



Введение в PostgreSQL



Что такое PostgreSQL

Кто использует

Основные свойства

История развития

Форки PostgreSQL

Разработчики

Версии системы

Цикл разработки

Поддержка

Новые возможности 9.5

Что такое PostgreSQL



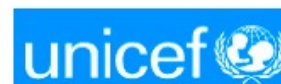
PostgreSQL – это свободно распространяемая объектно-реляционная СУБД (ORDBMS)

Поддержка ANSI SQL (1992...2011),
а также NoSQL (key-value, JSON, JSONB)

Web: <http://www.postgresql.org>

Лицензия: BSD, MIT-like

Кто использует



Надежность и устойчивость

Кроссплатформенность

Параллельная работа при большой нагрузке

Масштабируемость

Расширяемость

Доступность

Независимость

Поддержка

До 1985

Сетевые и иерархические БД

1970. Эдгар Кодд. Разработка реляционной модели данных

IBM System R

Ingres. Майкл Стоунбрейкер

Oracle

1985–1993

POSTGRES

В Беркли, под руководством Стоунбрейкера

На опыте Ingres, но созданная заново

Постреляционная

Принципы расширяемости

Последняя версия 4.2

Лицензия BSD и открытый код

1994–1995

Добавление SQL. Andrew Yu и Jolly Chen

Postgres95

1996–н.в.

PostgreSQL

Управление проектом: PostgreSQL Global Development Group
(PGDG)

Форки PostgreSQL

Большое количество форков

Postgres Plus Advanced Server (EnterpriseDB)

GreenPlum

Everest (Yahoo)

AsterData (Teradata)

JustOneDB

HadoopDB (Hadapt)

...

Коммерческие продукты

Системы специального назначения

Проверка и реализация идей

PostgreSQL Global Development Group

Управляющий комитет (Core Team)

Основные разработчики (Major Contributors)

Разработчики (Contributors)

Вадим Михеев (core team, сейчас отошел от проекта)

- Многоверсионность (MVCC)

- Очистка (vacuum)

- Журнал упреждающей записи (WAL)

- Вложенные запросы

- Триггеры

Олег Бартунов, Федор Сигаев, Александр Коротков
(основные разработчики)

- Поддержка локализации

- Полнотекстовый поиск

- Слабоструктурированные данные

- Новые методы индексации

- Различные расширения

Нумерация версий PostgreSQL: 9.4.4

Мажорная версия – 9.4, минорная – 4

Мажорная версия – добавление новой функциональности

Минорная версия – исправление ошибок

Для использования рекомендуется последняя минорная версия

Цикл разработки мажорной версии ~ 1 год

Предложение и обсуждение новых патчей –
в списке рассылки `pgsql-hackers`

Commitfest – регистрация и отслеживание статуса
предложенных патчей

Проверка патчей другими разработчиками

Бета-версии, Release Candidate

Выпуск новой версии

PGDG выполняет поддержку мажорных версий
в течении 5 лет

Для обращения с проблемой используются списки рассылки

Коммерческая поддержка

EnterpriseDB (Северная Америка)

2ndQuadrant (Европа)

Постгрес Профессиональный (Россия)

НОВЫЕ ВОЗМОЖНОСТИ 9.5

Row level security

pgaudit

GROUPING SETS, CUBE, ROLLUP

UPSERT

SKIP LOCKED

Block Range индексы (BRIN)

Abbreviated keys

pg_rewind

Сжатие WAL-файлов

...



Авторские права

Курс «Администрирование PostgreSQL 9.4. Базовый курс» разработан в компании Postgres Professional (2015 год).

Авторы: Егор Рогов, Павел Лузанов

Использование материалов курса

Некоммерческое использование материалов курса (презентации, демонстрации) разрешается без ограничений. Коммерческое использование возможно только с письменного разрешения компании Postgres Professional. Запрещается внесение изменений в материалы курса.

Обратная связь

Отзывы, замечания и предложения направляйте по адресу:
edu@postgrespro.ru

Отказ от ответственности

Компания Postgres Professional не несет никакой ответственности за любые повреждения и убытки, включая потерю дохода, нанесенные прямым или косвенным, специальным или случайным использованием материалов курса. Компания Postgres Professional не предоставляет каких-либо гарантий на материалы курса. Материалы курса предоставляются на основе принципа «как есть» и компания Postgres Professional не обязана предоставлять сопровождение, поддержку, обновления, расширения и изменения.

Что такое PostgreSQL
Кто использует
Основные свойства
История развития
Форки PostgreSQL
Разработчики
Версии системы
Цикл разработки
Поддержка
Новые возможности 9.5

Что такое PostgreSQL



PostgreSQL – это свободно распространяемая объектно-реляционная СУБД (ORDBMS)

Поддержка ANSI SQL (1992...2011),
а также NoSQL (key-value, JSON, JSONB)

Web: <http://www.postgresql.org>

Лицензия: BSD, MIT-like

3

PostgreSQL — наиболее полнофункциональная, свободно распространяемая СУБД с открытым кодом. Разработанная в академической среде, за долгую историю сплотившая вокруг себя широкое сообщество разработчиков, эта СУБД обладает всеми возможностями, необходимыми большинству заказчиков. PostgreSQL активно применяется по всему миру для создания критичных бизнес-систем, работающих под большой нагрузкой.

Кто использует



PostgreSQL имеет широкую **клиентскую базу**. Среди крупнейших пользователей PostgreSQL можно отметить: Skype (сейчас Microsoft), Yahoo, Instagram, BASF, Afilias и другие.

Позиции PostgreSQL сильны и в государственном секторе. Например, во Франции уже много лет действует государственная программа по переводу государственных учреждений на PostgreSQL. Среди результатов этой программы можно выделить:

Национальный фонд семейных пособий (CNAF) во Франции (данные по 30 млн человек; миллиард запросов в день)

Национальная метеослужба Франции (размер самой крупной БД – 3,5 ТБ)

leboncoin.fr (250 млн просмотров страниц в день; уникальных посетителей: в день – 5 млн, в месяц – 18 млн; 600000+ новых объявлений в день; 25 млн актуальных объявлений)

PostgreSQL получил широкое распространение и в России. Вот лишь некоторые из российских компаний, где активно используется PostgreSQL: Rambler, Yandex, Avito, Headhunter.

Надежность и устойчивость

Кроссплатформенность

Параллельная работа при большой нагрузке

Масштабируемость

Расширяемость

Доступность

Независимость

Поддержка

Надежность и устойчивость. PostgreSQL, на примере многих проектов, работает без единого сбоя и при больших нагрузках на протяжении нескольких лет.

Кроссплатформенность. PostgreSQL поддерживает все виды Unix, включая Linux, FreeBSD, Solaris, HP-UX, Mac OS X, а также MS Windows.

Параллельная работа при большой нагрузке. PostgreSQL использует многоверсионность (MVCC) для обеспечения надежной и быстрой работы большого количества одновременных транзакций.

Масштабируемость. PostgreSQL использует современную архитектуру многоядерных процессоров.

Расширяемость. PostgreSQL позволяет добавлять новую функциональность, в том числе и новые типы данных, без остановки сервера и своими силами.

Доступность. Лицензия BSD, не накладывает никаких ограничений на коммерческое использование и не требует лицензионных выплат. Вы можете даже продавать PostgreSQL под своим именем!

Независимость. PostgreSQL не принадлежит ни одной компании, развивается международным сообществом, в том числе и российскими разработчиками. Независимость PostgreSQL означает независимость вашего бизнеса от вендора и сохранность инвестиций.

Поддержка. Сообщество PostgreSQL предоставляет квалифицированную и быструю помощь. Коммерческие компании предлагают свои услуги по всему миру.

До 1985

Сетевые и иерархические БД

1970. Эдгар Кодд. Разработка реляционной модели данных

IBM System R

Ingres. Майкл Стоунбрейкер

Oracle

Краткую историю PostgreSQL можно найти в [документации](#). Из нее следует, что современный проект PostgreSQL ведет происхождение из проекта POSTGRES, который разрабатывался под руководством Майкла Стоунбрейкера (Michael Stonebraker), профессора Калифорнийского университета в Беркли (UCB). В марте 2015 года Майкл Стоунбрейкер стал лауреатом премии Тьюринга за фундаментальный вклад в принципы и практики, лежащие в основаниях современных систем управления базами данных.

А началось все с разработки реляционной модели Тедом Коддом (Ted Codd) из IBM в 1969, 1970 годах. В то время имелось две альтернативные модели баз данных – сетевая и иерархическая. Появление же реляционной модели привело к необходимости её практического применения.

Обращаясь к истории, можно отметить две ветви развития реляционных баз данных - одна следует из "System R", которая разрабатывалась в IBM в начале 70-х, и другая из проекта "INGRES", которым руководил Стоунбрейкер приблизительно в тоже время. Немного в стороне стоит Oracle, взлет которой во многом связан с коммерческим талантом Эллисона быть в нужном месте и в нужное время, как сказал Стоунбрейкер в одном из своих интервью, хотя она вместе с IBM сыграла большую роль в создании и продвижении SQL.

INGRES, вполне в духе Беркли, развивалась как открытая база данных, коды которой распространялись на лентах практически бесплатно (оплачивались почтовые расходы и стоимость ленты).

1985–1993

POSTGRES

В Беркли, под руководством Стоунбрейкера

На опыте Ingres, но созданная заново

Постреляционная

Принципы расширяемости

Последняя версия 4.2

Лицензия BSD и открытый код

7

Доступность INGRES и очень либеральная лицензия BSD, а также творческая деятельность, способствовали появлению большого количества реляционных баз данных и Стоунбрейкер лично способствовал их появлению.

Проект Postgres возник как результат осмысления ошибок Ingres и желания преодолеть ограниченность типов данных, за счет возможности определения новых типов данных.

Работа над проектом началась в 1985 и до 1988 было опубликовано несколько статей, описывающих модель данных, язык запросов POSTQUEL, и хранилище Postgres. POSTGRES иногда еще относят к так называемым постреляционным СУБД. Ограниченность реляционной модели всегда являлась предметом критики, хотя все понимали, что это является следствием ее простоты и ее заслугой. Однако, проникновение компьютерных технологий во все сферы жизни требовали новых приложений, а от баз данных – поддержки новых типов данных и возможностей, например, поддержка наследования, создание и управление сложными объектами.

Первая версия была выпущена в 1989 году. Нужно отметить, что коды Ingres и Postgres не имели ничего общего. После выпуска в 1993 году версии 4.2 проект был закрыт.

1994–1995

Добавление SQL. Andrew Yu и Jolly Chen

Postgres95

1996–н.в.

PostgreSQL

Управление проектом: PostgreSQL Global Development Group (PGDG)

Несмотря на официальное закрытие проекта, открытый код и BSD лицензия подвигли выпускников Беркли Andrew Yu и Jolly Chen в 1994 году взяться за его дальнейшее развитие. В 1995 году они заменили язык запросов POSTQUEL на общепринятый SQL, проект получил название Postgres95, изменилась нумерация версий, был создан веб сайт проекта и появились много новых пользователей.

К 1996 году стало ясно, что название «Postgres95» не выдержит испытание временем и было выбрано новое имя – «PostgreSQL», которое отражает связь с оригинальным проектом POSTGRES и приобретением SQL. Также, вернули старую нумерацию версий, таким образом новая версия стартовала как 6.0. Проект вырос и управление на себя взяла небольшая вначале группа инициативных пользователей и разработчиков, которая и получила название PGDG (PostgreSQL Global Development Group). Дальнейшее развитие проекта полностью [документировано](#) и отражено в архивах списка рассылки [pgsql-hackers](#).

Большое количество форков

Postgres Plus Advanced Server (EnterpriseDB)

GreenPlum

Everest (Yahoo)

AsterData (Teradata)

JustOneDB

HadoopDB (Hadapt)

...

Коммерческие продукты

Системы специального назначения

Проверка и реализация идей

PostgreSQL имеет большое количество форков.

Форки часто делают для проверки и реализации крупных идей, когда не получается все успеть за один релизный цикл (~ 1год). Идея проверяется, тестируется и через какое-то время возвращается в основную ветку PostgreSQL.

Многие коммерческие базы данных построены на основе PostgreSQL.

EnterpriseDB в 2014 году со своим продуктом Postgres Plus Advanced Server (построенным на базе PostgreSQL) включена в [Gartner Magic Quadrant](#) в квадрат лидеров среди операционных СУБД за 2014 год.

ЕМС с GreenPlum в 2013 включена в [Gartner Magic Quadrant](#) в квадрат лидеров среди БД для хранилищ данных (data warehousing).

Есть форки PostgreSQL и в России. Например, СУБД “Заря” используется для систем специального назначения.

PostgreSQL Global Development Group

Управляющий комитет (Core Team)

Основные разработчики (Major Contributors)

Разработчики (Contributors)

Глобальная группа разработчиков PGDG

Все основные решения о планах развития и выпусках новых версий принимаются Управляющим комитетом (Core team) проекта.

В настоящий момент он состоит из шести человек.

Помимо этого, выделяется внушительная группа основных (major) разработчиков, внесших существенный вклад в развитие PostgreSQL, а также просто разработчиков.

Состав группы разработчиков со временем меняется, появляются новые члены, кто-то уходит от проекта. [Актуальный список разработчиков](#) поддерживается на официальном сайте PostgreSQL.

Вадим Михеев (core team, сейчас отошел от проекта)

- Многоверсионность (MVCC)
- Очистка (vacuum)
- Журнал упреждающей записи (WAL)
- Вложенные запросы
- Триггеры

Олег Бартунов, Федор Сигаев, Александр Коротков
(основные разработчики)

- Поддержка локализации
- Полнотекстовый поиск
- Слабоструктурированные данные
- Новые методы индексации
- Различные расширения

11

Вклад российских разработчиков в PostgreSQL весьма значительный. Это один из самых крупных глобальных open source проектов, с таким широким российским представительством.

Огромную роль в становлении и развитии PostgreSQL сыграл, входивший в Управляющий комитет, Вадим Михеев (сейчас отошел от проекта). Он является автором таких важнейших частей системы как: многоверсионное управление одновременным доступом (MVCC), система очистки (Vacuum), журнал транзакций (WAL), вложенные запросы, триггеры.

В настоящий момент среди основных разработчиков проекта сразу трое представителей из России (Олег Бартунов, Федор Сигаев и Александр Коротков). Среди крупных направлений выполненных ими работ можно выделить: локализацию PostgreSQL (поддержка национальных кодировок, включая Unicode), систему полнотекстового поиска, работу со слабо-структурированными данными (hstore, json, jsonb), новые методы индексации (GiST, GIN, SP-GiST). Они являются авторами большого количества популярных расширений.

Компания «Постгрес Профессиональный», одной из своих важных задач ставит существенное увеличение российского представительства в глобальной группе разработчиков PostgreSQL.

Нумерация версий PostgreSQL: 9.4.4

Мажорная версия – 9.4, минорная – 4

Мажорная версия – добавление новой функциональности

Минорная версия – исправление ошибок

Для использования рекомендуется последняя минорная версия

Номер версии системы состоит из трех чисел, разделенных точкой. Например, 9.4.4.

Первые два числа представляют собой мажорную версию PostgreSQL. В приведенном примере это 9.4. Мажорные версии выпускаются примерно раз в год и всегда включают новую функциональность.

Последнее третье число (4) – минорная версия. Минорные версии содержат только исправления ошибок. После выхода очередной минорной версии рекомендуется перейти на нее.

Цикл разработки мажорной версии ~ 1 год
Предложение и обсуждение новых патчей –
в списке рассылки [pgsql-hackers](#)
Commitfest – регистрация и отслеживание статуса
предложенных патчей
Проверка патчей другими разработчиками
Бета-версии, Release Candidate
Выпуск новой версии

Цикл работы над новой мажорной версией PostgreSQL обычно составляет около года. Немного упрощая, процесс выглядит следующим образом.

Соблюдая определенные [правила](#), любой желающий может отправить на рассмотрение свой патч. Для обсуждения патчей используется список рассылки [pgsql-hackers](#). Если патч прошел обязательную процедуру проверки другими разработчиками, то он включается в новый релиз. Для удобного отслеживания предложенных патчей и их статуса, существует специальный сайт [commitfest](#).

Несколько раз в течении релизного цикла выпускаются бета версии. Выпуск бета версий обычно совмещают с проведением конференций PGDG.

В некоторый момент объявляется этап замораживания кода (code freeze), после которого патчи с новой функциональностью не принимаются, а допускается только исправление или улучшение кода. Иногда, в процессе работы над новой версией вскрываются или исправляются старые ошибки, а наиболее критичные из них исправляются и в предыдущих версиях (backporting). По мере накопления таких исправлений принимается решение о выпуске новой стабильной версии, которая совместима со старой. Например, 9.4.3 - является bugfix-ом для версии 9.4.

Ближе к концу цикла выпускается Release Candidate, а вскоре выходит и новая мажорная версия PostgreSQL.

PGDG выполняет поддержку мажорных версий
в течении 5 лет

Для обращения с проблемой используются списки рассылки

Коммерческая поддержка

EnterpriseDB (Северная Америка)

2ndQuadrant (Европа)

Постгрес Профессиональный (Россия)

PostgreSQL Global Development Group выполняет поддержку мажорных версий системы в течении 5 лет с момента выпуска. Поддержка, как и разработка, осуществляется через списки рассылки. **Корректно** оформленное сообщение об ошибке имеет все шансы на скорейшее решение. Не нужно ждать недели или даже месяцы – нередко случаи, когда патчи с исправлением ошибок выпускаются в течении суток.

Помимо поддержки сообществом разработчиков, ряд **компаний** осуществляет коммерческую поддержку PostgreSQL. Среди них можно выделить **EnterpriseDB** (в Северной Америке) и **2ndQuadrant** (в Европе).

Коммерческая поддержка доступна и в России. Компания «Постгрес Профессиональный» оказывает услуги по поддержке решений на базе PostgreSQL в режиме 24x7.

Row level security
pgaudit
GROUPING SETS, CUBE, ROLLUP
UPSERT
SKIP LOCKED
Block Range индексы (BRIN)
Abbreviated keys
pg_rewind
Сжатие WAL-файлов
...

PostgreSQL постоянно и активно развивается. Осенью 2015 года выйдет версия 9.5. Вот лишь некоторые из новых возможностей, которые появятся в этой версии:

- Row level security. Предоставление доступа не к таблице целиком, а к отдельным строкам. Эта возможность также известна как Virtual Private Database или Fine-grained access control.
- pgaudit. Расширение для детального аудита выполняемых операций (DML и DDL) в базе данных. Будет полезным для регистрации аудиторского следа.
- GROUPING SETS, CUBE, ROLLUP. Эта функциональность SQL теперь поддерживается и в PostgreSQL
- UPSERT. Новая команда для вставки строки в таблицу, если такой строки еще нет. Либо выполняет update для существующей строки.
- SKIP LOCKED. Позволяет пропустить заблокированные строки при выборке командой `SELECT ... FOR UPDATE SKIP LOCKED`
- Block Range индексы (BRIN). Индексы с новым методом доступа по диапазонам страниц особенно полезны для очень больших таблиц.
- Abbreviated keys. Ускорение сортировки текстовых данных. Например, создание индекса для типа данных `text` проходит в 3 раза быстрее, чем в версии 9.4
- Расширение `pg_rewind`. При использовании репликации ведущий-ведомый, позволяет быстро синхронизировать упавший ведущий сервер с ведомым, на который произошло переключение после сбоя.
- Сжатие WAL-файлов.