

# POSTGRESMEN

Компания «Постgresмен» (<http://postgresmen.ru>) с 2007 предоставляет услуги по настройке производительности, поддержке, администрированию баз данных PostgreSQL в крупных проектах, обслуживающих десятки миллионов пользователей онлайн. Компания сотрудничает с разработчиками ядра PostgreSQL, ежегодно представляя широкой публике возможность пообщаться со звездами мировой величины – разработчиками ядра СУБД PostgreSQL, а также техническими специалистами крупнейших проектов со всего мира.

«Обмен опытом с экспертами мирового уровня – необходимое условие для того, чтобы мы могли говорить о том, что наша ИТ-индустрия, особенно та ее часть, что опирается на продукты Open Source, качественно росла из года в год, – считает основатель компаний “Постgresмен” Николай Самохвалов. – Мы искренне рады, что в 2012 году на конференции Highload++ с ее тысячью посетителей нам с коллегами из PostgreSQL-Consulting.com удалось организовать целый день интереснейших докладов по PostgreSQL. А выход данной книги – еще один значительный шаг для популяризации этой замечательной системы в русскоязычной среде».



# Содержание

Об авторах.....	12
О рецензентах .....	13
<b>Предисловие .....</b>	<b>14</b>
О чем эта книга.....	14
Что вам понадобится .....	17
Для кого написана книга.....	17
Условные обозначения .....	18
Сообщение об ошибках .....	18
Незаконное воспроизведение содержимого книги.....	19
Вопросы .....	19
<b>ГЛАВА 1.</b>	
<b>Первые шаги.....</b>	<b>20</b>
Введение.....	20
Представляем PostgreSQL 9 .....	20
Где взять PostgreSQL .....	23
Соединение с сервером PostgreSQL .....	25
Открытие доступа для сетевых/удаленных пользователей.....	28
Использование графических инструментов администрирования .....	30
Использование инструмента для написания запросов и скриптов psql .....	34
Безопасная смена пароля .....	38
Как избежать хранения пароля в коде .....	39
Использование сервисного файла соединений .....	40
Если установить соединение не удается.....	41
<b>ГЛАВА 2.</b>	
<b>Исследование базы данных.....</b>	<b>44</b>
Введение.....	44
Как узнать версию сервера.....	44
Определение времени работы сервера .....	46
Где находятся файлы сервера БД .....	47

Где находится журнал сообщений сервера .....	49
Где находится системный идентификатор БД .....	52
Список баз данных на сервере .....	53
Сколько таблиц в базе данных .....	56
Сколько места на диске занимает база данных .....	59
Сколько места на диске занимает таблица .....	59
Какие таблицы являются самыми большими .....	60
Сколько строк в таблице .....	61
Быстрый подсчет строк в таблице .....	62
Понятие объектных зависимостей .....	67
<b>ГЛАВА 3.</b>	
<b>Конфигурация .....</b>	<b>70</b>
Введение .....	70
«Правильная» документация (RTFM) .....	70
Планирование новой базы данных .....	72
Изменение параметров в программах .....	73
Текущие конфигурационные настройки .....	75
Параметры, не определяемые по умолчанию .....	76
Обновление файла параметров .....	78
Настройка параметров для определенных групп пользователей .....	79
Основной список параметров для проверки конфигурации сервера .....	80
Добавление внешних модулей в PostgreSQL .....	82
Запуск сервера в энергосберегающем режиме .....	84
<b>ГЛАВА 4.</b>	
<b>Контроль над сервером .....</b>	<b>86</b>
Введение .....	86
Запуск сервера БД вручную .....	87
Быстрая и безопасная остановка сервера .....	88
Экстренная остановка сервера .....	89
Перезагрузка конфигурационных файлов .....	89
Быстрый перезапуск сервера .....	91
Запрет на создание новых соединений .....	93
Создание ограничения по одной сессии на пользователя .....	94
Принудительное отключение пользователей .....	96
Проектирование с учетом multi-tenancy .....	97
Использование множественных схем .....	99
Выделение для пользователя собственной базы данных .....	101
Запуск нескольких серверов на одной системе .....	103
Настройка пула соединений .....	104

**ГЛАВА 5.****Таблицы и данные ..... 108**

Введение.....	108
Выбор имени для объекта БД .....	108
Управление объектами с именами, заключенными в кавычки .....	110
Принудительная установка единого имени, единого определения столбцов.....	112
Обнаружение и удаление дубликатов .....	116
Предотвращение появления строк-дубликатов .....	119
Поиск уникального ключа для выборки данных .....	124
Генерирование тестовых данных .....	126
Случайные данные по примеру.....	129
Подгрузка данных из электронных таблиц .....	132
Подгрузка данных из файлов .....	135

**ГЛАВА 6.****Безопасность ..... 139**

Введение.....	139
Как забрать у пользователя права доступа к таблице.....	140
Как предоставить пользователю доступ к таблице.....	142
Создание нового пользователя .....	144
Временный запрет соединений для пользователя.....	145
Удаление пользователя без потери его данных.....	146
Проверка безопасности паролей пользователей.....	148
Назначение определенным пользователям ограниченных прав суперпользователя .....	150
Аудит изменений DDL .....	153
Аудит изменений данных .....	154
Интеграция с LDAP .....	158
Соединение с использованием SSL .....	159
Шифрование секретных данных .....	162

**ГЛАВА 7.****Администрирование баз данных ..... 166**

Введение.....	166
Написание скрипта, который срабатывает полностью или не работает вообще .....	167
Написание скрипта psql, который прерывается после первой ошибки .....	170
Действия с несколькими таблицами .....	171
Добавление и удаление столбца .....	176



Изменение типа данных столбца .....	178
Добавление и удаление схем.....	181
Перемещение объектов между схемами.....	183
Добавление и удаление табличных пространств.....	184
Перемещение объектов между табличными пространствами....	188
Доступ к объектам, находящимся в других БД PostgreSQL.....	191
Обновление представления .....	196

## ГЛАВА 8.

### **Мониторинг и диагностика ..... 202**

Введение.....	202
Подключен ли пользователь .....	205
Что выполняет пользователь .....	207
Активен или заблокирован запрос .....	209
Почему запрос блокирован.....	211
Как прервать сессию .....	212
Принятие решения о сомнительной подготовленной транзакции .....	215
Использует ли кто-либо определенную таблицу .....	215
Когда в последний раз использовали таблицу .....	217
Сколько места на диске будут занимать временные данные .....	220
Почему запросы выполняются медленно.....	222
Обнаружение ошибок и сообщение о них .....	225
Подведение итога записей в журнале об ошибках за день .....	227

## ГЛАВА 9.

### **Регулярное обслуживание ..... 230**

Введение.....	230
Управление автоматическим обслуживанием базы данных ...	231
Предотвращение «автозаморозки» и повреждения страниц.....	236
Предупреждение «свертывания» счетчика транзакций .....	238
Удаление ранее подготовленных транзакций.....	240
Для любителей временных таблиц.....	243
Обнаружение и исправление раздутых таблиц и индексов.....	244
Обслуживание индексов.....	249
Обнаружение неиспользуемых индексов .....	252
Безопасное удаление ненужных индексов .....	254
Планирование обслуживания .....	255

## ГЛАВА 10.

### **Производительность и согласованность ..... 258**

Введение.....	258
Как найти медленные запросы SQL .....	259

Как получить регулярную статистику из представлений pg_stat* .....	262
Как установить причины замедления работы SQL .....	263
Как уменьшить число возвращаемых строк .....	268
Как упростить сложный SQL .....	269
Как ускорить выполнение запросов, не переписывая их .....	275
Почему запрос не использует индекс .....	279
Как заставить запрос использовать индекс .....	280
Использование оптимистичных блокировок .....	282
Сообщения о проблемах с производительностью .....	284
<b>ГЛАВА 11.</b>	
<b>Резервное копирование и восстановление .....</b>	<b>285</b>
Введение .....	285
Что такое восстановление после сбоя и как его контролировать .....	286
Планирование резервного копирования .....	288
Горячее логическое резервное копирование одной базы данных .....	291
Горячее логическое резервное копирование всех баз данных .....	293
Горячее логическое резервное копирование всех таблиц из табличного пространства .....	294
Резервное копирование определений объектов базы данных .....	296
Автономное горячее физическое резервное копирование базы данных .....	297
Горячее физическое резервное копирование и непрерывное архивирование .....	300
Восстановление всех баз данных .....	303
Восстановление до определенной временной точки .....	307
Восстановление удаленной или поврежденной таблицы .....	310
Восстановление удаленного или поврежденного табличного пространства .....	312
Восстановление удаленной или поврежденной базы данных .....	314
Улучшение производительности резервного копирования и восстановления .....	315
Инкрементальное и дифференциальное резервное копирование и восстановление .....	318

**ГЛАВА 12.****Репликация и обновления ..... 321**

Репликация: основные термины .....	322
Лучшие способы репликации.....	326
Файловая репликация с передачей журналов .....	328
Настройка потоковой журнальной репликации .....	332
Управление репликацией с передачей журналов .....	337
Управление режимом горячего резервирования .....	341
Выборочная репликация с помощью Londiste 3.0.....	346
Выборочная репликация с помощью Slony 2.0 .....	351
Балансировка нагрузки с помощью pgpool-II 3.0.....	356
Минорное обновление.....	359
Мажорные обновления на месте.....	360
Мажорные онлайн-обновления с помощью инструментов репликации .....	362

## Об авторах

Саймон Ригс является ведущим разработчиком проекта PostgreSQL, а также техническим директором компании 2ndQuadrant, предоставляющей круглосуточную поддержку пользователям PostgreSQL по всему миру.

Саймон внес значительный вклад в развитие PostgreSQL, поскольку его работа напрямую связана с построением архитектуры баз данных и решением возникающих проблем. Среди добавленных им возможностей можно назвать восстановление точки во времени (Point in Time Recovery), репликацию с режимом с серверами теплого резерва (Warm Standby replication), режим горячего резерва (Hot Standby), асинхронную фиксацию транзакций (Asynchronous Commit), партиционирование (Partitioning). Его решения, позволяющие повысить производительность и усовершенствовать настройку, в огромном количестве встречаются в коде и документации по PostgreSQL.

Саймон также работал в Oracle, Teradata и DB2, имеет множество сертификатов. В качестве технического руководителя он трудился в банковской сфере, области телекоммуникаций и разработки программного обеспечения. Одно из его исследований было опубликовано Королевским обществом.

Ханну Кросинг является старшим консультантом в 2ndQuadrant и техническим советником в Ambient Sound Investments. В компании Skype Technologies он был архитектором баз данных и проектировал модуль для репликации Skytools и методы масштабирования. Уже 12 лет является разработчиком PostgreSQL.

## О рецензентах

Габриэль Бартолини, опытный разработчик открытого программного обеспечения, уже более десяти лет создает приложения для Linux/Unix на C и C++ и специализируется на поисковых машинах и веб-аналитике в больших базах данных.

Габриэль защитил диссертацию по статистике в Университете Флоренции. Анализируя веб-трафик в Австралии и Италии, он приобрел большой опыт в вопросах сбора, анализа и хранения данных.

Габриэль является консультантом в компании 2ndQuadrant и активным членом международного сообщества PostgreSQL.

В настоящее время Габриэль живет в Прато, небольшом, но оживленном городке на севере Тосканы в Италии. Вторым домом он считает австралийский Мельбурн, где он учился в Университете Монаша и работал в области ИТ.

Увлекается футболом и игрой на электрогитаре.

---

Благодарю свою семью, особенно Кэти, которая постоянно подбадривала меня, напоминая, что учиться никогда не поздно.

---

Димитрий Фонтеин входит в сообщество PostgreSQL и пишет программы с открытым кодом уже более десяти лет. Он разрабатывал такие промышленные решения, как `pgloader` (ETL), PostgreSQL `prefix indexing` (телефонная маршрутизация), `pg_staging` (поддержка окружения из резервных копий), `preprepare` (позволяет легко использовать подготовленные выражения в пуле соединений (connection pooler)) и патчи для обновлений и дополнений к PostgreSQL.

Кроме того, он принимал участие в разработке Skytools и системы репликации Londiste и создал поддержку PHP для системы организации асинхронной работы с очередями событий PgQ.

Его конек – улучшение системной архитектуры и дизайна с целью сокращения времени обслуживания и повышения доступности.

В настоящее время Димитрий работает старшим консультантом в компании 2ndQuadrant во Франции. Профессиональный опыт и отличное знание PostgreSQL позволяют ему вести активную работу в сообществе.

# Предисловие

PostgreSQL – это продвинутый сервер баз данных, который доступен на множестве платформ. Заработав отличную репутацию благодаря высокой производительности, стабильности и широким возможностям, PostgreSQL быстро набирает популярность, становясь одной из самых широко используемых СУБД в мире. Один из первых проектов с открытым кодом, в котором участвует множество разработчиков со всего мира, является абсолютно бесплатным. И это работает!

Преимущества свободного программного обеспечения с открытым кодом очевидны: вы вольны устанавливать, использовать и распространять PostgreSQL, и за это не нужно никому платить. Главное же заключается в том, что СУБД PostgreSQL способна долго и бесперебойно работать при минимальной поддержке со стороны администратора. В общем, стоимость использования PostgreSQL довольно мала.

Эта книга поможет вам управлять рабочими базами данных PostgreSQL. Здесь вы найдете рекомендации ведущего разработчика функционала репликации и восстановления PostgreSQL и системного архитектора Skype – самого успешного стартапа, созданного с помощью PostgreSQL. Данное руководство должно стать настольным для разработчиков, имеющих дело с «боевыми» базами данных и использующих Java, Python, Ruby, .Net в любой среде разработки.

В книге представлены ответы на общие вопросы и решения проблем, почерпнутые авторами из личного опыта работы с СУБД PostgreSQL – в качестве пользователей, разработчиков и преподавателей.

Каждая задача разбивается на простые части, или рецепты, проиллюстрированные кодом и сопровождаемые подробными объяснениями, как и почему это работает. Таким образом, руководство может быть полезным как новичкам, так и специалистам. Кроме того, здесь отражены все новые возможности PostgreSQL 9.

## О чем эта книга

В первой главе вы познакомитесь с PostgreSQL 9, узнаете, как скачать и установить PostgreSQL, как подключиться к серверу PostgreSQL, открыть доступ для сетевых/удаленных пользователей, как использовать визуальные инструменты администрирования, с помощью каких инструментов писать скрипты и запросы psql, как безопасно

менять свой пароль, использовать служебный файл соединений и выяснить причины потери соединения.

Вторая глава поможет вам определить версию сервера БД и время его работы. Вы поймете, где находятся файлы сервера БД, журнал сообщений сервера и системный идентификатор СУБД. Вы сможете просмотреть список баз данных на сервере и узнаете, сколько таблиц в базе данных, сколько места на диске занимают БД и отдельные таблицы, какие таблицы являются самыми большими, сколько строк в таблице и как определить зависимости объектов.

В третьей главе рассматривается, как читать документацию (RTFM), планировать новую базу данных, изменять параметры в программах, как узнавать текущие конфигурационные настройки, параметры, которые не являются значениями по умолчанию, как обновлять файл параметров, как устанавливать параметры для определенных групп пользователей. Также описываются простой список для проверки конфигурации сервера, добавление на сервер PostgreSQL внешних модулей и запуск сервера в режиме экономии энергии.

В четвертой главе вы найдете информацию о ручном запуске сервера БД, быстрой и безопасной остановке сервера, остановке сервера в экстренных случаях, перезагрузке конфигурационных файлов сервера, быстрым перезапуске сервера, запрете на установку новых соединений, установке ограничения по одной сессии на каждого пользователя и отключении пользователей от системы. Простые рецепты помогут вам организовать работу в команде, применять множественные схемы, выделять для пользователей приватные базы данных, запускать несколько серверов БД на одной системе и настраивать пул соединений.

Из пятой главы вы узнаете, как выбирать имена для объектов баз данных, как управлять объектами с квотированными именами, присваивать одно имя, одно определение для разных столбцов, обнаруживать и удалять строки-дубликаты, предотвращать появление строк-дубликатов, как находить уникальный ключ для набора данных, генерировать тестовые данные, случайные данные, подгружать данные из динамических таблиц и из неструктурированных файлов.

В шестой главе содержатся рецепты восстановления доступа пользователя к таблице, предоставления пользователю прав доступа к таблице, создания нового пользователя, временного ограничения соединения для пользователя, удаления пользователя без потери его данных, проверки безопасности пароля пользователей, делегирования ограниченных прав суперпользователя для определенных

пользователей, аудита изменений DDL, аудита изменений данных, интеграции с LDAP, установки соединения с использованием SSL и шифровании важных данных.

В седьмой главе рассказывается, как писать скрипт, в котором все строки должны выполняться или, наоборот, не выполняться, как создать скрипт в `psql`, который вызывается при первой ошибке, как выполнять действия сразу с несколькими таблицами, добавлять/удалять столбцы, изменять тип данных столбца, добавлять/удалять схемы, перемещать объекты между схемами, добавлять/удалять табличные пространства, перемещать объекты между табличными пространствами, получать доступ к объектам из других баз данных PostgreSQL и делать представления обновляемыми.

Рецепты, содержащиеся в восьмой главе, позволяют понять, подключен ли пользователь, что он запускает, является ли он активным или его заблокировали, кто его заблокировал, использует ли кто-либо определенную таблицу и когда он в последний раз это делал, сколько места на диске занимают временные данные, почему ваши запросы замедляются. Кроме того, вы узнаете, как обнаружить ошибку и сообщить о ней, как создавать ежедневный итоговый отчет на основе журнализации ошибок, как прервать определенную сессию и что делать с сомнительной подготовленной транзакцией.

Из девятой главы вы узнаете, как контролировать автоматическую поддержку базы данных, как избежать «автозаморозки» и повреждения страниц, обертывания транзакций, как удалять старые подготовленные транзакции, действия для «тяжелых»<sup>1</sup> пользователей со временными таблицами, как обнаруживать и исправлять раздутые таблицы и индексы, как поддерживать индексы, находить неиспользуемые индексы, аккуратно удалять ненужные индексы и планировать обслуживание.

В десятой главе описано, как находить медленные SQL-параметры, получать регулярную статистику из представлений `pg_stat*`, выявлять причины замедления SQL, уменьшать количество возвращаемых строк, упрощать сложный SQL, ускорять запросы без переписывания. Вы узнаете, почему запрос не использует индекс и как заставить запрос его задействовать. Кроме того, рассматривается использование оптимистичных блокировок и сообщение о проблемах с производительностью.

В одиннадцатой главе описываются резервное копирование и восстановление. Многие признают важность резервного копирования, но мало кто думает об этом всерьез. Из этой главы вы получите важную

<sup>1</sup> Речь идет о пользователях, генерирующих «тяжелые» (сильно нагружающие систему) запросы и/или о пользователях с плохим каналом связи. – Прим. ред.

информацию о том, как создавать резервную копию и как восстанавливать из нее базу данных PostgreSQL, узнаете о планировании резервного копирования, горячем логическом резервном копировании как одной, так и всех баз данных, горячем логическом резервном копировании всех таблиц в табличном пространстве, резервном копировании определений объектов БД, отдельном горячем физическом резервном копировании БД, горячем физическом резервном копировании и непрерывном архивировании. Кроме того, в этой главе рассматривается восстановление всех баз данных, восстановление к какому-либо моменту времени, восстановление удаленной или поврежденной таблицы, табличного пространства, улучшение производительности резервного копирования/восстановления и инкрементальное/дифференциальное резервное копирование и восстановление.

Двенадцатая глава посвящена репликации и обновлениям. Репликация – это не волшебство, но вещь удивительная (еще удивительнее, когда она работает). В этой главе дается определение репликации, описываются лучшие приемы, настройка репликации, основанной на трансфере файлов журнала, управление режимом hot standby, выборочная репликация с использованием Londiste 3.0, выборочная репликация с использованием Slony 2.0, балансирование загрузки с помощью pgpool II 3.0, обновление до нового минорного релиза (например, с 9.0.0 до 9.0.1), мажорные обновления на месте (например, с 8.4 на 9.0 или с 9.0 на 9.1) и мажорные обновления онлайн с помощью инструментов репликации.

## Что вам понадобится

Для работы с книгой вам понадобится следующее программное обеспечение:

- PostgreSQL 9.0 Server Software;
- psql client utility (part of 9.0);
- pgAdmin3 1.12.

## Для кого написана книга

Руководство предназначено системным администраторам, администраторам баз данных, архитекторам, разработчикам, а также всем, кто планирует работать с базами данных или уже занимается этим. В первую очередь она подойдет тем, у кого уже есть какой-либо технический опыт.

## Условные обозначения

Кодовые слова в тексте выделены **моноширинным** шрифтом. Например: «В PostgreSQL 9.0 вам больше не понадобится утилита pg\_Standby, поскольку многие ее функции теперь выполняет непосредственно сервер».

Блоки кода выглядят так:

```
CREATE USER repuser
    SUPERUSER
    LOGIN
    CONNECTION LIMIT 1
    ENCRYPTED PASSWORD 'changeme' ;
```

Фрагменты кода, на которые требуется обратить особенное внимание, выделены **полужирным моноширинным** шрифтом:

```
SELECT *
FROM mytable
WHERE (col1, col2, ... ,colN) IN (SELECT col1, col2, ... ,colN
FROM mytable
GROUP BY
col1, col2, ... ,colN HAVING count(*) > 1);
```

Все, что вводится или выводится в командной строке, также записывается **полужирным моноширинным** шрифтом:

```
$ postgres --single -D /full/path/to/datadir postgres
```

Новые термины и важные слова выделены *курсивом*.

Слова, которые вы можете видеть на экране, например в меню или диалоговых окнах, в тексте выделены следующим образом: «В Query есть замечательная опция **Visual Explain** и **Graphical Query Builder**».

Предупреждения, важные замечания, советы и подсказки вынесены в такие блоки.

## Сообщение об ошибках

Хотя мы делаем все возможное для того, чтобы не допустить ошибок в наших изданиях, вы можете встретить в тексте какие-либо неточности. В таком случае мы будем рады, если вы сообщите нам об этом. Пожалуйста, зайдите на веб-страницу <http://www.packtpub.com/support>, выберите название книги, в которой вам встретилась ошибка, и опишите ее, заполнив соответствующую форму. Как только эта информация будет проверена, мы опубликуем ее в разделе «Errata» (Ошибки) на сайте. Уже найденные ошибки можно просмотреть,

выбрав название интересующей вас книги в разделе <http://www.packtpub.com/support>.

Сообщения об ошибках в русскоязычном издании книги можно оставить на сайте издательства «ДМК Пресс»:

<http://www.dmk-press.ru/contacts1/contacts/>

## **Незаконное воспроизведение содержимого книги**

Пиратские копии – повсеместная проблема. Если вам встретились незаконным образом растиражированные экземпляры данной книги (в любом формате), пожалуйста, сообщите источник публикации, написав письмо по адресу [dm@dmk-press.ru](mailto:dm@dmk-press.ru) или [copyright@packtpub.com](mailto:copyright@packtpub.com).

## **Вопросы**

Вы можете присылать любые вопросы, касающиеся данной книги, по адресу [dm@dmk-press.ru](mailto:dm@dmk-press.ru) и [questions@packtpub.com](mailto:questions@packtpub.com). Мы постараемся решить возникшие проблемы.



# ГЛАВА 1.

## Первые шаги

Рассматриваемые темы:

- Представляем PostgreSQL 9
- Где взять PostgreSQL
- Соединение с сервером PostgreSQL
- Открытие доступа для сетевых/удаленных пользователей
- Использование графических инструментов администрирования
- Использование инструмента для написания запросов и скриптов psql
- Безопасная смена пароля
- Как избежать хранения пароля в коде
- Использование сервисного файла соединений
- Если установить соединение не удается

## Введение

PostgreSQL – это мощная система управления базами данных (СУБД), использующаяся для выполнения самых разных задач. Путь к овладению ею непрост, но любое путешествие должно начинаться с первого шага.

Для начала рассмотрим самое первое соединение с базой. Для многих оно оказывается непреодолимым препятствием, поэтому разберем его более подробно: сначала остановимся на доступе для удаленных пользователей, а оттуда доберемся до инструментов администрирования с графическим интерфейсом.

Также проанализируем инструмент запросов psql, который будет использоваться для подгрузки нашей базы данных.

Кроме того, здесь представлено несколько полезных рецептов, которые пригодятся вам в будущем.

## Представляем PostgreSQL 9

PostgreSQL – это продвинутый сервер для баз данных SQL, доступный на множестве платформ.

Одно из очевидных преимуществ СУБД PostgreSQL заключается в ее открытости: вы можете совершенно бесплатно устанавливать, использовать и распространять ее. Но самое главное – PostgreSQL славится своей стабильностью, и во многих случаях необходима лишь минимальная поддержка. Эта система очень недорога.

Возможности PostgreSQL огромны, ведь она разрабатывается и улучшается на протяжении вот уже 20 лет – начиная с рождения в Группе исследований проблем баз данных Калифорнийского университета Беркли. Сегодня свой вклад вносит огромная армия разработчиков кода, архитекторов, администраторов БД и преподавателей – многие из них используют PostgreSQL на своем основном месте работы. Некоторые, в том числе и авторы этой книги, работают в компаниях, специализирующихся на создании сервисов для PostgreSQL. Но нет такой компании, которая могла бы считать себя владельцем этой СУБД, и никто не требует и не предполагает, что вы где-то будете указывать, что используете PostgreSQL.

Ключевые особенности PostgreSQL:

- высокий уровень соответствия стандартам языка SQL вплоть до стандарта SQL:2008;
- архитектура «клиент-сервер»;
- рассчитанная на большое количество параллельных операций архитектура: чтение и запись в базу данных не мешают друг другу;
- большие возможности конфигурирования и работы с различными типами приложений;
- отличная масштабируемость и производительность в сочетании с широкими возможностями настройки.

## **Чем отличается PostgreSQL от других СУБД**

При разработке PostgreSQL ставились следующие задачи:

- создание надежного высококачественного программного кода, сопровождаемого подробными комментариями;
- сведение к минимуму необходимости в обслуживании системы;
- соответствие стандартам SQL, интероперабельность, обратная совместимость кода;
- высокая производительность, безопасность и доступность.

Многих удивляет, что по функциональности PostgreSQL ближе к Oracle или Microsoft SQL Server, чем к MySQL. Да и связывает MySQL с PostgreSQL лишь то, что оба эти проекта обладают открытым исходным кодом, в то время как основные свойства и философия различаются в корне.

Одной из ключевых особенностей Oracle – начиная с Oracle 7 – была «изоляция сnapshotов» (snapshot isolation), позволявшая читающим и пишущим процессам не блокировать друг друга, а работать независимо. Как ни удивительно, PostgreSQL – первая СУБД, в которой эта особенность заложена изначально. В PostgreSQL она называется MVCC (Multi-Version Concurrency Control – мультиверсионная система управления параллельным выполнением операций). Далее вы сможете ознакомиться с ней подробнее.

PostgreSQL – это СУБД для широкого использования. Вы настраиваете базу данных так, чтобы вам было удобно с ней работать. PostgreSQL представляет для этого огромное количество возможностей. Вы можете использовать как модель нормализованной базы данных, так и расширения для работы с массивами и record-подтипаами, или же положиться на полностью динамическую схему, используя расширение `hstore`. Кроме того, PostgreSQL позволяет создавать серверные функции на одном из десятка различных языков программирования.

Одно из полезных свойств PostgreSQL – расширяемость. Вы можете добавлять свои типы данных, операторы, типы индексов и функциональные языки. Например, можно переопределить поведение части системы и изменить выполнение команд, используя систему плагинов, или же добавить новый оптимизатор.

Все эти свойства открывают для архитекторов программного обеспечения огромные возможности – без лишних проблем создавать и поддерживать приложения на протяжении длительного времени. К сожалению, нельзя уместить в одну книгу все, что могло бы пригодиться разработчикам. Данное руководство посвящено администрированию, техническому обслуживанию и резервному копированию.

В самом начале, когда PostgreSQL была только системой для исследования в академических кругах, основные разработки сводились к добавлению новых впечатляющих функций. За последние 15 лет было переписано и усовершенствовано огромное количество строк кода, что сделало PostgreSQL одним из самых больших и стабильных серверов, готовых для оперативного использования.

Возможно, вы где-то читали, что PostgreSQL работает медленнее, чем Ваша Любимая СУБД (какой бы она ни была). Мы, авторы

этой книги, можем сообщить, что на протяжении последних шести лет трудились над улучшением производительности и наша команда добилась больших успехов в создании высокопроизводительного и масштабируемого сервера. Это дает проекту PostgreSQL большее пространство для развития.

Кто использует PostgreSQL? Среди крупных компаний можно назвать Apple, BASF, Genentech, IMDB.com, Skype, NTT, Yahoo и The National Weather Service. По данным Европейской комиссии, число загрузок PostgreSQL за год доходит до миллиона, и многие пользователи БД считают PostgreSQL достойной альтернативой их привычной СУБД.

Осталось упомянуть еще одну вещь. При рождении PostgreSQL получила имя Postgres, поэтому в некоторых ключевых частях проекта вы можете встретить именно это слово. Например, база данных по умолчанию называется postgres, и при установке программного обеспечения это слово часто выступает в качестве имени пользователя. Именно поэтому многие сокращают PostgreSQL до Postgres и считают эти два названия взаимозаменяемыми.

PostgreSQL читается «пост-грес-ку-эл», а Postgres – «пост-грес».

Некоторые используют еще и название Postgre, но это только вносит дополнительную путаницу. Двух названий вполне достаточно – пожалуйста, используйте только их!

## Где взять PostgreSQL

PostgreSQL на 100 % является свободным программным обеспечением с открытым кодом.

Вы можете свободно использовать, изменять и распространять PostgreSQL любым удобным для вас способом. Лицензия на применение PostgreSQL является улучшенной версией лицензии *BSD (Berkeley Distribution Software)* и в настоящее время называется *TPL (The PostgreSQL License)*.

## Как это сделать

PostgreSQL уже используется во многих пакетах приложений, так что, возможно, она уже установлена на ваших серверах. Многие дистрибутивы Linux включают в себя PostgreSQL как часть основной системы, или же PostgreSQL содержится на диске с этими дистрибутивами.

Единственная проблема в том, что, скорее всего, прилагаемая версия PostgreSQL не окажется самой последней. Точнее, она будет

самой свежей на момент выхода дистрибутива операционной системы. Как правило, нет смысла держаться именно за эту версию – такой подход не дает какой-либо дополнительной стабильности, а более новые версии СУБД так же хорошо поддерживаются различными Linux-дистрибутивами.

Если у вас до сих пор нет копии дистрибутива PostgreSQL или имеющаяся версия уже не нова, вы можете скачать исходники или бинарные пакеты для различных операционных систем на сайте <http://www.postgresql.org/download/>.

Конкретный способ установки зависит от используемой вами платформы, и универсального рецепта здесь нет. Просто следуйте инструкциям, прилагаемым к вашей версии дистрибутива – этого будет достаточно. Мы сознательно не описываем здесь процесс установки, чтобы не исказить информацию, приведенную специально в помощь вам.

Если вы хотите получать сообщения об обновлениях и последних новостях по электронной почте, можно подписаться на рассылку, описывающую обновления от всех производителей, которые поддерживают PostgreSQL. Ежемесячно вам будет приходить несколько писем о новых версиях ядра PostgreSQL, приложениях, конференциях, а также и другая информация для пользователей. Поддерживать связь с разработчиками порой небесполезно.

Подробнее о данной рассылке можно узнать по адресу <http://archives.postgresql.org/pgsql-announce/>.

## Как это работает

Многие удивляются: неужели все действительно бесплатно? Не придется ли в итоге платить? Кто стал бы трудиться даром?

Программные продукты с открытым кодом, такие как PostgreSQL, создаются сообществом разработчиков и пользователей. Из множества маленьких песчинок вырастает большая гора. Среди этих людей есть и любители, и профессионалы, ведь проектирование приложений часто бывает полезным не только разработчикам, но и их работодателям.

Хотите верьте, хотите нет – это просто работает!

## И еще

Не забудьте, что PostgreSQL – это не только ядро. Множество веб-сайтов предлагают дополнения, расширения и инструменты для PostgreSQL. В блогах можно обнаружить массу полезной информации

Конец ознакомительного фрагмента.  
Приобрести книгу можно  
в интернет-магазине «Электронный универс»  
([e-Univers.ru](http://e-Univers.ru))