

# PostgreSQL

## 修炼之道

从小工到专家

唐成 著

---

Guide To Expert, The Pragmatic PostgreSQL

---

- PostgreSQL中国社区资深数据库专家、沃趣科技首席数据库架构师撰写，PostgreSQL数据库领域经典著作
- 系统讲解PostgreSQL技术内幕，深入分析PostgreSQL特色功能，包含大量来自实际生产环境的经典案例和经验总结



# PostgreSQL

## 修炼之道

从小工到专家

---

Guide To Expert, The Pragmatic PostgreSQL

---

唐成 著



机械工业出版社  
China Machine Press

## 图书在版编目 (CIP) 数据

PostgreSQL 修炼之道：从小工到专家 / 唐成著. —北京：机械工业出版社，2015.4  
(数据库技术丛书)

ISBN 978-7-111-49872-8

I. P… II. 唐… III. 关系数据库系统 IV. TP311.132.3

中国版本图书馆 CIP 数据核字 (2015) 第 063966 号

---

# PostgreSQL 修炼之道：从小工到专家

---

出版发行：机械工业出版社（北京市西城区百万庄大街 22 号 邮政编码：100037）

责任编辑：杨绣国 陈佳媛

责任校对：董纪丽

印刷：北京瑞德印刷有限公司

版次：2015 年 4 月第 1 版第 1 次印刷

开本：186mm×240mm 1/16

印张：33.5

书号：ISBN 978-7-111-49872-8

定价：79.00 元

凡购本书，如有缺页、倒页、脱页，由本社发行部调换

客服热线：(010) 88378991 88361066

投稿热线：(010) 88379604

购书热线：(010) 68326294 88379649 68995259

读者信箱：hzsj@hzbook.com

版权所有·侵权必究

封底无防伪标均为盗版

本书法律顾问：北京大成律师事务所 韩光 / 邹晓东

## 为什么要写这本书

PostgreSQL 数据库是目前功能最强大的开源数据库，它基本包含了其他所有商业或开源的数据库中能找到的功能，甚至还包含了一些商业数据库中没有的功能。它是最接近工业标准 SQL92 的查询语言，并且正在实现新的功能以兼容最新的 SQL 标准：SQL2003。PostgreSQL 也获得数个奖项，曾三次被评为 Linux Journal 杂志编辑评选的“最佳数据库奖”（2000 年、2003 年和 2004 年），并获 2004 年度的 Linux 新媒体最佳数据库系统奖。

PostgreSQL 目前在国外很流行，特别是近两年，使用 PostgreSQL 数据库的公司越来越多，如提供网络电话功能的 skype 和著名的图片分享网站 Instagram。2012 年，美国联邦机构全面转向 PostgreSQL 阵营；法国也正推动政府机构积极采用 PostgreSQL 数据库取代商业数据库；世界最大的 CRM 软件服务提供商 Salesforce 同样开始大量使用 PostgreSQL。在 DB-Engine 发布的历年数据库排名中，PostgreSQL 自从 2013 年 3 月上升到第四名后，一直稳定在第四名，排在很多知名的商业数据库如 DB2、Sybase 之前，也排在所有 NoSQL 数据库如 Cassandra、Redis 等之前。

虽然在国外使用 PostgreSQL 数据库的人很多，但在国内，PostgreSQL 中文的学习资料并不多，因此我就想到写一本关于 PostgreSQL 的书，让国内更多的人加入到学习 PostgreSQL 数据库的队伍中来。

## 读者对象

适合阅读本书的用户：

- 数据库入门者。学习本书和相关的数据库知识，可以让一个对数据库了解不是很深的数据库爱好者成为数据库专家。
- 非 PostgreSQL 数据库的 DBA。可以让非 PostgreSQL 的 DBA 快速掌握 PostgreSQL 数据库相关知识，成为一名合格的 PostgreSQL DBA。
- PostgreSQL DBA。本书的一些章节对熟悉 PostgreSQL 数据库的 DBA 也有很大指导作

用，可以使读者的知识更充实。

- 开发人员。通过此书可以快速掌握 PostgreSQL 数据库方面的知识，提高开发人员的数据库水平。

## 如何阅读本书

本书分为四大部分，分别为准备篇、基础篇、提高篇和第三方开源软件及架构篇。准备篇是为没有数据库基础的读者准备的，如果你已经具备了一定的数据库基础，可以跳过其中的一些内容。基础篇介绍了 PostgreSQL 数据库中的一些基础内容，学完此篇可以完成基本的 PostgreSQL 数据库的日常操作。提高篇讲解了一些更深的内容，如 PostgreSQL 的一些技术内幕、特色功能、优化等方面的内容，仔细阅读此篇可使你早日成为 PostgreSQL 数据库高手。第三方开源软件及架构篇讲解了与 PostgreSQL 数据库配套使用的一些常用的开源软件及架构设计方面的内容，通过阅读此篇，可以开阔大家的眼界，提高数据库架构设计能力。

本书中有大量的例子，读者边阅读此书边按例子进行实际的操作，将获得最佳的学习效果。

## 勘误和支持

由于作者的水平有限，编写的时间也很仓促，书中难免会出现一些错误或者不准确的地方，不妥之处恳请读者批评指正。你可以将书中的错误，遇到的问题及宝贵意见发送邮件至我的邮箱 [chengdata@gmail.com](mailto:chengdata@gmail.com)，我很期待听到你们的真挚反馈。

## 致谢

首先要感谢国内 PostgreSQL 数据库的爱好者，他们已经整理了很多 PostgreSQL 的文章，翻译了 PostgreSQL 的官方手册，让我可以站在前人的肩膀上。大家可以在 PostgreSQL 在中国的维基主页 <http://www.pgsql.org/mwiki/index.php/> 上看到前人的成果。

感谢机械工业出版社华章公司的编辑杨绣国老师，感谢她一年多来始终支持我的写作，她的鼓励和帮助引导我能顺利完成全部书稿。

最后要感谢我的妻子，她一直支持和鼓励我，让我能坚持把这本书写完。

谨以此书，献给众多热爱 PostgreSQL 的朋友们。

唐成 (osdba)

中国，杭州，2015 年 1 月

前言

## 第一篇 准备篇

### 第 1 章 PostgreSQL 简介..... 2

- 1.1 什么是 PostgreSQL ..... 2
  - 1.1.1 PostgreSQL 概述..... 2
  - 1.1.2 PostgreSQL 的发展历史..... 2
  - 1.1.3 PostgreSQL 数据库的优势..... 3
  - 1.1.4 PostgreSQL 应用现状和发展趋势..... 4
- 1.2 PostgreSQL 数据库与其他数据库的对比 ..... 4
  - 1.2.1 PostgreSQL 与 MySQL 数据库的对比 ..... 4
  - 1.2.2 PostgreSQL 与 Oracle 数据库的对比 ..... 6
- 1.3 小结..... 6

### 第 2 章 PostgreSQL 安装与配置..... 7

- 2.1 从发行版本安装 ..... 7
  - 2.1.1 在 Debian 或 Ubuntu 下的安装..... 7
  - 2.1.2 在 Redhat、CentOS 或 Fedora 下的安装 ..... 9
  - 2.1.3 在 Windows 下的安装 ..... 12
  - 2.1.4 发行版安装总结 ..... 16
- 2.2 从源码安装 ..... 16

- 2.2.1 编译安装过程介绍 ..... 16
- 2.2.2 下载源代码 ..... 17
- 2.2.3 编译及安装 ..... 18
- 2.2.4 安装后的配置 ..... 20
- 2.2.5 创建数据库簇 ..... 21
- 2.2.6 安装 contrib 目录下的工具 ..... 21
- 2.2.7 启动和停止数据库 ..... 21
- 2.2.8 编译安装时的常见问题及解决方法 ..... 22
- 2.3 安装技巧介绍 ..... 24
  - 2.3.1 在 Redhat、CentOS 下使用二进制包安装较新版本的方法..... 24
  - 2.3.2 如何使用较大的数据块提高 I/O 性能 ..... 25
- 2.4 PostgreSQL 的简单配置 ..... 25
  - 2.4.1 修改监听的 IP 和端口 ..... 25
  - 2.4.2 与数据库 log 相关的参数 ..... 25
  - 2.4.3 内存参数的设置 ..... 26
- 2.5 小结..... 26

### 第 3 章 SQL 语言入门 ..... 27

- 3.1 SQL 语句语法简介 ..... 27
  - 3.1.1 语句的分类 ..... 27
  - 3.1.2 词法结构 ..... 27
- 3.2 DDL 语句 ..... 28
  - 3.2.1 建表语句 ..... 28
  - 3.2.2 删除表语句 ..... 30
- 3.3 DML 语句..... 30

3.3.1	插入语句	30	5.1.1	类型的分类	54
3.3.2	更新语句	31	5.1.2	类型输入与转换	55
3.3.3	删除语句	31	5.2	布尔类型	56
3.4	查询语句	31	5.2.1	布尔类型解释	56
3.4.1	单表查询语句	31	5.2.2	布尔类型的操作符	58
3.4.2	过滤条件的查询	32	5.3	数值类型	59
3.4.3	排序	32	5.3.1	数值类型解释	59
3.4.4	分组查询	33	5.3.2	整数类型	59
3.4.5	表 join	34	5.3.3	精确的小数类型	59
3.5	其他 SQL 语句	36	5.3.4	浮点数类型	60
3.5.1	INSERT INTO... SELECT 语句	36	5.3.5	序列类型	61
3.5.2	UNION 语句	36	5.3.6	货币类型	61
3.5.3	TRUNCATE TABLE 语句	37	5.3.7	数学函数和操作符	62
3.6	小结	37	5.4	字符串类型	64
			5.4.1	类型解释	64
			5.4.2	字符串函数和操作符	65
			5.5	二进制数据类型	67
			5.5.1	二进制数据类型解释	67
			5.5.2	二进制数据类型转义表示	67
			5.5.3	二进制数据类型的函数	68
			5.6	位串类型	69
			5.6.1	位串类型解释	69
			5.6.2	位串类型的使用	69
			5.6.3	位串的操作符及函数	70
			5.7	日期 / 时间类型	71
			5.7.1	日期 / 时间类型详解	71
			5.7.2	日期输入	72
			5.7.3	时间输入	73
			5.7.4	特殊值	75
			5.7.5	函数和操作符列表	75
			5.7.6	时间函数	77
			5.7.7	extract 和 date_part 函数	80
			5.8	枚举类型	81
			5.8.1	枚举类型的使用	81
			5.8.2	枚举类型的说明	82
			5.8.3	枚举类型的函数	83
			5.9	几何类型	84
			5.9.1	几何类型概况	84
			5.9.2	几何类型的输入	84
			5.9.3	几何类型的操作符	89
<b>第二篇 基础篇</b>					
<b>第 4 章</b>	<b>psql 工具的使用介绍</b>	40			
4.1	psql 介绍	40			
4.2	psql 的简单使用	40			
4.3	psql 的常用命令	42			
4.3.1	\d 命令	42			
4.3.2	指定字符集编译的命令	45			
4.3.3	\pset 命令	46			
4.3.4	\x 命令	46			
4.3.5	执行存储在外部文件中的 SQL 命令	47			
4.3.6	显示信息的命令	48			
4.3.7	更多的命令	49			
4.4	psql 的使用技巧和注意事项	50			
4.4.1	历史命令与补全的功能	50			
4.4.2	自动提交方面的技巧	50			
4.4.3	如何得到 psql 中命令实际执行的 SQL	51			
4.5	小结	53			
<b>第 5 章</b>	<b>数据类型</b>	54			
5.1	类型介绍	54			

5.9.4	几何类型的函数	97	5.17.2	pg_lsn 类型	143
5.10	网络地址类型	98	<b>第 6 章 逻辑结构管理</b> 145		
5.10.1	网络地址类型概况	98	6.1	数据库逻辑结构介绍	145
5.10.2	inet 与 cidr 类型	98	6.2	数据库基本操作	145
5.10.3	macaddr 类型	101	6.2.1	创建数据库	145
5.10.4	网络地址类型的操作符	101	6.2.2	修改数据库	146
5.10.5	网络地址类型的函数	102	6.2.3	删除数据库	147
5.11	复合类型	103	6.2.4	常见问题及解答	147
5.11.1	复合类型的定义	103	6.3	模式	148
5.11.2	复合类型的输入	104	6.3.1	模式的定义	148
5.11.3	访问复合类型	105	6.3.2	模式的使用	148
5.11.4	修改复合类型	105	6.3.3	公共模式	150
5.11.5	复合类型的输入与输出	106	6.3.4	模式的搜索路径	150
5.12	XML 类型	107	6.3.5	模式的权限	151
5.12.1	XML 类型的输入	107	6.3.6	模式的移植性	151
5.12.2	字符集的问题	108	6.4	表	152
5.12.3	XML 类型的函数	109	6.4.1	创建表	152
5.13	JSON 类型	114	6.4.2	表的存储属性	154
5.13.1	JSON 类型简介	115	6.4.3	临时表	156
5.13.2	JSON 类型的输入与输出	115	6.4.4	默认值	158
5.13.3	JSON 类型的操作符	116	6.4.5	约束	159
5.13.4	JSON 类型的函数	118	6.4.6	修改表	163
5.13.5	JSON 类型的索引	121	6.4.7	表继承及分区表	167
5.14	Range 类型	125	6.4.8	分区表	168
5.14.1	Range 类型简介	125	6.5	触发器	173
5.14.2	创建 Range 类型	126	6.5.1	创建触发器	173
5.14.3	Range 类型的输入与输出	127	6.5.2	语句级触发器与行级触发器	175
5.14.4	Range 类型的操作符	130	6.5.3	BEFORE 触发器与 AFTER 触发器	177
5.14.5	Range 类型的函数	130	6.5.4	删除触发器	178
5.14.6	Range 类型的索引和约束	131	6.5.5	触发器的行为	179
5.15	数组类型	132	6.5.6	触发器函数中的特殊变量	180
5.15.1	数组类型的声明	132	6.6	事件触发器	180
5.15.2	如何输入数组值	133	6.6.1	创建事件触发器	183
5.15.3	访问数组	135	6.6.2	修改事件触发器	186
5.15.4	修改数组	137	6.7	表空间	186
5.15.5	数组的操作符	138	6.7.1	表空间的定义	186
5.15.6	数组的函数	139	6.7.2	表空间的使用	186
5.16	伪类型	142	6.8	视图	187
5.17	其他类型	143			
5.17.1	UUID 类型	143			



6.8.1	视图的定义	187
6.8.2	创建视图	188
6.8.3	可更新视图	189
6.9	索引	191
6.9.1	索引简介	191
6.9.2	索引的分类	192
6.9.3	创建索引	192
6.9.4	并发创建索引	193
6.9.5	修改索引	196
6.9.6	删除索引	196
6.10	用户及权限管理	197
6.10.1	用户和角色	197
6.10.2	创建用户和角色	198
6.10.3	权限的管理	199
6.10.4	函数和触发器的权限	202
6.10.5	权限的总结	202
6.10.6	权限的示例	202
6.11	事务、并发、锁	203
6.11.1	ACID	203
6.11.2	DDL 事务	204
6.11.3	事务的使用	204
6.11.4	SAVEPOINT	205
6.11.5	事务隔离级别	206
6.11.6	两阶段提交	207
6.11.7	锁机制	209
6.11.8	死锁及防范	212
6.11.9	表级锁命令 LOCK TABLE	213
6.11.10	行级锁命令	213
6.11.11	锁的查看	214

## 第7章 PostgreSQL 的核心架构

7.1	应用程序的访问接口	221
7.1.1	访问接口总体图	221
7.1.2	不同编辑语言的 PostgreSQL 驱动介绍	222
7.2	进程及内存结构	223
7.2.1	进程和内存架构图	223
7.2.2	主进程 Postmaster	224
7.2.3	SysLogger (系统日志) 进程	224
7.2.4	BgWriter (后台写) 进程	225

7.2.5	WalWriter (预写式日志写) 进程	225
7.2.6	PgArch (归档) 进程	225
7.2.7	AutoVacuum (自动清理) 进程	225
7.2.8	PgStat (统计数据收集) 进程	226
7.2.9	共享内存	226
7.2.10	本地内存	226
7.3	目录结构	227
7.3.1	安装目录的结构	227
7.3.2	数据目录的结构	227
7.3.3	表空间的目录	228

## 第8章 服务管理

8.1	服务的启停和创建	229
8.1.1	启停方法	229
8.1.2	pg_ctl	230
8.1.3	信号	234
8.1.4	postgres 及单用户模式	234
8.2	服务配置介绍	235
8.2.1	配置参数	235
8.2.2	连接配置项	237
8.2.3	内存配置项	240
8.2.4	预写式日志的配置项	241
8.2.5	错误报告和日志项	243
8.3	访问控制配置文件	246
8.3.1	pg_hba.conf 文件	247
8.3.2	认证方法介绍	248
8.3.3	认证方法实战	249
8.4	备份和还原	249
8.4.1	逻辑备份	249
8.4.2	pg_dump 命令	250
8.4.3	pg_restore 命令	254
8.4.4	pg_dump 和 pg_restore 使用举例	257
8.4.5	物理备份	258
8.4.6	使用 LVM 快照进行热备份	259
8.5	常用的管理命令	261
8.5.1	查看系统信息的常用命令	261

8.5.2	系统维护常用命令	267
-------	----------	-----

### 第三篇 提高篇

<b>第 9 章</b>	<b>PostgreSQL 中执行计划</b>	270
9.1	执行计划的解释	270
9.1.1	EXPLAIN 命令	270
9.1.2	EXPLAIN 输出结果解释	271
9.1.3	EXPLAIN 使用示例	272
9.1.4	全表扫描	275
9.1.5	索引扫描	275
9.1.6	位图扫描	275
9.1.7	条件过滤	276
9.1.8	Nestloop Join	277
9.1.9	Hash Join	277
9.1.10	Merge Join	278
9.2	与执行计划相关的配置项	279
9.2.1	ENABLE_* 参数	279
9.2.2	COST 基准值参数	279
9.2.3	基因查询优化的参数	280
9.2.4	其他执行计划配置项	281
9.3	统计信息的收集	282
9.3.1	统计信息收集器的配置项	282
9.3.2	SQL 执行的统计信息输出	283
9.3.3	手工收集统计信息	283
<b>第 10 章</b>	<b>PostgreSQL 中的技术内幕</b>	285
10.1	表中的系统字段	285
10.1.1	oid	286
10.1.2	ctid	288
10.1.3	xmin、xmax、cmin、cmax	289
10.2	多版本并发控制	290
10.2.1	多版本并发控制的原理	290
10.2.2	PostgreSQL 中的多版本并发控制	291
10.2.3	PostgreSQL 多版本的优劣分析	293
10.3	物理存储结构	293
10.3.1	PostgreSQL 中的术语	293
10.3.2	数据块结构	293
10.3.3	Tuple 结构	294
10.3.4	数据块空闲空间管理	296
10.3.5	可见性映射表文件	298
10.4	技术解密	298
10.4.1	Index-only scans	298
10.4.2	Heap-Only Tuples	300
<b>第 11 章</b>	<b>PostgreSQL 的特色功能</b>	302
11.1	规则系统	302
11.1.1	SELECT 规则	302
11.1.2	更新规则	303
11.1.3	规则和权限	306
11.1.4	规则和命令状态	307
11.1.5	规则与触发器的比较	308
11.2	模式匹配和正则表达式	308
11.2.1	PostgreSQL 中的模式匹配和正则表达式介绍	308
11.2.2	传统 SQL 的 LIKE 操作符	309
11.2.3	SIMILAR TO 正则表达式	310
11.2.4	POSIX 正则表达式	312
11.2.5	模式匹配函数 substring	313
11.3	listen 与 notify	315
11.3.1	listen 与 notify 的简单示例	315
11.3.2	listen 与 notify 的相关命令	316
11.3.3	listen 与 notify 的使用详解	317
11.4	索引的特色	320
11.4.1	表达式上的索引	320
11.4.2	部分索引	320
11.4.3	GiST 索引	323
11.4.4	SP-GiST 索引	325
11.4.5	GIN 索引	326
11.5	序列的使用	328
11.5.1	序列的创建	328
11.5.2	序列的使用及相关的函数	329
11.5.3	常见问题及解答	331
11.6	咨询锁的使用	333
11.6.1	咨询锁的定义	333
11.6.2	咨询锁的函数及使用	333

11.6.3	常见问题及解答	337	13.1.3	流复制	372
11.7	SQL/MED	338	13.1.4	Standby 的运行原理	373
11.7.1	SQL/MED 的介绍	338	13.1.5	创建 Standby 的步骤	373
11.7.2	外部数据包装器对象	339	13.2	pg_basebackup 命令行工具	374
11.7.3	外部服务器对象	340	13.2.1	pg_basebackup 介绍	374
11.7.4	用户映射对象	341	13.2.2	pg_basebackup 的命令行参数	375
11.7.5	外部表对象	341	13.2.3	pg_basebackup 使用示例	376
11.7.6	file_fdw 使用实例	342	13.3	异步流复制 Hot Standby 的示例	377
11.7.7	postgres_fdw 使用实例	345	13.3.1	配置环境	377
<b>第 12 章</b>	<b>数据库优化</b>	<b>347</b>	13.3.2	主数据库的配置	378
12.1	优化准则和方法	347	13.3.3	在 Standby 上生成基础备份	378
12.1.1	优化准则	347	13.3.4	启动 Standby	379
12.1.2	优化方法	348	13.4	同步流复制的 Standby 数据库	380
12.2	硬件知识	348	13.4.1	同步流复制的架构	380
12.2.1	CPU 及服务器体系结构	348	13.4.2	同步复制的配置	381
12.2.2	内存	349	13.4.3	配置实例	381
12.2.3	硬盘	350	13.5	检查备库及流复制情况	383
12.3	文件系统及 I/O 调优	352	13.5.1	检查异步流复制的情况	383
12.3.1	文件系统的崩溃恢复	352	13.5.2	检查同步流复制的情况	384
12.3.2	Ext2 文件系统	353	13.5.3	视图 pg_stat_replication 详解	385
12.3.3	Ext3 文件系统	353	13.5.4	查看各库的状态	385
12.3.4	Ext4 文件系统	354	13.6	Hot Standby 的限制	387
12.3.5	XFS 文件系统	355	13.6.1	Hot Standby 的查询限制	387
12.3.6	Barriers I/O	355	13.6.2	Hot Standby 的查询冲突处理	389
12.3.7	I/O 调优的方法	356	13.7	恢复配置详解	390
12.4	性能监控	359	13.7.1	归档恢复配置的配置项	390
12.4.1	数据库性能视图	359	13.7.2	Recovery Target 配置	391
12.4.2	Linux 监控工具	362	13.7.3	Standby Server 配置	391
12.5	数据库配置优化	364	13.8	流复制的注意事项	392
12.5.1	内存配置优化	364	13.8.1	wal_keep_segments 参数的配置	392
12.5.2	关于双缓存的优化	366	13.8.2	vacuum_defer_cleanup_age 参数的配置	392
12.5.3	vacuum 中的优化	367			
12.5.4	预写式日志写优化	369			
<b>第 13 章</b>	<b>Standby 数据库的搭建</b>	<b>371</b>	<b>第四篇</b>	<b>第三方开源软件及架构篇</b>	
13.1	Standby 数据库原理	371	<b>第 14 章</b>	<b>PgBouncer</b>	<b>394</b>
13.1.1	PITR 原理	371	14.1	PgBouncer 介绍	394
13.1.2	WAL 日志归档	372			

14.2	PgBouncer 中的概念	395	16.2.3	安装 DBI 及 DBD::Pg	428
14.3	PgBouncer 的安装方法	395	16.2.4	安装 DBIx-Safe	429
14.4	PgBouncer 的简单使用	395	16.2.5	安装 Bucardo 源码包	429
14.4.1	简单配置方法	395	16.3	Bucardo 同步配置	429
14.4.2	启动 PgBouncer	396	16.3.1	示例环境	429
14.4.3	停止 PgBouncer	397	16.3.2	配置同步的简要过程	430
14.4.4	查看连接池信息	397	16.3.3	bucardo_ctl install	431
14.5	PgBouncer 的配置文件详解	399	16.3.4	bucardo_ctl add db	433
14.5.1	“[databases]”部分的配置项	399	16.3.5	bucardo_ctl add table	433
14.5.2	“[pgbouncer]”部分的配置项	399	16.3.6	bucardo_ctl add herd	433
14.5.3	用户密码文件	403	16.3.7	bucardo_ctl add sync	433
			16.3.8	bucardo_ctl start	434
<b>第 15 章</b>	<b>Slony-I 的使用</b>	404	16.4	Bucardo 的日常维护	434
15.1	Slony-I 中的概念	404	16.4.1	Bucardo 的触发器日志清理	434
15.1.1	集群	404	16.4.2	临时停止和启动同步的方法	435
15.1.2	节点	405	16.4.3	新增表到同步的方法	435
15.1.3	复制集合	405	16.4.4	移除某个表或序列的方法	435
15.1.4	数据原始生产者、数据提供者 和数据订阅者	405	<b>第 17 章</b>	<b>PL/Proxy 的使用</b>	436
15.1.5	slon 守护程序	405	17.1	PL/Proxy 中的概念	436
15.1.6	slonik 配置程序	405	17.1.1	PL/Proxy 的定义	436
15.2	Slony-I 复制的限制	405	17.1.2	PL/Proxy 的特性说明	438
15.3	在 Windows 下使用 pgAdminIII		17.2	PL/Proxy 安装及配置	438
	安装配置 Slony-I	406	17.2.1	编译安装	438
15.3.1	Windows 下安装 Slony-I	406	17.2.2	安装规划	439
15.3.2	Windows 配置 Slony-I 同步 实例	407	17.2.3	配置过程	439
15.4	在 Linux 下安装配置 Slony-I	419	17.3	PL/Proxy 的集群配置详解	443
15.4.1	编译安装 Slony-I	419	17.3.1	Cluster configuration API 方式	444
15.4.2	配置 Slony-I 复制	421	17.3.2	SQL/MED 方式配置集群	446
<b>第 16 章</b>	<b>Bucardo 的使用</b>	426	17.4	PL/Proxy 语言详解	446
16.1	Bucardo 中的概念	426	17.4.1	CONNECT	447
16.1.1	Bucardo 介绍	426	17.4.2	CLUSTER	447
16.1.2	Bucardo FAQ	426	17.4.3	RUN ON	447
16.1.3	Bucardo 同步中定义的概念	427	17.4.4	SPLIT	448
16.2	Bucardo 的安装方法	427	17.4.5	TARGET	448
16.2.1	Bucardo 的安装步骤	427	17.5	PL/Proxy 的一个高可用方案	448
16.2.2	安装 Test-Simple、ExtUtils- MakeMaker、version	428	17.5.1	方案介绍	448
			17.5.2	方案架构	449
			17.5.3	具体实施步骤	449

**第 18 章 pgpool-II 的使用** ..... 458

- 18.1 pgpool-II 中的概念 ..... 458
  - 18.1.1 pgpool-II 的定义 ..... 458
  - 18.1.2 pgpool-II 的架构 ..... 459
  - 18.1.3 pgpool-II 的工作模式 ..... 460
  - 18.1.4 pgpool-II 的程序模块 ..... 461
- 18.2 pgpool-II 安装方法 ..... 462
  - 18.2.1 源码安装 ..... 462
  - 18.2.2 安装 pgpool\_regclass ..... 463
  - 18.2.3 建立 insert\_lock 表 ..... 463
  - 18.2.4 安装 C 语言函数 ..... 463
- 18.3 pgpool-II 配置快速入门 ..... 463
  - 18.3.1 pgpool-II 的配置文件及启停方法 ..... 464
  - 18.3.2 复制和负载均衡的示例 ..... 466
  - 18.3.3 使用流复制的主备模式的示例 ..... 467
  - 18.3.4 show 命令 ..... 468
- 18.4 pgpool-II 高可用配置方法 ..... 471
  - 18.4.1 pgpool-II 高可用切换及恢复的原理 ..... 471
  - 18.4.2 pgpool-II 的健康检查 ..... 473
  - 18.4.3 复制和负载均衡模式的高可用示例 ..... 473
  - 18.4.4 使用流复制的主备模式下的  
高可用示例 ..... 488
- 18.5 pgpool-II 的总结 ..... 492

**第 19 章 Postgres-XC 的使用** ..... 493

- 19.1 Postgres-XC 中的概念 ..... 493
  - 19.1.1 Postgres-XC 的定义 ..... 493
  - 19.1.2 Postgres-XC 的特点 ..... 493
  - 19.1.3 Postgres-XC 的性能 ..... 494
  - 19.1.4 Postgres-XC 的组件 ..... 494
- 19.2 Postgres-XC 的安装 ..... 495
  - 19.2.1 源码安装方法 ..... 495
  - 19.2.2 Postgres-XC 目录及程序说明 ..... 496
- 19.3 配置 Postgres-XC 集群 ..... 497
  - 19.3.1 集群规划 ..... 497

- 19.3.2 初始化 GTM ..... 498
- 19.3.3 初始化 GTM 的备库 ..... 498
- 19.3.4 初始化 GTM Proxy ..... 499
- 19.3.5 初始化 Coordinators、数据  
节点 ..... 499
- 19.3.6 启动集群 ..... 500
- 19.3.7 停止集群 ..... 501
- 19.3.8 配置集群节点信息 ..... 502
- 19.4 Postgres-XC 的使用 ..... 503
  - 19.4.1 建表详解 ..... 503
  - 19.4.2 使用限制 ..... 509
  - 19.4.3 重新分布数据 ..... 510
  - 19.4.4 增加 Coordinator 节点的方法 ..... 512
  - 19.4.5 移除 Coordinator 节点的方法 ..... 513
  - 19.4.6 增加 Datanode 节点的方法 ..... 513
  - 19.4.7 移除 Datanode 节点的方法 ..... 514

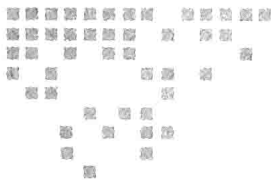
**第 20 章 高可用性方案设计** ..... 516

- 20.1 高可用架构基础 ..... 516
  - 20.1.1 各种高可用架构介绍 ..... 516
  - 20.1.2 服务的可靠性设计 ..... 517
  - 20.1.3 数据可靠性设计 ..... 517
- 20.2 基于共享存储的高可用方案 ..... 517
  - 20.2.1 SAN 存储的方案 ..... 517
  - 20.2.2 DRBD 的方案 ..... 518
- 20.3 WAL 日志同步或流复制同步的  
方案 ..... 519
  - 20.3.1 持续复制归档的 standby 的  
方法 ..... 519
  - 20.3.2 异步流复制的方案 ..... 519
  - 20.3.3 基于同步流复制方案 ..... 519
- 20.4 基于触发器的同步方案 ..... 520
  - 20.4.1 方案的特点 ..... 520
  - 20.4.2 基于触发器方案的同步软件  
介绍 ..... 520
- 20.5 基于语句中间件的高可用方案 ..... 520
  - 20.5.1 方案的特点 ..... 520
  - 20.5.2 基于语句中间件的开源软件  
介绍 ..... 521

第一篇 *Part 1*

# 准备篇

- 第 1 章 PostgreSQL 简介
- 第 2 章 PostgreSQL 安装与配置
- 第 3 章 SQL 语言入门



Chapter 1

第 1 章

# PostgreSQL 简介

本章将着重介绍 PostgreSQL 数据库的相关知识，让没有接触过 PostgreSQL 的读者对它有一个初步的了解。

## 1.1 什么是 PostgreSQL

### 1.1.1 PostgreSQL 概述

PostgreSQL 数据库是目前功能最强大的开源数据库，支持丰富的数据类型（如 JSON 和 JSONB 类型、数组类型）和自定义类型。而且它提供了丰富的接口，可以很容易地扩展它的功能，如可以在 GiST 框架下实现自己的索引类型等，它还支持使用 C 语言写自定义函数、触发器，也支持使用流行的语言写自定义函数，比如其中的 PL/Perl 提供了使用 Perl 语言写自定义函数的功能，当然还有 PL/Python、PL/Tcl，等等。

### 1.1.2 PostgreSQL 的发展历史

- 前身 Ingres：PostgreSQL 的前身是伯克利源于 1977 年的 Ingres 项目。这个项目是由著名的数据库科学家 Michael Stonebraker 领导的。1982 年，Michael Stonebraker 离开伯克利大学，把 Ingres 商业化，使之成为 Relational Technologies 公司的一个产品。后来 Relational Technologies 被 Computer Associates (CA) 收购。Ingres 是一个非关系型的数据库。
- 伯克利的 Postgres 项目：20 世纪 80 年代，数据库系统中的一个主要问题是数据关系维护。在 1985 年 Michael Stonebraker 回到伯克利后，为了解决 Ingres 中的数据关系维护问题，启动了一个后 Ingres (post-Ingres) 的项目，这就是 POSTGRES 的开端。

POSTGRES 项目是由防务高级研究项目局 (DARPA)、陆军研究办公室 (ARO)、国家自然科学基金 (NSF) 以及 ESL 公司共同赞助的。从 1986 年开始, Michael Stonebraker 教授发表了一系列论文, 探讨了新的数据库的结构设计和扩展设计。第一个“演示性”系统在 1987 年便可使用了, 并且在 1988 年的数据管理国际会议 (ACM-SIGMOD) 上展出。1989 年 6 月发布了版本 1 给一些外部的用户使用。由于源代码维护的时间日益增加, 占用了太多本应用于数据库研究的时间, 为减少支持的负担, 伯克利的 POSTGRES 项目在版本 4.2 时正式终止。

- ❑ Postgres95: 在 1994 年, 来自中国香港的两名伯克利的研究生 Andrew Yu 和 Jolly Chen 向 POSTGRES 中增加了现在 SQL 语言的解释器, 将 Postgres 改名为 Postgres95, 随后将 Postgres95 源代码发布到互联网上供大家使用。它成为一个开放源码的 POSTGRES 代码的继承者。
- ❑ PostgreSQL6.X: 到了 1996 年, 很明显地看出“Postgres95”这个名字已经经不起时间的考验。于是起了一个新名字 PostgreSQL, 为 Postgres 与 SQL 的缩写, 即增加了 SQL 功能的 Postgres 的意思。同时版本号也沿用伯克利 POSTGRES 项目的顺序, 从 6.0 开始。
- ❑ PostgreSQL7.1: PostgreSQL 7.1 是继 6.5 版本之后又一个有巨大变化的版本, 首先它引入了预写式日志的功能。这样, 事务就拥有了完善的日志机制, 可以提供更好的性能, 还可以实现更优良的备份和灾难恢复的能力 (比如联机热备份和宕机后的自动恢复)。其次是不再限制文本类型的数据段长度, 这从很大程度上解决了 PostgreSQL 大对象的问题。
- ❑ Postgres8.X: 该版本可以在 Windows 下运行, 它具有一些新的特性。比如具有事务保存点功能、改变字段的类型、表空间、即时恢复 (即时恢复允许对服务器进行连续的备份。既可以恢复到失败那个点, 也可以恢复到以前的任意事务) 等功能。并且开始支持 Perl 服务器端编程语言。
- ❑ PostgreSQL9.X: 进入 9.X 版本, 标识着 PostgreSQL 进入了黄金发展阶段。2010 年 9 月 20 日发布了 PostgreSQL 9.0, 大大增强了复制的功能 (replication), 比如增加了流复制功能 (stream replication) 和 HOT standby 功能。从 9.0 开始, 可以很方便地搭建主从数据库。2011 年 9 月 12 日发布了 PostgreSQL9.1, 在该版本中增加了同步复制功能 (synchronous replication); 2012 年 9 月 10 日发布了 PostgreSQL9.2, 增加了级连复制的功能。

### 1.1.3 PostgreSQL 数据库的优势

PostgreSQL 有以下优势:

- ❑ PostgreSQL 数据库是目前功能最强大的开源数据库, 它是最接近工业标准 SQL92 的查询语言, 并且正在实现新的功能以兼容最新的 SQL 标准: SQL2003。
- ❑ 稳定可靠: PostgreSQL 是唯一能做到数据零丢失的开源数据库。有报道称国外的部分银行也在使用 PostgreSQL 数据库。



- ❑ 开源省钱：PostgreSQL 数据库是开源的、免费的，而且是 BSD 协议，在使用和二次开发上基本没有限制。
- ❑ 支持广泛：PostgreSQL 数据库支持大量的主流开发语言，包括 C、C++、Perl、Python、Java、Tcl，以及 PHP 等。
- ❑ PostgreSQL 社区活跃：PostgreSQL 基本上每三个月推出一个补丁版本，这意味着已知的 BUG 很快会被修复，有应用场景的需求也会及时得到响应。

### 1.1.4 PostgreSQL 应用现状和发展趋势

PostgreSQL 目前在国外很流行，特别是近几年使用 PostgreSQL 数据库的公司越来越多。比如，日本电信（NTT）大量使用 PostgreSQL 替代 Oracle 数据库，并且在 PostgreSQL 之上二次开发了 Postgres-XC，Postgres-XC 是对使用者完全兼容 PostgreSQL 接口的 share-nothing 架构的数据库集群。网络电话公司 Skype 也大量使用 PostgreSQL，并贡献了一些与 PostgreSQL 数据库配套的开源软件：

- ❑ PL/Proxy：PostgreSQL 中的数据水平拆分软件
- ❑ pgQ：使用 PostgreSQL 的消息队列软件
- ❑ Londiste：用 C 语言实现的在 PostgreSQL 数据库之间进行逻辑同步的软件

全球最大的 CRM 软件服务提供商 Salesforce 也开始使用 PostgreSQL，并招募了 PostgreSQL 内核开发者 Tom lane。

2012 年，美国联邦机构全面转向 PostgreSQL 阵营；法国也正积极推动政府机构采用 PostgreSQL 数据库，从而取代商业数据库。

在国内，越来越多的公司开始使用 PostgreSQL，如斯凯网络（股票代码：MOBI）的后台数据库基本使用的都是 PostgreSQL 数据库，去哪儿网（qunar.com）也大量使用了 PostgreSQL 数据库。

主流的云服务提供商如亚马逊、阿里云的 RDS（关系型数据库服务）同样提供了 PostgreSQL 的支持。

更多的使用 PostgreSQL 数据库的情况可以见 PostgreSQL 官方网站（<http://www.postgresql.org/about/users/>）。

## 1.2 PostgreSQL 数据库与其他数据库的对比

### 1.2.1 PostgreSQL 与 MySQL 数据库的对比

可能有人会问，既然已经有一个人气很高的开源数据库 MySQL 了，为什么还要使用 PostgreSQL？这主要是因为在一些应用场景中，使用 MySQL 有以下几个缺点：

- ❑ 功能不够强大：MySQL 的多表连接查询方式只支持“Nest Loop”，不支持“hash join”和“sort merge join”。不仅如此，还有很多 SQL 语法它也不支持，子查询性能比