. 命令

表 54.1. pgctl 命令

| 命令              | 描述                |
|-----------------|-------------------|
| pg_connect      | 打开一个到后端服务器的连接     |
| pg_disconnect   | 关闭一个连接            |
| pg_conndefaults | 获取连接选项及其默认值       |
| pg_exec         | 向后端发送一个查询         |
| pg_result       | 操作查询的结果           |
| pg_select       | 循环遍历 SELECT 语句的结果 |
| pg_listen       | 为 NOTIFY 消息建立回调   |
| pg_lo_creat     | 创建一个对象            |
| pg_lo_open      | 打开一个对象            |
| pg_lo_close     | 关闭一个对象            |
| pg_lo_read      | 读取一个对象            |
| pg_lo_write     | 写入一个对象            |
| pg_lo_lseek     | 在对象中定位到某个位置       |
| pg_lo_tell      | 返回对象的当前寻位位置       |
| pg_lo_unlink    | 删除一个对象            |
| pg_lo_import    | 把 Unix 文件导入为对象    |
| pg_lo_export    | 把对象导出到 Unix 文件    |

这些命令将在后续页面中进一步描述。

pg\_lo\* 例程是 Postgres 对象特性的接口。这些函数的设计模仿标准 Unix 文件系统接口中的类似文件系统函数。pg\_lo\* 例程应当在 BEGIN/END 事务块内使用，因为 pg\_lo\_open 返回的文件描述符只在当前事务内有效。pg\_lo\_import 和 pg\_lo\_export 必须在 BEGIN/END 事务块内使用。

## 54.2. 示例

下面是使用这些例程的一个小例子：

```
# getDBs :
#   get the names of all the databases at a given host and port number
#   with the defaults being the localhost and port 5432
#   return them in alphabetical order
proc getDBs { {host "localhost"} {port "5432"} } {
    # datnames is the list to be result
```

```
    set conn [pg_connect template1 -host $host -port $port]
    set res [pg_exec $conn "SELECT datname FROM pg_database ORDER BY
datname"]
    set ntups [pg_result $res -numTuples]
    for {set i 0} {$i < $ntups} {incr i} {
lappend datnames [pg_result $res -getTuple $i]
    }
    pg_result $res -clear
    pg_disconnect $conn
    return $datnames
}
```

## 54.3. pgtcl 命令参考信息

## pg\_connect

pg\_connect — 打开一个到后端服务器的连接

### 大纲

```
pg_connect -conninfo connectOptions
pg_connect dbName [-host hostName]
               [-port portNumber] [-tty pqTTY]
               [-options optionalBackendArgs]
```

#### 输入（新风格）

*connectOptions*

一个连接选项字符串，每个选项都写成 keyword = value 的形式。

#### 输入（旧风格）

*dbName*

指定一个有效的数据库名。

[-host *hostName*]

指定 *dbName* 所在后端服务器的主机名。

[-port *portNumber*]

指定 *dbName* 所在后端服务器的 IP 端口号。

[-tty *pqTTY*]

指定用于后端可选调试输出的文件或 tty。

[-options *optionalBackendArgs*]

指定 *dbName* 所在后端服务器的选项。

#### 输出

*dbHandle*

成功时返回一个数据库连接的句柄。句柄以 "pgsql" 前缀开头。

### 描述

pg\_connect 打开一个到 Postgres 后端的连接。

有两种可用的语法。在较老的一种中，每个可能的选项在 pg\_connect 语句中都有一个单独的选项开关。在较新的一种中，提供的是单个选项字符串，其中可以包含多个选项值。关于较新语法中可用选项的信息参见 pg\_conndefaults。

### 用法

XXX thomas 1997-12-24

## pg\_disconnect

pg\_disconnect — 关闭一个到后端服务器的连接

### 大纲

```
pg_disconnect dbHandle
```

### 输入

*dbHandle*

指定一个有效的数据库句柄。

### 输出

None

### 描述

`pg_disconnect` 关闭一个到 Postgres 后端的连接。

## pg\_conndefaults

pg\_conndefaults — 获取关于默认连接参数的信息

### 大纲

pg\_conndefaults

### 输入

无。

### 输出

*option list*

结果是一个列表，描述可能的连接选项及其当前默认值。  
列表中的每个条目是如下格式的子列表：

{optname label dispchar dispsize value}

其中 optname 可用作 pg\_connect -conninfo 中的一个选项。

### 描述

pg\_conndefaults 返回关于 pg\_connect -conninfo 中可用的连接选项及每个选项当前默认值的信息。

### 用法

pg\_conndefaults

## pg\_exec

pg\_exec — 向后端发送一个查询字符串

### 大纲

```
pg_exec dbHandle queryString
```

#### 输入

*dbHandle*

指定一个有效的数据库句柄。

*queryString*

指定一个有效的 SQL 查询。

#### 输出

*resultHandle*

如果 Pgtcl 未能获得后端响应，将返回一个 Tcl 错误。否则，会创建一个查询结果对象并返回其句柄。可以把这个句柄传给 `pg_result` 以获取查询的结果。

### 描述

`pg_exec` 向 Postgres 后端提交一个查询并返回结果。查询结果句柄以连接句柄开头，再加上一个句点和一个结果编号。

注意，没有 Tcl 错误并不证明查询成功！后端返回的错误消息会作为带有失败状态的查询结果被处理，而不是让 `pg_exec` 产生 Tcl 错误。

## pg\_result

pg\_result — 获取关于查询结果的信息

### 大纲

```
pg_result resultHandle resultOption
```

#### 输入

*resultHandle*

一个查询结果的句柄。

*resultOption*

指定若干可能选项之一。

#### 选项

-status

结果的状态。

-error

错误消息（如果状态表示出错）；否则为空字符串。

-conn

产生该结果的连接。

-oid

如果命令是 INSERT，为所插入元组的 OID；否则为空字符串。

-numTuples

查询返回的元组数。

-numAttrs

每个元组中的属性数。

-assign arrayName

把结果赋给一个数组，使用形如 (tupno,attributeName)的下标。

-assignbyidx arrayName ?appendstr?

把结果赋给一个数组，用第一个属性的值和其余属性的名作为键。如果给出了 *appendstr*，则把它追加到每个键之后。简而言之，每个元组中除第一个字段外的所有字段都存入数组，使用形如 (firstFieldValue,fieldNameAppendStr)的下标。

-getTuple tupleNumber

以列表形式返回所指元组的各字段。元组编号从零开始。

`-tupleArray tupleNumber arrayName`

把该元组的各字段存入数组 `arrayName`，以字段名为索引。元组编号从零开始。

`-attributes`

返回元组各属性名的列表。

`-lAttributes`

返回子列表的列表，每个元组属性对应一个 {name ftype fsize}。

`-clear`

清除该查询结果对象。

## 输出

结果取决于所选的选项，如上所述。

## 描述

`pg_result` 返回关于由先前的 `pg_exec` 创建的查询结果的信息。

查询结果可以保留任意长的时间，但用完之后，务必通过执行 `pg_result -clear` 将其释放。否则会造成内存泄漏，Pgtcl 最终会抱怨你创建了太多查询结果对象。

## pg\_select

pg\_select — 循环遍历 SELECT 语句的结果

### 大纲

```
pg_select dbHandle queryString
         arrayVar queryProcedure
```

#### 输入

*dbHandle*

指定一个有效的数据库句柄。

*queryString*

指定一个有效的 SQL select 查询。

*arrayVar*

用于存放返回元组的数组变量。

*queryProcedure*

对找到的每个元组运行的过程。

#### 输出

*resultHandle*

返回结果是一个错误消息或一个查询结果的句柄。

### 描述

`pg_select` 向 Postgres 后端提交一个 SELECT 查询，并对结果中的每个元组执行一段给定的代码。*queryString* 必须是一条 SELECT 语句。其他语句都会返回错误。*arrayVar* 变量是循环中使用的数组名。对每个元组，*arrayVar* 都会以字段名作为数组下标填入该元组的字段值。然后执行 *queryProcedure*。

### 用法

如果表 "table" 有字段 "control" 和 "name"（或许还有其他字段），下面这样写即可：

```
pg_select $pgconn "SELECT * from table" array {
  puts [format "%5d %s" array(control) array(name)]
}
```

## pg\_listen

pg\_listen — 设置或更改异步 NOTIFY 消息的回调

### 大纲

```
pg_listen dbHandle notifyName callbackCommand
```

#### 输入

*dbHandle*

指定一个有效的数据库句柄。

*notifyName*

指定要开始或停止监听的 notify 条件名。

*callbackCommand*

如果提供且非空，给出匹配的通知到达时要执行的命令字符串。

#### 输出

None

### 描述

`pg_listen` 创建、更改或取消对来自 Postgres 后端的异步 NOTIFY 消息的监听请求。带 *callbackCommand* 参数时，会建立请求，或替换已存在请求的命令字符串。不带 *callbackCommand* 参数时，则取消先前的请求。

`pg_listen` 请求建立之后，每当带有给定名称的 NOTIFY 消息从后端到达时，就执行指定的命令字符串。任何 Postgres 客户端应用程序发出引用该名称的 NOTIFY 命令时都会发生这种情况。（注意，该名称可以是数据库中已存在关系的名称，但不必一定如此。）命令字符串在 Tcl 空闲循环中执行。这是用 Tk 编写的应用程序的正常空闲状态。在非 Tk 的 Tcl shell 中，可以执行 `update` 或 `vwait` 来进入空闲循环。

使用 `pg_listen` 时，不应直接调用 SQL 语句 LISTEN 或 UNLISTEN。Pgctl 会替你发出这些语句。但如果想自己发送 NOTIFY 消息，可以用 `pg_exec` 调用 SQL 的 NOTIFY 语句。

# pg\_lo\_creat

pg\_lo\_creat — 创建一个大对象

## 大纲

```
pg_lo_creat conn mode
```

## 输入

*conn*

指定一个有效的数据库连接。

*mode*

指定大对象的访问模式

## 输出

*objOid*

所创建大对象的 oid。

## 描述

pg\_lo\_creat 创建一个 Inversion 大对象。

## 用法

mode 可以是 INV\_READ、INV\_WRITE、INV\_ARCHIVE 的任意 OR 组合。OR 分隔字符是 "|"。

```
[pg_lo_creat $conn "INV_READ|INV_WRITE"]
```

## pg\_lo\_open

pg\_lo\_open — 打开一个大对象

### 大纲

```
pg_lo_open conn objOid mode
```

#### 输入

*conn*

指定一个有效的数据库连接。

*objOid*

指定一个有效的大对象 oid。

*mode*

指定大对象的访问模式

#### 输出

*fd*

供后续 pg\_lo\* 例程使用的文件描述符。

### 描述

pg\_lo\_open 打开一个 Inversion 大对象。

### 用法

模式可以是 "r"、"w" 或 "rw"。

# pg\_lo\_close

pg\_lo\_close — 关闭一个大对象

## 大纲

```
pg_lo_close conn fd
```

## 输入

*conn*

指定一个有效的数据库连接。

*fd*

供后续 pg\_lo\* 例程使用的文件描述符。

## 输出

无

## 描述

pg\_lo\_close 关闭一个 Inversion 大对象。

## 用法

# pg\_lo\_read

pg\_lo\_read — 读取一个大对象

## 大纲

```
pg_lo_read conn fd bufVar len
```

## 输入

*conn*

指定一个有效的数据库连接。

*fd*

pg\_lo\_open 返回的大对象文件描述符。

*bufVar*

指定一个有效的缓冲区变量，用于容纳大对象片段。

*len*

指定大对象片段的最大允许大小。

## 输出

无

## 描述

pg\_lo\_read 从大对象中读取至多 *len* 字节到名为 *bufVar* 的变量中。

## 用法

*bufVar* 必须是有效的变量名。

## pg\_lo\_write

pg\_lo\_write — 写入一个大对象

### 大纲

```
pg_lo_write conn fd buf len
```

### 输入

*conn*

指定一个有效的数据库连接。

*fd*

pg\_lo\_open 返回的大对象文件描述符。

*buf*

指定要写入大对象的有效字符串变量。

*len*

指定要写入字符串的最大大小。

### 输出

无

### 描述

pg\_lo\_write 从变量 *buf* 向大对象写入至多 *len* 字节。

### 用法

*buf* 必须是要写入的实际字符串，而不是变量名。

# pg\_lo\_lseek

pg\_lo\_lseek — 在大对象中定位到某个位置

## 大纲

```
pg_lo_lseek conn fd offset whence
```

### 输入

*conn*

指定一个有效的数据库连接。

*fd*

pg\_lo\_open 返回的大对象文件描述符。

*offset*

指定以字节计的从零开始的偏移量。

*whence*

whence 可以是 "SEEK\_CUR"、"SEEK\_END"、"SEEK\_SET"

### 输出

无

## 描述

pg\_lo\_lseek 定位到距大对象开头 *offset* 字节处。

## 用法

*whence* 可以是 "SEEK\_CUR"、"SEEK\_END" 或 "SEEK\_SET"。

## pg\_lo\_tell

pg\_lo\_tell — 返回大对象的当前寻位位置

### 大纲

```
pg_lo_tell conn fd
```

#### 输入

*conn*

指定一个有效的数据库连接。

*fd*

pg\_lo\_open 返回的大对象文件描述符。

#### 输出

*offset*

以字节计的从零开始的偏移量，适合作为 pg\_lo\_lseek 的输入。

### 描述

pg\_lo\_tell 返回当前距大对象开头的 *offset*（以字节计）。

### 用法

# pg\_lo\_unlink

pg\_lo\_unlink — 删除一个大对象

## 大纲

```
pg_lo_unlink conn lobjId
```

## 输入

*conn*

指定一个有效的数据库连接。

*lobjId*

大对象的标识符。XXX Is this the same as objOid in other calls?? - thomas 1998-01-11

## 输出

无

## 描述

`pg_lo_unlink` 删除指定的大对象。

## 用法

## pg\_lo\_import

pg\_lo\_import — 从文件导入一个大对象

### 大纲

```
pg_lo_import conn filename
```

#### 输入

*conn*

指定一个有效的数据库连接。

*filename*

Unix 文件名。

#### 输出

无 XXX Does this return a lojId? Is that the same as the objOid in other calls? thomas - 1998-01-11

### 描述

`pg_lo_import` 读取指定的文件并把内容放入一个大对象。

### 用法

`pg_lo_import` 必须在 BEGIN/END 事务块内调用。

# pg\_lo\_export

pg\_lo\_export — 把大对象导出到文件

## 大纲

```
pg_lo_export conn lobjId filename
```

### 输入

*conn*

指定一个有效的数据库连接。

*lobjId*

大对象标识符。XXX Is this the same as the objOid in other calls?? thomas - 1998-01-11

*filename*

Unix 文件名。

### 输出

无 XXX Does this return a lobjId? Is that the same as the objOid in other calls? thomas - 1998-01-11

## 描述

pg\_lo\_export 把指定的大对象写入一个 Unix 文件。

## 用法

pg\_lo\_export 必须在 BEGIN/END 事务块内调用。

---

# 第 55 章 ODBC 接口

Tim Goeke  
Thomas Lockhart

## 注意

背景信息最初由 Tim Goeke<sup>1</sup> 提供

ODBC（开放数据库互连）是一个抽象的 API，它让你能够编写可与各种 RDBMS 服务器互操作的应用。ODBC 在前端应用和数据库服务器之间提供了一个与产品无关的接口，让用户或开发者能够编写在不同厂商的服务器之间可移植（可传输）的应用。

## 55.1. 背景

ODBC API 在后端对应一个 ODBC 兼容的数据源。它可以是任何东西，从一个文本文件到 Oracle 或 Postgres RDBMS。

后端访问来自 ODBC 驱动，或者说来自允许访问数据的厂商特定驱动。  
psqlODBC 就是这样一种驱动，此外还有其他可用的驱动，例如 OpenLink 的 ODBC 驱动。

一旦你编写了一个 ODBC 应用，只要数据库模式相同，无论后端数据库来自哪个厂商，你应该都能连接到任何后端数据库。

例如，你可以有一台 MS SQL Server 服务器和一台数据完全相同的 Postgres 服务器。Using ODBC，你的 Windows 应用发出的调用将完全相同，后端数据源（在 Windows 应用看来）看起来也一样。

Insight Distributors<sup>2</sup> 为核心 psqlODBC 发行版提供积极且持续的支持。他们提供一份 FAQ<sup>3</sup>、对代码基础的持续开发，并积极参与 interfaces 邮件列表<sup>4</sup>。

## 55.2. Windows 应用

在现实世界中，驱动之间的差异以及 ODBC 支持水平的不同削弱了 ODBC 的潜力：

- Access、Delphi 和 Visual Basic 都直接支持 ODBC。
- 在 C++（例如 Visual C++）下，你可以使用 C++ 的 ODBC API。
- 在 Visual C++ 中，你可以使用 CRecordSet 类，它把 ODBC API 封装在一个 MFC 4.2 类中。如果你在 Windows NT 下进行 Windows C++ 开发，这是最容易的路线。

### 55.2.1. 编写应用

“如果我为 Postgres 编写一个应用，我能否通过 ODBC 调用访问 Postgres 服务器来编写它，还是说只有当 MS SQL Server 或 Access 之类的其他数据库程序需要访问数据时才这么做？”

---

<sup>1</sup><mailto:tgoeke@xpressway.com>

<sup>2</sup><http://www.insightdist.com/>

<sup>3</sup><http://www.insightdist.com/psqlodbc/>

<sup>4</sup><mailto:interfaces@postgresql.org>

ODBC API正是该走的路。关于Visual C++编程，你可以在 Microsoft 的网站或你的VC++文档中找到更多信息。

Visual Basic 和其他 RAD 工具有 Recordset 对象，它们直接使用 ODBC访问数据。使用数据感知控件，你可以（非常迅速地）快速链接到 ODBC后端数据库。

玩玩 MS Access 会帮你弄清楚这一切。试着使用 File->Get External Data。

### 提示

你必须先设置一个 DSN。

## 55.3. Unix 安装

ApplixWare有一个至少在部分平台上受到支持的ODBC数据库接口。在 Linux 下，已经演示过 ApplixWare v4.4.1 通过 Postgres发行版中包含的 psqlODBC驱动与 Postgres v6.4 配合工作。

### 55.3.1. 构建驱动

关于psqlODBC驱动（或任何 ODBC驱动）首先要注意的一点是：在 ODBC驱动要使用的系统上必须存在一个驱动管理器。Unix 上有一种免费软件的ODBC驱动管理器叫做iodbc，可以从网上多处获取，包括 AS200<sup>5</sup>。安装iodbc的说明超出了本文档的范围，但iodbc压缩 .shar 文件内有一个README，它应该能解释如何让它跑起来。

话虽如此，你能为你的平台找到的任何驱动管理器都应该支持 psqlODBC驱动或任何 ODBC驱动。

psqlODBC的 Unix 配置文件最近被大幅重构，以便在受支持的平台上轻松构建，并为将来支持其他 Unix 平台留出余地。驱动的新配置和构建文件应当让在受支持平台上构建驱动变成一个简单的过程。目前这包括 Linux 和 FreeBSD，但我们希望其他用户能贡献必要的信息，迅速扩大可构建该驱动的平台数量。

实际上有两种构建驱动的方法，取决于你以何种方式得到它，而这些差别归根结底只是在哪里以及如何运行 configure和make。驱动可以在一个独立的、仅客户端的安装中构建，也可以作为 Postgres主发行版的一部分来构建。如果你在多个异构平台上都有ODBC客户端应用，独立安装更方便。当目标客户端与服务器相同，或客户端和服务端有相似的运行时配置时，集成安装更方便。

具体来说，如果你得到的psqlODBC驱动是 Postgres发行版的一部分（下文称之为“集成”构建），那么你将在 Postgres发行版的顶层源目录中与其余库一起配置和构建ODBC驱动。如果你得到的是独立打包的驱动，那么就在你解开驱动源代码的目录中运行 configure 和 make。

### 集成安装

这个安装过程适用于集成安装。

1. 为src/configure指定 --with-odbc命令行参数：

```
% ./configure --with-odbc
```

<sup>5</sup><http://www.as220.org/FreeODBC/iodbc-2.12.shar.Z>

```
% make
```

## 2. 重新构建Postgres发行版：

```
% make install
```

一旦配置完成，ODBC驱动就会被构建并安装到为 Postgres系统其他组件定义的区域中。安装范围的 ODBC配置文件会被放到 Postgres 目标树（`POSTGRES_DIR`）的顶层目录。这可以在 `make` 命令行上覆盖，即

```
% make ODBCINST=filename install
```

## v6.4 之前的集成安装

如果你的Postgres安装早于 v6.4、你手头有原始源代码树、而且你想使用最新版本的 ODBC驱动，那么可以试试这种安装形式。

1. 把输出的 tar 文件复制到目标系统并解开到一个干净的目录。

2. 在包含源代码的目录中输入：

```
% ./configure
% make
% make POSTGRES_DIR=PostgresTopDir install
```

3. (可选的) 如果你想把组件安装到不同的树中，可以显式指定各个目的地：

```
% make BINDIR=bindir LIBDIR=libdir HEADERDIR=headerdir ODBCINST=
instfile install
```

## 独立安装

独立安装不与正常的Postgres发行版集成，也不在其上构建。它最适合为多个没有本地安装Postgres源代码树的异构客户端构建 ODBC驱动。

独立安装的库和头文件的默认位置分别是 `/usr/local/lib`和 `/usr/local/include/iodbc`。还有另一个系统范围的配置文件，它会被安装为`/share/odbcinst.ini`（如果`/share`存在）或`/etc/odbcinst.ini`（如果`/share`不存在）。

### 注意

把文件安装到`/share`或`/etc` 需要系统 root 权限。Postgres的大多数安装步骤没有这个要求，你可以改选一个可由你的非 root Postgres超级用户账号写入的其他目的地。

1. 独立安装发行版可以从Postgres发行版构建，也可以从非 Unix 源代码的当前维护者 Insight Distributors<sup>6</sup> 获取。

把 zip 或 gzip 压缩的 tar 文件复制到一个空目录。如果使用的是 zip 包，用命令解开它

```
% unzip -a packagename
```

`-a`选项是必要的，用于去掉源文件中的 DOS CR/LF 对。

<sup>6</sup><http://www.insightdist.com/psqlodbc>

如果你拿到的是 gzip 压缩的 tar 包，只需运行

```
tar -xzf packagename
```

- (可选的) 要从Postgres主源代码树创建一个完整独立安装的 tar 文件：

2. 配置Postgres主发行版。

3. 创建 tar 文件：

```
% cd interfaces/odbc
% make standalone
```

4. 把输出的 tar 文件复制到目标系统。如果使用 ftp，务必以二进制文件方式传输。

5. 把 tar 文件解开到一个干净的目录。

6. 配置独立安装：

```
% ./configure
```

配置可以带选项进行：

```
% ./configure --prefix=rootdir --with-odbc=inidir
```

其中--prefix把库和头文件安装到目录 *rootdir/lib*和 *rootdir/include/iodbc*，而--with-odbc把odbcinst.ini安装到指定的目录。

注意这两个选项也可以在集成构建中使用，但要小心：在集成构建中使用时，--prefix 也会作用于你的Postgres安装的其余部分。--with-odbc只作用于配置文件 *odbcinst.ini*。

7. 编译并链接源代码：

```
% make ODBCINST=instdir
```

你也可以在'make'命令行上覆盖安装的默认位置。这只适用于库文件和头文件的安装。由于驱动需要知道 *odbcinst.ini* 文件的位置，试图覆盖指定其安装目录的环境变量多半会给你带来麻烦。最安全的做法就是让驱动把 *odbcinst.ini* 文件安装在默认目录，或你在 './configure'命令行上用 --with-odbc 指定的目录。

8. 安装源代码：

```
% make POSTGRES_DIR=targettree install
```

要分别覆盖库和头文件的安装目录，你需要在 `make install` 命令行上传入正确的安装变量。这些变量是LIBDIR、HEADERDIR和 ODBCINST。在 `make` 命令行上覆盖 POSTGRES\_DIR会让LIBDIR和 HEADERDIR以你指定的新目录为根。ODBCINST独立于POSTGRES\_DIR。

下面是显式指定各个目的地的做法：

```
% make BINDIR=bindir LIBDIR=libdir HEADERDIR=headerdir install
```

例如，在（用过./configure和 make之后）输入

```
% make POSTGRES_DIR=/opt/psqlodbc install
```

会让库和头文件分别安装到/opt/psqlodbc/lib和 /opt/psqlodbc/include/iodbc目录。

而命令

```
% make POSTGRES DIR=/opt/psqlodbc HEADER DIR=/usr/local install
```

应该会让库安装到 /opt/psqlodbc/lib, 头文件安装到 /usr/local/include/iodbc。如果结果不如预期, 请联系某位维护者。

## 55.4. 配置文件

~/.odbc.ini 包含用户为 psqlODBC驱动指定的访问信息。该文件使用Windows注册表文件典型的约定, 不过尽管有这个限制, 还是可以让它工作的。

.odbc.ini文件有三个必需的节。第一个是 [ODBC Data Sources], 它是你想访问的每个数据库的任意名称和描述的列表。第二个必需的节是数据源说明, 每个数据库都会有一个这样的节。每个节必须用 [ODBC Data Sources]中给出的名称标记, 并且必须包含下列条目:

```
Driver = POSTGRES DIR/lib/libpsqlodbc.so
Database=DatabaseName
Servername=localhost
Port=5432
```

### 提示

记住Postgres数据库名通常是单个词, 不带任何形式的路径名。  
Postgres服务器管理对数据库的实际访问, 你只需要从客户端指定名字。

还可以插入其他条目来控制显示的格式。第三个必需的节是 [ODBC], 它必须包含InstallDir关键字, 还可以包含其他选项。

下面是一个.odbc.ini文件的例子, 展示三个数据库的访问信息:

```
[ODBC Data Sources]
DataEntry = Read/Write Database
QueryOnly = Read-only Database
Test = Debugging Database
Default = Postgres Stripped
```

```
[DataEntry]
ReadOnly = 0
Servername = localhost
Database = Sales
```

```
[QueryOnly]
ReadOnly = 1
Servername = localhost
Database = Sales
```

```
[Test]
```

```
Debug = 1
CommLog = 1
ReadOnly = 0
Servername = localhost
Username = tgl
Password = "no$way"
Port = 5432
Database = test

[Default]
Servername = localhost
Database = tgl
Driver = /opt/postgres/current/lib/libpsqlodbc.so

[ODBC]
InstallDir = /opt/applix/axdata/axshlib
```

## 55.5. ApplixWare

### 55.5.1. 配置

ApplixWare必须被正确配置，才能访问 Postgres的ODBC 软件驱动。

#### 启用ApplixWare数据库访问

这些说明针对的是Linux上的 ApplixWare 4.4.1 版。更详细的信息请参阅Linux Sys Admin在线图书。

1. 你必须修改`axnet.cnf`，让 `elfodbc`能找到`libodbc.so`（ODBC驱动管理器）共享库。这个库随 ApplixWare 发行版一起提供，但`axnet.cnf` 需要被修改为指向正确的位置。

以 root 身份编辑文件 `applixroot/applix/axdata/axnet.cnf`。

- a. 在`axnet.cnf`的底部，找到以

```
#libFor elfodbc /ax/...
```

- b. 把这一行改为

```
libFor elfodbc applixroot/applix/axdata/axshlib/lib
```

这会告诉`elfodbc`到这个目录中查找 ODBC支持库。如果你把 `applix` 安装在了其他地方，请相应地修改路径。

2. Create `.odbc.ini` as 按照上面的说明创建。你可能还想加上标志

```
TextAsLongVarchar=0
```

加到 `.odbc.ini`的数据库特定部分，这样文本字段就不会显示为`**BLOB**`。

#### 测试ApplixWare ODBC 连接

1. 启动Applix Data

2. 选择你感兴趣的Postgres数据库。
  - a. 选择Query->Choose Server。
  - b. 选择ODBC，然后点击Browse。你在.odbc.ini中配置的数据库应该会被显示出来。确保Host: field为空（如果不为空，axnet 会试图联系另一台机器上的 axnet 来查找数据库）。
  - c. 在Browse弹出的框中选择数据库，然后点击OK。
  - d. 在登录标识对话框中输入用户名和口令，然后点击OK。

你应该会在数据窗口左下角看到“Starting elfodbc server”。如果你得到一个错误对话框，请参见下面的调试一节。
3. 'Ready'消息会出现在数据窗口左下角。这表示你现在可以输入查询了。
4. 从 Query->Choose tables 中选择一张表，然后选择 Query->Query 来访问数据库。表中前 50 行左右应该会出现。

## 55.5.2. 常见问题

在试图通过Applix Data建立 ODBC连接时可能出现下列消息：

Cannot launch gateway on server

elfodbc找不到libodbc.so。检查你的axnet.cnf。

Error from ODBC Gateway: IM003::[iODBC][Driver Manager]Specified driver could not be loaded

libodbc.so找不到 .odbc.ini中列出的驱动。核对设置。

Server: Broken Pipe

驱动进程由于某种其他问题已经终止。你可能没有最新版本的 Postgres ODBC包。

setuid to 256: failed to launch gateway

ApplixWare v4.4.1 的九月版（第一个在 Linux 下正式支持 ODBC的版本）在用户名长度超过八（8）个字符时会出问题。问题描述由 Steve Campbell<sup>7</sup> 提供。

### 作者

由Steve Campbell<sup>8</sup> 于 1998-10-20 供稿。

axnet程序的安全系统看起来有点可疑。axnet代表用户做事，在一个真正的多用户系统上它确实应该以 root 权限运行（这样它才能在每个用户的目录中读写）。不过我不太愿意推荐这样做，因为我们完全不知道这会打开什么安全漏洞。

## 55.5.3. 调试ApplixWare ODBC 连接

调试连接问题的一个好工具是 Unix 系统实用程序 strace。

<sup>7</sup>mailto:scampbell@lear.com

<sup>8</sup>mailto:scampbell@lear.com

## 用strace调试

1. 启动 applixware。
2. 对 axnet 进程启动strace。例如，如果

```
ps -aucx | grep ax
```

显示

```
cary    10432   0.0   2.6  1740    392   ?   S   Oct  9   0:00 axnet
cary    27883   0.9  31.0 12692   4596   ?   S   10:24   0:04 axmain
```

然后运行

```
strace -f -s 1024 -p 10432
```

3. 检查 strace 输出。

### 来自 Cary 的提示

ApplixWare的许多错误消息会发往 `stderr`，但我不确定`stderr` 被送到了哪里，所以strace是查明这一点的手段。

例如，在得到 “Cannot launch gateway on server” 之后，我对 axnet 运行了 strace，得到

```
[pid 27947] open("/usr/lib/libodbc.so", O_RDONLY) = -1 ENOENT
        (No such file or directory)
[pid 27947] open("/lib/libodbc.so", O_RDONLY) = -1 ENOENT
        (No such file or directory)
[pid 27947] write(2, "/usr2/applix/axdata/elfodbc:
        can't load library 'libodbc.so'\n", 61) = -1 EIO (I/O error)
```

所以发生的事情是 `applix elfodbc` 在搜索 `libodbc.so` 但找不到它。这就是为什么需要修改 `axnet.cnf`。

## 55.5.4. 运行ApplixWare演示

要完成ApplixWare Data Tutorial，你需要创建教程中引用的示例表。用于创建这些表的 ELF 宏试图在许多数据库列上使用 NULL 条件，而Postgres目前不允许这样做。

要绕过这个问题，你可以这样做：

### 修改ApplixWare演示

1. 把 `/opt/applix/axdata/eng/Demos/sqldemo.am` 复制到一个本地目录。
2. 编辑这份 `sqldemo.am` 的本地副本：
  - a. 搜索 `'null_clause = "NULL"`
  - b. 把它改为 `null_clause = ""`
3. 启动Applix Macro Editor。

4. 从Macro Editor中打开 sqldemo.am 文件。
5. 选择File->Compile and Save。
6. 退出Macro Editor。
7. 启动Applix Data。
8. 选择\*->Run Macro
9. 输入值 "sqldemo"，然后点击OK。

你应该能在数据窗口的状态行（左下角）看到进度。

10. 现在你应该可以访问演示表了。

## 55.5.5. 有用的宏

你可以把数据库登录和口令信息加到标准的 Applix 启动宏文件中。下面是一个~/axhome/macros/login.am文件的例子：===== ApplixWare必须被正确配置，才能访问 Postgres的 ODBC 软件驱动。

### 启用ApplixWare数据库访问

注意，这些说明针对的是Linux上的 ApplixWare 4.4.1 版。更详细的信息请参阅Linux Sys Admin在线图书。

1. 你必须修改axnet.cnf，让 elfodbc能找到libodbc.so（ODBC驱动管理器）共享库。这个库随 ApplixWare 发行版一起提供，但axnet.cnf 需要被修改为指向正确的位置。

以 root 身份编辑文件 *applixroot/applix/axdata/axnet.cnf*。

- a. 在axnet.cnf的底部，找到以

```
#libFor elfodbc /ax/...
```

- b. 把这一行改为

```
libFor elfodbc applixroot/applix/axdata/axshlib/lib
```

这会告诉elfodbc到这个目录中查找 ODBC支持库。如果你把 applix 安装在了其他地方，请相应地修改路径。

2. Create .odbc.ini as 按照上面的说明创建。你可能还想加上标志

```
TextAsLongVarchar=0
```

加到.odbc.ini的数据库特定部分，这样文本字段就不会显示为\*\*BLOB\*\*。

### 测试ApplixWare ODBC 连接

1. 启动Applix Data
2. 选择你感兴趣的Postgres数据库。
  - a. 选择Query->Choose Server。

- b. 选择ODBC，然后点击Browse。你在.odbc.ini中配置的数据库应该会被显示出来。确保Host: field为空（如果不为空，axnet 会试图联系另一台机器上的 axnet 来查找数据库）。
- c. 在Browse弹出的框中选择数据库，然后点击OK。
- d. 在登录标识对话框中输入用户名和口令，然后点击OK。

你应该会在数据窗口左下角看到“Starting elfodbc server”。如果你得到一个错误对话框，请参见下面的调试一节。

3. 'Ready'消息会出现在数据窗口左下角。这表示你现在可以输入查询了。
4. 从 Query->Choose tables 中选择一张表，然后选择 Query->Query 来访问数据库。表中前 50 行左右应该会出现。

## 55.5.6. 常见问题

在试图通过Applix Data建立 ODBC连接时可能出现下列消息：

Cannot launch gateway on server

elfodbc找不到libodbc.so。检查你的axnet.cnf。

Error from ODBC Gateway: IM003:::[iODBC][Driver Manager]Specified driver could not be loaded

libodbc.so找不到 .odbc.ini中列出的驱动。核对设置。

Server: Broken Pipe

驱动进程由于某种其他问题已经终止。你可能没有最新版本的 Postgres ODBC包。

## 55.5.7. 调试ApplixWare ODBC 连接

调试连接问题的一个好工具是 Unix 系统实用程序 strace。

### 用strace调试

1. 启动 applixware。
2. 对 axnet 进程启动strace。例如，如果

```
ps -aucx | grep ax
```

显示

```
cary    10432   0.0   2.6  1740    392  ?   S   Oct  9   0:00 axnet
cary    27883   0.9  31.0 12692   4596  ?   S   10:24   0:04 axmain
```

然后运行

```
strace -f -s 1024 -p 10432
```

3. 检查 strace 输出。

### 来自 Cary 的提示

ApplixWare的许多错误消息会发往 `stderr`，但我不确定`stderr` 被送到了哪里，所以`strace`是查明这一点的手段。

例如，在得到“Cannot launch gateway on server”之后，我对 `axnet` 运行了 `strace`，得到

```
[pid 27947] open("/usr/lib/libodbc.so", O_RDONLY) = -1 ENOENT
        (No such file or directory)
[pid 27947] open("/lib/libodbc.so", O_RDONLY) = -1 ENOENT
        (No such file or directory)
[pid 27947] write(2, "/usr2/applix/axdata/elfodbc:
        can't load library 'libodbc.so'\n", 61) = -1 EIO (I/O error)
```

所以发生的事情是 `applix elfodbc` 在搜索 `libodbc.so` 但找不到它。这就是为什么需要修改 `axnet.cnf`。

## 55.5.8. 运行ApplixWare演示

要完成ApplixWare Data Tutorial，你需要创建教程中引用的示例表。用于创建这些表的 ELF 宏试图在许多数据库列上使用 NULL 条件，而Postgres目前不允许这样做。

要绕过这个问题，你可以这样做：

### 修改ApplixWare演示

1. 把`/opt/applix/axdata/eng/Demos/sqldemo.am` 复制到一个本地目录。
2. 编辑这份`sqldemo.am`的本地副本：
  - a. 搜索 `'null_clause = "NULL"`
  - b. 把它改为 `null_clause = ""`
3. 启动Applix Macro Editor。
4. 从Macro Editor中打开 `sqldemo.am` 文件。
5. 选择`File->Compile and Save`。
6. 退出Macro Editor。
7. 启动Applix Data。
8. 选择`*->Run Macro`
9. 输入值“`sqldemo`”，然后点击OK。

你应该能在数据窗口的状态行（左下角）看到进度。

10. 现在你应该可以访问演示表了。

## 55.5.9. 有用的宏

你可以把数据库登录和口令信息加到标准的 `Applix` 启动宏文件中。下面是一个 `~/axhome/macros/login.am` 文件的例子：

```
macro login
    set_set_system_var@("sql_username@", "tgl")
    set_system_var@("sql_passwd@", "no$way")
endmacro
```

### 小心

对于任何包含用户名和口令信息的文件，你都应当小心设置文件保护。

## 55.5.10. 支持的平台

`psqlODBC`已在Linux上构建并测试过。有报告称在FreeBSD和Solaris上也成功了。对于其他已经支持Postgres的平台，基本代码没有已知的限制。

---

# 第 56 章 JDBC 接口

## 作者

由 Peter T. Mount<sup>1</sup> 编写，他是 JDBC 驱动的作者。

JDBC 是 Java 1.1 及更高版本的核心 API。它为兼容 SQL 的数据库提供了一组标准接口。

Postgres 提供一个 4 类 JDBC Driver。4 类表示该驱动用纯 Java 编写，并使用数据库自身的网络协议通信。因此，该驱动是平台无关的。一旦编译完成，该驱动就可以在任何平台上使用。

## 56.1. 构建 JDBC 接口

### 56.1.1. 编译驱动

驱动的源码位于源码树的 `src/interfaces/jdbc` 目录中。要编译，只需进入该目录并输入：

```
% make
```

完成后，你会在当前目录中找到归档文件 `postgresql.jar`。这就是 JDBC 驱动。

## 注意

你必须使用 `make` 而不是 `javac`，因为驱动出于性能原因使用了一些动态加载技术，而 `javac` 无法应对。`Makefile` 会生成该 jar 归档。

### 56.1.2. 安装驱动

要使用驱动，需要把 jar 归档 `postgresql.jar` 包含在 `CLASSPATH` 中。

示例：

我有一个应用程序使用 JDBC 驱动访问一个包含天体数据的大型数据库。该应用程序和 `jdbc` 驱动安装在 `/usr/local/lib` 目录中，而 `java jdk` 安装在 `/usr/local/jdk1.1.6` 中。

要运行该应用程序，我可以使使用：

```
export CLASSPATH = \ /usr/local/lib/finder.jar:/usr/local/lib/postgresql.jar:. java uk.org.  
retep.finder.Main
```

载入驱动的内容将在本章稍后介绍。

## 56.2. 为 JDBC 准备数据库

由于 Java 只能使用 TCP/IP 连接，Postgres `postmaster` 必须以 `-i` 标志运行。

---

<sup>1</sup>[peter@retep.org.uk](mailto:peter@retep.org.uk)

另外，必须配置 `pg_hba.conf` 文件。它位于 `PGDATA` 目录中。在默认安装中，该文件只允许通过 UNIX 域套接字访问。要让 JDBC 驱动连接同一个 `localhost`，你需要加上类似这样的内容：

```
host all 127.0.0.1 255.255.255.255 password
```

这样，从本地机器就可以用 JDBC 访问所有数据库。

JDBC Driver 支持 `trust`、`ident`、`password` 和 `crypt` 认证方法。

## 56.3. 使用驱动

本节并不是一份完整的 JDBC 编程指南，但应该能帮助你入门。更多信息请参阅标准的 JDBC API 文档。

另外，也可以看看源码中附带的例子。这里使用的是其中的基本示例。

## 56.4. 导入 JDBC

任何使用 JDBC 的源码都需要导入 `java.sql` 包，使用：

```
import java.sql.*;
```

### 重要

不要导入 `postgresql` 包。如果你这样做，你的源代码将无法编译，因为 `javac` 会产生混淆。

## 56.5. 载入驱动

在连接数据库之前，你需要载入驱动。有两种可用的方法，哪种最好取决于你的代码。

第一种方法中，你的代码使用 `Class.forName()` 方法隐式地载入驱动。对于 `Postgres`，你可以使用：

```
Class.forName(postgresql.Driver);
```

这会载入驱动，并且在载入时，驱动会自动向 JDBC 注册自身。

注意：`forName()` 方法可能抛出一个 `ClassNotFoundException`，所以如果驱动不可用，你需要捕获它。

这是最常用的方法，但它把你的代码限制为只使用 `Postgres`。如果你的代码将来可能访问另一个数据库，并且你不使用我们的扩展，那么建议使用第二种方法。

第二种方法在 JVM 启动时使用 `-D` 参数把驱动作为参数传递给它。

示例：

```
% java -Djdbc.drivers=postgresql.Driver example.ImageViewer
```

在这个例子中，JVM 会尝试把载入驱动作为其初始化的一部分。完成后，就会启动 Image Viewer。

现在，这种方法是更好的一种，因为它允许你的代码在不重新编译的情况下与其他数据库一起使用。唯一还需要改变的是 URL，这是接下来要介绍的内容。

最后一点。当你的代码随后尝试打开一个 Connection，并且你得到一个被抛出的 No driver available SQLException 时，这可能是驱动不在类路径中，或者参数中的值不正确造成的。

## 56.6. 连接到数据库

在 JDBC 中，数据库由一个 URL（统一资源定位符）表示。对于 Postgres，它采用下列形式之一：

- jdbc:postgresql:database
- jdbc:postgresql://host/database
- jdbc:postgresql://host:port/database

其中：

*host*

服务器的主机名。默认为“localhost”。

*port*

服务器监听的端口号。默认为 Postgres 的标准端口号（5432）。

*database*

数据库名。

要连接，你需要从 JDBC 获得一个 Connection 实例。为此，你可以使用 DriverManager.getConnection() 方法：

```
Connection db = DriverManager.getConnection(url,user,pwd);
```

## 56.7. 发出查询并处理结果

每当你想向数据库发出 SQL 语句时，你都需要一个 Statement 实例。一旦你有了一个 Statement，就可以使用 executeQuery() 方法发出查询。这将返回一个 ResultSet 实例，其中包含全部结果。

### 56.7.1. 使用 Statement 接口

使用 Statement 接口时必须考虑以下几点：

- 你可以按需多次使用一个 Statement 实例。你可以在打开连接后立即创建一个，并在该连接的整个生命周期中使用它。你必须记住，每个 Statement 只能存在一个 ResultSet。

- 如果你需要在处理一个 `ResultSet` 期间执行另一个查询，只需创建并使用另一个 `Statement` 即可。
- 如果你使用线程，而且有多个线程在使用数据库，就必须为每个线程使用一个单独的 `Statement`。如果你打算使用它们，请参阅本文档后面介绍线程和 `Servlet` 的章节，其中涵盖了一些重要的内容。

## 56.7.2. 使用 `ResultSet` 接口

使用 `ResultSet` 接口时必须考虑以下几点：

- 在读取任何值之前，你必须调用 `next()`。如果有结果它会返回 `true`，但更重要的是，它会让该行做好被处理的准备。
- 按照 `JDBC` 规范，你应该只访问一个字段一次。遵守这条规则是最安全的，尽管目前 `Postgres` 驱动允许你按需多次访问一个字段。
- 一旦用完一个 `ResultSet`，你必须通过调用 `close()` 来关闭它。
- 一旦你用创建某个 `ResultSet` 的 `Statement` 请求另一个查询，当前打开的实例就会被关闭。

下面是一个例子：

```
Statement st = db.createStatement();
ResultSet rs = st.executeQuery(select * from mytable);
while(rs.next()) {
    System.out.print(Column 1 returned );
    System.out.println(rs.getString(1));
}
rs.close();
st.close();
```

## 56.8. 执行更新

要执行一个更新（或任何其他不返回结果的 `SQL` 语句），只需使用 `executeUpdate()` 方法：

```
st.executeUpdate(create table basic (a int2, b int2));
```

## 56.9. 关闭连接

要关闭数据库连接，只需对 `Connection` 调用 `close()` 方法：

```
db.close();
```

## 56.10. 使用大对象

在 `Postgres` 中，大对象（也称为 `blob`）用于存放数据库中无法存储在普通 `SQL` 表中的数据。它们以一个“表/索引”对的形式存储，并通过一个 `OID` 值从你自己的表中引用。

使用大对象有两种方法。第一种是标准的 `JDBC` 方式，这里介绍的就是它。另一种使用我们自己 `api` 的扩展，它把 `libpq` 的大对象 `API` 呈现给 `Java`，提供了比标准方式更好的大对象访问能力。在内部，驱动正是使用该扩展来提供大对象支持的。

在 JDBC 中，访问它们的标准方式是使用 `ResultSet` 中的 `getBinaryStream()` 方法和 `PreparedStatement` 中的 `setBinaryStream()` 方法。这些方法让大对象表现为一个 Java 流，使你可以使用 `java.io` 包和其他包来操纵该对象。

例如，假设你有一个表存储图像的文件名，并且你有一个包含该图像本身的大对象：

```
create table images (imgname name,imgoid oid);
```

要插入一幅图像，你可以使用：

```
File file = new File(myimage.gif);
FileInputStream fis = new FileInputStream(file);
PreparedStatement ps = conn.prepareStatement(insert into images values
    (?,?));
ps.setString(1,file.getName());
ps.setBinaryStream(2,fis,file.length());
ps.executeUpdate();
ps.close();
fis.close();
```

在这个例子中，`setBinaryStream` 从一个流向大对象传输指定数量的字节，并把 `OID` 存入持有对它的引用的字段中。

取回一幅图像甚至更容易（我这里使用 `PreparedStatement`，但 `Statement` 同样可以使用）：

```
PreparedStatement ps = con.prepareStatement(select oid from images
    where name=?);
ps.setString(1,myimage.gif);
ResultSet rs = ps.executeQuery();
if(rs!=null) {
    while(rs.next()) {
        InputStream is = rs.getBinaryInputStream(1);
        // use the stream in some way here
        is.close();
    }
    rs.close();
}
ps.close();
```

这里你可以看到大对象是作为一个 `InputStream` 取回的。你还会注意到，我们在处理结果中的下一行之前关闭了流。这是 JDBC 规范的一部分，该规范规定，当调用 `ResultSet.next()` 或 `ResultSet.close()` 时，返回的任何 `InputStream` 都会被关闭。

## 56.11. Postgres 对 JDBC API 的扩展

Postgres 是一个可扩展的数据库系统。你可以向后端添加自己的函数，这些函数随后可以在查询中被调用，你甚至可以添加自己的数据类型。

现在，由于这些是我们独有的设施，我们通过一组扩展 API 从 Java 支持它们。标准驱动核心中的某些特性（如大对象）实际上就是使用这些扩展实现的。

访问扩展

要访问其中一些扩展，你需要使用 `postgresql.Connection` 类中的一些额外方法。这种情况下，你需要把 `Driver.getConnection()` 的返回值强制转换。

例如：

```
Connection db = Driver.getConnection(url,user,pass);

// later on
Fastpath fp = ((postgresql.Connection)db).getFastpathAPI();

Class postgresql.Connection

java.lang.Object
|
+----postgresql.Connection

public class Connection extends Object implements Connection
```

这些是用来访问我们的扩展的额外方法。这里没有列出 `java.sql.Connection` 已定义的方法。

```
public Fastpath getFastpathAPI() throws SQLException
```

返回当前连接的 `Fastpath` API。

注意：这不是 JDBC 的一部分，而是允许访问 `postgresql` 后端本身的函数。

它主要供 `LargeObject` API 使用

最好的使用方式如下：

```
import postgresql.fastpath.*;
...
Fastpath fp = ((postgresql.Connection)myconn).getFastpathAPI();
```

其中 `myconn` 是一个到 `postgresql` 的已打开连接。

返回值：

允许访问 `postgresql` 后端函数的 `Fastpath` 对象。

抛出： `SQLException`

首次初始化时由 `Fastpath` 抛出

```
public LargeObjectManager getLargeObjectAPI() throws SQLException
```

返回当前连接的 `LargeObject` API。

注意：这不是 JDBC 的一部分，而是允许访问 `postgresql` 后端本身的函数。

最好的使用方式如下：

```
import postgresql.largeobject.*;
...
LargeObjectManager lo =
((postgresql.Connection)myconn).getLargeObjectAPI();
```

其中 `myconn` 是一个到 `postgresql` 的已打开连接。

返回值:

实现该 API 的 `LargeObject` 对象

抛出: `SQLException`

首次初始化时由 `LargeObject` 抛出

```
public void addDataType(String type,
                        String name)
```

它允许客户端代码为 `postgresql` 较独特的数据类型之一添加处理器。通常, 驱动不认识的数据类型会由 `ResultSet.getObject()` 以 `PGObject` 实例的形式返回。

此方法允许你编写一个继承 `PGObject` 的类, 并告诉驱动要使用的类型名和类名。

它的缺点是, 每次建立连接时都必须调用此方法。

注意: 这不是 JDBC 的一部分, 而是一个扩展。

最好的使用方式如下:

```
...
((postgresql.Connection)myconn).addDataType("mytype", "my.class.name"-
);
...
```

其中 `myconn` 是一个到 `postgresql` 的已打开连接。

处理类必须继承 `postgresql.util.PGObject`

参见:

`PGObject`

`Fastpath`

`Fastpath` 是 `libpq C` 接口中的一个 API, 允许客户端机器在数据库后端上执行一个函数。大多数客户端代码不需要使用它, 但之所以提供, 是因为大对象 API 会用到它。

要使用它, 需要用下面这行导入 `postgresql.fastpath` 包:

```
import postgresql.fastpath.*;
```

然后, 在你的代码中需要获得一个 `FastPath` 对象:

```
Fastpath fp = ((postgresql.Connection)conn).getFastpathAPI();
```

它会返回一个与数据库连接关联的实例, 你可以用它发出命令。这里必须把 `Connection` 转换为 `postgresql.Connection`, 因为 `getFastpathAPI()` 是我们自己的方法, 不属于 JDBC。

一旦有了 `Fastpath` 实例, 就可以使用 `fastpath()` 方法执行一个后端函数。

```
Class postgresql.fastpath.Fastpath
```

```
java.lang.Object
```

```
|
+----postgresql.fastpath.Fastpath
```

```
public class Fastpath
```

```
extends Object
```

这个类实现 Fastpath api。

这是一种从 Java 应用内执行嵌入在 postgresql 后端中的函数的手段。

它基于 src/interfaces/libpq/fe-exec.c 文件

参见：

FastpathFastpathArg, LargeObject

## 方法

```
public Object fastpath(int fnid,
                        boolean resulttype,
                        FastpathArg args[]) throws SQLException
```

向 PostgreSQL 后端发送一次函数调用

参数：

fnid - 函数 id  
 resulttype - 如果结果是整数则为 true, 其他结果为 false  
 args - 传递给 fastpath 的 FastpathArg 参数

返回值：

无数据时为 null, 整数结果时为 Integer, 否则为 byte[]

抛出：SQLException

发生数据库访问错误时。

```
public Object fastpath(String name,
                        boolean resulttype,
                        FastpathArg args[]) throws SQLException
```

按名称向 PostgreSQL 后端发送一次函数调用。

注意：

过程名到函数 id 的映射必须已经存在, 通常来自早先对 addfunction() 的调用。这是推荐的调用方式, 因为函数 id 在后端的不同版本之间可能改变。关于其工作方式的示例, 参见 postgresql.LargeObject

参数：

name - 函数名  
 resulttype - 如果结果是整数则为 true, 其他结果为 false  
 args - 传递给 fastpath 的 FastpathArg 参数

返回值：

无数据时为 null, 整数结果时为 Integer, 否则为 byte[]

抛出: SQLException  
名称未知或发生数据库访问错误时。

参见:  
LargeObject

```
public int getInteger(String name,  
                      FastpathArg args[]) throws SQLException
```

这个便捷方法假定返回值是 Integer

参数:  
name - 函数名  
args - 函数参数

返回值:  
整数结果

抛出: SQLException  
发生数据库访问错误或没有结果时。

```
public byte[] getData(String name,  
                      FastpathArg args[]) throws SQLException
```

这个便捷方法假定返回值是二进制数据

参数:  
name - 函数名  
args - 函数参数

返回值:  
包含结果的 byte[] 数组

抛出: SQLException  
发生数据库访问错误或没有结果时。

```
public void addFunction(String name,  
                        int fnid)
```

它向我们的查找表中添加一个函数。

用户代码应使用基于查询的 addFunctions 方法, 而不是硬编码 oid。函数的 oid 并不保证保持不变, 即使在同一版本的不同服务器上也是如此。

参数:  
name - 函数名  
fnid - 函数 id

```
public void addFunctions(ResultSet rs) throws SQLException
```

它接受一个包含两列的 ResultSet。第 1 列是函数名, 第 2 列是 oid。

它会读取整个 ResultSet, 把值装入函数表。

记住：调用之后要 `close()` 该 `resultset`!!

关于函数名查找的实现说明：

PostgreSQL 把函数 `id` 及其对应名称存储在 `pg_proc` 表中。为了在本地加速，不是在需要时逐个查询该表，而是使用一个 `Hashtable`。同时，只有需要的函数才会被录入该表，以使连接时间尽可能快。

`postgresql.LargeObject` 类在启动时执行一次查询，并把返回的 `ResultSet` 传给这里的 `addFunctions()` 方法。

完成之后，`LargeObject api` 就按名称引用这些函数。

不要以为手工把它们转换成 `oid` 就能行。目前是可以，但在开发过程中它们可能改变（V7.0 期间曾有相关讨论），因此这样实现是为了避免将来不必要的麻烦。

参数：

`rs - ResultSet`

抛出：SQLException

发生数据库访问错误时。

参见：

`LargeObjectManager`

```
public int getID(String name) throws SQLException
```

返回与名称关联的函数 `id`

如果没有为该名称调用过 `addFunction()` 或 `addFunctions()`，则会抛出 `SQLException`。

参数：

`name - 要查找的函数名`

返回值：

供 `fastpath` 调用的函数 `ID`

抛出：SQLException

函数未知时。

```
Class postgresql.fastpath.FastpathArg
```

```
java.lang.Object
```

```
|
```

```
+---postgresql.fastpath.FastpathArg
```

```
public class FastpathArg extends Object
```

每次 `fastpath` 调用都需要一个参数数组，其数量和类型取决于所调用的函数。

这个类实现提供此能力所需的方法。

关于使用示例，参见 `postgresql.largeobject` 包

参见:

Fastpath, LargeObjectManager, LargeObject

### 构造器

```
public FastpathArg(int value)
```

构造一个由整数值构成的参数

参数:

value - 要设置的 int 值

```
public FastpathArg(byte bytes[])
```

构造一个由字节数组构成的参数

参数:

bytes - 要存储的数组

```
public FastpathArg(byte buf[],
                    int off,
                    int len)
```

构造一个由字节数组一部分构成的参数

参数:

buf - 源数组  
off - 数组内的偏移  
len - 要包含的数据长度

```
public FastpathArg(String s)
```

构造一个由字符串构成的参数。

参数:

s - 要存储的字符串

### 几何数据类型

PostgreSQL 有一组能把几何特性存入表中的数据类型，涵盖单点、直线和多边形。

我们用 `postgresql.geometric` 包在 Java 中支持这些类型。

它包含继承 `postgresql.util.PGobject` 类的各个类。关于如何实现你自己的数据类型处理器，参见那个类。

```
Class postgresql.geometric.PGbox
```

```
java.lang.Object
```

```
|
```

```
+----postgresql.util.PGobject
```

```
|
```

```
+----postgresql.geometric.PGbox
```

```
public class PGbox extends PGobject implements Serializable,
Cloneable
```

它表示 postgresql 中的 box 数据类型。

#### 变量

```
public PGpoint point[]
```

这是该矩形的两个角点。

#### 构造器

```
public PGbox(double x1,
              double y1,
              double x2,
              double y2)
```

参数:

x1 - 第一个 x 坐标  
y1 - 第一个 y 坐标  
x2 - 第二个 x 坐标  
y2 - 第二个 y 坐标

```
public PGbox(PGpoint p1,
              PGpoint p2)
```

参数:

p1 - 第一个点  
p2 - 第二个点

```
public PGbox(String s) throws SQLException
```

参数:

s - PostgreSQL 语法的矩形定义

抛出: SQLException

如果定义无效

```
public PGbox()
```

必需的构造器

#### 方法

```
public void setValue(String value) throws SQLException
```

此方法设置该对象的值。子类应当覆盖它，但仍应调用它。

参数:

value - 对象值的字符串表示

抛出: SQLException

如果值对该类型无效则抛出

Overrides:  
setValue in class PGObject

```
public boolean equals(Object obj)
```

参数:  
obj - 要比较的对象

返回值:  
如果两个矩形相同则为 true

Overrides:  
equals in class PGObject

```
public Object clone()
```

必须覆盖此方法才能克隆该对象

Overrides:  
clone in class PGObject

```
public String getValue()
```

返回值:  
按 postgresql 所期望语法表示的 PGbox

Overrides:  
getValue in class PGObject

```
Class postgresql.geometric.PGcircle
```

```
java.lang.Object
```

```
  |
  +----postgresql.util.PGobject
      |
      +----postgresql.geometric.PGcircle
```

```
public class PGcircle extends PGobject implements Serializable,
Cloneable
```

它表示 postgresql 的 circle 数据类型, 由一个点和一个半径组成

变量

```
public PGpoint center
```

这是圆心点

```
public double radius
```

这是半径

构造器

```
public PGcircle(double x,  
                double y,  
                double r)
```

参数:

x - 圆心的坐标  
y - 圆心的坐标  
r - 圆的半径

```
public PGcircle(PGpoint c,  
                double r)
```

参数:

c - 描述圆心的 PGpoint  
r - 圆的半径

```
public PGcircle(String s) throws SQLException
```

参数:

s - PostgreSQL 语法的圆定义。

抛出: SQLException  
转换失败时

```
public PGcircle()
```

此构造器由驱动使用。

## 方法

```
public void setValue(String s) throws SQLException
```

参数:

s - PostgreSQL 语法的圆定义。

抛出: SQLException  
转换失败时

Overrides:  
setValue in class PGObject

```
public boolean equals(Object obj)
```

参数:

obj - 要比较的对象

返回值:  
如果两个矩形相同则为 true

Overrides:  
equals in class PGObject

```
public Object clone()
```

必须覆盖此方法才能克隆该对象

```
Overrides:
    clone in class PGObject
```

```
public String getValue()
```

返回值:

按 postgresql 所期望语法表示的 PGcircle

```
Overrides:
    getValue in class PGObject
```

```
Class postgresql.geometric.PGline
```

```
java.lang.Object
```

```
|
```

```
+----postgresql.util.PGobject
```

```
|
```

```
+----postgresql.geometric.PGline
```

```
public class PGline extends PGObject implements Serializable,
Cloneable
```

它实现由两点构成的 line。后端目前尚未实现 line 类型，但这个类保证了在实现之后我们就能直接使用它。

变量

```
public PGpoint point[]
```

这就是这两个点。

构造器

```
public PGline(double x1,
               double y1,
               double x2,
               double y2)
```

参数:

```
x1 - 第一个点的坐标
y1 - 第一个点的坐标
x2 - 第二个点的坐标
y2 - 第二个点的坐标
```

```
public PGline(PGpoint p1,
               PGpoint p2)
```

参数:

```
p1 - 第一个点
p2 - 第二个点
```

```
public PGline(String s) throws SQLException
```

参数:

s - PostgreSQL 语法的圆定义。

抛出: SQLException

转换失败时

```
public PGline()
```

驱动所需

方法

```
public void setValue(String s) throws SQLException
```

参数:

s - PostgreSQL 语法的线段定义

抛出: SQLException

转换失败时

Overrides:

setValue in class PGObject

```
public boolean equals(Object obj)
```

参数:

obj - 要比较的对象

返回值:

如果两个矩形相同则为 true

Overrides:

equals in class PGObject

```
public Object clone()
```

必须覆盖此方法才能克隆该对象

Overrides:

clone in class PGObject

```
public String getValue()
```

返回值:

按 postgresql 所期望语法表示的 PGline

Overrides:

getValue in class PGObject

```
Class postgresql.geometric.PGline
```

```
java.lang.Object
```

```
|
+----postgresql.util.PGobject
|
+----postgresql.geometric.PGlseg
```

```
public class PGLseg extends PGobject implements Serializable,
Cloneable
```

它实现由两点构成的 lseg (线段)

变量

```
public PGpoint point[]
```

这就是这两个点。

构造器

```
public PGLseg(double x1,
               double y1,
               double x2,
               double y2)
```

参数:

```
x1 - 第一个点的坐标
y1 - 第一个点的坐标
x2 - 第二个点的坐标
y2 - 第二个点的坐标
```

```
public PGLseg(PGpoint p1,
               PGpoint p2)
```

参数:

```
p1 - 第一个点
p2 - 第二个点
```

```
public PGLseg(String s) throws SQLException
```

参数:

```
s - PostgreSQL 语法的圆定义。
```

抛出: SQLException  
转换失败时

```
public PGLseg()
```

驱动所需

方法

```
public void setValue(String s) throws SQLException
```

参数:

s - PostgreSQL 语法的线段定义

抛出: SQLException

转换失败时

Overrides:

setValue in class PObject

```
public boolean equals(Object obj)
```

参数:

obj - 要比较的对象

返回值:

如果两个矩形相同则为 true

Overrides:

equals in class PObject

```
public Object clone()
```

必须覆盖此方法才能克隆该对象

Overrides:

clone in class PObject

```
public String getValue()
```

返回值:

按 postgresql 所期望语法表示的 PGlseg

Overrides:

getValue in class PObject

```
Class postgresql.geometric.PGpath
```

```
java.lang.Object
```

```
|
```

```
+----postgresql.util.PObject
```

```
|
```

```
+----postgresql.geometric.PGpath
```

```
public class PGpath extends PObject implements Serializable,
Cloneable
```

它实现 path (一条多段线, 可以是闭合的)

变量

```
public boolean open
```

如果路径是开放的则为 true, 闭合则为 false

```
public PGpoint points[]
```

定义该路径的各个点

#### 构造器

```
public PGpath(PGpoint points[],
              boolean open)
```

参数:

points - 定义该路径的 PGpoint 数组  
open - 如果路径是开放的则为 true, 闭合则为 false

```
public PGpath()
```

驱动所需

```
public PGpath(String s) throws SQLException
```

参数:

s - PostgreSQL 语法的圆定义。

抛出: SQLException

转换失败时

#### 方法

```
public void setValue(String s) throws SQLException
```

参数:

s - PostgreSQL 语法的路径定义

抛出: SQLException

转换失败时

Overrides:

setValue in class PGObject

```
public boolean equals(Object obj)
```

参数:

obj - 要比较的对象

返回值:

如果两个矩形相同则为 true

Overrides:

equals in class PGObject

```
public Object clone()
```

必须覆盖此方法才能克隆该对象

Overrides:

clone in class PGObject

```
public String getValue()
```

返回按 postgresql 所期望语法表示的该多边形

Overrides:

getValue in class PObject

```
public boolean isOpen()
```

如果路径是开放的则返回 true

```
public boolean isClosed()
```

如果路径是闭合的则返回 true

```
public void closePath()
```

把路径标记为闭合

```
public void openPath()
```

把路径标记为开放

```
Class postgresql.geometric.PGpoint
```

```
java.lang.Object
```

```
|
```

```
+----postgresql.util.PObject
```

```
|
```

```
+----postgresql.geometric.PGpoint
```

```
public class PGpoint extends PObject implements Serializable,  
Cloneable
```

它实现 java.awt.Point 的一个版本，只是用 double 表示坐标。

它映射到 postgresql 的 point 数据类型。

变量

```
public double x
```

该点的 X 坐标

```
public double y
```

该点的 Y 坐标

构造器

```
public PGpoint(double x,  
double y)
```

参数:

x - 坐标  
y - 坐标

```
public PGpoint(String value) throws SQLException
```

当某点嵌入其他几何类型的定义中时, 主要由那些类型调用此构造器。

参数:

value - PostgreSQL 语法的该点定义

```
public PGpoint()
```

驱动所需

方法

```
public void setValue(String s) throws SQLException
```

参数:

s - PostgreSQL 语法的该点定义

抛出: SQLException

转换失败时

Overrides:

setValue in class PGObject

```
public boolean equals(Object obj)
```

参数:

obj - 要比较的对象

返回值:

如果两个矩形相同则为 true

Overrides:

equals in class PGObject

```
public Object clone()
```

必须覆盖此方法才能克隆该对象

Overrides:

clone in class PGObject

```
public String getValue()
```

返回值:

按 postgresql 所期望语法表示的 PGpoint

Overrides:

getValue in class PGObject

```
public void translate(int x,  
                     int y)
```

按给定的量平移该点。

参数:

x - 在 x 轴上要加的整数增量  
y - 在 y 轴上要加的整数增量

```
public void translate(double x,  
                     double y)
```

按给定的量平移该点。

参数:

x - 在 x 轴上要加的 double 增量  
y - 在 y 轴上要加的 double 增量

```
public void move(int x,  
                int y)
```

把该点移动到给定的坐标。

参数:

x - 整数坐标  
y - 整数坐标

```
public void move(double x,  
                double y)
```

把该点移动到给定的坐标。

参数:

x - double 坐标  
y - double 坐标

```
public void setLocation(int x,  
                      int y)
```

把该点移动到给定的坐标。相关描述参见 `java.awt.Point`

参数:

x - 整数坐标  
y - 整数坐标

参见:

`Point`

```
public void setLocation(Point p)
```

把该点移动到给定的 `java.awt.Point`。相关描述参见 `java.awt.Point`

参数:

p - 要移动到的点

参见:

Point

Class postgresql.geometric.PGpolygon

java.lang.Object

|

+----postgresql.util.PGobject

|

+----postgresql.geometric.PGpolygon

public class PGpolygon extends PGobject implements Serializable,  
Cloneable

它实现 PostgreSQL 中的 polygon 数据类型。

变量

public PGpoint points[]

定义该多边形的各个点

构造器

public PGpolygon(PGpoint points[])

用一个 PGpoint 数组创建多边形

参数:

points - 定义该多边形的各个点

public PGpolygon(String s) throws SQLException

参数:

s - PostgreSQL 语法的圆定义。

抛出: SQLException

转换失败时

public PGpolygon()

驱动所需

方法

public void setValue(String s) throws SQLException

参数:

s - PostgreSQL 语法的多边形定义

抛出: SQLException

转换失败时

```

    Overrides:
        setValue in class PGObject

    public boolean equals(Object obj)

```

参数:

obj - 要比较的对象

返回值:

如果两个矩形相同则为 true

```

    Overrides:
        equals in class PGObject

    public Object clone()

```

必须覆盖此方法才能克隆该对象

```

    Overrides:
        clone in class PGObject

    public String getValue()

```

返回值:

按 postgresql 所期望语法表示的 PGpolygon

```

    Overrides:
        getValue in class PGObject

```

## 大对象

标准 JDBC 规范支持大对象。但那个接口能力有限，而 PostgreSQL 提供的 api 允许像访问本地文件一样随机访问对象内容。

postgresql.largeobject 包向 Java 提供了 libpq C 接口的大对象 API。它由两个类组成：负责创建、打开和删除大对象的 LargeObjectManager，以及处理单个对象的 LargeObject。

```

Class postgresql.largeobject.LargeObject

java.lang.Object
|
+----postgresql.largeobject.LargeObject

public class LargeObject extends Object

```

这个类实现 postgresql 的大对象接口。

它提供运行该接口所需的基本方法，另有一对方法为该对象提供 InputStream 和 OutputStream 类。

通常，客户端代码会使用 ResultSet 中的 getAsciiStream、getBinaryStream 或 getUnicodeStream 方法，或 PreparedStatement 中的 setAsciiStream、setBinaryStream 或 setUnicodeStream 方法来访问大对象。

然而有时需要对大对象的更底层访问，这是 JDBC 规范不支持的。

关于如何访问大对象或如何创建大对象，参见 `postgresql.largeobject.LargeObjectManager`。

参见：

`LargeObjectManager`

## 变量

```
public static final int SEEK_SET
```

表示从文件开头定位

```
public static final int SEEK_CUR
```

表示从当前位置定位

```
public static final int SEEK_END
```

表示从文件末尾定位

## 方法

```
public int getOID()
```

返回值：

该大对象的 `OID`

```
public void close() throws SQLException
```

此方法关闭该对象。调用之后不得再对该对象调用方法。

抛出： `SQLException`

发生数据库访问错误时。

```
public byte[] read(int len) throws SQLException
```

从该对象读取一些数据，并以 `byte[]` 数组返回

参数：

`len` - 要读取的字节数

返回值：

包含所读数据的 `byte[]` 数组

抛出： `SQLException`

发生数据库访问错误时。

```
public void read(byte buf[],
                 int off,
                 int len) throws SQLException
```

从该对象读取一些数据到既有数组中

参数:

buf - 目标数组  
off - 数组内的偏移  
len - 要读取的字节数

抛出: SQLException

发生数据库访问错误时。

```
public void write(byte buf[]) throws SQLException
```

把一个数组写入该对象

参数:

buf - 要写入的数组

抛出: SQLException

发生数据库访问错误时。

```
public void write(byte buf[],  
                  int off,  
                  int len) throws SQLException
```

把数组中的一些数据写入该对象

参数:

buf - 目标数组  
off - 数组内的偏移  
len - 要写入的字节数

抛出: SQLException

发生数据库访问错误时。

```
public void seek(int pos,  
                 int ref) throws SQLException
```

设置该对象内的当前位置。

这类似于标准 C 库中的 `fseek()` 调用, 允许随机访问大对象。

参数:

pos - 对象内的位置  
ref - `SEEK_SET`、`SEEK_CUR` 或 `SEEK_END` 之一

抛出: SQLException

发生数据库访问错误时。

```
public void seek(int pos) throws SQLException
```

设置该对象内的当前位置。

这类似于标准 C 库中的 `fseek()` 调用, 允许随机访问大对象。

参数:

pos - 对象内相对开头的位置

抛出: SQLException

发生数据库访问错误时。

```
public int tell() throws SQLException
```

返回值:

对象内的当前位置

抛出: SQLException

发生数据库访问错误时。

```
public int size() throws SQLException
```

此方法效率不高, 因为要知道对象的大小, 只能定位到末尾、记录当前位置、再回到原位置。

将来会找到更好的方法。

返回值:

大对象的大小

抛出: SQLException

发生数据库访问错误时。

```
public InputStream getInputStream() throws SQLException
```

返回该对象的一个 InputStream。

随后这个 InputStream 可用于任何需要 InputStream 的方法。

抛出: SQLException

发生数据库访问错误时。

```
public OutputStream getOutputStream() throws SQLException
```

返回该对象的一个 OutputStream

随后这个 OutputStream 可用于任何需要 OutputStream 的方法。

抛出: SQLException

发生数据库访问错误时。

```
Class postgresql.largeobject.LargeObjectManager
```

```
java.lang.Object
```

```
|
```

```
+----postgresql.largeobject.LargeObjectManager
```

```
public class LargeObjectManager extends Object
```

这个类实现 postgresql 的大对象接口。

它提供允许客户端代码在数据库中创建、打开和删除大对象的方法。打开对象时会返回一个 `postgresql.largeobject.LargeObject` 实例，其方法随后允许访问该对象。

这个类只能由 `postgresql.Connection` 创建

要访问这个类，使用下面这段代码：

```
import postgresql.largeobject.*;
Connection conn;
LargeObjectManager lobj;
... 打开连接的代码 ...
lobj = ((postgresql.Connection)myconn).getLargeObjectAPI();
```

通常，客户端代码会使用 `ResultSet` 中的 `getAsciiStream`、`getBinaryStream` 或 `getUnicodeStream` 方法，或 `PreparedStatement` 中的 `setAsciiStream`、`setBinaryStream` 或 `setUnicodeStream` 方法来访问大对象。

然而有时需要对大对象的更底层访问，这是 JDBC 规范不支持的。

关于如何操作大对象内容，参见 `postgresql.largeobject.LargeObject`。

参见：

`LargeObject`

变量

```
public static final int WRITE
```

此模式表示我们想写一个对象

```
public static final int READ
```

此模式表示我们想读一个对象

```
public static final int READWRITE
```

此模式是默认值，表示我们想对一个大对象进行读写访问

方法

```
public LargeObject open(int oid) throws SQLException
```

它根据 `OID` 打开一个已有大对象。此方法假定需要 `READ` 和 `WRITE` 访问（默认）。

参数：

`oid` - 大对象的

返回值：

提供对该对象访问的 `LargeObject` 实例

抛出： `SQLException`

出错时

```
public LargeObject open(int oid,  
                        int mode) throws SQLException
```

它根据 OID 打开一个已有大对象

参数:

oid - 大对象的  
mode - 打开模式

返回值:

提供对该对象访问的 LargeObject 实例

抛出: SQLException

出错时

```
public int create() throws SQLException
```

它创建一个大对象并返回其 OID。

新对象的属性默认为 READWRITE。

返回值:

新对象的 oid

抛出: SQLException

出错时

```
public int create(int mode) throws SQLException
```

它创建一个大对象并返回其 OID

参数:

mode - 描述新对象不同属性的位掩码

返回值:

新对象的 oid

抛出: SQLException

出错时

```
public void delete(int oid) throws SQLException
```

它删除一个大对象。

参数:

oid - 描述要删除的对象

抛出: SQLException

出错时

```
public void unlink(int oid) throws SQLException
```

它删除一个大对象。

它与 `delete` 方法相同，之所以提供是因为 C API 使用 `unlink`。

参数：

`oid` - 描述要删除的对象

抛出：SQLException

出错时

### 对象序列化

PostgreSQL 不是普通的 SQL 数据库。它比大多数其他数据库可扩展得多，并且支持它独有的面向对象特性。

其中一个后果是，可以让一个表引用另一个表中的某一行。例如：

```
test=> create table users (username name,fullname text);
CREATE
test=> create table server (servername name,adminuser users);
CREATE
test=> insert into users values ('peter','Peter Mount');
INSERT 2610132 1
test=> insert into server values ('maidast',2610132::users);
INSERT 2610133 1
test=> select * from users;
username|fullname
-----+-----
peter   |Peter Mount
(1 row)

test=> select * from server;
servername|adminuser
-----+-----
maidast   | 2610132
(1 row)
```

好，上面的例子表明我们可以把表名用作字段，该行的 `oid` 值就存储在这个字段中。

这与 Java 有什么关系？

在 Java 中，只要对象的类实现了 `java.io.Serializable` 接口，就可以把对象存储到流中。这一过程称为对象序列化，可用于把复杂对象存入数据库。

而在 JDBC 之下，你得用 `LargeObject` 来存储它们。然而，无法对这些对象执行查询。

`postgresql.util.Serialize` 类所做的事，就是提供一种把对象存成表、再从表中取回该对象的手段。大多数情况下你不需要直接访问这个类，而是使用 `PreparedStatement.setObject()` 和 `ResultSet.getObject()` 方法。这些方法会把对象的类名与数据库中的表比对。找到匹配时，就假定该对象是一个序列化对象，并从那个表取回它。在此过程中，如果对象还包含其他序列化对象，它会沿树递归。

听起来复杂？其实比我写的简单——只是难以解释。

唯一需要直接访问这个类的场合，是使用 `create()` 方法。驱动本身不使用它们，而是根据你要序列化的 Java 对象或类，向数据库发出一条或多条 `"create table"` 语句。

哦，最后一件事。如果你的对象里有这样一行：

```
public int oid;
```

那么，当对象从表中取回时，它会被设为表内的 `oid`。之后，如果对象被修改并重新序列化，既有条目会被更新。

如果没有 `oid` 变量，那么对象被序列化时总是插入表中，表中任何既有条目都会保留。

序列化之前把 `oid` 设为 0 也会导致对象被插入。这使得可以在数据库中复制一个对象。

```
Class postgresql.util.Serialize

java.lang.Object
|
+----postgresql.util.Serialize

public class Serialize extends Object
```

这个类利用 PostgreSQL 的面向对象特性来存储 Java 对象。做法是把 Java 类名映射为数据库中的一个表。新表中的每个条目代表该类的一个序列化实例。由于每个条目都有一个 `OID`（对象标识符），这个 `OID` 可以被包含在另一个表中。这太复杂，不在此展示，主文档中会有更详细的说明。

构造器

```
public Serialize(Connection c,
                  String type) throws SQLException
```

它创建一个实例，可用于对 PostgreSQL 表中的 Java 对象做序列化/反序列化。

方法

```
public Object fetch(int oid) throws SQLException
```

它根据 `OID` 从表中取回一个对象

参数：

`oid` - 对象的 `oid`

返回值：

与 `oid` 对应的对象

抛出：SQLException

出错时

```
public int store(Object o) throws SQLException
```

它把一个对象存入表中并返回其 `OID`。

如果对象有一个名为 `OID` 的 `int` 且其值  $> 0$ ，就用该值作 `OID`，表将被更新。如果 `OID` 的值为 0，则会新建一行，并且 `OID` 的值会被设置到对象中。这使对象在数据

库中的值可以更新。如果对象没有名为 `OID` 的 `int`，则直接存储对象。但若对象随后被取回、修改并再次存储，它的新状态会被追加到表中，而不会覆盖旧条目。

参数：

- - 要存储的对象（必须实现 `Serializable`）

返回值：

已存储对象的 `oid`

抛出： `SQLException`

出错时

```
public static void create(Connection con,
                          Object o) throws SQLException
```

驱动不使用此方法；它根据一个可序列化的 `Java` 对象创建表。应当在序列化任何对象之前使用它。

参数：

- `c` - 到数据库的连接
- - 表所基于的对象

抛出： `SQLException`

出错时

返回值：

与 `oid` 对应的对象

抛出： `SQLException`

出错时

```
public int store(Object o) throws SQLException
```

它把一个对象存入表中并返回其 `OID`。

如果对象有一个名为 `OID` 的 `int` 且其值  $> 0$ ，就用该值作 `OID`，表将被更新。如果 `OID` 的值为 `0`，则会新建一行，并且 `OID` 的值会被设置到对象中。这使对象在数据库中的值可以更新。如果对象没有名为 `OID` 的 `int`，则直接存储对象。但若对象随后被取回、修改并再次存储，它的新状态会被追加到表中，而不会覆盖旧条目。

参数：

- - 要存储的对象（必须实现 `Serializable`）

返回值：

已存储对象的 `oid`

抛出： `SQLException`

出错时

```
public static void create(Connection con,
                          Object o) throws SQLException
```

驱动不使用此方法；它根据一个可序列化的 `Java` 对象创建表。应当在序列化任何对象之前使用它。

参数:

- c - 到数据库的连接
- o - 表所基于的对象

抛出: SQLException

出错时

```
public static void create(Connection con,
                          Class c) throws SQLException
```

驱动不使用此方法; 它根据一个可序列化的 Java 对象创建表。应当在序列化任何对象之前使用它。

参数:

- c - 到数据库的连接
- o - 表所基于的类

抛出: SQLException

出错时

```
public static String toPostgreSQL(String name) throws SQLException
```

它把 . 替换为 \_, 从而把 Java 类名转换为 postgresql 表名

因此, 类名中不能含有 \_。

另一个限制是, 完整类名 (含包名) 不能超过 31 个字符 (这是 PostgreSQL 强加的限制)。

参数:

name - 类名

返回值:

PostgreSQL 表名

抛出: SQLException

出错时

```
public static String toClassName(String name) throws SQLException
```

它把 \_ 替换为 ., 从而把 postgresql 表名转换为 Java 类名

参数:

name - PostgreSQL 表名

返回值:

类名

抛出: SQLException

出错时

工具类

postgresql.util 包包含主驱动内部使用的类以及其他扩展使用的类。

```
Class postgresql.util.PGmoney
```

```
java.lang.Object
```

```
    |  
    +----postgresql.util.PGobject
```

```
        |  
        +----postgresql.util.PGmoney
```

```
    public class PGmoney extends PGobject implements Serializable,  
Cloneable
```

它实现一个处理 PostgreSQL money 类型的类

变量

```
public double val
```

字段的值

构造器

```
public PGmoney(double value)
```

参数:

value - 字段的值

```
public PGmoney(String value) throws SQLException
```

当某点嵌入其他几何类型的定义中时，主要由那些类型调用此构造器。

参数:

value - PostgreSQL 语法的该点定义

```
public PGmoney()
```

驱动所需

方法

```
public void setValue(String s) throws SQLException
```

参数:

s - PostgreSQL 语法的该点定义

抛出: SQLException

转换失败时

Overrides:

setValue in class PGobject

```
public boolean equals(Object obj)
```

参数:

obj - 要比较的对象

返回值:

如果两个矩形相同则为 true

Overrides:

equals in class PObject

```
public Object clone()
```

必须覆盖此方法才能克隆该对象

Overrides:

clone in class PObject

```
public String getValue()
```

返回值:

按 postgresql 所期望语法表示的 PGpoint

Overrides:

getValue in class PObject

```
Class postgresql.util.PObject
```

```
java.lang.Object
```

```
|
```

```
+----postgresql.util.PObject
```

```
public class PObject extends Object implements Serializable,  
Cloneable
```

这个类用于描述 JDBC 标准所不知道的数据类型。调用 `postgresql.Connection` 可以让一个继承本类的类与一个命名类型关联。`postgresql.geometric` 包就是这样工作的。对任何没有自己处理器的类型, `ResultSet.getObject()` 都会返回这个类。因此, 任何 `postgresql` 数据类型都得到支持。

构造器

```
public PObject()
```

它由 `postgresql.Connection.getObject()` 调用来创建对象。

方法

```
public final void setType(String type)
```

此方法设置该对象的类型。

子类不应扩展它, 因此它是 `final` 的

参数:

type - 描述该对象类型的字符串

```
public void setValue(String value) throws SQLException
```

此方法设置该对象的值。它必须被覆盖。

参数:

value - 对象值的字符串表示

抛出: SQLException

如果值对该类型无效则抛出

```
public final String getType()
```

由于它在对象生命周期内不能改变, 因此是 final 的。

返回值:

该对象的类型名

```
public String getValue()
```

必须覆盖它, 以 postgresql 所要求的形式返回对象的值。

返回值:

该对象的值

```
public boolean equals(Object obj)
```

必须覆盖它才能比较对象

参数:

obj - 要比较的对象

返回值:

如果两个矩形相同则为 true

Overrides:

equals in class Object

```
public Object clone()
```

必须覆盖此方法才能克隆该对象

Overrides:

clone in class Object

```
public String toString()
```

这里已定义它, 因此用户代码不必覆盖。

返回值:

按 postgresql 所期望语法表示的该对象的值

Overrides:

toString in class Object

```
Class postgresql.util.PGtokenizer  
  
java.lang.Object  
|  
+----postgresql.util.PGtokenizer  
  
public class PGtokenizer extends Object
```

这个类用于对 postgres 的文本输出做分词。

本可以用 StringTokenizer 来做，但我们需要处理 '(' ')' '[' ']' '<' 和 '>' 的嵌套，几何数据类型会用到它们。

它主要由几何类使用，但对解析 postgresql 自定义数据类型的任何输出也有用。

参见：

PGbox, PGcircle, PGIseg, PGpath, PGpoint, PGpolygon

### 构造器

```
public PGtokenizer(String string,  
                   char delim)
```

创建一个分词器。

参数：

string - 包含各词的字符串  
delim - 分割各词的单个字符

### 方法

```
public int tokenize(String string,  
                   char delim)
```

用一个新字符串和/或新分隔符重置该分词器。

参数：

string - 包含各词的字符串  
delim - 分割各词的单个字符

```
public int getSize()
```

返回值：

可用词的数量

```
public String getToken(int n)
```

参数：

n - 词的编号 (0 ... getSize()-1)

返回值：

词的值

```
public PGtokenizer tokenizeToken(int n,
                                char delim)
```

它基于我们的某个词返回一个新的分词器。几何数据类型用它处理嵌套的词（通常是 PGpoint）。

参数：

n - 词的编号 (0 ... getSize()-1)  
 delim - 要使用的分隔符

返回值：

基于该词的一个新 PGtokenizer 实例

```
public static String remove(String s,
                            String l,
                            String t)
```

移除一个字符串的前导/尾随字符串

参数：

s - 源字符串  
 l - 要移除的前导字符串  
 t - 要移除的尾随字符串

返回值：

不含前导/尾随字符串的字符串

```
public void remove(String l,
                   String t)
```

移除所有词的前导/尾随字符串

参数：

l - 要移除的前导字符串  
 t - 要移除的尾随字符串

```
public static String removePara(String s)
```

移除字符串开头和结尾的 ( 和 )

参数：

s - 要从中移除的字符串

返回值：

不含 ( 或 ) 的字符串

```
public void removePara()
```

移除所有词开头和结尾的 ( 和 )

返回值：

不含 ( 或 ) 的字符串

```
public static String removeBox(String s)
```

移除字符串开头和结尾的 [ 和 ]

参数:

s - 要从中移除的字符串

返回值:

不含 [ 或 ] 的字符串

```
public void removeBox()
```

移除所有词开头和结尾的 [ 和 ]

返回值:

不含 [ 或 ] 的字符串

```
public static String removeAngle(String s)
```

移除字符串开头和结尾的 < 和 >

参数:

s - 要从中移除的字符串

返回值:

不含 < 或 > 的字符串

```
public void removeAngle()
```

移除所有词开头和结尾的 < 和 >

返回值:

不含 < 或 > 的字符串

```
Class postgresql.util.Serialize
```

它已在前面“对象序列化”一节中说明。

```
Class postgresql.util.UnixCrypt
```

```
java.lang.Object
```

```
|
```

```
+----postgresql.util.UnixCrypt
```

```
public class UnixCrypt extends Object
```

这个类使我们能够在口令经网络流发送时对其加密

包含用于加密口令以及与 Unix 加密口令比对的静态方法。

原始源码见 John Dumas 的 Java Crypt 页面。

<http://www.zeh.com/local/jfd/crypt.html>

方法

```
public static final String crypt(String salt,
                                String original)
```

根据明文口令和一个“盐”加密口令。

参数：

`salt` - 一个两字符字符串，表示用于以多种方式迭代加密引擎的盐。如果要生成新加密，此值应当随机化。`original` - 要加密的口令。

返回值：

由 2 字符盐后接加密口令组成的字符串。

```
public static final String crypt(String original)
```

根据明文口令加密口令。此方法用 `'java.util.Random'` 类生成随机盐。

参数：

`original` - 要加密的口令。

返回值：

由 2 字符盐后接加密口令组成的字符串。

```
public static final boolean matches(String encryptedPassword,
                                    String enteredPassword)
```

检查 `enteredPassword` 加密后是否等于 `encryptedPassword`。

参数：

`encryptedPassword` - 已加密口令。假定前两个字符是盐。该字符串与 Unix `/etc/passwd` 文件中的形式相同。`enteredPassword` - 用户（或其他途径）输入的口令。

返回值：

如果口令应视为正确则为 `true`。

在多线程或 `Servlet` 环境中使用驱动

许多 JDBC 驱动都有一个问题：同一时刻只能有一个线程使用一个连接——否则一个线程可能在另一个线程接收结果时发出查询，这对数据库引擎是坏事。

PostgreSQL 6.4 为整个驱动带来了线程安全。6.3.x 中标准 JDBC 是线程安全的，但 Fastpath API 不是。

因此，如果你的应用使用多线程（大多数像样的应用都会），就不必为确保同一时刻只有一个线程使用数据库而操心复杂方案。

如果一个线程在另一个线程正在使用连接时试图使用它，它会等那个线程完成当前操作。

若是标准 SQL 语句，操作就是发送语句并（完整地）取回任何 `ResultSet`。

若是 Fastpath 调用（即从 `LargeObject` 读取一块），就是发送并取回该块的时间。

对应用和小程序来说这没问题，但可能给 `servlet` 带来性能问题。

对 `servlet` 而言，连接可能承受重载。如果多个线程执行查询，每个都会停顿，这可能不是你想要的。

为解决此问题，建议创建一个连接池。

每当线程需要使用数据库时，它向一个管理器类请求 `Connection`。管理器把一个空闲连接交给线程并标记为忙。如果没有空闲连接，就新开一个。

线程用完后把连接还给管理器，管理器可以关闭它或把它放回池中。管理器还会检查连接是否仍存活，若已死则从池中移除。

所以在 `servlet` 场景下，用单连接还是连接池由你决定。池的好处是线程不会受单一网络连接造成的瓶颈影响；坏处是服务器负载增加，因为每个 `Connection` 都会创建一个后端。

这取决于你和你的应用需求。

## 56.12. 延伸阅读

如果你还没有读过，我建议你阅读 JDBC API 文档（随 Sun 的 JDK 提供）以及 JDBC 规范。两者都可以在 JavaSoft 的网站<sup>2</sup>上找到。

我自己的网站<sup>3</sup>包含本文档未收录的更新信息，并且还提供 v6.4 及更早版本的预编译驱动。

---

<sup>2</sup><http://www.javasoft.com>

<sup>3</sup><http://www.retep.org.uk>

---

## 部分 VI. 开发者指南

开发者指南包括对设计决策的讨论以及对未来发展的建议。

---

# 目录

|                                               |     |
|-----------------------------------------------|-----|
| 57. 架构 .....                                  | 534 |
| 57.1. Postgres 体系结构概念 .....                   | 534 |
| 58. 数据库系统中的遗传查询优化 .....                       | 537 |
| 58.1. 将查询处理看成是一个复杂的优化问题 .....                 | 537 |
| 58.2. 遗传算法 (GA) .....                         | 537 |
| 58.3. Postgres 中的遗传查询优化 (GEQO) .....          | 538 |
| 58.4. Postgres GEQO 的未来实现任务 .....             | 538 |
| 58.4.1. 基本改进 .....                            | 538 |
| 58.4.2. 进一步改进 .....                           | 539 |
| 59. 前端/后端协议 .....                             | 540 |
| 59.1. Overview .....                          | 540 |
| 59.2. Protocol .....                          | 540 |
| 59.2.1. Startup .....                         | 541 |
| 59.2.2. Query .....                           | 542 |
| 59.2.3. 函数调用 .....                            | 543 |
| 59.2.4. Notification Responses .....          | 544 |
| 59.2.5. Cancelling Requests in Progress ..... | 544 |
| 59.2.6. Termination .....                     | 545 |
| 59.3. Message Data Types .....                | 545 |
| 59.4. 消息格式 .....                              | 545 |
| 60. Postgres 信号 .....                         | 553 |
| 61. gcc 默认优化 .....                            | 555 |
| 62. 后端接口 .....                                | 556 |
| 62.1. BKI 文件格式 .....                          | 556 |
| 62.2. 通用命令 .....                              | 556 |
| 62.3. 宏命令 .....                               | 557 |
| 62.4. 调试命令 .....                              | 557 |
| 62.5. 示例 .....                                | 558 |
| 63. 页文件 .....                                 | 559 |
| 63.1. 页结构 .....                               | 559 |
| 63.2. 文件 .....                                | 559 |
| 63.3. 缺陷 .....                                | 560 |

---

# 第 57 章 架构

## 57.1. Postgres 体系结构概念

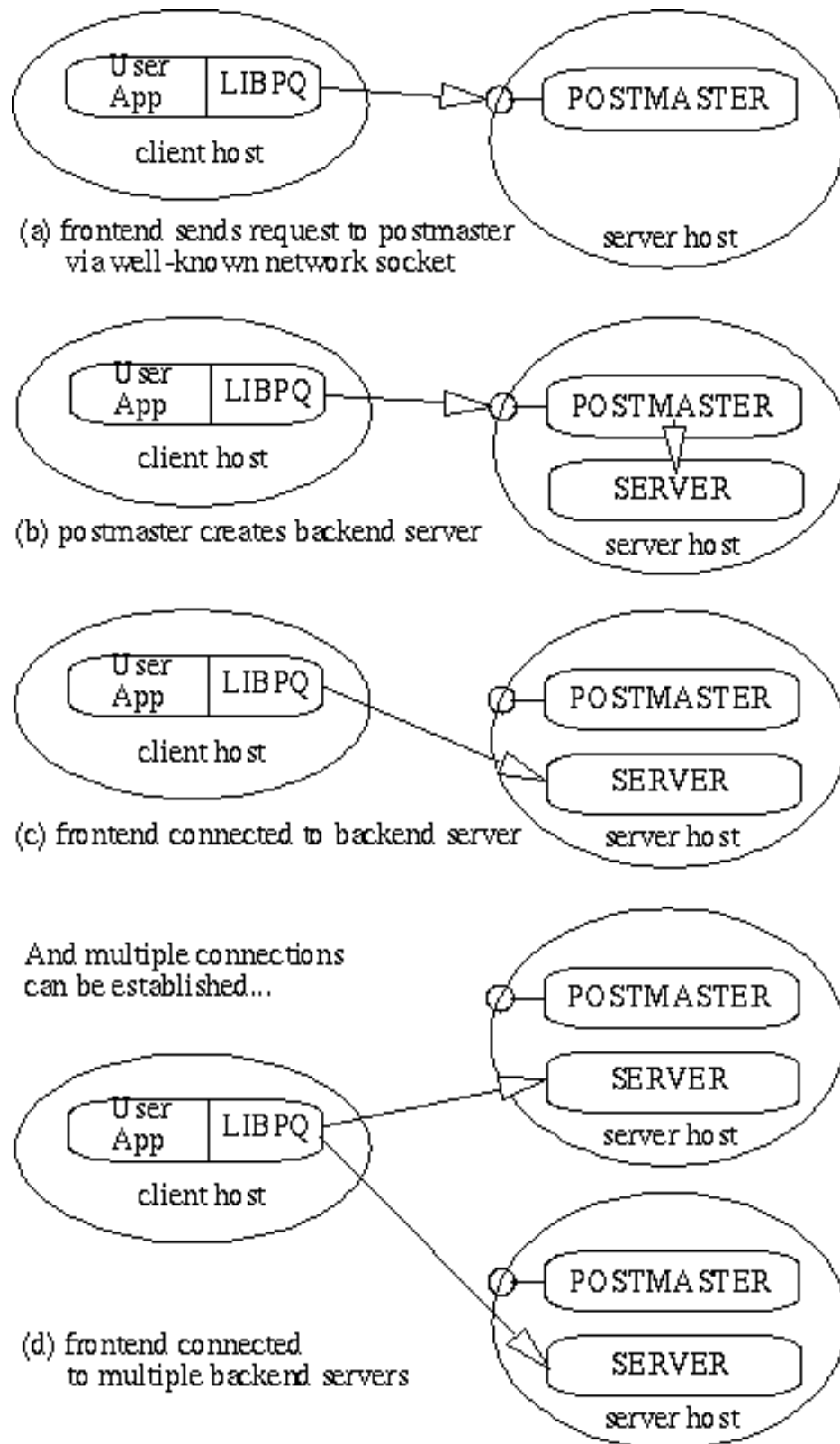
在继续之前，你应该先了解 Postgres 的基本体系结构。理解 Postgres 各部分之间如何交互，会让下一章的内容更加清晰。用数据库行话来说，Postgres 使用一种简单的“每用户一个进程”客户端/服务器模型。一个 Postgres 会话由下列相互协作的 UNIX 进程（程序）组成：

- 一个监督守护进程（postmaster），
- 用户的前端应用程序（例如 psql 程序），以及
- 一个或多个后端数据库服务器（postgres 进程本身）。

单个 postmaster 管理同一台主机上给定的一组数据库。这样一组数据库称为一个安装（installation）或站点（site）。

希望访问某个安装中特定数据库的前端应用程序调用相应的库。该库把用户请求通过网络发送给 postmaster（图 57.1(a)），后者再启动一个新的后端服务器进程（图 57.1(b)）

图 57.1. 连接是如何建立的



并把前端进程连接到这个新的服务器（图 57.1(c)）。从那一刻起，前端进程与后端服务器之间的通信不再经过 `postmaster` 干预。因此，`postmaster` 始终在运行并等待请求，而前端和后端进程则按需生灭。`libpq` 库允许单个前端与多个后端进程建立多条连接。但是，前端应用程序仍然是单线程进程。`libpq` 目前不支持多线程的前端/后端连接。这种体系结构的一个含义是：`postmaster` 和后端总是运行在同一台机器（数据库服务器）上，而前端应用程序可以运行在任何地方。你应当牢记这一点，因为在客户端机器上可以访问的文件，在数据库服务器机器上可能无法访问（或者只能用不同的文件名访问）。你还应当知道，`postmaster` 和 `postgres` 服务器以 Postgres “超级用户” 的用户 id 运行。注意，Postgres 超级用户并不需要是一个特殊的用户（例如名为 “postgres” 的用户）。而且，Postgres 超级用户绝对不应该是 UNIX 超级用户 “root”！无论如何，与某个数据库相关的所有文件都应当属于这个 Postgres 超级用户。

# 第 58 章 数据库系统中的遗传查询优化

Martin Utesch, University of Mining and Technology

## 作者

由 Martin Utesch<sup>1</sup> 为德国弗赖贝格矿业和技术大学自动控制研究所编写。

## 58.1. 将查询处理看成是一个复杂的优化问题

在所有关系操作符中，最难处理和优化的是连接。随着查询中join数目的增加，可能的查询计划数量会呈指数增长。为了处理单个连接而支持多种连接方法（例如 Postgres 中的嵌套循环、索引扫描和归并连接）来处理单个 join，以及作为关系访问路径的多种索引（例如 Postgres 中的 r-树、b-树和哈希），也进一步增加了优化工作量。

当前的Postgres查询优化器实现会在可选策略空间中执行近似穷举搜索。这种查询优化技术不足以支持诸如人工智能这类需要大量查询的数据库应用领域。

德国弗赖贝格矿业和技术大学自动控制研究所在尝试将 Postgres 用作一个用于电网维护的基于知识的决策支持系统后端时遇到了一些问题。该 DBMS 需要为该基于知识系统的推理机处理大型 join 查询。

在探索可能查询计划空间内部的性能困难，产生了开发一种新优化技术的需求。

下文我们提出把遗传算法的实现作为数据库查询优化问题的一种可选方案。

## 58.2. 遗传算法（GA）

GA 是一种通过确定的、随机化的搜索进行工作的启发式优化方法。优化问题的可能解集合被视为由若干个体组成的种群。个体对其环境的适应程度由其适应度指定。

一个个体在搜索空间中的坐标由染色体表示，本质上是一组字符串。基因是染色体的一个片段，它编码某个待优化参数的值。基因的典型编码可以是二进制或整数。

通过模拟重组、变异和选择这些进化操作，可以找到平均适应度高于前代的新一代搜索点。

根据"comp.ai.genetic" FAQ 的说法，再怎么强调也不过分：GA并不是为了求解问题而进行的纯粹随机搜索。GA会使用随机过程，但其结果显然并非随机的（优于随机）。

Structured Diagram of a GA:

```
P(t)      generation of ancestors at a time t
P'(t)     generation of descendants at a time t
```

```
+=====+
|>>>>>>>>> Algorithm GA <<<<<<<<<<<<<|
+=====+
| INITIALIZE t := 0                               |
+=====+
```

<sup>1</sup>utesch@aut.tu-freiberg.de

```

| INITIALIZE P(t) |
+=====+
| evaluate FITNESS of P(t) |
+=====+
| while not STOPPING CRITERION do |
|   +-----+ |
|   | P'(t) := RECOMBINATION{P(t)} | |
|   +-----+ |
|   | P''(t) := MUTATION{P'(t)} | |
|   +-----+ |
|   | P(t+1) := SELECTION{P''(t) + P(t)} | |
|   +-----+ |
|   | evaluate FITNESS of P''(t) | |
|   +-----+ |
|   | t := t + 1 | |
+=====+

```

## 58.3. Postgres 中的遗传查询优化（GEQO）

GEQO模块旨在以类似于旅行商问题（TSP）的方式解决查询优化问题。

可能的查询计划被编码为整数串。每个串表示查询中从一个关系到下一个关系的 join 顺序。例如，查询树

```

      /\
     /\ 2
    /\ 3
   4  1

```

会被编码为整数串 '4-1-3-2'，这意味着先连接关系 '4' 和 '1'，再连接 '3'，最后连接 '2'；其中 1、2、3、4 是 Postgres 中的 relid。

GEQO模块的部分内容改编自 D. Whitley 的 Genitor 算法。

Postgres 中 GEQO 实现的具体特征是：

- 采用稳态 GA（替换种群中适应度最低的个体，而不是整代替换），能够快速收敛到更优的查询计划。这对于在合理时间内处理查询至关重要；
- 采用边重组交叉，它特别适合于在利用 GA 求解 TSP 时将边损失保持在较低水平；
- 不使用变异作为遗传操作符，因此无须借助修复机制来生成合法的 TSP 回路。

与 Postgres 查询优化器实现相比，GEQO 模块为 Postgres DBMS 带来以下益处：

- 通过非穷举搜索处理大型 join 查询；
- 改进了查询计划的代价大小估算，因为不再需要计划合并（GEQO 模块把一个查询计划的代价作为一个个体来评估）。

## 58.4. Postgres GEQO 的未来实现任务

### 58.4.1. 基本改进

#### 58.4.1.1. 改进查询处理完成后的内存释放

对于大型 join 查询，遗传查询优化所花费的计算时间似乎只是 Postgres 通过例程 MemoryContextFree（文件 backend/utils/mmgr/mcxt.c）释放内存所用时间的小分数。调

试显示它卡在了例程 `OrderedElemPop` (文件 `backend/utils/mmgr/oset.c`) 的一个循环中。使用普通 Postgres 查询优化算法处理长查询时也会出现同样的问题。

#### 58.4.1.2. 改进遗传算法参数设置

在文件 `backend/optimizer/geqo/geqo_params.c` 的例程 `gimme_pool_size` 和 `gimme_number_generations` 中, 我们必须为参数设置找到一种折中, 以满足两个相互竞争的需求:

- 查询计划的最优性
- 计算时间

#### 58.4.1.3. 为整数溢出寻找更好的解决方案

在文件 `backend/optimizer/geqo/geqo_eval.c` 的例程 `geqo_joinrel_size` 中, 目前对 `MAXINT` 溢出的临时做法是把 Postgres 整数值 `rel->size` 设为其对数。对 `backend/nodes/relation.h` 中 `Rel` 的修改肯定会对整个 Postgres 实现产生严重影响。

#### 58.4.1.4. 为内存耗尽寻找解决方案

当查询涉及的关系超过 10 个时可能发生内存耗尽。在文件 `backend/optimizer/geqo/geqo_eval.c` 中, 例程 `gimme_tree` 会被递归调用。也许我忘了正确释放某些东西, 但我不知道是什么。当然, `join` 的 `rel` 数据结构会随着装进去的关系越来越多而不断增长。欢迎建议 :-)

### 58.4.2. 进一步改进

在 Postgres 中启用浓密查询树处理; 这可能会改进查询计划的质量。

## 参考文献

GEQ 算法的参考信息。

The Hitch-Hiker's Guide to Evolutionary Computation . Jörg Heitkötter和David Beasley. InterNet resource .

`comp.ai.genetic`<sup>1</sup> 中的 FAQ 见 `Encore`<sup>2</sup>。

The Design and Implementation of the Postgres Query Optimizer . Z. Fong. University of California, Berkeley Computer Science Department .

文件 `planner/Report.ps` (在 'postgres-papers' 发行版中)。

Fundamentals of Database Systems . R. Elmasri和S. Navathe. The Benjamin/Cummings Pub., Inc. .

<sup>1</sup>[news://comp.ai.genetic](http://news://comp.ai.genetic)

<sup>2</sup>[ftp://ftp.Germany.EU.net/pub/research/softcomp/EC/Welcome.html](http://ftp://ftp.Germany.EU.net/pub/research/softcomp/EC/Welcome.html)

---

# 第 59 章 前端/后端协议

Phil Thompson

## 注意

由 Phil Thompson 编写。协议 2.0 的更新由 Tom Lane 完成。

Postgres 使用一种基于消息的协议在frontend和backend 之间通信。The protocol is implemented over TCP/IP and also on Unix sockets. Postgres v6.3 introduced version numbers into the protocol. This was done in such a way as to still allow connections from earlier versions of frontends, but this document does not cover the protocol used by those earlier versions.

This document describes version 2.0 of the protocol, implemented in Postgres v6.4 and later.

Higher level features built on this protocol (for example, how `libpq` passes certain environment variables after the connection is established) are covered elsewhere.

## 59.1. Overview

The three major components are the frontend (running on the client) and the postmaster and backend (running on the server). The postmaster and backend have different roles but may be implemented by the same executable.

A frontend sends a startup packet to the postmaster. This includes the names of the user and the database the user wants to connect to. The postmaster then uses this, and the information in the `pg_hba.conf(5)` file to determine what further authentication information it requires the frontend to send (if any) and responds to the frontend accordingly.

The frontend then sends any required authentication information. Once the postmaster validates this it responds to the frontend that it is authenticated and hands over the connection to a backend. The backend then sends a message indicating successful startup (normal case) or failure (for example, an invalid database name).

随后的通信是前端与后端之间交换的查询和结果包。postmaster 不再参与普通的查询/结果通信。(However, the postmaster is involved when the frontend wishes to cancel a query currently being executed by its backend. Further details about that appear below.)

When the frontend wishes to disconnect it sends an appropriate packet and closes the connection without waiting for a response for the backend.

Packets are sent as a data stream. The first byte determines what should be expected in the rest of the packet. The exception is packets sent from a frontend to the postmaster, which comprise a packet length then the packet itself. The difference is historical.

## 59.2. Protocol

This section describes the message flow. There are four different types of flows depending on the state of the connection: startup, query, function call, and termination. There are

also special provisions for notification responses and command cancellation, which can occur at any time after the startup phase.

## 59.2.1. Startup

Startup is divided into an authentication phase and a backend startup phase.

Initially, the frontend sends a StartupPacket. The postmaster uses this info and the contents of the pg\_hba.conf(5) file to determine what authentication method the frontend must use. The postmaster then responds with one of the following messages:

### ErrorResponse

然后 postmaster 立即关闭连接。

### AuthenticationOk

然后 postmaster 移交给后端。postmaster 不再参与通信。

### AuthenticationKerberosV4

然后前端必须与 postmaster 进行 Kerberos V4 认证对话（此处不描述）。如果成功，postmaster 以 AuthenticationOk 响应，否则以 ErrorResponse 响应。

### AuthenticationKerberosV5

然后前端必须与 postmaster 进行 Kerberos V5 认证对话（此处不描述）。如果成功，postmaster 以 AuthenticationOk 响应，否则以 ErrorResponse 响应。

### AuthenticationUnencryptedPassword

然后前端必须发送一个 UnencryptedPasswordPacket。如果密码正确，postmaster 以 AuthenticationOk 响应，否则以 ErrorResponse 响应。

### AuthenticationEncryptedPassword

然后前端必须发送一个 EncryptedPasswordPacket。如果密码正确，postmaster 以 AuthenticationOk 响应，否则以 ErrorResponse 响应。

如果前端不支持 postmaster 请求的认证方法，那么它应当立即关闭连接。

After sending AuthenticationOk, the postmaster attempts to launch a backend process. Since this might fail, or the backend might encounter a failure during startup, the frontend must wait for the backend to acknowledge successful startup. The frontend should send no messages at this point. The possible messages from the backend during this phase are:

### BackendKeyData

This message is issued after successful backend startup. It provides secret-key data that the frontend must save if it wants to be able to issue cancel requests later. The frontend should not respond to this message, but should continue listening for a ReadyForQuery message.

### ReadyForQuery

后端启动成功。前端现在可以发出查询或函数调用消息。

### ErrorResponse

后端启动失败。发送此消息后连接被关闭。

### NoticeResponse

发出了警告消息。前端应显示该消息，但继续等待 ReadyForQuery 或 ErrorResponse。

ReadyForQuery 消息与后端在每个查询周期之后发出的消息相同。Depending on the coding needs of the frontend, it is reasonable to consider ReadyForQuery as starting a query cycle (and then BackendKeyData indicates successful conclusion of the startup phase), or to consider ReadyForQuery as ending the startup phase and each subsequent query cycle.

## 59.2.2. Query

查询周期由前端向 backend 发送 Query 消息发起。然后后端根据查询命令串的内容发送一个或多个响应消息，最后发送 ReadyForQuery 响应消息。ReadyForQuery informs the frontend that it may safely send a new query or function call.

The possible response messages from the backend are:

### CompletedResponse

An SQL command completed normally.

### CopyInResponse

后端准备好把数据从前端复制到关系。前端然后应发送 CopyDataRows 消息。后端随后将以带 "COPY" 标签的 CompletedResponse 消息响应。

### CopyOutResponse

后端准备好把数据从关系复制到前端。然后它发送 CopyDataRows 消息，再发送带 "COPY" 标签的 CompletedResponse 消息。

### CursorResponse

查询是 insert(l)、delete(l)、update(l)、fetch(l) 或 select(l) 命令之一。如果事务已被中止，则后端发送带 "\*ABORT STATE\*" 标签的 CompletedResponse 消息。否则发送以下响应。

对于 insert(l) 命令，后端随后发送一条带 "INSERT *oid rows*" 标签的 CompletedResponse 消息，其中 *rows* 是插入的行数，*oid* 在 *rows* 为 1 时是所插入行的对象 ID，否则 *oid* 为 0。

对于 delete(l) 命令，后端随后发送一条带 "DELETE *rows*" 标签的 CompletedResponse 消息，其中 *rows* 是删除的行数。

对于 update(l) 命令，后端随后发送一条带 "UPDATE *rows*" 标签的 CompletedResponse 消息，其中 *rows* 是更新的行数。

对于 fetch(l) 或 select(l) 命令，后端发送一条 RowDescription 消息。然后对返回给前端的每一行，跟着一条 AsciiRow 或 BinaryRow 消息（取决于是否指定了二进制游标）。最后，后端发送一条带 "SELECT" 标签的 CompletedResponse 消息。

### EmptyQueryResponse

识别到一个空查询串。（需要特别区分这种情况是历史原因。）

#### ErrorResponse

发生了一个错误。

#### ReadyForQuery

查询串的处理完成。发送单独的消息来指示这一点，因为查询串可能包含多条 SQL 命令。(CompletedResponse marks the end of processing one SQL command, not the whole string.) ReadyForQuery will always be sent, whether processing terminates successfully or with an error.

#### NoticeResponse

发出了与查询有关的警告消息。通知是其他响应之外的附加内容，即后端将继续处理该命令。

每当期待任何其他类型的消息时，前端必须准备好接受 ErrorResponse 和 NoticeResponse 消息。

Actually, it is possible for NoticeResponse to arrive even when the frontend is not expecting any kind of message, that is, the backend is nominally idle. (In particular, the backend can be commanded to terminate by its postmaster. In that case it will send a NoticeResponse before closing the connection.) It is recommended that the frontend check for such asynchronous notices just before issuing any new command.

此外，如果前端发出任何 listen(l) 命令，那么它必须准备好在任何时间接受 Notification Response 消息；见下文。

### 59.2.3. 函数调用

A Function Call cycle is initiated by the frontend sending a FunctionCall message to the backend. The backend then sends one or more response messages depending on the results of the function call, and finally a ReadyForQuery response message. ReadyForQuery informs the frontend that it may safely send a new query or function call.

The possible response messages from the backend are:

#### ErrorResponse

发生了一个错误。

#### FunctionResultResponse

函数调用已执行并返回了结果。

#### FunctionVoidResponse

函数调用已执行但没有返回结果。

#### ReadyForQuery

函数调用的处理完成。无论处理成功还是出错，ReadyForQuery 都将被发送。

#### NoticeResponse

发出了与函数调用有关的警告消息。通知是其他响应之外的附加内容，即后端将继续处理该命令。

每当期待任何其他类型的消息时，前端必须准备好接受 ErrorResponse 和 NoticeResponse 消息。Also, if it issues any listen(l) commands then it must be prepared to accept Notification Response messages at any time; see below.

## 59.2.4. Notification Responses

If a frontend issues a listen(l) command, then the backend will send a NotificationResponse message (not to be confused with NoticeResponse!) whenever a notify(l) command is executed for the same notification name.

Notification responses are permitted at any point in the protocol (after startup), except within another backend message. Thus, the frontend must be prepared to recognize a NotificationResponse message whenever it is expecting any message. Indeed, it should be able to handle NotificationResponse messages even when it is not engaged in a query.

NotificationResponse

A notify(l) command has been executed for a name for which a previous listen(l) command was executed. Notifications may be sent at any time.

值得指出的是，listen 和 notify 命令中使用的名字不必与 SQL 数据库中关系（表）的名字有任何关系。Notification names are simply arbitrarily chosen condition names.

## 59.2.5. Cancelling Requests in Progress

During the processing of a query, the frontend may request cancellation of the query by sending an appropriate request to the postmaster. The cancel request is not sent directly to the backend for reasons of implementation efficiency: we don't want to have the backend constantly checking for new input from the frontend during query processing. Cancel requests should be relatively infrequent, so we make them slightly cumbersome in order to avoid a penalty in the normal case.

To issue a cancel request, the frontend opens a new connection to the postmaster and sends a CancelRequest message, rather than the StartupPacket message that would ordinarily be sent across a new connection. The postmaster will process this request and then close the connection. For security reasons, no direct reply is made to the cancel request message.

A CancelRequest message will be ignored unless it contains the same key data (PID and secret key) passed to the frontend during connection startup. If the request matches the PID and secret key for a currently executing backend, the postmaster signals the backend to abort processing of the current query.

The cancellation signal may or may not have any effect --- for example, if it arrives after the backend has finished processing the query, then it will have no effect. If the cancellation is effective, it results in the current command being terminated early with an error message.

The upshot of all this is that for reasons of both security and efficiency, the frontend has no direct way to tell whether a cancel request has succeeded. It must continue to wait for the backend to respond to the query. Issuing a cancel simply improves the odds that the current query will finish soon, and improves the odds that it will fail with an error message instead of succeeding.

Since the cancel request is sent to the postmaster and not across the regular frontend/backend communication link, it is possible for the cancel request to be issued by any process, not just the frontend whose query is to be canceled. This may have some benefits

of flexibility in building multiple-process applications. It also introduces a security risk, in that unauthorized persons might try to cancel queries. The security risk is addressed by requiring a dynamically generated secret key to be supplied in cancel requests.

### 59.2.6. Termination

The normal, graceful termination procedure is that the frontend sends a Terminate message and immediately closes the connection. On receipt of the message, the backend immediately closes the connection and terminates.

不优雅的终止可能由于任一端的软件故障（即核心转储）而发生。

如果前端或后端看到连接意外关闭，应当清理并终止。The frontend has the option of launching a new backend by recontacting the postmaster, if it doesn't want to terminate itself.

## 59.3. Message Data Types

本节描述消息中使用的基本数据类型。

`Intn(i)`

一个按网络字节序的  $n$  位整数。如果指定了  $i$ ，它就是字面值。Eg. `Int16`, `Int32(42)`.

`LimStringn(s)`

恰好  $n$  字节的字符数组，解释为以 `'\0'` 结尾的字符串。如果空间不足则省略 `'\0'`。如果指定了  $s$ ，它就是字面值。例如 `LimString32`、`LimString64("user")`。

`String(s)`

常规的以 `'\0'` 结尾的 C 字符串，没有长度限制。即使缓冲区不够大而必须丢弃某些字符，前端也应始终读取完整的字符串。

#### 注意

8193 字节是允许的最大尺寸吗？

If  $s$  is specified it is the literal value. Eg. `String`, `String("user")`.

`Byten(c)`

恰好  $n$  字节。如果指定了  $c$ ，它就是字面值。例如 `Byte`、`Byte1('\n')`。

## 59.4. 消息格式

本节描述每条消息的详细格式。每条消息都可以由 frontend (F)、postmaster/backend (B) 或两者 (F & B) 发送。

`AsciiRow (B)`

`Byte1('D')`

标识此消息为 ASCII 数据行。（先前的 `RowDescription` 消息定义了行中字段的数量及其数据类型。）

Byte<sub>n</sub>

一个位图，行中每个字段一位。第 1 个字段对应第 1 个字节的第 7 位（MSB），第 2 个字段对应第 1 个字节的第 6 位，第 8 个字段对应第 1 个字节的第 0 位（LSB），第 9 个字段对应第 2 个字节的第 7 位，依此类推。如果对应字段的值不是 NULL，则每个位被置位。如果字段数不是 8 的倍数，位图中最后一个字节的剩余部分就被浪费了。

然后，对每个非 NULL 值的字段，有下列内容：

Int32

Specifies the size of the value of the field, including this size.

Byte<sub>n</sub>

以 ASCII 字符指定字段本身的值。n 是上面的尺寸减 4。字段数据中没有尾随的 '\0'；如果前端需要，必须自己添加一个。

AuthenticationOk (B)

Byte1('R')

标识此消息为认证请求。

Int32(0)

指定认证成功。

AuthenticationKerberosV4 (B)

Byte1('R')

标识此消息为认证请求。

Int32(1)

指定需要 Kerberos V4 认证。

AuthenticationKerberosV5 (B)

Byte1('R')

标识此消息为认证请求。

Int32(2)

指定需要 Kerberos V5 认证。

AuthenticationUnencryptedPassword (B)

Byte1('R')

标识此消息为认证请求。

Int32(3)

指定需要未加密的密码。

### AuthenticationEncryptedPassword (B)

Byte1('R')

标识此消息为认证请求。

Int32(4)

指定需要加密的密码。

Byte2

加密密码时使用的盐。

### BackendKeyData (B)

Byte1('K')

标识此消息为取消密钥数据。如果前端希望以后能够发出 CancelRequest 消息，就必须保存这些值。

Int32

此后端的进程 ID。

Int32

此后端的密钥。

### BinaryRow (B)

Byte1('B')

标识此消息为二进制数据行。（先前的 RowDescription 消息定义了行中字段的数量及其数据类型。）

Byte<sub>n</sub>

一个位图，行中每个字段一位。第 1 个字段对应第 1 个字节的第 7 位（MSB），第 2 个字段对应第 1 个字节的第 6 位，第 8 个字段对应第 1 个字节的第 0 位（LSB），第 9 个字段对应第 2 个字节的第 7 位，依此类推。如果对应字段的值不是 NULL，则每个位被置位。如果字段数不是 8 的倍数，位图中最后一个字节的剩余部分就被浪费了。

然后，对每个非 NULL 值的字段，有下列内容：

Int32

Specifies the size of the value of the field, excluding this size.

Byte<sub>n</sub>

以二进制格式指定字段本身的值。n 是上面的尺寸。

### CancelRequest (F)

Int32(16)

包的尺寸（字节）。

Int32(80877102)

取消请求码。该值经过选择，使其高 16 位包含 1234，低 16 位包含 5678。(To avoid confusion, this code must not be the same as any protocol version number.)

Int32

目标后端的进程 ID。

Int32

目标后端的密钥。

CompletedResponse (B)

Byte1('C')

标识此消息为完成响应。

String

命令标签。这通常是（但不总是）标识完成了哪条 SQL 命令的单个词。

CopyDataRows (B & F)

这是一个行流，其中每行以 Byte1('\n') 结束。然后跟随序列 Byte1('\\')、Byte1('.')、Byte1('\n')。

CopyInResponse (B)

Byte1('G')

标识此消息为 Start Copy In 响应。前端现在必须发送 CopyDataRows 消息。

CopyOutResponse (B)

Byte1('H')

标识此消息为 Start Copy Out 响应。此消息后面将跟随 CopyDataRows 消息。

CursorResponse (B)

Byte1('P')

标识此消息为游标响应。

String

游标的名称。如果游标是隐式的，则为“blank”。

EmptyQueryResponse (B)

Byte1('I')

标识此消息为对空查询串的响应。

String("")

未使用。

EncryptedPasswordPacket (F)

Int32

包的尺寸（字节）。

String

加密（使用 crypt()）的密码。

ErrorResponse (B)

Byte1('E')

标识此消息为错误。

String

错误消息本身。

FunctionCall (F)

Byte1('F')

标识此消息为函数调用。

String("")

未使用。

Int32

指定要调用的函数的对象 ID。

Int32

指定提供给函数的参数数量。

然后，对每个参数，有下列内容：

Int32

Specifies the size of the value of the argument, excluding this size.

Byte<sub>n</sub>

以二进制格式指定字段本身的值。n 是上面的尺寸。

FunctionResultResponse (B)

Byte1('V')

标识此消息为函数调用结果。

Byte1('G')

指定返回了非空结果。

Int32

指定结果的值的尺寸，不包括此尺寸本身。

Byte $n$

以二进制格式指定结果本身的值。 $n$  是上面的尺寸。

Byte1('0')

未使用。（严格说来，FunctionResultResponse 和 FunctionVoidResponse 是同一个东西，只是消息的某些部分是可选的。）

FunctionVoidResponse (B)

Byte1('V')

标识此消息为函数调用结果。

Byte1('0')

指定返回了空结果。

NoticeResponse (B)

Byte1('N')

标识此消息为通知。

String

通知消息本身。

NotificationResponse (B)

Byte1('A')

标识此消息为通知响应。

Int32

发出通知的后端进程的进程 ID。

String

引发此通知的条件名称。

Query (F)

Byte1('Q')

标识此消息为查询。

String

查询串本身。

ReadyForQuery (B)

Byte1('Z')

标识消息类型。每当后端准备好进入新的查询周期时发送 ReadyForQuery。

RowDescription (B)

Byte1('T')

标识此消息为行描述。

Int16

指定一行中的字段数（可以为零）。

Then, for each field, there is the following:

String

指定字段名。

Int32

指定字段类型的对象 ID。

Int16

指定类型尺寸。

Int32

指定类型修饰符。

StartupPacket (F)

Int32(296)

包的尺寸（字节）。

Int32

协议版本号。高 16 位是主版本号。低 16 位是次版本号。

LimString64

数据库名，如果为空则默认为用户名。

LimString32

用户名。

LimString64

要由 postmaster 传递给后端的任何附加命令行参数。

LimString64

未使用。

LimString64

后端应用于调试消息的可选 tty。

Terminate (F)

Byte1('X')

标识此消息为终止。

UnencryptedPasswordPacket (F)

Int32

包的尺寸（字节）。

String

未加密的密码。

# 第 60 章 Postgres 信号

Massimo Dal Zotto

注意

由 Massimo Dal Zotto<sup>1</sup> 贡献

Postgres 使用下列信号在 postmaster 和后端之间通信：

表 60.1. Postgres 信号

| 信号      | postmaster 动作                   | 服务器动作                   |
|---------|---------------------------------|-------------------------|
| SIGHUP  | kill(*,sighup)                  | read_pg_options         |
| SIGINT  | die                             | cancel query            |
| SIGQUIT | kill(*,sigterm)                 | handle_warn             |
| SIGTERM | kill(*,sigterm), kill(*,9), die | die                     |
| SIGPIPE | ignored                         | die                     |
| SIGUSR1 | kill(*,sigusr1), die            | quickdie                |
| SIGUSR2 | kill(*,sigusr2)                 | async notify (SI flush) |
| SIGCHLD | reaper                          | ignored (alive test)    |
| SIGTTIN | ignored                         |                         |
| SIGTTOU | ignored                         |                         |
| SIGCONT | dumpstatus                      |                         |
| SIGFPE  |                                 | FloatExceptionHandler   |

注意

“kill(\*,signal)” 表示向所有后端发送一个信号。

与旧信号处理方式相比的主要变化是：用 SIGQUIT 而不是 SIGHUP 来处理 warn，用 SIGHUP 重新读取 pg\_options 文件，以及把发送给 postmaster 的 SIGHUP、SIGTERM、SIGUSR1 和 SIGUSR2 重定向到所有活动后端。这样，发送给 postmaster 的这些信号就可以自动发送给所有后端，而无需知道它们的 pid。要关闭 postgres，只需向 postmaster 发送一个 SIGTERM，它会自动停止所有后端。

SIGUSR2 信号还用于防止 SI 缓存表溢出，这种情况发生在某个后端长时间不处理 SI 缓存时。当一个后端检测到 SI 表在 70% 满时，它就向 postmaster 发送一个信号，后者会唤醒所有空闲的后端并让它们刷写缓存。

程序员对信号的典型用法可能如下：

```
# stop postgres
kill -TERM $postmaster_pid
```

<sup>1</sup><mailto:dz@cs.unitn.it>

```
# kill all the backends
kill -QUIT $postmaster_pid

# kill only the postmaster
kill -INT $postmaster_pid

# change pg_options
cat new_pg_options > $DATA_DIR/pg_options
kill -HUP $postmaster_pid

# change pg_options only for a backend
cat new_pg_options > $DATA_DIR/pg_options
kill -HUP $backend_pid
cat old_pg_options > $DATA_DIR/pg_options
```

# 第 61 章 gcc 默认优化

Brian Gallew

## 注意

由 Brian Gallew<sup>1</sup> 贡献

把 gcc 配置成默认使用某些标志只是编辑 `/usr/local/lib/gcc-lib/platform/version/specs` 文件的一件简单事。这个文件的格式相当简单。文件分成若干节，每节三行。第一行是 `"*section_name:"`（例如 `"*asm:"`）。第二行是标志列表，第三行为空。

最容易的修改是把想要的默认标志追加到相应节的列表中。作为示例，假设我在一台 '486 上运行 linux，gcc 2.7.2 安装在默认位置。在文件 `/usr/local/lib/gcc-lib/i486-linux/2.7.2/specs` 的第 13 行处我找到下面这一节：

```
- -----SECTION-----  
*cc1:
```

```
- -----SECTION-----
```

如你所见，没有任何默认标志。如果我希望 C 代码的编译总是使用 `"-m486 -fomit-frame-pointer"`，我会把它改成这样：

```
- -----SECTION-----  
*cc1:  
- -m486 -fomit-frame-pointer
```

```
- -----SECTION-----
```

如果我还想为另一台闲置的较旧 linux 机器生成 386 代码，就必须把它改成这样：

```
- -----SECTION-----  
*cc1:  
%{!m386:-m486} -fomit-frame-pointer
```

```
- -----SECTION-----
```

这样就会总是省略帧指针，并且在命令行上未指定 `-m386` 时生成 486 优化的代码。

实际上你可以用 `specs` 文件做相当多的定制。但始终要记住，这些更改是全局的，会影响系统的所有用户。

<sup>1</sup><mailto:geek+@cmu.edu>

---

## 第 62 章 后端接口

后端接口 (BKI) 文件是这样一些脚本：它们是运行在特殊"引导"模式下的 Postgres 后端的输入，这种模式使后端能够在数据库系统尚不存在的情况下执行数据库功能。因此 BKI 文件可以用于最初创建数据库系统。initdb 正是使用 BKI 文件来完成这件事：创建数据库系统。不过，initdb 的 BKI 文件是内部生成的。它使用 `global1.bki.source` 和 `local1.template1.bki.source` 这两个文件来生成它们，这两个文件可以在 Postgres 的"库"目录中找到。它们作为安装 Postgres 的一部分被安装到那里。这些 .source 文件作为 Postgres 构建过程的一部分，由一个名为 `genbki` 的构建程序构建。`genbki` 接收 Postgres 源文件作为输入，这些源文件同时充当 `genbki` 的输入，用以构建表以及描述这些表的 C 头文件。

相关信息可以在下列文档中找到：initdb、createdb，以及 SQL 命令 `CREATE DATABASE`。

### 62.1. BKI 文件格式

Postgres 后端按下面的描述解释 BKI 文件。这段说明在手边有一份 `global1.bki.source` 文件作为示例时会更容易理解。（如上面所解释的，这个 .source 文件并不完全是一个 BKI 文件，但你还是能猜出得到的 BKI 文件会是什么样子。）

命令由命令名后跟以空格分隔的参数组成。以 "\$" 开头的命令参数会被特殊处理。如果 "\$\$" 是前两个字符，那么第一个 "\$" 被忽略，该参数随后按正常方式处理。如果 "\$" 后面跟的是空格，那么它被当作一个 NULL 值。否则，"\$" 后面的字符被解释为一个宏的名字，使该参数被替换为该宏的值。如果这个宏未定义，则是一个错误。

宏用

```
define macro macro_name = macro_value
```

定义，用

```
undefine macro macro_name
```

取消定义，重新定义使用与 `define` 相同的语法。

下面是通用命令和宏命令的列表。

### 62.2. 通用命令

`OPEN classname`

打开名为 `classname` 的类以便进一步操作。

`CLOSE [classname]`

关闭名为 `classname` 的已打开类。如果 `classname` 尚未打开，则这是一个错误。如果没有给出 `classname`，则关闭当前打开的类。

`PRINT`

打印当前打开的类。

`INSERT [OID=oid_value] (value1 value2 ...)`

用 `value1`、`value2`、等作为属性值，用 `oid_value` 作为其 OID，向已打开的类中插入一个新的实例。如果 `oid_value` 不为 "0"，则该值将被用作该实例的对象标识符。否则，这是一个错误。

INSERT (*value1 value2 ...*)

同上，但由系统生成一个唯一的对象标识符。

CREATE *classname* (*name1 = type1 [,name2 = type2[,...]]*)

创建一个名为 *classname* 的类，其属性在圆括号中给出。

OPEN (*name1 = type1 [,name2 = type2[,...]]*) AS *classname*

打开一个名为 *classname* 的类用于写入，但不把它的存在记录到系统目录中。（这主要是为了帮助引导。）

DESTROY *classname*

销毁名为 *classname* 的类。

DEFINE INDEX *indexname* ON *class\_name* USING *amname* (*opclass attr | (function (attr))*)

用 *amname* 访问方法在名为 *classname* 的类上创建一个名为 *indexname* 的索引。要索引的字段称为 *name1*、*name2* 等，而要使用的操作符集合分别为 *collection\_1*、*collection\_2* 等。

### 注意

这最后一句话没有引用示例中的任何东西。应该改成有意义的内容。 -  
1998-08-04

Thomas

## 62.3. 宏命令

DEFINE FUNCTION *macro\_name* AS *rettype function\_name(args)*

为一个名为 *macro\_name* 的函数定义函数原型，该函数的类型为 *rettype* 的值由执行带参数 *args* 的 *function\_name* 计算得出，参数以类似 C 的方式声明。

DEFINE MACRO *macro\_name* FROM FILE *filename*

定义一个名为 *macro\_name* 的宏，它的值从名为 *filename* 的文件中读出。

## 62.4. 调试命令

### 注意

关于调试命令的这一节在原始文档中是被注释掉的。Thomas 1998-08-05

r

随机打印打开的类。

m -1

切换时间信息的显示。

m 0

把检索设置为当前时间（now）。

m 1 Jan 1 01:00:00 1988

把检索设置为指定时间的快照。

m 2 Jan 1 01:00:00 1988, Feb 1 01:00:00 1988

把检索设置为指定时间范围的快照。如果想要无界的时间范围，任一时间都可以用空格代替。

&A *classname natts name1 type1 name2 type2 ...*

向类 *classname* 添加 *natts* 个名为 *name1*、*name2*、等的属性，其类型分别为 *type1*、*type2*、等。

&RR *oldclassname newclassname*

把 *oldclassname* 类重命名为 *newclassname*。

&RA *classname oldattname newattname classname oldattname newattname*

把名为 *classname* 的类中名为 *oldattname* 的属性重命名为 *newattname*。

## 62.5. 示例

下面的命令集将创建“pg\_opclass”类，它包含 *int\_ops* 集合，作为一个 OID 为 421 的对象，然后打印该类并关闭它。

```
create pg_opclass (opcname=name)
open pg_opclass
insert oid=421 (int_ops)
print
close pg_opclass
```

# 第 63 章 页文件

对数据库文件默认页格式的描述。

本节概述 Postgres 类使用的页格式。用户定义的访问方法不必使用这种页格式。

在下面的说明中，假定一个字节（byte）包含 8 位。此外，术语项（item）指存储在 Postgres 类中的数据。

## 63.1. 页结构

下表展示了普通 Postgres 类和 Postgres 索引类（例如 B-tree 索引）中的页面是如何组织的。

表 63.1. 页面布局示例

| 项                    | 描述 |
|----------------------|----|
| itemPointerData      |    |
| filler               |    |
| itemData...          |    |
| 未分配空间                |    |
| ItemContinuationData |    |
| 特殊空间                 |    |
| "ItemData 2"         |    |
| "ItemData 1"         |    |
| ItemIdData           |    |
| PageHeaderData       |    |

每个页面的前 8 个字节由页头（PageHeaderData）组成。在页头中，前三个 2 字节整数字段（lower、upper 和 special）分别表示到未分配空间开始处、到未分配空间结束处以及到特殊空间开始处的字节偏移。特殊空间是页面末尾的一个区域，它在页面初始化时分配，并包含某种访问方法特有的信息。页头的最后 2 个字节 opaque 编码页面大小和页面内部碎片的信息。每个页面都存储页面大小，因为缓冲池中的帧在一个类内可以逐帧细分为大小相等的页面。内部碎片信息用来帮助确定何时应进行页面重组。

页头之后是项标识符（ItemIdData）。新的项标识符从未分配空间的前四个字节分配。由于项标识符在被释放之前从不移动，其索引可用来指示一个项在页面上的位置。事实上，Postgres 创建的每个指向项的指针（ItemPointer）都由一个帧号和一个项标识符的索引组成。项标识符包含到项开始处的字节偏移、以字节为单位的长度以及一组影响其解释的属性位。

项本身存储在从未分配空间末尾向后分配的空间中。通常不对项进行解释。但当项太长而无法放在单个页面上或希望对项进行分片时，项会被切分，每一片按下述方式作为不同的项处理。第一片到倒数第二片放在项延续结构（ItemContinuationData）中。该结构包含指向下一片的 itemPointerData 和这一片本身。最后一片按正常方式处理。

## 63.2. 文件

data/

共享（全局）数据库文件的位置。

`data/base/`

本地数据库文件的位置。

## 63.3. 缺陷

页格式将来可能改变，以提供对大对象更高效的访问。

本节包含的细节不足以对编写新的访问方法提供任何帮助。

---

## 部分 VII. 附录

其他相关信息。

---

# 目录

|                              |     |
|------------------------------|-----|
| A. 文档 .....                  | 563 |
| A.1. 文档路线图 .....             | 563 |
| A.2. 文档源码 .....              | 563 |
| A.2.1. 文档结构 .....            | 564 |
| A.3. 文档项目 .....              | 564 |
| A.3.1. 样式与约定 .....           | 565 |
| A.3.2. 编写工具 .....            | 565 |
| A.4. 构建文档 .....              | 565 |
| A.5. v6.4 的硬拷贝生成 .....       | 566 |
| A.5.1. RTF 清理过程 .....        | 566 |
| A.6. 工具集 .....               | 567 |
| A.6.1. Linux 上的 RPM 安装 ..... | 567 |
| A.6.2. 手动安装工具 .....          | 567 |
| A.7. 备用工具集 .....             | 571 |
| B. 联系方式 .....                | 572 |
| SQL 参考文献 .....               | 573 |

---

# 附录 A. 文档

Thomas Lockhart  
Tom Ivar Helbekkmo

文档的目的是让 Postgres 更容易学习、使用和开发。文档集应当描述 Postgres 的系统、语言和接口。它应当能够回答常见问题，并让用户能够自行找到这些答案，而不必求助于邮件列表支持。

## A.1. 文档路线图

Postgres 有四种主要的文档格式：

- 纯文本，用于安装前信息。
- HTML，用于在线浏览和参考。
- 硬拷贝，用于深入阅读和参考。
- 手册页，用于快速参考。

**表 A.1. Postgres 文档产品**

| 文件                      | 描述                           |  |
|-------------------------|------------------------------|--|
| ./COPYRIGHT             | 版权声明                         |  |
| ./INSTALL               | 安装说明（文本由 sgml->rtf->text 生成） |  |
| ./README                | 入门信息                         |  |
| ./register.txt          | make 期间的注册消息                 |  |
| ./doc/bug.template      | 缺陷报告模板                       |  |
| ./doc/postgres.tar.gz   | 集成文档（HTML）                   |  |
| ./doc/programmer.ps.gz  | 程序员指南（Postscript）            |  |
| ./doc/programmer.tar.gz | 程序员指南（HTML）                  |  |
| ./doc/reference.ps.gz   | 参考手册（Postscript）             |  |
| ./doc/reference.tar.gz  | 参考手册（HTML）                   |  |
| ./doc/tutorial.ps.gz    | 入门介绍（Postscript）             |  |
| ./doc/tutorial.tar.gz   | 入门介绍（HTML）                   |  |
| ./doc/user.ps.gz        | 用户指南（Postscript）             |  |
| ./doc/user.tar.gz       | 用户指南（HTML）                   |  |

有可供安装的手册页，Postgres 源代码树各处还有大量纯文本 README 类型的文件。

## A.2. 文档源码

文档源码包括纯文本文件、手册页和 html。不过，Postgres 的新文档大多将使用标准通用标记语言（SGML）DocBook<sup>1</sup> 文档类型定义（DTD）编写。现有文档的大部分已经或将被转换为 SGML。

---

<sup>1</sup><http://www.ora.com/davenport/>

SGML 的目的是让作者能够指定文档的结构和内容（例如使用 DocBook DTD），并由文档样式定义这些内容如何被渲染成最终形式（例如使用 Norm Walsh 的样式表）。

文档积累自多个来源。随着我们把现有文档整合成一个连贯的文档集，较旧的版本将逐渐过时，并会从发行包中删除。但是，这不会立即发生，也不会同时发生在所有文档上。为了简化过渡，并帮助引导开发者和作者，我们定义了一份过渡路线图。

下面是 v6.5 的文档计划：

- 开始为用户指南和管理员指南编译索引信息。
- 为用户指南编写更多涵盖参考页之外内容的节。这将包括入门信息，以及对典型设计问题的处理方法建议。
- 把现有手册页中的信息合并到参考页和用户指南中。把手册页压缩成提示性信息，并带上指向主文档集的引用。
- 把新的 sgml 参考页转换为新的手册页，替换现有的手册页。
- 为了可移植性，把所有源图形转换为 CGM 格式文件。目前我们大多只有 Applix Graphics 源文件，可以从生成 .gif 输出。有一幅图只有 .gif 和 .ps 形式，应当重画或删除。

## A.2.1. 文档结构

目前有五份用 DocBook 编写的独立文档。每份文档都有一个容器源文档，它定义 DocBook 环境和其他文档源文件。这些主源文件位于 `doc/src/sgml/`，文档使用的许多其他源文件也在哪里。主源文件有：

`postgres.sgml`

这是集成文档，把所有其他文档作为部分包含在内。它以 HTML 格式生成输出，因为浏览器界面让你只需点击就能在全部文档之间轻松跳转。其他文档同时提供 HTML 和硬拷贝两种格式。

`tutorial.sgml`

入门教程，带示例。不包括编程主题，旨在帮助不熟悉 SQL 的读者。这是“入门”文档。

`user.sgml`

用户指南。包括数据类型和用户级接口的信息。这是放置“为什么”类信息的地方。

`reference.sgml`

参考手册。包括 Postgres SQL 语法。这是放置“怎么做”类信息的地方。

`programming.sgml`

程序员指南。包括 Postgres 可扩展性以及编程接口的信息。

`admin.sgml`

管理员指南。包括安装说明和发行说明。

## A.3. 文档项目

打包的文档有 HTML 和 Postscript 两种格式。它们作为标准 Postgres 安装的一部分提供。我们在这里讨论如何使用文档源码以及如何生成文档包。

## 注意

这是三年来 Postgres 新文档的第一个发行版。内容和环境都还在变动和演化之中。

SGML 的目的是让作者能够指定文档的结构和内容（例如使用 DocBook DTD），并由文档样式定义这些内容如何被渲染成最终形式（例如使用 Norm Walsh 的样式表）。

参阅 Introduction to DocBook<sup>2</sup>，那里有一份不错的 DocBook 特性“快速入门”摘要。DocBook Elements<sup>3</sup> 为 DocBook 的特性提供了强大的交叉参考。

本文档集使用若干工具构建，包括 James Clark 的 jade<sup>4</sup> 和 Norm Walsh 的 Modular DocBook Stylesheets<sup>5</sup>。

目前，硬拷贝是通过把 jade 输出的富文本格式（RTF）导入 ApplixWare 做少量格式修正，然后导出为 Postscript 文件来生成的。

TeX<sup>6</sup> 是 jade 输出的一种受支持格式，但当时未被使用，原因有几点，包括在提交硬拷贝之前无法做小的格式修正，以及 TeX 样式表对表格的支持普遍不足。

### A.3.1. 样式与约定

DocBook 有一组丰富的标签和构造，其中直接而明显地适用于良构文档的百分比出人意料地高。Postgres 文档集最近才改写为 SGML，不久将来会从文档集中选出若干节，作为 DocBook 用法的示范性示例加以维护。此外，下面还会给出 DocBook 标签的简短摘要。

### A.3.2. 编写工具

当前的 Postgres 文档集是使用纯文本编辑器（或 emacs/psgml，见下文）编写的，内容用 SGML 标记。

SGML 和 DocBook 的开源编写工具并不多。最常见的工具集是带 psgml 功能扩展的 emacs/xemacs 编辑软件包。在某些系统（例如 RedHat Linux）上，典型的完整安装会包含这些工具。

#### A.3.2.1. emacs/psgml

使用 emacs/psgml 时，处理这些分别保存书籍各部分的文件，有一种方便的方式：在编辑时插入适当的 DOCTYPE 声明。例如，当前这个源码文件是一章附录，因此可以将它指定为 DocBook 文档的“appendix”实例，把第一行写成这样：

```
!doctype appendix PUBLIC "-//Davenport//DTD DocBook V3.0//EN"
```

这意味着任何读取 SGML 的东西都能正确处理它，而且我可以用“nsgmls -s docguide.sgml”验证该文档。

## A.4. 构建文档

GNU make 用于从 DocBook 源码构建文档。有一些环境定义可能需要为你的安装设置或修改。Makefile 会寻找 doc/./src/Makefile 和（隐式地）寻找 doc/./src/Makefile.custom 来获取环境信息。在我的系统上，src/Makefile.custom 看起来像

<sup>2</sup><http://nis-www.lanl.gov/~rosalia/mydocs/docbook-intro.html>

<sup>3</sup><http://www.ora.com/homepages/dtdparse/docbook/3.0/>

<sup>4</sup><http://www.jclark.com/jade/>

<sup>5</sup><http://www.berkshire.net/~norm/docbook/dsssl>

<sup>6</sup><http://sunsite.unc.edu/pub/packages/TeX/systems/unix/>

```
# Makefile.custom
# Thomas Lockhart 1998-03-01

POSTGRES DIR= /opt/postgres/current
CFLAGS+= -m486
YFLAGS+= -v

# documentation

HSTYLE= /home/tgl/SGML/db107.d/docbook/html
PSTYLE= /home/tgl/SGML/db107.d/docbook/print
```

其中 HSTYLE 和 PSTYLE 分别决定 HTML 和硬拷贝（打印）样式表的 docbook.dsl 路径。这些样式表文件名是针对 Norm Walsh 的 Modular Style Sheets 的；如果使用其他样式表，可以把 HDSL 和 PDSL 定义为样式表的完整路径和文件名，就像上面为 HSTYLE 和 PSTYLE 所做的那样。在许多系统上，这些样式表会出现在安装于 /usr/lib/sgml/、/usr/share/lib/sgml/ 或 /usr/local/lib/sgml/ 的软件包中。

HTML 文档包可以通过输入以下命令从 SGML 源码生成

```
% cd doc/src
% make tutorial.tar.gz
% make user.tar.gz
% make admin.tar.gz
% make programmer.tar.gz
% make postgres.tar.gz
% make install
```

这些包可以从主文档目录安装，输入

```
% cd doc
% make install
```

## A.5. v6.4 的硬拷贝生成

硬拷贝 Postscript 文档的生成方法是：先把 SGML 源码转换为 RTF，然后导入 ApplixWare-4.4.1。经过少量清理（见下一节）后，把输出“打印”到一个 postscript 文件。

有些插图经过重画，以避免硬拷贝文档中出现位图 GIF 文件。有一幅关于系统目录的图足够复杂，没有时间重画。它用以下命令转换以适应页面：

```
% convert -monochrome -v -geometry 500x500 '>' catalogs.ps catalogs.gif
% convert -v -crop 400x500 catalogs.gif catalogs-cropped.gif
```

### A.5.1. RTF 清理过程

生成 Postscript 硬拷贝时必须处理若干事项：

#### Applixware RTF 清理

Applixware 在导入由 jade/MSS 生成的 RTF 方面似乎做得不完整。特别是，所有文本都被赋予“Header1”样式属性标签，尽管文本格式本身是可以接受的。另外，目录页码引用的不是表中所列的节，而是指向 ToC 自身所在的页。

1. 输入以下命令生成 RTF 输入

```
% cd doc/src/sgml
% make tutorial.rtf
```

2. 在 Applix Words 中打开一个新文档，然后导入 RTF 文件。
3. 打印现有的目录，供下面几步标记使用。
4. 把插图插入文档。用居中边距按钮把每幅图在页面上居中。

并非所有文档都有插图。你可以在 SGML 源文件中 grep 字符串 “Graphic”，来找出文档中可能带有插图的部分。有少数插图在文档的不同部分重复出现。

5. 通读整个文档，调整分页和表格列宽。
6. 如果存在书目，Applix Words 似乎会把第一个标题之后的所有其余文本都标上下划线属性。选中所有其余文本，用下划线按钮关闭下划线，然后显式地为每个文档名和书名加下划线。
7. 通读整个文档，在 ToC 硬拷贝上标出每个 ToC 条目的实际页码。
8. 用正确的值替换 ToC 中右对齐的错误页码。这只需要每份文档几分钟。
9. 把文档保存为 Applix Words 原生格式，以便日后更容易地进行最后时刻的编辑。
10. 把文档导出为 Postscript 格式的文件。
11. 用 gzip 压缩 Postscript 文件。把压缩后的文件放入 doc 目录。

## A.6. 工具集

我们记录了处理文档所需各种工具的两中安装方法的经验。第一种是在 Linux 上从 RPM 安装，另一种是从各工具的原始发行包进行通用安装。下面将分别介绍。

我们了解到这些工具可能还有其他打包发行形式。FreeBSD 似乎提供了它们。请把软件包状态报告到 docs 邮件列表，我们会把相关信息收录在这里。

### A.6.1. Linux 上的 RPM 安装

安装 Jade 及相关软件包的 RPM<sup>7</sup>。

### A.6.2. 手动安装工具

这里简要过一遍获取和安装软件的过程，你需要这些软件来用 Emacs 编辑 DocBook 源码，并用 Norman Walsh 的 DSSSL 样式表处理它，创建 HTML 和 RTF。

这些说明不包括 sgml-tools 软件包中新的 jade/DocBook 支持。作者们自该软件包采用 DocBook 以来还没有试过它，但它几乎肯定是一个好的候选。

#### A.6.2.1. 先决条件

你需要：

- 一个可用的 GCC 2.7.2 安装

---

<sup>7</sup><ftp://ftp.cygnum.com/pub/home/rosalia/>

- 一个可用的 Emacs 19.19 或更高版本安装
- 一个 UNIX 的 unzip 程序用来解包

你必须获取：

- James Clark 的 Jade 1.1 版<sup>8</sup>
- DocBook 3.0 版<sup>9</sup>
- Norman Walsh 的 Modular Stylesheets 1.19 版<sup>10</sup>
- Lennart Staflin 的 PSGML 1.0.1 版<sup>11</sup>

重要 URL：

- Jade 网页<sup>12</sup>
- DocBook 网页<sup>13</sup>
- Modular Stylesheets 网页<sup>14</sup>
- PSGML 网页<sup>15</sup>
- Steve Pepper 的 Whirlwind Guide<sup>16</sup>
- Robin Cover 的 SGML 软件数据库<sup>17</sup>

## A.6.2.2. 安装 Jade

### 安装 Jade

1. 阅读上面列出的 URL 处的安装说明。
2. 把发行包解压到一个合适的位置。使用的命令大致是

```
unzip -aU jade1_1.zip
```

3. Jade 不是用 GNU Autoconf 构建的，所以你需要自己编辑 Makefile。既然 James Clark 已经为此准备好了他的套件，一个好办法是在 Jade 发行包的主目录下建一个构建目录（也许以你的机器架构命名），把主目录中的 Makefile 文件复制进去，在那里编辑它，然后在那里运行 make。

不过，Makefile 确实需要编辑。主目录中有一个名为 Makefile.jade 的文件，构建 Jade（而不是构建 Jade 所基于的 SGML 解析器套件 SP）时，可以用 `make -f Makefile.jade` 来使用它。不过我们建议你不要那样做，因为需要修改的内容超出了 Makefile.jade 的范围，反正你也得编辑其中一个。

<sup>8</sup>[http://ftp.jclark.com/pub/jade/jade1\\_1.zip](http://ftp.jclark.com/pub/jade/jade1_1.zip)

<sup>9</sup><http://www.ora.com/davenport/docbook/current/docbk30.zip>

<sup>10</sup><http://nwalsh.com/docbook/dsssl/db119.zip>

<sup>11</sup><http://ftp.lysator.liu.se/pub/sgml/psgml-1.0.1.tar.gz>

<sup>12</sup><http://www.jclark.com/jade/>

<sup>13</sup><http://www.ora.com/davenport/>

<sup>14</sup><http://nwalsh.com/docbook/dsssl/>

<sup>15</sup>[http://www.lysator.liu.se/projects/about\\_psgml.html](http://www.lysator.liu.se/projects/about_psgml.html)

<sup>16</sup><http://www.infotek.no/sgmltool/guide.htm>

<sup>17</sup><http://www.sil.org/sgml/publicSW.html>

通读 Makefile，阅读 James 的说明并按需编辑。有若干变量需要设置。下面收集汇总了最重要的一些变量及其典型值：

```
prefix = /usr/local
XDEFINES = -DSGML_CATALOG_FILES_DEFAULT=\"/usr/local/share/sgml/
catalog\"
XLIBS = -lm
RANLIB = ranlib
srcdir = ..
XLIBDIRS = grove spgrove style
XPROGDIRS = jade
```

注意其中指定了查找 SGML 支持文件默认 catalog 的位置 -- 你可能想把它改成更适合你自己安装的值。如果你的系统不需要上面这些数学库和 ranlib 命令的设置，就让它们保持 Makefile 中的原样。

4. 输入 make 构建 Jade 和各种 SP 工具。
5. 软件构建完成后，make install 会完成显而易见的那一步。

### A.6.2.3. 安装 DocBook DTD 工具包

#### 安装 DocBook DTD 工具包

1. 你会想把构成 DocBook DTD 工具包的文件放在你构建 Jade 时让它查找的目录中；如果你遵循了上面的建议，那就是 /usr/local/share/sgml/。除了 DocBook 实际文件外，你还需要放好一个 catalog 文件，用于把文档类型规范和外部实体引用映射到该目录中的实际文件。你还需要 ISO 字符集映射，可能还有一或多个版本的 HTML。

安装各种 DTD 和支持文件并建立 catalog 文件的一种办法是，把它们全部收进上面提到的目录，用一个名为 CATALOG 的文件描述它们全部，然后创建文件 catalog 作为指向前者的 catalog 指针，给它这样一行内容：

```
CATALOG /usr/local/share/sgml/CATALOG
```

2. CATALOG 文件随后应包含三类行。第一类是（可选的）SGML 声明，即：

```
SGMLDECL docbook.dcl
```

接下来，对 DTD 和实体文件的各项引用必须解析。对于 DocBook 文件，这些行看起来像这样：

```
PUBLIC "-//Davenport//DTD DocBook V3.0//EN" docbook.dtd
PUBLIC "-//USA-DOD//DTD Table Model 951010//EN" cals-tbl.dtd
PUBLIC "-//Davenport//ELEMENTS DocBook Information Pool V3.0//EN"
dbpool.mod
PUBLIC "-//Davenport//ELEMENTS DocBook Document Hierarchy V3.0//
EN" dbhier.mod
PUBLIC "-//Davenport//ENTITIES DocBook Additional General Entities
V3.0//EN" dbgenent.mod
```

3. 当然，包含这些内容的文件随 DocBook 工具包一起提供。注意这些行每行的最后一项都是文件名，这里没有带路径。你当然可以把文件放在主 SGML 目录的子目录中，并修改 CATALOG

文件中的引用。DocBook 还引用 ISO 字符集实体，所以你需要获取并安装它们（它们可从多个来源获得，通过上面列出的 URL 很容易找到），并为它们全部配上 catalog 条目，例如：

```
PUBLIC "ISO 8879-1986//ENTITIES Added Latin 1//EN" ISO/ISolat1
```

注意这里的文件名包含目录名，表明我们把 ISO 实体文件放在了名为 ISO 的子目录中。同样，你获取的实体包应附带正确的 catalog 条目。

#### A.6.2.4. 安装 Norman Walsh 的 DSSSL 样式表

##### 安装 Norman Walsh 的 DSSSL 样式表

1. 阅读上面列出的 URL 处的安装说明。
2. 要安装 Norman 的样式表，只需把发行包解压到合适的位置。一个合适的位置是 /usr/local/share，这会把工具包放到 /usr/local/share/docbook 目录树下。命令大致是

```
unzip -aU db107.zip
```

3. 测试安装的一种办法是构建 PostgreSQL 手册的 HTML 和 RTF 形式。

- a. 要构建 HTML 文件，进入 SGML 源目录 doc/src/sgml，执行

```
jade -t sgml -d /usr/local/share/docbook/html/docbook.dsl -D ..  
/graphics postgres.sgml
```

book1.htm 是输出的顶层节点。。

- b. 要生成可导入你喜欢的字处理系统并打印的 RTF 输出，输入：

```
jade -t rtf -d /usr/local/share/docbook/print/docbook.dsl -D ..  
/graphics postgres.sgml
```

#### A.6.2.5. 安装 PSGML

##### 安装 PSGML

1. 阅读上面列出的 URL 处的安装说明。
2. 解开发行文件，运行 configure、make 和 make install，把字节编译文件和 info 库放到位置。
3. 然后把以下几行加入你的 /usr/local/share/emacs/site-lisp/site-start.el 文件，让 Emacs 在需要时正确载入 PSGML：

```
(setq load-path  
      (cons "/usr/local/share/emacs/site-lisp/psgml" load-path))  
(autoload 'sgml-mode "psgml" "Major mode to edit SGML files." t)
```

4. (可选的) 如果你想在编辑 HTML 时也使用 PSGML，还要加上：

```
(setq auto-mode-alist  
      (cons '("\\.s?html?$\\'" . sgml-mode) auto-mode-alist))
```

5. (可选的) 使用 PSGML 有一点需要注意: 它的作者假定你的主 SGML DTD 目录是 `/usr/local/lib/sgml`。如果像本章示例中那样使用 `/usr/local/share/sgml`, 你就必须对此作出补偿。
  - a. (可选的) 你可以设置 `SGML_CATALOG_FILES` 环境变量。
  - b. (可选的) 你可以定制你的 PSGML 安装 (它的手册说明了方法)。
  - c. (可选的) 你甚至可以在编译和安装 PSGML 之前编辑源文件 `psgml.el`, 把硬编码的路径改成你自己的默认值。

### A.6.2.6. 安装 JadeTeX

如果你愿意, 还可以安装 JadeTeX, 用 TeX 作为 Jade 的格式化后端。注意它仍是相当粗糙的软件, 生成的打印输出不如 RTF 后端的质量。不过它仍然工作得不错, 尤其是对不使用表格的较简单文档而言; 而且由于 JadeTeX 和样式表都在持续改进, 它将来肯定会越来越好。

要安装和使用 JadeTeX, 你需要可用的 TeX 和 LaTeX2e 安装, 包括受支持的 `tools` 和 `graphics` 宏包、Babel、AMS 字体和 AMS-LaTeX、PSNFSS 扩展及 “35 种字体” 配套包、用于生成 PostScript 的 `dvips` 程序、宏包 `fancyhdr`、`hyperref`、`minitoc`、`url` 和 `ot2enc`, 当然还有 JadeTeX 本身。所有这些都可以在邻近的 CTAN 站点找到。

JadeTeX 在撰写本文时并没有附带多少安装指南, 但有一个 `makefile` 展示了需要什么。它还包含一个 `cooked` 目录, 你会在其中找到它所需的部分宏包 -- 但不是全部, 也不完整, 至少我们上次查看时是这样。

在构建 `jadetex.fmt` 格式文件之前, 你可能需要编辑 `jadetex.ltx` 文件, 把 Babel 的配置改为适合你所在地区的形式。要改的那一行看起来像

```
\RequirePackage[german,french,english]{babel}[1997/01/23]
```

显然, 你应该只列出你确实需要、并且已为 Babel 配置好的语言。

JadeTeX 可用后, 你就应该能为 PostgreSQL 手册生成并排版 TeX 输出, 办法是 (如上所述, 在 `doc/src/sgml` 目录中) 给出命令

```
jade -t tex -d /usr/local/share/docbook/print/docbook.dsl -D ../
graphics postgres.sgml
jadetex postgres.tex
jadetex postgres.tex
dvips postgres.dvi
```

当然, 这样做时, TeX 会在第二轮运行中停下来, 并告诉你它的容量已被超出。就我们所知, 这是 JadeTeX 生成交叉引用信息的方式造成的。当然, TeX 可以用更大的数据结构尺寸编译。具体细节因安装而异。

## A.7. 备用工具集

`sgml-tools v2.x` 现在支持 `jade` 和 `DocBook`。它可能成为使用 SGML 工作的首选工具集, 但我们还没有机会评估这个新软件包。

---

## 附录 B. 联系方式

- Thomas Lockhart<sup>1</sup> 负责 SQL 标准兼容性和文档。

---

<sup>1</sup><http://alumni.caltech.edu/~lockhart>

---

# SQL 参考文献

这里列出了一些关于 SQL 和 Postgres 的参考文献与读物。

## SQL 参考书

SQL 特性的参考文本。

The Practical SQL Handbook. Bowman et al, 1993. Using Structured Query Language. 3. Judith Bowman、Sandra Emerson和Marcy Damovsky. ISBN 0-201-44787-8. 1996. Addison-Wesley. 版权 © 1997 Addison-Wesley Longman, Inc..

A Guide to the SQL Standard. Date and Darwen, 1997. A user's guide to the standard database language SQL. 4. C. J. Date和Hugh Darwen. ISBN 0-201-96426-0. 1997. Addison-Wesley. 版权 © 1997 Addison-Wesley Longman, Inc..

Understanding the New SQL. Melton and Simon, 1993. A complete guide. Jim Melton和Alan R. Simon. ISBN 1-55860-245-3. 1993. Morgan Kaufmann. 版权 © 1993 Morgan Kaufmann Publishers, Inc..

### Abstract

SQL 特性的易读参考。

## PostgreSQL 专属文档

本节收录相关文档。

The PostgreSQL. Administrator's Guide. The Administrator's Guide. Thomas Lockhart. 1998-10-01. The PostgreSQL Global Development Group.

The PostgreSQL. Developer's Guide. The Developer's Guide. Thomas Lockhart. 1998-10-01. The PostgreSQL Global Development Group.

The PostgreSQL. Programmer's Guide. The Programmer's Guide. Thomas Lockhart. 1998-10-01. The PostgreSQL Global Development Group.

The PostgreSQL. Tutorial Introduction. The Tutorial. Thomas Lockhart. 1998-10-01. The PostgreSQL Global Development Group.

The PostgreSQL. User's Guide. The User's Guide. Thomas Lockhart. 1998-10-01. The PostgreSQL Global Development Group.

The Postgres95. User Manual. Yu and Chen, 1995. A. Yu和J. Chen. . Sept. 5, 1995. University of California, Berkeley CA.

## 会议论文和期刊文章

本节收录文章和简报。

Partial indexing in POSTGRES : research project. Olson, 1993. Nels Olson. 1993. UCB Engin T7.49. 1993 O676. University of California, Berkeley CA.

- A Unified Framework for Version Modeling Using Production Rules in a Database System. Ong and Goh, 1990. L. Ong和J. Goh. April, 1990. ERL Technical Memorandum M90/33. University of California, Berkeley CA.
- The Postgres. Data Model. Rowe and Stonebraker, 1987. L. Rowe和M. Stonebraker. VLDB Conference, Brighton, England, Sept. 1987.
- Generalized partial indexes<sup>1</sup>. . P. Seshadri和A. Swami. Eleventh International Conference on Data Engineering, March 1995. 1995. Cat. No.95CH35724. IEEE Computer Society Press.
- The Design of Postgres. . Stonebraker and Rowe, 1986. M. Stonebraker和L. Rowe. Conference on Management of Data, Washington DC, May 1986.
- The Design of the Postgres. Rules System. Stonebraker, Hanson, Hong, 1987. M. Stonebraker、E. Hanson和C. H. Hong. Conference on Data Engineering, Los Angeles, CA, Feb. 1987.
- The Postgres. Storage System. Stonebraker, 1987. M. Stonebraker. VLDB Conference, Brighton, England, Sept. 1987.
- A Commentary on the Postgres. Rules System. Stonebraker et al, 1989. M. Stonebraker、M. Hearst和S. Potamianos. Record 18(3), Sept. 1989.
- The case for partial indexes (DBMS)<sup>2</sup>. Stonebraker, M, 1989b. M. Stonebraker. Record 18(no.4):4-11, Dec. 1989.
- The Implementation of Postgres. . Stonebraker, Rowe, Hirohama, 1990. M. Stonebraker、L. A. Rowe和M. Hirohama. Transactions on Knowledge and Data Engineering 2(1), March 1990.
- On Rules, Procedures, Caching and Views in Database Systems. Stonebraker et al, ACM, 1990. M. Stonebraker和et al. Conference on Management of Data, June 1990.

---

<sup>1</sup><http://simon.cs.cornell.edu/home/praveen/papers/partindex.de95.ps.Z>

<sup>2</sup><http://s2k-ftp.CS.Berkeley.EDU:8000/postgres/papers/ERL-M89-17.pdf>